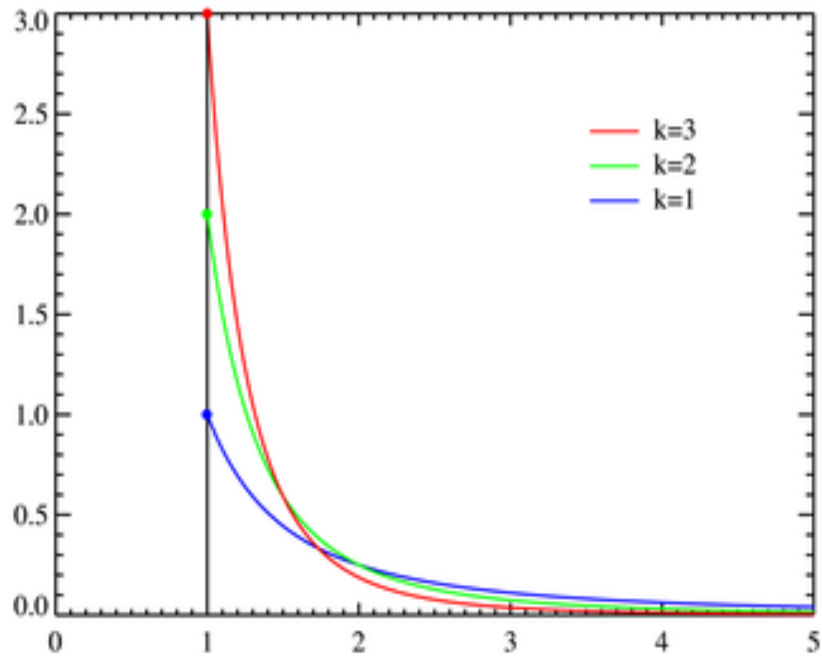


Survey of Consumer Finances, 2004



Pareto distribution:

$$\frac{k}{x_m} \left(\frac{x}{x_m} \right)^{-k-1} = k (x_m)^k x^{-k-1}$$

0.0.1 On old mechanisms possibly underlying a Pareto distribution of wealth (Cantelli-1921)

Suppose a variable determining wealth (e.g., talent, age), which we denote α , is exponentially distributed.

$$N(\alpha) = pe^{-p\alpha}$$

Suppose wealth increases exponentially with α :

$$w = e^{g\alpha}, \quad g \geq 0 \tag{1}$$

Therefore, we can solve for $\alpha = g^{-1} \ln w$, operate a change of variables and express the distribution of wealth as

$$N(w) = N(g^{-1} \ln w) \frac{d\alpha}{dw}$$

that is,

$$N(w) = \frac{p}{g} w^{-\left(\frac{p}{g} + 1\right)}$$

It is a Pareto distribution with exponent $\frac{p}{g}$.

A similar mechanism which makes wealth Pareto distributed is one in which the factor α is represented by age. At any time t , in this economy, the distribution of the population by age $t - s$ implied by the demographic structure of the economy is in fact

$$N(t - s) = pe^{-p(t-s)}$$

Moreover, abstracting from the complications of inheritance, each optimal consumption-savings choices imply a wealth accumulation process results in wealth increasing exponentially with age.

Frechet (1939) Use Laplace distribution for talents/skills:

$$N(s) = e^{-a|s|}, \quad s = (-\infty, +\infty)$$

Income:

$$\begin{aligned} w &= e^{bs}, \quad s = b^{-1} \ln w \\ \frac{ds}{sw} &= b^{-1} \frac{1}{w} \end{aligned}$$

Edgeworth translation technique:

$$N(w) = b^{-1} \frac{1}{w} e^{-a|\frac{\ln w}{b}|} = \begin{cases} b^{-1} \frac{1}{w} e^{-\frac{a}{b} \ln w} & \text{if } w \geq 1 \\ b^{-1} \frac{1}{w} e^{-\frac{a}{b} \ln w} & \text{if } w \leq 1 \end{cases} = \begin{cases} b^{-1} w^{-\frac{a}{b}-1} & \text{if } w \geq 1 \\ b^{-1} w^{\frac{a}{b}-1} & \text{if } w \leq 1 \end{cases}$$

Note: for $w \leq 1$, $N(w)$ is increasing

Pareto in Blanchard's Model Here θ is the discount rate, p is the constant death probability and also the fair annuity premium, r is the return on wealth. For wealth to remain bounded a standard assumption (See Blanchard (JPE, 1985)) is $r < p + \theta$, where $p + \theta$ is the effective augmented discount rate of the agent with constant death probability p .

