NBER WORKING PAPER SERIES

SKEWED WEALTH DISTRIBUTIONS: THEORY AND EMPIRICS

Jess Benhabib
Alberto Bisin

Working Paper 21924
http://www.nber.org/papers/w21924

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
January 2016, Revised May 2017

Thanks to Mariacristina De Nardi, Steven Durlauf, Raquel Fernandez, Luigi Guiso, Dirk Krueger, Mi Luo, Ben Moll, Thomas Piketty, Alexis Toda, Shenghao Zhu, and Gabriel Zucman. We thank the Washington Center for Equitable Growth for financial support. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2016 by Jess Benhabib and Alberto Bisin. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Skewed Wealth Distributions: Theory and Empirics
Jess Benhabib and Alberto Bisin
NBER Working Paper No. 21924
January 2016, Revised May 2017
JEL No. B16,E13,E21,E24,E25

ABSTRACT

Invariably across a cross-section of countries and time periods, wealth distributions are skewed to the right displaying thick upper tails, that is, large and slowly declining top wealth shares. In this survey we categorize the theoretical studies on the distribution of wealth in terms of the underlying economic mechanisms generating skewness and thick tails. Further, we show how these mechanisms can be micro-founded by the consumption-saving decisions of rational agents in specific economic and demographic environments. Finally we map the large empirical work on the wealth distribution to its theoretical underpinnings.

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

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

F. S. Fitzgerald: The rich are different from you and me.

E. Hemingway: Yes, they have more money.[1]

1 Introduction

Income and wealth distributions are skewed to the right, displaying thick upper tails, that is, large and slowly declining top wealth shares. Indeed, these statistical properties essentially determine wealth inequality and characterize wealth distributions across a large cross-section of countries and time periods, an observation which has lead Vilfredo Pareto, in the *Cours d'Economie Politique* (1897), to suggest what Samuelson (1965) enunciated as the "Pareto's Law:"

> 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.[2]

The distribution, which now takes his name, is characterized by the cumulative distribution function

$$F(x) = 1 - \left(\frac{x_m}{x}\right)^{\alpha} \text{ for } x \in [x_m, \infty) \text{ and } x_m, \alpha > 0. \tag{1}$$

The "law" has in turn led to much theorizing about the possible economic and sociological factors generating skewed thick-tailed wealth and earnings distributions. Pareto himself initiated a lively literature about the relation between the distributions of earnings and wealth, i) whether the skewness of the wealth distribution could be the result of a skewed distribution of earnings, and ii) whether a skewed thick-tailed distribution of earnings could be derived from first principles about skills and talent. A subsequent literature exploited instead results in the mathematics of stochastic processes to derive these properties of distributions of wealth from the mechanics of accumulation.

Recently, with the distribution of earnings and wealth becoming more unequal, there has been a resurgence of interest in the various mechanisms that can generate the statistical properties of earnings and wealth distributions, resulting in new explorations, new data, and a revival of interest in older theories and insights. The book by Thomas Piketty (2014) has successfully taken some of this new data to the general public.[3]

[1] This often cited dialogue is partially apocryphal, see http://www.quotecounterquote.com/2009/11/rich-are-different-famous-quote.html?m=1

[2] The "law," here enunciated for income, was seen by Pareto as applying more precisely to both labor earnings and wealth.

[3] For an extensive discussion and some criticism of Piketty (2014), see Blume and Durlauf (2015); see also Acemoglu and Robinson (2015), Krusell and Smith (2014), and Ray (2014).

In this survey we concentrate only on wealth, discussing the distribution of earnings only inasmuch as it contributes to the distribution of wealth. More specifically, we aim at i) categorizing the theoretical studies on the distribution of wealth in terms of the underlying economic mechanism generating skewness and thick tails; ii) showing how these mechanisms can be micro-founded by the consumption-saving decisions of rational agents in specific economic and demographic environments; and finally we aim at iii) mapping the large empirical work on the wealth distribution to its theoretical underpinnings, with the ultimate objective of measuring the relative importance of the various mechanisms in fitting the data.[4]

In the following we first define what it is meant by skewed thick-tailed distributions and refer to some of the available empirical evidence to this effect regarding the distribution of wealth. We then provide an overview and analysis of the literature on the wealth distribution, starting from various fundamental historical contributions. In subsequent sections we explore various models of wealth accumulation which induce stationary distributions of wealth that are skewed and thick-tailed. Finally, we report on how various insights and mechanisms from theoretical models are combined to describe the empirical distributions of wealth.

1.1 Skewed and thick-tailed wealth distributions

A distribution is skewed (to the right) when it displays an asymmetrically long upper tail and hence large top wealth shares. The thickness of the tail refers instead to its rate of decay: thick (a.k.a. fat) tails decay as power laws, that is, more slowly than e.g., exponentially.

Formally, thick tails are defined as follows. Let a measurable function R defined on $(0,\infty)$ be *regularly varying with tail index* $\alpha \in (0,\infty)$ if

$$\lim_{x\to\infty} \frac{R(tx)}{R(x)} = t^{-\alpha}, \quad \forall t > 0.^{5}$$

Then, a differentiable cumulative distribution function (cdf) $F(x)$ has a power-law tail with index α if its counter-cdf $1 - F(x)$ is *regularly varying* with index $\alpha > 0$. We say that

> A distribution is thick-tailed if its cumulative $F(x)$ has a power-law tail with some index $\alpha \in (0,\infty)$.

A standard example is the Pareto distribution in (1). A distribution with a power-law tail has integer moments equal to the highest integer below α.[6] We also say that

[4] For an excellent survey of the mechanisms generating power laws in Economics and Finance, see Gabaix (2009).

[5] For $t > 1$, it is *slowly varying* if $\alpha = 0$ and *rapidly varying* if $\alpha = \infty$; see Resnick (1987), p.13-16.

[6] The Cauchy distribution, for instance, has a tail index of 1 and has no mean or higher moments.

a distribution is thin-tailed if it has all its moments, that is, $\alpha = \infty$: e.g., the normal, lognormal, exponential distributions are thin-tailed. Obviously, the smaller is α, the "thicker" is the tail.

As we noted, consistent with the "Pareto law," distributions of wealth are generally skewed and thick-tailed in the data, over countries and time. Skewness in the U.S. since the 60's is documented e.g., by Wolff (1987, 2004): the top 1% of the richest households in the U.S. hold over 33% of wealth; see also Kuhn and Ríos-Rull (2016).[7] Thick tails for the distributions of wealth are also well documented. Indeed, the top end of the wealth distribution in the U.S. obeys a power law (more specifically, a Pareto law): Using the richest sample, the Forbes 400, for the period 1988-2003, Klass et al. (2007) estimate a tail index equal to 1.49. Vermeulen (2015) adjusts estimates of the tail index for non-response rates for the very rich by combining the Forbes 400 list with the Survey of Consumer Finances and other data sets. He obtains estimates of the tail index in the range of $1.48 - 1.55$ for the U.S. Thick tails are also documented, for example by Clementi and Gallegati (2004) for Italy from 1977 to 2002, by Dagsvik and Vatne (1999) for Norway in 1998, and by Vermeulen (2015) for several European countries; see his Table 8.

2 Historical overview

In this section we briefly identify several foundational studies regarding the distribution of wealth. Indeed these studies introduce the questions and also the methods which a large subsequent literature picks up and develops.

2.1 Skewed earnings

The main question at the outset, since Pareto himself, is how to obtain a skewed thick-tailed distribution of wealth. Pareto assumed that a skewed distribution of labor earnings would map into a skewed distribution of wealth, focusing then on the determinants of skewed distributions of earnings. Pareto and a rich literature in his steps in turn explored whether some heterogeneity in the distribution of talents could produce a skewed labor earnings distribution.[8] Along similar lines Edgeworth (1917) proposed the *method of translation,* which consists in identifying distributions of talents coupled with mappings

[7] Kuhn and Ríos-Rull (2016) also report detailed statistics on the recent distribution on income (which include labor earnings and returns to wealth) by fractiles and Gini coefficients for the U.S., updated in 2013. Atkinson, Piketty, and Saez (2011) present an extensive historical survey of the evolution of top income across countries. A related literature investigates whether consumption is less unequal than income or wealth. Recent studies however show that consumption inequality closely tracks earnings inequality. See Aguiar and Bils (2011) and Attanasio, Hurst, and Pistaferri (2012).

[8] See Pareto (1897), notes to No. 962, p. 416.

from talents to earnings that, through a simple change of variable, yield appropriately skewed distributions of earnings.

More formally, the *method of translation* can be simply introduced. Suppose labor earnings y are constant over time and depend on an individual characteristic s according to a monotonic map g: $y = g(s)$. Suppose s is distributed according to the law f_s in the population. Therefore, from the standard change of variables for distributions, the distribution of labor earnings is:

$$f_y(y) = f_s\left(g^{-1}(y)\right)\frac{ds}{dy}.$$

For instance, if the map g is exponential, $y = e^{gs}$, and if f_s is an exponential distribution, $f_s(s) = pe^{-ps}$, the distribution of y is $f_y(y) = pe^{-p\frac{1}{g}\ln y}\frac{1}{g}\frac{1}{y} = \frac{p}{g}y^{-(\frac{p}{g}+1)}$, a power law distribution.

2.1.1 Models of skewed earnings

Several models of the determination of earnings have been proposed in the literature, which produce a skewed distribution induced by basic heterogeneities of productivity and talent. They link, through the *method of translation*, the thickness of the tail of the distribution of earnings to various different properties of the labor market.

Talent. The simplest application is due to the mathematician F.P. Cantelli (1921, 1929) and then refined by D'Addario (1943). Suppose talent, denoted by s, is exponentially distributed: $f_s(s) = pe^{-ps}$. Suppose also earnings y increase exponentially in talent: $y(s) = e^{gs}$, $g \geq 0$. As we have shown above, by a change of variables, $f_y(y) = \frac{p}{g}y^{-\left(\frac{p}{g}+1\right)}$, a power law distribution with exponent $\alpha = \frac{p}{g}$.[9]

Inspired by Edgeworth's (1896, 1898, 1899) critical comment of Pareto's work, that the lower earnings brackets does not follow a Pareto distribution, Frechet's (1939) model produces a hump-shaped distribution of earnings, with a left tail more akin to a log-normal than a power law. Indeed, suppose that the distribution of talent follows a Laplace distribution, $f_s(s) = \frac{1}{2}pe^{-p|s|}$, $s = (-\infty, +\infty)$. Maintaining earnings which increase exponentially in talent, $s = g^{-1}\ln y$ and $\frac{ds}{dy} = g^{-1}\frac{1}{y}$, we obtain, by *translation*,

$$f_y(y) = \frac{1}{2}\frac{p}{g}\frac{1}{y}e^{-p|\frac{\ln y}{g}|} = \begin{cases} \frac{1}{2}\frac{p}{g}y^{-(\frac{p}{g}+1)} & if \quad y \geq 1 \\ \frac{1}{2}\frac{p}{g}y^{\frac{p}{g}-1} & if \quad y \leq 1 \end{cases}.$$

[9] In fact Cantelli (1921, 1929) also provides a rationale for a negative exponential distribution of talent. Drawing on arguments by Boltzman and Gibbs, he shows that, if total talent is fixed, the most likely distribution of talent across a large number of individuals drawing earnings according to a multinomial probability from equally likely earnings bins is approximated by an exponential.

The distribution of earnings $f_y(y)$ is then a power law with exponent $\alpha = \frac{p}{g}$ above the median (normalized to 1) and it is increasing below the median as long as $\alpha > 1$.

Schooling. Suppose acquiring human capital involves i) an opportunity cost of time evaluated at discount rate $\frac{1}{1+r}$, and ii) a non-monetary marginal cost c, a measure of ability. Let h denote human capital, identified with years of schooling, and let $y(h)$ denote labor earnings for an agent with human capital h. Then, the competitive equilibrium condition in the labor market is

$$y(h)e^{-rh} = y(0).$$

If the marginal cost of acquiring human capital through schooling, c, is exponentially distributed, $f_c(c) = pe^{-pc}$, so are years of schooling, h, in equilibrium. Then the same *transformation* algebra used for talent in the previous example implies then that y has a distribution even more skewed than a power law with exponent $\alpha = \frac{p}{r}$. This is essentially Mincer's (1958) schooling model;[10] see also Roy (1950) for an extension to multi-dimensional ability.[11]

Span of control. Let an entrepreneur with talent s be characterized by the opportunity to hire n agents at wage x to produce with production function $f(n, s) = sn^{\alpha}$. This entrepreneur's earnings $y(s)$ will satisfy

$$y(s) = \max_{n \geq 0} \begin{cases} x & if \quad n = 0 \\ sn^{\alpha} - xn & else \end{cases}.$$

It follows that, for any $n > 0$, $y(s) = A(x)s^{\frac{1}{1-\alpha}}$, where $y(s)$ is a convex function which amplifies differences in talent s. If s is uniformly distributed with support $[1, b]$, transforming variables produces a truncated power law distribution of earnings:

$$f_y(y) = \frac{B}{b-1} y^{-\alpha},$$

over support $\left[C, b^{\frac{1}{1-\alpha}} C\right]$, where B and C are constants depending on the parameters x, α.[12]

[10] In Mincer (1958)'s analysis, however, ability, and hence human capital, are normally distributed for $h \geq 0$. As a consequence, y has a log-normal distribution in the tail, since $\ln y = \ln \underline{y} + rs$.

[11] More specifically, Roy (1950) postulates that human capital depends on an index of ability composed of the sum of several multiplicative i.i.d. components (intelligence, perseverance, originality, health etc). If these are normally distributed, or under assumptions for the Central limit theorem to apply, earnings are approximately lognormal. However as Roy notes, if components of talent are correlated, the distribution of earnings is more skewed than log-normal (see Roy's reference to Haldane, 1942).

[12] If s is instead exponentially distributed, the *transformation* generates a Weibull distribution of earnings (with decreasing density).

