Wealth distribution and social mobility in the US: A quantitative approach

Jess Benhabib — NYU and NBER
Alberto Bisin — NYU and NBER
Mi Luo — NYU

First draft: July 2015; this draft: September 2017*

Abstract

We quantitatively identify the factors that drive wealth dynamics in the U.S. and are consistent with its skewed cross-sectional distribution and with social mobility. We concentrate on three critical factors: i) skewed earnings, ii) differential saving across wealth levels, and iii) stochastic idiosyncratic returns to wealth. All of these are fundamental for matching both distribution and mobility. The stochastic process for returns which best fits the cross-sectional distribution of wealth and social mobility in the U.S. shares several statistical properties with those of the returns to wealth uncovered by Fagereng et al. (2017) from tax records in Norway.

Key Words: wealth distribution; thick tails; inequality; social mobility
JEL Numbers: E13, E21, E24

*Thanks to seminar audiences at Duke, NYU, Minneapolis Fed, SED-Warsaw, Lake Baikal Summer School, SAET-Cambridge, University College of London, Wharton School, NBER Summer Institute. Special thanks to Alberto Alesina, Fernando Alvarez, Orazio Attanasio, Laurent Calvet, Tim Christensen, Tom Cogley, Mariacristina De Nardi, Pat Kehoe, Dirk Krueger, Per Krusell, Konrad Menzel, Ben Moll, Andrew Newman, Tom Sargent, Ananth Shesadri, Rob Shimer, Kevin Thom, Gianluca Violante, Daniel Xu, Fabrizio Zilibotti. Special thanks to Luigi Guiso, for many illuminating discussions and for spotting a mistake in a previous version, to the Editor and the referees for their exceptional work on the paper. Generous financial support from the Washington Center for Equitable Growth is gratefully aknowledged.
Corresponding author's email: alberto.bisin@nyu.edu.

1 Introduction

Wealth in the U.S. is unequally distributed, with a Gini coefficient of 0.82. It is skewed to the right, and displays a thick, right tail: the top 1% of the richest households in the United States hold over 33.6% of wealth.[1] At the same time, the U.S. is characterized by a non-negligible social mobility, with an inter-generational Shorrocks mobility index in the range of $.88 - .98$.[2]

This paper attempts to quantitatively identify the factors that drive wealth dynamics in the U.S. and are consistent with the observed cross-sectional distribution of wealth and with the observed social mobility. We first develop a macroeconomic model displaying various distinct wealth accumulation factors. Once we allow for an explicit demographic structure, the model delivers implications for social mobility as well as for the cross-sectional distribution. We then match the moments generated by the model to several empirical moments of the observed distribution of wealth as well as of the social mobility matrix. While the model is very stylized and parsimonious, it allows us to identify various distinct wealth accumulation factors through their distinct role on inequality and mobility: savings rates which increase steeply with wealth, e.g., deliver the thick tails of the wealth distribution but also imply too little intergenerational mobility relative to the data.

Many recent studies of wealth distribution and inequality focus on the relatively difficult task of explaining the thickness of the upper tail. We shall concentrate mainly on three critical factors previously shown, typically in isolation from each other, to affect the tail of the distribution, empirically and theoretically. First, a skewed and persistent distribution of *stochastic earnings* translates, in principle, into a wealth distribution with similar properties. A large literature in the context of Aiyagari-Bewley economies has taken this route, notably

[1]See Diaz-Gimenez et al. (2011), Table 6, elaborating data from the Survey of Consumer Finances (SCF) 2007.

[2]The range is computed, respectively, from the social mobility matrix in Charles and Hurst (2003), Table 2, from PSID data, and the matrix we construct from SCF 2007-2009 data, in Section 3.2; see footnotes 19-20 for more detailed discussions of the different methodologies adopted.

Castañeda et al. (2003) and Kindermann and Krueger (2015).[3] Another factor which could contribute to generating a skewed distribution of wealth is *differential saving rates* across wealth levels, with higher saving and accumulation rates for the rich. In the literature this factor takes the form of non-homogeneous bequests, bequests as a fraction of wealth that are increasing in wealth; see for example Cagetti and De Nardi (2016).[4] Finally, stochastic idiosyncratic returns to wealth, or *capital income risk*, has been shown to induce a skewed distribution of wealth, in Benhabib et al. (2011); see also Quadrini (2000), which focuses on entrepreneurial risk.[5]

Allowing rates of return on wealth to be increasing in wealth might also add to the skewedness of the distribution. This could be due e.g., to the existence of economies of scale in wealth management, as in Kacperczyk et al. (2015), or to fixed costs of holding high return assets, as in Kaplan et al. (2016); see Saez and Zucman (2016), Fagereng et al. (2016, 2017) and Piketty (2014, p. 447) for evidence about the relationship between returns and wealth.

While all these factors possibly contribute to produce skewed wealth distributions, their relative importance remains to be ascertained.[6] In our quantitative analysis we find that all the factors we study, stochastic earnings, differential savings, and capital income risk, have a fundamental role in generating the thick right tail of the wealth distribution and sufficient social mobility in the wealth accumulation process. We also identify a distinct role for these factors. Capital income risk and differential savings both contribute to generating the thick tail. But while differential savings at the same time reduces social mobility, capital income

[3]Several papers in the literature include a stochastic length of life (typically, "perpetual youth") to complement the effect of skewed earnings on wealth. We do not include this in our model as it has counterfactual demographic implications.

[4]See also Piketty (2014), which directly discusses the saving rates of the rich.

[5]Krusell and Smith (1998) instead introduce stochastic discount factors. However, such discount factors are non-measurable. Micro data allowing estimates of capital income risk are instead rapidly becoming more available; see e.g., the tax records for Norway studied by Fagereng et al. (2016, 2017) and the Swedish data studied by Bach et al. (2017).

[6]Other possible factors which qualitatively would induce skewed wealth distributions include a precautionary savings motive for wealth accumulation. In fact, the precautionary motive, by increasing the savings rate at low wealth levels under borrowing constraints and random earnings, works in the opposite direction of savings rates increasing in wealth. We do not exploit this channel for simplicity, assuming that life-cycle earnings profiles are random across generations but deterministic within lifetimes.

risk appears not to have major role in affecting mobility. On the other hand, stochastic earnings have a limited role in filling the tail of the wealth distribution but are fundamental in inducing enough mobility in the the wealth process. Finally, a rate of return of wealth increasing in wealth itself is also apparently supported in our estimates, mostly by contributing to fill the tail of the wealth distribution (though, without directly observing return data, this mechanism is somewhat poorly identified).

The rest of the paper is structured as follows. Section 2 lays out the theoretical framework. Section 3 explains our quantitative approach and data sources we use. Section 4 shows the baseline results with the model fit for both targeted and un-targeted moments. Section 5 presents several counterfactual exercises, where we re-estimate the model shutting down one factor at a time. Section 6 reports on a robustness exercise, allowing for non-stationarity of the wealth distribution and measuring the transition speed our model delivers. Section 7 concludes.

2 Wealth dynamics and stationary distribution

Most models of the wealth dynamics in the literature focus on deriving skewed distributions with thick tails, e.g., Pareto distributions (power laws).[7] While this is also our aim, we more generally target the whole wealth distribution and its intergenerational mobility properties. To this end we study a simple micro-founded model - a standard macroeconomic model in fact - of life-cycle consumption and savings. While very parsimounious, the model exploits the interaction of the factors identified in the Introduction that tend to induce skewed wealth distributions: stochastic earnings, differential saving and bequest rates across wealth levels, and stochastic returns on wealth.

Each agent's life span is finite and deterministic, T years. Every period t, consumers

[7]See Benhabib and Bisin (2017) for an extensive survey of the theoretical and empirical literature on the wealth distribution.

choose consumption c_t and accumulate wealth a_t, subject to a no-borrowing constraint. Consumers leave wealth a_T as a bequest at the end of life T. Each agent's preferences are composed of a per-period utility from consumption, $u(c_t)$, at any period $t = 1, \dots, T$, and a warm-glow utility from bequests at T, $e(a_T)$. Their functional forms display Constant Relative Risk Aversion:

$$u(c_t) = \frac{{c_t}^{1-\sigma}}{1-\sigma}, \quad e(a_T) = A\frac{{a_T}^{1-\mu}}{1-\mu}.$$

Wealth accumulates from savings and bequests. Idiosyncratic rates of return r and lifetime labor earnings profiles $w = \{w_t\}_{t=1}^{T}$ are drawn from a distribution at birth, possibly correlated with those of the parent, deterministic within each generation.[8] We emphasize that r and w are stochastic over generations only: agents face no uncertainty within their life span. Lifetime earnings profiles are hump-shaped, with low earnings early in life. Borrowing constraints limit how much agents can smooth lifetime earnings.

Let $\beta < 1$ denote the discount rate. Let $V_t(a_t)$ denote the present discounted utility of an agent with wealth a_t at the beginning of period t. Given initial wealth a_0, earnings profile w, and rate of return r, each agent's maximization problem, written recursively, then is:

$$\begin{array}{ll} V_t(a_{t-1}) = & \max_{c_t, a_{t+1}} u(c_t) + \beta V_{t+1}(a_{t+1}) \\ s.t. & a_t = (1+r)a_{t-1} - c_t + w_t \\ & 0 \leq c_t \leq a_t, \;\; t = 1, \dots, T-1 \\ V_T(a_T) = & \frac{1}{\beta} e(a_T) \end{array}$$

The solution of the recursive problem can be represented by a map

$$a_T = g\left(a_0; r, w\right).$$

[8]As we noted, assuming deterministic earning profiles amounts to disregarding the role of intra-generational life-cycle uncertainty and hence of precautionary savings. While the assumption is motivated by simplicity, see Keane and Wolpin (1997), Huggett et al. (2011), and Cunha et al. (2010) for evidence that the life-cycle income patterns tend to be determined early in life.

Following Benhabib et al. (2011), we exploit the map $g(.)$ as the main building block to construct the stochastic wealth process across generations. Adding an apex n to indicate the generation and slightly abusing notation, we denote with $\{r^n, w^n\}_n$ the stochastic process over generations for the rate of return on wealth r and earnings w. We assume it is a finite irreducible Markov Chain. We assume also that r^n and w^n are independent, though each is allowed to be serially correlated, with transition $P\left(r^n \mid r^{n-1}\right)$ and $P\left(w^n \mid w^{n-1}\right)$. The life-cycle structure of the model implies that the initial wealth of the n'th generation coincides with the final wealth of the $n-1$'th generation: $a^n = a_0^n = a_T^{n-1}$. We can then construct a stochastic difference equation for the initial wealth of dynasties, induced by $\{r^n, w^n\}_n$, mapping a^{n-1} into a^n:

$$a^n = g\left(a^{n-1}; r^n, w^n\right).$$

This difference equation in turn induces a stochastic process $\{a^n\}_n$ for initial wealth a.

It can be shown that, under our assumptions, the map $g(.)$ can be characterized as follows:

if $\mu = \sigma$, then $g\left(a_0; r, w\right) = \alpha(r, w)a_0 + \beta(r, w)$;

if $\mu < \sigma$, then $\frac{\partial^2 g}{\partial {a_0}^2}\left(a_0; r, w\right) > 0$.