$$\begin{aligned}\dot{w} &= (r + p)w + y - c \\ c(s, t) &= (p + \theta)(w + h) \\ h &= \int_t^\infty y(s, v) e^{\int_t^v (r(u) + p) du} dv \\ h &= y \int_t^\infty e^{-(r+p)(v-t)} dv = -y(r+p)^{-1} e^{-(r+p)(v-t)} \Big|_t^\infty \\ &= y(r+p)^{-1}\end{aligned}$$

$$\begin{aligned}\dot{w} &= (r + p)w + y - (p + \theta)(w + h) \\ &= (r - \theta)w + y - (p + \theta)h \\ &= (r - \theta)w + y - (p + \theta)y(r+p)^{-1} \\ &= (r - \theta)w + y \left(\frac{r - \theta}{r + p} \right)\end{aligned}$$

$$w(t) = \left(w(0) + \frac{y}{r + p} \right) e^{(r-\theta)t} - \frac{y}{r + p}$$

so for large w the agent wealth grows at rate $r - \theta$. If $w(0) = 0$, t is age:

$$\begin{aligned}w(t) &= \left(\frac{y}{r + p} \right) e^{(r-\theta)t} - \frac{y}{r + p} \\ &= \frac{y}{r + p} \left(e^{(r-\theta)t} - 1 \right)\end{aligned}$$

$$\begin{aligned}w(t) + \frac{y}{r + p} &= \left(\frac{y}{r + p} \right) e^{(r-\theta)t} \\ \ln \left(\frac{w(t)(r + p)}{y} + 1 \right) &= (r - \theta)t \\ t &= \frac{\ln \left(\frac{(r+p)w}{y} + 1 \right)}{r - \theta}\end{aligned}$$

Transform variables:

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

$$N(t) = pe^{-pt}, \int_0^{\infty} pe^{-pt} = 1$$

$$\begin{aligned} N(w) &= N(t(w)) \frac{dt}{dw} \\ &= pe^{-p \frac{\ln\left(\frac{(r+p)w}{y} + 1\right)}{r-\theta}} \left[\left(\frac{(r+p)w}{y} + 1 \right)^{-1} \frac{(r+p)}{(r-\theta)y} \right] \end{aligned}$$

$$\begin{aligned} N(w) &= pe^{\ln\left(\frac{(r+p)w}{y} + 1\right) \frac{-p}{r-\theta}} \\ &\quad \cdot \left[\left(\frac{(r+p)w}{y} + 1 \right)^{-1} \frac{(r+p)}{(r-\theta)y} \right] \end{aligned}$$

$$\begin{aligned} N(w) &= p \left(\frac{(r+p)w}{y} + 1 \right)^{\frac{-p}{r-\theta}} \\ &\quad \cdot \left[\left(\frac{(r+p)w}{y} + 1 \right)^{-1} \frac{(r+p)}{(r-\theta)y} \right] \end{aligned}$$

Pareto density in $W = \frac{(r+p)w}{y} + 1$:

$$\begin{aligned} N(w) &= \left[\frac{p(r+p)}{(r-\theta)y} \right] \left(\frac{(r+p)w}{y} + 1 \right)^{\frac{-p}{r-\theta}-1} \\ N(0) &= \left[\frac{p(r+p)}{(r-\theta)y} \right] \end{aligned}$$

with Pareto exponent $\frac{p}{r-\theta}$. It is greater than 1 under the assumption $r < p + \theta$.
Check that population integrates to 1:

$$\begin{aligned} &\int_0^\infty N(w) dw \\ &= \left[\frac{p(r+p)}{(r-\theta)y} \right] \left[\left(-\frac{r-\theta}{p} \right) \left(\frac{y}{r+p} \right) \left(\frac{(r+p)w}{y} + 1 \right)^{\frac{-p}{r-\theta}} \right]_0^\infty \\ &= \left[\frac{p(r+p)}{(r-\theta)y} \right] \left(\frac{y}{r+p} \right) \left(\frac{r-\theta}{p} \right) = 1 \end{aligned}$$

If we allow incomes y to grow over time at rate g , and discount wealth by g , then the Pareto exponent becomes $\frac{p}{(r-g)-\theta}$.

0.1 Multiplicative Talent (Roy-1950)

Incomes w are linearly dependent on talent S ;

$$w = aS$$

Talent is the product of attributes, s_i where each attribute is drawn from an *iid* distribution.

$$w = as_1s_2\dots s_n$$

$$\ln w = \ln a + \ln s_1 + \ln s_2\dots + \ln s_n$$

So incomes are log linear. You can also study skewness when the s_i are not independent.

0.1.1 Kalecki (1945)

Idea: Kill random walk or variance exploding: make random return depend on firm size.

Let deviation of firm size, $X_{t+1}^i = R_t^i X_t^i$ where R_t^i is a random variable. In logs

$$\ln X_{t+1}^i = \ln X_t^i + \ln R_t^i = \ln X_0^i + \sum_{j=0}^t \ln R_j^i$$

so $\ln X_{t+1}^i$ is approximately normal. Assume variance of firm size remains constant.