Assortative matching. Suppose the expected output of firms, $E(Y)$, is determined by an "O-Ring" production function, as in Kremer (1993):

$$E[Y] = k^a(h_1 h_2 \ldots h_m)mB,$$

where k denotes capital, h_i is the human capital of the worker the firm assigns to task i, m is the total number of tasks, and B is a firm productivity parameter. We look for a competitive equilibrium of the labor market in which earnings do not depend on tasks. At such an equilibrium, $y(h)$ represents workers' earnings as a function of their human capital. Firms then choose $h_1, h_2 ... h_m$, and k to maximize

$$\max \; E(Y) - y(h_1) - y(h_2) - ...y(h_m) - rk.$$

Because of the complementarity between the human capital of workers in different tasks which characterize the O'Ring production, that is, because $\partial^2 E[y]/\partial h_i \partial h_j > 0$, in equilibrium workers of the same human capital will be matched assortatively. Letting $h_i = h$, for $i = 1, \ldots, m$, the first order conditions for profit maximization imply then $mBh^{m-1}(ah^m Bm/r)^{a/(1-a)} - \frac{dy(h)}{dh} = 0$, a differential equation whose solution is

$$y(h) = (1-a)(h^m B)^{1/(1-a)}(am/r)^{a/(1-a)}.$$

The equilibrium earnings function $y(h)$ is homogeneous of degree $m/(1-a) > 1$ in h: small differences in skills h translate into large differences in earnings y. Indeed $y(h)$ is a convex function, so that labor earnings y are skewed to the right even if h is distributed symmetrically.[13] Consider again for instance the case in which h is uniformly distributed: $f_h(h) = \frac{1}{b}$, $0 \leq h \leq b$. Then, by *transformation*,

$$f_y(y) = \frac{1}{b} C y^{-\left(\frac{a-1}{m} - 1\right)},$$

a truncated power law over support $\left[0, b^{\frac{m}{1-a}} D\right]$, where C and D are constant depending on parameters a, B, m, r.

Hierarchical production (Lydall (1959))**.** Suppose production is structured in hierarchical levels, $1, \ldots, I$, where lower indexes correspond to lower positions in the hierarchy to which a higher number of people, $n_i > n_{i+1}$ are assigned. Suppose also that the technology requires $n_i = \gamma n_{i+1}$, for some $\gamma > 1$. Finally, suppose earnings at level $i+1$, y_{i+1} are proportional to earnings in the contiguous lower level i (this could be the case,

[13] Since the production function exhibits decreasing returns to scale, firms will have positive profits. But even if redistributed to the agents in general equilibrium, these profits do not constitute labor earnings but rather capital income.

e.g., if higher level workers manage lower level ones): $y_{i+1} = q\gamma y_i$, with $q \leq 1,\ q\gamma \geq 1$. It follows then that

$$\ln\left(\frac{n_{i+1}}{n_i}\right) = -\frac{\ln\gamma}{\ln\gamma + \ln q}\ln\left(\frac{y_{i+1}}{y_i}\right).$$

In the discrete distribution we have constructed, n_i is the number of agents with earnings y_i. It is clear that a discrete power law distribution, $n_i = B(y_i)^{-\frac{\ln\gamma}{\ln\gamma+\ln q}}$, for some constant B and $\frac{\ln\gamma}{\ln\gamma+\ln q} \geq 1$. [14]

2.1.2 Thickness of the distribution of earnings

The models of skewed earnings surveyed in Section 2.1.1 link the exponent α to various structural parameters characterizing the labor market that produces earnings. We review in this section the implications of these models regarding the thickness of earnings distributions.

In the Talent model, $\alpha = \frac{p}{g}$ and hence the earnings distribution is thicker when the earnings map is steeper in talent (g is high), or when the density of talent decreases relatively slowly (p is small). In the schooling model, human capital replaces talent in the determination of the thickness of earnings distribution and $\alpha \geq \frac{p}{r}$. The earnings distribution is then thicker when the earnings map is steeper in human capital, that is, when the rate of return r is high and agents need to be compensated more for the opportunity costs of accumulating the human capital. It is also thicker when the density of human capital decreases relatively slowly (p is small).

In the Span of control model, instead, earnings distributions are thicker the lower are the decreasing returns in production (the lower is α). A related result holds in the Assortative Matching model. In this case $\alpha = \frac{1-a}{m}$ and earnings distributions are thicker the lower are the decreasing returns in production (the higher is a), and the more specialized is human capital (the higher are the number of tasks m).

Earnings are distributed like power laws with exponent α in these models. In general, however, a power law (for example a Pareto distribution) is well defined over an unbounded support only for $\alpha \geq 0$. Otherwise the distribution does not have a finite integral unless its support is truncated, that is, defined on a bounded support. This is the case for the distributions of earnings we obtained in the Span of Control and Assortative Matching production models.[15] In all these cases in fact the density of the distribution is a power function with exponent < 1 over finite support. The implied thickness is larger than the thickness of any power law with exponent $\alpha > 0$.[16]

[14] Note that if $q = 1$, we get Zipf's Law.

[15] We thanks Francois Geerolf for this observation.

[16] More precisely in this case we say a truncated power law $F^T(y)$ over $[a, b]$, $b > a$, is thicker than a power law $F(y)$ as there exists an $\varepsilon' > \epsilon > 0$ such that $F^T(b-\varepsilon) > F(b-\varepsilon)$ and where $F^T(b) = F(\infty)$, normalized to 1 without loss of generality.

A related more recent literature has developed which obtains thick-tailed earnings endogenously. Along the lines of the Span of Control model, Gabaix and Landier (2008) exploit assortative matching between firms and their executives to produce a Pareto distribution of the earnings of executives. More specifically, in Gabaix and Landier (2008) the more talented executives are matched with larger firms, which results in executive earnings y increasing in firm size S: $y(S) = S^{\beta},\ S \geq S_{\min} > 0, \beta \geq 0$. Suppose firm size is Pareto distributed with exponent $\gamma > 1$, $f(S) = QS^{-\gamma}$. Then, by *transformation*, earnings are also Pareto, with exponent $\alpha = \frac{\gamma-1}{\beta}$:

$$f_y(y) = f(S(y))\frac{dS}{dy} = \frac{Q}{\beta}y^{-\left(\left(\frac{\gamma-1}{\beta}\right)+1\right)}.$$

This model induces thicker earnings the thicker is the distribution of firms size (the smaller is γ) and the steeper are earnings as a function of size (the higher is β). Interestingly, the distribution of earnings is power law even if earnings are concave in size S, that is, if $\beta < 1$. Finally, in the Hierarchical Production model, $\alpha = \frac{\ln\gamma}{\ln\gamma+\ln q} - 1 \geq 0$ since $\gamma \geq \gamma q \geq 1$ and thickness increases with the depth of the hierarchical structure, γ, and the steepness of the earning map with respect to the hierarchical level, q.

Geerolf (2016) obtains instead power law earnings in a model of one-dimensional knowledge or skill hierarchies (rather than task specialization) with workers and layers of management endogenously sorted, incorporating span of control and assortative matching within the firm.

2.2 Stochastic returns to wealth

The literature focusing on the factors determining skewed thick-tailed earnings distribution tended to disregard the properties of wealth accumulation. Motivated by the empirical fact that wealth generally tends to be much more skewed than earnings, an important question for the subsequent literature has been whether a stochastic process describing the accumulation of wealth could amplify the skewness of the earnings distribution. Alternatively, could skewed wealth distributions become skewed due to factors unrelated to skewed earnings distributions? Several accumulation processes have been proposed to study these questions.

Indeed Champernowne (1953) introduces a wealth accumulation process which contracts on average, but, due to stochastic returns on wealth, nonetheless induces a stationary distribution of wealth with a thick tail. More specifically Champernowne (1953) divides wealth into bins,[17] with a bottom bin from which it is only possible to move up, acting as a reflecting barrier. While the overall average drift is assumed to be negative, there are positive probabilities for moving up to the higher bins. Champernowne (1953)

[17] In fact Champernowne (1953) applied the process to earnings rather than wealth, but the logic of the result is invariant.

shows that this stochastic process generates a Pareto distribution of wealth. Formally, the wealth bins, indexed by $i = 0, 1, 2, 3, ...$, are defined by their lower boundaries:

$$w(i) = w(0)\, e^{ai}, \quad i = 1, 2, 3... \tag{2}$$

and $w(0) > 0$ is the lowest bin. With the exception of the lowest bin, the probability for moving up (resp. down) a bin is p_1 (resp. p_{-1}), while the probability of staying in place is p_0, with $p_{-1} + p_0 + p_1 = 1$. The number of people at bin $j = 0, 1, 2..$ at time t, n_t^i, is given by

$$\begin{aligned} n_{t+1}^0 &= p_{-1} n_t^1 + (p_0 + p_{-1})\, n_t^i, \\ n_{t+1}^i &= p_1 n_t^{i-1} + p_{-1} n_t^{i+1} + p_0 n_t^i, \quad i \geq 1; \end{aligned}$$

where the adding up constraint is $\sum_{i=0}^{\infty} n_t^i = \sum_{i=0}^{\infty} n_{t+1}^i = n$. The stationarity condition, that the number of people moving away from a bin must be offset by those incoming at each t, takes then a simple form,

$$p_{-1} n^{i+1} - (p_{-1} + p_1) n^i + p_1 n^{i-1} = 0, \; i \geq 1.$$

Champernowne shows, as it can be verified by direct substitution, that this condition implies that a stationary wealth distribution must satisfy $n^i = q \left(\frac{p_1}{p_{-1}}\right)^i$, for q appropriately chosen. Letting $\frac{p_1}{p_{-1}} = e^{-\lambda}$, and after a transformation of variables using equation (2),

$$\begin{aligned} n^i &= q \left(\frac{p_1}{p_{-1}}\right)^i = q e^{-\frac{\lambda}{a} \ln \frac{w(i)}{w(0)}} \left(\frac{1}{a} (w(i))^{-1}\right) \\ &= = \frac{q}{a} \frac{w(0)^{-\frac{\lambda}{a}}}{w(i)^{\frac{\lambda}{a}+1}}; \end{aligned}$$

which defines a Pareto distribution, with exponent $\alpha = \frac{\lambda}{a}$ and $\sum_{i=0}^{\infty} n^i = n$.[18] Champernowne (1953) also shows that a stationary wealth distribution exists if and only if $p_1 < p_{-1}$ (that is, wealth is contracting on average).

Champernowne's approach, foreshadowing the subsequent mathematical results of Kesten (1973), is at the core of a large literature exploiting the mathematics of wealth accumulation processes with a stochastic rate of return of the form:

$$w_{t+1} = \begin{cases} r_{t+1} w_t & \text{for } r_{t+1} w_t > \underline{w} \\ w_t & \text{for } r_{t+1} w_t \leq \underline{w}, \end{cases}$$

[18] Champernowne also considered a two sided Pareto distribution with two-sided tails, one relating to low incomes and one to high incomes. To obtain this, he eliminated the reflecting barrier, imposing instead a form of "non-dissipation:" a negative drift for bins above a threshold bin and a positive one for lower bins.

where $r_t \geq 0$ and i.i.d., and $\underline{w} > 0$. We discuss several examples in the next section. Importantly, Champernowne's result that stationarity requires wealth to be contracting on average holds robustly, as these processes induce a stationary distribution for w_t if $0 < E(r_t) < 1$. Furthermore, for the stationary distribution to be Pareto it is required that $prob\,(r_t > 1) > 0$, an assumption also implicit in the accumulation process postulated by Champernowne.

2.3 Explosive wealth accumulation

One central issue in this literature is the stationarity of the wealth distribution. Indeed skewed wealth distributions can be easily obtained for explosive wealth accumulation processes over time, but these processes do not necessarily converge to a stationary wealth distribution.

As the simplest example, consider the wealth accumulation equation:

$$w_{t+1} = r_{t+1} w_t$$

(the economy has no labor earning, $y_t = 0$, for simplicity and without loss of generality). The wealth process is non-stationary, trivially, when rate of return is deterministic, $r_{t+1} = r$, and $r > 1$. But this is also the case if r_{t+1} is Normal i.i.d. and $E(r_t) > 1$. The wealth process satisfies then what is generally referred to as *Gibrat's Law:*[19] at each finite time t, it induces a log-normal distribution around it mean at t, with a mean and variance increasing and exploding in t ,

$$\ln w_t = \ln w_0 + \sum_{j=0}^{t-1} \ln r_j.$$

The variance of wealth explodes and no stationary distribution of wealth exists.[20]

This logic clearly illustrates that an expanding wealth accumulation process can coexist with stationary wealth distribution only in conjunction with some other mechanism to tame the tendency of these processes to become non-stationary. Consistently, in Wold and Whittle (1957) it is a *birth and death* process which tames the possible non-stationarity and induces a Pareto distribution for wealth.[21]

[19] From Gibrat (1931).

[20] Economic forces might however produce a stationary distribution of wealth that tames the exploding variance resulting from proportional growth. Kalecki (1945) proposed to this effect a mean rate of return appropriately decreasing in wealth, e.g., $\ln r_t = -\alpha \ln w_t + z_t$. The resulting negative correlation between r_t and w_t could induce a constant variance in the distribution of wealth. It is straightworward to show that this is in fact the case if z_t is i.i.d. and $\alpha = \frac{\sum(\ln r_t)^2}{2\sum\left(\ln w_t^i\right)^2}$. Benhabib (2014a) obtains the same result by means of progressive taxation of capital income. This line of argument has not been much followed recently because a decreasing net rate of return in wealth appears counterfactual.

[21] An early version of a related birth and death model giving rise to a skewed distribution was also proposed by Rutherford (1955).

Consider an economy with a constant explosive rate of return on wealth, $r > 1$, and no earnings, $y = 0$. In each period individuals die with probability γ, in which case their wealth is divided at inheritance between $n > 1$ heirs in an Overlapping Generations framework. The accumulation equation for this economy is therefore

$$w_{t+1} = \begin{cases} rw_t & \text{with prob. } 1-\gamma \\ \frac{1}{n}w_t & \text{with prob. } \gamma \end{cases}$$

and population grows at the rate $\gamma(n-1)$. By working out the master equation for the density of the stationary wealth distribution associated to this stochastic process (after normalizing by population growth), $f_w(w)$, and guessing $f_w(w) = w^{\alpha-1}$, Wold and Whittle (1957) verify that a solution exists for α satisfying $\frac{r}{\gamma}\alpha = n(1-n^{-\alpha})$. The tail α depends then directly on the ratio of the rate of return to the mortality rate, $\frac{r}{\gamma}$; see Wold and Whittle (1957), Table 1, p. 584. To guarantee that the stationary wealth distribution characterized by density $f_w(w)$ is indeed a Pareto law, Wold and Whittle (1957) need to formally introduce a lower bound for wealth $\underline{w} \geq 0$. Such lower bound effectively acts as a reflecting barrier: below $\underline{w}$ the wealth accumulation process is arbitrarily specified so that those agents whose inheritance falls below $\underline{w}$ are replaced by those crossing $\underline{w}$ from below, keeping the population above $\underline{w}$ growing at the rate $\gamma(n-1)$.

The *birth and death* mechanism introduced by Wold and Whittle (1957) is at the core of a large recent literature on wealth distribution which we discuss in subsequent sections. In particular, to guarantee stationarity all these models need to introduce, besides birth and death, a mean-reverting force (e.g., some form of reflecting barrier) to ensure that the children's initial wealth is not proportional to the final wealth of their parents for all the agents in the economy. Furthermore, the sign of the dependence of the Pareto tail on r and γ also turns out to be a robust implication of this class of models; see the discussion in Section 2.1.3.

