NBER WORKING PAPER SERIES

WEALTH DISTRIBUTION AND SOCIAL MOBILITY IN THE US: A QUANTITATIVE APPROACH

Jess Benhabib
Alberto Bisin
Mi Luo

Working Paper 21721
http://www.nber.org/papers/w21721

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
November 2015

Thanks to seminar audiences at Duke, NYU, Minneapolis Fed, SED-Warsaw, Lake Baikal Summer School, SAET - Cambridge, University College of London. Special thanks to Orazio Attanasio, Mariacristina De Nardi, Pat Kehoe, Ben Moll, Tom Sargent, Gianluca Violante, Daniel Xu. Corresponding author's email: alberto.bisin@nyu.edu. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Wealth Distribution and Social Mobility in the US: A Quantitative Approach
Jess Benhabib, Alberto Bisin, and Mi Luo
NBER Working Paper No. 21721
November 2015
JEL No. E21

ABSTRACT

This paper attempts to quantitatively identify the factors that drive wealth dynamics in the U.S. and are consistent with the observed skewed cross-sectional distribution of wealth and with social mobility in wealth. We concentrate on three critical factors: 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 factor are necessary for matching both distribution and mobility, with a distinct role for each. Stochastic earnings avoid poverty traps and allow for upward mobility near the borrowing constraints, as capital income risk has relatively small effects at low levels of wealth. Saving and bequest rate differentials help match the top tail but they reduce social mobility, inducing the rich get richer accumulating at higher rates. Capital income risk also contributes to the thick top tail of wealth while allowing for social mobility, especially in terms of speeding up downward mobility.

Jess Benhabib
Department of Economics
New York University
19 West 4th Street, 6th Floor
New York, NY 10012
and NBER
jess.benhabib@nyu.edu

Alberto Bisin
Department of Economics
New York University
19 West 4th Street, 5th Floor
New York, NY 10012
and NBER
alberto.bisin@nyu.edu

Mi Luo
New York University
19 West 4th Street 6FL
New York, NY 10012
ml3376@nyu.edu

1 Introduction

Wealth in the U.S. is unequally distributed, with a Gini coefficient of 0.78. 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% of wealth. At the same time, the U.S. is characterized by a relatively fast social mobility, with a Shorrocks mobility index in the range of $0.67 - 0.88$.[1] 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.

Many recent studies of wealth distribution and measures of inequality in the cross sectional distribution focus on the upper tail. We shall concentrate 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 *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 Castañeda, Ana, Javier Díaz-Gimínez, and José-Víctor Ríos-Rull (2003) and Kindermann and Krueger (2015). 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 Nardi (2006), and the recent work of Piketty (2014) discussing the saving rates of the rich directly. Finally, stochastic returns on 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 exploits stochastic returns of entrepreneurs to the same effect.[2] While all these factors

[1] 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. The U.S. range is from mobility matrices across 5 years from Klevmarken et al. (2003) for 1994-1999 (see Table 9, p. 342), and generational transitions estimated by Charles and Hurst (2003) using 1984-89 cohort for parents, Tables 2 and 5.

[2] This factor is also related to stochastic heterogeneous discount factors or heterogeneous stochastic impa-

contribute to produce skewed wealth distributions, their relative importance remains to be ascertained.[3]

The quantitative analysis we pursue consist in matching the moments generated by a macroeconomic model of wealth dynamics to the empirical moments of the observed distribution of wealth and social mobility matrix. An advantage of working with formal macroeconomic models is that, once we allow for an explicit demographic structure, we obtain implications for social mobility as well as the cross-sectional distribution. Indeed these factors have distinct effects on social mobility, which then helps to identify their relative importance in driving wealth dynamics.

A novel contribution of our approach in matching the moments of wealth distribution is that we simultaneously target moments of wealth mobility across generations, in particular we target some of the elements of the empirical wealth transition matrix. We explore how the various mechanisms that shape the distribution of wealth also contribute to social mobility in wealth, that is we try to jointly explain the distribution of wealth and the movements within it across generations. This is important for our estimation because certain mechanisms that can deliver the fat tails of wealth distribution, may also imply too little intergenerational mobility relative to the data.

We shall argue that all of the three factors are necessary for matching both distribution and mobility. Each of the factors seems to have a distinct role. Stochastic earnings avoid poverty traps and allow for upward mobility near the borrowing constraints as random returns on capital or capital income risk have relatively small effects at low levels of wealth. Saving rate differentials help match the top tail but they reduce social mobility as the rich

tience adopted by Krusell and Smith (1998). However, such discount factors are non-measurable and hence we prefer to restrict our analysis to capital income risk. Several papers in the literature include stochastic length of life (typically, "perpetual youth") models, to tame explosive accumulation path. We do not include this in our model as it has manifestly counterfactual demographic consequences: the rich are those agents who turn out to live relatively longer, as a (modeling) consequence have children later in life, and leave larger estates.

[3]See Hubmer et al. (2015) for a related attempt at this, though they have not looked at intergenerational wealth mobility. The paper is not yet available and hence we refer to later drafts for a discussion of similarities and differences.

get richer accumulating at higher rates. Stochastic returns on wealth, or capital income risk, also contributes to the thick top tail while allowing for social mobility, especially in terms of speeding up downward mobility.[4]

The rest of the paper is structured as follows. Section 2 lays out the theoretical framework, and 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 untargeted moments. Section 5 is our decomposition exercise where we shut down each mechanism at one time and re-estimate the model. Section 6 is an attempt to check the speed of transition our model delivers. The last section 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 distribution (power laws).[5] While this is also our aim, we more generally target the whole wealth distribution and its intergenerational mobility properties by building a simple micro-founded model - a standard macroeconomic model in fact - of life-cycle consumption and savings. The model exploits the interaction of the three 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

[4]Another possible factor which qualitatively would induce skewed wealth distributions but which we do not address directly is a rate of return on wealth which increases in wealth. It could serve as an important quantitative factor to explain the observed wealth distribution, working in a similar way to savings rates that increase in wealth. Such dependence of rates of return on wealth however has not been empirically documented in U.S. data; see Saez and Zucman (2016) who study returns on large public portfolios. Finally, precautionary savings could also contribute to distributional characteristics and the lower tail of wealth distribution under borrowing constraints, as precautionary motives increase savings rates especially of the poor for whom income risk is more important. 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 and the concavity of the consumption function. We do not exploit this channel for simplicity, assuming that life-cycle earnings profiles are random across generations but deterministic within lifetimes.

[5]We only discuss here those models which are directly relevant to our present analysis, referring to Benhabib and Bisin (2015) for an extensive survey of the theoretical and empirical literature on the wealth distribution.

wealth. These factors also have different implications on wealth mobility.