In the first case, $\mu = \sigma$, the savings rate is $\alpha(r, w)$ and it is independent of wealth. In this case, the wealth process across generations is represented then by a linear stochastic difference equation in wealth, which has been closely studied in the math literature; see de Saporta (2005). Indeed, if $\mu = \sigma$, under general conditions,[9] the stochastic process $\{a^n\}_n$ has a stationary distribution whose tail is independent of the distribution of earnings and asymptotic to a Pareto law:

$$Pr(a > \underline{a}) \sim Q\underline{a}^{-\gamma},$$

[9]More precisely, the tail of earnings must be not too thick and furthermore $\alpha(r^n, w^n)$ and $\beta(r^n, w^n)$ must satisfy the restrictions of a *reflective process*; see Grey (1994), Hay et al. (2011), and Benhabib et al. (2011), for a related application.

where $Q \geq 1$ is a constant and $\lim_{N\to\infty} E\left(\prod_{n=0}^{N-1}\left(\alpha(r^{-n},w^{-n})\right)^{\gamma}\right)^{\frac{1}{N}} = 1$.[10]

If instead, keeping σ constant, $\mu < \sigma$, differential savings rate emerge, increasing with wealth. In this case, a stationary distribution might not exist; but if it does,

$$Pr(a > \underline{a}) \geq Q(\underline{a})^{-\gamma},$$

and hence it displays a thick tail.

Finally, the model is straightforwardly extended to allow for the Markov states of the stochastic process for r to depend on the initial wealth of the agent a. In this case, the intergenerational wealth dynamics have properties similar to the $\mu < \sigma$ case: a stationary distribution might not exist; but if it does, it displays a thick tail.

3 Quantitative analysis

The objective of this paper, as we discussed in the Introduction, consists in measuring the relative importance of various factors which determine the wealth distribution and the social mobility matrix in the U.S. The three factors are stochastic earnings, differential saving and bequest rates across wealth levels, and stochastic returns on wealth. These are represented in the model by the properties of the dynamic process and the distribution of (r^n, w^n) and by the parameters μ and σ, which imply differential savings (the rich saving more) when $\mu < \sigma$.

3.1 Methodology

We estimate the parameters of the model described in the previous section using a Method of Simulated Moments (MSM) estimator: i) we fix (or externally calibrate) several parameters

[10] While a denotes initial wealth, it can be shown that when the distribution of initial wealth has a thick tail, the distribution of wealth also does; see Benhabib et al. (2011) for the formal result.

of the model; ii) we select some relevant moments of the wealth process as target in the estimation; and iii) we estimate the remaining parameters by matching the targeted moments generated by the stationary distribution induced by the model and those in the data. The quantitative exercise is predicated then on the assumption that the wealth and social mobility observed in the data are generated by a stationary distribution.[11]

More formally, let θ denote the vector of the parameters to be estimated. Let m_h, for $h = 1, \ldots, H$, denote a generic empirical moment; and let $d_h(\theta)$ the corresponding moment generated by the model for given parameter vector θ. We minimize the deviation between each targeted moment and the corresponding simulated moment. For each moment h, define $F_h(\theta) = d_h(\theta) - m_h$. The MSM estimator is

$$\hat{\theta} = \arg\min_{\theta} \mathbf{F}(\theta)' W \mathbf{F}(\theta).$$

where $\mathbf{F}(\theta)$ is a column vector in which all moment conditions are stacked, i.e. $\mathbf{F}(\theta) = [F_1(\theta), \ldots, F_H(\theta)]^T$. The weighting matrix in the baseline is a diagonal matrix with identical weights for all but the last moment of both the wealth distribution and the mobility moments, which are overweighted (10 times). This is according to the prior that matching the tail of the distribution is a fundamental objective of our exercise.[12] Also, the objective function is highly nonlinear in general and therefore, following Guvenen (2016), we employ a global optimization routine for the MSM estimation.[13]

In our quantitative exercise we proceed as follows.

i) We fix $\sigma = 2$, $T = 36$, $\beta = 0.97$ per annum. We feed the model with a stochastic process

[11] Very few studies in the literature deal with the transitional dynamics of wealth and its speed of transition along the path, though this issue has been put at the forefront of the debate by Piketty (2014). Notable and very interesting exceptions are Gabaix et al. (2016), Kaymak and Poschke (2016), and Hubmer et al. (2017). We extend the analysis to possibly non-stationary distributions in Section 6 as a robustness check. Our preliminary results are encouraging, in the sense that the model seems to be able to capture the transitional dynamics with parameters estimates not too far from those obtained under stationarity.

[12] It is also a reasonable approximation to optimal weighting: an efficient two-step estimation with the optimal weighting matrix produce no relevant changes on estimated parameters nor on fit; see Appendix C.4 for details. See Altonji and Segal (1996) for a justification for the adoption of an identity weighting matrix.

[13] See Appendix A.1-2.

for individual earnings profiles, w^n, and its transition across generations, $P\left(w^n \mid w^{n-1}\right)$. Both the earning process and its transition are taken from data; respectively from the PSID and the federal income tax records studied by Chetty et al. (2014).

ii) We target as moments:

- the bottom 20%, $20-39\%$, $40-59\%$, $60-79\%$, $80-89\%$, $90-94\%$, $95-99\%$, and the top 1% wealth percentiles; and

- the diagonal of the social mobility Markov chain transition matrix defined over the same percentile ranges as states.

iii) We estimate:

- the preference parameters μ, A; and

- a parameterization of the stochastic process for r defined by 5 states r_i and 5 diagonal transition probabilities, $P(r^n = r_i \mid r^{n-1} = r_i)$, $i = 1, \ldots, 5$, restricting instead the 5×5 transition matrix to display constantly decaying off-diagonal probabilities except for the last row for which we assume constant off-diagonal probabilities.[14]

In total, therefore, we target 15 moments and we estimate 12 parameters.

Finally, in Section 4.4 we study the case in which the Markov states of the stochastic process for r depend on the initial wealth a of the agent. This adds one parameter to the estimation.

3.2 Data

Our quantitative exercise requires data for labor earnings, wealth distribution, and social mobility.

[14]Formally, $P(r^n = r_i \mid r^{n-1} = r_j) = P(r^n = r_i \mid r^{n-1} = r_i)e^{-\lambda j}$, $i = 1, 2, 3, 4$, $j \neq i$, λ such that $\sum_{j=1}^{5} P(r^n = r_i \mid r^{n-1} = r_j) = 1$; and $P(r^n = r_5 \mid r^{n-1} = r_j) = \frac{1}{4}\left(1 - P(r^n = r_5 \mid r^{n-1} = r_5)\right)$. We adopt a restricted specification in order to reduce the number of parameters we need to estimate. This particular specification performs better than one with constant off-diagonal probabilities as well as one with decaying off-diagonal probabilities in all rows.

Labor earnings. We use 10 deterministic life-cycle household-level earnings profiles at different deciles, as estimated by Heathcote et al. (2010) from the Panel Study of Income Dynamics (PSID), 1967-2002.[15] These profiles are drawn in Figure 1.[16] In our quantitative exercise we collapse earnings levels into six-year averages, as in Table 1.

Table 1: Life-cycle earnings profiles

Age range / %	0-10	10-20	20-30	30-40	40-50	50-60	60-70	70-80	80-90	90-100
1 [25-30]	9.760	19.95	26.85	33.05	39.02	45.05	51.40	59.16	70.23	100.3
2 [31-36]	11.55	24.01	32.58	40.33	47.70	54.85	65.10	73.06	87.21	138.1
3 [37-42]	12.06	25.20	34.96	43.95	52.42	60.70	69.42	80.37	97.51	169.5
4 [43-48]	12.81	26.42	36.46	45.55	54.37	63.09	72.89	85.09	103.5	182.4
5 [49-54]	11.74	24.66	33.56	42.23	51.18	60.34	70.63	82.78	101.4	183.4
6 [55-60]	8.222	19.08	26.78	34.39	42.96	51.91	61.65	74.35	93.42	180.4

Notes: Earnings are in thousand dollars.

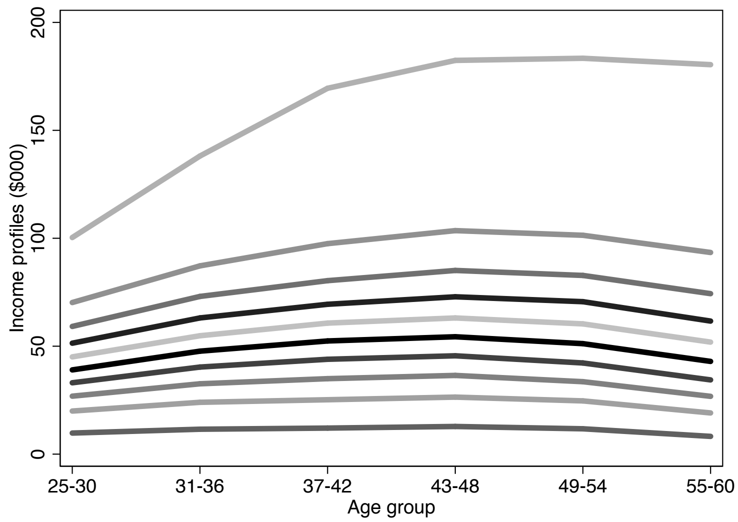
The intergenerationl transition matrix for earnings we use is from Chetty et al. (2014). The data in Chetty et al. (2014) refers to the 1980-82 U.S. birth cohort and their parental income. We reduce it to a ten-state Markov chain.[17]

[15] We detrend life-cycle earning profiles by conditioning out year dummies in a log-earnings regression; see Appendix B.1 for the details of the procedure.

[16] The panel data on earnings from the U.S. Social Security Administration (SSA) are not yet generally available. However, the crucial aspect of earnings data, for our purposes, is that they are far from skewed enough to account by themselves for the skewedness of the wealth distribution. This is in fact confirmed on SSA data directly by Guvenen et al. (2016), Section 7.2.II, and by De Nardi, Fella, and Paz-Pardo (2016); see also Hubmer, Krusell, and Smith (2017).

[17] See Appendix B.2 for details.

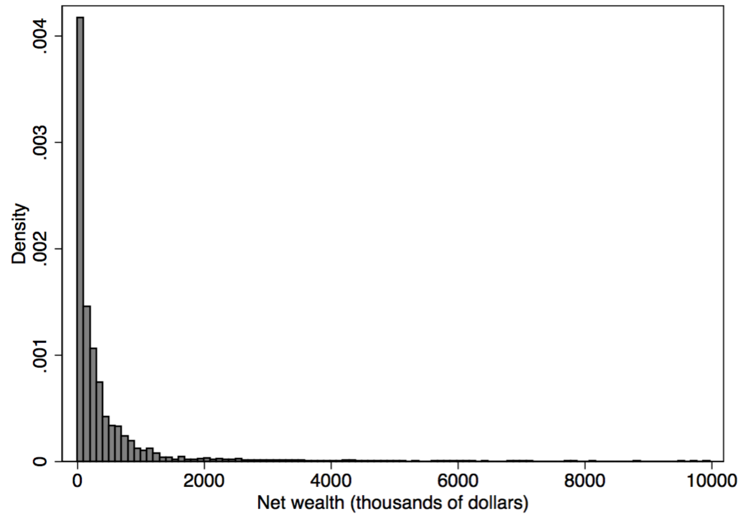
Figure 1: Life-cycle earnings profiles by deciles



Wealth distribution. We use wealth distribution data from the Survey of Consumer Finances (SCF) 2007.[18] The wealth variable we use is *net wealth*, the sum of net financial wealth and housing, minus any debts. The distribution is very skewed to the right. We take the percentile shares from the cleaned version in Díaz-Giménez et al. (2011). Figure 2 displays the histogram of the wealth distribution.

