

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

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

First draft: July 2015; this draft: December 2015\*

## Abstract

This paper attempts to quantitatively identify the factors that drive wealth dynamics in the U.S. and are consistent with its observed skewed cross-sectional distribution and social mobility. 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. All of these factors are necessary for matching both distribution and mobility, with a distinct role in inducing wealth accumulation near the borrowing constraints, contributing to the thick top tail of wealth, and affecting upward and/or downward social mobility.

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. Special thanks to Orazio Attanasio, Mariacristina De Nardi, Pat Kehoe, Per Krusell, Ben Moll, Tom Sargent, Gianluca Violante, Daniel Xu. Special thanks to Luigi Guiso for spotting a mistake in a previous version. Corresponding author's email: alberto.bisin@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.<sup>1</sup>

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 the cross-sectional distribution. We then match the moments generated by the model to several empirical moments of the observed distribution of wealth and social mobility matrix. Indeed distinct wealth accumulation factors have qualitatively different effects on social mobility, which then helps to identify their relative importance in driving wealth dynamics, as e.g., certain factors that deliver the thick tails of the wealth distribution also imply too little intergenerational mobility relative to the data.<sup>2</sup>

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 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

---

<sup>1</sup>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.

<sup>2</sup>See Hubmer et al. (2015) for a related exercise. This paper is not yet available and hence we refer to later drafts for a detailed discussion of similarities and differences; however, like the rest of the literature, it does not look at intergenerational wealth mobility.

Castañeda et al. (2003) and Kindermann and Krueger (2015).<sup>3</sup> 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 (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.<sup>4</sup> While all these factors contribute to produce skewed wealth distributions, their relative importance remains to be ascertained.

Other possible factors which qualitatively would induce skewed wealth distributions but which we do not address directly include a rate of return on wealth which increases in wealth and a precautionary savings motive for wealth accumulation. A dependence of rates of return on wealth would have similar effects as differential savings rates. We do not include this factor in our analysis since it has not been empirically documented in U.S. data; see Saez and Zucman (2016) who study returns on large public portfolios. A precautionary savings motive could also contribute to explaining the thick tail of the wealth distribution. 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.

Our main results document that all of the three factors we study, stochastic earnings, differential savings rates, and capital income risk, are necessary for matching both distribu-

---

<sup>3</sup>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 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.

<sup>4</sup>This factor is also related to stochastic heterogeneous discount factors or heterogeneous stochastic impatience 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.

tion and mobility. Furthermore, our analysis allows us to distinguish the contributions of these factors in driving wealth accumulation. Each of them has 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.

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 our attempt to extend the analysis allowing for non-stationarity of the wealth distribution and to measure 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).<sup>5</sup> 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

---

<sup>5</sup>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 distribution and mobility.

Each agent's life span is finite and deterministic,  $T$  years. Every period, consumers choose consumption  $c$  and accumulate wealth  $a$ , 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 accumulates from savings and bequests. Idiosyncratic rates of returns  $r$  and life-time labor earnings profiles  $\{w_t\}_{t=1}^T$  are drawn from a distribution at birth, possibly correlated with those of the parent, deterministic within each generation life.<sup>6</sup> We emphasize that  $r$  and  $\{w_t\}$  are stochastic over generations only, not 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 an agent with wealth  $a_t$  at the beginning of period  $t$ . The agent's maximization problem, written recursively, then is

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

The agent faces no uncertainty within his life span, hence no expectation is taken in the

---

<sup>6</sup>Assuming deterministic earning profiles amounts to disregarding the role of intra-generational life-cycle uncertainty and hence of precautionary savings; see Keane and Wolpin (1997), Huggett et al. (2011), and Cunha et al. (2010) for evidence that the life-cycle income positions tend to be determined early in life.