$$\begin{aligned} \frac{1}{n} \sum_i (\ln X_t^i + \ln R_t^i)^2 &= \frac{1}{n} \sum_i (\ln X_t^i)^2 = M \\ 2 \sum_i (\ln X_t^i \ln R_t^i) &= - \sum_i (\ln R_t^i)^2 \end{aligned}$$

and we may assume a negative linear relation

$$\begin{aligned} \ln R_t &= -\alpha \ln X_t + z, \\ \alpha &= \frac{\sum (\ln R_t^i)^2}{2 \sum (\ln X_t^i)^2}, \quad z \text{ iid} \\ \sum_i (\ln X_t^i \ln R_t^i) &= -\alpha \sum_i (\ln X_t^i)^2 \\ &= -\frac{\sum (\ln R_t^i)^2}{2 \sum (\ln X_t^i)^2} \sum_i (\ln X_t^i)^2 \\ \sum_i (\ln X_t^i \ln R_t^i) &= -\frac{1}{2} \sum_i (\ln R_t^i)^2 \end{aligned}$$

Then the limiting distribution is normal:

$$\ln X_{t+1}^i = (1 - \alpha) \ln X_{t+1}^i + z; \quad X_\infty = \alpha^{-1} z$$

Infinite horizon, progressive taxation Small agents max CRRA utility, linear stochastic technology, r is log normal, take tax rate τ as given as a function of time, save fraction $(1 - \lambda_t) = \beta^{\frac{1}{\sigma}} (r(1 - \tau_t))^{\frac{1-\sigma}{\sigma}}$. Take $k_0 > 0$. Taxes are progressive: $\tau_t = 1 - (k_t)^{-a} \geq 0$.

$$\begin{aligned} (1 - \lambda_t) &= \beta^{\frac{1}{\sigma}} (r(1 - \tau_t))^{\frac{1-\sigma}{\sigma}} \\ (1 - \lambda_t) &= \beta^{\frac{1}{\sigma}} \left(r(k_t)^{-a} \right)^{\frac{1-\sigma}{\sigma}} \end{aligned}$$

Note: a has to be appropriately constrained so that $0 < \lambda_t < 1$.

$$\begin{aligned} \tau_t &= 1 - (k_t)^{-a} \\ k_{t+1} &= \beta^{\frac{1}{\sigma}} \left(r \left((k_t)^{-a} \right) \right)^{\frac{1-\sigma}{\sigma}} (1 - \tau_t) r k_t \\ x_{t+1} &= \ln(1 - \tau_t) + \ln \beta^{\frac{1}{\sigma}} \left(r \left((k_t)^{-a} \right) \right)^{\frac{1-\sigma}{\sigma}} r + x_t \\ x_{t+1} &= \ln(k_t)^{-a} + \ln(r\beta)^{\frac{1}{\sigma}} + \ln(k_t)^{-a \left(\frac{1-\sigma}{\sigma} \right)} + x_t \\ x_{t+1} &= \ln(r\beta)^{\frac{1}{\sigma}} + (1 - a)x_t - a \left(\frac{1 - \sigma}{\sigma} \right) x_t \\ x_{t+1} &= \ln(r\beta)^{\frac{1}{\sigma}} + \left(1 - \frac{a}{\sigma} \right) x_t \\ x &= \frac{\ln(r\beta)^{\frac{1}{\sigma}}}{1 - a\sigma^{-1}} \end{aligned}$$

Note: zero taxes, that is $a = 0$, makes x_t a random walk and variance blows up.

0.2 Mincer (1958)

Schooling increases income, but there is an opportunity cost in terms of lost time. Arbitrage implies where s is years of schooling, y is income without schooling, \tilde{y} is income with schooling, V is lifetime income to infinity.

$$V = \frac{y}{1+r} = e^{-rs} \frac{\tilde{y}}{1+r}$$

$$\ln \tilde{y} = \ln y + rs$$

If years of schooling is normally (?) distributed (truncated at zero?) (due to self-selection based on the distribution of ability) around a positive mean, log of incomes are normally distributed.

In many cases, the theoretical modeling of the wealth distribution got very mechanical, and engineering or physics-like in fact, leading Mincer (1958) to plead for explicit microfoundations and more explicit determinants of earnings and wealth distributions:

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

Distribution of profits with uniformly distributed talent-Span of Control x is Talent, w is wage, n is labor (elastic), π is profit. Output is xn^α

$$\text{Max}_x xn^\alpha - wn$$

$$\begin{aligned} \alpha xn^{\alpha-1} - w &= 0 \\ n &= \left(\frac{w}{\alpha x}\right)^{\frac{1}{\alpha-1}} \end{aligned}$$

Profit

$$\begin{aligned} \pi &= x \left(\frac{w}{\alpha x}\right)^{\frac{\alpha}{\alpha-1}} - w \left(\frac{w}{\alpha x}\right)^{\frac{1}{\alpha-1}} \\ \pi &= Ax^{\frac{1}{1-\alpha}}, \quad x = \left(\frac{\pi}{A}\right)^{1-\alpha} \end{aligned}$$

Distribution of talent is uniform

$$f(x) = b, \quad 0 \leq x \leq b^{-1}$$

$$f(x(\pi)) = f(x) \frac{dx}{d\pi} = (bA^{\alpha-1} (1-\alpha)) \pi^{-\alpha}$$

So profits are a power law over $0 \leq \left(\frac{\pi}{A}\right)^{1-\alpha} \leq b^{-1}$, even if talent is uniform.

