Distribution of Wealth: Mechanisms

F. S. Fitzgerald: "The rich are different from you and me."
E. Hemingway: "Yes, they have more money."

Jess Benhabib, Alberto Bisin, Mi Luo
The Question

Which factors drive quantitatively the cross-sectional distribution of wealth in the U.S.?
Which factors drive, most notably, its skewed, thick, - asymptotically linear in the log-log plot - right tail?
Distributions of income and wealth which are very concentrated with thick right tails have been well documented over time and across countries: 
U.K. - Atkinson (2001),
Japan - Moriguchi-Saez (2005),
France - Piketty (2001),
U.S. - Piketty-Saez (2003),
Canada - Saez-Veall (2003),
Italy - Clementi-Gallegati (2004),
Norway - Dagsvik-Vatne (1999)
Historical notes on stationarity

Vilfredo Pareto introduced in the *Cours d’Economie Politique* (1897) the distribution which takes his name

$$ f(a) = \frac{\beta (a_{\min})^\beta}{a^{\beta+1}} \sim a^{-\beta-1}, \quad a \geq a_{\min} > 0 $$

$$ F(a) = 1 - \left( \frac{a_{\min}}{a} \right)^\beta $$

to represent empirical wealth distributions, characterized by heavy right tails:

$$ \lim_{a \to \infty} e^{\lambda a} Pr(a > a) = \infty, \quad \text{for all } \lambda > 0 $$

"Pareto’s Law," enunciated e.g., by Samuelson (1965):

In all places and all times, the distribution of income remains the same. Neither institutional change nor egalitarian taxation can alter this fundamental constant of social sciences.

F. S. Fitzgerald: "The rich are different from..."
What drives the (stationary) wealth distribution?

What are the possible driving factors of the wealth distribution? A few possible factors include:

- Skewed/persistent earnings, non-homogeneous bequests, differential savings, stochastic length of life/dynasty, the infamous $r > g$, (persistent) capital income risk, stochastic discount rates, ...  
- Theoretical models will let us categorize, organize, select, factors to put to data.

F. S. Fitzgerald: "The rich are different from me."  
E. Hemingway: "Yes, they have more money."
“Classic” dynamic models of Pareto distributions

- Kalecki (1945), Champernowne (1953), Rutherford (1955), Simon (1955), Wold-Whittle (1957), ...
- Stochastic processes which generate power laws in these “classic” models are essentially exogenous; the same can be said for a large recent literature on this topic in Econophysics.
Theoretical models

- The wealth accumulation equation is:

\[ a_{t+1} = r_{t+1} a_t + y_{t+1} - c_{t+1} \]

- If the consumption function is linear, \( c_t = \psi a_t + \chi \), the wealth accumulation equation now is:

\[ a_{t+1} = (r_{t+1} - \psi) a_t + (y_{t+1} - \chi) \]  \hspace{1cm} (1)
Consider economies with constant rate of return on wealth, \( r_t = r \) and linear consumption function:

\[
a_{t+1} = (r_{t+1} - \psi) a_t + (y_{t+1} - \chi)
\]  

(2)

Suppose the accumulation is contractive, \( r - \psi < 1 \), and \{\( y_t - \chi \)\} is i.i.d. with tail index \( \alpha \), then the accumulation equation 2 induces an ergodic stationary distribution for wealth \( a \) with tail \( \alpha \). [Grey, 1994].

In Grey’s Thm \((r_{t+1} - \psi)\) can be stochastic, as long as it does not "overwhelm" \((y_{t+1} - \chi)\). (Technically, \( E((r - \psi)^\alpha) < 1 \), \( E((r - \psi)^\beta) < \infty \) for \( \beta > \alpha \).)

However, if instead \( E((r_t - \psi)^{\alpha'}) = 1 \) and \( \alpha' < \alpha \), then the tail index of the stationary distribution of wealth will be \( \alpha' \). (Kesten).

Generalizations also allow persistent (finite Markov chain) \((r_{t+1} - \psi)\) and \((y_{t+1} - \chi)\) as well as correlations between them.

Basically the distribution of wealth inherits the tail of \((y_{t+1} - \chi)\).
Basic Aiyagari Models

- Basic Aiyagari Models with homothetic preferences are not linear but asymptotically linear.
- Under perfect capital markets and a neoclassical production technologies, these models generate well-defined stationary distributions, but the precautionary motive for savings declines at high wealth levels that are far away from the borrowing constraint.
- So Grey’s result for the tail carries over, unless additional features are introduced.
Carroll, Slacalek, and Tokuoka, *AER P&P*, 2014:

"However, the wealth heterogeneity in the $\beta$-Point (common discount rate) model essentially just replicates heterogeneity in permanent income.....; for example the Gini coefficient for permanent income measured in the Survey of Consumer Finances of roughly 0.5 is similar to that for wealth generated in the (common discount rate) model. Since the empirical distribution of wealth (which has the Gini coefficient of around 0.8) is considerably more unequal than the distribution of income (or permanent income), the setup only captures part of the wealth heterogeneity in the data, especially at the top."