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

$$a_T = g(a_0; r, w),$$

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  $P(r^n, w^n \mid r^{n-1}, w^{n-1})$  such that (abusing notation):

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

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 stochastic process for  $(r^n, w^n)$ , and mapping  $a^{n-1}$  into  $a^n$ :

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

where the map  $g(\cdot)$  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:

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

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

In the first case,  $\mu = \sigma$ , the savings rate is independent of wealth. If instead  $\mu < \sigma$  differential savings rate emerge, increasing with wealth. In the first case the intergenerational wealth dynamics is then governed by a linear stochastic difference equation in wealth, which has been closely studied in the math literature; see de Saporta (2005). Indeed, if  $\mu = \sigma$ , the distribution of earnings is not too thick,<sup>7</sup> 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 independent of the distribution of earnings and asymptotic to a Pareto law,

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

where  $Q \geq 1$  is a constant and  $\lim_{N \rightarrow \infty} E \left( \prod_{n=0}^{N-1} (\alpha(r^{-n}, w^{-n}))^\gamma \right)^{\frac{1}{N}} = 1$ . If instead, keeping  $\sigma$  constant,  $\mu < \sigma$ , a stationary distribution might not exist; but if it does,

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

### 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 savings, and capital income risk. 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  $\mu$  and  $\sigma$ , which imply 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.

---

<sup>7</sup>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).

### 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 extend the analysis to possibly non-stationary distributions in Section 6 as a robustness check.<sup>8</sup>

We estimate the parameters of the stochastic process described in the previous section using a Method of Simulated Moments (MSM) estimator: we fix (or externally calibrate) several parameters of the model ; we select some relevant moments as target in the estimation; and we estimate the remaining parameters by matching the targeted moments generated by the model and those in the data. Specifically, we fix  $\sigma = 2$ ,  $T = 36$ ,  $\beta = 0.97$  per annum. We also fix the stochastic process for individual earnings profiles (estimated from the PSID) and its transition across generations following Chetty et al. (2014).<sup>9</sup> The moments we target are (i) the bottom 20%, 20 – 39%, 40 – 59%, 60 – 79%, 80 – 89%, 90 – 94%, 95 – 99%, and top 1% wealth percentiles; and (ii) the diagonal of the social mobility Markov chain transition matrix for the bottom 25%, 25 – 49%, 50 – 74%, 75 – 89%, 90 – 94%, 95 – 99%, and top 1% states. We estimate  $\mu$ ,  $A$ , a five-state Markov Chain grid for  $r^n$ , and a restricted form of the transition matrix consisting in leaving diagonal elements free and imposing equal probabilities off the diagonal.<sup>10</sup> In total, therefore, we target 15 moments and we estimate 12 parameters.

Let  $\theta$  denote the vector of the parameters to be estimated. Let  $m_n$  for  $n = 1, \dots, N = 15$  denote a generic empirical moment and  $d_n(\theta)$  be the corresponding model moment that is

---

<sup>8</sup>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.

<sup>9</sup>The data in Chetty et al. (2014) refers to the 1980-82 U.S. birth cohort and their parental income. We only use their data for the transition matrix of our earnings process. Originally, it is a 100-state Markov chain: each percentile of income distribution. We reduce it to a ten-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.

<sup>10</sup>This is just to reduce the number of parameters we need to estimate. We also experimented with exponentially decreasing off-diagonal cells, and results are very similar.



simulated for given parameter vector  $\theta$ . We simulate the entire wealth process of 100,000 individuals, and we minimize the deviation between each targeted moment and the corresponding simulated moment. For each moment  $n$ , define

$$F_n(\theta) = d_n(\theta) - m_n.^{11}$$

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.

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

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.

We discuss identification in Section 5 through counterfactual re-estimations, shutting off the wealth accumulation factors one at a time. Intuitively, however, the two preference parameters, in particular  $\mu$ , determine the degree of heterogeneity in saving rates driven by the bequest motives. The farther  $\mu$  deviates to the left of  $\sigma$ , the more heterogeneity we have among agents' saving rates. The five parameters for the  $r$  grid values and the corresponding probabilities along the transition matrix diagonal are crucial for generating not only different wealth accumulation, hence inequality, but also the social mobility across generations.<sup>12</sup>

---

<sup>11</sup>Note that all our moments fall between  $[0, 1]$ , thus we are not too worried about large variation in the scales of moments.