Lydall, Econometrica 1959, Hierarchies: Here i is level of supervisor, with $i = 1$ the lowest and most numerous, and ratio of people at adjacent levels = n

$$\frac{y_i}{y_{i+1}} = n$$

Wages proportional with factor p to supervisees' wage bill:

$$x_{i+1} = p(nx_i) = (pn)x_i$$

Now

$$\begin{aligned} \lambda_i &= \frac{\ln y_{i+1} - \ln y_i}{(\ln x_{i+1} - \ln x_i)} = \frac{\ln n^{-1}}{\ln np} = -\frac{\ln n}{\ln np}, \text{ a constant} \\ \ln y_{i+1} - \ln y_i &= -\lambda (\ln x_{i+1} - \ln x_i) \\ \ln \left(\frac{y_{i+1}}{y_i} \right) &= \lambda \ln \left(\frac{x_{i+1}}{x_i} \right) \end{aligned}$$

Relation to Pareto:

$$y = bx^{-\alpha-1}$$

$$\begin{aligned} \ln y_i &= \ln b - (\alpha + 1) \ln x_i \\ \ln y_{i+1} - \ln y_i &= -(\alpha + 1) (\ln x_{i+1} - \ln x_i) \end{aligned}$$

$$\begin{aligned} -\alpha &= \lambda + 1 = -\frac{\ln n}{\ln np} + 1 = \frac{-\ln n + \ln np}{\ln np} = \frac{\ln p}{\ln np} \\ \alpha + 1 &= -\lambda = \frac{\ln n}{\ln np}; \quad np = n^{\frac{1}{\alpha+1}} \end{aligned}$$

If n is technologically fixed, p varies inversely with α : higher \rightarrow lower α , more inequality.

Kremer, "The O-Ring Theory of Economic Development" The space shuttle Challenger exploded due to a malfunction in one of its millions of components, the O-rings. Each of the many components of a product or of the many tasks involved in a project must be done right for the whole to function properly. Such a production function may be represented by

$$E[y] = k^a (q_1 q_2 \dots q_n) n B$$

where

- $E[y]$ = expected output
- k = capital
- n = total number of tasks
- B = benefits per worker with one unit of capital if all tasks are performed perfectly
- q_i = probability that worker i performs perfectly

The variable q_i is an index of worker i 's skill level. The important feature about this production function is complementarity between workers' skills. The marginal product of a worker is higher if he works with a higher quality worker. Formally,

$$\partial^2 E[y] / \partial q_i \partial q_j > 0$$

In the space shuttle example, having precise instruments and high-tech equipments does not mean the the shuttle is going to fly if other components (e.g., the O-rings) are defective. It is this complementarity rather than the explicit functional form of the production function that is driving the results of the paper.

A producer chooses $q_1 q_2 \dots q_n$, and k to maximize

$$E[y] - w(q_1) - \dots - w(q_n) - rk$$

where $w(q)$ is the market wage function. The FOCs are

$$\begin{aligned} (\prod_{j \neq i} q_j) n B k^a - w'(q_i) &= 0 \\ (\prod_j q_j) n B a k^{a-1} - r &= 0 \end{aligned}$$

Because of complementarity, a firm with high q_j will be willing to bid more for q_i . In equilibrium, workers of the same skill will be matched together. We let this common skill within a firm be represented by q .

Use the second equation to solve for k , and substitute the result into the first equation. We get

$$n q^{n-1} B (a q^n n B / r)^{a/(1-a)} - w'(q) = 0$$

Another firm that chooses workers with any other value of q must also observe a equation like this. The above equation, which is true for all q , is a differential equation in q . The solution is

$$w(q) = (1 - a)(q^n B)^{1/(1-a)} (an/r)^{a/(1-a)}$$

This model has several interesting implications:

- The wage function $w(q)$ is homogeneous of degree $n/(1 - a)$ in q . So a 1 percent increase in q leads to a $n/(1 - a) > 1$ percent increase in wage. Small differences in q translates into large differences in w .
- Good firms hire good workers and pay high wages. Bad firms hire bad workers and pay low wages. It is an equilibrium situation to have heterogeneous firms hiring workers of different quality and producing products of different quality.
- Good workers are matched to good workers, and bad workers are matched to bad workers in equilibrium. You don't see high-power law firms hiring cheap secretaries because a slight typing mistake might cost the firm millions of dollars. Positive assortative matching explains the observed positive correlation of wages among workers working in the same firm.
- Marginal product is not unambiguously defined for a single worker; it depends on who his co-workers are. Firms do not just offer a wage function and hire whoever comes forward. Instead, they active select workers to ensure good matching.
- The distribution of wages is more skewed than the distribution of ability. Since $w(q)$ is a convex function, w is skewed to the right if q is symmetric. Indeed $\log(w)$ is symmetric.