[18]As noted, the wealth distribution in our methodology is to be interpreted as stationary. Choosing 2007 avoids the non-stationary changes due to the Great Recession.

Figure 2: Wealth distribution in the SCF 2007 (weighted)



Notes: Net wealth, from 2007 SCF, truncated at \$0 on the left and at \$10 million on the right.

Table 2 displays the wealth share moments we use.

Table 2: Wealth distribution moments

Share of wealth	0-20	20-40	40-60	60-80	80-90	90-95	95-99	99-100
	-0.002	0.001	0.045	0.112	0.120	0.111	0.267	0.336

Social mobility. As for wealth transition across generations, we estimate an inter-generational mobility matrix from the 2007-2009 SCF 2-year panel as follows. We first construct age-dependent 2-year transition matrices for age groups running from $30-31$ to $66-67$.[19] We then multiply these age-dependent 2-year transition matrices for all age groups, to construct the intergenerational social mobility matrix, which we report in Table 3.[20]

[19] Because of limited sample dimension, we average the left and right matrices obtained using, respectively, the left-middle ages and the middle-right ages to define the age group in the 2-year panel; for instance, the $30-31$ age-group is constructed using the average of the transitions of the $29-30$ and the $31-32$ groups in the data.

[20] Effectively, this construction computes the transition matrix for a synthetic agent over his/her age profile. It accounts for the wealth transitions along the whole working life of agents. As a consequence, it accounts for any transition induced by bequests (as well as in-vivos transfers) the agents receive in this period. This is a defining element of our quantitative strategy, since the model relies importantly on the bequest motive. Indeed this is the reason why we preferred not to adopt the inter-generational social mobility matrix estimated by Charles and Hurst (2003) on PSID data. This matrix is in fact constructed by means

Table 3: Intergenerational social mobility transition matrix

Percentile	0-20	20-40	40-60	60-80	80-90	90-95	95-99	99-100
0-20%	0.223	0.222	0.215	0.187	0.081	0.038	0.029	0.006
20-40%	0.221	0.220	0.215	0.188	0.082	0.039	0.029	0.006
40-60%	0.208	0.209	0.210	0.194	0.090	0.046	0.036	0.008
60-80%	0.199	0.201	0.207	0.198	0.095	0.052	0.040	0.009
80-90%	0.175	0.178	0.197	0.207	0.110	0.067	0.054	0.012
90-95%	0.182	0.184	0.200	0.205	0.106	0.062	0.050	0.011
95-99%	0.125	0.125	0.166	0.216	0.141	0.114	0.094	0.021
99-100%	0.086	0.084	0.142	0.228	0.170	0.143	0.121	0.028

It displays substantial social mobility: the Shorrocks mobility index is .98.[21]

4 Estimation results

The baseline estimation results are reported in Section 4.1, Table 4. The targeted simulated moments of the estimated model are reported and compared to their counterpart in the data in Section 4.2, Table 5. Some independent evidence which bears on the fit of the model is discussed in Section 4.3.

of pairs of simultaneously alive parent and child and, as a consequence, cannot account for bequests.

[21]Formally, for a square mobility transition matrix A of dimension m, the Shorrocks index given by $s(A) = \frac{m-\sum_j a_{jj}}{m-1} \in (0,1)$, with 0 indicating complete immobility. By construction, mobility matrices have Shorrock indexes increasing as the transition step gets long (indeed the index converges to 1 as the step goes to ∞). This explains in part the high index associated to the inter-generational matrix we construct. Also, and most importantly, measurement error in wealth can by itself induce spurious mobility in the matrix; Jappelli and Pistaferri (2006) discuss this issue with regards to consumption mobility and account explictly for measurement error in the analysis. We leave this for a future analysis of social mobility per se. The qualitative properties of social mobility we obtain are similar to those we obtain from Kennickell and Starr-McCluer (1997)'s 6-years transition matrix from SCF (1983-89). In this case, the inter-generational (36-years) matrix is constructed by raising the 6-years matrix to the $6-th$ power; see .3 for details. Our method, besides using more recent data, exploits the more precise information contained in age-dependent transitions. Similar properties also hold when adopting Klevmarken et al. (2003) and Charles and Hurst (2003) estimates with the PSID data (though Charles and Hurst (2003) mobility matrix displays moderatly less mobility, with a Shorrocks index of .88); see Appendix B.3 where these matrices are reported.

4.1 Parameter estimates

The upper part of Table 4 reports the estimates of the preference parameters. The lower part of Table 4 reports the estimated state space and diagonal of the transition matrix of the 5-state Markov process for r we postulate. It also reports, to ease the interpretation of the estimates, the implied mean and standard deviation of the process, $E(r)$, $\sigma(r)$; as well as its auto-correlation, $\rho(r)$, computed fitting an $AR(1)$ on simulated data from the estimated process.[22]

Table 4: Parameter estimates: Baseline

	Preferences				
	σ	μ	A	β	T
	[2]	0.5653	0.0004	[0.97]	[36]
		(0.0260)	(0.0002)		
	Rate of return process				
state space	0.0010	0.0087	0.0253	0.0532	0.0850
	(0.0001)	(0.0013)	(0.0019)	(0.0123)	(0.0062)
transition diagonal	0.0224	0.2698	0.1371	0.2746	0.0224
	(0.3189)	(0.6096)	(0.0710)	(0.1463)	(0.2672)
statistics	$\mathbb{E}(r)$	$\sigma(r)$	$\rho(r)$		
	3.00%	2.68%	0.1751		
	(0.85%)	(0.51%)	(0.1656)		

Notes: Standard errors in (); fixed parameters in [].

The curvature parameter μ is statistically significant and so is the bequest intensity parameter A (though barely). As for the rate of return process r, each element of the state space is very precisely estimated; and while the parameters of the transition diagonal, individually taken, are statistically insignificant, most importantly, the mean $\mathbb{E}(r)$ and the variance $\sigma(r)$ of the rate of return process are. The correlation $\rho(r)$ is not surprisingly imprecisely estimated (because the transition matrix is in-and-of itself imprecisely estimated and because the auto-correlation parameter is not a statistics pertaining directly to the r process but is estimated by fitting an AR(1) process on simulated data). A Quandt

[22]The full transition matrix for r is reported in Appendix C.1. The standard errors, also reported in the Table, are obtained by bootstrapping; details are in Appendix A.3.

Likelihood Ratio (QLR) test against the null hypothesis that the rate of return process is a constant r squarely rejects the null.

4.2 Model fit

The simulations of our estimated model seem to capture the targeted moments reasonably well. Table 5 compares the moments in the data with those obtained simulating the model.

Table 5: Model fit: Baseline

	0-20	20-40	40-60	60-80	80-90	90-95	95-99	99-100
	Wealth distribution							
Data	-0.002	0.001	0.045	0.112	0.120	0.111	0.267	0.336
Model	0.047	0.074	0.107	0.102	0.105	0.071	0.155	0.339
	Social mobility							
Data	0.223	0.220	0.210	0.198	0.110	0.062	0.094	0.028
Model	0.228	0.207	0.200	0.193	0.102	0.048	0.047	0.036

The simulated wealth distribution is less skewed than the data's: too much wealth is concentrated in the bottom, especially the bottom 40%. This is in part due to the borrowing constraint necessarily inducing non-negative wealth holdings throughout the agents' lifetime. A more detailed modeling of financial markets than possible in a parsimonious specification like ours would possibly improve on this dimension. Most importantly, however, we match rather precisely the top 1% share, the moment which previous literature has found hardest to match; see the discussion in Benhabib and Bisin (2015). Furthermore, while we underestimate the $90-99\%$ share, we will see in Section 4.4 that allowing the return process r to depend on wealth improves somewhat our fit on this margin. Finally, we match quite accurately the social mobility moments we target (the diagonal of the social mobility matrix); in the top 10% of the distribution, we over-estimate the probability of staying in the top 1% but under-estimate the probability of staying in the $90-99\%$ percentile.

4.3 Discussion and interpretation

We discuss and interpret here the estimates we obtain. We also put them in the context of independent evidence which bears on non-targeted moments regarding savings, bequests, rates of return, and wealth mobility.

Differential savings and bequests. Our estimates point to the existence of the differential saving factor as a component of the observed wealth dynamics in the U.S. Indeed, our estimate of μ is 0.5653, which is significantly lower than 2, the value of σ we fixed; therefore $\mu < \sigma$ and, as we noted, savings out of wealth increase with wealth itself: the rich save proportionally more than the poor.

Of course, the strength of this factor depends on the intensity parameter A as well.To better evaluate the quantitative role of differential savings and bequests in our estimation, we calculate the average savings rates implied by our model at the estimated parameters and compare them with the empirical values calculated by Saez and Zucman (2016) using 2000-2009 data on wealth accumulation with the capitalized income tax method; see Table 6. Interestingly, the implied (year-to-year synthetic) savings rate schedule shares its main characteristic feature with the one reported by Saez and Zucman (2016): it is very steep (even steeper in fact) - rates range from slightly negative (-3% of the bottom 90%) to 51% for the top 1% of the population.

Table 6: Savings rates

	0-90	90-99	99-100
Our estimates	-3%	22%	51%
Saez and Zucman (2016)	-4%	9%	35%

To gain a more precise sense of the mechanism driving differential savings, we also look at bequests, since in our model differential savings are mostly motivated by a bequest motive.[23] The distribution of bequests implied by our model at the estimated parameters is very

[23] The bequest motive stands on relative solid grounds: it is well documented that retirees do not run down their wealth as predicted by the classical life-cycle consumption-savings model (Poterba et al., 2011).

skewed, mapping closely the stationary wealth distribution. This is consistent with Health Retirement Survey (HRS) data studied by Hurd and Smith (2003). In particular, retirement savings in the data do not decline along the age path and, furthermore, this pattern is more accentuated for the 75% percentile, as our estimates also imply.[24] Bequests implied by the model are about 13% of GDP, higher than its empirical counterpart: Wang (2016) estimates them to be between 2.4% to 4.7% of GDP, using the HRS data; see also Hendricks (2002). On the other hand bequests in the model should more correctly be interpreted to include at least part of inter-vivos transfers, which can account for the remaining 10% of GDP. Indeed, Cox (1990) and Gale and Scholz (1994) estimate inter-vivos tranfer to be about the same order of magnitude as bequests, while Luo (2017), working with SCF (2013) data, has them close to 13% of GDP.

Returns to wealth. The wealth accumulation process in our estimates indicates a substantial role of capital income risk as a factor driving wealth and mobility. Indeed the rate of return on wealth displays a standard deviation which is significantly different than 0. The standard deviation $\sigma(r) = 2.68\%$ is however smaller than previous direct estimates. This is the case, e.g., for the return estimates by Case and Shiller (1989) and Flavin and Yamashita (2002) on the housing market, by Campbell and Lettau (1999), Campbell et al. (2001) on individual stocks of publicly traded firms, and by Moskowitz and Vissing-Jørgensen (2002) on private equity and entrepreneurship. A wide dispersion in returns to wealth is also documented by Fagereng et al. (2017) and Bach et al. (2017) using, respectively Norwegian and Swedish data.