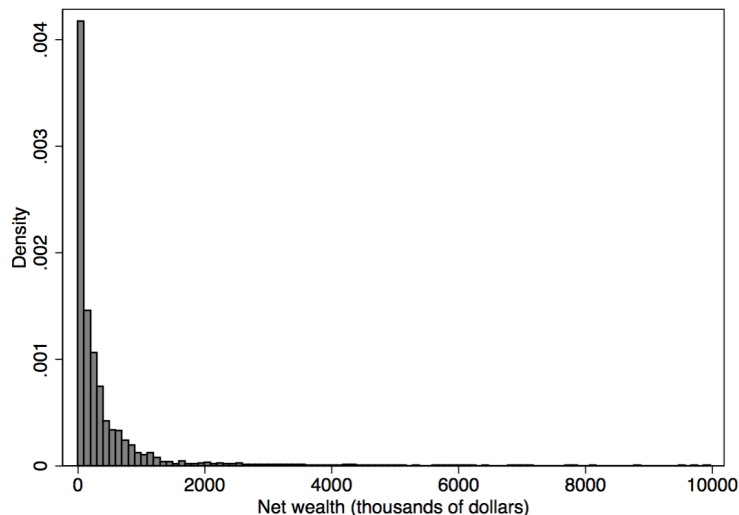
<sup>12</sup>Note that we feed the earnings process directly from the data into the estimated model.

## 3.2 Data

We discuss our careful choice of data for the targeted moments first, turning then to the labor earnings data.

**Targeted moments.** 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 (before the Great Recession). Figure 1 displays the histogram for the wealth distribution, truncated at 0 on the left and ten million dollars on the right. Here the variable is net wealth, defined as the sum of net financial wealth and housing minus any debts. The wealth distribution is very skewed to the right. We take the fractile shares from the cleaned version in Díaz-Giménez et al. (2011) for [bottom 20%, 20 – 39%, 40 – 59%, 60 – 79%, 80 – 89%, 90 – 94%; 95 – 99%, top 1%] and report them in the first row of Table 3.

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 net wealth in this figure, 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.<sup>13</sup> 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 1% wealth share in its transition matrix.<sup>14</sup>

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

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

Observe that there is quite a lot mobility in wealth holdings across one generation already. In the estimation we are only matching the diagonal of the above matrix.

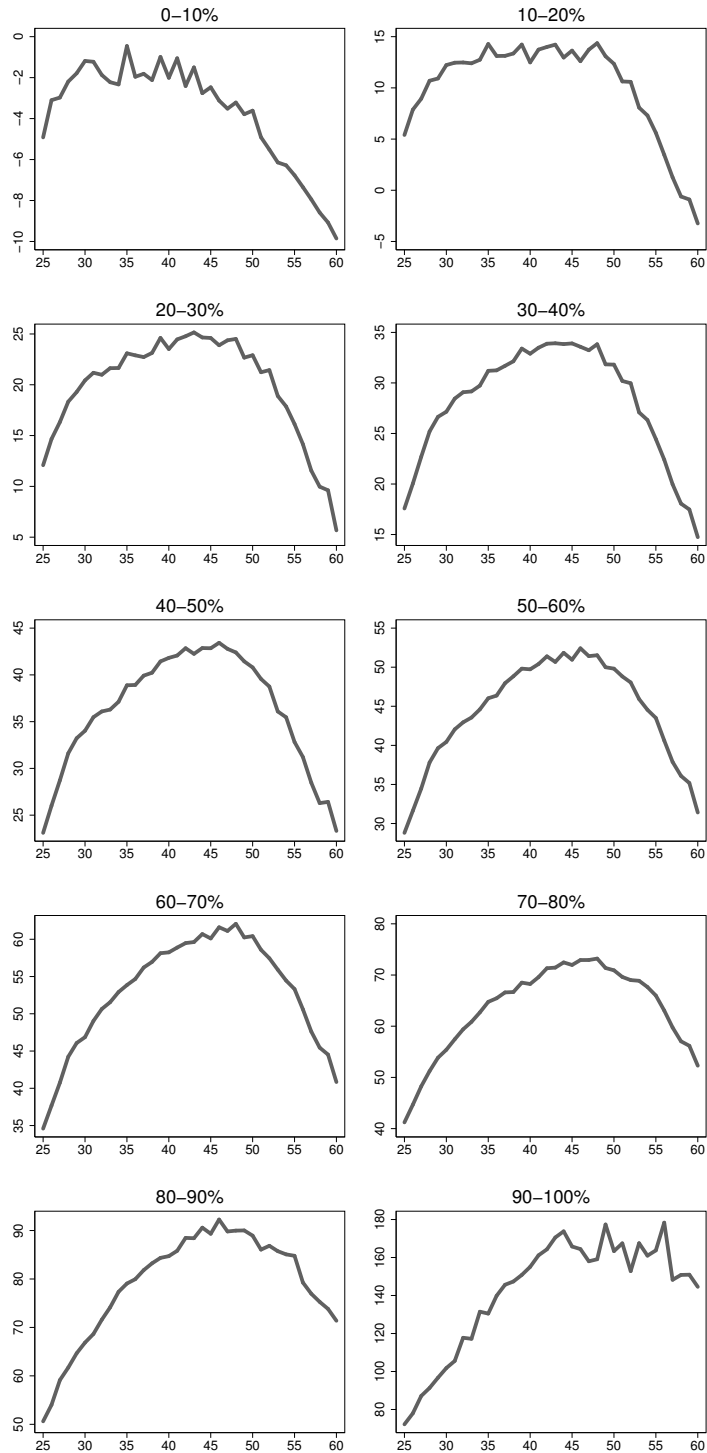
**Labor earnings data.** We use ten deterministic life-cycle household-level earnings profiles at different deciles, estimated from the Panel Study of Income Dynamics (PSID) and drawn in Figure 2. We use household-level labor earnings, and do not distinguish between single or couple households. We relegate more details of the data choice and estimations to Appendix B.

