ECON0057 Lecture 6

Income and Wealth Distribution

Alan Olivi

UCL

2024

Introduction

- In previous lectures, models that approximately replicated income and wealth distribution
- In these models top wealth share driven mostly by top income share: we need an absurdly high income state (top 1% earnings 1000 times median earning, at least an order of magnitude too large)
- Need other mechanisms to generate realistic tails of wealth distribution (random growth)
- Important to understand impact of fiscal policies

Table of Contents

Distribution of Income

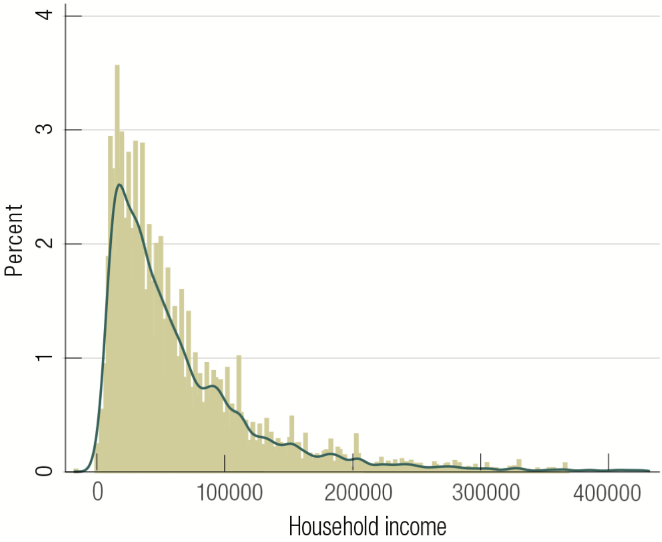
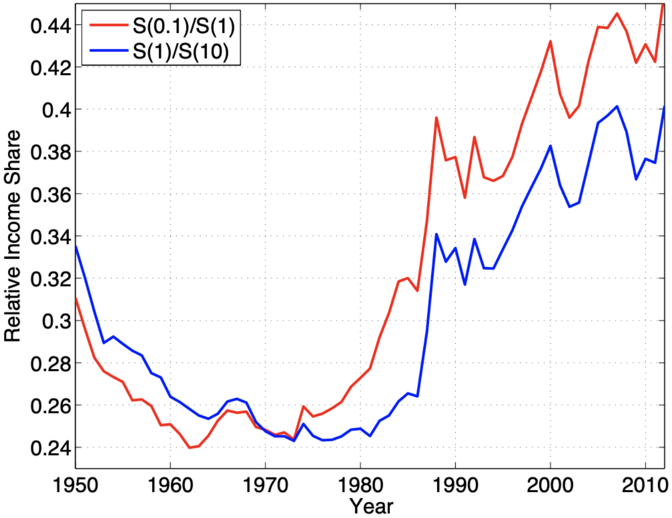


Figure 6. **Histogram of the 2013 income distribution (2013 USD).**

Top Income Shares



Upper Tails: Pareto Distributions

- Pareto (1896): top income/wealth shares well described by **power law** distributions

  $\rightarrow$ number of people with an income (or wealth) X greater than x is:

$$P(X > x) = cx^{-\alpha}$$

- Tail goes to 0 more slowly (as a power law) than normal or log normal distributions

  $\rightarrow$ Fat tailed distribution

- The lower α the fatter the tail ($\alpha > 1$ for the mean to be well defined)

- Often we only care about the **asymptotic behavior of the tail**

  $\rightarrow$ X is an asymptotic power law if

$$P(X > x) \sim cx^{-\alpha} \quad as \quad x \rightarrow \infty$$

  (meaning that $\lim_{x\rightarrow\infty} P(X > x)/cx^{-\alpha} = 1$)

Power Laws and Fractal Inequality

- Power laws describe **fractal inequality**:

  $\rightarrow$ if top 1% n times richer than top 10% then top 0.1% n times richer than top 1%

- To see this, let's compute the average income above some level x^*:

$$\bar{x} = \mathbb{E}(x \mid x > x^*) = \frac{\int_{x^*}^{\infty} x \alpha c x^{-\alpha-1} dx}{c(x^*)^{-\alpha}} = \frac{\alpha}{\alpha - 1} x^*$$