Such comparisons require however great caution. First of all, in our model, r is assumed constant throughout each agent's lifetime, disregarding the whole variation across the life-cycle. The rate of return we estimate should ideally be then compared with the permanent

[24]Our model does not have a role for accidental bequests. Therefore, while the literature on retirement savings distinguishes between precautionary saving motives for uncertain medical expenses (De Nardi et al., 2010), uncertain and potentially large long-term care expenses (Ameriks et al., 2015a), family needs (Ameriks et al., 2015b) and the genuine bequest motive, we necessarily lump all these into aggregate bequests.

Table 7: Rate of return process

Statistics	$E(r)$	$\sigma(r)$	$\rho(r)$
Our estimates	3.00%	2.68%	0.17
Fagereng et al. (2017)	2.98%	2.82%	0.1

Notes: Fagereng et al. (2017)'s permanent component has zero-mean by construction: we report their mean of returns.

components of individual returns across generations, which are hardly available. Furthermore, rate of returns heterogeneity in the data is in part a consequence of differences in the risk composition of investment portfolio, which also we disregard in the model; see Calvet and Sodini (2014) and Bach et al. (2017) for evidence in Swedish data. For our purposes, therefore, the most appropriate outside validation perspective is provided by Fagereng et al. (2017), in that their Norwegian administrative data allows them to estimate the permanent components of individual returns across generations and to control for portfolio composition. In this comparison, the consistency of our estimates with the Fagereng et al. (2017)'s data is striking; see Table 7.[25]

Social mobility. The implied non-targeted moments (the off-diagonal cells) of the social mobility matrix we obtain align quite well with the mobility we constructed from the SCF data.[26] With regards to the inter-generational flows from the top 1% of the distribution of wealth (the last row of the mobility matrix), however, our model over-estimates the probability of staying in the top 1%, as we already noted, but compensates this by generally overestimating the churn: e.g., the probability that children of parents in the top 1% move to the bottom 40% is 28.5% in the estimated model but 17% in the data.

[25] Fagereng et al. (2017) also find rate of returns increasing in wealth. We shall discuss this in the next section.

[26] We report the whole mobility matrix we obtain in Appendix C.3

4.4 Rate of return dependent on wealth

A positive correlation between the rate of return on wealth and wealth has been documented by Piketty (2014)'s analysis of university endowments, see especially p. 447, and by Fagereng et al. (2017)'s careful study of Norwegian administrative data.[27] Such a correlation of course does not imply that the rate of return increases with wealth. Even in the context of our model, agents with relatively high wealth would have experienced on average high realizations of the rate of return r. Indeed, for the simulated model at the parameters estimates in the previous section, a fractile regression between r and wealth a produces a small but strongly significant coefficient of .01 (standard error .0004).

Allowing rates of return on wealth to be increasing in wealth might however add to the skewedness of the distribution. In this section we therefore extend our analysis to allow for the rate of return process r to depend on wealth, explicitly introducing a dependence of the stochastic rate of return r on wealth percentiles. The functional form we introduce allows for r to depend on wealth a as follows:

$$r = r_0 + b \times p(a) \tag{1}$$

where $p(a) = 1, 2, \ldots, 8$ numbers the wealth percentiles we identify as moments and r_0 is a 5-state Markov process as in the baseline model for r. Note that this formulation maps a positive slope b into a convex relationship between r and a.[28] We then estimate the parameters of our model as well as the wealth dependence parameter b that enters the stochastic rate of return process. The results of our estimation are reported in Tables 8 and 9.

[27]See also Kacperczyk et al. (2015). On the other hand no correlation is apparent in Saez and Zucman (2016). Also, Bach et al. (2017) find that the correlation is largely due, in the Swedish administrative data they observe, to the portfolio composition by risk class changing with wealth.

[28]This formulation implies a standard deviation for r which is increasing in wealth, as documented by Fagereng et al. (2017) for Norwegian data.

Table 8: Parameter estimates: r dependendent on wealth

	Preferences				
	σ	μ	A	β	T
	[2]	0.6019	0.0002	[0.97]	[36]
		(0.0791)	(0.0003)		
	Rate of return process				
state space	0.0005	0.0106	0.0213	0.0501	0.0584
	(0.0015)	(0.0000)	(0.0153)	(0.0309)	(0.0012)
transition diagonal	0.1531	0.4552	0.1621	0.0314	0.0192
	(0.5653)	(0.0847)	(0.2001)	(0.0637)	(0.3461)
wealth dependence, b	0.008				
	(0.1389)				
statistics	$\mathbb{E}(r_0)$	$\sigma(r_0)$	$\rho(r_0)$	$\mathbb{E}(r)$	$\sigma(r)$
	2.26%	1.96%	0.2109	3.91%	2.08%
	(0.62%)	(0.33%)	(0.1641)		

Notes: Standard errors in (); fixed parameters in [].

Table 9: Model fit; r dependent of wealth

	0-20	20-40	40-60	60-80	80-90	90-95	95-99	99-100
	Wealth distribution							
Data	-0.002	0.001	0.045	0.112	0.120	0.111	0.267	0.336
Model	0.028	0.067	0.101	0.108	0.113	0.080	0.164	0.340
	Social mobility							
Data	0.223	0.220	0.210	0.198	0.110	0.062	0.094	0.028
Model	0.231	0.207	0.224	0.200	0.095	0.060	0.043	0.054

The estimate of the parameter b, which captures the dependence of the rate of return on wealth is positive. The point estimate implies that going from the bottom 20% to the top 1% in the wealth distribution would increase the annual return by 5.6 percentage points, from 3% to 8.6%. While b is unsurprisingly not well-identified, it is reassuring that the point estimates of the preference parameters are not much changed when we allow for r to depend on wealth with respect to the baseline. Furthermore, the fit of the wealth distribution is somewhat improved: while the distribution of wealth implied by the model is still less skewed than the data's, we continue to match precisely the top 1% share and, most importantly, we improve in matching all shares in the top 40%, though somewhat marginally. With regards to social mobility, this specification loses fit on the top 1%, as the rich face substantially higher rates of return and hence remain even more stably in the top wealth brackets, but otherwise the fit is maintained.

Fagereng et al. (2017) also estimate the dependence of the rate of return r on wealth, their rich and detailed Norwegian data set allowing them to do so precisely, directly controlling for the effects of a variety of factors like age, education and portfolio composition. Their findings provide stronger evidence of dependence than ours, with average returns within generations significantly increasing in wealth. Orders of magnitude are once again strikingly close, with returns in Fagereng et al. (2017) ranging from about 3% to 8% going from lower to higher wealth percentiles; see Figure 11(b).

5 Counterfactual estimates

In this section we perform a set of counterfactual estimations of the model, under restricted conditions. More in detail, we perform three sets of counterfactuals, corresponding to shutting down each of the three main factors which can drive the distribution of wealth: (1) capital income risk, (2) differential savings rates, and (3) stochastic earnings.

The objective of this counterfactual analysis is twofold. First of all we aim at gauging the relative importance of the various mechanisms we identified as possibly driving the distribution of wealth. In particular, we aim at a better understanding of which mechanism mostly affects which dimension of the wealth distribution and mobility. Second, we interpret the counterfactuals as informal tests of identification of these mechanisms, lack of identification implying that shutting down one or more of the mechanism has limited effects on the fit for the targeted moments.

5.1 Re-estimation results

We examine the counterfactual estimates in detail in the following. The estimated parameters are in Table 10.[29] Table 11 reports the fit of the estimates.

In the counterfactual with no stochastic earnings we feed the model an average earnings path. The resulting estimates of the preference parameters and of the rate of return process r reveal a minor strengthening of the savings factor (through an increase in A, though compensated by an increase in μ as well) and especially of capital income risk (both the mean and the auto-correlation of r are increased, while the standard deviation is slighly smaller). Interestingly, in this case the estimate does not miss much in matching the top 1% of the wealth distribution but misses more dramatically the $95-99\%$ and overall produces a distribution which is much less skewed than the baseline's (and the data's), displaying more mass on the bottom 80%. This is an indication that stochastic earnings is not as much relevant a factor in filling the tail of the wealth distribution, but much more in producing churning, facilitating the escape from low levels of wealth close to the borrowing constraint and from the top as well. This is confirmed in the match of social mobility: with no stochastic earnings, notably, the top 1% has a dramatically higher probability of staying in the same cell than in the baseline (20.5% vs. 2.8% in the baseline) and so does the bottom 20% (31.6% vs. 22.8%).

[29]We report only the mean, standard deviation and auto-correlation statistics for r, to save space. The estimates for the state space and the diagonal of the transition matrix are in Appendix C.2

Table 10: Parameter estimates: Counterfactuals

	Preferences				
	σ	μ	A	β	T
baseline	[2]	0.5653	0.0004	[0.97]	[36]
		(0.0260)	(0.0002)		
constant r	[2]	0.6008	0.0008	[0.97]	[36]
		(0.0780)	(0.1292)		
constant w	[2]	0.9232	0.0073	[0.97]	[36]
		(0.0037)	(0.0002)		
$\mu = 2$	[2]	2	0.0001	[0.97]	[36]
		-	(0.0001)		
	Rate of return process				
	$\mathbb{E}(r)$	$\sigma(r)$	$\rho(r)$		
baseline	3.00%	2.68%	0.1751		
	(0.85%)	(0.51%)	(0.1656)		
constant r	3.01%				
	(0.02%)				
constant w	3.39%	2.59%	0.1814		
	(0.82%)	(0.38%)	(0.1517)		
$\mu = 2$	2.78%	2.71%	0.2186		
	(0.86%)	(0.72%)	(0.1588)		

Notes: Standard errors in (); fixed parameters in [].

In the counterfactual with homogeneous saving rates, we set $\mu = 2$, that is, we set the curvature parameter of the bequest utility equal to the curvature of consumption utility, so that agents with different wealth save at the same rate. In terms of the estimates, preferences for bequests are greatly reduced: most of the action is left to capital income risk, via the larger auto-correlation of r. In this case, contrary to the previous counterfactual with no stochastic earnings, the model misses to match the top 1% of the wealth share, which is greatly reduced. Apart from the last percentile, the simulated wealth distribution is less skewed than the data, as usual. With respect to social mobility, it is noteworthy that homogeneous savings induces higher mobility out of the top 1% share, indicating an important trade-off in the role of differential savings: it contributes fundamentally to fill the tail of the wealth distribution at the cost of substantially reducing social mobility.