---

<sup>13</sup>Note that the 1983-1989 SCF is a panel data on a smaller scale.

<sup>14</sup>The qualitative implications regarding social mobility in Kennickell and Starr-McCluer (1997)'s estimates are robust: the matrix obtained by Klevmarcken 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.

Figure 2: Life-cycle earnings profiles by deciles



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

The earnings levels used in our quantitative exercise are collapsed 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]	-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 household earnings cleaned by Heathcote et al. (2010). Earnings 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 hence no negative positions of wealth.

## 4 Estimation results

The baseline estimation results are reported in Table 2. The targeted simulated moments of the estimated model are reported and compared to their counterpart in the data in Tables 3-4. Some independent evidence which bears on the fit of the model is discussed in Section 4.3.

### 4.1 Parameter estimates

Several features of the estimates are important. The upper part of Table 2 shows the estimates for the preference parameters. Note that the estimate of  $\mu$ , the elasticity of the CRRA utility for bequests, is 1.1860, which is significantly lower than 2, the assumed elasticity of the CRRA utility for consumption. Therefore  $\mu < \sigma$  and as we noted, 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.<sup>15</sup> Of course, the strength of the bequest motive has to be evaluated jointly with the intensity parameter  $A$  as well. Here we have a low estimated  $A$  of 0.0312.

Table 2: Parameter estimates: baseline

<i>Preferences</i>	<b>Parameters</b>				
	$\sigma$	$\mu$	$A$	$\beta$	$T$
	[2]	1.1860 (0.0077)	0.0312 (0.1276)	[0.97]	[36]
<i>Rate of return</i>					
$r$ grid	0.0024 (0.0004)	0.0143 (0.0192)	0.0234 (0.0089)	0.0665 (0.0106)	0.0741 (0.0089)
prob. grid	0.1992 (0.1243)	0.3876 (0.1602)	0.4043 (0.1984)	0.2520 (0.1772)	0.0414 (0.0136)
stationary prob. grid	0.1812	0.2300	0.2436	0.1940	0.1513

Notes: [] indicates fixed parameters; () indicates standard errors computed with numerical derivatives for the parameter estimates.  $\sigma$  is the CRRA elasticity of consumption,  $\mu$  is the CRRA elasticity of bequest, and  $A$  is the intensity of bequest.  $\beta$  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.<sup>16</sup> The first two rows report the values estimated for these five states together with standard errors, the next two rows report the probabilities on the diagonal of the transition matrix, while the last row reports the corresponding probabilities for the stationary distribution of the Markov chain. The rate of return estimate should be interpreted as annual, growth detrended, real, and after-tax. Its mean is 3.35%, and its standard deviation is a 2.73%. While the estimate

<sup>15</sup>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 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.