Each agent's life span is finite and deterministic, T years. Consumers choose consumption c and savings every period, subject to a no-borrowing constraint. Consumers leave a bequest at the end of life and get a warm-glow utility. The per-period utility from consumption, $u(c)$, and bequests, $e(a)$, are CRRA. Their functional forms, respectively, are

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

Wealth a accumulates from savings and bequests. Idiosyncratic rates of returns r and lifetime labor earnings profiles $\{w_t\}$ are drawn from a distribution at birth, possibly correlated with those of the parent, deterministic within each generation life.[6] We emphasize that r and $\{w_t\}$ are stochastic over generations only and can be interpreted as heterogeneity within the life cycle. Lifetime earnings profiles are hump-shaped, with low earnings early in life. As a consequence, borrowing constraints limit how much agents can smooth lifetime earnings.

Let $V_t(a_t)$ denote the discounted expected utility of agent with wealth a_t at the beginning of period t. The agents' maximization problem, written recursively, then is

$$\begin{aligned} V_t(a_t) &= \max_{c_t, a_{t+1}} u(c_t) + \beta V_{t+1}(a_{t+1}) \\ s.t.\ a_{t+1} &= (1+r)(a_t - c_t) + w_t, \\ 0 \leq c_t \leq a_t, &\quad for\ t \in [0, T] \\ V_{T+1}(a_{T+1}) &= \frac{1}{\beta} e(a_{T+1}) \end{aligned}$$

Note that within life, there is no uncertainty, hence no expectation is taken. The solution of

[6] Assuming deterministic earning profiles amounts to disregarding the role of intra-generational life-cycle uncertainty and hence of precautionary savings; see Huggett et al. (2011), and Cunha et al. (2010) for evidence that the life-cycle income positions tend to be determined early in life.

the recursive problem can be represented by a map

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

which we use to construct the intergenerational wealth dynamics process. Let apex n denote the generation. The process for the rate of return on wealth and earnings processes over generation n, (r^n, w^n) is a finite irreducible Markov Chain with transition $P\left(r^n, w^n \mid r^{n-1}, w^{n-1}\right)$ such that (abusing notation):

$$\begin{aligned} P\left(r^n \mid r^{n-1}, w^{n-1}\right) &= P\left(r^n \mid r^{n-1}\right), \\ P\left(w^n \mid r^{n-1}, w^{n-1}\right) &= P\left(w^n \mid w^{n-1}\right) \end{aligned}$$

Also, 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 construct then a stochastic difference equation for the initial wealth of dynasties, induced by the (forcing) stochastic process for (r^n, w^n), and mapping a^{n-1} into a^n:

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

where the map $g(.)$ represents indeed the solution of the life-cycle consumption-saving problem.

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

Case 1. $\mu = \sigma$, $g\left(a_0; r, w\right) = \alpha(r, w)a_0 + \beta(r, w)$;

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

In the first case $\mu = \sigma$, the intergenerational wealth dynamics is governed by a linear stochastic difference equation in wealth, which has been closely studied in the math literature (see de Saporta, Benoîte, 2005). Indeed, if $\mu = \sigma$ and $(\alpha(r^n, w^n), \beta(r^n, w^n))$ satisfy the restrictions of a reflective process (see Benhabib et al., 2011 for details), the tail of the stationary distribution of wealth, a^n is asymptotic to a Pareto law (where $Q \geq 1$ is a constant)

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

where $\lim_{N\to\infty} E\left(\prod_{n=0}^{N-1} (\alpha(r^{-n}, w^{-n}))^{\gamma}\right)^{\frac{1}{N}} = 1$.

If instead, keeping σ constant, $\mu < \sigma$, a stationary distribution might not exist; but if it does,

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

If $\mu = \sigma$, the restrictions of a reflective process which induce a limit stationary distribution of a require that the contractive and expansive components of the effective rate of return tend to balance, i.e., that the distribution of $\alpha(r^n, w^n)$ display enough mass on $\alpha(r^n, w^n) < 1$ as well some as on $\alpha(r^n, w^n) > 1$; and that effective earnings $\beta(r^n, w^n)$ be positive and bounded, hence acting as a reflecting barrier (these are the restrictions for a *reflective process*). In the general case, $\mu < \sigma$, saving rates and bequests tend to increase with initial wealth; as a consequence the model displays a distinct expansive tendency acting against the stationarity of a_n.

The stochastic properties of labor income risk, $\beta(r^n, w^n)$, have no effect on the tail on the long-run stationary distribution of wealth, if it exists, as long as they are not very thick.[7] Heavy tails in the stationary distribution require that the economy has sufficient capital income risk: if $\mu = \sigma$, for instance, an economy with limited capital income risk, where $\alpha(r^n, w^n) \leq \tilde{\alpha} < 1$ and where $\tilde{\beta}$ is the upper bound of $\beta(r^n, w^n)$, has a stationary

[7]This statement is not circular: the precise condition is that the tail of earnings be less thick than the tail implied by capital income risk under no earnings; see Grey (1994) and Hay et al. (2011).

distribution of wealth bounded above by $\frac{\tilde{\beta}}{1-\tilde{\alpha}}$. As long as a stationary distribution exists, wealth inequality (e.g., the Gini coefficient of the tail) increases with i) the capital income risk agents face in the economy, as measured by a "mean preserving spread" on the distribution of $\alpha(r^n, w^n)$; ii) the bequest motive A, iii) a smaller μ.

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 the earnings distribution, capital income risk, and differential savings. These are represented in the model by the properties of the dynamic process and the distribution of (r^n, w^n) and by the parameter μ, which implies differential savings (the rich saving more) when $\mu < \sigma$. We assume in the following analysis however that r^n and w^n are independent, though each is allowed to be serially correlated.

3.1 Methodology

The main assumption of the quantitative exercise is that the wealth and social mobility data observed in the U.S. are generated by a stationary distribution. We consider it a reasonable first step, but extend it to the case without assuming stationarity in Section 6 as a robustness check.[8]

In detail, we estimate the parameters of the described stochastic process using a Simulated Moments (MSM) estimator: we fix several parameters of the model (externally calibrated); we select some relevant moments; and we estimate the remaining parameters by matching the moments generated by the model and those in the data. Specifically, we fix $\sigma = 2$, $T = 36$,

[8]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. (2015), Kaymak and Poschke (2015), and Hubmer et al. (2015). Our preliminary results in Section 6 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.

$\beta = 0.97$ per annum, the stochastic process for individual income and its transition across generations, following Chetty et al. (2014).[9] The empirical targets are: (i) the following wealth percentiles: bottom 20%, $20 - 39\%$, $40 - 59\%$, $60 - 79\%$, $80 - 89\%$, $90 - 94\%$, $95 - 99\%$, and top 1% (eight moments), and (ii) the diagonal of the social mobility Markov chain transition matrix with states for bottom 25%, $25-49\%$, $50-74\%$, $75-89\%$, $90-94\%$, $95 - 99\%$, and top 1% (seven moments) as the moments to match. We estimate μ, A, a 5-state Markov Chain grid for r^n, and a restricted form of the social mobility matrix consisting in leaving diagonal elements free and imposing equal probabilities off the diagonal (twelve parameters summarized in vector θ).