- To remedy this, Carroll, Slacalek, and Tokuoka (2014) introduce heterogenous discount rates into the Aiyagari model.
Basic Aiyagari Models, Cont’d

Earnings alone are typically not enough

- For example Diaz-Gimenez, Mas-Pijoan, Rios-Rull (J.Mon Ec., 2003) or Davila, Hong, Krussell and Rios-Rull (Econometrica, 2012) use an earnings process where 6% of top wage earners make 46 times that of the median earner. This is one of the more moderately skewed calibrations.

- In the World Top Income Database of Piketty and Saez, 5% of top incomes average $367,000 in 2013 (not all of which is labor earnings) whereas the median incomes are about $50,000, a factor of 7.5, not 46.

- The exceptional earnings state at the upper tail then affects the whole distribution and is critical for generating the match to the empirical skew in the wealth distribution.

- So other additional features than earnings may be needed to microfound the wealth distribution.
Combining asymptotic linearity with random returns, it is possible to overcome results of Grey (1994) and others cited above to produce wealth tails fatter than those of earnings, as in Benhabib, Bisin and Zhu (2016) (using a generalization of Kesten results by Mirek (2011)).

Under some regularity conditions, the unique stationary distribution for wealth in the Aiyagari-Bewley model augmented with stochastic heterogeneous returns is unbounded above and has a fat tail.

We’ll discuss this (Kesten) approach in more detail in the context of our estimated model.

As we will see later, some earnings variability may be essential however.
To be explained as well: Social mobility

- Most studies of the wealth distribution center on the tail - hence on measures of inequality in the cross sectional distribution.
- But an advantage of working with formal macro models is that - once we allow for an explicit demographic structure - we also obtain implications for social mobility.
Theoretical models — Explanatory factors

What does it take to fit the distribution of wealth (that is, to obtain Pareto tails) in a standard macro model (that is, micro-founded):

- **Factor 1:** Skewed/persistent distribution of earnings - Diaz-Gimenez, Pijoan-Mas, and Rios Rull (2003); Castaneda, Diaz-Gimenez, Rios-Rull (2003; Kindermann and Krueger (2014).

- **Factor 2:** Differential saving rates across wealth levels - Atkinson (1973); Non-homogeneous bequests - Cagetti and DeNardi (2006), Piketty (2014)

- **Factor 3:** **Capital income risk** - Benhabib, Bisin, Zhu (2012); (2016), Entrepreneurship - Quadrini (2000), Cagetti and DeNardi (2003); Stochastic discount - Krusell and Smith (1988).
Another approach to get thick right tails in wealth is to introduce heterogeneous but fixed discount rates to create heterogeneity in impatience and in marginal propensities to consume, as in Carroll, Slajek and Tokuoka (2014b).

Such features that introduce additional heterogeneity across agents can generate fat tails in wealth, provided the distribution does not explode.

Heterogeneous marginal propensities to consume and savings rates can also be introduced through fixed costs of portfolio adjustment for high-return illiquid assets. Together with hand-to-mouth consumers, sufficient numbers of households with zero illiquid wealth can be maintained via "perpetual youth" demographics, generating heterogeneity in wealth accumulation across households: See Kaplan, Moll and Violante (2015).
Carroll, Slajek and Tokuoka (2014b) models introduce a constant probability of death for agents, replacing the dead by injecting new-born agents at low levels of wealth in order to prevent distributions from exploding over time.

Carroll, Slacalek, and Tokuoka, AER P&P, 2014: "The perpetual-youth mechanism of Blanchard (1985): To ensure that the ergodic cross-sectional distribution of permanent income exists, households die stochastically with a constant intensity and are replaced with newborns earning permanent income equal to the population mean. When the probability of dying is large enough, it outweighs the effect of permanent shocks and ensures that the ergodic distribution of income exists (and has a finite variance)."
Unlike the case where all agents are infinitely lived, variable life-spans can also produce differential sojourn times in the high earnings states in otherwise standard Aiyagari models. This leads to variation in wealth accumulation rates across agents.

Those who end up not only having long working-life spans, but are also lucky enough to spend a good deal of their working life in the high earnings states, end up working longer hours, and saving at higher rates in the high earnings states for precautionary and for retirement reasons.

