Skewed Wealth Distributions: Theory and Empirics

Jess Benhabib
New York University

Alberto Bisin
New York University and NBER

First draft: June 2015; This draft: January 2016.*

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 mechanism 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 underpinning.

Key Words: Wealth distribution; wealth inequality.
JEL Numbers: E13, E21, E24

*Thanks to Mariacristina De Nardi, Steven Durlauf, Raquel Fernandez, Luigi Guiso, Mi Luo, Ben Moll, Thomas Piketty, Shenghao Zhu, and Gabriel Zucman.

1 Introduction

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

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

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 "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 skewedness 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 with stochastic returns.

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]

In this survey we aim at i) categorizing the theoretical studies on the distribution of wealth in terms of the underlying economic mechanism generating skewness and thick

[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 (2014), Krusell and Smith (2014), and Ray (2014).

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 underpinning, with the ultimate objective of measuring the relative importance of the various mechanisms in fitting the data.

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 historical and recent literature on wealth distribution. In subsequent sections we explore various stochastic 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 can be 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, while thin tails decay at a faster rate.

More formally, thick tails are defined as follows. Let a real function L be *regularly varying with index* $\alpha \in (0, \infty)$ if

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

Then, a distribution with a differentiable cumulative distribution function (cdf) $F(x)$ and counter-cdf $1 - F(x)$ is a power-law with tail index α if $1 - F(x)$ is regularly varying with index $\alpha > 0$. In this case we say the distribution is thick tailed with tail α. The standard example of a thick-tailed distribution is the Pareto distribution. A distribution on the contrary thin-tailed if $lim_{x \to \infty} \frac{1-F(tx)}{1-F(x)} = \infty, \forall t > 0$. Standard examples are the Normal and the Lognormal distributions.[4] Intuitively, α captures the number of finite moments of the distribution ($\alpha = \infty$ means the distribution has all moments). Obviously, the larger α, the thinner is the tail.

As we argued, 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. Thick tails for the distributions of wealth are also well documented, for example by Clementi-Gallegati (2004) for Italy from 1977 to 2002, and by Dagsvik-Vatne (1999)

[4] A distribution is instead heavy-tailed if $lim_{x \to \infty} \frac{1-F(tx)}{1-F(x)} = 0, \forall t > 0$. A standard example is the Cauchy distribution. If the distribution $F(x)$ has bounded support on the other hand it is necessarily thin-tailed and typically has all its moments.

for Norway in 1998. Indeed, the top end of the wealth distribution obeys a power law (more specifically, a Pareto law): Using the richest sample of the U.S., the Forbes 400, during 1988-2003 Klass et al. (2007) estimates a tail index equal to 1.49.

1.2 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.

The main question at the outset, since Pareto himself, is how to obtain a skewed thick-tailed distribution of wealth. In this respect Pareto explored whether some heterogeneity in the distribution of talents could produce the observed skewness of the labor earnings distribution.[5] His underlying idea was that a skewed distribution of labor earnings would then map into a skewed distribution of wealth. Along similar lines Edgeworth (1917) proposed the *method of translation,* which consists in identifying distributions of talents coupled with mappings 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 $s = g^{-1}(y)$ and the distribution of labor earnings is:

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

The simplest and most direct application of this *method* is due to the mathematician F.P. Cantelli (1921, 1929), who showed that if earnings exponentially increase in talent, and talent has a negative exponential distribution, *then using the above formula*, earnings would follow a Pareto distribution.[6] As we shall see in the next section, this class of arguments has since been developed by a large literature modeling the distribution of labor earnings to obtain better qualitative fits. For instance, inspired by Edgeworth's (1896, 1898, 1899) critical comment of Pareto's work that the lower earnings brackets does not follow a Pareto distribution, Maurice Frechet (1939) showed that a Laplace distribution for talents induces a unimodal density function of earnings; that is, Pareto

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

[6] In fact Cantelli (1921, 1929), drawing on arguments by Boltzman and Gibbs, also gave a derivation showing that if total talent was 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 distribution.

above the median, but increasing below it.[7] More generally, richer models of the determination of earnings have been developed. We shall study several examples in the next section.

By concentrating on the factors determining skewed thick-tailed earnings distribution, this literature tended to disregard the properties of the map between earnings and wealth. 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 wealth distribution that is induced by skewed earnings, or could even lead to skewed wealth distributions without the help of skewed earnings distributions. In this respect, several accumulation processes have been introduced and studied, focussing on central issue of the stationarity of the wealth distribution. Indeed skewed wealth distributions can be easily obtained for expansionary wealth accumulation processes over time, but these processes do not necessarily converge to a stationary wealth distribution.

A simple description of the issues involved is obtained for economies in which the rate of return is stochastic.[8] As the simplest example, consider proportional growth with Normal i.i.d. rate of returns r_t over time:

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

(the economy has no labor earning, $y_t = 0$, for simplicity). This process satisfies what is generally referred to as *Gibrat's Law:* it induces a log-normal distribution at each finite time t, with a variance increasing and exploding in t ,

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

It follows that the distribution of $\frac{1}{t}\ w_t$ converges to a log-normal distribution as t increases. But the variance of wealth explodes and no stationary distribution of wealth existsr.

Economic forces might however produce a stationary distribution of wealth that tames the exploding variance resulting from proportional growth. Kalecki (1945) proposed

[7] If instead talent were normally distributed, then earnings would be log-normally distributed, providing a better fit to the lower earning brackets than the Pareto distribution, but not a good fit for the upper brackets. Roy (1950) obtains a normally distributed talent by assuming it composed of several multiplicative i.i.d. normal components (intelligence, peseverence, originality etc). Under standard assumptions for the Central limit theorem to apply, income would approximately be lognormally distributed.

[8] Related issues with stationarity arise, for instance, in economies with high rate of interest and high and increasing savings rates out of wealth; or in economies characterized by a rate of interest increasing in wealth. We discuss several of these models in the next section.

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, would induce a constant variance in the distribution of wealth.[9] This line of argument has not been much followed recently because a decreasing net rate of return in wealth appears counterfactual. Models in which the rate of return increases in wealth, e.g., because of entrepreneurship opportunities, on the other hand, exacerbate the non-stationarity implicit in Gibrat's Law.[10]

This logic then 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 taming the possible non-stationarity and derive a Pareto distribution for wealth in the context of an economy with an expansionary wealth accumulation process.[11]

Consider an economy with a constant expansionary 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, and guessing $f_w = w^{\alpha-1}$, 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 growth of the economy as a whole, $\frac{r}{\gamma}$; see Wold and Whittle (1957), Table 1, p. 584. To guarantee that the stationary wealth distribution characterized by density f_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 exploiting birth and death processes

[9] It is straightforworward 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.