2.4 Microfoundations

The theoretical models of skewed earnings in this early literature, as well as models of stochastic accumulation, often tend to be very mechanical, engineering- or physics-like in fact. This was duly noted and repeatedly criticized at various times in the literature. Assessing his "method of translation," Edgeworth (1917) defensively writes:

> It is now to be added that our translation has the advantage of simplicity. Not dealing with differential equations, it is more accessible to practitioners not conversant with the higher mathematics.

Most importantly, these models were criticized for lacking explicit micro-foundations and more explicit determinants of earnings and wealth distributions. Mincer (1958) writes:

> From the economist's point of view, perhaps the most unsatisfactory feature of the stochastic models, which they share with most other models of personal income distribution, is that they shed no light on the economics of the distribution process. Non-economic factors undoubtedly play an important role in the distribution of incomes. Yet, unless one denies the relevance of rational optimizing behavior to economic activity in general, it is difficult to see how the factor of individual choice can be disregarded in analyzing personal income distribution, which can scarcely be independent of economic activity.

Similarly, Becker and Tomes (1979) were also critical of models of inequality by economists like Roy (1950) or Champernowne (1953) for having neglected the intergenerational transmission of inequality by assuming that stochastic processes largely determine inequality through distributions of luck and abilities. They complain that:

> "[...] mechanical" models of the intergenerational transmission of inequality that do not incorporate optimizing responses of parents to their own or to their children's circumstances greatly understate the contribution of endowments and thereby understate the influence of family background on inequality.

The criticisms by Mincer and Becker and Tomes were especially influential. Beginning in the 1990s, they lead economists to work with micro-founded models of stochastic processes of wealth dynamics and optimizing heterogenous agents.

3 Theoretical Mechanisms for the Skewed Distribution of Wealth

In this section we identify the distinct theoretical mechanisms responsible for thick-tailed distributions of wealth. Various combinations of these mechanisms drive the modern theoretical and especially empirical literature attempting to account for the shape of wealth distribution. We follow the structure of the historical contributions laid out in the previous section. We start with models that describe the wealth distribution as induced by the distribution of labor earnings $\{y_t\}$. We then introduce models of skewed thick-tailed wealth distributions driven by individual wealth processes which contract on average down to a reflecting barrier, but expand with positive probability due to random rates of return $\{r_t\}$. Such models can be considered variations and extensions of Champernowne (1953). We finally study models in which skewed thick-tailed wealth distributions are obtained by postulating expansive accumulation patterns on the part of at least a subclass of agents in the economy. As noted, these models by themselves may

not induce a stationary wealth distribution and are therefore often accompanied by birth and death processes which indeed re-establish stationarity. These are in effect variations on Wold and Whittle (1957). We then discuss models where preferences induce savings rates that increase in wealth and can contribute to generating thick tails in wealth, with expanding wealth checked again by birth and death processes (or by postulating finite lives).

These models are generally micro-founded, so that assumptions on preferences (including bequests), financial markets, and demographics guarantee that wealth accumulation is the outcome of savings behavior which constitutes the solution of an optimal dynamic consumption-savings problem. Formally, consider an economy in which i) wealth at the end of time t, w_t, can only be invested in an asset with return process $\{r_{t+1}\}$; and ii) the earning process is $\{y_{t+1}\}$. Let c_{t+1} denote consumption at $t+1$, so that savings at $t+1$ is $y_{t+1} - c_{t+1}$. The wealth accumulation equation is then:

$$w_{t+1} = r_{t+1} w_t + y_{t+1} - c_{t+1}. \tag{3}$$

3.1 Linear savings

Suppose consumption (hence savings) is linear in wealth, $c_{t+1} = \psi w_t + \chi_{t+1}$, and assume $\psi,\ \chi_{t+1} \geq 0$. For these economies, equation (3) becomes:

$$w_{t+1} = (r_{t+1} - \psi)\, w_t + \left(y_{t+1} - \chi_{t+1}\right). \tag{4}$$

In this section we show that, while the environments and underlying assumptions of most micro-foundations of wealth accumulation models do not induce an exact linear consumption function, this is a very useful benchmark to establish some of the basic properties of wealth accumulation processes. Consider economies populated by agents with identical Constant Relative Risk Aversion (CRRA) preferences over consumption at any date t,

$$u(c_t) = \frac{c_t^{1-\sigma}}{1-\sigma},$$

who discount utility at a rate $\beta < 1$. We maintain the assumption that wealth at any time can only be invested in an asset paying constant return r. We distinguish in turn between infinite horizon and overlapping generations economies.

3.1.1 Infinite horizon

Consider an infinite horizon Bewley-Aiyagari economy.[22] Under CRRA preferences, each agent's consumption-savings problem must satisfy a borrowing constraint and $\beta E(r_t) < 1$. The borrowing constraint together with stochastic earnings generates a precautionary motive for saving and accumulation and acts as a lower reflecting barrier for assets.

Consider first the case in which the rate of return is deterministic, $r_t = r$. [23] The consumption function $c(w_t)$ is concave and the marginal propensity to consume declines with wealth, as the precautionary motive for savings declines with higher wealth levels far away from the borrowing constraint. While the model is non-linear, the consumption function is asymptotically linear in wealth:

$$\lim_{w_t \to \infty} \frac{c(w_t)}{w_t} = \psi.^{24}$$

The additive component of consumption, χ_{t+1} in (4), can be characterized at the solution of the consumption-savings problem. It reflects the fraction of discounted sum of earnings consumed, as well as precautionary savings. Indeed, the optimal choice of χ_{t+1} depends on the the stochastic process for $\{y_t\}$, for example for an $AR(1)$ process on its persistence, and on the volatility of its innovations. In the very stylized case in which $y_t \geq 0$ is deterministic, growing at some rate λ, and where $\lambda\beta r = 1$, with CRRA or with Quadratic utility, we have $\chi_{t+1} = y_{t+1}$. When the income process $\{y_t\}$ is stochastic, optimal savings include a precautionary component that can depend on wealth w_t.[25] This is the case under CRRA utility for example that belongs to the decreasing absolute risk aversion class, even though consumption and savings are asymptotically linear in wealth in this case.[26]

[22] From Bewley (1983) and Aiyagari (1994); see also Huggett (1993). These economies represent some of the most popular approaches of introducing heterogeneity into the representative infinitely-lived consumer; see Aiyagari (1994) and the excellent survey and overview of the recent literature of Quadrini and Rios-Rull (1997).

[23] These economies easily extend to include production. In fact, under a neoclassical production function, the marginal product of capital converges to r at the steady state and $\beta r < 1$ holds because capital also provides insurance against sequences of bad shocks.

[24] See Benhabib, Bisin and Zhu (2015) for a formal proof.

[25] In some specifications, consumption decisions c_t are taken at the beginning of the period before earnings y_t are realized. Then in an optimizing framework current earnings realizations would not affect current consumption.

[26] Some further intuition can be developed if we use to quadratic utility, $\left(-\left(b_1 c_t - b_2 (c_t)^2\right)\right.$, which yields certainty equivalence, as well as analytical results with linear consumption and savings functions, as in (4) above. Linear consumption policies obtained with quadratic preferences give rise to a wealth accumulation process that is stationary (rather than a random walk) if $\beta r > 1$, and under certainty equivalence precautionary savings that depend on wealth levels are avoided.(See Zeldes, 1989, for differences in consumption policies under quadratic and CRRA preferences.) If we assume that earnings are *iid*, and that consumption c_t is chosen at the beginning of the period before the earnings y_t are

More generally, as far as the right tail of wealth is concerned, the asymptotic linearity of $c(w_t)$ guarantees that equation (4) approximates wealth accumulation in the economy. The condition $r - \psi < 1$ is an implication of $\beta r < 1$ under CRRA preferences. With constant $(r - \psi)$, the right tail of the wealth distribution is therefore the same as that of the stationary distribution of $\{y_t - \chi_t\}$. Therefore, $\{\chi_t\}$ determines the divergence between the right tails of wealth and earnings. Specifically,for example if $\chi_t = y_t$ for all t, the distribution of wealth does not have a thick tail (the tail index is ∞).[27] Alternatively, as discussed further below in section 3.2, if χ_t is a constant that just shifts the distribution of $\{y_t\}$ to the left, the right tail of wealth will be no thicker than the right tail of earnings y_t.

The more general case in which returns are stochastic has essentially the similar micro-foundations. Infinite horizon economies with CRRA preferences and borrowing constraints still display a concave, asymptotically linear consumption function, as the precautionary motive dies out for large wealth levels.[28]

3.1.2 Overlapping generations (OLG).

Let n denote a generation (living for a length of time T). A given intra-generation earnings profile, $\{y_n\}_t$, can be mapped into lifetime earnings, y_n. Also, a lifetime rate of return factor r_n can be constructed from the endogenous consumption and bequest pattern.[29] The initial wealth of each dynasty maps then into a bequest T periods later, which becomes the initial condition for the next generation. The inter-generational wealth accumulation equation is linear in this economy, that is equation (4) holds inter-generationally: $w_{n+1} = (r_n - \psi) w_n + \left(y_{n+1} - \chi_{n+1}\right)$. The details of these arguments and closed form solutions are derived in Benhabib, Bisin, and Zhu (2011).

Importantly, because of the OLG structure, no restriction is required on $E(r_n)$, nor borrowing constraints need be imposed.

---

observed, then $\chi_t = aE(y_t) - k$ where a and k are positive constants and k goes to zero as b_1 goes to zero. A disadvantage of quadratic utility however is that for large wealth and therefore consumption above the "bliss point," marginal utility can become negative, creating complications. For an excellent recent treatment of the Markovian income process see Light (2016).

[27] When the earnings distribution is a finite Markov chain however it is necessarily thin-tailed and typically all its moments exist.

[28] The wealth distribution in this class of economies has been studied, for example, by Benhabib, Bisin and Zhu (2015), and Achdou, Han, Lasry, Lions, Moll (2016), in discrete and continuous time, respectively and solved numerically by Nirei and Souma (2016).

[29] The constructiuon is simpler under the assumption that the rate of return is constant in t, though generally stochastic over generations n.

3.2 Skewed Earnings

A general characterization of the stationary distribution for $\{w_t\}$ induced by Eq. 4 will be introduced in Section 3.3.2, Theorem 3, due to Grey (1994). In this section, however, we study the simple special case where $r_t = r$, a deterministic constant.

Theorem 1 *Suppose $0 < r - \psi < 1$ and $\{y_t\}$ has a stationary distribution with a thick tail with tail-index β. Then the accumulation equation, (4), induces an ergodic stationary distribution for wealth with right tail index α not thicker than β: $\alpha \geq \beta$.*

More precisely, the stationary distribution of wealth has a right tail-index equal to the right tail-index of the stationary distribution of the stochastic process $\{y_t - \chi_t\}$. However if $\chi_t \geq 0$ the tail index of wealth matching that of the stationary distribution of $\{y_t - \chi_t\}$ can be no thicker than the distribution of earnings.[30] In other words, under our assumptions for contracting economies with constant rates of return and linear consumption with $\chi_t \geq 0$, the statistical properties of the right tail of the wealth distribution are directly inherited from those of the distribution of earnings. As a consequence, the tail of the wealth distribution cannot be thicker than the tail of the distribution of earnings.

The wealth distribution in economies with heterogenous agents and (exogenous) stochastic earnings has been studied, for example, by Diaz et al (2003), Castaneda, Diaz-Jimenez, and Rios-Rull (2003).

3.3 Stochastic returns to wealth

An important contribution to the study of stochastic processes which has turned out to induce many applications to the theoretical analysis of wealth distributions is a result which obtains for the linear accumulation equation, (4), when the rate of return r_t follows a well-defined stochastic process.

Equation (4) defines a *Kesten process* if i) (r_t, y_t) are independent and $i.i.d$ over time; and if ii) satisfies:[31]

$$y > 0, \;\; 0 < E(r_t) - \psi < 1, \text{ and } prob\,(r_t - \psi > 1) > 0,$$

for any $t \geq 0$.

[30] Let $f(y_t)$ be the density of earnings y_t with a thick tail. If $\chi_t \geq 0$ and $s_t = y_t - \chi_t$, then $h(s_t) = f(y_t)$ is a left shift of the density $f(y_t)$. So if $f'(y_t) \leq 0$ in the tail,then $h(s_t + \chi_t) = h(y_t) \leq f(y_t)$, and the tail of $y_t - \chi_t$ is no thicker than that of y_t. From the definition of power laws in section 1.1, $\lim_{y\to\infty} \frac{f(qy)}{f(y)} = q^{-\alpha}$ for $q, > 0, \alpha \geq 0$ and so $\frac{f(qy)}{f(y)} < 1$ (> 1) if $q > 1$ (< 1). Then indeed $\lim_{y\to\infty} f'(y) < 0$ so for some x, $f'(y) \leq 0$ for $y \geq x$.

[31] Some other regularity conditions are required; see Benhabib, Bisin, Zhu (2011) for details.

These assumptions guarantee, respectively, that earnings act as a reflecting barrier in the wealth process and that wealth is contracting on average, while expanding with positive probability.

The stationary distribution for w_t can then be characterized as follows.