  Define x_p the income of the top p% we have

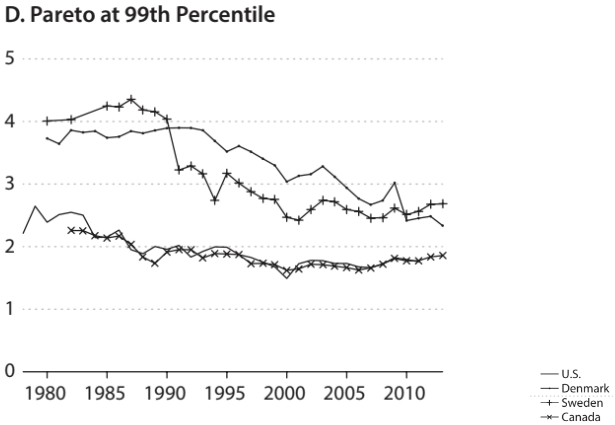
$$x_p = \left(\frac{p/100}{c}\right)^{-1/\alpha} \quad so \quad \frac{\bar{x}_{0.01}}{\bar{x}_{0.1}} = \frac{\bar{x}_{0.1}}{\bar{x}_1} = \frac{\bar{x}_1}{\bar{x}_{10}} = 10^{1/\alpha}$$

- In the same way defining $S(p)$ the share of the top p% we have

$$S(p) = \left(\frac{p}{100}\right)^{1-1/\alpha} \quad so \quad \frac{S(0.01)}{S(0.1)} = \frac{S(0.1)}{S(1)} = \frac{S(1)}{S(10)} = 0.1^{1-1/\alpha}$$

  e.g. for $\alpha = 2$ the top 1% share is $S(0.1) = 10\%$

- Captures key property of top incomes, how do we check that more precisely?
  - Regress $\mathbb{E}(x \mid x > x^*)$ on x^* gives the coeff $\alpha/\alpha - 1$
  - Or regress $ln(p)$ on $ln(x_p)$, since $ln(p) = -\alpha ln(x_p) + ln(c))$



Gives an α around 2 for US

Tails of Wealth distribution

A bit Noisier

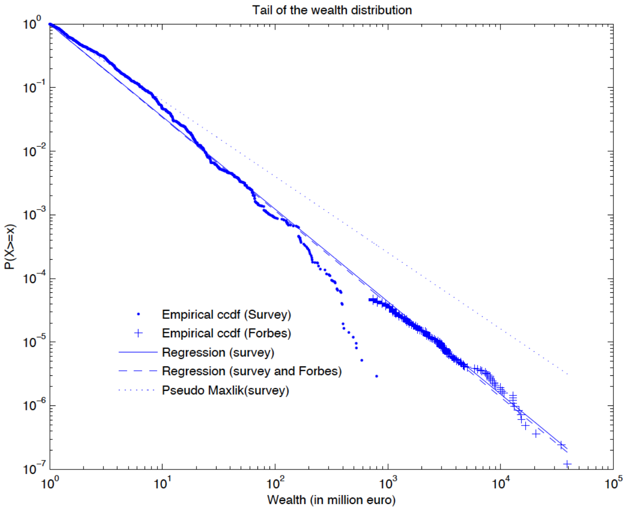


Figure 3: *Tail of the wealth distribution: USA*

TABLE A2

Estimates of Pareto tail index

thresholds: 3, 5 and 10 million euro

	Pseudo max.likelihood			Regression method					
				excluding Forbes			including Forbes		
	$\geq$ 10M	$\geq$5M	$\geq$3M	$\geq$ 10M	$\geq$5M	$\geq$3M	$\geq$ 10M	$\geq$5M	$\geq$3M
USA	1.63	1.56	1.53	1.85	1.80	1.74	1.51	1.54	1.55
	0.14	0.08	0.06	0.09	0.07	0.05	0.01	0.01	0.01
France	1.47	1.43	1.76	1.47	1.61	1.60	1.36	1.42	1.45
	0.38	0.21	0.18	0.30	0.26	0.18	0.05	0.03	0.02
UK	1.47	2.19	2.06	1.79	1.94	2.08	1.41	1.49	1.58
	-	-	-	-	-	-	-	-	-
Spain	1.55	1.46	1.68	1.80	1.77	1.68	1.53	1.58	1.58
	0.20	0.17	0.31	0.37	0.22	0.15	0.08	0.07	0.05

Notes: Mean estimate using all five implicates. Standard errors below mean estimate.

Table of Contents

- To match the data, a good theory of income and wealth distribution should:
  1. Generate a stationary distribution of wealth and income following a power law (asymptotically)
  2. The wealth distribution should have fatter tails than the income distribution
- Under relatively weak assumptions, **random growth processes** can generate pareto tails
  - random growth processes are ubiquitous
  - used to explain size of cities, size of firms etc.
  - see Gabaix (2009) for an intro.