[10] This is the case also for the class of models discussed in Section 2.4, in which the savings rate increases in wealth, e.g., because of non-homogeneous preferences for bequests.

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

which we discuss in subsequent sections. In particular, besides birth and death, all these models need to introduce some form of reflecting barrier to guarantee stationarity, so 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 2.4.

Expansionary wealth accumulation processes are not necessary to obtain skewed wealth distributions, however. Indeed Champernowne (1953) introduces a wealth accumulation process which contracts on average, but nonetheless he obtainis a stationary distribution of wealth with a thick tail. More specifically Champernowne (1953) divides wealth into many bins,[12] with a bottom bin from which it is only possible to move up, hence inducing 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) shows that this stochastic process with downward drift and a reflecting barrier generated a Pareto distribution of wealth. Formally the wealth bins $j = 1, 2, 3$ are defined by their lower boundaries,

$$w(j) = w(0)\, e^{aj}, \quad j = 1, 2, 3...$$

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

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

where the adding up constraint to n people is $\sum_i^\infty n_t^i = n$. 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,

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

Champernowne shows that this condition implies that a stationary wealth distribution must satisfy $n^j = q\left(\frac{p_1}{p_{-1}}\right)^j$, for q appropriately chosen. After a transformation of variables,

$$n^j = \frac{q}{a} \frac{w(j)^{-\left(\frac{\lambda}{a}+1\right)}}{w(0)^{\frac{\lambda}{a}}}$$

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

which defines a Pareto distribution, with $\sum_0^\infty n^j = n$.[13] 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:

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

where $r_t \geq 0$ and i.i.d., and $\underline{w} > 0$; examples include Quadrini (1999, 2000), Benhabib, Bisin and Zhu (2011, 2016). Achdou, Lasry, Lions and Moll (2014), Gabaix, Lasry, Lions and Moll (2015). 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.

1.2.1 Calling for micro-foundations

Theoretical models of skewed earnings in the 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

[13] 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.

> 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 of Mincer and Becker and Tomes were especially influential. Beginning in the 1990es, they lead economists to work with micro-founded models of stochastic processes of wealth dynamics and optimizing heterogenous agents. For instance, Bewley-Aiyagari economies with heterogenous agents and stochastic earnings[14] have since been used to derive stationary wealth distributions; see for example Diaz et al (2003), Castaneda et al (2003), and Benhabib, Bisin and Zhu (2016). Blanchard-Yaari's perpetual-youth model, exploiting a birth-death process with constant probability of death, can be considered a recent simplified micro-founded version of Wold and Whittle. Castaneda et al (2003) and Carroll, Slajek and Tokeu (2014b) also make use of microfounded versions of the perpetual youth model combined with skewed random earnings. Champernowne's insights on income or wealth transitions and stochastic returns have been translated into fully microfounded models with heterogenous agents facing stochastic rates of return and labor incomes acting as a reflecting barrier; see for example Quadrini (2000), Benhabib, Bisin and Zhu (2011), (2016), Achdou, Lasry, Lions and Moll (2014), or Gabaix, Lasry, Lions and Moll (2015) for heterogenous stochastic entrepreneurial returns, Krusell and Smith(1998) for stochastic heterogenous discount rates. Microfounded models with non-homogeneity in savings rates and bequest functions have been employed to generate fat tailed wealth distributions; see Atkinson (1971) or Cagetti and DeNardi (2006).

In the following sections we will place the recent work offering mechanisms to explain the empirical distributions of wealth in historical context and we will attempt an integration of some of the theoretical mechanisms to account for the shape of wealth distribution and for the empirics of wealth mobility.

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

2 Theoretical Mechanisms for the Skewed Distribution of Wealth

Suppose wealth at time t, w_t, can only be invested in an asset with return process $\{r_t\}$. Let $\{y_t\}$ denote the labor earnings process. Let c_t denote consumption at t. The wealth accumulation equation is then:

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

What characteristics of the earnings and wealth accumulation processes are responsible for producing skewed wealth distributions? We distinguish several classes of models, each emphasizing a different economic mechanism producing skewed wealth distributions. We follow the structure laid out in the previous historical section, in terms of the foundational work we have identified. In the next section we start with models that describe the skewed distribution of labor earnings $\{y_t\}$. Under reasonable micro-foundations, in fact, skewed earning distributions translate into skewed wealth distributions. We also survey models which generate skewed earnings distributions. This section exploits heavily Edgeworth's "method of translation" and Cantelli's early work.

Alternatively, models of skewed thick-tailed wealth distributions are driven by individual wealth processes which shrink on average down to a reflecting barrier, but expand with positive probability due to random rates of return $\{r_t\}$ and labor incomes $\{y_t\}$. These models can be considered variations and extensions of Champernowne (1953).

We then study models in which skewed 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 also discuss models where preferences induce savings rates that increase in wealth and can contribute to generating thick right tails in wealth, with expanding wealth checked by finite lives or age independent death probabilities, i.e. perpetual youth demographics.

We restrict our analysis to models which can be micro-founded. We discuss and identify assumptions on preferences (including bequests), financial markets, and demographics so that consumption c_{t+1} is the solution of an optimal dynamic consumption-savings problem.

2.1 Skewed Earnings

While the environments and underlying assumptions of most micro-foundations of wealth accumulation models do not induce consumption functions that are linear in wealth, it is useful to postulate a linear consumption function, $c_t = \psi w_t + \chi(y_t)$, as a benchmark to establish some of the basic properties of wealth accumulation processes with stochastic earnings.[15] Consider also the case of a deterministic, indeed constant, rate of return on wealth, $r_t = r$. For these economies, Equation (1) becomes:

$$w_{t+1} = (r_{t+1} - \psi)\, w_t + (y_{t+1} - \chi(y_{t+1})) \tag{2}$$

The results for characterizing the stationary distribution for $\{w\}$ for this linear equation 2 will be given in section 2.3, Theorem 3, due to Grey (1994). We will however start with the simple case of a deterministic constant $r_t = r$, a very special case of Theorem 3, which we state for expository convenience as Theorem 1 below:

Theorem 1 *Suppose $r - \psi < 1$ and $\{y_t\}$ has a thick-tailed stationary distribution with tail-index α. Then the accumulation equation 2 induces an ergodic stationary distribution for wealth with tail not thicker than (bounded below by) α.*

More precisely, the stationary distribution of wealth has tail equal to the tail of (the stationary distribution of) $\{y_t - \chi(y_t)\}$. The behavior of the additive component of consumption $\chi(y_t)$, in a micro-founded model depends on the persistence of the stochastic process of earnings (see the next section). However, $\chi(y_t) \geq 0$, and hence the tail index of wealth, matching that of $(y_{t+1} - \chi(y_{t+1}))$, can diverge from earnings, but only in the direction of being thinner. In other words, for contracting economies with constant rates of return and linear consumption, the statistical properties of the tail of the wealth distribution are directly inherited from those of the distribution of earnings and hence the tail of the wealth distribution cannot be thicker than the tail of the distribution of earnings.