They populate the tail of the wealth distribution, as discussed by Kaymak and Poschke (2015, p. 37).
For example Kaymak and Poschke (2015) calibrate expected working lives to 45 years, with a constant exponential decay rate into retirement of $\mu = 0.022 = 1/45$.

This however implies that at the stationary distribution 11% of the working population has been working for at least 100 years. (See also Kaplan, Violante and Moll (2015) for the same calibration).

A subset of those will spend long years in the extraordinary state, will build large wealth holdings, and will populate the tail of the wealth distribution.

In fact, Kaplan, Moll and Violante (2015) also make clear that they cannot match the very top tail of the wealth distribution with earnings alone, and suggest introducing alternative mechanisms like stochastic returns (see their footnote 32).

So without unrealistically skewed labor earnings, or extreme life-span variability, agents may not be able to accumulate enough wealth to populate the tail of wealth distribution via saved earnings alone.
By contrast, in a recent paper De Nardi, Fella, and Pardo (2016), adapt earnings data from Guvenen, Karahan, Ozkan, and Song (2015), and introduce it into a finite-life OLG model.

They note that earnings processes derived from data, including the one that they use,

"...when introduced in a standard quantitative model of consumption and savings over the life cycle, generate a much better fit of the wealth holdings of the bottom 60% of people, but vastly underestimate the level of wealth concentration at the top of the wealth distribution" (p.37).

They repeatedly stress that this can be due to many important mechanisms (bequests, entrepreneurship, inter-vivos transfers, etc.) from which their model abstracts, but which are discussed in Cagetti and De Nardi (2006, 2007, 2008) and De Nardi (2004).
Savings Rates Increasing in Wealth

- A related important mechanism to explain the skewness of wealth distribution is a savings rate that increases in wealth. The early work of Kaldor (1957, 1961), Pasinetti (1962) and Stiglitz (1969) emphasized this feature, in terms of discrete savings classes.
- Atkinson (1971) obtains a savings rate that increases continuously in wealth via the bequest motive. Consider Overlapping Generation economies with constant rate of return on wealth, finitely lived agents, and warm glow preferences for bequests:

\[ \nu(a_T) = A \frac{(a_T)^{1-\eta}}{1 - \eta} \]

- For these economies it is straightforward to show that, if the intertemporal consumption elasticity \( \sigma > \eta \), the propensity to consume out of wealth decreases in wealth. The increased concavity of consumption induces more skewness in wealth with respect to the linear case or the Bewley-Aiyagari model.

F. S. Fitzgerald: "The rich are different from you and me." E. Hemingway: "Yes, they have more money."
Cagetti and De Nardi (2008), who explicitly introduce non-homogeneous bequest motives:

\[ v(a_T) = A \left( 1 + \frac{a_T}{\gamma} \right)^{1-\sigma}, \]

where \( \gamma \) measures how much bequests increase with wealth. Savings therefore also increase in wealth due to the bequest motive.

The Cagetti and De Nardi (2008) approach combines non-homogeneity in bequests in an OLG model with skewed earnings.

A challenge for differential savings rate models may be that the rich dynasties may stay rich, so downward mobility may be too low relative to the data.
Returns as a Function of Wealth

- A further mechanism for increasing inequality of course is the rich making higher returns, having $r_t = r(a_t)$, for some increasing function $r$. (Interestingly, Kalecki (1937) did the opposite, $r$ decreasing in $a$ to tame Gibrat’s law.)
- In our empirical assessment at the end we will see that both stochastic earnings and savings rates increasing in wealth will play an important role in explaining the skewness of wealth distribution.
Back to wealth accumulation with stochastic returns

Some recent micro-founded models with stochastic returns:
1. Krusell and Smith (1997): *Stochastic heterogeneous discounting*
3. Benhabib, Bisin and Zhu (2011): *Finite life OLG with bequests*
Kesten results for

The linear accumulation equation for wealth,
\[ a_{t+1} = (r_{t+1} - \psi) a_t + (y_{t+1} - \chi) \]
defines a *Kesten process* if it satisfies the following (with some other regularity conditions):

1. \((r_t, y_t)\) are independent and i.i.d over time; and for any \(t \geq 0\):
2. \(0 < (y_t - \chi) < \infty\), (reflecting barrier)
3. \(0 < E(r_t) - \psi < 1\), (contracting on average)
4. \(\text{prob} (r_t - \psi > 1) > 0\). (escapes possible)

The stationary distribution for \(a_t\) can then be characterized as follows.