$$\begin{aligned} w(q) &= Cq^{\frac{n}{1-a}}; \quad q = (C^{-1}w)^{\frac{1-a}{n}} \\ \ln w(q) &= (1 - a)(q^n B)^{1/(1-a)} (an/r)^{a/(1-a)} \\ &= C + n/(1 - a) \ln q \end{aligned}$$

If the distribution of q is uniform: Distribution of talent is uniform

$$\begin{aligned} f(q) &= b, \quad 0 \leq q \leq b^{-1} \\ f(q(w)) \frac{dq}{dw} &= bC^{-1} \left(\frac{1-a}{n} \right) w^{(\frac{1-a}{n}-1)} \end{aligned}$$

where $\frac{1-a}{n} < 1$. So wages are a power law over $0 \leq (C^{-1}w)^{\frac{1-a}{n}} \leq b^{-1}$, even if talent is uniform.

Champernowne (1953, ECMA) Divide incomes into income bins, and assume income bins in geometric partition:

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

$$j = \frac{1}{a} \ln y(j) - \ln y(0) = \frac{1}{a} \ln \frac{y(j)}{y(0)}, \quad \frac{dj}{d(y(j))} = \frac{1}{a} (y(j))^{-1}$$

where $y(0) > 0$ is the lowest bin. Probabilities for moving up a bin is p_1 , down a bin p_{-1} and staying in place is p_0 , except you cannot move down from the lowest bin. The number of people at bin $i = 0, 1, 2, \dots$ at time t , n_t^i is given by

$$n_{t+1}^i = p_1 n_t^{i-1} + p_{-1} n_t^{i+1} + p_0 n_t^i, \quad i \geq 1$$

$$n_{t+1}^0 = p_{-1} n_t^1 + (p_0 + p_{-1}) n_t^0,$$

where the adding up constraint is

$$\sum_i^{\infty} n_t^i = n$$

so n is the total number of people, and where $p_{-1} + p_0 + p_1 = 1$. The second equation above reflects the fact that it is not possible to transition down from $y(0)$.

For a stationary equilibrium the number of people moving away from a bin must be offset by those incoming at each t :

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

This gives a difference equation whose solutions are $n^j = 0, 1$, and $\left(\frac{p_1}{p_{-1}}\right)^j$. The constraint $\sum_i^{\infty} n_t^i = n$ is satisfied by

$$n^j = q \left(\frac{p_1}{p_{-1}}\right)^j$$

only by $n^j = \left(\frac{p_1}{p_{-1}}\right)^j$, if and only if $p_1 < p_{-1}$ and $q > 0$ is appropriately chosen. The requirement $p_1 < p_{-1}$ requires incomes to contract on average, a non-dissipative system with a reflecting barrier at $y(0)$ which we will also see in the Kesten approach in the next section.

Let $\lambda = -\ln\left(\frac{p_1}{p_{-1}}\right) > 0$ so, performing a transformation of variables,

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

which is in Pareto format, with $\sum_0^\infty n^j = n$.

Champernowne also considered a two sided Pareto distribution with two-sided tails, one relating to low incomes and one to high incomes. He eliminated the reflecting barrier, the requirement that it was only possible to go up from the lowest $y(0)$ income bin. Instead, he again assumed an infinite sequence of income bins, but now for the low income bins he modified the assumption of "non-dissipation": the drift was still down for bins above a threshold bin, but for lower bins the average drift was up. This "offset the the continual dispersal of incomes."

0.3 Levy, M. (2003) "Are Rich People Smarter?" JET

The stochastic multiplicative wealth accumulation process is given by

$$W_i(t+1) = \tilde{\lambda}W_i(t)$$

where $W_i(t)$ is the wealth of investor i at time t and $\tilde{\lambda}$ represents the total return, which is a random variable drawn from some distribution $f(\lambda)$.

For people at the high-wealth range, changes in wealth are mainly due to financial investment, and are therefore typically multiplicative. For people at the lower wealth range, changes in wealth are mainly due to labor income and consumption, which are basically additive rather than multiplicative. Here we are only interested in modeling wealth dynamics in the high-wealth range.

There are many ways one could model the boundary between these two regions. We start by considering the most simple model in which there is a sharp boundary between the two regions.

As the stochastic multiplicative process describes the dynamics only at the higher wealth range, we introduce a threshold wealth level, W_0 , above which the dynamics are multiplicative.

$$W_i(t) \geq W_0$$

A natural way to define the lower bound is in terms of the average wealth. We define the lower bound,

$$W_0 = \omega N^{-1} \sum_1^N W_i(t)$$

where N is the number of investors and $\omega < 1$ is a threshold in absolute terms.

When individuals' wealth changes they may cross the boundary between the upper and lower wealth regions. As we do not model the dynamics at the lower wealth range, and we assume that the market has reached an equilibrium in which the flow of people across the boundary is equal in both directions, i.e. the number of people participating at the process remains constant. The above assumption simplifies the analysis, but the results presented here are robust to the relaxation of this assumption.