<sup>16</sup>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.

of the mean appears reasonably in line with previous direct estimates, the standard deviation is substantially smaller. This is the case for instance when our estimates are compared with 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 the standard deviation of rates of return to private equity and entrepreneurship. Such comparisons require however great caution. We assume in fact that  $r$  stays constant throughout each agent's lifetime, therefore disregarding the whole variation across the life-cycle. The rate of return we estimate should be ideally be compared with the permanent components of individual returns across generations.<sup>17</sup> Interestingly, the rate of return process we estimate is close to i.i.d. across generations. This is shown by the fact that, in the stationary distribution, the probability for each state is estimated close to 0.2.<sup>18</sup>

Most of the parameters are estimated quite precisely with a  $t$  statistic greater than 2. To elaborate, the curvature parameter  $\mu$  is statistically very significant and, though the bequest intensity parameter  $A$  is not, the bequest motives are jointly significant. Likewise, while a very small subset of the  $r$  grid values or probabilities are statistically insignificant, the parameters for the whole process are jointly significant.

## 4.2 Model fit

The simulations of our estimated model seem to capture the targeted moments quite well.

---

<sup>17</sup>Inter-generational return data is hardly available, with the exception of administrative data collected in Norway currently studied by L. Guiso and L. Pistaferri.

<sup>18</sup>Note however that we impose the off-diagonal elements in the estimated transition matrix to be equal.

Table 3: Wealth fractiles: baseline

<i>Share of wealth</i>	<b>Distributional moments</b>								Gini
	0-19	20-39	40-59	60-79	80-89	90-94	95-99	99-100	
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

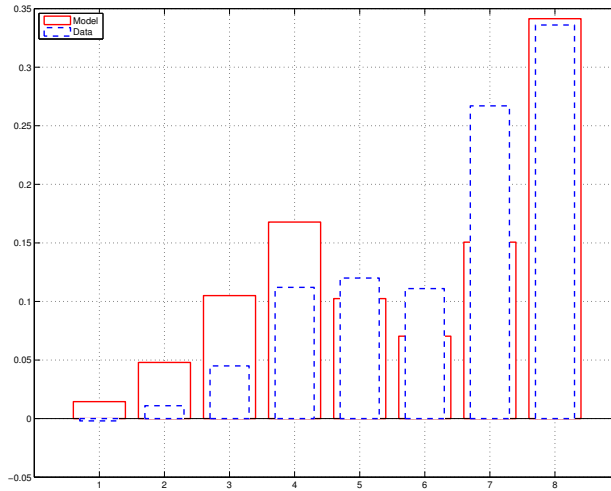
Table 4: Transition matrix: baseline

Share of wealth	<b>Mobility moments</b>						
	0-24	25-49	50-74	75-89	90-94	95-99	99-100
<i>Data</i>							
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						
<i>Our Simulation</i>							
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						

While this is apparent in Tables 3-4, we also report the fit for the targeted moments in Figures 3-4 to facilitate the reading. The first figure plots the empirical wealth share from the 2007 SCF (in dashed blue) against the simulated shares from the model (in solid 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%.

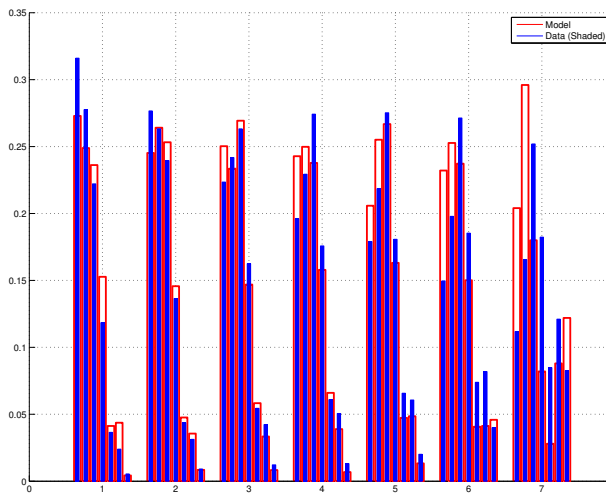


Figure 3: Wealth fractiles: baseline



Notes: On the horizontal axis we mark eight bars corresponding to the eight wealth distribution shares. The red bars with solid outlines are model simulated moments, while the blue bars with dashed outlines 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. The red hollow bars are model simulated moments, while the blue shaded bars are for empirical moments.

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, against which it is interesting to check how our estimates fare.

**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 on wealth accumulation with the capitalized income tax method; 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 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

Share of wealth	Fractile		
	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.