Table 11: Model fit: Counterfactuals

	0-20	20-40	40-60	60-80	80-90	90-95	95-99	99-100
				Wealth distribution				
Data	-0.002	0.001	0.045	0.112	0.120	0.111	0.267	0.336
Model								
(1) Baseline	0.047	0.074	0.107	0.102	0.105	0.071	0.155	0.339
(2) Constant r	0.068	0.107	0.150	0.219	0.160	0.110	0.119	0.068
(3) Constant w w	0.119	0.131	0.149	0.154	0.094	0.055	0.051	0.247
(4) $\mu = 2$	0.070	0.112	0.161	0.229	0.159	0.104	0.123	0.042
				Social mobility				
Data	0.223	0.220	0.210	0.198	0.110	0.062	0.094	0.028
Model								
(1) Baseline	0.228	0.207	0.200	0.193	0.102	0.048	0.047	0.036
(2) Constant r	0.204	0.234	0.201	0.135	0.163	0.058	0	0.036
(3) Constant w	0.316	0.161	0.192	0.141	0.214	0.084	0.096	0.205
(4) $\mu = 2$	0.228	0.216	0.200	0.198	0.099	0.043	0.049	0.025

In the counterfactual with no capital income risk, we re-estimate the model under the constraint that the rate of return is constant. The estimate of the rate of return we obtain in this case is 3.01%, just above its mean in the baseline. Though in our baseline estimate the implied savings rate is already too high (see Table 6), the differential savings factor compensates the lack of capital income risk to produce some skewedness in the wealth distribution and hence this counterfactual estimate produces a much higher bequest motive (associated to an even more excessive savings rate): while μ is essentially unchanged, the estimated relative preference for bequests A is doubled. Nonetheless, the estimate with r constant dramatically misses in matching the top 1% of the wealth distribution, which is reduced to about 1/5 of the baseline (and the data). More generally, the wealth distribution implied by the model is much less skewed, displaying a smaller fraction of wealth concentrated on top 5%, which is shifted to the whole rest of the distribution (but especially in the middle $40 - 90\%$). In terms of social mobility, restricting the estimate to a constant r does not remarkably change the fit on social mobility: it still produces too small a probability of transition away from the top 1% though it slightly accentuates the excess churning in the top 10% (especially in the $95 - 99\%$ cell).

In summary, all the factors we study in our quantitative analysis, stochastic earnings, differential savings, and capital income risk, are well-identified as crucial for generating the thick right tail of the wealth distribution and sufficient mobility. The factors seems to have a distinct role. Capital income risk and differencial savings both contribute in a fundamental manner to generating the thick tail. But while differential savings at the same time reduces social mobility, capital income risk appears not to have major role in affecting mobility. On the other hand, stochastic earnings have a limited role in filling the tail of the wealth distribution but are fundamental in inducing enough mobility in the the wealth process, both by limiting poverty traps at the bottom and favoring the churn at the top.[30]

6 Transitional dynamics of the wealth distribution

Our quantitative analysis so far is predicated on the assumption that the observed distribution of wealth is a stationary distribution, in the sense that our estimates are obtained by matching the data with the moments of the stationary distribution generated by the model. In this section we instead begin studying the implications of our model for the transitional dynamics of the distribution of wealth.

The exercise we perform is as follows: using the observed SCF 1962-1963 distribution

[30]In apparent contrast with our results, several previous papers in the literature have obtained considerable success in matching the wealth distribution in the data with simulated models fundamentally driven by the stochastic earnings mechanism; see e.g., Castañeda et al. (2003), Díaz et al. (2003), Dávila et al. (2012), Kindermann and Krueger (2015), Kaymak and Poschke (2016). These simulated models however appear driven by extreme assumptions either about the skewness of earnings (adding an *awesome* state) or about the working life of agents. For instance, Díaz et al. (2003) postulate an earning process where roughly 6% of the top earners have 46 times the labor endowment of the median (in the World Top Income Database 2013-14 the average income of the top 5% is no more that 7.5 times the median income); Kaymak and Poschke (2016)'s calibration implies a working life-span of over 100 years, at the stationary distribution, for 11% of the working population. See Benhabib et al. (2017) and Benhabib and Bisin (2017) for detailed discussions of these issues.

of wealth as initial condition,[31] we estimate the parameters of the model by matching the implied distribution after 72 years (two iterations of the model) with the observed SCF 2007 distribution and the transition matrix adopted in the previous quantitative analysis.[32]

The fundamental feature of the change in the wealth distribution from 1962-1963 to 2007, in our data, is the substantial increase in inequality. The top 1% share, for instance, goes from 24.2% to 33.6%; the top 5% from 43.2% to 60.3%. In this respect, our new estimate shows that such a dramatic increase in wealth inequality can be obtained within the confines of our simple model, by exploiting the explanatory power of capital income risk and differential savings; see Gabaix et al. (2016) for related results. The new parameter estimates we obtain show in fact a larger bequest motive (a larger A, though compensated by a larger μ), with respect to their counterparts in the benchmark model, and a rate of return process with higher mean and volatility. This induces a simulated distribution of wealth for 2007 which, with respect to the data, is even more skewed at the top. Strikingly, the bottom 40% of the distribution is very well matched, better than in our baseline. All in all, the match in this exercise is quite successful and the skewedness of the simulated distribution more closely matches the data than even our baseline. This is obtained at the cost of not matching well the social mobility, by severely undestimating mobility, that is, underestimating the probability that children move away from their parents' wealth cell.

[31] More precisely, these data is from precursor surveys to the SCF: the 1962 Survey of Financial Characteristics of Consumers and the 1963 Survey of Changes in Family Finances. See $http://www.federalreserve.gov/econresdata/scf/scf_6263.htm.$ for a discussion. Differences in methodology and quality notwithstanding, these data provide a useful benchmark as initial condition to the recent wealth dynamics.

[32] While the analysis does not require nor imposes any stationarity of the distribution of wealth over time, it does postulate that the model structure and parameter values stay constant after 1962. Importantly, we do not feed in the analysis the observed fiscal policy reforms since the $'60's$. Doing so should improve the fit.

Table 12: Parameter estimates: Transitional dynamics

	Preferences				
	σ	μ	A	β	T
	[2]	1.4291	0.0177	[0.97]	[36]
		(0.0169)	(0.0043)		
	Rate of return process				
state space	0.0011	0.0155	0.0182	0.0677	0.0990
	(0.0001)	(0.0082)	(0.0082)	(0.0026)	(0.0088)
transitional diagonal	0.0909	0.2643	0.2587	0.1401	0.0105
	(0.4308)	(1.1170)	(0.2116)	(0.1108)	(0.0119)
statistics	$\mathbb{E}(r)$	$\sigma(r)$	$\rho(r)$		
	3.29%	3.19%	0.1296		
	(1.00%)	(0.62%)	(0.1500)		

Notes: Standard errors in (); fixed parameters in [].

Table 13: Model fit: Transitional dynamics

	0-20	20-40	40-60	60-80	80-90	90-95	95-99	99-100
Data: SCF 1962-63	0.009	0.043	0.094	0.173	0.142	0.115	0.190	0.242
Data: SCF 2007	-0.002	0.001	0.045	0.112	0.120	0.111	0.267	0.336
Model	0.000	0.001	0.010	0.045	0.115	0.179	0.297	0.354
	Social mobility							
Data	0.223	0.220	0.210	0.198	0.110	0.062	0.094	0.028
Model	0.663	0.535	0.493	0.494	0.297	0.175	0.271	0.168

7 Conclusions

We estimate a parsimonious macroeconomic model of the distribution of wealth in the U.S. While we assign special emphasis on the tail of the distribution, the model performs rather well in fitting the whole distribution of wealth in the data. Importantly, the model is also successful in fitting the social mobility of wealth in the data. Parameter estimates, especially the rate of return of wealth process, compare very closely to independent observations.

Our analysis allows us to distinguish the contributions of three critical factors driving wealth accumulation: a skewed and persistent distribution of earnings, differential saving and bequest rates across wealth levels, and capital income risk in entrepreneurial activities. All of these three factors are necessary and empirically relevant in matching both distribution and mobility, with a distinct role for each, which we identify.

Finally, we begin studying the implications of the model for the transitional dynamics of the distribution of wealth. The estimates are qualitatively similar to those in the baseline, and our model delivers fast transitional dynamics. While more work is obviously necessary in this respect, our results are quite encouraging.

References

Altonji, Joseph G and Lewis M Segal, "Small-Sample Bias in GMM Estimation of Covariance Structures," *Journal of Business & Economic Statistics*, July 1996, *14* (3), 353–366.

Ameriks, John, Joseph Briggs, Andrew Caplin, Matthew D. Shapiro, and Christopher Tonetti, "Long-Term Care Utility and Late in Life Saving," 2015.

_ , _ , _ , Mi Luo, Matthew D. Shapiro, and Christopher Tonetti, "Measuring and Modeling Inter-generational Wealth Transfers: the Precautionary Transfer Motives," 2015.

Bach, Laurent, Laurent E. Calvet, and Paolo Sodini, "Rich Pickings? Risk, Return, and Skill in the Portfolios of the Wealthy," 2017.

Benhabib, Jess, Alberto Bisin, and Mi Luo, "Earnings Inequality and Other Determinants of Wealth Inequality," *American Economic Review: Papers & Proceedings*, 2017, *107* (5), 593–597.

_ , _ , and Shenghao Zhu, "The Distribution of Wealth and Fiscal Policy in Economies With Finitely Lived Agents," *Econometrica*, 01 2011, *79* (1), 123–157.

_ and _ , "Skewed Wealth Distributions: Theory and Empirics," *Journal of Economic Literature*, 2017.

Cagetti, Marco and Mariacristina De Nardi, "Entrepreneurship, Frictions, and Wealth," *Journal of Political Economy*, October 2006, *114* (5), 835–870.

Calvet, Laurent E. and Paolo Sodini, "Twin Picks: Disentangling the Determinants of Risk-Taking in Household Portfolios," *Journal of Finance*, 04 2014, *69* (2), 867–906.

Campbell, John Y. and Martin Lettau, "Dispersion and Volatility in Stock Returns: An Empirical Investigation," NBER Working Papers 7144, National Bureau of Economic Research, Inc May 1999.

_ , _ , Burton G. Malkiel, and Yexiao Xu, "Have Individual Stocks Become More Volatile? An Empirical Exploration of Idiosyncratic Risk," *The Journal of Finance*, 2001, *56* (1), 1–43.

Case, Karl E. and Robert J. Shiller, "The Efficiency of the Market for Single-Family Homes," *American Economic Review*, March 1989, *79* (1), 125–37.

Castañeda, Ana, Javier Díaz-Gimínez, and José-Víctor Ríos-Rull, "Accounting for the U.S. Earnings and Wealth Inequality," *Journal of Political Economy*, August 2003, *111* (4), 818–857.

Charles, Kerwin Kofi and Erik Hurst, "The Correlation of Wealth across Generations," *Journal of Political Economy*, December 2003, *111* (6), 1155–1182.

Chetty, Raj, Nathaniel Hendren, Patrick Kline, and Emmanuel Saez, "Where is the land of Opportunity? The Geography of Intergenerational Mobility in the United States," *The Quarterly Journal of Economics*, 2014, *129* (4), 1553–1623.

Cox, Donald, "Intergenerational Transfers and Liquidity Constraints," *Quarterly Journal of Economics*, 1990, *105*, 187–217.

Cunha, Flavio, James J. Heckman, and Susanne M. Schennach, "Estimating the Technology of Cognitive and Noncognitive Skill Formation," *Econometrica*, 05 2010, *78* (3), 883–931.

Dávila, Julio, Jay H. Hong, Per Krusell, and José-Víctor Ríos-Rull, "Constrained Efficiency in the Neoclassical Growth Model With Uninsurable Idiosyncratic Shocks," *Econometrica*, November 2012, *80* (6), 2431–2467.

De Nardi, Mariacristina, Eric French, and John B. Jones, "Why Do the Elderly Save? The Role of Medical Expenses," *Journal of Political Economy*, 02 2010, *118* (1), 39–75.