Kesten's Theorem (1973)

Consider the process the *reduced form* wealth accumulation process:

$$W_{t+1} = R_t W_t + Y_t$$

with $R_t > 0$, R_t, Y_t iid.

Under which conditions on R_t, Y_t can the stationary distribution of W_t have fatter tail than Y_t?

Result: Suppose $\mathbb{E}(R^\alpha) = 1$, $\mathbb{E}(|Y|^\alpha) < \infty$ with $\alpha > 0$ and a unique non-degenerate stationary distribution W exists, then W follows an asymptotic power law with coefficient α

Intuition: Suppose that Y is constant. The stationary distribution satisfies $W = RW + Y$ so:

$$P(W > w) = \int P(W > \frac{w - Y}{r})\pi(r)dr$$

Plugging $P(W > w) = cw^{-\alpha}$ we get:

$$cw^{-\alpha} \approx \int c(w - Y)^{-\alpha} r^\alpha \pi(r)dr \approx c(w - Y)^{-\alpha}\mathbb{E}(r^\alpha) \approx cw^{-\alpha}\mathbb{E}(r^\alpha)$$

for $w \gg Y$. So if $\mathbb{E}(r^\alpha) = 1$ the equation is approximately satisfied and the Pareto coefficient of W is α.

Kesten: Technical Discussion (Optional)

- The formal theorem by Kesten requires additional assumptions:
  - The distribution of $Y/(R-1)$ is non degenerate
  - R is a non lattice distribution (R does not take discrete values $r_i = a + b\,i$)
  - $\mathbb{E}(R^{\alpha} log^{+} R) < \infty$
- Those assumptions ensures that unique non-degenerate stationary distribution W exists
- For example if R and Y are constant, the stationary distribution would be degenerate equal to $Y/(R-1)$

  $\rightarrow$ This is forbidden by the first assumption
- If Y is stochastic and $R = 1$, W_t would be a random walk (no stationary distribution)

  $\rightarrow$ This is forbidden by the first assumption
- The condition $\mathbb{E}(R^{\alpha} log^{+} R) < \infty$ is more complex, it ensures that W_t does not grow too fast

Kesten: Discussion

Result: Suppose $\mathbb{E}(R^\alpha) = 1$, $\mathbb{E}(|Y|^\alpha) < \infty$ with $\alpha > 0$ and a unique non-degenerate stationary distribution W exists, then W follows an asymptotic power law with coefficient α

- If Y is a power law with coefficient $\beta > \alpha$, then $\mathbb{E}(|Y|^\alpha) = \beta/(\beta - \alpha) < \infty$
- Inversely $\mathbb{E}(|Y|^\alpha) < \infty$ implies that the upper tail of Y is thinner than α
- Therefore, if $\mathbb{E}(R^\alpha) = 1$ and $\mathbb{E}(|Y|^\alpha) < \infty$, the tail of the wealth distribution is fatter than the tail of the income distribution
- Note that the tail of W does not depend on Y: stronger income inequality does not lead to stronger wealth inequality (in the tail)
- With Kesten processes, households reach the top of the distribution thanks to high return investments (high R) not because of good income shocks
- Can standard HA model generate fatter tail for the wealth distribution?

In standard HA models R is not stochastic what can go wrong?

Result: If $\mathbb{E}(R^{\beta}) < 1$, (and 1.there is $\gamma > \beta$ such that $\mathbb{E}(R^{\gamma}) < \infty$, 2.$\mathbb{E}(log^{+}|Y|) < \infty$), then:

$$P(Y > y) \sim cy^{-\beta} \Leftrightarrow P(W > w) \sim \tilde{c}w^{-\beta}$$

To see what this result says, suppose that $0 < R < 1$, in that case we cannot have $\mathbb{E}(R^{\alpha})$ with $\alpha > 0$ so the only way that W can be pareto is if Y is itself pareto, in that case the tail of W is at most as fat as the tail of Y.

Kesten and Grey delineate 2 cases: either there's enough dispersion in R and in that case the growth of W is dominated by the stochastic growth component RW and 1) the tails of W can be fatter than the tails of Y 2) Y does not matter much for the tails behavior of W, or R does not have enough dispersion and in that case the growth of W is entirely determined by Y at the tails in particular the tail of W cannot be fatter the tail of Y.