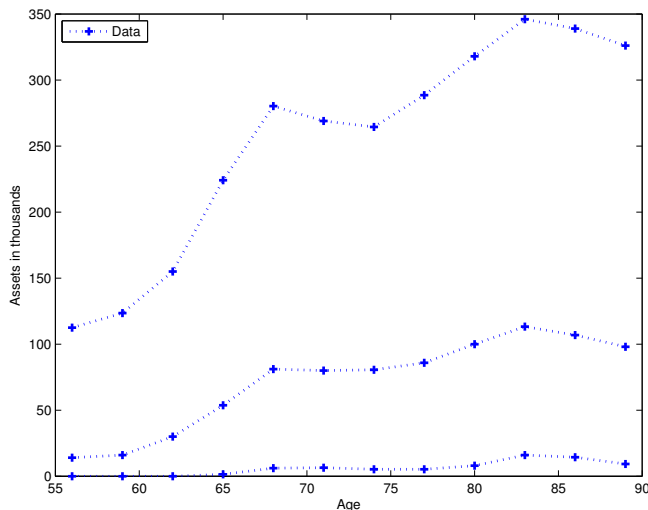
**Bequests.** The distribution of bequests in our model maps closely the stationary wealth distribution. It is also consistent with the distribution of actual and expected bequests in various micro-level data. For example, Hurd and Smith (2003) use the HRS to characterize such distributions and find the bequests are very skewed, just as the cross-sectional wealth.

A crucial implication of our estimates is that the savings rate, part of which is driven by the bequest motive, is increasing in wealth.<sup>19</sup> 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.<sup>20</sup>

<sup>19</sup>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).

<sup>20</sup>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.

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.

**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: <sup>21</sup>

---

<sup>21</sup>Charles and Hurst (2003) report two versions of the matrix estimates, adjusting logs of parental and child wealth by conditioning on different sets of variables. We use the one in which conditioning is limited to age, which is more consistent with our model.

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

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

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 (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 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. In the counterfactual with no capital income risk, we re-estimate the model under a constant rate of return. The estimated parameters are in Table 8, the simulation moments in Table 6-7, Line 2.<sup>22</sup> The differential savings factor does not substitute for capital income risk, as the higher estimated  $\mu$  implies. However the saving rate is higher due to a higher relative preference for bequests reflected in the estimate for  $A$ . Nonetheless, the model now misses completely in matching 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,

---

<sup>22</sup>The standard errors for the parameters in the counterfactual re-estimations tend to be large. In other words, the surface of the objective function is flat under the counterfactual restrictions, while it is not in our baseline estimation. This indicates that i) all the three factors are necessary to deliver an overall good match for the targeted moments; ii) when we shut down one of the factors, the other two end up being close substitutes in producing the overall (not as good) fit. We nonetheless base some of our comments on the point estimates.

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 bottom 25% too high.

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. The estimated parameters are in Table 9, the simulation moments in Table 6-7, 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. Once again, the model misses to match the top 1% of the wealth fractiles, which is reduced (though less than in the counterfactual with constant rate of return of wealth) to about 1/5th of the baseline (and the data). Apart 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.

In the counterfactual with no stochastic earnings, we set a constant labor earnings profile. We experiment with both a low and a high profile.<sup>23</sup> The estimated parameters are in Table 10-11, the simulation moments in Table 6-7, Line 4-5. The differential savings mechanism now, in either case, substitutes for the stochastic earning factor, as the lower estimated  $\mu$  implies. On the other hand the preference for bequests decreases in the low profile case. 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

---

<sup>23</sup>The high earnings profile is the 3rd highest one (the top two yield a nonstationary distribution) and the low earnings profile is the median one.