Master equation:

$$P(W_{t+1}) = P(W_t) + \int_{-\infty}^{\infty} P(W_t/\lambda) f(\lambda) d\lambda - \int_{-\infty}^{\infty} P(W_t) f(\lambda) d\lambda$$

But $\int_{-\infty}^{\infty} P(W_t) f(\lambda) d\lambda = P(W_t)$ so

$$P(W_{t+1}) = \int_{-\infty}^{\infty} P(W_t/\lambda) f(\lambda) d\lambda$$

$P(W) = CW^{-\alpha-1}$ as a stationary distribution:

$$CW^{-\alpha-1} = \int_{-\infty}^{\infty} CW^{-\alpha-1} \lambda^{\alpha+1} f(\lambda) d\lambda$$

where α solves

$$1 = \int_{-\infty}^{\infty} \lambda^{\alpha+1} f(\lambda) d\lambda$$

But note that this is nothing other than the Kesten Theorem for the characterization of fat tails in linear stochastic systems with multiplicative noise, where the additive term serves as a reflecting barrier (see Benhabib, Bisin and Zhu, Econometrica, 2011).

Note here however the problem with the lower bound. In range $[W_0, \lambda W_0)$, $W/\lambda < W_0$ for $\lambda > 1$ in the support of $f(\lambda)$. So if $\lambda_M = \text{Max } \lambda$ in the support of f , Pareto works for $[\lambda_M W_0, \infty)$? What's the economics? Notice that we must have $W_0 > 0$ for the distribution to be well defined. What if f has infinite support? We will have to get back to Double Pareto distributions on the one hand and also to study the full fledged Kesten theorems.

Stochastic Returns and Kesten Accumulation: $w_{t+1} = r_t w_t + y_t$ Here the rhs is after consumption, $c_t = \lambda w_t + q_t$ has been subtracted from capital and labor income, so (r_t, y_t) and accumulation is defined to account for that.¹

Definition 1 *The accumulation equation for wealth, (??), defines a Kesten process if it satisfies the following:*

1. (r_t, y_t) are independent and i.i.d over time; and for any $t \geq 0$:
2. $0 < y_t < \infty$, (but see below)
3. $0 < E(r_t) < 1$,
4. $\text{prob}(r_t > 1) > 0$.

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

Theorem 2 (Kesten) *A Kesten process displays an ergodic stationary distribution which has Pareto tail:*

$$\lim_{w \rightarrow \infty} \text{prob}(w_t \geq w) \sim k w^{-\alpha}, k > 0$$

where $\alpha > 1$ satisfies $E(r_t)^\alpha = 1$.

Two tails: If y_t can take negative values in addition to the right tail you can have a left tail such that $\text{prob}(w_t \leq -w) \sim q w^{-\alpha}$, $q > 0$.

For an nonlinear extension due to Mirek (2011, "Heavy tail phenomenon and convergence to stable laws for iterated Lipschitz maps," Probability Theory and Related Fields, 151, 705-34.) For an application to a Bewley model with borrowing constraints so $w > 0$ that act as a reflecting barrier see Benhabib, Bisin and Zhu (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>

¹Some other regularity conditions for microfoundations are required; see Benhabib, Bisin, Zhu (2011) for details.

Double Pareto, Reed (2001) GBM, Gibrat and Random lifetime.

Another mechanism to obtain more skewedness is to consider the case in which multiplicative accumulation is associated to random lifetime. Reed (Economics Letters, 2001, v. 74, pp. 15-19) has

$$dw = rwdt + \sigma w d\omega$$

with initial state W_0 , a scalar, $w_0 = \log W_0$ and $r \geq 0$ constant. Then the distribution of $\log w_T$ is:

$$\begin{aligned} \log w_T &= N\left(w_0 + \left(r - \frac{\sigma^2}{2}\right)T, \sigma^2 T\right) \\ \log w_T &= \frac{1}{\sigma w \sqrt{2\pi T}} e^{-\left(\frac{(\ln w - (w_0 + (r - \frac{\sigma^2}{2})T))^2}{2\sigma^2 T}\right)} \end{aligned}$$

He then considers the case in which T is exponentially distributed,

$$f_T = p e^{-pT}$$

and computes

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

with solution:

$$f_W = \left\{ \begin{array}{ll} \frac{\alpha\beta}{\alpha+\beta} \left(\frac{W}{W_0}\right)^{\beta-1} & W < W_0 \\ \frac{\alpha\beta}{\alpha+\beta} \left(\frac{W}{W_0}\right)^{-\alpha-1} & W \geq W_0 \end{array} \right\}$$

where $(\alpha, -\beta)$ solve the quadratic:

$$\frac{\sigma^2}{2} z^2 + \left(r - \frac{\sigma^2}{2}\right) z - p = 0$$

Mitzenmacher (Internet Mathematics, vol 1, No. 3, pp. 305-334, 2004, pp.241-

242) implicitly assumes $\log w_0 = 0$ or $W_0 = 1$, $r = \frac{\sigma^2}{2}$, $\sigma = 1$. So for Mitzenmacher the distribution of w is

$$f_w = \int_0^\infty p e^{-pT} \frac{1}{w\sqrt{2\pi T}} e^{-\left(\frac{(\ln w)^2}{2T}\right)} dT$$

with solution²

$$f_W = \begin{cases} \sqrt{\frac{p}{2}} W^{(-\sqrt{2p}-1)} & \text{for } W \geq 1 \\ \sqrt{\frac{p}{2}} W^{(\sqrt{2p}-1)} & \text{for } W \leq 1 \end{cases}$$