2.1.1 Micro-foundations

In this section we show how the case considered above, with linear consumption, constitutes the relevant benchmark case, once explicit micro-foundations are taken into account, to study the tail of the wealth distribution. 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.

Infinite horizon. Consider an infinite horizon Bewley-Aiyagari economy. Each agent's consumption-savings problem must satisfy a borrowing constraint and $\beta r < 1$. The borrowing constraint together with stochastic earnings generates a precautionary motive

[15] The parameter ψ and the map $\chi(y_t)$ generally depend on the properties of the stochastic process $\{r_t, y_t\}$.

for saving and accumulation and acts as a lower reflecting barrier for assets.[16] 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:

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

Furthermore, the additive component of consumption, $\chi(y_{t+1})$ can be characterized at the solution of the consumption-savings problem. Indeed, varying the stochastic process for y_t, say from a highly persistent $AR1$ to a finite Markov chain with a lower implicit persistence, affects the optimal choice of $\chi(y_{t+1})$. To develop some intuition consider the very stylized case in which $y_t \geq 0$ is deterministic, growing at some rate λ, and where $\lambda\beta r = 1$. In that case $\chi(y_{t+1}) = y_{t+1}$. Alternatively if y_t is $i.i.d.$, then $\chi(y_{t+1}) = E(y_t) = \bar{y}$ (for simplicity taking $\beta r = 1$).

As long as the right tail of wealth is concerned, therefore, the asymptotic linearity of $c(w_t)$, guarantees that Equation 2 approximates wealth accumulation in the economy and that Theorem 1 applies.[17] The condition $r - \psi < 1$ is an implication of $\beta r < 1$. With constant $(r - \psi)$, the tail of the wealth distribution is therefore the same as that of $(y_{t+1} - \chi(y_{t+1}))$. Therefore, $\chi(y_{t+1})$ determines the divergence between the tails of wealth and earnings (how much thinner is the tail of wealth than that of earnings). Specifically, if $\chi(y_{t+1}) = y_{t+1}$, the distribution of wealth does not have a thick tail (the tail index is ∞).[18] While, alternatively if $\chi(y_{t+1}) = E(y_t) = \bar{y}$ the tail of wealth coincides with the tail of earnings y_t. Intuitively, the higher is persistence, the thinner is the tail of wealth with respect to that of earnings.

Overlapping generations (OLG). Let n denote a generation (living for a length of time T). A given an intra-generation earnings profile, $\{y_{tn}\}_t$, 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. If the rate of return is constant in t, then also r_n is constant, say $r_n = n$. The initial wealth of each dynasty maps then into a bequest T periods later, which becomes the initial condition for the next generation. Without borrowing constraints[19] the inter-generational wealth accumulation equation is linear, that is Equation 2 holds inter-generationally: $w_{n+1} = (r - \psi)\, w_n + (y_{n+1} - \chi)$. As

[16]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 $r\beta < 1$ holds because capital also provides insurance against sequences of bad shocks.

[17]See Benhabib, Bisin and Zhu (2016) for a formal demonstration that the Aiyagari model with CRRA preferences is asmptotically linear.

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

[19]Because of the OLG structure $\beta r > 1$ is compatible with stationarity, and it guarantees that agents

a consequence, Theorem 1 applies and the tail index of $\{y_n - \chi\}$ translates into wealth in the stationary distribution as long as $r - \psi < 1$. The details of these arguments are exposed in Benhabib, Bisin, and Zhu (2011).

2.1.2 Models of skewed earnings

Several models of incomes or earnings y_t which produce a skewed distribution have been proposed in the literature that derive from basic heterogeneities of productivity and talent. Most models can typically be represented in terms of the *translation method* introduced by Edgeworth (1917) that was discussed in the previous section.

Talent. Suppose s denotes talent, which is exponentially distributed: $f_s(s) = pe^{-ps}$. Suppose also earnings y increase exponentially in talent: $y = e^{gs}, \quad g \geq 0$.[20] Then $f_y(y) = \frac{p}{g} y^{-\left(\frac{p}{g}+1\right)}$ and it is Pareto with exponent $\frac{p}{g}$. This is the model for the determination of earnings introduced by Cantelli (1921) and then refined by D'Addario (1943).

Frechet (1939) objected that the empirical wealth distribution was hump-shaped, and that the lower range of income was more log-normal than Pareto. He supposed instead that the distribution of talent follows a Laplace distribution, $f_s(s) = e^{-a|s|}, \quad s = (-\infty, +\infty)$ while income was exponentially increasing in $s : y = e^{bs}$, so that $s = b^{-1} \ln y$ and $\frac{ds}{sy} \quad = b^{-1}\frac{1}{y}$. Using the Edgeworth translation technique he obtained a distribution that is a power law above the median of 1:

$$f_y(y) = b^{-1}\frac{1}{y}e^{-a\left|\frac{\ln y}{b}\right|} = \begin{cases} b^{-1}\frac{1}{y}e^{-\frac{a}{b}\ln y} = b^{-1}y^{-\frac{a}{b}-1} & if \quad y \geq 1 \\ b^{-1}\frac{1}{y}e^{\frac{a}{b}\ln y} = b^{-1}y^{\frac{a}{b}-1} & if \quad y \leq 1 \end{cases}.$$

Note that for $y \leq 1, \quad f_y(y)$ is increasing if $\frac{a}{b} > 1$.

Schooling. Suppose schooling increases income, but at an opportunity cost in terms of lost time at discount rate $\frac{1}{1+r}$, and at a non-monetary marginal cost c, a measure of ability. Let s denote years of schooling and let $y(s)$ denote labor earnings with s years of schooling. Then, the optimal choice of schooling, implies that

$$y(s) \geq \underline{y}e^{rs}, \ s \geq s' \rightarrow y(s) \geq y(s')$$

If the marginal cost of schooling, c, is exponentially distributed, $f_c(c) = pe^{-pc}$, so are years of schooling s. The same *transformation* algebra used for talent in the previous example implies then that y has a distribution even more skewed than the Pareto distribution with exponent $\frac{p}{r}$. This is essentially Mincer's (1958) schooling model.[21]

---

do not borrow even if allowed to do so. Otherwise a borrowing constraint can be imposed and the analysis is straightforward.

