The Dynamics of Inequality

Xavier Gabaix
NYU Stern

Jean-Michel Lasry
Dauphine

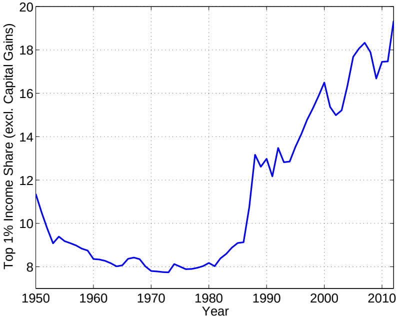
Pierre-Louis Lions
Collège de France

Benjamin Moll
Princeton

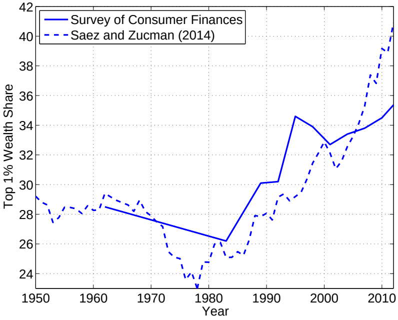
NBER Summer Institute
Working Group on Income Distribution and Macroeconomics

July 15, 2015

Question



(a) Top Income Inequality



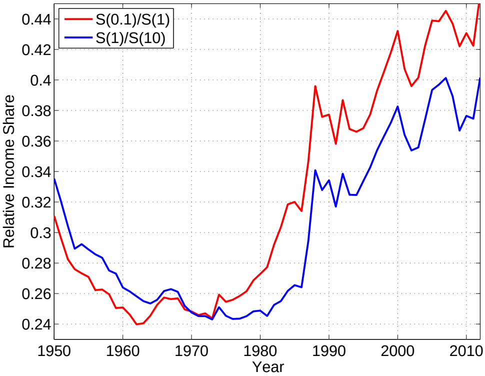
(b) Top Wealth Inequality

- In U.S. past 40 years have seen (Piketty, Saez, Zucman & coauthors)
  - rapid rise in top income inequality
  - rise in top wealth inequality (rapid? gradual?)
- **Why?**

Question

- **Main fact** about **top inequality** (since Pareto, 1896): upper tails of income and wealth distribution follow **power laws**
- Equivalently, top inequality is **fractal**
  1. ... top 0.01% are X times richer than top 0.1%,... are X times richer than top 1%,... are X times richer than top 10%,...
  2. ... top 0.01% share is fraction Y of 0.1% share,... is fraction Y of 1% share, ... is fraction Y of 10% share,...

Evolution of “Fractal Inequality”



- $\frac{S(p/10)}{S(p)}$ = fraction of top p% share going to top (p/10)%
  - e.g. $\frac{S(0.1)}{S(1)}$ = fraction of top 1% share going to top 0.1%
- Paper: same exercise for wealth

This Paper

- **Starting point:** existing theories that explain top inequality **at point in time**
  - differ in terms of underlying economics
  - but share basic mechanism for generating power laws: **random growth**

- **Our ultimate question:** which specific economic theories can also explain observed **dynamics** of top inequality?
  - income: e.g. falling income taxes? superstar effects?
  - wealth: e.g. falling capital taxes (rise in after-tax $r - g$)?

- **What we do:**
  - study **transition dynamics** of cross-sectional distribution of income/wealth in theories with random growth mechanism
  - contrast with data, **rule out** some theories, **rule in** others

Main Results

- Transition dynamics of standard random growth models **too slow** relative to those observed in the data
  - analytic formula for speed of convergence
  - transitions particularly slow in **upper tail** of distribution
  - jumps cannot generate fast transitions either
- Two parsimonious deviations that generate **fast transitions**
  1. heterogeneity in mean growth rates
  2. "superstar shocks" to skill prices
- Both only consistent with particular economic theories
- Rise in top **income** inequality due to
  - ~~simple tax stories, stories about Var(permanent earnings)~~
  - **rise of "superstar" entrepreneurs or managers**
- Rise in top **wealth** inequality due to
  - ~~increase in $r - g$ due to falling capital taxes~~
  - **rise in saving rates/RoRs of super wealthy**

Literature: Inequality and Random Growth

- **Income distribution**
  - Champernowne (1953), Simon (1955), Mandelbrot (1961), Nirei (2009), Toda (2012), Kim (2013), Jones and Kim (2013), Aoki and Nirei (2014),...
- **Wealth distribution**
  - Wold and Whittle (1957), Stiglitz (1969), Cowell (1998), Nirei and Souma (2007), Benhabib, Bisin, Zhu (2012, 2014), Piketty and Zucman (2014), Piketty and Saez (2014), Piketty (2015)
- **Dynamics** of income and wealth distribution
  - Blinder (1973), but no Pareto tail
  - Aoki and Nirei (2014)
- **Power laws are everywhere** ⇒ results useful there as well
  - firm size distribution (e.g. Luttmer, 2007)
  - city size distribution (e.g. Gabaix, 1999)
  - ...