Grey and Bewley

Given the result of Kesten and Grey, if we have a simple wealth accumulation equation $W_{t+1} = R_t W_t + Y_t$, we need $\mathbb{E}(R^\alpha) = 1$ for $\alpha < \beta$ where β is the pareto coefficient of the income distribution to generate realistic inequalities

How is it related to Bewley models? In the standard HA model, where the interest rate is constant, it can be shown that when income is iid, we have $\lim_\infty c(w)/w = \psi < R$ (see Benhabib, Bisin and Zhu) so we have approximately $w_{t+1} = R_t w_t + y_t - c_t = (R_t - \psi) w_t + y_t - \chi_t$: we're exactly in Grey's case ($0 < R - \psi < 1$) and the model cannot generate fatter tail for wealth than income.

The result was extended by Stachurski & Toda (2019), for any Markov process for income: it's not possible in general to generate realistic distributions in standard HA models.

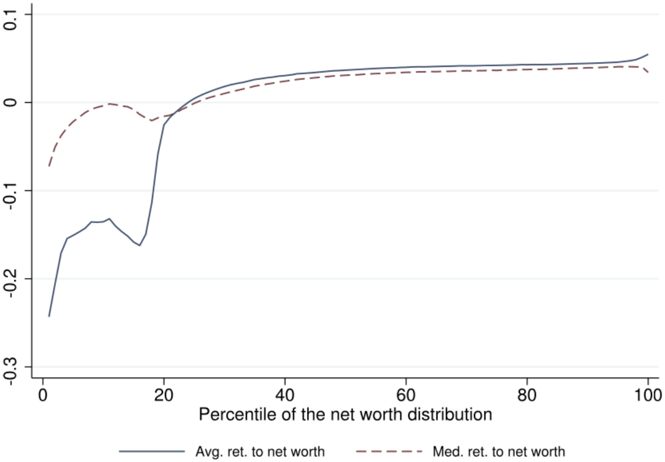
It's also worth noting that the mathematical result (Kesten) can be a bit restrictive for economic applications (iid processes) but were extended by Saporta (2005) an Roitherstein (2007) to cover some Markov processes

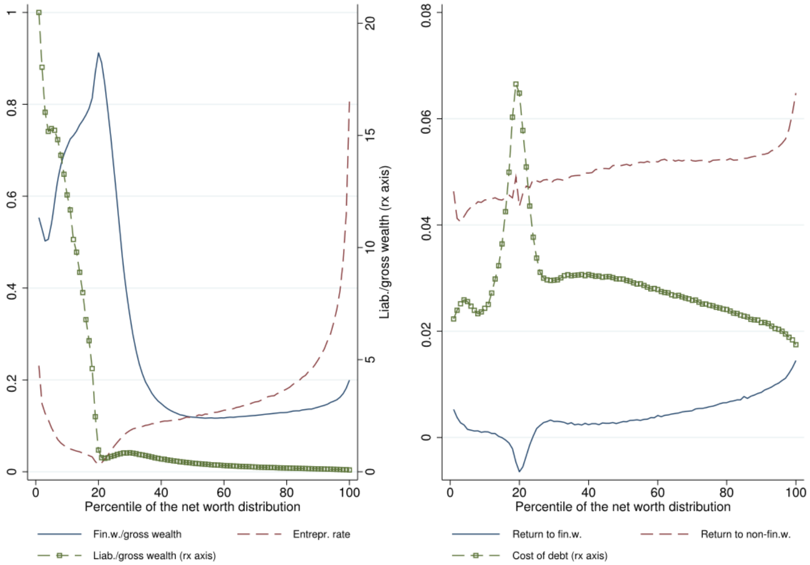
Table of Contents

Evidence of Idiosyncratic Returns: Fagereng, Guiso, Malacrino & Pistaferri

- Fagereng et al. (2020): Norwegian population tax record data from 1993 to 2013
- Norwegian residents pay a wealth tax (third-party report), tax records include:
  - Capital income
  - Detailed information on asset holdings
- Findings: massive returns heterogeneity
  - Returns strongly correlated with wealth
  - Returns persistent across life cycle and across generations