We can also specify that the initial condition W_0 is the wealth at which new agents are inserted to replace those who die or are terminated, as a boundary condition. In fact the microfoundations of Benhabib, Bisin and Zhu (Macroeconomic Dynamics, <http://dx.doi.org/10.1017/S1365100514000066>, Published online: 09 April 2014) does this, but it is more general than Reed in some sense because it allows estates and inheritance: not all agents are re-inserted at a fixed treshold, only those whose inheritance falls below the treshold. Since some inheritances do fall below the treshold, we do have a Birth and Death process.

Also, we could have the insertion of agents not at a point W_0 , but at an initial distribution: we could use Reed (2003, The Pareto law of incomes - an explanation and an extension, *Physica A* 319:469-486), who generalizes the initial condition W_0 to allow the initial state to be a log normal distribution instead of W_0 . Similar results, adjusting constants, hold. Reed says:

Suppose that the distribution of starting incomes, $X_0(t)$ say, at time t , is lognormally distributed and also that it evolves as

$$dX_0 = \mu_0 X_0 dt + \sigma_0 X_0 d\omega$$

so that

$$\log(X_0(t)) \sim N\left(a + \left(\mu_0 - \frac{\sigma_0^2}{2}\right)t, b^2 + \sigma_0^2 t\right)$$

where a and b^2 are the mean and variance of $\log(\text{starting income})$ at some initial reference time $t = 0$. With these assumptions it can easily be shown that the current income, $X = X|T$, of a randomly selected individual, who entered the

²Setting $T = u^2$, and remembering to use $\frac{dt}{du} = 2u$ for the change of variables, we get:

$$f_w = \frac{2}{w\sqrt{2\pi}} \int_0^\infty p e^{-pu^2 - \left(\frac{(\ln w)^2}{2u^2}\right)} dT$$

and from integral tables,

$$\int_0^\infty e^{-pu^2 - \frac{(\ln w)^2}{2u^2}} du = \frac{1}{2} \sqrt{\frac{\pi}{p}} e^{-2\sqrt{\frac{1}{2}p(\ln w)^2}}$$

Substituting this yields two different behaviors, depending on $w \geq 1$ or $w \leq 1$ (for the root of $\sqrt{(\ln w)^2}$).

workforce T years ago (at time $T - \tau$) will be log-normally distributed with $\log(X)$ having mean

$$E(\log(X)) = a + (\mu_0 - \frac{\sigma_0^2}{2})(T - \tau) + (\mu - \frac{\sigma^2}{2})(T - \tau) = A_0 + (\tilde{\mu} - \frac{\sigma^2}{2})T$$

and variance

$$V(\log(X)) = b^2 - \sigma_0^2(\tau - T) + \sigma^2 T = B_0^2 T - \tilde{\sigma}^2 T$$

where $A_0 = a + (\mu_0 - \frac{\sigma_0^2}{2})T$ and $B_0^2 = b^2 + \sigma^2 \tau$ are the mean and variance of the log of current starting income (i.e at time $t = \tau$), and

$$\tilde{\mu} = \mu - \mu_0 \quad \text{and} \quad \tilde{\sigma}^2 = \sigma^2 - \sigma_0^2$$

Kesten vs Gibrat in Continuous Time Models where the agents wealth exhibit growth proportional to wealth due to accumulation (with or without partial inheritance) can have stationary distributions with the introduction of Poisson or exponential death rates or termination times, making them into birth and death processes. The underlying mechanism is Gibrat: average growth rates are positive. By contrast in Kesten type models the dynamics are on average contracting, with the possibility of sequences of lucky draws leading to escape for the lucky few. Consider an accumulation process for each agent with wealth w :

$$dw = r(X) w dt + \sigma(X) d\omega$$

where $r(X)$, $\sigma(X) > 0$, and $d\omega$ Brownian motion. Note first, that the additive term, in contrast to the previous section, does not have w^i . We can consider it as labor earnings minus the affine part of consumption, and consider $r(X)$ as the return on wealth net of the part of consumption proportional to wealth, as in some microfounded models (Benhabib, Bisin, Zhu. (2011)). X an exogenous random variable, a finite Markov chain. The usual Kesten assumptions require $E(r(x)) < 0$, and $\Pr(r(X) > 0) > 0$. Under some additional technical assumptions we have, as in the discrete time Kesten models, for $\alpha > 0$:

$$\begin{aligned} \lim_{w \rightarrow \infty} \text{prob}(w_t \geq w) &\sim kw^{-\alpha}, k > 0 \\ \lim_{w \rightarrow \infty} \text{prob}(w_t \leq -w) &\sim kw^{-\alpha}, k > 0 \end{aligned}$$

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

Note that $\lim_{w \rightarrow \infty} \text{prob}(w_t \leq -w) \sim kw^{-\alpha}, k > 0$ also holds because $d\omega$ can take both positive and negative values however.