Let m_n for $n = 1, \ldots, N = 15$ denote a generic empirical moment, and $d_n(\theta)$ be the corresponding model moment that is simulated for a given vector of model parameters, θ. We simulate the entire wealth process of 100,000 individuals, and we minimize the deviation between each data target and the corresponding simulated moment. For each moment n, define

$$F_n(\theta) = d_n(\theta) - m_n$$

Note that all our moments fall between $[0, 1]$, thus we are not too worried about large variation in the scales of moments. 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.

$$\mathbf{F}(\theta) = [F_1(\theta), \ldots, F_N(\theta)]^T$$

[9]The data in Chetty et al. (2014) refers to the 1980-82 U.S. birth cohort and their parental income. Originally, it is a 100-state Markov chain: each percentile of income distribution. We reduce it to a 10-state Markov chain; see Appendix B for the deciles, the transition matrix, and a detailed discussion of several issues with our measure of individual income.

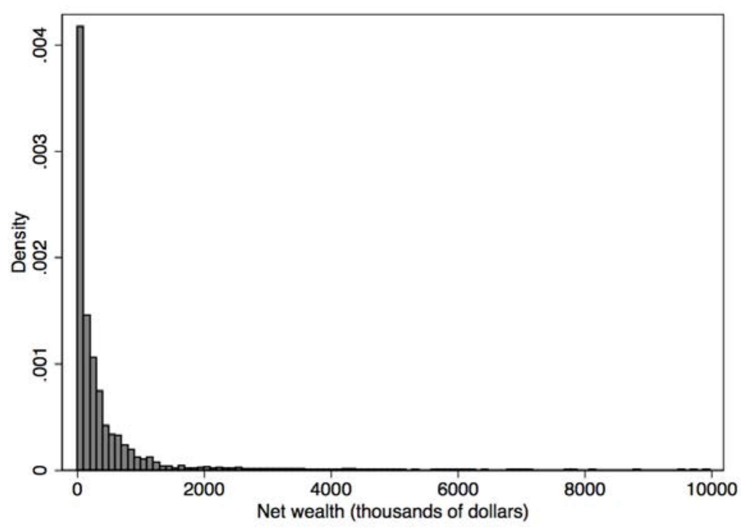
We choose an identity matrix for the weighting matrix in the baseline, $W = \mathbf{I}_N$. The objective function is highly nonlinear in general, therefore we employ a global optimization routine following Guvenen (2016) for the MSM estimation. Further details on the estimation can be found in Appendix A.

3.2 Data

We discuss the choice of output data for the targeted moments first, then the input data of labor income processes.

Output data. Matching the model and data generated moments requires wealth distribution and social mobility data. We take wealth distribution data from the Survey of Consumer Finances (SCF) 2007. Figure 1 displays the histogram for the wealth distribution, truncated at 0 on the left and ten million dollars on the right. Again the wealth distribution is very skewed to the right. We take the fractile shares from the cleaned version in Díaz-Giménez et al. (2011).

Figure 1: Wealth distribution in the SCF 2007 (weighted)



Notes: Data source is the 2007 SCF. Net wealth is defined as the sum of net financial wealth and housing. We restrict the sample to between 0 and 10 million negative wealth in this plot, but when we calculate the wealth fractile shares we do not apply those restrictions.

As for wealth transition across generations we take the six-year transition matrix (1983-1989) in Kennickell and Starr-McCluer (1997) also computed from SCF data. The states of the matrix are [bottom 25%, $25-49\%$, $50-74\%$, $75-89\%$, $90-94\%$, top $2-5\%$, top 1%]. The main reason for using this estimate is that it is the only estimate to our best knowledge that has a state for the top wealth share in its transition matrix.[10]

We transform the matrix into a 36 years transition (by raising it to the power 6), and obtain:

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

In the estimation we are only matching the diagonal of the above matrix. We are also only estimating the diagonal elements of the rates of return process, and we impose the off-diagonal cells of each row in the transition matrix for the r process to be equal. This assumption brings down the number of parameters we need to estimate.[11]

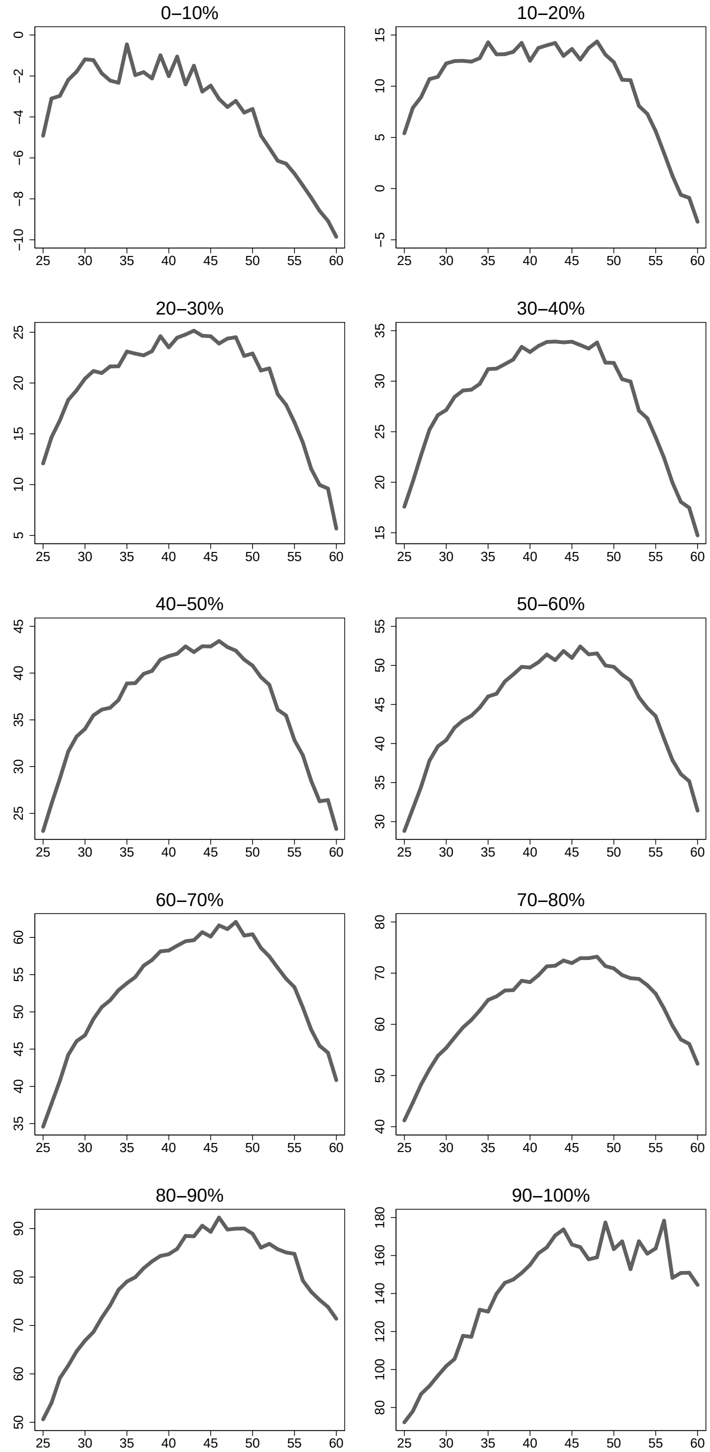
Input data. We use ten deterministic life-cycle household-level income profiles at different deciles, estimated from the Panel Study of Income Dynamics (PSID), drawn in Figure 2:[12]

[10] The qualitative implications regarding social mobility in Kennickell and Starr-McCluer (1997)'s estimates are robust: the matrix obtained by Klevmarken et al. (2003) with the PSID data is qualitatively similar; see Appendix C where the matrix is reported. Most importantly, the matrix estimated by Charles and Hurst (2003) to capture the intergenerational transmission in wealth exploiting information contained in the PSID about parent-child pairs is also similar; we discuss this point in detail in Section 4.3.