[20] Note that an increase in g is like an increase in the skill-premium.

[21] In Mincer (1958)'s analysis, however, schooling is normally distributed for $s \geq 0$ and y has hence a log-normal distribution in the tail, since $\ln y = \ln \underline{y} + rs$.

Span of control. Consider the income y of an entrepreneur with talent s who has the opportunity to hire n agents at wage x to produce with production function $f(n, s) = sn^{\alpha}$. This entrepreneur will solve the following problem:

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

It follows that, for any $n > 0$, $y = A(x)s^{\frac{1}{1-\alpha}}$, for some $A(x)$.

Multi-dimensional human capital skills (Roy (1950)). Consider the following multiplicative labor earnings model: y is linear in human capital h, which is in turn the product of i.i.d. components h_i, $i = 1, \ldots, n$:

$$y = ah, \quad h = [h_1 h_2 ... h_m]^{\frac{1}{m}}.$$

It follows that

$$\ln y = \ln a + \frac{1}{n} [\ln h_1 + \ln h_2 ... + \ln h_m],$$

and, by the Central limit theorem, earnings have a log-normal distribution $m \to \infty$.[22]

Multi-dimensional human capital skills and assortative matching. Consider the income y of a worker of human capital h engaged in production for a firm. The Expected output of the firm, $E(y)$, is determined by the "O-Ring" production function, as in Kremer (1993):

$$E[y] = k^a (h_1 h_2 ... h_m) n B$$

where k denotes capital, h_i is probability that worker in task i performs, an index of worker i's skills, m denotes the total number of tasks, and B is a firm productivity parameter. A firm chooses $h_1, h_2 ... h_m$, and k to maximize

$$Max \ E(y) - y(h_1) - y(h_2) ... y(h_m) - rk$$

where $y(h)$ is the market earnings function, inducing the workers' income as a function of their talent.

Because of complementarity between workers' skills, $\partial^2 E[y] / \partial h_i \partial h_j > 0$, ceteris paribus, a firm with high h_j will be willing to bid more for h_i. In equilibrium, therefore workers of the same talent will be matched together, that is, talent will be matched assortatively. We let this common skill within a firm be represented by h. At equilibrium then, first order conditions for profit maximization imply

$$mh^{m-1} B (ah^m m/r)^{a/(1-a)} - \frac{dy(h)}{dh} = 0;$$

[22]If the h_i's are not independent but positively correlated, the skewness of the distribution of y can increase. See Haldane (1942).

a differential equation whose solution is

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

In equilibrium $y(h)$ is homogeneous of degree $m/(1-a) > 1$ in h: small differences in skills h translate into large differences in income y. Indeed $y(h)$ is a convex function; so that labor earnings y are skewed to the right even if h is distributed symmetrically.[23] Consider for instance the case in which h is uniformly distributed: $f_h(h) = b, \qquad 0 \leq h \leq b^{-1}$. Then, by the transformation method,

$$f_y(y)\frac{dh}{dy} = bC^{-1}\left(\frac{1-a}{m}\right)y^{\left(\frac{1-a}{m}-1\right)}$$

where $C = (1-a)B^{1/(1-a)}(am/r)^{a/(1-a)}$ and $\frac{1-a}{m} < 1$. Earnings are then distributed as a power law over $0 \leq (C^{-1}y)^{\frac{1-a}{m}} \leq b^{-1}$, even if skills are distributed uniformly.

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 higher number of people, $n_i > n_{i+1}$ are assigned. Assume that technologically, $n_i = \gamma n_{i+1}$, for some $\gamma > 1$. Also wage earnings at level $i+1$, x_{i+1} are proportional to the wage bill in the contiguous lower level i (e.g., suppose higher level workers manage lower level ones): $x_{i+1} = q\gamma x_i$. It follows that

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

In the discrete distribution we have constructed, n_i is the number of agents with wage earnings x_i. It is clear that a discrete Pareto distribution, $n_i = b(x_i)^{-(\alpha+1)}$ satisfies this condition for $\alpha + 1 = \frac{\ln\gamma}{\ln\gamma q}$. Replicating the economy to make the bins i smaller produces a Pareto distribution of wage earnings in the limit.

2.2 Stochastic returns

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 2 when the rate of return r_t follows a well-defined stochastic process. Its relevance for wealth distributions is that it can generate skewed and thick-tailed distributions even when the distribution of r_t is neither skewed nor fat-tailed.

The accumulation equation 2 defines a *Kesten process* if (r_t, y_t) are independent and $i.i.d$ over time; and if for any $t \geq 0$ it satisfies the following:[24]

[23] 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.

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

$$0 < (y_t - \chi_t(y_t)) < \infty, \;\; 0 < E(r_t) - \psi < 1, \text{ and } prob\,(r_t - \psi > 1) > 0.$$

These assumptions guarantee, respectively, that the wealth process has a reflecting barrier and that it is contracting on average, while expanding with positive probability.

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

Theorem 2 (Kesten) *Suppose the accumulation Equation 2 defines a* Kesten process *and $\{y_t\}$ is thin tailed. Then the induced wealth process displays an ergodic stationary distribution which has Pareto tail α:*

$$\lim_{\bar{w}\to\infty} prob(w \geq \bar{w})\bar{w}^{\alpha} \sim k,$$

where k is a constant and $\alpha > 1$ satisfies $E(r_t - \psi)^{\alpha} = 1$.

Allowing for negative earning shocks in Kesten processes induces a double Pareto distribution, thereby allowing the process to deal at least in part with Edgeworth's criticism of Pareto:

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

with $C_1 = C_2 > 0$ under regularity assumptions.

Most importantly, for the study of wealth accumulation, recent results extend the characterization result for *generalized Kesten processes* where (r_t, y_t) is allowed to be a general Markov process, hence r_t correlated with y_t and both auto-correlated over time.[25]

2.2.1 Micro-foundations

The micro-foundations for economies with stochastic returns are essentially the same as those with constant r_t delineated in the previous section. 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, and the results of Kesten (1973) for characterizing tails, as generalized by Mirek (2011), still apply; see Benhabib, Bisin and Zhu (2016) and Achdou, Han, Lasry, Lions, Moll (2016) for formal proofs in discrete and continuous time, respectively. Similarly, the micro-foundations in terms of inter-generational accumulation for OLG economies naturally extend to stochastic r_t; this is in fact the case in Benhabib, Bisin and Zhu (2011).