De Nardi, Mariacristina, Giulio Fella, and Gonzalo Paz-Pardo, "The Implications of Richer Earnings Dynamics for Consumption and Wealth," 2016.

de Saporta, Benoîte, "Tail of the stationary solution of the stochastic equation with Markovian coefficients," *Stochastic Processes and their Applications*, 2005, *115* (12), 1954 – 1978.

Díaz, Antonia, Josep Pijoan-Mas, and José-Victor Ríos-Rull, "Precautionary savings and wealth distribution under habit formation preferences," *Journal of Monetary Economics*, September 2003, *50* (6), 1257–1291.

Díaz-Giménez, Javier, Andrew Glover, and José-Víctor Ríos-Rull, "Facts on the distributions of earnings, income, and wealth in the United States: 2007 update," *Quarterly Review*, 2011.

Fagereng, Andreas, Luigi Guiso, Davide Malacrino, and Luigi Pistaferri, "Heterogeneity in Returns to Wealth and the Measurement of Wealth Inequality," *American Economic Review: Papers & Proceedings*, 2016, *106* (5), 651–655.

_ , _ , _ , and _ , "Heterogeneity and Persistence in Returns to Wealth," 2017.

Flavin, Marjorie and Takashi Yamashita, "Owner-Occupied Housing and the Composition of the Household Portfolio," *American Economics Review*, 2002, pp. 345–362.

Gabaix, Xavier, Jean-Michel Lasry, Pierre-Louis Lions, and Benjamin Moll, "The Dynamics of Inequality," *Econometrica*, 2016, *84* (6), 2071–2111.

Gale, William G. and John K. Scholz, "Intergenerational Transfers and the Accumulation of Wealth," *The Journal of Economic Perspectives*, 1994, *8* (4), 145–160.

Grey, D. R., "Regular Variation in the Tail Behaviour of Solutions of Random Difference Equations," *Ann. Appl. Probab.*, 02 1994, *4* (1), 169–183.

Guvenen, Fatih, *Quantitative Economics with Heterogeneity: An A-to-Z Guidebook*, Princeton University Press, 2016.

_ , Fatih Karahan, Serdar Ozkan, and Jae Song, "What Do Data on Millions of U.S. Workers Reveal about Life-Cycle Earnings Dynamics?," 2016.

Hay, Diana, Reza Rastegar, and Alexander Roitershtein, "Multivariate linear recursions with Markov-dependent coefficients," *Journal of Multivariate Analysis*, 2011, *102* (3), 521 – 527.

Heathcote, Jonathan, Fabrizio Perri, and Giovanni L. Violante, "Unequal We Stand: An Empirical Analysis of Economic Inequality in the United States: 1967-2006," *Review of Economic Dynamics*, January 2010, *13* (1), 15–51.

Hendricks, Lutz, "Bequests and Retirement Wealth in the United States," 2002.

Hubmer, Joachim, Per Krusell, and Anthony A. Smith, "The historical evolution of the wealth distribution: A quantitative-theoretic investigation," 2017.

Huggett, Mark, Gustavo Ventura, and Amir Yaron, "Sources of Lifetime Inequality," *American Economic Review*, December 2011, *101* (7), 2923–54.

Hurd, Michael and James P. Smith, "Expected Bequests and Their Distribution," Working Papers 03-10, RAND Corporation Publications Department April 2003.

Jappelli, Tullio and Luigi Pistaferri, "Intertemporal Choice and Consumption Mobility," *Journal of the European Economic Association*, 2006, *4* (1), 75–115.

Kacperczyk, Marcin, Jaromir Nosal, and Luminita Stevens, "Investor Sophistication and Capital Income Inequality," 2015.

Kaplan, Greg, Benjamin Moll, and Giovanni L. Violante, "Monetary Policy According to HANK," 2016.

Kaymak, Bariş and Markus Poschke, "The evolution of wealth inequality over half a century: The role of taxes, transfers and technology," *Journal of Monetary Economics*, 2016, *77*, 1 – 25. Inequality, Institutions, and Redistributionheld at the Stern School of Business,New York University,April 24-25, 2015.

Keane, Michael P. and Kenneth I. Wolpin, "The Career Decisions of Young Men," *Journal of Political Economy*, June 1997, *105* (3), 473–522.

Kennickell, Arthur B. and Martha Starr-McCluer, "Household Saving and Portfolio Change: Evidence from the 1983-89 SCF Panel," *Review of Income and Wealth*, December 1997, *43* (4), 381–99.

Kindermann, Fabian and Dirk Krueger, "High Marginal Tax Rates on the Top 1%? Lessons from a Life Cycle Model with Idiosyncratic Income Risk," 2015.

Klevmarken, Anders, Joseph P. Lupton, and Frank P. Stafford, "Wealth Dynamics in the 1980s and 1990s: Sweden and the United States," *Journal of Human Resources*, 2003, *38* (2).

Krusell, Per and Anthony A. Smith, "Income and Wealth Heterogeneity in the Macroeconomy," *Journal of Political Economy*, October 1998, *106* (5), 867–896.

Luo, Mi, "Unintended Effects of Estate Taxation on Wealth Inequality," 2017.

Moskowitz, Tobias J. and Annette Vissing-Jørgensen, "The Returns to Entrepreneurial Investment: A Private Equity Premium Puzzle?," *American Economic Review*, September 2002, *92* (4), 745–778.

Piketty, Thomas, *Capital in the 21st Century*, Cambridge, MA: Harvard University Press, 2014.

Poterba, James, Steven Venti, and David Wise, "The Composition and Drawdown of Wealth in Retirement," *Journal of Economic Perspectives*, Fall 2011, *25* (4), 95–118.

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

Saez, Emmanuel and Gabriel Zucman, "Wealth Inequality in the United States since 1913: Evidence from Capitalized Income Tax Data," *The Quarterly Journal of Economics*, 2016.

Wang, Kang, "Bequests in the US: Patterns, Motives, and Tax Policy," 2016.

Supplemental Appendix

A. Methods

A.1 Numerical solution

We solve the model for value functions and policy functions with the *Collocation method* in Miranda and Fackler (2004).

A.1.1 Problem

Each agent's recursive problem in the baseline case is

$$\begin{aligned}
V_t(a,r,w) &= \max_c \mathbf{1}\{t<T\}\{u(c)+\beta V(a',r,w,t+1)\}+\mathbf{1}\{t=T\}\{u(c)+e(a')\} \\
s.t. & \\
a' &= (1+r)(a-c)+w \\
c &\leq a \\
c &\geq 0
\end{aligned}$$

where we have explicitly allowed for the dependence on (r,w).

The problem can be written as

$$\begin{aligned}
V_1(a,r,w) &= \max_{c\in[0,a]} u(c)+\beta V_2((1+r)(a-c)+w,r,w) \\
V_2(a,r,w) &= \max_{c\in[0,a]} u(c)+\beta V_3((1+r)(a-c)+w,r,w) \\
&\vdots \\
V_{T-1}(a,r,w) &= \max_{c\in[0,a]} u(c)+\beta V_T((1+r)(a-c)+w,r,w) \\
V_T(a,r,w) &= \max_{c\in[0,a]} u(c)+e((1+r)(a-c)+w)
\end{aligned}$$

The parameters are: $\{\beta,T,u(c),e(a)\}$. Set $T=6$ for simplicity and we can decrease β to account for the longer length of periods.

A.1.2 Collocation

The state space is $s=(a,z)$; $z=(r,w)$ is the exogenous state which has transition matrix $P=P_r\otimes P_w$ across generations, but is constant for each generation. The state space for z is discrete and so is enumerated

$k = 1, \ldots, K$, where $K = N_r \times N_w$. Let $s = (s_1, s_2)$ and the choice variable $x = c$. The choice is consumption $x \in B(s)$, where

$$B(s) = [0, a]$$

Re-writing this as a system of six value functions

$$\begin{aligned} V_1(s) &= \max_{x \in B(s)} F_1(s, x) + \beta V_2([(1+r)(s_1 - x) + w, s_2]) \\ &\vdots \\ V_T(s) &= \max_{x \in B(s)} F_2(s, x) \end{aligned}$$

This is the system we will solve.

Approximation: Take $V_1, \ldots, V_T$ and approximate them on J collocation nodes $s_1, \ldots, s_J$ with a spline with J coefficients $c^1 = (c_1^1, \ldots, c_J^1)$, $c^2, \ldots, c^T$ and linear basis ϕ_j.

$$\begin{aligned} V_1(s_i) &= \sum_{j=1}^{J} c_j^1 \phi_j(s_i) \\ &\vdots \\ V_T(s_i) &= \sum_{j=1}^{J} c_j^T \phi_j(s_i) \end{aligned}$$

Let $c = (c^1, \ldots, c^T)$ and let $v_1(c^1) = [V_1(s_1), \ldots, V_1(s_J)]'$ and $v_2(c^2), \ldots, v_T(c^T)$ similarly defined for a given c. With $v(c) = [v_1(c^1)', \ldots, v_J(c^J)']'$ then

$$\begin{aligned} v_1(s) &= \Phi c^1 \\ &\vdots \\ v_T(s) &= \Phi c^T \end{aligned}$$

this is the *Collocation equation.*

Substituting the interpolants into the value functions

$$\begin{aligned} \sum_{j=1}^{J} c_j^1 \phi_j(s_i) &= \max_{x \in B(s_i)} F_1(s_i, x) + \beta \sum_{j=1}^{J} c_j^2 \phi_j([(1+r)(s_{i,1} - x) + w, s_{i,2}]) \\ \sum_{j=1}^{J} c_j^2 \phi_j(s_i) &= \max_{x \in B(s_i)} F_1(s_i, x) + \beta \sum_{j=1}^{J} c_j^3 \phi_j([(1+r)(s_{i,1} - x) + w, s_{i,2}]) \\ &\vdots \end{aligned}$$

$$\sum_{j=1}^{J} c_j^T \phi_j(s_i) \quad = \quad \max_{x \in B(s_i)} F_2(s_i, x)$$

The stacked system of value functions is

$$\begin{aligned}
\Phi(s)c^1 &= F_1(s, x(s)) + \beta\Phi([(1+r)(s_1 - x(s)) + w, s_2])c^2 =: v_1(c^2) \\
\Phi(s)c^2 &= F_1(s, x(s)) + \beta\Phi([(1+r)(s_1 - x(s)) + w, s_2])c^3 =: v_2(c^3) \\
&\vdots \\
\Phi(s)c^T &= F_2(s, x(s))
\end{aligned}$$

The zero system would be $\tilde{\Phi}(s)c - v(c) = 0$, where $\tilde{\Phi}$ is a block diagonal matrix of $\Phi's$.

A.2 Estimation

The estimation procedure we use, described below, is adapted from Guvenen (2016). The global stage is a multi-start algorithm where candidate parameter vectors are uniform Sobol (quasi-random) points. We typically take about 10,000 initial Sobol points for pre-testing and select the best 200 points (i.e., ranked by objective value) for the multiple restart procedure. The local minimization stage is performed with the Nelder-Mead's downhill simplex algorithm (which is slow but performs well on non-linear objectives). Within one evaluation, we draw 100,000 individuals randomly and simulate their entire wealth process initiated with zero wealth and the lowest earnings profile.