Theorem 2 (Kesten) *Suppose the accumulation equation, (4), defines a* Kesten process *and* $\{y_t\}$ *has a thin right tail. Then the induced wealth process displays an ergodic stationary distribution with Pareto tail* α*, where* $\alpha > 1$ *solves*

$$E\left(r_t - \psi\right)^{\alpha}) = 1.^{32}$$

A stochastic rate of return to wealth can generate a skewed and thick-tailed distribution of wealth even when neither the distribution of r_t nor the distribution of earnings are thick-tailed.[33]

An heuristic sketch for a proof of Kesten (1973) in a very simple case can be given along the lines of Gabaix (1999, Appendix)). Consider the special case in which i) $\left(y_{t+1} - \chi_{t+1}\right)$ is constant, equal to $\bar{y} > 0$; and ii) $\lambda_t = \left(r_{t+1} - \psi\right)$ is *i.i.d.*, and $E\left(r_{t+1} - \psi\right) < 1$. If $\bar{y} > 0$, and $w_0 \geq 0$, then $w_t \geq \bar{y}$. Then the master equation for the dynamics can be written as:

$$P\left(w_{t+1}\right) = \int_0^{\infty} \frac{P\left(w_t/\lambda - \bar{y}\right)}{\lambda} f\left(\lambda\right) d\lambda$$

where $P\left(w_t\right)$ is the density of w_t, to be solved for. For large w we can ignore $\bar{y}$ which becomes insignificant relative to w, and conjecture that we can approximate the stationary distribution with $P\left(w\right) = Cw^{-\alpha-1}$, a power law over $[\bar{y}, \infty)$. Then for large w at the stationary distribution, $Cw^{-\alpha-1} = \int_0^{\infty} C\left(w^{-\alpha-1}\lambda^{\alpha+1}\right) f\left(\lambda\right) d\lambda$, where α solves $1 = \int_0^{\infty} \lambda^{\alpha+1} f\left(\lambda\right) d\lambda$. This is Kesten's result in this simplified case: the tail index a solves $E\left(\lambda^{\alpha+1}\right) = 1$. Note that, since $\int_0^{\infty} f\left(\lambda\right) d\lambda = 1$, for a solution with $\alpha + 1 > 0$

[32] Allowing for negative earning shocks, so that $\Pr\left(\left(r_t - \psi\right) < 0\right) > 0$, and without borrowing constraints, Kesten processes induce a two-sided Pareto distribution,

$$\lim_{\bar{w}\to\infty} prob(w > \bar{w})\bar{w}^{\alpha} = C_1, \quad \lim_{\bar{w}\to\infty} prob(w < -\bar{w})\bar{w}^{\alpha} = C_2$$

with $C_1 = C_2 > 0$ under regularity assumptions (see Roitershtein (2007), Theorem 1.6). This extension addressing at least in part Edgeworth's criticism of Pareto, was anticipated by Champernowne (1953) (see footnote 18); see also Benhabib and Zhu (2008), as well as Alfarano, Milakovic, Irle, and Kauschke (2012), Benhabib, Bisin, and Zhu (2016a); Toda (2012).

[33] This result is generalized by Mirek (2011) to apply to asymptotically linear accumulation equations. This is important in this context because asymptotic linearity is the property generally obtained in micro-founded models, as we have shown in Section 3.1. Furthermore, for the study of wealth distributions, recent results extend the characterization result for *generalized Kesten processes* where (r_t, y_t) may be driven by a Markov process, hence r_t can be correlated with y_t, and furthermore both r_t and y_t can be auto-correlated over time (see Roitershtein (2007)). In this case, α solves $\lim_{N\to\infty} E\left(\prod_{n=0}^{N-1}\left(r_{-n} - \psi\right)^{\alpha}\right)^{1/N} = 1$.

we need $\Pr(\lambda > 1) > 0$. Thus for large w the stationary distribution is approximated by a power-law with index a if there is a reflecting barrier $\bar{y} > 0$, $E(\lambda) < 1$, and the probability of growth is positive, that is $\Pr(\lambda > 1) > 0$.

The Kesten result has important implications for a characterization of the tail of the induced distribution of wealth, depending on the stochastic properties of the rate of return process r_t. More specifically, it can be shown that the distribution of wealth has a thicker tail (the α which solves $E((r_t - \psi)^\alpha) = 1$ is lower) the more variable is r_t, in terms of second order stochastic dominance; see Benhabib, Bisin, Zhu (2011), Proposition 1.

Nirei and Souma (2007) used Kesten processes to study wealth accumulation and its tail in a model with stochastic returns that is not microfounded. Wealth distribution of economies with stochastic returns in microfounded models has been studied, for example, in discrete time, by Quadrini (2000), Benhabib, Bisin, and Zhu (2011, 2015, 2016), Fernholz (2016), and Wälde (2016). Krusell and Smith (1998), have studied a related economy with stochastic heterogenous discount rates.[34]

3.3.1 Stochastic Returns in Continuous Time

The Kesten result in Theorem 2 can also be extended to continuous time under different sets of assumptions. We survey them in the following.[35]

The stationary distribution of wealth is a power law when the wealth accumulation process is defined by:

$$dw = r(X)\, w dt + \sigma(X)\, d\omega, \tag{5}$$

where X is an exogenous Markov jump process, $E(r(X)) < 0$, $\Pr(r(X) > 0) > 0$, $\sigma(X) > 0$, and $d\omega$ is a Brownian motion.[36] Here $r(X)$ can be interpreted as the stochastic net rate of return on wealth, and $E(r(X)) < 0$ assures that the process is contractionary on average.

The stationary distribution of wealth is a power law also when the wealth accumulation process is a generalized "geometric" Ornstein-Uhlenbeck (OU) process:

$$dw = (\mu - rw)dt + \sigma w d\omega, \tag{6}$$

where $\mu, r, \sigma > 0$.[37] In this case, while the drift $(\mu - rw)$ becomes negative for large w,

[34] See also Angeletos and Calvet (2005, 2006), Angeletos (2007) and Panousi (2008).

[35] We keep this section rather informal as the study of wealth distribution is mostly developed in discrete time in the economics literature. But we carefully reference the relevant results in mathematics.

[36] See Saporta and Yao (2005).

[37] More precisely, the stationary distribution induced by Equation (6) is an inverse Gamma, $f(w) = \left(\frac{\sigma^2}{2\mu}\right)^{-\left(\frac{2r}{\sigma^2}+1\right)} \left(\Gamma\left(\frac{2r}{\sigma^2}+1\right)\right)^{-1} w^{-\frac{2r}{\sigma^2}-1} e^{-\frac{2\mu}{\sigma^2}w}$, where Γ is the gamma function. Since $e^{-\frac{2\mu}{\sigma^2}w} \to 1$ as $w \to \infty$, the tail index of the stationary distribution of this process is $\frac{2r}{\sigma^2}$. More generally, see Borkovec and Klüpperberg (1998), p. 68, and Fasen, Klüpperberg and Lindner (2006), p.113, for a characterization

the drift $\sigma w d\omega$ is multiplicative in wealth and hence acts like a stochastic return on w.[38]

Finally, the stationary distribution of wealth is a power law also when the wealth accumulation process is a standard OU process

$$dw = (\mu - rw)dt + \sigma d\nu, \tag{8}$$

driven by a Levy jump process with positive increments $d\nu$ rather than a Brownian motion.[39]

The wealth distribution of economies with stochastic returns has been studied in continuous time, for example by Benhabib and Zhu (2008), Achdou, Lasry, Lions, and Moll (2014), Gabaix, Lasry, Lions, and Moll (2015), Aoki and Nirei (2015), Benhabib, Bisin and Zhu (2016). In particular, Gabaix, Lasry, Lions and Moll (2015) study stochastic processes of the type given by Equation (5). They apply Laplace transforms methods to characterize the speed of convergence of the distribution of wealth to the stationary distribution in response to changes in underlying parameters.

3.3.2 Stochastic returns and skewed earnings

We have seen in Theorem 1 that linear (or asymptotically linear) wealth accumulation processes in economies with deterministic returns and skewed thick right-tailed distributions of earnings induce wealth distributions with right tails at most as thick as those of earnings (Theorem 1). We have also seen that when returns are stochastic and earnings are thin tailed, the stationary wealth distribution can have thick tails (Theorem 2). A natural question is what happens in economies with both stochastic returns and thick-tailed earnings. How thick is the tail of the wealth distribution in this case? The result, from Grey (1994), is the following.

Theorem 3 *Suppose $(r_t - \psi)$ and $(y_t - \chi_t)$ are both random variables, independent of w_t. Suppose the accumulation equation 4 defines a* Kesten process *and $(y_t - \chi_t)$ has a thick right-tailed with tail-index $\beta > 0$. Then,*

If $E\left((r_t - \psi)^{\beta}\right) < 1$, and $E\left((r_t - \psi)^{\gamma}\right) < \infty$ for some $\gamma > \beta$, under some regularity assumptions, the right-tail of the stationary distribution of wealth will be β.

---

of heavy-tailed stationary distributions induced by

$$dw = (\mu - rw)dt + \sigma w^{\gamma} d\omega, \tag{7}$$

for $\gamma \geq 0.5$. Cox, Ingersoll and Ross (1985) model interest rates as driven by a proces as (7) with $\gamma = 0.5$. In this connection see also Conley et al (1997). For further applications of these results in economics, see Luttmer (2012, 2016)).

[38] The standard OU process, with drift $\sigma d\omega$, induces a Gaussian stationary distribution for w .

[39] See Barndorff- Nielsen and Shepard (2001) or Fasen, Klüpperberg and Lindner (2006)).

If instead $E\left((r_t - \psi)^{\gamma}\right) = 1$ for $\gamma < \beta$, then the right-tail index of the stationary distribution of wealth will be $\alpha = \gamma$.

Theorem 3 makes clear that the right-tail index of the wealth distribution induced by Equation 4 is either γ, which depends on the stochastic properties of returns, or β, the right-tail of earnings $\{y_t - \chi_t\}$).[40] Note that if we assume $\chi_t \geq 0$, the right tail of $\{y_t - \chi_t\}$ is no thicker than that of $\{y_t\}$.[41] Then it is never the case that, for a stochastic process describing the accumulation of wealth, the tail index of earnings could amplify the right-tail index of the wealth distribution; it's either the accumulation process or the skewed earnings which determine the thickness of the right- tail of the wealth distribution.

These asymptotic results on right tails however do not specify the wealth level at which the right tail starts, and in principle the tails could be very far to the right, raising the question of their empirical relevance when data is finite. Section 4.1 however indicates that indeed it is very hard to get actual earnings distributions to produce the top wealth shares in the data, unless top earnings are augmented to induce thickness in the distribution of earnings largely in excess of that which can be documented in earnings data.

3.4 Explosive wealth accumulation

Even without a skewed distribution of earnings and a Kesten process for w_t, a skewed distribution of wealth might be obtained if a non-contracting process for (r_t, y_t) is postulated which does not satisfy the Kesten conditions. Equation (4) defines a *explosive process* if i) (r_t, y_t)

$$y > 0, \;\; 1 < E(r_t) - \psi$$

for any $t \geq 0$.

Theorem 4 *Suppose the accumulation equation, (4), defines an explosive process. Then the induced wealth process is non-stationary, independently of the distribution of y_t.*

This is the case, for instance, if i) the rate of return is deterministic, $r_t = r > 1$;[42]; and if ii) returns to wealth follow a stochastic process inducing an accumulation equation following Gibrat's Law.

While general results are nor available for non-linear accumulation equations, it is straightforward that a non-stationary distribution of wealth is also induced when i)

[40] See Ghosh et al. (2010) for extensions of Grey (1994) to random, Markov-dependent (persistent), and correlated coefficients $(y_t - \chi)$ and $(r_t - \psi)$; and see Hay et al. (2011) for the multivariate case.

[41] See the discussion in Section 3.2 and footnote 3.2.

[42] Even if only for a sub-class of the agents in the economy.

consumption is strictly concave (hence savings strictly convex) increasing in wealth, that is $c_{t+1} = \psi(w_{t+1})w_{t+1}$, $\psi(w) > 0$, $\psi'(w) < 0$ and/or ii) the rate of return on wealth is increasing in wealth, that is $r_{t+1} = r_{t+1}(w_{t+1})$, with $r_{t+1}(w)$ increasing in w such that $\lim_{w\to\infty} r_{t+1}(w - \psi(w)) > 1$.

3.4.1 Birth and death processes

As we noted discussing Wold and Whittle (1957) in Section 1.2, a number of birth and death mechanisms can be super-imposed onto explosive economies to generate a skewed stationary distribution of wealth.

The simplest micro-founded model which illustrates the power of birth and death processes to tame the non-stationarity of wealth accumulation is Blanchard (1985).[43] The economy in the model is characterized by a deterministic explosive rate of return t and perpetual youth, that is, constant mortality rate p. Indeed, the only stochastic variable generating wealth heterogeneity is the Poisson death rate. Agents receive constant earnings y, face a constant return on wealth r and a fair rate p from an annuity on their accumulated wealth. The discount is θ but agents discount the future at rate $\theta + p$, reflecting their mortality rate. [44] Consumption is linear in $w + h$, where $h = \frac{y}{r+p}$ is the present discounted value of earnings, and wealth w_t satisfies

$$w_t = \left(\frac{y}{r+p}\right)\left(e^{(r-\theta)t} - 1\right).^{45}$$

It is assumed that dying agents are replaced with newborns, so population size is constant normalized at p^{-1} and the age density is exponential: $n(t) = pe^{-pt}$. Newborns start life with exogenous initial wealth $\underline{w}$.[46] In this model, therefore, wealth is increasing in age t and age is distributed exponentially. Applying Edgeworth *translation* method, the distribution of wealth is:

$$f_w(w) = \left(p\frac{(r+p)}{y(r-\theta)}\right)\left(\frac{w}{y}(r+p)+1\right)^{-\left(\frac{p}{r-\theta}+1\right)},$$

which is a power law in the tail, that is, for large w, with exponent $\alpha = \frac{p}{r-\theta} > 0$.[47]

[43] Castaneda, Diaz-Jimenez, and Rios-Rull (2003) and Carroll, Slajek, and Tokeu (2014b) also make use of microfounded versions of the perpetual youth model combined with skewed random earnings.

[44] Several recent papers use features of the perpetual youth model to obtain thick tails; see for example Benhabib and Bisin (2006), Benhabib, Bisin and Zhu (2016a), Piketty and Zucman (2015), and Jones (2015), Toda (2014, 2015).

[45] See Benhabib and Bisin (2006) and Benhabib and Zhu (2009), where the full optimization dynamics is spelled out in a more general stochastic continuous time model.

[46] See Benhabib and Bisin (2006) for the endogenous determination of $\underline{w}$ via a social security system funded by taxation.

[47] Note the stationary distribution of wealth will be well-defined with a positive Pareto exponent if $r > \theta$, that is as long as the return r is smaller than the effective discount rate: $r < \theta + p$.

In this mode, the thickness of the tail of the distribution therefore increases with the rate of return r and decreases with the death rate p, just as in Wold and Whittle (1957). Death rates can check unbounded growth and induce a stationary tail in the distribution. The re-insertion of (at least some) newborns at a wealth level $\underline{w}$ independent of their parents' wealth (a reflecting barrier) is crucial, however, as it is in Wold and Whittle (1957). [48] The model implies that wealth will be correlated with age (or, in extensions with bequests, with the average life-span of ancestors).[49]

In continuous time for an accumulation process following Gibrat's Law, it is still the case that a birth-death process can re-establish stationarity of the wealth distribution. This is clearly demonstrated by Reed (2001). Consider exponentially distributed death times, with re-insertion at initial wealth w_0.[50] Assuming wealth evolves in continuous time, with a constant positive drift (rate of return) r, and a geometric Brownian motion as diffusion, Reed (2001) obtains a log-normal distribution for wealth w_T, where T denotes the time of death. Assuming T is exponentially distributed, $f_T = pe^{-pT}$, and integrating, he obtains:

$$f_w(w) = \int_0^\infty pe^{-pT} \frac{1}{\sigma w \sqrt{2\pi T}} e^{-\left(\frac{\left(\ln w - \left(w_0 + \left(r - \frac{\sigma^2}{2}\right)T\right)\right)^2}{2\sigma^2 T}\right)} dT \tag{9}$$

with solution

$$f_w(w) = \begin{cases} \frac{\alpha\beta}{\alpha+\beta}\left(\frac{w}{w_0}\right)^{\beta-1}, & \text{for } w < w_0 \\ \frac{\alpha\beta}{\alpha+\beta}\left(\frac{w}{w_0}\right)^{-\alpha-1}, & \text{for } w \geq w_0 \end{cases}$$

where $(\alpha, -\beta)$ solve the quadratic $\frac{\sigma^2}{2}z^2 + \left(r - \frac{\sigma^2}{2}\right)z - p = 0$. Note that the density of wealth $f_w(w)$ is increasing in wealth for $w < w_0$, if $\beta > 1$. As Reed (2001) notes, this is

[48] The re-insertion of newborns at a wealth level corresponding to a fixed fraction of the wealth of their parents at death however would simply dilute the growth rate on average, but would be insufficient to guarantee stationarity.