**Theorem (Kesten)**

A Kesten process displays an ergodic stationary distribution which has Pareto tail:

\[ \lim_{a \to \infty} \text{prob}(a_t \geq a) a^\alpha \sim k, \]

where \(\alpha > 1\) satisfies \(E(r_t - \psi)^\alpha = 1\).
Kesten results, Cont’d

- It should be noted that allowing for positive probabilities for negative additive shocks, \((y_{t+1} - \chi)\), in Kesten processes also induce a Double Pareto distribution:

\[
\lim_{a \to \infty} \text{prob}(a_t > a) a^\alpha = C_1 > 0
\]

and

\[
\lim_{a \to \infty} \text{prob}(a_t < a) a^\alpha = C_2 > 0
\]

with \(C_1 = C_2\) under regularity assumptions.

- Recent results extend the characterization result for generalized Kesten processes where \((r_t, y_t)\) is allowed to be a general Markov process, hence \(r_t\) correlated with \(y_t\) and both auto-correlated over time. In this case, \(\alpha > 1\) satisfies \(\lim_{n \to \infty} \left(E \prod_{j=0}^{n} (r_{-j})^\alpha\right) = 1\).

- Mirek (2011) extends Kesten results to asymptotically linear models.
Kesten results, continuous time: Saporta and Yao (2005).

Consider an accumulation process for each agent with wealth $a$:

$$da = r(X) \, adt + \sigma(X) \, d\omega$$

where $r(X), \sigma(X) > 0$, and $d\omega$ Brownian motion, which we can view as labor earnings minus the affine part of consumption. Consider $r(X)$ as the return on wealth net of the part of consumption proportional to wealth. Let $X$ be an exogenous a finite Markov chain. The usual Kesten assumptions require $E(r(x)) < 0$, and $\Pr(r(X) > 0) > 0$. Under some additional technical assumptions we have, as in the discrete time Kesten models, for $\alpha > 0$:

$$\lim_{a \to \infty} \text{prob}(a_t \geq a) a^\alpha \sim k, \quad k > 0$$

$$\lim_{a \to \infty} \text{prob}(a_t \leq -a) a^\alpha \sim k, \quad k > 0$$

- In fact, the results of Gabaix, Moll, Lasry and Lions (2015) can be obtained using this method. (They use pde's instead.)
An Estimation

The figure below displays the histogram for the wealth distribution, truncated at 0 on the left and ten million dollars on the right.

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.
Construct a finite life model with bequests nesting stochastic earnings, stochastic returns, wealth possibly increasing in wealth.

To estimate parameters target the distribution of wealth as well as wealth mobility.

The model is a generalization of Atkinson (1973) to incorporate the features above.
Each agent’s life is finite and deterministic: $T$ years. Consumer of dynasty $j$ chooses consumption $\{c_{j,t}\}$ and savings each period, subject to a no-borrowing constraint. Consumers also choose a bequest $e_{j,T}^n$.

We abstract from precautionary savings, so wage profiles and rates of return are drawn from distributions at the beginning of working-life. (See Huggett, Ventura and Yaron (2011), Keane and Wolpin (1997) or Cunha, Heckman and Schennack (2010)).
Model, Cont’d

- Assumption: Consumer of dynasty \( j \), generation \( n \), draw a lifetime return \( r^n_j \), a deterministic earnings profile \( \{y^n_{j,t}\}_{0}^{T} \) parameterized by \( y^n_{j,0} \), and maximizes utility (single heir, no estate tax for simplicity):

\[
V^n_j(a^n_{j,0}) = \max \left\{ c^n_{j,t} \right\} \sum_{t=0}^{T} \frac{(c^n_{j,t})^{1-\sigma}}{1-\sigma} + A\left(\frac{e^n_{j,T}}{1-\mu}\right) \quad \text{for} \ t \in [0, T]
\]

\[
s.t. \quad a^n_{j,t+1} = (1 + r^n_j)(a^n_{j,t} - c^n_{j,t}) + y^n_{j,t} \\
0 \leq c^n_{j,t} \leq a^n_{j,t}
\]

- If \( \mu < \sigma \), savings rates increase in wealth.
- Connecting generations: \( e^n_{j,T} = a^{n+1}_{j,0} \).
The solution of the recursive problem can be represented by a map

$$a_T = g(a_0; r, y).$$

Furthermore:

The map $g$ satisfies the following:

- If $\mu = \sigma$, $g(a_0; r, y) = \alpha(r, y)a_0 + \beta(r, y)$.
- If $\mu < \sigma$, $\frac{\partial^2 g}{\partial a_0^2}(a_0; r, y) > 0$. 