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 6: Wealth fractiles: counterfactual re-estimations

Share of wealth	<b>Distributional moments</b>							
	0-19	20-39	40-59	60-79	80-89	90-95	95-99	99-100
<i>Data</i> (SCF 2007)	-0.002	0.001	0.045	0.112	0.120	0.111	0.267	0.336
<i>Simulation</i>								
(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) $\mu = 2$	0.024	0.069	0.148	0.234	0.160	0.136	0.162	0.068
(4) Const. low $w$	0.153	0.174	0.165	0.168	0.157	0.093	0.034	0.057
(5) Const. high $w$	0.054	0.065	0.089	0.134	0.114	0.095	0.164	0.286

Table 7: Diagonal of transition matrix: counterfactual re-estimations

Share of wealth	<b>Mobility moments</b>						
	0-24	25-49	50-74	75-89	90-94	95-99	99-100
<i>Data</i>	0.316	0.263	0.263	0.176	0.066	0.082	0.083
<i>Simulation</i>							
(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) $\mu = 2$	0.283	0.230	0.258	0.167	0.032	0.077	0.102
(4) Const. low $w$	0.326	0.265	0.248	0.089	0	0.180	0
(5) Const. high $w$	0.149	0.292	0.135	0.245	0.149	0.418	0.508



Table 8: Parameter estimates: no stochastic rates of return

<b>Markov chain regime</b>					
<i>Preferences</i>	$\sigma$	$\mu$	$A$	$\beta$	$T$
	[2]	1.4969 (0.0017)	0.3148 (0.0090)	[0.97]	[36]
<i>Rate of return</i>					
<i>r</i> grid	0.034 (0.022)				
prob. grid	-				

Notes: The objective value is 0.2365.

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

<b>Parameters</b>					
<i>Preferences</i>	$\sigma$	$\mu$	$A$	$\beta$	$T$
	[2]	[2]	1.9794 (0.8955)	[0.97]	[36]
<i>Rate of return</i>					
<i>r</i> grid	0.007 (0.404)	0.014 (0.318)	0.026 (0.041)	0.067 (0.149)	0.091 (0.225)
prob. grid	0.061 (1.549)	0.411 (2.403)	0.499 (1.917)	0.148 (1.719)	0.137 (0.793)

Notes: The objective value is 0.1246.

Table 10: Parameter estimates: constant low  $w$  profile

<b>Parameters</b>					
<i>Preferences</i>	$\sigma$	$\mu$	$A$	$\beta$	$T$
	[2]	0.9681 (0.0077)	0.0145 (0.0182)	[0.97]	[36]
<i>Rate of return</i>					
<i>r</i> grid	0.003 (0.013)	0.009 (0.010)	0.030 (0.067)	0.046 (0.128)	0.088 (0.027)
prob. grid	0.045 (0.174)	0.418 (0.402)	0.395 (0.819)	0.301 (0.449)	0.183 (0.434)

Notes: The objective value is 0.2306.

Table 11: Parameter estimates: constant high  $w$  profile

<i>Preferences</i>	<b>Parameters</b>				
	$\sigma$	$\mu$	$A$	$\beta$	$T$
	[2]	0.2434 (0.8433)	0.0935 (0.1889)	[0.97]	[36]
<i>Rate of return process</i>					
$r$ grid	0.005 (0.170)	0.015 (0.002)	0.019 (0.188)	0.050 (0.396)	0.068 (1.541)
prob. grid	0.172 (1.028)	0.070 (1.315)	0.039 (0.426)	0.023 (2.262)	0.091 (2.201)

Notes: The objective value is 0.1415.

In summary, all the factors we study in our quantitative analysis, stochastic earnings, differential savings, and capital income risk, are crucial for generating the thick 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. These simulated models how-

ever appear driven by extreme assumptions either about the skewness of earnings (adding an *awesome* state) or about the working life of agents.

Kindermann and Krueger (2015), for instance, effectively estimate the (tail of the) earning process, not directly from earnings data but rather to fit the wealth distribution data (as we do for  $r$ ). Their estimates imply that, at the stationary distribution of earnings, the top 0.25% of agents earn on average somewhere between 400 to 600 times the median income, that is, at least \$20,000,000 (median earnings are about \$50,000). While substantial uncertainty pervades the data on top earners, this number appears implausibly high. In fact the top 0.1% of households have average incomes of about \$4,000,000 (out of which \$1,637,000 is earnings, defined as wages, salaries, and pensions), which is of course more than the average earnings of the top 0.25% (these calculations are according to Piketty-Saez's World Top Income Database for 2013-2014, which is not top-coded).