[49] If we allow earnings y to grow exogenously at rate g the stationary distribution of wealth discounted at the the rate g will still have the same Pareto tail exponent. $\frac{p}{r-\theta}$. The discounted wealth of an agent of age t is now $\tilde{w} = e^{-gt}w = y\left(\frac{1}{r+p-g}\right)\left(e^{(r-\theta)t} - 1\right)$ and the *discounted* distribution of wealth is given by $f_w(w) = p\frac{(r+p-g)}{y(r-\theta)}\left(\frac{\tilde{w}}{y}(r+p-g)+1\right)^{-\left(\frac{p}{r-\theta}+1\right)}$, assuming that growing earnings discounted at the effective return, $r+p$, do not explode, that is $r+p > g$ Thus for a given growth rate of earnings g, increasing r results in thicker wealth tails. For a related discussion of the effects of r versus g on income distribution see in particular Piketty and Zuchman (2015).

[50] Reed (2003) generalizes the initial condition to allow the initial state w_0 to be a log-normally distributed random variable.

a hump-shaped Double-Pareto distribution[51], which captures the stylized fact that the distribution of wealth is increasing in the left tail.[52,53]

3.4.2 Non-homogeneous savings and/or returns

A model characterized by a savings rate $1-\psi(w)$ which is increasing in wealth has been studied by Atkinson (1971) in an OLG economy with constant rate of return on wealth, finitely lived agents, and warm glow preferences for bequests given by

$$v(w_T) = A\frac{\left(w_{T+1}^n\right)^{1-\mu}}{1-\mu},$$

where w_{T+1} is the end of life wealth, that is, bequests.[54] For these economies it is straightforward to show that, if the curvature of consumption in the instantaneous utility function, σ, is greater than the curvature in the bequest function, μ, propensity to consume out of wealth, $\frac{c(w_t)}{w_t}$, is decreasing in wealth, and therefore savings rates are increasing in wealth.[55]

[51] See also Benhabib and Zhu (2008), Toda (2014), and Benhabib, Bisin and Zhu (2016a) for micro-foundations of wealth accumulation processes driven by geometric Brownian motion and contained by constant death probabilities generating the Double-Pareto distributions.

[52] Note that with insertion at $w_0 > 0$, the stationary distribution results of Reed (2001) should hold even if $r < 0$.

[53] A particularly simple solution can be obtained with simplifying assumptions, following Mitzenmacher (2004), pp. 241- 242. Suppose $w_0 = 1$ and $r = \frac{\sigma^2}{2}$, $\sigma = 1$. Substituting these in 9, setting $T = u^2$, and remembering to use $\frac{dT}{du} = 2u$ in the change of variables, integral tables yield:

$$f_w(w) = \begin{cases} \sqrt{\frac{p}{2}}w^{\left(\sqrt{2p}-1\right)} & \text{for } w \leq 1 \\ \sqrt{\frac{p}{2}}w^{\left(-\sqrt{2p}-1\right)} & \text{for } w \geq 1 \end{cases}$$

[54] A related class of models with potentially explosive wealth dynamics is characterized by heterogeneous savings rates appearing in the early work of Kaldor (1957, 1961), Pasinetti (1962) and Stiglitz (1969). A recent example of this approach is Carroll, Slajek and Tokuo (2014b). Notably to generate a stationary distribution, Carroll, Slajek and Tokuo (2014b) also introduce a constant probability of death, with reinsertion at exogenous low levels of wealth, as in Blanchard's model.

[55] Atkinson's approach using bequest functions more elastic than the utility of consumption is explored in Benhabib, Bisin and Luo (2015) to study a model that nests stochastic earnings, stochastic returns and savings rates increasing in wealth. To the same effect, De Nardi (2004) and to Cagetti and De Nardi (2008) explicitly introduce non-homogenous bequest motives:

$$v(w_{T+1}) = A\left(1+\frac{w_T+1}{\gamma}\right)^{1-\sigma},$$

where γ measures how much bequests increase with wealth. See also Laitner (2001), for a model with heterogeneity in the strength of the bequest motive; and Roussanov (2010) for status concerns in accumulation incentives, distribution of asset holdings, and mobility.

4 Empirical evidence

In our theoretical survey, we identified three basic mechanisms that can contribute to generate wealth distributions that have thick tails: skewed earnings, stochastic returns on wealth, and explosive wealth accumulation. Here we focus on the same mechanisms to analyze the empirical literature on the wealth distribution. This is very useful to understand how thick-tailed wealth distributions are or are not obtained in the data, even though many of the classic models in the recent literature are hybrid models that contain more than one of these mechanisms to generate thick tails in wealth.

4.1 Skewed earnings

A general view of the stylized facts regarding the distribution of earnings is helpful to introduce the main issues regarding how much skewed earnings can contribute to explain the thick-tail in the wealth distribution.

Earnings distributions are generally skewed and thick-tailed. In the U.S. this is well documented by Piketty and Saez (2003) and especially by Guvenen, Karahan, Ozkan, and Song (2016) in their detailed study of the Social Security Administration panel data covering 1978 to 2013; see also De Nardi, Fella, and Pardo (2016) for an overview. Across countries, Atkinson (2002), Moriguchi-Saez (2005), Piketty (2001), and Saez-Veall (2003) document skewed distributions of earnings with relatively large top shares consistently over the last century, respectively, in the U.K., Japan, France, and Canada. Thick upper tails are also documented, for example, by Nirei and Souma (2005) in Japan from 1960 to 1999, by Clementi-Gallegati (2004) for Italy from 1977 to 2002, and by Dagsvik-Vatne (1999) for Norway in 1998.

Most importantly, however, earnings distributions display thinner upper tails than the wealth distribution. The tail indices of earnings reported by Badel, Dayl, Huggett and Nybom (2016) are about 2 for the US and Canada, and close to 3 for Sweden. Corresponding tail indices for wealth are about 1.5 for the U.S.(Vermeulen (2015)) , 1.4 for Canada and 1.7 for Sweden (Cowell, 2011).[56]

The fact that the distribution of wealth has a thicker tail than the distribution of earnings has important implications. In this case, in fact Theorem 1 suggests that the distribution of earnings cannot by itself explain the thick tail of the wealth distribution, and Theorem 3 strikingly implies that the distribution of earnings won't even partially contribute to explain the thickness of the tail of the wealth distribution; the burden for

[56] Gini coefficients, often used as a proxy measure of the thickness of the tail , are available for earnings distributions for a number of countries; see the special volume of *Review of Economic Dynamics*, Krueger, Perri, Pistaferri, Violante (2010). They can be compared with the Gini coefficients for wealth given by Davies, Sandstrom, Shorrocks and Wolff (2011). In all cases wealth Gini's are higher than earnings Gini's. More specifically, for the 9 countries for which we have Gini coefficients for both earnings and wealth (Canada, UK, Germany, Italy, Spain, Sweden, Russia, Mexico and the U.S.) the average ratio of the wealth Gini to the earnings Gini is 1.73.

explaining the thick tails of wealth distribution will have to rely on other factors, like the stochastic returns on wealth and/or explosive wealth accumulation. Indeed, recent empirical studies of the wealth distribution driven by earnings consistently find this to be the case. Working with the standard Aiyagari-Bewley model with stochastic labor earnings and borrowing constraints, Carroll, Slajek and Tokuo (2014b) note that "... the wealth heterogeneity [...] model essentially just replicates heterogeneity in permanent income (which accounts for most of the heterogeneity in total income)." Relatedly, De Nardi et al (2016), adapt earnings data from Guvenen, Karahan, Ozkan, and Song (2015), which they introduce into a finite-life OLG model. They note that earnings processes derived from data, including the one that they use, generate a much better fit of the wealth holdings of the bottom 60% of people, but generates too little wealth concentration at the top of the wealth distribution (See De Nardi et al, (2016), p. 44).[57]

A careful account of those studies that do successfully match the distribution of wealth with skewed earning also provides evidence which is consistent with the implications of Theorem 1 and Theorem 3 above. These studies in fact all estimate some specific stochastic properties of the distribution of earnings to fit the distribution of wealth. More specifically, these studies introduce an additional state to the stochastic process for earnings in order to match the chosen moments of the wealth distribution. The estimated state, appropriately called *awesome state* in the literature, invariably induces thickness in the distribution of earnings largely in excess of that which can be documented in earnings data. In other words, these results can be interpreted to suggest that, if earnings were the main determinant of the thickness in the tail of the distribution of wealth, a much thicker distribution of the tail of earnings relative to the tail of actual earnings data would be required to fit the wealth data. For example, Castaneda et al. (2003) estimate the properties of an *awesome state* in a rich overlapping-generation model with various demographic and life-cycle features. It requires the top 0.039% earners have about $1,000$ times the average labor endowment of the bottom 61%, while this ratio, even for the top .01%, is of the order of 200 in the World Wealth and Income Database (WWID) by Facundo Alvaredo, Anthony B. Atkinson, Thomas Piketty, Emmanuel Saez, and Gabriel Zucman (2016).[58] Similarly, Diaz et al. (2003) estimate that the top 6% earn more than 40 times the labor earnings of the bottom 50%, while the top 5% of households in WWID earn about 5 times the median. Finally, in Kindermann and Krueger (2014) earnings are endogenously driven by a seven state Markov chain for labor productivity. In their stationary distribution, 0.036% of the population is in the *awesome* productivity state with average earnings of about 20 million dollars when calibrated to median earnings,

[57] Furthermore, while the precautionary savings motive is the driving force of the Aiyagari-Bewley model, Guvenen and Smith (2014) note that "... the amount of uninsurable lifetime income risk that individuals perceive is substantially smaller than what is typically assumed in calibrated macroeconomic models with incomplete markets."

[58] We use WWID earnings data ,which is not top-coded, for 2014. The argument is not much changed even when considering average income, excluding capital gains.

about 3 times the earnings reported in the WWID for the same share of the population.[59]

Perpetual youth demographics and random working life-spans that introduce age or life-span heterogeneity across agents can complement skewed earnings to produce some additional dispersion in wealth accumulation. For example even though their "awesome" earnings state is less extreme than in the above cited literature, Kaymak and Poschke (2015) calibrate expected working lives to 45 years, as in Castaneda et al (2003). This however implies a substantial fraction of agents with an unbounded and excessive working life-span at the stationary distribution: over 100 working-years for 11% of the working population. Of these 11% a subset spend a lot of years in high earnings states to populate the tail of the wealth distribution.[60] The thick right tail of the wealth distributions will then have dynasties with long average life-spans spent in high earnings states.

4.2 Stochastic returns to wealth

Data on stochastic returns are relatively hard to find. This is in part because of the conceptual difficulties involved in mapping r_t with a measure of idiosyncratic rate of return on wealth, or capital income risk. The available systematic evidence suggests however that the idiosyncratic component of capital income risk is composed mainly of returns to ownership of principal residence and unincorporated private business equity, and of returns on private equity.[61] Also, capital income risk appears to be a significant component of individuals returns on wealth. Case and Shiller (1989) document a large standard deviation, of the order of 15%, of yearly capital gains or losses on owner-occupied housing. Similarly, Flavin and Yamashita (2002) measure the standard deviation of the return on housing, at the level of individual houses, from the 1968-92 waves of the Panel Study of Income Dynamics (PSID), obtaining a similar number, 14%. Returns on private equity have an even higher idiosyncratic dispersion across household.[62] Over the years 1953-1999, Moskowitz and Vissing-Jorgensen (2002) find the average returns on private equity, conditional on survival, of about 13% (Table 6). The distribution of returns from private equity investment to households, obtained from the 1989 SCF, even conditional on survival, is very dispersed, especially compared to the dispersion on public equity. Moskowitz and Vissing-Jorgensen (2002) note that their "Figure 2 shows that

[59] We thank the authors for a personal communication which clarified some issues in these calculations.

[60] In fact De Nardi et al (2016), working with earnings data, also introduce stochastic but finite life-spans, but with death certain by age 86. The age heterogeneity with such finite lives however, as the authors note, generates much too litle concentration of wealth at the top.

[61] Principal residence and private business equity plus investment real estate account for, respectively, 28.2% and 27% of household wealth in the U.S. according to the 2001 *Survey of Consumer Finances* ; see Wolff, 2006, Table 5, and also Bertaut and McCluer (2002), Table 2). Quadrini (2000) also extensively documents the role of returns to entrepreneurial talent in wealth accumulation.

[62] This is a consequence of the fact that private equity is highly concentrated: 75% of all private equity is owned by households for which it constitutes at least 50% of their total net worth, as documented by Moskowitz and Vissing-Jorgensen (2002).

the distribution of entrepreneurial returns is highly skewed with a fat right tail."