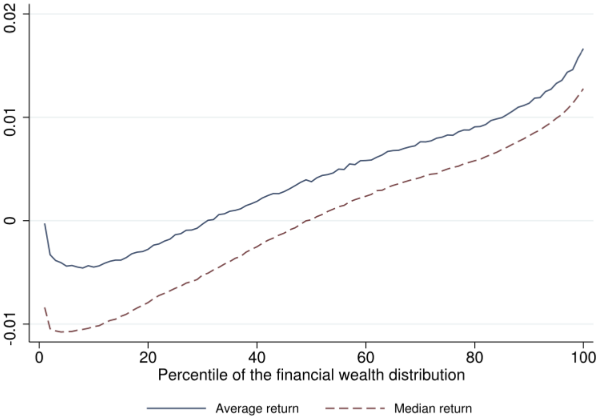
Returns on Net Worth



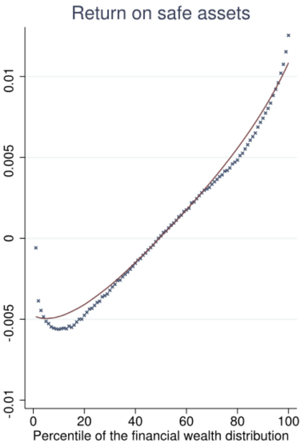


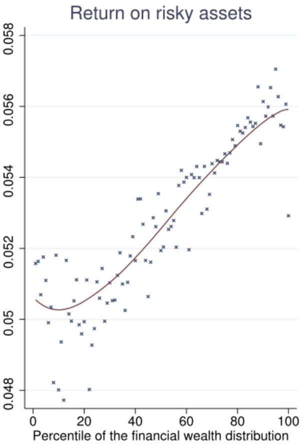
1
0.8
0.6
0.4
0.2
0
20
15
10
5
0
Liab./gross wealth (rx axis)
0 20 40 60 80 100
Percentile of the net worth distribution
Fin.w./gross wealth
Entrepr. rate
Liab./gross wealth (rx axis)
0.08
0.06
0.04
0.02
0
0 20 40 60 80 100
Percentile of the net worth distribution
Return to fin.w.
Return to non-fin.w.
Cost of debt (rx axis)

Returns on Financial Wealth

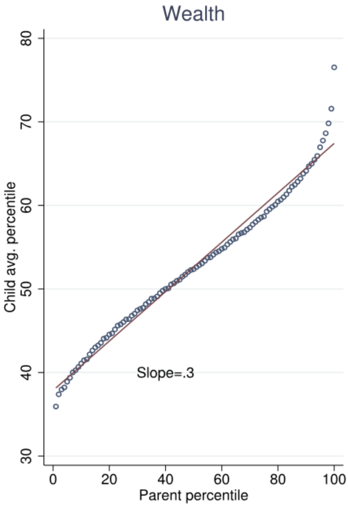


Returns on Financial Wealth by Asset Class





Intergenerational Mobility



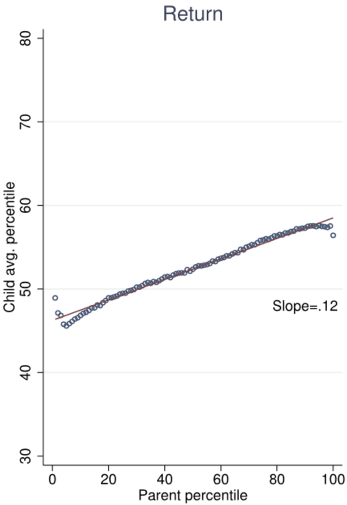


Table of Contents

Goal of the paper: provide a simple model of intergenerational wealth accumulation. The n^th household in a dynasty has a return to wealth r_n an income profile $y_n(t)$ and solves:

$$\sup \sum_0^T \beta^t \frac{c(t)^{1-\sigma}}{1-\sigma} + \beta^T A \frac{a_n(T)^{1-\mu}}{1-\mu}$$

$$s.t. \quad a_n(t+1) = (1+r_n)a_n(t) + w_n(t) - c(t), \quad a_n(0) = a_{n-1}(T)$$

The returns r_n and wage profiles w_n are Markov (correlated across generations), they provide the stochastic growth that we saw in Kesten.