A.3 Standard errors

We use numerical derivatives to calculate the standard errors for the parameters in all the estimates. The procedure is standard. The variance-covariance matrix for parameter estimates is given by

$$Q(W) = \left[\frac{\partial b(\theta_0)}{\partial(\theta)}' W \frac{\partial b(\theta_0)}{\partial(\theta)}\right]^{-1},$$

where $\frac{\partial b(\theta_0)}{\partial(\theta)}$ is the derivative of the vector of moments with respect to the parameter vector. We calculate the derivatives numerically, i.e. perturbing θ and calculating new vector moments. Standard errors will then be the square roots of the diagonal elements of $Q(W)$.

We use bootstrapping to generate the standard errors for the statistics related to the return process, e.g. its mean, standard deviation, and autocorrelation coefficient. The procedure is standard.

We take the parameter values for generating the return process as given, i.e. the values for the five Markov states and the diagonal matrix of the transition matrix (hence the whole Markov transition matrix), then generate the return process a sufficiently large number of times. We then calculate the mean, standard errors and the autocorrelation coefficient directly using these series of the return processes.

B. Data

B.1 Labor earnings

The labor earnings data we use are adapted from the PSID, as cleaned by Heathcote et al. (2010) - Sample C in their labeling.

We adopt the following procedure to obtain life-cycle age profiles, conditioning out year effects: i

1. Import household labor earnings - Sample C in Heathcote et al. (2010). The exact variable is *redlabinc*, i.e. household labor income (head + wife for couples). Keep households with head aged between 25 and 60 (inclusive). Label this variable *inc*.

2. Take log of the household labor earnings, $\log(inc)$. Drop all the observations with zero labor earnings. Record the mean of $\log(inc)_{it}$ in the initial year (2002) as $\overline{\log(inc)}_{2002}$.

3. Regress $\log(inc)_{it}$ against a full set of year dummies, denoting residuals ϵ_{it}:

$$\log(inc)_{it} = \overline{\log(inc)}_{2002} + year_{1967-2001} + \epsilon_{it}.$$

4. Generate predicted log earnings as:

$$\widehat{\log(inc)}_{it} = \overline{\log(inc)}_{2002} + \epsilon_{it};$$

   and predicted earnings as:

$$\exp(\widehat{\log(inc)})_{it}.$$

5. Construct, with the generated predicted earnings, age-dependent decile values as follows:[1]

   (a) Calculate decile values of earnings for each age;

   (b) Calculate average decile earnings for each six-year age bin.

[1] This procedure maintains the distributional ranking of households across the life cycle and allows them to move across bins during the life-cycle.

B.2 Inter-generational labor earnings transitions

Chetty et al. (2014) construct a 100 by 100 transition matrix linking parental family income with child's income - see `http://equality-of-opportunity.org/images/online_data_tables.xls`, Online Table 1. The main sample they use is the Statistics of Income (SOI) annual cross-sections from 1980 to 1982 cohorts for children, linking children to their parents by using population tax records spanning 1996-2012. We in turn collapse this matrix into a 10 by 10 transition matrix, which we associate to labor earnings.[2]

B.3 Inter-generational wealth mobility

The inter-generational wealth mobility matrix we use is estimated from the 2007-2009 SCF 2-year panel as follows. The procedure is the follllowing. Construct age-dependent 2-year transition matrices for age groups running from $30-31$ to $66-67$.[3] Multiply then these age-dependent 2-year transition matrices for all age groups, to construct the intergenerational social mobility matrix.

For comparison we report on previous estimates of wealth mobility in the literature:

1. Kennickell and Starr-McCluer (1997) estimate a seven-state (bottom 25, 25-49, 50-74, 75-89, 90-94, top 2-5, top 1) six-year transition matrix from the 1983-89 SCF panel - Table 7:

$$T_{KS,6} = \begin{bmatrix} 0.672 & 0.246 & 0.063 & 0.018 & 0.001 & 0.000 & 0.000 \\ 0.246 & 0.495 & 0.190 & 0.042 & 0.019 & 0.007 & 0.000 \\ 0.066 & 0.192 & 0.480 & 0.208 & 0.037 & 0.016 & 0.000 \\ 0.021 & 0.082 & 0.329 & 0.418 & 0.113 & 0.036 & 0.002 \\ 0.011 & 0.071 & 0.212 & 0.301 & 0.225 & 0.177 & 0.004 \\ 0.000 & 0.028 & 0.164 & 0.104 & 0.180 & 0.430 & 0.094 \\ 0.000 & 0.031 & 0.024 & 0.061 & 0.045 & 0.247 & 0.593 \end{bmatrix}$$

When raised to the power of 6, the 36-year transition matrix is:

[2] Chetty et al. (2014) also construct average income levels for both parent and child at age 29-30 - Online Table 2 - but do not provide life cycle data.

[3] Because of limited sample dimension, average the left and right matrices obtained using, respectively, the left-middle ages and the middle-right ages to define the age group in the 2-year panel.

$$T_{KS,36} = \begin{bmatrix} 0.316 & 0.278 & 0.222 & 0.118 & 0.037 & 0.024 & 0.005 \\ 0.276 & 0.263 & 0.240 & 0.137 & 0.044 & 0.031 & 0.009 \\ 0.224 & 0.242 & 0.263 & 0.163 & 0.054 & 0.042 & 0.012 \\ 0.196 & 0.229 & 0.274 & 0.176 & 0.061 & 0.051 & 0.013 \\ 0.179 & 0.219 & 0.275 & 0.181 & 0.066 & 0.061 & 0.020 \\ 0.150 & 0.198 & 0.271 & 0.185 & 0.074 & 0.082 & 0.040 \\ 0.112 & 0.166 & 0.252 & 0.182 & 0.085 & 0.121 & 0.083 \end{bmatrix}$$

2. Klevmarken et al. (2003) estimate a five-state (quintiles) five-year transition matrix from 1994- 1999 PSID data - Table 6:

$$T_{KLS,5} = \begin{bmatrix} 0.583 & 0.273 & 0.099 & 0.031 & 0.015 \\ 0.267 & 0.435 & 0.223 & 0.058 & 0.016 \\ 0.087 & 0.208 & 0.419 & 0.232 & 0.055 \\ 0.048 & 0.079 & 0.193 & 0.481 & 0.200 \\ 0.014 & 0.022 & 0.051 & 0.200 & 0.713 \end{bmatrix}$$

3. Charles and Hurst (2003) estimate a five-state (quintiles) inter-generational transition matrix from PSID data - using parent-child pairs in which (a) the parents were in the survey in 1984–89 and were alive in 1989, (b) the child was in the survey in 1999, (c) the head of the parent family was not retired and was between the ages of 25 and 65 in 1984, (d) the child was between ages 25 and 65 in 1999, and (e) both the child and the parent had positive wealth when measured - Table 5:[4]

$$T_{CH,gen} = \begin{bmatrix} 0.23 & 0.21 & 0.18 & 0.21 & 0.17 \\ 0.25 & 0.17 & 0.19 & 0.21 & 0.19 \\ 0.20 & 0.25 & 0.20 & 0.20 & 0.15 \\ 0.15 & 0.17 & 0.21 & 0.21 & 0.25 \\ 0.17 & 0.20 & 0.22 & 0.17 & 0.24 \end{bmatrix}$$

C. Additional results

C.1 Full transition matrix for r in the baseline

The parameterization of the stochastic process for r we use is defined by 5 states r_i and 5 diagonal

[4] Table 5 contains another transition matrix obtained by conditioning parental and child wealth for age, income, and portfolio choice, which is not relevant for our purposes.

transition probabilities, $P(r^n = r_i \mid r^{n-1} = r_i)$, $i = 1, \ldots, 5$, restricting instead the 5×5 transition matrix as follows: $P(r^n = r_i \mid r^{n-1} = r_j) = P(r^n = r_i \mid r^{n-1} = r_i)e^{-\lambda j}$, $i = 1, 2, 3, 4$, $j \neq i$, λ such that $\sum_{j=1}^{5} P(r^n = r_i \mid r^{n-1} = r_j) = 1$; and $P(r^n = r_5 \mid r^{n-1} = r_j) = \frac{1}{4}\left(1 - P(r^n = r_5 \mid r^{n-1} = r_5)\right)$. For readers' convenience, we report here the full transition matrix for the return process r in the baseline estimation.

$$\begin{bmatrix} 0.0224 & 0.5072 & 0.2631 & 0.1365 & 0.0708 \\ 0.2867 & 0.2698 & 0.2867 & 0.1126 & 0.0442 \\ 0.1156 & 0.3158 & 0.1371 & 0.3158 & 0.1156 \\ 0.0439 & 0.1118 & 0.2848 & 0.2746 & 0.2848 \\ 0.2444 & 0.2444 & 0.2444 & 0.2444 & 0.0224 \end{bmatrix}$$

C.2 Counterfactual estimates

Appendix C - Table 1: Parameter estimates: Constant r

	Preferences				
	σ	μ	A	β	T
	[2]	0.6008	0.0008	[0.97]	[36]
		(0.0780)	(0.1292)		
	Rate of return process				
$\mathbb{E}(r)$	3.01%				
	(0.02%)				

Notes: Standard errors in (); fixed parameters in [].

Appendix C - Table 2: Parameter estimates: Constant w

	Preferences				
	σ	μ	A	β	T
	[2]	0.9232	0.0073	[0.97]	[36]
		(0.0037)	(0.0002)		
	Rate of return process				
state space	0.0011	0.0156	0.0290	0.0603	0.0789
	(0.0004)	(0.0004)	(0.0013)	(0.0013)	(0.0015)
transition diagonal	0.1121	0.0317	0.1713	0.1890	0.0556
	(0.2019)	(0.0465)	(0.1262)	(0.0388)	(0.0152)
statistics	$\mathbb{E}(r)$	$\sigma(r)$	$\rho(r)$		
	3.39%	2.59%	0.1814		
	(0.82%)	(0.38%)	(0.1517)		

Notes: Standard errors in (); fixed parameters in [].

Appendix C - Table 3: Parameter estimates: $\mu = 2$

	Preferences				
	σ	μ	A	β	T
	[2]	2	0.0001	[0.97]	[36]
		-	(0.0001)		
	Rate of return process				
state space	0.0028	0.0111	0.0251	0.0602	0.0945
	(0.0002)	(0.0003)	(0.0052)	(0.0230)	(0.0082)
transition diagonal	0.3156	0.4941	0.4684	0.0574	0.0227
	(0.1460)	(0.1746)	(0.1876)	(0.2479)	(0.1061)
statistics	$\mathbb{E}(r)$	$\sigma(r)$	$\rho(r)$		
	2.78%	2.71%	0.2186		
	(0.86%)	(0.72%)	(0.1588)		

Notes: Standard errors in (); fixed parameters in [].