The most important contributions to the measurement of stochastic returns on wealth consists however in the recent studies of administrative data in Norway and Sweden. Fagereng, Guiso, Malacrino and Pistaferri (2016), in particular, using Norwegian administrative data, provide a systematic analysis of the stochastic properties of returns on wealth. They find 3.7% average returns on overall wealth, with a standard deviation of 6.1%. They also document that such heterogeneity in returns is not simply the reflection of differences in portfolio allocations between risky and safe assets mirroring heterogeneity in risk aversion and actually identify the idiosyncratic component of the lifetime rate of return on wealth across the population.[63] This measure, conditioning away within lifetime risk, is the most accurate measure to be mapped to r_t, especially in OLG models as e.g., Benhabib, Bisin, Luo (2016). This measure of returns to wealth also exhibits substantial heterogeneity. For example, for 2013 they find that the average (median) return varies significantly across households, with a standard deviation of 2.8 percentage points. Bach, Calvet, Sodini (2015), on Swedish administrative data, also find a substantial heterogeneity in returns to wealth. In particular they document large differences in returns across wealth classes: households in the top 1% of the wealth distribution, e.g., earn 4.1% more than median wealth households. They attribute this heterogeneity in large part to different portfolio strategies (riskier for wealthier individuals).[64]

Several recent studies allow for stochastic returns to wealth to successfully match the observed thick tail of the wealth distribution, consistently with the implications of Theorem 2. Importantly, the calibrated (and, in one case, even the estimated) stochastic properties of returns are quite close to those documented in the data we just discussed. More specifically, Quadrini (2000) calibrates his rich model of entrepreneurial activity and returns to PSID and SCF data on private businesses, consistently with Moskowitz and Vissing-Jorgensen (2002). Cagetti and De Nardi (2006) build on the entrepreneurial model of Quadrini (2000), and also calibrate their model to SCF data.[65] Benhabib, Bisin and Zhu (2011), using the methods of Kesten (1973), Saporta (2005) and Roitherstein (2007), formally obtain a thick-tailed wealth distributions in OLG models with finite lives. They calibrate them explicitly r_t to Moskowitz and Vissing-Jorgensen's (2002) data.

Finally Benhabib, Bisin, Luo (2016) explicitly estimate the stochastic properties of the Markov process for r_t to match the distribution of wealth. Interestingly, the mean and standard deviation of estimated returns, 2.76% and 2.54%, respectively, closely match

[63] For a study of possible genetic factors that can induce differences in differences wealth accumulation and portfolio choices, see Barth, Papageorge and Thom (2017).

[64] For a recent study that combines asset rikiness with differences in investor sophistication and endogenous participation in financial markets to explain the U.S. asset ownership dynamics and capital income dispersion see Kacperczyk et al., (2015).

[65] Their return to wealth for entrepreneurs is larger than in the data, with a median as high as 49% (which however includes all entrepreneurial income and does not correct for the survival bias.)

those estimated by Fagereng, Guiso, Malacrino and Pistaferri (2015) for the idiosyncratic component of the lifetime rate of return on wealth with Norwegian administrative data.

4.3 Explosive wealth accumulation

Various modelling features, can induce explosive wealth accumulation if not curtailed by birth and death, or decreasing returns, or fiscal policies, or other mechanisms that tame explosive accumulation. They are used in the empirical literature to help match the tail of the distribution of wealth. Cagetti and De Nardi (2008) notably center on the role of voluntary (as opposed to accidental) bequests to populate the tail of the wealth distribution with accumulated returns from entrepreneurial activities. Crucially, as we argue in Section 3.4.2, the preferences for bequests that they adopt (see footnote 55) induce a saving rate that increases in wealth.[66] Similarly, Atkinson's approach to non-homogeneous bequest is adopted by Benhabib, Bisin and Luo (2015) to structurally identify the empirical relevance of a savings rate increasing in wealth in order to match the distribution of wealth. They indeed estimate a curvature parameter (the inverse of the elasticity of substitution with CRRA functional forms) in their bequest function significantly smaller than the curvature parameter of consumption in the instantaneous utility function (1.01 and 2, respectively), implying savings rates which are substantially increasing in wealth.

Another factor that can potentially induce explosive wealth accumulation, and which plays an important role in the empirical literature on the wealth distribution is rates of return that increase in wealth. A positive correlation between returns to wealth and wealth is reasonably documented in the data, though a causal interpretation requires caution. First of all, reverse causation is certainly at least in part present: individual with higher returns, especially lifetime returns, e.g., due to personal ability, turn out to be wealthier, other things equal. Furthermore, wealthier individuals will generally hold in riskier portfolios, hence receiving higher returns as a remuneration for risk. Nonetheless, even after consideration to these confounding factors, Fagereng, Guiso, Malacrino and Pistaferri (2016) find returns significantly increasing in wealth: the difference between the median returns for individuals in the 90th and the 10th wealth percentile is about 1.8% in their data. Bach, Calvet, and Sodini (2015) find higher returns on large wealth portfolios although, as we noted, little of this difference holds in their data after conditioning for risk. Related evidence is due to Piketty (2013), showing that returns to capital endowments of U.S. universities increase with the size of endowments (Table 12.2). On the other hand, averaged over the period 1980-2012, estimates of Saez and Zucman (2016, online appendix, Tables B29, B30, and B31) show mildly increasing pre-tax returns in wealth, but flat or mildly decreasing post-tax returns in wealth.

More recent studies have also highlighted the role of undiversified portfolios, and

[66] For savings rates increasing in wealth see also Dynan, Skinner and Zeldes (2004) and Carroll (2000).

especially portfolio compositions that can depend on wealth levels. Changes in prices of asset classes that generating capital gains and losses can then differentially affect returns across wealth classes and the distribution of wealth. Thus not only returns to wealth are heterogenous when portfolio compositions differ, but they may vary systematically across wealth levels. For example higher middle wealth classes may be invested heavily in housing while the very top wealth groups may be more heavily invested in stocks and equity. Stock or housing booms and busts may then affect wealth shares and wealth distribution especially at the top . Garbinti et al (2016), looking at French historical data from 1870 to the present, show how the dynamic evolution of wealth distribution in France reflects the changes in the prices of different asset classes. Similarly Gomez (2016) and Kuhn et al (2017) study the effects of capital gains and changes in asset prices of the US distribution of wealth.

Non-homogeneous (increasing) returns to wealth have been exploited empirically by Kaplan, Moll and Violante (2015) and Benhabib, Bisin, Luo (2016). In Kaplan, Moll and Violante (2015) returns of wealth are increasing in wealth, and are endogenously obtained in a model with fixed costs of portfolio adjustments for high-return illiquid assets.[67] To produce thick tails in wealth however, they complement increasing returns with perpetual youth demographics calibrated as in Kaymak and Poschke (2015); a calibration which, we noted in the previous section, induces excessive working life-spans. Benhabib, Bisin, Luo (2016) directly estimate instead a specification of the rate of return to wealth process which is allowed to depend on wealth. They find a relatively weak but significant positive dependence, which induces a correlation between returns and wealth, in equilibrium, of the same order of magnitude as the correlation documented by Fagereng, Guiso, Malacrino and Pistaferri (2016).

4.4 On the relative importance of the various mechanisms for thick-tailed wealth

Our focus on three basic mechanisms that can contribute to generating wealth distributions has a pedagogic motivation in that it clarifies the relationship between the theoretical and empirical studies on the distribution of wealth, and identifies the main forces underlying simulations and calibrations. But distinguishing these mechanisms and evaluating their relative importance in driving wealth accumulation and the thick-tails in the distribution of wealth has also important normative implications. Modelling choices, in particular whether the source of stochastic incomes is solely shocks to labor earnings, or whether heterogeneity in rates of return also plays a role, can have significant policy consequences.

[67] See also Mengus and Pancrazi (2016) for a model with participation costs to state-contingent asset markets; and Guvenen (2006), for a model where a majority of households do not participate in stock markets due to an elasticity of intertemporal substitution that increases with wealth rather than fixed costs of participation.

Empirically, Benhabib, Bisin and Luo (2015) structurally assess the relevance of the various mechanisms that generate thick-tailed wealth distributions by estimating a model that nests them.[68] The results give a good match to wealth distribution and mobility. Benhabib, Bisin and Luo (2015) then estimate separate counterfactuals shutting down one mechanism at a time. Their findings indicate that all the three mechanisms that they focus on are important: stochastic earnings prevent poverty traps, or too many of the poor from getting stuck close to the borrowing constraints; stochastic returns assures downward mobility as well as a thick tail to match the wealth distribution; and a savings rate increasing in wealth is essential to match the tail of the wealth distribution.

4.5 Stochastic returns and the effects of tax policy

The identification of the various factors which possibly explain the thick tail of the distribution of wealth is not just relevant in and of itself. It also has important implications for the effects of policy; in particular regarding whether estate or capital taxes have an effect on wealth inequality across generations?

To address this issue, Becker and Tomes (1979) constructed an OLG model with two period lives. They introduced altruistic investments by parents in the earning ability of their children, and the transmission of earnings ability through spillovers from parents to children within families, as well as from average abilities in the economy. They also introduced a random element of luck in earnings ability, but without any capital income or rate of return risk. In this dynamic setup where choices of consumption and altruistic investments in children are optimized, they concluded that progressive and redistributive taxation may have unintended consequences for inequality: the effect of estate taxes in the transmission of inequality may be offset if parents respond by adjusting their net bequests and investments in their children. They then concluded:

> "Although increased redistribution within a progressive tax-subsidy system initially narrows inequality, the new long-run equilibrium position may well have greater inequality because parents reduce their investments in children. Perhaps this conflict between initial and long-run effects helps explain why the large growth in redistribution during the last 50 years has had very modest effects on inequality. "

Along similar lines, Castaneda, Diaz-Gimenez, and Rios-Rull (2003) and Cagetti and De Nardi (2007) also found very small (or even perverse) effects of eliminating bequest taxes in their calibrations that have a skewed distribution of earnings but no capital income risk. Laitner (2001) on the other hand introduced heterogeneity in the strength of the bequest motive, or the in the degree of intergenerational altruism. Family

[68] As we noted, they adopt the OLG model in Benhabib, Bisin and Zhu (2011), extended to allow for a savings rate increasing in wealth, via non-homogeneous bequests as in Atkinson (1971).

earnings are stochastic and drawn from a distribution. In a standard two-period OLG model Laitner could then match the top tail of the US wealth distribution in the data. However Laitner (2001)[69] showed that matching the top tail of wealth distribution is possible only if a large fraction (95%) of families are not altruistic and care only about their own consumption, while the rest have an altruistic bequest motive; it is not possible to match the top tail of wealth if everyone is equally altruistic. As Laitner points out, a small group of altruistic families that are lucky enough to get rich through high earnings perpetuate their dynasties' fortunes with large estates, fattening the top tail and generating substantial wealth inequality in the process. Introducing estate taxes can then have a significant impact and reduce wealth inequality, putting altruistic families on a closer footing to the non-altruistic ones.

Alternatively, estate and capital taxes can have a significant impact on wealth inequality and its transmission across generations in the presence of random and idiosyncratic rates of return on wealth, without relying on heterogeneity in altruistic preferences. Benhabib, Bisin and Zhu (2011) introduce stochastic idiosyncratic returns across generations. Parents derive utility from after tax bequests, and therefore also adjust their bequests in response to estate taxes. Nevertheless Benhabib, Bisin and Zhu (2011) show that when idiosyncratic rates of return across generations are a significant source of wealth inequality, reducing estate taxes, or amplifying the heterogeneity of after-tax bequests by reducing estate taxes, or for that matter decreasing capital income taxes, can significantly increase wealth inequality in the top tail of the distribution of wealth. This result holds whether rates of return across generations are *iid* or persistent, and arise from the multiplicative effect of random returns on wealth, as opposed to the additive effects of saved earnings. The change in wealth inequality in response to changes in estate or capital income taxes then are not offset by the parental adjustments of bequests. Reducing estate taxes can significantly increase wealth inequality if returns are stochastic across generations. Therefore to asses the full effects of estate or capital taxation policies on wealth distribution, it is important to explicitly model the idiosyncratic variability of rates of return.[70]

5 Conclusions

Various mechanisms which can lead to wide swings in the distribution of wealth over the long-run fall outside the scope of this survey. Some of these have been informally highlighted by Piketty (2014). First of all, the distribution of wealth in principle depends

[69] See also Laitner (2002) for a brief overview of OLG models with altruistic bequests.

[70] See also Guvenen, Kambourov Kuruscu Ocampo and Chen (2015) who study the differential effects of wealth versus capital income taxation under return heterogeneity, and Hubmer, Krusell and Smith (2016) and Nirei and Aoki (2016), Aoki and Nirei (forthcoming) on the effect of recent changes in taxes on wealth distribution..

on fiscal policy, while political economy considerations suggest that the determination of fiscal policy in turn depends on the distribution of wealth, specifically on wealth inequality. This link is, strangely enough, poorly studied in the literature. A related interesting mechanism, which did not receive much formal attention in the literature but has been introduced by Pareto (who in turn borrows it from Mosca (1896), however) goes under the heading of "circulations of the elites". It refers to the cyclical overturn of political elites who lose political power because of social psychology considerations, e.g., the lack of socialization to attitudes like ambition and enterprise, in part due to selective pressures weakening dominant elites. Alternatively wars and depressions can destroy wealth, or social interest groups whose political power and fortunes rise can appropriate economic advantages, and increase various forms of redistribution towards themselves.

References

Acemoglu, D., and J. A. Robinson (2015): "The Rise and Decline of General Laws of Capitalism." Journal of Economic Perspectives, 29, 3-28.

Achdou, Y., Han, J., J-M., Lions, P-L, and B. Moll (2014): "Heterogeneous Agent Models in Continuous Time," http://www.princeton.edu/~moll/HACT.pdf

Aguiar, Mark, and Mark Bils. (2015) "Has Consumption Inequality Mirrored Income Inequality," *American Economic Review* 105, 2725-2756.

Aiyagari, S.R. (1994): "Uninsured Idiosyncratic Risk and Aggregate Savings," *Quarterly Journal of Economics,* 109 (3), 659-684.

Alfarano, S., M. Milakovic, A. Irle, and J. Kauschke (2012): "A statistical equilibrium model of competitive firms," *Journal of Economic Dynamics and Control*, 36(1), 136-149.

Alvaredo, F., Atkinson, A. B., Piketty, T., Saez, E. (2013): "The Top 1 Percent in International and Historical Perspective," *Journal of Economic Perspectives*, 27, 3-20.

Alvaredo, F., Atkinson, A. B., Piketty, T., Saez, E., Zucman, G., (since 2011): *World Wealth and Income Database,* http://52.209.180.1/

Angeletos, G. (2007): "Uninsured Idiosyncratic Investment Risk and Aggregate Saving", *Review of Economic Dynamics*, 10, 1-30.

Angeletos, G. and L.E. Calvet (2005): "Incomplete-market dynamics in a neoclassical production economy," *Journal of Mathematical Economics*, 41(4-5), 407-438.

Angeletos, G. and L.E. Calvet (2006): "Idiosyncratic Production Risk, Growth and the Business Cycle", *Journal of Monetary Economics*, 53, 1095-1115.

Aoki, Shuhei, and Makoto Nirei (forthcoming): "Zipf's Law, Pareto's Law, and the Evolution of Top Incomes in the U.S.," *American Economic Journal: Macro.*

Arrow, K. (1987): "The Demand for Information and the Distribution of Income," *Probability in the Engineering and Informational Sciences,* 1, 3-13.

Atkinson, A.B. (1971): "Capital taxes, the redistribution of wealth and individual savings," *Review of Economic Studies,* 38(2), 209-27.