[11] We also experimented with exponentially decreasing off-diagonal cells, and results are very similar.

[12] We use household-level labor income, and do not distinguish between single or couple households. More details are provided in Appendix B.

Figure 2: Life-cycle income profiles by deciles



Notes: Data source same as in Table 1. However in plotting this figure we do not necessarily restrict people to have positive earnings.

The income levels used in our quantitative exercise are collapsed into six-year averages, as in Table 1. In all computations we assume an initial distribution of wealth concentrated on zero assets and an initial distribution of income concentrated on the lowest labor earnings decile.

Table 1: Life-cycle income 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]	-2.689	9.356	16.87	23.23	29.47	35.48	41.71	49.12	59.52	87.90
2 [31-36]	-1.683	12.90	21.88	29.78	37.10	44.21	52.06	61.69	75.01	123.5
3 [37-42]	-1.733	13.48	23.84	32.88	41.35	49.64	57.95	68.42	84.67	153.8
4 [43-48]	-2.734	13.59	24.54	33.73	42.76	51.46	60.73	72.46	90.04	165.5
5 [49-54]	-4.973	10.47	20.95	29.68	38.81	47.98	57.98	69.65	87.23	165.2
6 [55-60]	-8.226	1.047	11.31	19.63	28.21	37.60	47.20	59.23	77.07	156.5

Notes: Data source is the PSID cleaned by Heathcote et al. (2010). Income levels are in thousand dollars. In the actual quantitative exercise we replace negative earnings levels for the first decile (first column) with a small value (0.001), as our theoretical model does not allow for borrowing.

4 Estimation results

4.1 Parameter estimates

The baseline results are reported in Table 2. Several features of the estimates are important. The upper part of the table shows the estimates for the preference parameters. Note that the elasticity of the CRRA utility for bequests, $\hat{\mu}$ is 1.1860, which is significantly lower than that of the CRRA utility for normal consumption. $\mu < \sigma$ implies, as we noted, that savings out of wealth increase with wealth itself: the rich save proportionally more than the poor. In other words, our estimates point to the existence of the differential saving factor as a component of the observed wealth dynamics in the U.S.[13] Of course, the strength of the

[13] As we noted, differential saving can in principle, for a theoretical standpoint, make it impossible for a stationary distribution of wealth. But our estimates are predicated on the assumption that such distribution

bequest motive has to be evaluated jointly with the intensity parameter A as well. Here we have an estimated A of 0.0312, which is very low.

Table 2: Parameter estimates: baseline

	Parameters				
Preferences	σ	μ	A	β	T
	[2]	1.1860	0.0312	[0.97]	[36]
		(0.0077)	(0.1276)		
Rate of return					
r grid	0.0024	0.0143	0.0234	0.0665	0.0741
	(0.0010)	(0.0470)	(0.0217)	(0.0261)	(0.0218)
prob. grid	0.1992	0.3876	0.4043	0.2520	0.0414
	(0.1243)	(0.1602)	(0.1984)	(0.1772)	(0.0136)
stationary prob. grid	0.1812	0.2300	0.2436	0.1940	0.1513

Notes: [] indicates fixed parameters, standard errors computed with numerical derivatives for the parameter estimates in (). σ is the CRRA elasticity of consumption, μ is the CRRA elasticity of bequest, and A is the intensity of bequest. β is the annual discount factor, and T is the number of working periods. The return process follows a standard Markov chain. The values for the r grid is for an annual return. The whole matrix is reported in Appendix A. The objective value in the baseline is 0.0295. All the above notations remain the same throughout parameter estimates tables in the remainder of the paper.

The bottom part of Table 2 lays out the estimates for the rate of return process. Recall that we set the process as a Markov chain with five states - we do not impose any restriction on the process, such as assuming an AR(1) or any other distributional assumptions, in order to allow for sufficient freedom in the estimation.

Most of the parameters are estimated quite precisely with a t statistic greater than 2. To elaborate, the important curvature parameter μ is statistically very significant, though the bequest intensity parameter A is not. The bequest motives are jointly significant. Likewise, some of the r grid values or probabilities are statistically insignificant, yet most of them and parameters for the whole process are jointly significant.

The first row of the bottom part are the values estimated for these five states, the second

exists. In practice this must limit the possible strength of this factor. We will gauge at this issue better when we discuss counterfactuals in the next section.

row are the probabilities on the diagonal of the transition matrix, while the third row are the corresponding probabilities for the stationary distribution of the Markov chain. The mean annual rate is 3.35%, and the standard deviation is 40.15%.[14] Furthermore, the process is close to i.i.d. in the stationary distribution, i.e., the probability for each state is close to 0.2. This should be interpreted as a real, after-tax, growth-detrended rate. We find reassuring that it is quite consistent with previous estimates by Campbell and Lettau (1999), Campbell et al. (2001) and Moskowitz and Vissing-Jørgensen (2002), obtained from data on rates of return to private equity and entrepreneurship. This is particularly striking if we note that our estimate is obtained to match the wealth distribution and social mobility, with no direct return data.

4.2 Model fit

The simulations of our estimated model seem to capture the targeted moments quite well, as shown in Tables 3-4.

Table 3: Wealth fractiles: baseline

	Distributional moments								
Share of wealth	0-19	20-39	40-59	60-79	80-89	90-94	95-99	99-100	Gini
Data (SCF 2007)	-0.002	0.001	0.045	0.112	0.120	0.111	0.267	0.336	0.816
Baseline	0.014	0.048	0.105	0.168	0.102	0.070	0.151	0.341	0.799

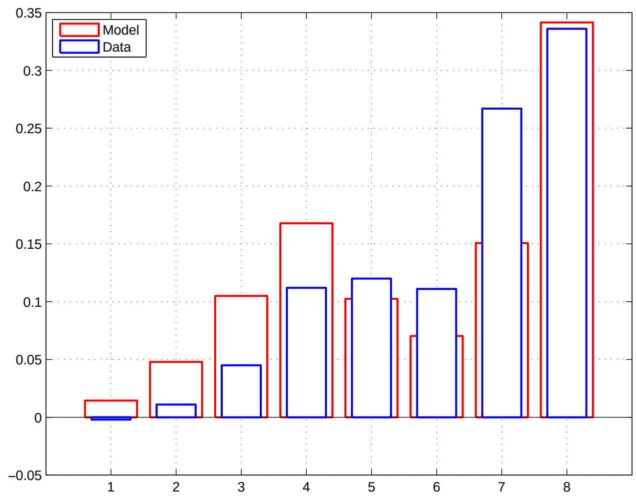
[14]Because we assume that one period is six years in the simulated model, we need to scale up the estimated standard deviation by $\sqrt{6}$ to convert to the actual standard deviation of an stochastic annual rate (i.e. one draw is constant throughout six years).