μ determines the strength of the bequest motive, when $\mu = \sigma$ as we'll see bequests are linear in initial wealth. When $\mu < \sigma$ bequest becomes a luxury "good" and is convex in initial wealth: richer households save more to donate more to their children

Simple case $\mu = \sigma$

- Euler equation $c^{-\sigma}(t) = \beta(1+r_n)c^{-\sigma}(t+1)$
- Terminal condition $c(T)^{-\sigma} = Aa(T)^{-\mu}$
- NPV budget constraint:

$$\frac{a(T)}{(1+r_n)^T} - a(0) = \sum_0^T \frac{c(t)}{(1+r_n)^t} + \frac{y(t)}{(1+r_n)^t}$$

Using $c(t) = (\beta(1+r_n))^{-\frac{T-t}{\sigma}} A^{\frac{1}{\sigma}} a(T)^{\frac{\mu}{\sigma}}$ we can directly see that when $\mu = \sigma$, $c(t)$ is a linear function of $a(T)$ so:

$$a_n(T) = a_{n+1}(0) = \alpha(r_n, w_n)a_n(0) + \beta(r_n, w_n)$$

Intergenerational wealth follows a stochastic growth equation.

When $\mu < \sigma$, c is concave in $a(T)$ we have non linear stochastic growth: $a_{n+1}(0) = g(a_n(0), r_n, w_n)$ where g is convex in $a_n(0)$

Calibration Parameters

Main parameters intergenerational transmission of rates r_n and bequest motive A and μ

TABLE 4—PARAMETER ESTIMATES: BASELINE

	Preferences				
	σ	μ	A	β	T
	[2]	0.5993 (0.0061)	0.0006 (0.0004)	[0.97]	[36]
	Rate of return process				
State space	0.0011 (0.0069)	0.0094 (0.0118)	0.0258 (0.0004)	0.0560 (0.0059)	0.0841 (0.0043)
Transition diagonal	0.0338 (0.6162)	0.2676 (0.5570)	0.1360 (0.0699)	0.2630 (1.3659)	0.0208 (0.2678)
Statistics	$E(r)$ 3.06% (0.02%)	$\sigma(r)$ 2.69% (0.01%)	$\rho(r)$ 0.103 (0.486)		

Note: Standard errors in parentheses; fixed parameters in brackets.

TABLE 5—MODEL FIT: BASELINE

	Wealth distribution							
Percentile	0–20	20–40	40–60	60–80	80–90	90–95	95–99	99–100
Wealth share (data)	−0.002	0.001	0.045	0.112	0.120	0.111	0.267	0.336
Wealth share (model)	0.049	0.077	0.111	0.110	0.110	0.076	0.142	0.325

	Social mobility				
Percentile	0–20	20–40	40–60	60–80	80–100
Transition diagonal (data)	0.36	0.24	0.25	0.26	0.36
Transition diagonal (model)	0.349	0.197	0.201	0.210	0.340

Fit wealth share and social mobility without counterfactual awesome states!

Comparison to FGMP

TABLE 7—RATE OF RETURN PROCESS

Statistics	$E(r)$	$\sigma(r)$	$\rho(r)$
Model estimates	3.06%	2.69%	0.103
Fagereng et al. (2017)	2.98%	2.82%	0.1

Note: Fagereng et al.'s (2017) permanent component has zero-mean by construction: we report their mean of returns.

Reasonable stochastic process for r_n

Why Does it Matter

Recall that impact of estate tax (and capital tax in KM) had relatively little direct impact on wealth distribution in CDR. This because the wealth distribution was mostly determined by the income distribution and the "awesome" states.

In this type of model (when $\mu = \sigma$) can simply rederive the growth equation as $a_{n+1}(0) = \alpha(r_n, w_n, 1-\tau)a_n(0) + \beta(r_n, w_n, 1-\tau)$ where $1-\tau$ is the estate tax. Can directly show that $\alpha(r_n, w_n, 1-\tau)$ is decreasing in τ.

When r_n is iid, the tails of the wealth distribution is directly determined by $\mathbb{E}(\alpha(r_n, w_n, 1-\tau)^\kappa) = 1$ (when r_n is markov we need $\lim \mathbb{E}(\alpha(r_0, w_0, 1-\tau)^\kappa \times ... \times \alpha(r_n, w_n, 1-\tau)^\kappa)^{\frac{1}{n+1}} = 1$ but the logic is the same)

Since $\mathbb{E}(\alpha(r_n, w_n, 1-\tau)^k)$ is log convex in k (moment generating function) and $\mathbb{E}(\alpha(r_n, w_n, 1-\tau)^0) = 1$, this means that $\partial_k \mathbb{E}(\alpha(r_n, w_n, 1-\tau)^k)|_{k=\kappa} > 0$ so if τ increases and $\alpha(r_n)$ decreases, κ has to increase to satisfy Kesten: the tails of the wealth distribution become thinner and the estate tax (same for capital tax) has a first order impact on wealth inequality. When wealth inequality determined by stochastic growth rather than income, taxes can have a large impact on top wealth!