[25] The Kesten results can also also extend to be continuous time, for an accumulation process defined by:

$$dw = r\,(X)\,wdt + \sigma\,(X)\,d\omega,$$

where $r(X)$, $\sigma\,(X) > 0$, and $d\omega$ is a Brownian motion; see Saporta and Yao (2005). See also Benhabib, Bisin, Zhu (2011), Gabaix, Lasry, Lions and Moll (2015) for applications.

2.3 Stochastic returns vs. skewed earnings

We have seen in Theorem 1 that linear (or asymptotically linear) wealth accumulation processes in economies with deterministic returns and skewed thick tailed distributions of earnings induce wealth distributions as 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* $(y_t - \chi_t(y_t))$ *are both random variables, independent of* w_t*. Suppose the accumulation Equation 2 defines a* Kesten process *and* $(y_t - \chi_t(y_t))$ *has a thick-tail with tail-index* $\alpha > 0$*. Then,*

If $E\left((r_t - \psi)^{\alpha}\right) < 1$*, and* $E\left((r_t - \psi)^{\beta}\right) < \infty$ *for* $\beta > \alpha$*, and under some regularity assumptions, the tail of the stationary distribution of wealth will be the same as that of* $\{y_t - \chi_t(y_t)\}$ *.*

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

Theorem 3 makes clear that the tail index of the wealth distribution induced by Equation 2 is either α', which depends on the stochastic properties of returns, or α, the tail of the earnings distribution.[26] In other words, it is never the case that a stochastic process describing the accumulation of wealth could amplify the skewness of the wealth distribution that is induced by skewed earnings; it's either the accumulation process or the skewed earnings which determine the thickness of the tail of the wealth distribution.

2.4 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 wealth accumulation in equation 2 explodes, that is, if $r_t - \psi > 1$ on average (for at least a sub-class of the agents in the economy), or if a non-contracting process for r_t is postulated which does not satisfy the Kesten conditions. In this case, however, the distribution of wealth will generally not converge to a limit distribution. As we noted discussing Wold and Whittle (1957) in Section 1,2, a number of birth and death mechanisms can be super-imposed onto economies with explosive savings to generate a skewed stationary distribution of wealth.

[26] See Ghosh et al (2010) for extensions of Grey (1994) to random, Markov-dependent (persistent) coefficients $(y_t - \chi)$ and $(r_t - \psi)$, that also allow for correlations between them.

To illustrate this in the context of a micro-founded model, we start with a very simple model introduced by Blanchard (1985), a deterministic economy characterized by explosive wealth accumulation and perpetual youth, that is, constant mortality rate.[27] 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. They discount the future at rate β. In this model agents' wealth grows proportionally at stationary equilibrium rates during lifetimes and across generations. Consumption is linear in $w + h$, where $h = \frac{y}{r+p}$ is the present discounted value of earnings and wealth w_t satisfies

$$w_t = \frac{p-\beta}{(r+p)(r-\beta)} y \left(e^{(r-\beta)t} - 1\right).$$[28]

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 $\underline{w}$.[29] The translation method, once age takes the position of talent, implies

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

which is Pareto in the tail, that is, for large w.

In this model, therefore, 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 $\underline{w}$ which is independent of their parents' wealth (a reflecting barrier) is crucial, however, as it is in Wold and Whittle (1957). In general, the re-insertion of newborns at a wealth level corresponding to a fixed fraction of the wealth of their parents at death would simply dilute the growth rate on average, but would be insufficient to guarantee stationarity.

Extending the model to allow for income y to grow at the exogenous rate γ, we can better relate the resulting Pareto exponent for wealth (discounted at the rage γ) with the one obtained by Wold and Whittle (1957). In the Blanchard model the exponent is $\frac{p}{r-\gamma-\beta} + 1$, which induces a thicker tail for higher $(r-\gamma)$; while in Wold and Whittle (1957), relatedly, the tail depends on r/γ.

Note also that the model implies that wealth will be correlated with age (or, in extensions with bequests, with the average lifespan of ancestors). The heterogeneity in

[27] Several recent papers use features of the perpetual youth model to obtain thick tails; see for example Benhabib and Bisin (2006), Piketty and Zucman (2013), and Jones (2015).

[28] See Benhabib and Bisin (2007) and Benhabib and Zhu (2009), where the full optimization dynamics is spelled out in a more general stochastic continuous time model.

[29] See Benhabib and Bisin (2007) for the endogenous determination of $\underline{w}$ via a social security system funded by taxation.

wealth across agents arises from heterogeneity in lifespans.

Stochastic rates of return might also induce non-stationarity if they follow a non-contracting stochastic process that does not satisfy the Kesten condition delineated in the previous section. As we saw in Section 1.2, this is the case, for instance, when the accumulation equation follows Gibrat's law. In this context as well, Reed (2001) shows that a different birth-death process can re-establish stationarity. More specifically, he studies exponentially distributed death times and generates a "Double-Pareto" distribution of wealth, with those dying replaced by newborns with initial wealth w_0.[30] Assuming wealth evolves in continuous time, with a constant positive drift (rate of return) r, and a geometric Brownian motion as diffusion,[31] 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:

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

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 can be increasing in wealth for $w < w_0$ if $\beta > 1$. As Reed (2001) notes, this is a hump-shaped "Double-Pareto" distribution, more appropriate to describe the empirical wealth distribution than a Pareto distribution.[32]

Another class of models with explosive wealth dynamics is characterized by heterogeneous savings rates, prominently appearing in the early work of Kaldor (1957, 1961), Pasinetti (1962) and Stiglitz (1969). A more recent example of this approach is Carroll, Slajek and Tokuo (2014b). Notably Carroll, Slajek and Tokuo (2014b) also introduce a constant probability of death for agents, replacing the dead by injecting new-born agents at low levels of wealth, as Blanchard's model.

Yet another mechanism to help generate thick tails in wealth is to directly produce a savings rate that increases in wealth. Atkinson (1971) obtains such a savings rate in an OLG economy with constant rate of return on wealth, finitely lived agents, and warm

[30] Reed (2003) generalizes the point initial condition w_0 to allow the initial state to be a log-normal distribution.

[31] See Benhabib, Bisin and Zhu (2014) for the micro-foundation of such accumulation process.