Table 4: Transition matrix: baseline

	Mobility moments						
Share of wealth	0-24	25-49	50-74	75-89	90-94	95-99	99-100
Data							
Diagonal	0.316	0.263	0.263	0.176	0.066	0.082	0.083
Top 1%	0.112	0.166	0.252	0.182	0.085	0.121	0.083
Shorrocks	0.959						
Our Simulation							
Diagonal	0.274	0.263	0.269	0.158	0.047	0.041	0.122
Top 1%	0.206	0.303	0.172	0.082	0.030	0.084	0.122
Shorrocks	0.971						

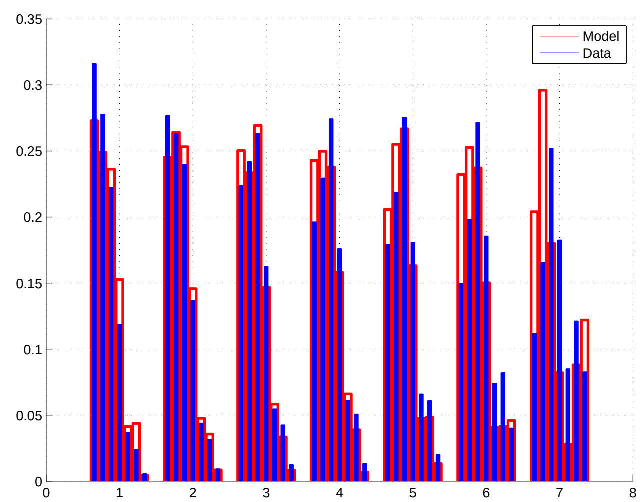
To facilitate the reading, we report the fit for the moments we have selected in Figures 3-4. The first one plots the empirical wealth share from the 2007 SCF (in blue) against the simulated shares from the model (in red). Due to the borrowing constraint, in our simulation agents display non-negative wealth holdings throughout their lifetime, and thus the simulated wealth distribution is less skewed than the data's. We match the top 1% share right on spot, yet somewhat miss the second top cell for the next $1 - 5\%$.

Figure 3: Wealth fractiles: baseline



Notes: On the horizontal axis we mark 8 bars corresponding to the 8 wealth distribution shares. The red bars are model simulated moments, while the blue bars are the empirical moments.

Figure 4: Mobility: baseline



Notes: On the horizontal axis we mark 7 clusters corresponding to the 7 rows of the transition matrix. Within each cluster, there are 7 bars corresponding to each column of the row. For example, the first bar of the second group refers to the $(2, 1)$ elements of the matrix.

Figure 4 plots the fit for mobility moments. Recall that we only explicitly target the diagonal elements of the Markovian transition matrix. We plot the whole matrix in this figure in order to get a sense how well we do on the (non-targeted) off-diagonal cells.

Careful reading of the figures indicates that we match the diagonal pretty well, and do a reasonable job for the off-diagonals. The last cluster, corresponding to the $7th$ row of the matrix, is the hardest to match. These are the probabilities of people in the top 1% either staying within the top 1% or falling down to other positions of the distribution. There is a lot of movement in the data: the probability of staying is a mere 8.3%, while the probability of moving down to the bottom 25% is a non-negligible 11.2%. Our simulated probability of staying is 12.2%, close enough to the empirical one; yet our estimate for the probability of falling down to the bottom is almost twice its empirical counterpart.

4.3 Independent evidence

We discuss here independent evidence which bears on the fit of the model with regards to savings, bequests, and wealth mobility. These are essentially moments we have not explicitly targeted, but we would like to show how well we fare in these regards.

Savings. In order to highlight the effect of the bequest motive on savings, we calculate the (non-targeted) savings rates for different wealth fractiles in our simulation and compare them with the empirical values calculated by Saez and Zucman (2016) using 2000-2009 data; see Table 5. Synthetic saving rates (defined by grouping everyone within a certain wealth fractile and calculating the ratio between changes in total wealth and total income of this group) are increasing in with wealth levels both in the data and in our simulation. While the simulation misses the saving rate of the top $10 - 1\%$, it does reasonably well for the top 1% and the bottom 90%.

Table 5: Synthetic savings rates comparison

	Fractile		
Share of wealth	Bottom 90	Top 10-1	Top 1
Data 2000-2009	-4%	9%	35%
Simulation	-5.65%	29.3%	42.2%

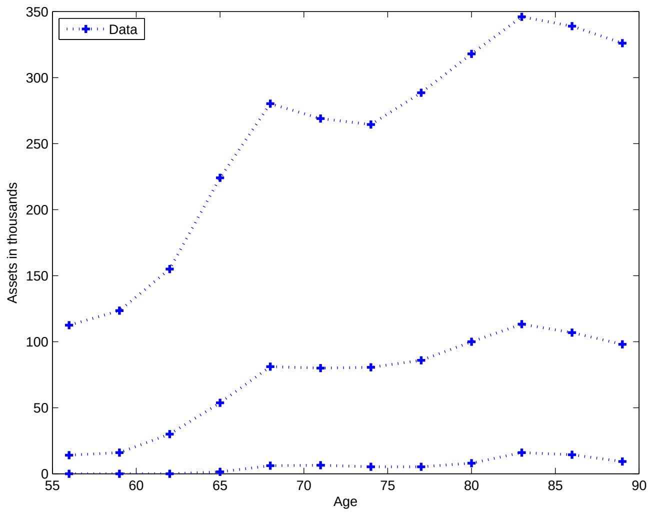
Notes: Data source is Saez and Zucman (2016). Synthetic saving rate for fractile p in year t is defined as $S_t^p = \frac{W_{t+1}^p - W_t^p}{Y_t^p}$, also adjusted for changes over time in price of assets in data.

Bequests. The distribution of bequests in our model, without mortality risks, maps closely the stationary wealth distribution. A crucial implication of our estimates is that the savings rate, part of which is driven by the bequest motive, is increasing in wealth.[15] To examine whether this feature is also a characteristics of the data, we examine the age profile of wealth. Figure 5 plots the wealth profiles for the U.S. population above age 55 based on the Health and Retirement Study (HRS). The three lines correspond to (from bottom to top) 25%, median, and 75% percentiles of the wealth distribution. We see that indeed retirement savings do not decline along the age path, and that this pattern is more accentuated for the 75% percentile, as our estimates imply.[16]

[15] 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).

[16] 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, Mariacristina, Eric French and John B. Jones, 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.