Atkinson, A.B. (2002): "Top Incomes in the United Kingdom over the Twentieth Century," mimeo, Nuffield College, Oxford.

Attanasio, O., E. Hurst, and L. Pistaferri (2015): "The Evolution of Income, Consumption, and Leisure Inequality in the US, 1980–2010,"Chap. 4 in *Improving the Measurement of Consumer Expenditures*, edited by Christopher D. Carroll, Thomas F. Crossley, and John Sabelhaus, University of Chicago Press.

Bach, L., Calvet, L., and P. Sodini (2015): "Rich Pickings? Risk, Return, and Skill in the Portfolios of the Wealthy," mimeo, Stockholm School of Economics.

Badel, A., Dayl, M., Huggett, M. and M. Nybom (2016): "Top Earners: Cross-Country Facts", mimeo.

Barndorff-Nielsen, O. E. and Shepard, N. (2001): "Non-Gaussian Ornstein–Uhlenbeck based models and some of their uses in financial economics," *J. Roy. Statist. Soc.* Ser. B 63, 167-241.

Barth, D., Papageorge, N. W. and Thom, K. (2017): "Genetic Ability, Wealth, and Financial Decision-Making," working paper, https://www.dropbox.com/s/7lxgughkq2ar6lf/GENX_TALK.pdf?dl=0

Becker, G.S. and N. Tomes (1979): "An Equilibrium Theory of the Distribution of Income and Intergenerational Mobility," *Journal of Political Economy*, 87, 6, 1153-1189.

Benhabib, J. and A. Bisin (2006): "The Distribution of Wealth and Redistributive Policies," New York University.

https://sites.google.com/site/jessbenhabib/working-papers

Benhabib, J. and S. Zhu (2008): "Age, Luck and Inheritance," NBER Working Paper No. 14128.

Benhabib, J., Bisin, A. and Zhu, S. (2011): "The Distribution of Wealth and Fiscal Policy in Economies with Finitely-Lived Agents" *Econometrica* (79) 2011, 123-157.

Benhabib, J., Bisin, A. and Zhu, S. (2016a): "The Distribution of Wealth in the Blanchard-Yaari Model," , *Macroeconomic Dynamics,* 20, 466-481.

Benhabib, Jess, "Wealth Distribution Overview," (2014a). NYU teaching slides.

http://www.econ.nyu.edu/user/benhabib/wealth%20distribution%20theories%20overview3.pdf

Benhabib, J., Bisin, A. and Zhu, S. (2015): "The Wealth Distribution in Bewley Models with Capital Income", *Journal of Economic Theory,* Part A, 489-515.

Benhabib, J., Bisin, A. and Luo, M., (2015): "Wealth distribution and social mobility in the US: A quantitative approach," NBER working paper w21721.

Bertaut, C. and M. Starr-McCluer (2002): "Household Portfolios in the United States", in L. Guiso, M. Haliassos, and T. Jappelli, Editor, *Household Portfolios*, MIT Press, Cambridge, MA.

Bewley, T. (1983): "A difficulty with the optimum quantity of money." *Econometrica*, 51:1485–1504,1983.

Blanchard, O. J. (1985): "Debt, Deficits, and Finite Horizons," *Journal of Political Economy,* 93, 223-247.

Boserup, Simon H., Wojciech Kopczuk, and Claus T. Kreiner (2015): "Born with a Silver Spoon: Danish Evidence on Intergenerational Wealth Formation from Cradle to Adulthood," University of Copenhagen and Columbia University, mimeo.

Blume, L. and Durlauf, S. (2015): "Capital in the Twentieth Century: A Critical Review Essay", Journal of Political Economy, 123, 749-777.

Borkovec, M. and Kluppelberg, C. (1998): "Extremal Behavior of Diffusions Models in Finance". *Extremes* 1 (1), 47–80.

Brandt, A. (1986): "The Stochastic Equation $Y_{n+1} = A_n Y_n + B_n$ with Stationary Coefficients," *Advances in Applied Probability*, 18, 211–220.

Burris, V. (2000): "The Myth of Old Money Liberalism: The Politics of the "Forbes" 400 Richest Americans", *Social Problems*, 47, 360-378.

Cagetti, C. and M. De Nardi (2005): "Wealth Inequality: Data and Models," Federal Reserve Bank of Chicago, W. P. 2005-10.

Cagetti, M. and M. De Nardi (2006): "Entrepreneurship, Frictions, and Wealth", *Journal of Political Economy*, 114, 835-870.

Cagetti, M. and C. De Nardi (2007): "Estate Taxation, Entrepreneurship, and Wealth," NBER Working Paper 13160.

Cagetti, C. and M. De Nardi (2008): "Wealth inequality: Data and models," *Macroeconomic Dynamics*, 12(Supplement 2), 285-313.

Case, K., and Shiller, R. (1989): "The Efficiency of the Market for Single-Family Homes", *American Economic Review*, 79, 125-137.

Castaneda, A., J. Diaz-Gimenez, and J. V. Rios-Rull (2003): "Accounting for the U.S. Earnings and Wealth Inequality," *Journal of Political Economy*, 111, 4, 818-57.

Carroll, C. D. (1997): "Buffer-Stock Saving and the Life Cycle/Permanent Income Hypothesis," *Quarterly Journal of Economics,* 112, 1–56.

Carroll, C. D. and Slacalek, J., and K. Tokuoka, (2014): "The Distribution of Wealth and the MPC: Implications of New European Data," *American Economic Review: Papers and Proceedings,* Volume 104, No. 5, pp. 107–11.

Carroll, C. D. and Slacalek, J., and K. Tokuoka, (2014b): "Wealth Inequality and the Marginal Propensity to Consume, " European Central Bank, Working Paper Series, NO 1655.

Carroll, C. D. (2000): "Why Do the Rich Save So Much?," in Joel B. Slemrod, ed., *Does Atlas Shrug?: The Economic Consequences of Taxing the Rich*, Harvard University Press, Cambridge, 465-484.

Champernowne, D.G. (1953): "A Model of Income Distribution," *Economic Journal*, 63, 318-51.

Chipman, J.S. (1976): "The Paretian Heritage," *Revue Europeenne des Sciences Sociales et Cahiers Vilfredo Pareto*, 14, 37, 65-171.

Cantelli, F.P. (1921): "Sulla deduzione delle leggi di frequenza da considerazioni di probabilita," *Metron*, 1, 83-91.

Cantelli, F.P. (1929): "Sulla legge di distribuzione dei redditi," *Giornale degli Economisti e Rivista di Statistica*, 4, 69, 850-852.

Clementi, F. and M. Gallegati (2005): "Power Law Tails in the Italian Personal Income Distribution," *Physica A: Statistical Mechanics and Theoretical Physics,* 350, 427-438.

Conley, T.G., Hansen, L.P, Luttmer, E.G.J., Scheinkman, J.A. (1997): "Short-Term Interest Rates as Subordinated Diffusions," Review of Financial Studies, 10, 525-577.

Cox, J.C., Ingersoll, J.E., and Ross, S.A. (1985): "A Theory of the Term structure of Interest Rates," Econometrica 53, 385-408.

Cowell F.A., (2011): "Inequality Among the Wealthy," mimeo, Centre for Analysis of Social Exclusion, LSE.

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

Dagsvik, J.K. and B.H. Vatne (1999): "Is the Distribution of Income Compatible with a Stable Distribution?," D.P. No. 246, Research Department, Statistics Norway.

Davies, J.B., S. Sandström, A. Shorrocks, E. Wolff (2011): "The Level and Distribution of Global Household Wealth," *Economic Journal,* 121(551), 223-254. http://onlinelibrary.wiley.com/doi/10.1111/j.1468-0297.2010.02391.x/epdf

De Nardi, M. (2004): "Wealth Inequality and Intergenerational Links," *Review of Economic Studies*, 71, 743-768.

De Nardi, M., Fella, G. and G. P. Pardo (2016): "The Implications of Richer Earnings Dynamics for Consumption, Wealth, and Welfare", NBER Working paper No, 21917.

Diaz, A, Josep Pijoan-Masb, J., and Rios-Rull, J-F, (2003): "Precautionary savings and wealth distribution under habit formation preferences,"*Journal of Monetary Economics* 50, 1257-1291.

Diaz-Gimenez, J., V. Quadrini, and J. V. Rios-Rull (1997): "Dimensions of Inequality: Facts on the U.S. Distributions of Earnings, Income, and Wealth," *Federal Reserve Bank of Minneapolis Quarterly Review*, 21(2), 3-21.

Díaz-Giménez, J., V. Quadrini, J. V. Ríos-Rull, and S. B. Rodríguez (2002): "Updated Facts on the U.S. Distributions of Earnings, Income, and Wealth", *Federal Reserve Bank of Minneapolis Quarterly Review*, 26(3), 2-35.

Dynan, K., Skinner, J. and Zeldes, S. P. (2004): "Do the Rich Save More?" *Journal of Political Economy*, 112, 397-444.

Edgeworth, F. Y., (1896): "Review notice concerning Vilfredo Pareto's "Courbe des Revenus," *Economic Journal*, 6, 666.

Edgeworth, F. Y., (1899): "On the Representation of Statistics by Mathematical Formulae," Part I, *Journal of the Royal Statistical Society*, 61, 670-700.

Edgeworth, F. Y., (1899): "On the Representation of Statistics by Mathematical Formulae," Part IV, *Journal of the Royal Statistical Society*, 62 , 534-555.

Elwood, P., S.M. Miller, M. Bayard, T. Watson, C. Collins, and C. Hartman (1997): "Born on Third Base: The Sources of Wealth of the 1996 Forbes 400," Unified for a Fair Economy. Boston.

Fagereng, Andreas, Mogstad, Magne and Marte Rønning (2015): "Why do wealthy parents have wealthy children?" Statistics Norway Discussion Paper.

Fagereng, A., L. Guiso, D. Malacrino, and L. Pistaferri (2015): "Wealth Return Dynamics and Heterogeneity," mimeo, Stanford University.

Fagereng, A., Guiso, L., Malacrino, D., and Pistaferri, L. (2016): “Heterogeneity in Returns to Wealth, and the Measurement of Wealth Inequality” *American Economic Review Papers and Proceedings*, 106, 651-655(5).

Fasen, V., Kluppelberg, C. & Lindner, A. (2006):. ¨"Extremal Behavior of Stochastic Volatility Models," in *Stochastic Finance*, M.D.R. Grossinho, A.N. Shiryaev, M. Esquivel & P.E. Oliviera, eds, Springer, New York, pp. 107-155.

Feenberg, D. and J. Poterba (2000): "The Income and Tax Share of Very High Income Household: 1960-1995," *American Economic Review*, 90, 264-70.

Feller, W. (1966): *An Introduction to Probability Theory and its Applications*, 2, Wiley, New York.

Fernholz, R.T. (2016): "A Model of Economic Mobility andthe Distribution of Wealth," Claremont McKenna College, mimeo.

Fiaschi, D. and M. Marsili (2009): "Distribution of Wealth and Incomplete Markets: Theory and Empirical Evidence," University of Pisa Working paper 83.

Flavin, M. and T. Yamashita (2002): "Owner-Occupied Housing and the Composition of the Household Portfolio", *American Economic Review*, 92, 345-362.

Fréchet, Maurice, (1939): "Sur les formules de repartition des revenus," *Revue de l'Institut International de Statistique*, 7, 32-38.

Fréchet, Maurice (1958): "Letter to the Editor", *Econometrica*, 26(4), 590-591.

Gabaix, Xavier, (1999): “Zipf 's Law for Cities: An Explanation,” *Quarterly Journal of Economics*, CXIV , 739 -767.

Gabaix, Xavier, (2009):. "Power Laws in Economics and Finance," Annual Review of Economics, Annual Reviews, vol. 1(1), pages 255-294, 05.

Gabaix, X., Lasry, J-M., Lions, P-L, and B. Moll, (2016): "The Dynamics of Inequality," *Econometrica* 84, 2071-2111.

Gabaix, X. and A. Landier (2008): "Why Has CEO Pay Increased So Much?" *The Quarterly Journal of Economics*, 123(1), 49-100.

Garbinti, B., Goupille-Lebret, J. and Piketty, T. (2016): "Accounting for Wealth Inequality Dynamics: Methods, Estimates and Simulations for France (1800-2014)," https://editorialexpress.com/cgi-bin/conference/download.cgi?db_name=EWM2016&paper_id=5

Geerolf, F. (2016): "A Theory of Pareto Distributions," UCLA, mimeo.

Ghosh, A.P., D. Hay, V. Hirpara, R. Rastegar, A. Roitershtein, A. Schulteis, J. Suh (2010): "Random linear recursions with dependent coefficients," *Statistics and Probability Letters*, 80, 1597-1605.

Gibrat, R. (1931): *Les InÃ©galitÃ©s Ã©conomiques,* Paris, Librairie du Recueil Sirey.

Goldie, C. M. (1991): "Implicit Renewal Theory and Tails of Solutions of Random Equations," *Annals of Applied Probability*, 1, 126–166.

Gomez, M. (2016): "Asset Prices and Wealth Inequality," working paper. Princeton.

http://www.princeton.edu/~mattg/files/jmp.pdf

Grey, D.R. (1994): "Regular variation in the tail behavior of solutions of random difference equations," *The Annals of Applied Probability*, 4)(1), 169-83.

Guvenen, F. (2006): "Reconciling conflicting evidence on the elasticity of intertemporal substitution:A macroeconomic perspective," *Journal of Monetary Economics*, 53, 1451–1472.

Guvenen, F. and A. A. Smith (2014): "Inferring Labor Income Risk and Partial Insurance From Economic Choices," *Econometrica*, 82, 2085–2129.

Guvenen, F. (2015): "Income Inequality and Income Risk: Old Myths vs. New Facts," Lecture Slides available at https://fguvenendotcom.iles.wordpress.com/2014/04/handout_inequality_risk_webversion_may20151.pdf.

Guvenen, F., G. Kambourov, B. Kuruscu, S. Ocampo, and D. Chen (2015): "Use It or Lose It: Efficiency Gains from Wealth Taxation,"

https://fguvenendotcom.files.wordpress.com/2014/04/handout_ggkoc_slides_web.pdf

Guvenen, F., Karahan, F., Ozkan, S. and J. Song (2015): "What Do Data on Millions of U.S. Workers Reveal about Life-Cycle Earnings Risk?," National Bureau of Economic Research, Inc. NBER Working Papers 20913

Haldane, J. B. S. (1942): "Moments of the Distributions of Powers and Products of Normal Variates" *Biometrika*, 32, 226-242.

Hay, D., R. Rastegar, and A. Roitershtein (2011): Multivariate linear recursions with Markov-dependent coefficients, J. Multivariate Anal. 102, 521 527.