[32] A particularly simple solution can be obtained with simplfying assumptions, following Mitzenmacher (2004), pp. 241- 242. Suppose $w_0 = 1$ and $r = \frac{\sigma^2}{2}$, $\sigma = 1$. Then

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

glow preferences for bequests given by

$$v(w_T) = A\frac{(w_T^n)^{1-\mu}}{1-\mu},$$

where w_T is the end of life wealth, that is, bequests. For these economies it is straightforward to show that, if the inter-temporal elasticity of substitution in consumption, σ, is greater than the elasticity of bequests, μ, so that the utility of consumption is flatter than that of bequests, the propensity to consume out of wealth, $\frac{c(w_t)}{w_t}$, is decreasing in wealth, and therefore savings rates are increasing in initial wealth.[33] A related approach is due to De Nardi (2004) and to Cagetti and De Nardi (2008), who explicitly introduce non-homogenous bequest motives:

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

where γ measures how much bequests increase with wealth.[34]

Finally another possible mechanism for generating wealth inequality of course is the rich making higher post-tax returns, where $r_t = r(w_t)$ is an increasing function r.[35]

3 Empirical evidence

in our theoretical survey, we identified three basic mechanisms that can contribute to generate wealth distributions that have thick tails: Stochastic earnings, stochastic returns, and exploding wealth accumulation. We here 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, 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.

[33] 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.

[34] To get thick tails in wealth, Cagetti and DeNardi (2008) use a hybrid model: in addition to non-homogenous bequest functions they introduce stochastic returns as in Quadrini (2000). See also Roussanov (2010) for role of status concerns in accumulation incentives, distribution of asset holdings, and mobility.

[35] The evidence for returns increasing in wealth is somewhat inconclusive. Averaged over the period 1980-2012, recent 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. The post-tax returns on wealth for the period 1960-1980 however are clearly decreasing, possibly due to higher capital income taxes during that period. Recent contributions, using administrative data from Scandinavian countries also show pre-tax returns increasing in wealth. Fagereng, Guiso, Malacrino and Pistaferri (2015, 2016) find returns significantly increasing in wealth only for high wealth classes, above the top 10%, in Norway. Bach, Calvet, and Sodini (2015) find higher retuns on large wealth portfolios for Sweden.

3.1 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; see Guvenen (2015), Guvenen, Karahan, Ozkan, and Song (2015), and De Nardi, Fella, and Parlo (2016) for detailed recent studies of earnings.
Earnings distributions are skewed. Atkinson (2002), Moriguchi-Saez (2005), Piketty (2001), Piketty-Saez (2003), 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, the U.S., and Canada.
Earnings distributions display thick upper tails; that is, slowly declining top shares. Thick upper tails (power law behavior) for the distributions of earnings are also well documented, for example by Nirei-Souma (2004) in the U.S. and Japan from 1960 to 1999, by Clementi-Gallegati (2004) for Italy from 1977 to 2002, and by Dagsvik-Vatne (1999) for Norway in 1998. The Gini coefficient of the distribution of earnings for the U.S. in 1992 is .57; see Diaz Gimenez-Quadrini-Rios Rull (1997) and Feenberg-Poterba (2000).
Earnings distributions display thinner upper tails than the wealth distribution. For example the earnings distribution Gini coefficient for the US is about 0.48 while it is about 0.8 for wealth.

All this implies that realistic distributions of earnings by themselves, without other complimentary mechanisms, have difficulty in generating the skewed wealth distributions we observe. This is because the asymptotic linearity of consumption, as we noted, implies that the tail index of $y_t - \chi(y_t)$ will be inherited by wealth. Indeed, 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)."[36]

For example Krueger and Kindermann (2014) show that a version of the Aiyagari-Bewley model with sufficiently skewed earnings can be made to match the empirical wealth distribution. But this requires excessive and empirically unrealistic tails for the distribution of earnings. Krueger and Kindermann (2014) use a seven state Markov chain model for earnings that implies, in the stationary distribution, that the average top 0.25% earn somewhere between 400 to 600 times the median income, implying earnings for the top 0.25% of at least $20,000,000. But consider the World Wealth and Income Database (WWID) by Facundo Alvaredo, Anthony B. Atkinson, Thomas Piketty, Emmanuel Saez, and Gabriel Zucman (2016), which is not top-coded. In the WWID, the top 0.1%

[36] 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."

average income in 2014, excluding capital gains, is \$4,128,000 and average earnings (wages, pensions and salaries) are \$1,716,000 while the median income is about \$50,000, a ratio of about 34 (it would be of course even lower if we considered the top 0.25% instead of the top 0.1%).

Relatedly, Castenada et al (2003) develop a very rich overlapping-generation model with life-cycle features, borrowing constraints as in Aiyagari-Bewley models, constant retirement and death probabilities independent of age, perpetual youth demographics with accidental inheritances, pensions, income and estate taxes, and persistent stochastic labor endowments. They carefully calibrate their model to the US economy to generate an excellent match to wealth distributions. Their 4-state labor endowment stochastic process is however highly skewed, so that at the stationary distribution for labor endowments, the top 0.039% earners have 1000 times the average labor endowment of the bottom 61%. Thus to attain a ratio of a 1000, if the bottom 61% earn \$32,000 a year on average[37], the top 0.0389% would have to have earnings of \$32,000,000, quite excessive according to the WWID. For top incomes from all sources the WWID gives, excluding capital gains, an average income for the top 0.01% in 2014 of \$17,180,000 and earnings (wages, pensions and salaries) of about \$6,000,000, which of course would be lower for the top 0.0389%. Life cycle/consumption-smoothing considerations arise naturally, so agents at the rare and somewhat persistent highest labor endowment state, sometimes called the "awesome" state, save at higher rates and accumulate wealth faster. These agents decumulate during retirement, but invariably some fraction die early into their retirement, leaving large accidental bequests. A similar mechanism is at work in Diaz et al (2003) who use a standard Aiyagari model with infinitely lived agents and three earnings states. At the top "awesome" state, the top 6% of the population earn 46 times the labor earnings of the median, a highly excessive skew relative to the earnings data. According again to the WWID, the 5% of top incomes average \$367,100 in 2013, of which only 69.44% or \$255,000 are labor earnings (wages, salaries plus pensions). The median incomes however are about \$50,000, a factor of 5.1, not 46.