Figure 5: Retirement savings profiles



Notes: Data source is HRS wave 10 fielded in 2010. Here wealth is defined as net worth of assets in the household, including net financial wealth and housing.

People have also described the distribution of actual and expected bequests with various micro-level data. For example, Hurd and Smith (2003) use the HRS to characterize such distributions and find the bequests are very skewed, just as cross-sectional wealth.

Wealth mobility. The mobility matrix we use from Kennickell and Starr-McCluer (1997) is based on estimates from a six-year panel data, not necessarily representing intergenerational mobility. As explained previously, we choose this matrix because its Markov chain has a state for the top 1% wealth share. However, we ask here whether our estimates are similar to other intergenerational wealth mobility matrix in the literature. In particular, we compare our results with the intergenerational wealth mobility matrix estimated by Charles and Hurst (2003) with the PSID panel data. They have a five-state Markov chain, with each state representing each quintile share of the wealth distribution. If we only adjust the logs of parental and child wealth for age, the estimated matrix is:[17]

[17] Charles and Hurst (2003) also report another version of the matrix estimates, adjusting logs of parental and child wealth for age, income, and portfolio choice. We use the one with less conditioning first because we do not have those elements in our model, and second these aspects such as portfolio choices could be endogenous.

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

It should be noted that this transition matrix is a doubly stochastic matrix by construction, as the states are even percentiles. As a consequence, the implied stationary distribution will be uniform, 0.2 in each cell. The transition matrix estimated by Charles and Hurst (2003) over only one generation is relatively close to the uniform stationary distribution, which suggests a high intergenerational wealth mobility: people are churning quite fast in their wealth ranking across generations.

Our model simulations generate the following corresponding transition matrix which is very close to what Charles and Hurst (2003) have estimated:

$$\hat{T}_{gen} = \begin{bmatrix} 0.20 & 0.22 & 0.17 & 0.22 & 0.18 \\ 0.20 & 0.22 & 0.20 & 0.18 & 0.19 \\ 0.19 & 0.21 & 0.21 & 0.19 & 0.19 \\ 0.21 & 0.17 & 0.21 & 0.20 & 0.21 \\ 0.20 & 0.18 & 0.20 & 0.20 & 0.22 \end{bmatrix}$$

In particular, the Shorrocks index for both matrices is exactly the same, 0.988.

5 Counterfactual estimates

5.1 Re-estimation results

In this section we perform a set of counterfactual estimates of the model, under restricted conditions, and the associated simulations. There are three sets of counterfactuals that we are interested in, corresponding to shutting down each of the three main mechanisms which can drive the distribution of wealth: (1) stochastic rates of returns, (2) bequest motives leading to differential saving rates, and (3) stochastic labor income.

The objective of this counterfactual analysis is twofold. First of all we aim at gauging (and possibly, measuring) the relative importance of the various mechanisms we identified as possibly driving the distribution of wealth. We also attempt at a better understanding of which mechanism mostly affects which dimension of the wealth distribution. 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.

We examine the counterfactual estimates in detail in the following. In the counterfactual with no stochastic rates of return, we re-estimate a constant rate of return. The estimated parameters are in Table 6, the simulation moments in Table 7-8, Line 2. The differential savings mechanism does not substitute for the random rate of return, as the higher estimated μ implies. However the saving rate is higher due to a higher relative preference for bequests reflected in A. Nonetheless, the model now misses completely to match the top 1% of the wealth fractiles, which is reduced to about 1/6th of the baseline (and the data). The simulated wealth distribution becomes less skewed but does not entirely collapse, it has less mass on the top 10% and more of the bottom 80%. The match in mobility is also off: the top 1-5% has a too low a probability of staying and the the bottom too high.

Table 6: Parameter estimates: no stochastic rates of return

	Markov chain regime				
Preferences	σ	μ	A	β	T
	[2]	1.4969	0.3148	[0.97]	[36]
Rate of return					
r grid	0.034				
prob. grid	-				

Notes: The objective value is 0.2365.

In the counterfactual with homogeneous saving rates, we set the curvature parameter of the bequest utility, μ, to 2 (which is the curvature value of the normal consumption utility), such that agents with different wealth would still save at the same rate. The estimated parameters are in Table 9, the simulation moments in Table 7-8, Line 3. In terms of the estimates, preferences for bequests are jacked way up, so as to increase the (now constant) saving rate in the economy. Also the rate of return has a slightly lower mean. Once again, the model misses to match the top 1% of the wealth fractiles, which is reduced less than in the counterfactual with constant rate of return of wealth, to about 1/3rd of the baseline (and the data). A part from the last percentile, the simulated wealth distribution is not really less skewed, displaying even excessive mass on the top 60%. The match in mobility is reasonably good, except that the top 1% has still a too low a probability of staying.

Table 7: Wealth fractiles: counterfactual re-estimations

	Distributional moments							
Share of wealth	0-19	20-39	40-59	60-79	80-89	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) Baseline	0.014	0.048	0.105	0.168	0.102	0.070	0.151	0.341
(2) Const. r	0.184	0.187	0.191	0.194	0.098	0.050	0.038	0.057
(3) $\mu = 2$	0.021	0.072	0.155	0.223	0.164	0.134	0.160	0.073
(4) Const. low w	0.153	0.174	0.165	0.168	0.157	0.093	0.034	0.057
(5) Const. high w	0.016	0.060	0.081	0.119	0.158	0.125	0.174	0.267

In the counterfactual with no stochastic labor income, we set a constant labor income profile. We experiment with both a low and a high profile. The estimated parameters are in Table 10-11, the simulation moments in Table 7-8, Line 4-5. The differential savings mechanism now, in either case, substitutes for the random rate of return, as the higher estimated μ implies. On the other hand the preference for bequests decreases. While the low w counterfactual completely misses the wealth distribution, which is way less skewed than the data, the high w does not do badly, and even the top 1%, while lower than in the baseline and the data, is better than in all the other counterfactuals. A possible interpretation of these results is that it is the level of the wage, in particular the low wage, which creates a problem with the ability of the model to generate the observed wealth distribution, because the poor get stuck and cannot afford to save. When the wage is high, even if it is constant and not stochastic, the poor can afford to save and move up to better populate the right tail, and the fit for the targeted moments of the wealth distribution is much better. The match in mobility is also not good. Interestingly, when w is low, the staying probability of the top 1% is way too low (actually 0), while it is way too high (more than three times that in the data) when w is high.

Table 8: Diagonal of transition matrix: counterfactual re-estimations

	Mobility moments						
Share of wealth	0-24	25-49	50-74	75-89	90-94	95-99	99-100
Data	0.316	0.263	0.263	0.176	0.066	0.082	0.083
Simulation							
(1) Baseline	0.274	0.263	0.269	0.158	0.047	0.041	0.122
(2) Const. r	0.368	0.257	0.257	0.158	0.008	0.038	0.090
(3) $\mu = 2$	0.276	0.255	0.275	0.188	0.030	0.055	0.177
(4) Const. low w	0.326	0.265	0.248	0.089	0	0.180	0
(5) Const. high w	0.537	0.375	0.284	0.191	0.104	0.223	0.280

Table 9: Parameter estimates: $\mu = 2$

	Parameters				
Preferences	σ	μ	A	β	T
	[2]	[2]	1.9794	[0.97]	[36]
Rate of return					
r grid	0.007	0.014	0.026	0.067	.091
prob. grid	0.061	0.411	0.499	0.148	0.137

Notes: The objective value is 0.1246.