Heathcote, J. (2008): "Discussion *Heterogeneous Life-Cycle Profiles, Income Risk, and Consumption Inequality*, by G. Primiceri and T. van Rens," Federal Bank of Minneapolis, mimeo.

Huggett, M. (1993): "The Risk-Free Rate in Heterogeneous-household Incomplete-Insurance Economies, *Journal of Economic Dynamics and Control*, 17, 953-69.

Huggett, M. (1996): "Wealth Distribution in Life-Cycle Economies," *Journal of Monetary Economics,* 38, 469-494.

Huggett, M., G. Ventura, and A. Yaron, (2011): "Sources of Lifetime Inequality," *American Economic Review*, 101 (7), 2923{54.

Hurst, E. and K.F. Kerwin (2003): "The Correlation of Wealth across Generations," *Journal of Political Economy*, 111, 1155-1183.

Jones, C. I. (2015): "Pareto and Piketty: The Macroeconomics of Top Income and Wealth Inequality," J*ournal of Economic Perspectives*, 29, 29–46

Kaldor, N. (1957): "A Model of Economic Growth", *Economic Journal*, 67(268): 591-624.

Kaldor, N. (1961): "Capital Accumulation and Economic Growth," in F. A. Lutz and D. C. Hague, eds. *The Theory of Capital*, St. Martins Press, pp. 177-222.

Kalecki, M. (1945): "On the Gibrat Distribution," *Econometrica* 13: 161–170.

Kacperczyk, M., Nosal, J., and Stevens, L. (2015): "Investor sophistication and capital income inequality," Narodowy Bank Polski, working paper 199.

Kaplan, G, Moll, B., and Gianluca Violante (2015): "Asset Illiquidity, MPC Heterogeneity and Fiscal Stimulus," in preparation.

Kaymak, Baris and Markus Poschke (2016): "The evolution of wealth inequality over half a century: the role of taxes, transfers and technology," *Journal of Monetary Economics*, Vol. 77(1).

Keane, M. and Wolpin, K. (1997): "The Career Decisions of Young Men", *Journal of Political Economy*, 105, 473-

Kesten, H. (1973): "Random Difference Equations and Renewal Theory for Products of Random Matrices," *Acta Mathematica,* 131, 207–248.

Klass, O.S., Biham, O., Levy, M., Malcai O., and S. Solomon (2007): "The Forbes 400, the Pareto Power-law and Efficient Markets, *The European Physical Journal B - Condensed Matter and Complex Systems,* 55(2), 143-7.

Kuhn, M. and V.J. Rios-Rull (2016): "2013 Update on the U.S. Earnings, Income, and Wealth Distributional Facts: A View from Macroeconomics," *Quarterly Review*, Federal Reserve Bank of Minneapolis, April, 1-75.

Kuhn, M, Schularick, M., and Steins, Ulrike I. (2017): "Income and Wealth Inequality in America,"

http://www.wiwi.uni-bonn.de/kuhn/paper/Wealthinequality_20March2017.pdf

Klevmarken, N.A., J.P. Lupton, and F.P. Stafford (2003): "Wealth Dynamics in the 1980s and 1990s: Sweden and the United States", *Journal of Human Resources*, XXXVIII, 322-353.

Krueger, D. and Kindermann, F., (2014: "High marginal tax rates on the top 1%? Lessons from a life-Cycle model with idiosyncratic income risk," NBER WP 20601 http://www.nber.org/papers/w20601

Krueger, D., F. Perri, L. Pistaferri, G.L. Violante (eds.) (2010): *Review of Economic Dynamics*, 13.

Krusell, P. and A. A. Smith (1998): "Income and Wealth Heterogeneity in the Macroeconomy," *Journal of Political Economy*, 106, 867-896.

Krusell, P. and A. A. Smith (2014): "Is Piketty's "Second Law of Capitalism" Fundamental? http://aida.wss.yale.edu/smith/piketty1.pdf

Hubmer, J., P. Krusell, and A.A. Smith, Jr. (2016): "The Historical Evolution of the Wealth Distribution: A Quantitative-Theoretic Investigation," NBER Working Paper, 23011.

Kuznets, S. (1955): 'Economic Growth and Economic Inequality,' *American Economic Review*, 45, 1-28.

Laitner, John (2001): "Inequality and Wealth Accumulation: Eliminating the Federal Gift and Estate Tax," In William G. Gale, James R. Hines, Jr., and Joel Slemrod, editors, *Rethinking Estate and Gift Taxation*, Brookings Institution Press, 258–292.

Laitner, John (2002): "Wealth Inequality and Altruistic Bequests," *American Economic Review, Papers and Proceedings*, 92, 270-273.

Lampman, R.J. (1962): *The Share of Top Wealth-Holders in National Wealth, 1922-1956*, Princeton, NJ, NBER and Princeton University Press.

Levy, M. (2005): " Market Efficiency, The Pareto Wealth Distribution, and the Levy Distribution of Stock Returns" in *The Economy as an Evolving Complex System* eds. S. Durlauf and L. Blume, Oxford University Press, USA.

Levy, M. and S. Solomon (1996): "Power Laws are Logarithmic Boltzmann Laws," *International Journal of Modern Physics*, C,7, 65-72.

Light, B. (2016): "Precautionary Saving in a Markovian Earnings Environment," mimeo, https://www.gsb.stanford.edu/sites/gsb/files/working-papers/precautionary_saving_in_a_marko

Loève, M. (1977): *Probability Theory*, 4th ed., Springer, New York.

Luttmer, E.G.J. (2012): "Slow Convergence in Economies with Firm Heterogeneity," Federal Reserve Bank of Minneapolis working paper 696.

Luttmer, E.G.J. (2016): "Further Notes on Micro Heterogeneity and Macro Slow Convergence," University of Minnesota, http://users.econ.umn.edu/~luttmer/research/MicroHeterogeneityMacroSlowConvergence.pdf

Mandelbrot, Benoit (1969). "Stable Paretian Random Functions and the Multiplicative Variation of lncome," *Econometrica*, 29, 517-543.

McKay, A. (2008): "Household Saving Behavior, Wealth Accumulation and Social Security Privatization," mimeo, Princeton University.

Mengus, E. and R. Pancrazi (2016): "Endogenous Partial Insurance and Inequality," mimeo, University of Warwick.

Meyn, S. and R.L. Tweedie (2009): *Markov Chains and Stochastic Stability*, 2nd edition, Cambridge University Press, Cambridge.

Mirek, M. (2011): "Heavy tail phenomenon and convergence to stable laws for iterated Lipschitz maps," Probability Theory and Related Fields, 151, 705-34.

Mitzenmacher, M. (2004): "A Brief History of Generative Models for Power Law and Lognormal Distributions", *Internet Mathematics*, vol 1, No. 3, pp. 305-334, 2004, pp.241-242)

Moriguchi, C. and E. Saez (2005): "The Evolution of Income Concentration in Japan, 1885-2002: Evidence from Income Tax Statistics," mimeo, University of California, Berkeley.

Mosca, G. (1896): *Elementi di scienza politica* , Fratelli Bocca, Rome. Translated into English by H. D. Kahn (1939), *The Ruling Class*, McGraw-Hill, New York.

Moskowitz, T. and A. Vissing-Jorgensen (2002): "The Returns to Entrepreneurial Investment: A Private Equity Premium Puzzle?", *American Economic Review*, 92, 745-778.

Nirei, M. and W. Souma (2007): "Two Factor Model of Income Distribution Dynamics," *Review of Income and Wealth*, 53(3), 440-459.

Nirei, M. and S. Aoki (2016): Pareto Distribution of Income in Neoclassical Growth Models," Review of Economic Dynamics, 20:25-42.

Panousi, V. (2008): "Capital Taxation with Entrepreneurial Risk," mimeo, MIT.

Pareto, V. (1897): *Cours d'Economie Politique*, II, F. Rouge, Lausanne.

Pareto, V. (1901): "Un' Applicazione di teorie sociologiche", *Rivista Italiana di Sociologia*. 5. 402-456, translated as *The Rise and Fall of Elites: An Application of Theoretical Sociology* , Transaction Publishers, New Brunswick, New Jersey, (1991).

Pareto, V. (1909): *Manuel d'Economie Politique*, V. Girard et E. Brière, Paris.

Piketty, T. (2003): "Income Inequality in France, 1901-1998," *Journal of Political Economy.* 111, 1004–1042.

Piketty, T. and E. Saez (2003): "Income Inequality in the United States, 1913-1998," *Quarterly Journal of Economics*, CXVIII, 1, 1-39.

Piketty , T. (2014): *Capital in the Twenty-First Century,* Harvard University Press.

Piketty, T. and Zucman, G. (2015): "Wealth and Inheritance in the Long Run," in *Handbook of Income Distribution,* Atkinson, A. B. and Bourguignon, F., eds., Elsevier, Amsterdam, 1303-1368.

Pasinetti, L. (1962): "The rate of profit and income distribution in relation to the rate of economic growth," *Review of Economic Studies*, 29, 267-279.

Primiceri, G. and T. van Rens (2006): "Heterogeneous Life-Cycle Profiles, Income Risk, and Consumption Inequality," CEPR Discussion Paper 5881.

Quadrini, V. and J.V. Ríos-Rull, (1997): "Understanding the U.S. Distribution of Wealth", *Federal Reserve Bank of Minneapolis Quarterly Review* Vol. 21, No. 2, pp. 22–36.

Quadrini, V. (1999): "The importance of entrepreneurship for wealth concentration and mobility," *Review of Income and Wealth*, 45, 1-19.

Quadrini, V. (2000): "Entrepreneurship, Savings and Social Mobility," *Review of Economic Dynamics*, 3, 1-40.

Ray, D. (2014): "Nit-Piketty: A Comment on Thomas Piketty's Capital in the Twenty First Century," http://www.econ.nyu.edu/user/debraj/Papers/Piketty.pdf.

Reed, W. J. (2001): "The Pareto, Zipf and other power laws", *Economics Letters* 74 (2001) 15–19.

Reed, W. J. (2003): The Pareto law of income — an explanation and an extension. *Physica A: Statistical Mechanics and its Applications*, 319, 469-486.

Resnick, Sidney (1987): *Extreme Values, Regular Variation, and Point Processes,* NewYork: Springer-Verlag.

Roitershtein, A. (2007): "One-Dimensional Linear Recursions with Markov-Dependent Coefficients," *The Annals of Applied Probability,* 17(2), 572-608.

Roussanov, N. (2010): "Diversification and its Discontents: Idiosyncratic and Entrepreneurial Risk in the Quest for Social Status", Journal of Finance, LXV, 1755-1788.

Roy, A. D.(1950): "The Distribution of Earnings and of Individual Output", *The Economic Journal*, 60, 489-505.

Rutherford, R.S.G. (1955): "Income Distribution: A New Model," *Econometrica*, 23, 277-94.

Saez, E. and M. Veall (2003): "The Evolution of Top Incomes in Canada," NBER Working Paper 9607.

Saez E. and Zucman, G., (2016): "Wealth Inequality in the United States since 1913: Evidence from Capitalized Income Tax Data" Quarterly Journal of Economics, forthcoming. Online appendix at http://gabriel-zucman.eu/files/SaezZucman2016QJEAppendix.pdf

P.A. Samuelson (1965): 'A Fallacy in the Interpretation of the Pareto's Law of Alleged Constancy of Income Distribution,' *Rivista Internazionale di Scienze Economiche e Commerciali*, 12, 246-50.

Saporta, B. (2004): "Etude de la Solution Stationnaire de l'´Equation $Y_{n+1} = a_n Y_n + b_n$, a Coefficients Aleatoires," (Thesis), at http://tel.archives-ouvertes.fr/docs/00/04/74/12/PDF/tel-00007666.pdf

Saporta, B. (2005): "Tail of the stationary solution of the stochastic equation $Y_{n+1} = a_n Y_n + b_n$ with Markovian coefficients," *Stochastic Processes and Application* 115(12), 1954-1978.

Saporta, B., and Yao, J.-F., (2005): "Tail of a linear diffusion with Markov switching, Ann. Appl. Probab. 15 , 992–1018.

Sargan, J.D. (1957): "The Distribution of Wealth", *Econometrica*, 25, pp. 568-590.

Stiglitz, J. (1969): "Distribution of income and wealth among individuals," *Econometrica*, 37(3), 382-397.

Stiglitz, J. (2015a): "Of income and wealth among individuals: Part 1. The wealth residual," NBER Working paper 21189.

Stiglitz, J. (2015b): "Of income and wealth among individuals: Part 2. Equilibrium wealth distributions," NBER Working paper 21190.

Stiglitz, J. (2015c): "Of income and wealth among individuals: Part 3. Life-cycle savings vs. inherited savings," NBER Working paper 21191.

Sornette, D. (2000): *Critical Phenomena in Natural Sciences,* Berlin, Springer Verlag.

Souma,W. and Nirei, M. (2005): "Empirical study and model of personal income," https://arxiv.org/pdf/physics/0505173.pdf

Storesletten, K., C.I. Telmer, and A. Yaron (2004): "Consumption and Risk Sharing Over the Life Cycle," *Journal of Monetary Economics*, 51(3), 609-33.

Toda, A. (2012): "The double power law in income distribution: Explanations and evidence," *Journal of Economic Behavior & Organization*, 84(1), 364-381.

Toda, A. (2014): "Incomplete market dynamics and cross-sectional distributions," *Journal of Economic Theory*, 154, 310–348.

Toda, A., K.J. Walsh (2015): "The Double Power Law in Consumption and Implications for Testing Euler Equations," *Journal of Political Economy*, 123(5), 1177-1200

Vermeulen, P. (2014): "How fat is the top tail of the wealth distribution?" European Central Bank, Working Paper Series no.1692.

Wälde, K., (2016): "Pareto-Improving Redistribution of Wealth –The Case of the NLSY 1979 Cohort," Johannes-Gutenberg University, Mainz, mimeo.

Wold, H.O.A. and P. Whittle (1957): "A Model Explaining the Pareto Distribution of Wealth," *Econometrica*, 25, 4, 591-5.

Wolff, E. (1987): "Estimates of Household Wealth Inequality in the U.S., 1962-1983," *The Review of Income and Wealth*, 33, 231-56.

Wolff, E. (2006): "Changes in Household Wealth in the 1980s and 1990s in the U.S.," in Edward N. Wolff, Editor, International Perspectives on Household Wealth, Elgar Publishing Ltd.

Zeldes, S. P. (1989): "Optimal Consumption with Stochastic Income: Deviations from Certainty Equivalence," *Quarterly Journal of Economics*, 275-298.

Zhu, S. (2010): "Wealth Distribution under Idiosyncratic Investment Risk", mimeo, NYU.