Even if the extraordinary or "awesome" earnings states are more modest, perpetual youth demographics and random working life-spans can complement skewed earnings by producing additional heterogeneity in wealth accumulation across agents. Unlike the case where all agents are infinitely lived, variable life-spans produce differential sojourn times in the high earnings states in otherwise standard Aiyagari-Bewley models. Those agents who end up not only having long working-life spans, but are also lucky enough to spend a good deal of their working life in the high earnings states, work longer hours and save at higher rates for precautionary and for retirement reasons. This leads to variation in wealth accumulation rates across agents and with bequests, across dynasties. Thus variable working life-spans produce a fraction of agents with high sojourn times in high

[37] See Table A2 from the Census Bureau for 2014 at: http://www.census.gov/content/dam/Census/library/publications/2015/demo/p60-252.pdf

earnings states which in turn imply a better fit to the wealth distribution without relying on extreme skewness in earnings. 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 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. The thick right tail of the wealth distributions will then have dynasties with long average life-spans spent in high earnings states. Kaplan, Moll and Violante (2015) feed leptokurtic innovations from Guvenen (2015) and Guvenen, Karahan, Ozkan, and Song (2015) to the earnings process into a model with fixed costs of portfolio adjustments for high-return illiquid assets in a perpetual youth model with annuities and no bequests. They use the same perpetual youth calibration of expected working lives of 45 years as Kaymak and Poschke (2015), again with 11% of the population working more than 100 years. Wealthy agents hold more of the illiquid assets that have fixed costs of liquidation, while some poorer agents hold zero illiquid wealth. A fraction of very long-lived agents end up spending a long time in high earnings states, hold high-return illiquid assets, earn higher returns, and accumulate more wealth during their lifetime. These agents populate the tail of the wealth distribution. So without unrealistically skewed labor earnings, or extreme life-span variability, it may not be possible for agents to accumulate enough wealth to populate the tail of wealth distribution via saved earnings alone." [38]

3.2 Stochastic returns

Data on stochastic return is relatively hard to find. This is in part because of the conceptual difficulties involved in mapping r_t in the data with a measure of idiosyncratic rate of return on wealth, or capital income risk. The most relevant component of capital income risk appear to be i) returns to ownership of principal residence and private business equity; and ii) returns on private equity. Some stylized facts are useful in this case as well.

[38] By contrast, in a recent paper De Nardi, Fella, and Pardo (2016), adapt earnings data from Guvenen, Karahan, Ozkan, and Song (2015), and introduce it into a finite-life OLG model. They note that earnings processes derived from data, including the one that they use, "when introduced in a standard quantitative model of consumption and savings over the life cycle, generate a much better fit of the wealth holdings of the bottom 60% of people, but vastly underestimate the level of wealth concentration at the top of the wealth distribution" (p.37). They repeatedly stress that this can be due to many important mechanisms (bequests, entrepreneurship, inter-vivos transfers, etc.) from which their model abstracts, but which are discussed in Cagetti and De Nardi (2006, 2007, 2008) and De Nardi (2004). In fact, Kaplan, Moll and Violante (2015) also make clear that they cannot match the very top tail of the wealth distribution with earnings alone, and suggest introducing alternative mechanisms like stochastic returns (see their footnote 32).

The idiosyncratic component of capital income risk appears to be significant. Returns to ownership of principal residence and private business equity account for, respectively, 28.2% and 27% of household wealth in the U.S., according to the 2001 *Survey of Consumer Finances* ; see Wolff, 2004 and Bertaut-Starr-McCluer, 2002). From a different angle, 67.7% of households own principal residence (16.8% own other real estate) and 11.9% of household own unincorporated business equity.[39]
The idiosyncratic component of capital income risk appears highly variable. 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, obtaining a similar number, 14%. Returns on private equity have an even higher idiosyncratic dispersion across household, 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 (Moskowitz and Vissing-Jorgensen, 2002). In the 1989 *SCF* studied by Moskowitz and Vissing-Jorgensen (2002), the median of the distribution of returns on private equity is 6.9%, while the first quartile is 0 and the third quartile is 18.6%; see Angeletos (2007) and Benhabib and Zhu (2008) for more evidence on the macroeconomic relevance of idiosyncratic capital income risk.

Recently Fagereng, Guiso, Malacrino and Pistaferri (2015), using Norwegian administrative data, find that returns to wealth exhibit substantial heterogeneity. For example, for 2013 they find that the average (median) return on overall wealth is 2.8% (1.9%), but varies significantly across households, with a standard deviation of returns of 4%. Furthermore, they note that "... heterogeneity in returns is not simply the reflection of differences in portfolio allocations between risky and safe assets mirroring heterogeneity in risk aversion. On the one hand, when regressing the returns to wealth on a set of covariates, [...] the portfolio share of risky assets has only little explanatory power. On the other hand, [...] returns heterogeneity [is present] even when [focusing] on safe assets, although the dominant source of returns heterogeneity originates from heterogeneity within risky assets."

In models where wealth distribution does not explode after discounting for an underlying growth trend, stochastic returns across agents contribute to generating thick-tailed wealth distributions.[40] Quadrini (2000) first explicitly introduced stochastic returns to

[39] Quadrini (2000) also extensively documents the role of returns to entrepreneurial talent in wealth accumulation. Evidently, the presence of moral hazard and other frictions render risk diversification, as well as concentrating each household's wealth in the hands of the household with the best investment technology, is hardly feasible.

[40] The possibility of stochastic returns generating thick tails in a basic Aiyagari-Bewley model was already implicit in the work of Krusell and Smith (1998) who matched the upper tail of wealth by introducing stochastic discount factors across agents (since it is discounted returns that matter for accumulation, that stochastic discount rates across agents operate like stochastic returns).

entrepreneurial abilities under market imperfections and financial constraints, and was able to produce skewed and thick-tailed wealth distributions. Cagetti and De Nardi (2006) built on the entrepreneurial model of Quadrini (2000) in an OLG model with bequests, with constant probabilities for retirement and death independent of age, and also produced thick-tailed distributions. These calibrated models with stochastic returns relied on simulations for obtaining thick-tails. Benhabib, Bisin and Zhu (2011), using the methods of Kesten (1973), Saporta (2005) and Roitherstein (2007), formally obtained thick-tailed wealth distributions in OLG models with finite lives, and simulated their model to match the data.

3.3 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, it 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. Consider in this respect the issue of transmission of inequality across generations, and its response to taxes. In this respect, Becker and Tomes constructed an OLG model with two period lives and introduced altruistic investments by parents in the earning ability of their children, as well as the transmission of earnings ability through spillovers from parents to children within families, and from average abilities in the economy. They also introduced a random element of luck in earnings ability. 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: 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 in models with a skewed distribution of earnings but no capital income risk. However Benhabib, Bisin and Zhu (2011) show that if the capital income risk component is a substantial fraction of idiosyncratic risk, reducing bequest taxes, or amplifying the heterogeneity of after-tax returns on capital by reducing capital income taxes, could have sizable effects in increasing wealth inequality in the top tail of the distribution of wealth, which may not show up in measurements of the Gini coefficient.