Table 10: Parameter estimates: constant low w profile

	Parameters				
Preferences	σ	μ	A	β	T
	[2]	0.9681	0.0145	[0.97]	[36]
Rate of return					
r grid	0.003	0.009	0.030	0.046	.088
prob. grid	0.045	0.418	0.395	0.301	0.183

Notes: The objective value is 0.2306.

Table 11: Parameter estimates: constant high w profile

	Parameters				
Preferences	σ	μ	A	β	T
	[2]	0.2434	0.0935	[0.97]	[36]
Rate of return process					
r grid	0.005	0.015	0.019	0.050	.068
prob. grid	0.172	0.070	0.039	0.023	0.091

Notes: The objective value is 0.1415.

In a nutshell, all these three mechanisms are crucial for generating the fat right tail of the wealth distribution and sufficient mobility. A high constant w reduces the fit the least. We cautiously interpret this result to imply that the stochastic earning mechanism is the least important in driving the observed distribution of wealth. Furthermore, each of the factors seems to have a distinct role. Stochastic earnings avoid poverty traps and allow for upward mobility near the borrowing constraints as random returns on capital or capital income risk have relatively small effects at low levels of wealth. Saving rate differentials help match the top tail but they reduce social mobility as the rich get richer accumulating at higher rates. Stochastic returns on wealth, or capital income risk, also contributes to the thick top tail while allowing for social mobility, especially in terms of speeding up downward mobility.

5.2 More on the earnings mechanism

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. This is the case, for instance, of Kindermann and Krueger (2015). The main difference between our analysis and Kindermann and Krueger (2015)'s is methodological. While we feed the model with a distribution of earnings obtained from the data, Kindermann and Krueger (2015) effectively estimate the (tail of the) earning process. More specifically, they use a seven state Markov chain for earnings constructed from data except in the highest state, which is instead estimated to fit the wealth distribution data. Their estimates imply that, at the stationary distribution of earnings, the average top 0.25% earns somewhere between 400 to 600 times the median income. Translated in dollars (the median earnings are about $\$50,000$), the earnings of the top 0.25% amount to at least $\$20,000,000$. While substantial uncertainty pervades the data on top earners, this number appears implausibly high: in fact the top 0.1% have average incomes of about $\$4,000,000$, out of which $\$1,637,000$ is earnings (wages, salaries, and pensions), according to the Piketty-Saez World Top Income Database at `http://topincomes.parisschoolofeconomics.eu/#Database:`, which is of course more than the average earnings of the top 0.25%.

Another successful simulation exercise driven by earnings is Castañeda, Ana, Javier Díaz-

Gimínez, and José-Víctor Ríos-Rull (2003). They develop a very rich overlapping-generation model with life-cycle features. Their 4-state stochastic process for labour earnings introduces a mechanism, sometimes referred to as the addition of an *awesome state*, to generate a very high skewness in the distribution.[18] Indeed, at the stationary distribution for labor earnings in the simulation, the top 0.039% earners have 1000 times the average labor endowment of the bottom 61%. Thus to attain a ratio of a 1000, if the bottom 61% earn $\$25,000$ on average, the top 0.0389% would have to earn $\$25,000,000$. This also appears implausibly high according to the Piketty-Saez World Income Database.

More generally, barring other mechanisms contributing to thick tails in wealth, the skewness of the earnings distributions tends to translate one-for-one to the distribution of wealth; see Benhabib and Bisin (2015). But in the data wealth is substantially more skewed than earnings,[19] making it difficult to match the data with stochastic earnings only, without the introduction of too *awesome* a state.

6 Transitional dynamics of the wealth distribution

As we have discussed, our quantitative analysis 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 of wealth as initial condition, we estimate the parameters of the model by matching the

[18]Díaz et al. (2003) use an alternative but also excessively skewed earning process relative to the earnings data (for example relative to administrative data reported by the World Top Income Database compiled by Piketty and Saez), where roughly 6% of the top earners have 46 times the labor endowment of the median. Dávila et al. (2012) match the distribution of wealth (their Section 5.2) with the same calibration as in Díaz et al. (2003).

[19]In the SCF 2007, the Pareto tail, an inverse measure of skewness of the distribution, is estimated to be 1.09 for wealth (net worth), 1.71 for total income, 2.13.

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. While the analysis does not require nor impose any stationarity of the distribution of wealth over time in the data, it does postulate that the model structure and parameter values stay constant after 1962.[20]

Table 12: Parameter estimates

Distribution + Mobility	Preferences				
Markov chain	σ	μ	A	β	T
	[2]	1.2923	0.0109	[0.97]	[36]
	Rate of return process				
r grid	0.0003	0.0091	0.0212	0.0540	0.0795
prob. grid	0.1596	0.4288	0.2299	0.2731	0.0263
stationary prob. grid	0.1792	0.2636	0.1955	0.2071	0.1546

Distribution only	Preferences				
Markov chain	σ	μ	A	β	T
	[2]	1.7610	1.2425	[0.97]	[36]
	Rate of return process				
r grid	0.0022	0.0132	0.0256	0.0496	0.0983
prob. grid	0.0445	0.4668	0.2207	0.4824	0.1086
stationary prob. grid	0.1442	0.2584	0.1768	0.2662	0.1545

Notes: The mean annual return is 3.35%, and its standard deviation is 40.15%. The objective value is 0.1035 in simulation (1) matching both distribution and mobility moments, and 0.5141 in simulation (2) matching only the distribution moments.

Reporting results in Tables 12-14, we distinguish two separate exercises we conducted. In the first (results in the upper panel of tables) we target both the moments of the wealth distribution and those of the transition matrix, as in our previous quantitative exercise. In the second (results in the lower panel) we only targets the moments of the wealth distribution.

We note that the fit with the fractals of the wealth distribution is reasonably close in

[20]Importantly, we do not feed in the analysis the observed fiscal policy reforms since the $'60's$. Doing so should improve the fit.

both cases. In particular, we match the top 1%, which is an indication of how well the model performs in the tail. As the second case only targets the distributional moments, it performs slightly better in this regard, as expected. The slightly worse distributional moments match in the first case is compensated by a better job with the mobility moments. As we can see, the simulated mobility moments in the second case are much larger than the empirical counterparts, yet in the first case they are not that off.