F. S. Fitzgerald: "The rich are different from you and me." E. Hemingway: "Yes, they have more money."
Earnings and Returns

The process for the rate of return of wealth and earnings processes over
generation $n$, $\{r^n_j, y^n_j\}$ is a finite irreducible Markov Chain with transition

$$P \left( r^n_j, y^n_j, 0 \mid r^{n-1}_j, y^{n-1}_j \right)$$

such that (abusing notation):

$$P \left( r^n_j \mid r^{n-1}_j, y^{n-1}_j \right) = P \left( r^n_j \mid r^{n-1}_j \right),$$

$$P \left( y^n_j, 0 \mid r^{n-1}_j, y^{n-1}_j \right) = P \left( y^n_j \mid y^{n-1}_j \right)$$

The life-cycle structure of the model implies that the initial wealth of the
$n$'th generation coincides with the final wealth of the $n - 1$'th generation:

$$a^n = a^n_0 = a_{T-1}^{n-1}.$$
We can construct then a stochastic difference equation for the initial wealth of dynasties, induced by the (forcing) stochastic process for \((r^n, y^n)\), and mapping \(a^{n-1}\) into \(a^n\):

\[
a^n = g \left( a^{n-1}; r^n, y^n \right),
\]

where the map \(g(.)\) represents the solution of the life-cycle consumption-saving problem.
Wealth dynamics across generations - cont. ed

• If $\mu = \sigma$, to induce a limit stationary distribution of $a_n$ it is required that the contractive and expansive components of the effective rate of return tend to balance.

• The distribution of $\alpha(r_n, y_n)$ display enough mass on $\alpha(r_n, y_n) < 1$ as well some as on $\alpha(r_n, y_n) > 1$; and that effective earnings $\beta(r_n, y_n)$ 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 can display a distinct expansive tendency acting against the stationarity of $a_n$. 

F. S. Fitzgerald: "The rich are different from me." E. Hemingway: "Yes, they have more money."
Wealth dynamics across generations

- The stochastic properties of labor income risk, $\beta(r^n, y^n)$, have no effect on the tail stationary distribution of wealth if it exists.

- 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 $\alpha(r^n, y^n) \leq \bar{\alpha} < 1$ and where $\bar{\beta}$ is the upper bound of $\beta(r^n, y^n)$, has a stationary distribution of wealth bounded above by $\frac{\bar{\beta}}{1-\bar{\alpha}}$.

- As long as a stationary distribution exists, wealth inequality (e.g., the Gini coefficient of the tail) increases with:
  - the capital income risk agents face in the economy, as measured by a "mean preserving spread" on the distribution of $\alpha(r^n, y^n)$,
  - the bequest motive $A$,
  - smaller $\mu$. 

F. S. Fitzgerald: "The rich are different from you and me.
E. Hemingway: "Yes, they have more money."
Quantitative exercise: Method of Simulated Moments

- We start with the assumption that the wealth and social mobility data observed in the U.S. are generated by a stationary distribution. Later we re-estimate dropping this assumption.
- In detail, our quantitative exercise is an application of the Simulated Method of Moments, whereby we fix several parameters of the model (externally calibrated), select some relevant moments, and estimate the remaining parameters by matching the moments generated by the model and those in the data.
Quantitative exercise

- Specifically, we fix $\sigma = 2$, $T = 36$, $\beta = 0.97$ per annum, the stochastic process for individual income and its transition across generations, following Chetty et al. (2014).
- We select the following wealth percentiles: bottom 20%, 20 – 40%, 40 – 60%, 60 – 80%, 90 – 95%, 95 – 99%, and top 1%, and the diagonal of the social mobility matrix, as the moments to match. We estimate $\mu, A$, 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.
We shall briefly discuss the input data of labor income processes and then the choice of output data for the targeted moments.

Matching the model and data generated moments requires wealth distribution and diagonal probabilities of the transition matrix for social mobility data.

We take wealth distribution data from the SCF, 2007.
Labor Income

Next we use ten deterministic life-cycle household-level income profiles at different quantiles, estimated from the PSID.

- Originally a 100-state Markov chain: each percentile of income distribution
- Reduce that to a 10-state Markov chain: each decile is a state for $y_0^n$. 