0.3.1 Neoclassical Models

With standard neoclassical aggregate production functions the marginal product of capital declines with capital, and the rate of return is the same across agents. Unlike models where Gibrat's law takes hold where variances and means of distributions explode over time, stationary distributions, driven simply by heterogeneity of earnings rather than random lifetimes can be established. In addition, introducing borrowing constraints, as in Aiyagari-Bewley models, provide a mechanism to limit unbounded borrowing, and induce a precautionary motive to save in order to insure against sequences of bad income shocks. The precautionary motive however decreases as wealth and is insufficient to generate fat-tailed wealth distributions. One alternative is to introduce highly skewed earnings, but the degree of skewness may not be sufficient to generate the fat tails of income distribution. Consider a simple linear model where an agent's wealth grows as

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

where r_t is common across agents and in a stationary distribution settles to a constant. Wealth heterogeneity reflects the heterogeneity in the histories of y_t . Of course in a microfounded model, even with simple homothetic CRRA preferences, the evolution of r_t may differ across agents as the fraction of wealth consumed can vary with wealth across agents, especially under borrowing constraints, and the same applies to the additive term y_t which represents stochastic earnings minus the affine part of consumption. If we ignore such complications and assume y_t is Gaussian *iid*, the stationary distribution of wealth, once aggregate capital and r_t has converged to its steady state value r , simply scales the distribution of y_t , so measures of inequality, like the Gini coefficient, the coefficient of variation, the tail exponent, or ratios of top to bottom income percentiles will be the same for wealth and earnings distributions.

However more can be said for linear models. If r_t as well as y_t have distributions, independent of w_t , but y_t has tails decaying at $\alpha > 0$, while $E((r_t)^\alpha) < 1$, and $E((r_t)^\beta) < \infty$ for $\beta > \alpha$, then under some regularity assumptions the tail of the stationary distribution of $\{w\}$ will be the same as that of $\{y\}_t$ ³ But that contradicts what we know, that wealth distribution has thicker tails than income.

The Bewley-Aiyagari model which is the workhorse of many DSGE models, with a neoclassical production function, stochastic labor earnings heterogeneous across agents without insurance markets for earnings and borrowing constraints. The borrowing constraints and stochastic incomes introduce a natural non-linearity acting as a reflecting barrier in conjunction with labor earnings. The precautionary motive for savings however is insufficient to generate fat tails

³(see Grey, D.R., 1994. Regular variation in the tail of solutions of random difference equations. *Ann. Appl. Probab.* 4, 169-183., and Ghosh, Arka P. Diana Haya, Vivek Hirpara, Reza Rastegar, Alexander Roitershtein, Ashley Schulteis, Jiyeon Suh. (2010). "Random linear recursions with dependent coefficients", *Statistics and Probability Letters* 80 (2010) 1597-1605.).

in wealth, and for CRRA preferences with the model converging to a steady state in wealth, the consumption and accumulation become asymptotically linear in wealth. If we relax the assumption that capital markets are perfect and rates of return can also differ across agents and agent portfolios or backyard technologies, the Kesten results can be implemented using their generalization due to Mirek, which apply to asymptotically linear models. Fat tails in wealth now obtain due to stochastic earnings as in their linear counterparts: since fat tails are for large wealth levels, asymptotic linearity of accumulation in wealth levels are sufficient if stochastic earnings and returns prevent poverty traps and assure ergodicity and some social mobility.⁴

Of course if we move to non-linear models thick tails can also arise where r_t can increase with wealth, or where the rich may save or bequeath an increasing fraction of their wealth (Cagetti, M. and M. De Nardi (2006), "Entrepreneurship, Frictions, and Wealth", *Journal of Political Economy*, 114, 835-870., and Atkinson, A. B. "Capital Taxes, the Redistribution of Wealth and Individual Savings", *The Review of Economic Studies*, Vol. 38, No.2. (Apr., 1971), pp. 209-227.). Alternatively the mean rate of return can increase in wealth, perpetuating the wealth of the rich. Such non-linearities also contribute to generating fat tails, but if their effect is too dominant, they will result in restricting the downward mobility of the rich across decades, while we observe some downward mobility. Borrowing constraints in Aiyagari-Bewley models on the other hand can create poverty traps and prevent the wealth-poor from accumulating wealth unless positive earning shocks allows them to escape the poverty trap and start accumulating. (See Benhabib, Bisin and Zhu (2016).) A combination of stochastic earnings, stochastic returns and non-linearities together may be considered to jointly explain the distribution of wealth and mobility.

⁴See Benhabib, Bisin and Zhu (2016) for a full treatment of the Aiyagari-Bewley model that also allows for heterogenous rates of return across agents. The consumption function is asymptotically linear and the results from the linear case apply for the right tail with high wealth.