This exercise also provides useful information about the implied speed of convergence of the model. Gabaix et al. (2015) show on the other hand that typically in continuous-time models the typical models wealth distribution dynamics deliver extremely slow convergence in simulations. To remedy this they introduce exogenous variability in the growth rate or drift term of their model. This helps fill the fat tail of the wealth distribution faster, as the few lucky agents drawing long sequences of high growth rates fill the tail faster. In our model stochastic returns play the same role and the tail is filled faster. Stochastic discount rates also help with faster convergence in Hubmer et al. (2015) play the same role.

Table 13: Wealth fractiles

Moments								
Share of wealth	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
Simulation								
(1) In 2 periods (72 yrs)	0	0	0.002	0.033	0.135	0.172	0.292	0.367
(2) In 2 periods (distr. only)	0	0.003	0.025	0.080	0.155	0.150	0.269	0.319

Table 14: Diagonal of transition matrix: life cycle

Moments							
Share of wealth	0-24%	25-49%	50-74%	75-89%	90-94%	95-99%	99-100%
Data	0.316	0.263	0.263	0.176	0.066	0.082	0.083
Simulation							
(1) In 2 periods (72 yrs)	0.347	0.289	0.483	0.333	0.097	0.064	0.200
(2) In 2 periods (distr. only)	0.490	0.549	0.623	0.541	0.295	0.357	0.164

7 Conclusions

We estimated a 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 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.

Our analysis allows us to distinguish the contribution 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. While more work is obviously necessary in this respect, our results are quite encouraging.

References

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.

Benhabib, Jess, Alberto Bisin, 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," 2015.

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

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.

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.

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 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.

Gabaix, Xavier, Jean-Michel Lasry, Pierre-Louis Lions, and Benjamin Moll, "The Dynamics of Inequality," 2015.

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.

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.

Hubmer, Joachim, Per Krusell, and Anthony A. Smith, "The historical evolution of the wealth distribution: A quantitative-theoretic investigation," 2015.

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.

Kaymak, Baris and Markus Poschke, "The evolution of wealth inequality over half a century: the role of taxes, transfers and technology," 2015.

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.

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.

Supplemental Online Appendix

A. Details of Estimation Method

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

The problem is

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

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 increase 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 the transition matrix $P=P_r\otimes P_w$. 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 five 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

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

$$\vdots$$

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

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 procedure

The estimation procedure is described as below in two steps, adapted fromGuvenen (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).

A.3 Additional results

We report the full mobility matrix here for the baseline:

$$\hat{T}_{36} = \begin{bmatrix}
0.274 & 0.247 & 0.238 & 0.152 & 0.041 & 0.043 & 0.005 \\
0.246 & 0.263 & 0.254 & 0.145 & 0.048 & 0.036 & 0.009 \\
0.252 & 0.233 & 0.269 & 0.147 & 0.058 & 0.033 & 0.008 \\
0.238 & 0.254 & 0.239 & 0.158 & 0.066 & 0.039 & 0.009 \\
0.201 & 0.266 & 0.262 & 0.164 & 0.047 & 0.049 & 0.013 \\
0.232 & 0.250 & 0.240 & 0.152 & 0.040 & 0.041 & 0.044 \\
0.207 & 0.303 & 0.172 & 0.082 & 0.030 & 0.084 & 0.122
\end{bmatrix}$$

The full matrices for all other cases are available upon request.

B. Input Data Sources

B.1 Labor income levels

The labor income data we use is adapted from the PSID data cleaned by Heathcote et al. (2010), specifically Sample C in their labeling. We only keep those aged between 25-60 inclusively. Then we construct the age-dependent decile values in the following order: this order corresponds to several implicit assumptions, the most important of which is that we allow people to move across bins during their life cycle.

1. for each age calculate the decile values of earnings;

2. for each age bin of six years, calculate the average decile earnings across these six years.

The above order maintains the distributional ranking of model agents across the life cycle.

B.2 Intergenerational labor income transitions

Chetty et al. (2014) provide a 100 by 100 transition matrix linking parental family income and child's income in their online data and tables, with each cell corresponding to share of each percentile of the income distribution.[1] The main sample they use is the Statistics of Income (SOI) annual cross-sections from 1980 to 1982 cohorts for children, and the authors link children to their parents using population tax records spanning 1996-2012 for parent family income. We collapse this big matrix into a 10 by 10 transition matrix, with each cell corresponding to share of each decile of the income distribution. Note that this matrix captures intergenerational transition in income.

Online table 2 of Chetty et al. (2014) also provide the average income levels for both parent and child. However, they are an average income around a particular age (29-30) for both parent and child rather than an average life cycle income. We would like our income profiles to capture the hump-shaped life cycle feature, thus calculate our own as explained in the last sub-section.

C. Output Data Sources

C.1 Wealth distributional moments

The wealth distributional moments are taken from Díaz-Giménez et al. (2011). Their calculations are more cleaned and serve as an official report. Many papers have used their numbers, e.g. in Kindermann and Krueger (2015). Other estimates are very close.

[1] See `http://equality-of-opportunity.org/images/online_data_tables.xls`, online table 1.

C.2 Intergenerational Wealth mobility moments

There are three papers, to the best of our knowledge, that estimate a transition matrix for wealth mobility. Kennickell and Starr-McCluer (1997) and Klevmarken et al. (2003) are both estimated using panel data, i.e. not necessarily transition across generations. The former paper used SCF panel and the latter used PSID panel. Charles and Hurst (2003) are a transition matrix for generations in particular. Please note the difference, though I try to argue they yield similar estimates.

Kennickell and Starr-McCluer (1997)calculate the six-year transition matrix from 1983 to 1989 for quartiles and top percentile ranges, and their results are quite similar to Klevmarken et al. (2003). The seven states are: bottom 25, 25-49, 50-74, 75-89, 90-94, top 2-5, top 1, respectively. Their estimates are (from 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 (i.e. 36-year transition matrix), we have

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

We see that the 36-year transition matrix does not necessarily reach the stationary distribution.

Klevmarken et al. (2003)calculate the five-year transition matrix from 1994 to 1999 for quintiles using PSID data. Note the states for the Markov chain are different. Their estimates are (from Table 6),

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

One potential issue with the above transition matrix is that it does not necessarily capture the *intergenerational* transmission in wealth. For that argument, let us look at the alternative transition matrix estimated by Charles and Hurst (2003).[2] There are two transition matrices in Table 5. If we only adjust the

[2]Sample selection: Their sample consists of all PSID 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. There were 1,491 such parent-child pairs.

logs of parental and child wealth for age, the matrix is:

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

If we adjust logs of parental and child wealth for "age, income, and portfolio choice," the corresponding matrix is:

$$T_{CH,gen,adj} = \begin{bmatrix} 0.36 & 0.29 & 0.16 & 0.12 & 0.07 \\ 0.26 & 0.24 & 0.24 & 0.15 & 0.12 \\ 0.16 & 0.21 & 0.25 & 0.24 & 0.15 \\ 0.15 & 0.13 & 0.20 & 0.26 & 0.26 \\ 0.11 & 0.16 & 0.14 & 0.24 & 0.36 \end{bmatrix}$$

with each cell corresponding to a quintile-to-quintile transition probability. Again note the differences in the states of the Markov chain.

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.

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).

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