F. S. Fitzgerald: "The rich are different from me." E. Hemingway: "Yes, they have more money."
Transition matrix for labor decile $y_0^n$

$$T_{\text{gen}} =$$

$$\begin{bmatrix}
0.209 & 0.157 & 0.133 & 0.111 & 0.093 & 0.077 & 0.065 & 0.057 & 0.051 & 0.048 \\
0.176 & 0.150 & 0.131 & 0.112 & 0.098 & 0.085 & 0.074 & 0.065 & 0.057 & 0.052 \\
0.162 & 0.150 & 0.131 & 0.114 & 0.100 & 0.089 & 0.078 & 0.068 & 0.059 & 0.049 \\
0.121 & 0.128 & 0.124 & 0.116 & 0.108 & 0.100 & 0.092 & 0.082 & 0.072 & 0.058 \\
0.095 & 0.106 & 0.113 & 0.114 & 0.111 & 0.108 & 0.102 & 0.095 & 0.085 & 0.068 \\
0.076 & 0.089 & 0.099 & 0.107 & 0.111 & 0.112 & 0.112 & 0.108 & 0.101 & 0.085 \\
0.061 & 0.075 & 0.087 & 0.098 & 0.108 & 0.114 & 0.117 & 0.119 & 0.116 & 0.106 \\
0.049 & 0.063 & 0.076 & 0.090 & 0.104 & 0.116 & 0.124 & 0.129 & 0.129 & 0.122 \\
0.038 & 0.050 & 0.063 & 0.079 & 0.095 & 0.110 & 0.126 & 0.139 & 0.151 & 0.149 \\
0.028 & 0.035 & 0.046 & 0.059 & 0.072 & 0.088 & 0.107 & 0.135 & 0.175 & 0.256 \\
\end{bmatrix}$$
## Life-cycle earnings profiles

<table>
<thead>
<tr>
<th>Age range / %</th>
<th>0-10</th>
<th>10-20</th>
<th>20-30</th>
<th>30-40</th>
<th>40-50</th>
<th>50-60</th>
<th>60-70</th>
<th>70-80</th>
<th>80-90</th>
<th>90-100</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 [31-36]</td>
<td>-1.68</td>
<td>12.90</td>
<td>21.88</td>
<td>29.78</td>
<td>37.10</td>
<td>44.21</td>
<td>52.06</td>
<td>61.69</td>
<td>75.01</td>
<td>123.5</td>
</tr>
<tr>
<td>3 [37-42]</td>
<td>-1.73</td>
<td>13.48</td>
<td>23.84</td>
<td>32.88</td>
<td>41.35</td>
<td>49.64</td>
<td>57.95</td>
<td>68.42</td>
<td>84.67</td>
<td>153.8</td>
</tr>
<tr>
<td>4 [43-48]</td>
<td>-2.73</td>
<td>13.59</td>
<td>24.54</td>
<td>33.73</td>
<td>42.76</td>
<td>51.46</td>
<td>60.73</td>
<td>72.46</td>
<td>90.04</td>
<td>165.5</td>
</tr>
<tr>
<td>5 [49-54]</td>
<td>-4.97</td>
<td>10.47</td>
<td>20.95</td>
<td>29.68</td>
<td>38.81</td>
<td>47.98</td>
<td>57.98</td>
<td>69.65</td>
<td>87.23</td>
<td>165.2</td>
</tr>
<tr>
<td>6 [55-60]</td>
<td>-8.23</td>
<td>1.04</td>
<td>11.31</td>
<td>19.63</td>
<td>28.21</td>
<td>37.60</td>
<td>47.20</td>
<td>59.23</td>
<td>77.07</td>
<td>156.5</td>
</tr>
</tbody>
</table>

F. S. Fitzgerald: "The rich are different from us."
E. Hemingway: "Yes, they have more money."
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.
Cross-sectional wealth distribution: shares in bottom 20%, 20-40%, 40-60%, 60-80%, 80-90%, 90-95%, 95-99%, and top 1% of net worth holdings in the 2007 SCF.

Kennickell and Starr-McCluer (1997)'s estimates are robust:

- i) the matrix obtained by Klevmarken et al. (2003) with the PSID data is qualitatively similar;
- ii) the matrix estimated by Charles and Hurst (2003) to capture the inter-generational transmission in wealth exploiting information contained in the PSID about parent-child pairs is also similar.

In the estimation we are only matching the diagonal of the above matrix, and we impose the off-diagonal cells of each row to be equal. This assumption brings down the number of parameters we need to estimate.
Mobility Matrix

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

Two components of capital income are particularly subject to idiosyncratic risk:

- Ownership of principal residence and private business equity, which account for, respectively, 28.2% and 27% of household wealth in the US according to the 2001 Survey of Consumer Finances (SCF).

- Case and Shiller (1989) documented a 15% standard deviation of yearly capital gains or losses on owner-occupied housing; Flavin and Yamashita (2002) find a 14% standard deviation on the return on housing, at the level of individual houses, from the 1968-92 waves of the Panel Study of Income Dynamics (PSID).
Capital income risk - what is it?