Another successful simulation exercise driven by skewed earnings is Díaz et al. (2003).<sup>24</sup> They also use an excessively skewed 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 \$367,000. Even if it were all earnings, it would correspond to no more than 7.5 times the median income. Similarly, in Castañeda et al. (2003), 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%. Even if the bottom 61% earns a low \$25,000 on average, the top 0.039% would have to earn \$25,000,000. This also appears implausibly high according to the World Top Income Database 2013-14.

A feature complementing skewed earnings in producing heterogeneity in wealth accumulation rates across agents is variable and random working life-span. Variable working life-spans produce a fraction of agents with high sojourn times in high earnings states, which implies in turn a better fit to the wealth distribution with less extreme skewness. For example even though their awesome state is less extreme than in the above cited literature,

---

<sup>24</sup>Dávila et al. (2012), Section 5.2, adopt this same calibration.

Kaymak and Poschke (2015) calibrate expected working lives to 45 years, with a constant exponential decay rate into retirement of  $\mu = 1/45$ . This implies a substantial fraction of agents with an excessive working life-span at the stationary distribution: over 100 years for 11% of the working population (of these 11%, those at the high earnings states populate the tail of the wealth distribution).<sup>25</sup>

We conclude that without extremely skewed labor earnings or extreme working life-span variability, it may not be possible for enough agents to accumulate enough wealth to populate the tail of wealth distribution via saved earnings alone. 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,<sup>26</sup> making it difficult to match the data with stochastic earnings only.

## 6 Transitional dynamics of the wealth distribution

As we have noticed, 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 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.<sup>27</sup>

---

<sup>25</sup>Kaplan et al. (2015) also use the same perpetual youth calibration as in Kaymak and Poschke (2015).

<sup>26</sup>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 for earnings.

<sup>27</sup>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

Table 12: Parameter estimates

Distribution + Mobility	Preferences				
Markov chain	$\sigma$	$\mu$	$A$	$\beta$	$T$
	[2]	1.2923	0.0109	[0.97]	[36]
	Rate of return process				
$r$ grid	0.000	0.009	0.021	0.054	0.080
prob. grid	0.160	0.429	0.230	0.273	0.026
stationary prob. grid	0.179	0.264	0.196	0.207	0.155

Distribution only	Preferences				
Markov chain	$\sigma$	$\mu$	$A$	$\beta$	$T$
	[2]	1.7610	1.2425	[0.97]	[36]
	Rate of return process				
$r$ grid	0.002	0.013	0.026	0.050	0.098
prob. grid	0.045	0.467	0.221	0.482	0.109
stationary prob. grid	0.144	0.258	0.177	0.266	0.155

Notes: 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 the 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 target the moments of the wealth distribution.

The estimated parameters are relatively close to those estimated in the baseline. We are then encouraged to conclude that the stationarity assumption imposed in the main quantitative analysis of this paper is not unreasonable.

We also note that the fit with the fractals of the wealth distribution is reasonably close in 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 mo-  
do not feed in the analysis the observed fiscal policy reforms since the '60's. Doing so should improve the fit.

ments match in the first case is compensated by a better job with the mobility moments (the simulated mobility moments in the second case are substantially larger than the empirical counterparts).

Finally, this exercise provides useful information about the implied speed of convergence of the model. In this respect, Gabaix et al. (2015) show that in continuous-time models the typical 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.<sup>28</sup>

Table 13: Wealth fractiles

Moments								
Share of wealth	0-20%	20-40%	40-60%	60-80%	80-90%	90-95%	95-99%	99-100%
<i>Data</i> (SCF 1962-63)	0.009	0.043	0.094	0.173	0.142	0.115	0.190	0.242
<i>Data</i> (SCF 2007)	-0.002	0.001	0.045	0.112	0.120	0.111	0.267	0.336
<i>Simulation</i>								
(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%
<i>Data</i>	0.316	0.263	0.263	0.176	0.066	0.082	0.083
<i>Simulation</i>							
(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

<sup>28</sup>A related mechanism plays the same role in producing faster convergence in Hubmer et al. (2015).

## 7 Conclusions

We estimate 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 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

- 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.
- 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-Giménez, 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-Víctor 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.
- 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,” 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.
- Kaplan, Greg, Benjamin Moll, and Giovanni L. Violante**, “Monetary Policy According to HANK,” 2015.
- Kaymak, Baris and Markus Poschke**, “The evolution of wealth inequality over half a century: the role of taxes, transfers and technology,” 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.
- 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.