Empirically, to assess the relevance of the various mechanisms that generate thick-tailed wealth distributions Benhabib, Bisin and Luo (2015) estimate a model that nests them. The authors use 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). Each agent's life is finite, T years. Each consumer of dynasty j chooses consumption $\{c_{j,t}\}$ and savings each period, subject to a no-borrowing constraint and a bequest $e^n_{j,T}$. Consumer of dynasty j, generation n, also draws a lifetime return r^n_j, and a deterministic earnings profile $\left\{w^n_{j,t}\right\}_0^T$ parameterized by $w^n_{j,0}$, and maximizes utility (single heir, no estate tax for simplicity). This abstracts from the precautionary savings motive during the life of a consumer, so that wage profiles and returns become stochastic only across generations.[41] The optimization problem of consumer j is:

$$V_j^n\left(a^n_{j,0}\right) = Max_{\left\{c^n_{j,t}\right\}} \sum_{t=0}^{T} \frac{\left(c^n_{j,t}\right)^{1-\sigma}}{1-\sigma} + A\frac{\left(w^n_{j,T}\right)^{1-\mu}}{1-\mu} \qquad for\ t \in [0,T]$$

$$\begin{aligned} s.t.\ w^n_{,j,t+1} &= (1+r^n_j)(w^n_{j,t} - c^n_{j,t}) + y^n_{j,t} \\ 0 &\leq c^n_{j,t} \leq w^n_{j,t} \end{aligned}$$

If $\mu > \sigma$, the utility of consumption is flatter than that of bequests, so is savings rates increase in wealth. Generations are connected through bequests and inheritances: $w^n_{j,T} = w^{n+1}_{j,0}$. The estimation is by Simulated Method of Moments, fixing several parameters of the model, selecting some relevant moments, and estimating the remaining parameters by matching the moments generated by the model and those in the data, specifically, setting $\sigma = 2$, $T = 36$, $\beta = 0.97$ per annum as well as the stochastic process for individual income and its transition across generations, following Chetty et al. (2014). The objective is to match several moments: the wealth percentiles, bottom 20%, $20-40\%$, $40-60\%$, $60-80\%$, $90-95\%$, $95-99\%$, and top 1%, as well as the diagonal of a wealth mobility matrix[42] by estimating the rate of return process and the bequest elasticity and the bequest intensity parameter A.[43] The results give a good match to wealth distribution and mobility. Benhabib, Bisin and Luo (2015) then estimate separate counterfactuals that sets a constant return r, a utility with bequest elasticity equal to that of consumption implying linearity of savings rates in wealth, and a constant low wage $\bar{w}$. Their findings indicate that all three features are important: stochastic incomes prevent 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. Furthermore

[41] There is evidence that in fact life-time earnings are already largely determined or forecastable early in life; see Huggett, Ventura and Yaron (2011), Cunha et al. (2010) or Keane and Wolpin (1997).

[42] The mobility matrix they use is from Klevmarken et al. (2003). Using other mobility transition matrices, as in Hurst and Kerwin (2003) makes little difference to the results.

[43] The mean and standard deviation of estimated returns closely match those estimated by Fagereng, Guiso, Malacrino and Pistaferri (2015) for Norwegian administrative data.

both stochastic returns and earnings are important to match wealth mobility, and an elastic bequest utility ($\mu = 1.1 < 2$) implying that savings rate increases in wealth is also essential to match the tail of the wealth distribution.

4 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.[44] First of all, the distribution of wealth in principle depends 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 changing political power and fortunes of social interest groups can appropriate economic advantages or can increase or decrease various forms of redistribution towards themselves.

[44]Some of these have been informally highlighted by Piketty (2014), however.

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

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

Aoki, Shuhei, and Makoto Nirei. 2015. "Zipf's Law, Pareto's Law, and the Evolution of Top Incomes in the U.S." Hitotsubashi University Working Paper.

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.

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.

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.

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", Manuscript, New York University.

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. (2014), “The Distribution of Wealth in the Blanchard-Yaari Model,” , *Macroeconomic Dynamics,* http://dx.doi.org/10.1017/S1365100514000066, online version 09 April 2014.

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. (2016). “The Wealth Distribution in Bewley Models with Capital Income”, forthcoming, *Journal of Economic Theory.*

http://www.econ.nyu.edu/user/benhabib/lineartailNBER-07-22-14.pdf

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

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.

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

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

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.

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.

De Nardi, M. 2004. “Wealth Inequality and Intergenerational Links.” *Rev. Econ. Studies* 71 (July): 743–68.

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.

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.

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.

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, Francesco Paolo, "Sulla deduzione delle leggi di frequenza da considerazioni di probabilita," *Metron*, 1 (April 1921), 83-91.

Cantelli, Francesco Paolo, "Sulla legge di distribuzione dei redditi," *Giornale degli Economisti e Rivista di Statistica* [4], 69 (November 1929), 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.

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

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.

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", working paper.

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.

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

Edgeworth, F. Y., "On the Representation of Statistics by Mathematical Formulae," Part I, *Journal of the Royal Statistical Society*, 61 (December 1898), 670-700.

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

Facundo Alvaredo, Anthony B. Atkinson, Thomas Piketty, Emmanuel Saez, and Gabriel Zucman (2016), *The World Wealth and Income Database*, http://www.wid.world

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.

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.

Frechet, Maurice, "Sur les formules de repartition des revenus," *Revue del'Institut International de Statistique*, 7 (No. 1, 1939), 32-38.

Fréchet, Maurice, :"Letter to the Editor", *Econometrica*, Vol. 26, No. 4 (Oct., 1958), pp. 590-591.

Gabaix, X., Lasry, J-M., Lions, P-L, and B. Moll,"The Dynamics of Inequality,"

http://www.princeton.edu/~moll/dynamics.pdf

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

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.

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

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. 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.files.wordpress.com/2014/04/handout_inequality_risk_webversion_may20151.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, Mark, Gustavo Ventura, and Amir 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.

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.

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.

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

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

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

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.

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

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.

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 (2004): "Two Factor Model of Income Distribution Dynamics," mimeo, Utah State University.

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.

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 Ríos-Rull, J-V, (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. A., (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.

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.

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.

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.

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. (2004): "Changes in Household Wealth in the 1980s and 1990s in the U.S.,"mimeo, NYU.

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