- In the 1989 SCF studied by Moskowitz and Vissing-Jorgensen (2002) and Bitler, Moskowitz and Vissing-Jorgensen (2005) both capital gains and earnings on private business equity exhibit very substantial variation, as does excess returns to private over public equity investment, even conditional on survival.
- Private equity is highly concentrated: 75% owned by households for which it constitutes at least 50% of their total net worth.
- See Campbell, Lettau, Malkiel, and Xu (2001) for firm level return volatility.
- See also Quadrini (2000) and Cagetti and De Nardi (2006) on stochastic entrepreneurial returns in the US.
Table: Parameter estimates: baseline

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

Notes: [] indicates fixed parameters, standard errors computed with numerical derivatives for the parameter estimates in (). $\sigma$ 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.

- Rates of return for a 5 state Markov chain (equal off-diagonals).
- Mean real, after-tax, growth-detrended annual rate is 3.35%, close to i.i.d. in the stationary distribution.
- We estimate a modest 2.74% annualized standard deviation of returns.
With counterfactuals using re-estimated parameters

Distributional moments

<table>
<thead>
<tr>
<th>Moments</th>
<th>Share of wealth</th>
<th>0-20%</th>
<th>20-40%</th>
<th>40-60%</th>
<th>60-80%</th>
<th>80-90%</th>
<th>90-95%</th>
<th>95-99%</th>
<th>99-100%</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Data (SCF 2007)</strong></td>
<td>-0.002</td>
<td>0.001</td>
<td>0.045</td>
<td>0.112</td>
<td>0.120</td>
<td>0.111</td>
<td>0.267</td>
<td>0.336</td>
<td></td>
</tr>
<tr>
<td><strong>Simulation</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(1) Life cycle</td>
<td>0.014</td>
<td>0.048</td>
<td>0.105</td>
<td>0.168</td>
<td>0.102</td>
<td>0.070</td>
<td>0.151</td>
<td>0.341</td>
<td></td>
</tr>
<tr>
<td>(2) Const. (r)</td>
<td>0.026</td>
<td>0.089</td>
<td>0.159</td>
<td>0.208</td>
<td>0.137</td>
<td>0.161</td>
<td>0.177</td>
<td>0.045</td>
<td></td>
</tr>
<tr>
<td>(3) (\mu = 2)</td>
<td>0.027</td>
<td>0.083</td>
<td>0.176</td>
<td>0.275</td>
<td>0.158</td>
<td>0.100</td>
<td>0.139</td>
<td>0.042</td>
<td></td>
</tr>
<tr>
<td>(4) Const. low (w)</td>
<td>0.137</td>
<td>0.134</td>
<td>0.228</td>
<td>0.246</td>
<td>0.107</td>
<td>0.029</td>
<td>0.084</td>
<td>0.036</td>
<td></td>
</tr>
<tr>
<td>(5) Const. high (w)</td>
<td>0.007</td>
<td>0.100</td>
<td>0.100</td>
<td>0.149</td>
<td>0.106</td>
<td>0.085</td>
<td>0.120</td>
<td>0.333</td>
<td></td>
</tr>
</tbody>
</table>

Mobility Diagonals

<table>
<thead>
<tr>
<th>Moments</th>
<th>Share of wealth</th>
<th>0-24%</th>
<th>25-49%</th>
<th>50-74%</th>
<th>75-89%</th>
<th>90-94%</th>
<th>95-99%</th>
<th>99-100%</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Data</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Simulation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(1) Life cycle</td>
<td>0.274</td>
<td>0.263</td>
<td>0.269</td>
<td>0.158</td>
<td>0.047</td>
<td>0.041</td>
<td>0.122</td>
<td></td>
</tr>
<tr>
<td>(2) Const. (r)</td>
<td>0.435</td>
<td>0.242</td>
<td>0.339</td>
<td>0.236</td>
<td>0</td>
<td>0.061</td>
<td>0.256</td>
<td></td>
</tr>
<tr>
<td>(3) (\mu = 2)</td>
<td>0.273</td>
<td>0.268</td>
<td>0.264</td>
<td>0.153</td>
<td>0.064</td>
<td>0.030</td>
<td>0.022</td>
<td></td>
</tr>
<tr>
<td>(4) Const. low (w)</td>
<td>0.332</td>
<td>0.341</td>
<td>0.357</td>
<td>0.103</td>
<td>0.395</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>(5) Const. high (w)</td>
<td>0.342</td>
<td>0.368</td>
<td>0.275</td>
<td>0</td>
<td>0</td>
<td>0.286</td>
<td>0.662</td>
<td></td>
</tr>
</tbody>
</table>

- Benchmark is the Life-Cycle Model

F. S. Fitzgerald: "The rich are different from..."
Synthetic savings rates

- Synthetic savings rates group everyone within a wealth fractile and calculate the ratio between changes in total wealth and total income.
- Using 2000-2009 data, Saez and Zuchman (2014) found that synthetic savings rates are indeed increasing with wealth levels.