C.3 Complete wealth mobility matrices

We report the complete wealth mobility matrix in the baseline:

$$\hat{T}_{36} = \begin{bmatrix} 0.228 & 0.216 & 0.170 & 0.201 & 0.101 & 0.042 & 0.038 & 0.005 \\ 0.225 & 0.207 & 0.201 & 0.178 & 0.101 & 0.044 & 0.039 & 0.005 \\ 0.206 & 0.203 & 0.200 & 0.192 & 0.107 & 0.042 & 0.040 & 0.009 \\ 0.193 & 0.203 & 0.212 & 0.193 & 0.111 & 0.041 & 0.042 & 0.006 \\ 0.188 & 0.185 & 0.228 & 0.199 & 0.102 & 0.048 & 0.042 & 0.008 \\ 0.171 & 0.175 & 0.223 & 0.207 & 0.127 & 0.048 & 0.043 & 0.005 \\ 0.164 & 0.140 & 0.221 & 0.210 & 0.130 & 0.072 & 0.047 & 0.015 \\ 0.151 & 0.130 & 0.245 & 0.187 & 0.158 & 0.065 & 0.029 & 0.036 \end{bmatrix}$$

The corresponding complete matrices for all the three counterfactual cases are:

1. constant r:

$$\hat{T}_{36,const\ r} = \begin{bmatrix} 0.204 & 0.165 & 0.281 & 0.180 & 0.120 & 0.003 & 0.048 & 0 \\ 0.203 & 0.234 & 0.243 & 0.149 & 0.115 & 0.034 & 0.021 & 0 \\ 0.200 & 0.263 & 0.201 & 0.160 & 0.116 & 0.011 & 0.049 & 0 \\ 0.172 & 0.234 & 0.258 & 0.135 & 0.140 & 0.004 & 0.057 & 0 \\ 0.245 & 0.082 & 0.244 & 0.212 & 0.163 & 0.010 & 0.043 & 0 \\ 0.176 & 0.154 & 0.256 & 0.167 & 0.141 & 0.058 & 0.013 & 0.035 \\ 0.166 & 0.150 & 0.246 & 0.201 & 0.117 & 0.052 & 0 & 0.069 \\ 0.127 & 0.146 & 0.273 & 0.109 & 0.255 & 0.055 & 0 & 0.036 \end{bmatrix}$$

2. constant w:

$$\hat{T}_{36,const\ w} = \begin{bmatrix} 0.316 & 0.228 & 0.192 & 0.188 & 0.073 & 0.004 & 0 & 0 \\ 0.278 & 0.161 & 0.141 & 0.306 & 0.105 & 0.009 & 0 & 0 \\ 0.294 & 0.153 & 0.192 & 0.263 & 0.075 & 0.023 & 0 & 0 \\ 0.120 & 0.215 & 0.173 & 0.141 & 0.101 & 0.179 & 0.072 & 0 \\ 0.085 & 0.247 & 0.202 & 0.051 & 0.214 & 0 & 0.196 & 0.005 \\ 0.138 & 0.147 & 0.281 & 0.088 & 0.157 & 0.084 & 0.021 & 0.085 \\ 0.227 & 0.069 & 0.285 & 0.116 & 0 & 0.126 & 0.096 & 0.082 \\ 0.155 & 0.110 & 0.160 & 0.215 & 0.055 & 0 & 0.100 & 0.205 \end{bmatrix}$$

3. $\mu = 2$:

$$\hat{T}_{36,\mu=2} = \begin{bmatrix} 0.228 & 0.191 & 0.194 & 0.213 & 0.094 & 0.035 & 0.035 & 0.009 \\ 0.210 & 0.216 & 0.185 & 0.207 & 0.098 & 0.042 & 0.037 & 0.005 \\ 0.200 & 0.213 & 0.200 & 0.199 & 0.098 & 0.041 & 0.039 & 0.011 \\ 0.192 & 0.188 & 0.222 & 0.198 & 0.114 & 0.039 & 0.040 & 0.008 \\ 0.214 & 0.179 & 0.205 & 0.210 & 0.100 & 0.045 & 0.032 & 0.016 \\ 0.163 & 0.189 & 0.194 & 0.240 & 0.124 & 0.032 & 0.031 & 0.015 \\ 0.178 & 0.155 & 0.248 & 0.222 & 0.095 & 0.042 & 0.049 & 0.011 \\ 0.110 & 0.178 & 0.239 & 0.252 & 0.104 & 0.049 & 0.043 & 0.025 \end{bmatrix}$$

The complete wealth mobility matrix in the estimate with r dependent on wealth is:

$$\hat{T}_{36,r(a)} = \begin{bmatrix} 0.231 & 0.206 & 0.209 & 0.178 & 0.098 & 0.036 & 0.037 & 0.005 \\ 0.205 & 0.219 & 0.180 & 0.194 & 0.097 & 0.056 & 0.043 & 0.007 \\ 0.197 & 0.172 & 0.224 & 0.207 & 0.099 & 0.045 & 0.040 & 0.018 \\ 0.181 & 0.216 & 0.198 & 0.202 & 0.103 & 0.054 & 0.041 & 0.007 \\ 0.203 & 0.179 & 0.196 & 0.214 & 0.095 & 0.062 & 0.042 & 0.010 \\ 0.173 & 0.203 & 0.182 & 0.208 & 0.134 & 0.057 & 0.039 & 0.004 \\ 0.193 & 0.206 & 0.192 & 0.203 & 0.089 & 0.064 & 0.034 & 0.019 \\ 0.159 & 0.165 & 0.239 & 0.222 & 0.080 & 0.085 & 0 & 0.051 \end{bmatrix}$$

Finally, the complete wealth mobility matrix in the estimate allowing for non-stationary transitional dynamics is:

$$\hat{T}_{36,ns} = \begin{bmatrix} 0.663 & 0.317 & 0.020 & 0.000 & 0 & 0 & 0 & 0 \\ 0.125 & 0.535 & 0.212 & 0.088 & 0.039 & 0.000 & 0.001 & 0 \\ 0.073 & 0.142 & 0.493 & 0.183 & 0.075 & 0.031 & 0.002 & 0 \\ 0.070 & 0.002 & 0.275 & 0.494 & 0.068 & 0.054 & 0.036 & 0.001 \\ 0.065 & 0.004 & 0 & 0.467 & 0.297 & 0.058 & 0.098 & 0.012 \\ 0.073 & 0.003 & 0 & 0.013 & 0.594 & 0.175 & 0.085 & 0.057 \\ 0.066 & 0.005 & 0.000 & 0 & 0.104 & 0.452 & 0.271 & 0.102 \\ 0.073 & 0 & 0.005 & 0 & 0 & 0.015 & 0.740 & 0.168 \end{bmatrix}$$

C.4 Efficient Method of Simulated Moments Estimate

The following describes the procedure we used to produce an optimal weighting matrix for the second step estimation of the two-step Method of Simulated Moments (MSM).

Optimal weighting matrix. We follow Gourieroux, Monfort, and Renault (1993) and calculate the variance-covariance matrix of the data moments by bootstrapping, respectively for the wealth distribution moments and the intergenerational wealth mobility moments. Note that in order to invert the variance-covariance matrix, we use seven wealth moments (dropping the first one) to avoid perfect collinearity. Denote the variance-covariance matrix of the wealth distribution moments as $\mathcal{V}_{T_1}$, and that of the wealth mobility moments as $\mathcal{V}_{T_2}$, where T_1 and T_2 are the number of observations in each of the two samples.[5] We assume that there is no correlation in the error structure between the two samples. The optimal weighting matrix W_{T_1,T_2} would be the inverse of the concatenated block-diagonal variance-covariance matrix, that is,

$$W_{T_1,T_2} = \begin{bmatrix} \mathcal{V}_{T_1} & \mathbf{0} \\ \mathbf{0} & \mathcal{V}_{T_1} \end{bmatrix}^{-1}$$

We bootstrap 10,000 times for each of the variance-covariance matrix, and for each bootstrap we use half of the original sample to calculate the bootstrapped sample moments. As the wealth distribution moments are much more precisely estimated than the mobility moments, the weights on the former are around 3 orders of magnitude higher than the latter.

MSM results. In the first step, we use the same matrix we use in the baseline as the weighting matrix. We denote the first-step estimate as $\hat{\theta}_1$. Using $\hat{\theta}_1$ as the initial guess, we repeat the estimation procedure with the new optimal weighting matrix calculated earlier, $\widehat{W_{T_1,T_2}}$. Denote the second-step estimate as $\hat{\theta}_2$.

Appendix C - Table 4: Wealth quintiles: MSM

Moments								
Share of wealth	0-20	20-40	40-60	60-80	80-90	90-95	95-99	99-100
Data (SCF 2007)	-0.002	0.001	0.045	0.112	0.120	0.111	0.267	0.336
Simulation								
(1) Benchmark	0.047	0.074	0.107	0.102	0.105	0.071	0.155	0.339
(2) Opt W	0.048	0.077	0.111	0.107	0.110	0.075	0.141	0.331

Appendix C - Table 5: Diagonal of transition matrix: MSM

Moments								
Share of wealth	0-20	20-40	40-60	60-80	80-90	90-95	95-99	99-100
Data	0.223	0.220	0.210	0.198	0.110	0.062	0.094	0.028
Simulation								
(1) Benchmark	0.228	0.207	0.200	0.193	0.102	0.048	0.047	0.036
(2) Opt W	0.222	0.200	0.185	0.188	0.100	0.052	0.036	0.122

[5] Note that we use two different data samples for calculating wealth distribution and mobility moments. The former comes from the SCF 2007 cross-sectional sample, while the latter comes from the SCF 2007-2009 panel subsample.

Appendix C - Table 6: Parameter estimates: MSM

	Preferences				
(1) Markov chain	σ	μ	A	β	T
	[2]	0.5603	0.0005	[0.97]	[36]
		(0.0407)	(0.0001)		
	Rate of return process				
r grid	0.0010	0.0088	0.0261	0.0529	0.0855
	(0.0002)	(0.0009)	(0.0027)	(0.0088)	(0.0049)
prob. grid	0.0224	0.2679	0.1357	0.2715	0.0230
	(0.0587)	(0.3697)	(0.2946)	(0.1849)	(0.0613)
statistics	$\mathbb{E}(r)$	$\sigma(r)$	$\rho(r)$		
	3.01%	2.68%	0.1782		
	(0.86%)	(0.52%)	(0.1683)		

Notes: Standard errors in (); fixed parameters in [].

References

Charles, Kerwin Kofi and Erik Hurst, "The Correlation of Wealth across Generations," *Journal of Political Economy*, December 2003, *111* (6), 1155–1182.

Chetty, Raj, Nathaniel Hendren, Patrick Kline, and Emmanuel Saez, "Where is the land of Opportunity? The Geography of Intergenerational Mobility in the United States," *The Quarterly Journal of Economics*, 2014, *129* (4), 1553–1623.

Díaz-Giménez, Javier, Andrew Glover, and José-Víctor Ríos-Rull, "Facts on the distributions of earnings, income, and wealth in the United States: 2007 update," *Quarterly Review*, 2011.

Guvenen, Fatih, *Quantitative Economics with Heterogeneity: An A-to-Z Guidebook*, Princeton University Press, 2016.

Heathcote, Jonathan, Fabrizio Perri, and Giovanni L. Violante, "Unequal We Stand: An Empirical Analysis of Economic Inequality in the United States: 1967-2006," *Review of Economic Dynamics*, January 2010, *13* (1), 15–51.

Kennickell, Arthur B and Martha Starr-McCluer, "Household Saving and Portfolio Change: Evidence from the 1983-89 SCF Panel," *Review of Income and Wealth*, December 1997, *43* (4), 381–99.

Klevmarken, Anders, Joseph P. Lupton, and Frank P. Stafford, "Wealth Dynamics in the 1980s and 1990s: Sweden and the United States," *Journal of Human Resources*, 2003, *38* (2).

Miranda, Mario J. and Paul L. Fackler, *Applied Computational Economics and Finance*, Vol. 1 of *MIT Press Books*, The MIT Press, June 2004.