<table>
<thead>
<tr>
<th>Fractile</th>
<th>Share of wealth</th>
<th>Bottom 90%</th>
<th>Top 10-1%</th>
<th>Top 1%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data 2000-2009</td>
<td>-4%</td>
<td>9%</td>
<td>35%</td>
<td></td>
</tr>
<tr>
<td>Simulation</td>
<td>-5.65%</td>
<td>29.3%</td>
<td>42.2%</td>
<td></td>
</tr>
</tbody>
</table>

Synthetic saving rates for fractile $p$ in year $t$ is defined as $S_t^p = \frac{W_{t+1}^p - W_t^p}{Y_t^p}$, adjusted for changes over time in asset prices in data.

The synthetic savings rates in simulations are increasing in wealth, in line with the Saez-Zuchman computations, although they are high for the top 10%-1%.
The three lines represent retirement savings profiles for the 25%, median and 75%. 

F. S. Fitzgerald: "The rich are different from me." E. Hemingway: "Yes, they have more money."
## Estimation without Stationarity Assumption

### Distribution and Mobility

<table>
<thead>
<tr>
<th>Moments</th>
<th>Share of wealth</th>
<th>0-20%</th>
<th>20-40%</th>
<th>40-60%</th>
<th>60-80%</th>
<th>80-90%</th>
<th>90-95%</th>
<th>95-99%</th>
<th>99-100%</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Data (SCF 1962-63)</strong></td>
<td>0-20%</td>
<td>0.009</td>
<td>0.043</td>
<td>0.094</td>
<td>0.173</td>
<td>0.142</td>
<td>0.115</td>
<td>0.190</td>
<td>0.242</td>
</tr>
<tr>
<td></td>
<td>20-40%</td>
<td>0.001</td>
<td>0.015</td>
<td>0.045</td>
<td>0.112</td>
<td>0.120</td>
<td>0.111</td>
<td>0.267</td>
<td>0.336</td>
</tr>
<tr>
<td><strong>Simulation</strong></td>
<td>In 2 periods (72 yrs)</td>
<td>0</td>
<td>0</td>
<td>0.002</td>
<td>0.033</td>
<td>0.135</td>
<td>0.172</td>
<td>0.292</td>
<td>0.367</td>
</tr>
<tr>
<td></td>
<td>In 2 periods (distr. only)</td>
<td>0</td>
<td>0.003</td>
<td>0.025</td>
<td>0.080</td>
<td>0.155</td>
<td>0.150</td>
<td>0.269</td>
<td>0.319</td>
</tr>
</tbody>
</table>

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

---

F. S. Fitzgerald: "The rich are different from..."
Estimation without Stationarity Assumption

<table>
<thead>
<tr>
<th>Distribution + Mobility</th>
<th>Preferences</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Markov chain</strong></td>
<td>$\sigma$</td>
</tr>
<tr>
<td></td>
<td>[2]</td>
</tr>
<tr>
<td>Rate of return process</td>
<td></td>
</tr>
<tr>
<td>$r$ grid</td>
<td>0.000</td>
</tr>
<tr>
<td>prob. grid</td>
<td>0.160</td>
</tr>
<tr>
<td>stationary prob. grid</td>
<td>0.179</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Distribution only</th>
<th>Preferences</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Markov chain</strong></td>
<td>$\sigma$</td>
</tr>
<tr>
<td></td>
<td>[2]</td>
</tr>
<tr>
<td>Rate of return process</td>
<td></td>
</tr>
<tr>
<td>$r$ grid</td>
<td>0.002</td>
</tr>
<tr>
<td>prob. grid</td>
<td>0.045</td>
</tr>
<tr>
<td>stationary prob. grid</td>
<td>0.144</td>
</tr>
</tbody>
</table>

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.

For the case of Distribution+Mobility the expected value of $r$ is 3.28% and its standard deviation is 2.84%
Conclusion

- We estimated a macroeconomic model of the distribution of wealth in the U.S.
- While emphasize the tail of the distribution, the model performs well in hitting the whole distribution of wealth in the data.
- Importantly, the model is also successful in hitting the social mobility of wealth in the data.
- Capital income risk and differential savings are fundamental factors in explaining wealth distribution and social mobility (in the U.S.).
- Variable earnings are also essential but by themselves are not enough.
- Capital income risk estimates are roughly consistent with observations regarding return on real estate and private business equity.

F. S. Fitzgerald: "The rich are different from
Conclusion II

Figure: Awaiting the re-birth of socialism