Plan

1. **Random growth theories** of top inequality
   - a simple theory of top income inequality
   - stationary distribution
2. **Slow transitions** in the baseline model
3. Models that generate **fast transitions**
   - heterogeneous mean growth rates
   - "superstar shocks" to skill prices

- **Today's presentation:** focus on top income inequality
- **Paper:** analogous results for top wealth inequality

A Random Growth Theory of Income Dynamics

- Continuous time
- Continuum of workers, heterogeneous in human capital h_{it}
- die/retire at rate δ, replaced by young worker with h_{i0}
- Wage is $w_{it} = \omega h_{it}$
- Human capital accumulation involves
  - investment
  - luck
- "Right" assumptions $\Rightarrow$ wages evolve as

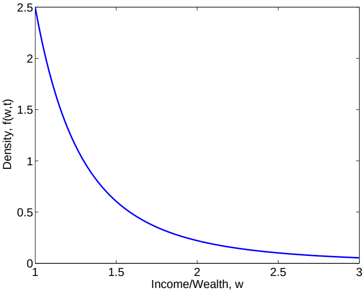
$$d \log w_{it} = \mu dt + \sigma dZ_{it}$$

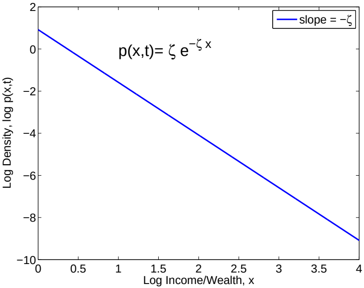
  - **growth rate** of wage w_{it} is **stochastic**
  - μ, σ depend on model parameters
  - Z_{it} = Brownian motion, i.e. $dZ_{it} \equiv \lim_{\Delta t \to 0} \varepsilon_{it} \sqrt{\Delta t}, \varepsilon_{it} \sim \mathcal{N}(0, 1)$
- A number of alternative theories lead to same reduced form

Stationary Income Distribution

- **Result:** The stationary income distribution has a Pareto tail

$$\Pr(\tilde{w} > w) \sim C w^{-\zeta}$$





- ... with tail exponent

$$\zeta = \frac{-\mu + \sqrt{\mu^2 + 2\sigma^2\delta}}{\sigma^2}$$

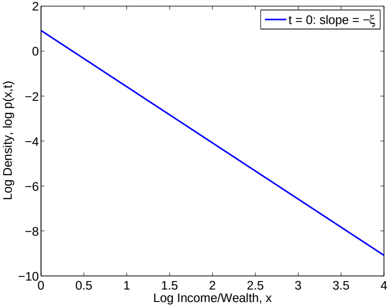
- Tail inequality $\eta = 1/\zeta$ increasing in μ, σ, decreasing in δ

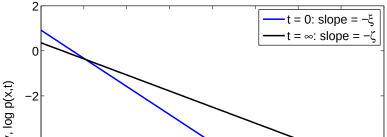
Other Theories of Top Inequality

- We confine ourselves to theories that generate power laws
  - random growth
  - models with superstars (assignment models) – more later
- Example of theories that do not generate power laws, i.e. do not generate **fractal feature** of top income inequality:
  - theories of rent-seeking (Benabou and Tirole, 2015; Piketty, Saez and Stantcheva, 2014)
  - someone should write that "rent-seeking $\Rightarrow$ power law" paper!

Transitions: The Thought Experiment

- $\sigma \uparrow$ leads to increase in stationary tail inequality
- But what about dynamics? Thought experiment:
  - suppose economy is in Pareto steady state
  - at $t = 0$, $\sigma \uparrow$. Know: in long-run $\rightarrow$ higher top inequality





Transitions: Tools

- Convenient to work with $x_{it} = \log w_{it}$

$$dx_{it} = \mu dt + \sigma dZ_{it}$$

- Need additional "friction" to ensure existence of stat. dist.
  - income application: death/retirement at rate δ
  - alternative: reflecting barrier
- Distribution $p(x, t)$ satisfies

$$p_t = -\mu p_x + \frac{\sigma^2}{2} p_{xx} - \delta p + \delta \delta_0$$

where $\delta_0 =$ Dirac delta function (point mass at $x = 0$)

- Useful to write in terms of differential operator $\mathcal{A}^*$

$$p_t = \mathcal{A}^* p + \delta \delta_0, \quad \mathcal{A}^* p = -\mu p_x + \frac{\sigma^2}{2} p_{xx} - \delta p$$

- $\mathcal{A}^* =$ "transition matrix" for continuous-state process

Average Speed of Convergence

- **Proposition:** $p(x,t)$ converges to stationary distrib. $p_\infty(x)$

$$||p(x,t) - p_\infty(x)|| \sim ke^{-\lambda t}$$

  - without reflecting barrier, rate of convergence is

$$\lambda = \delta$$

  - with reflecting barrier, rate of convergence is

$$\lambda = \frac{1}{2}\frac{\mu^2}{\sigma^2}\mathbf{1}_{\{\mu<0\}} + \delta$$

- For given amount of top inequality η, speed $\lambda(\eta,\sigma,\delta)$ satisfies

$$\frac{\partial\lambda}{\partial\eta} \leq 0, \quad \frac{\partial\lambda}{\partial\sigma} \geq 0, \quad \frac{\partial\lambda}{\partial\delta} > 0$$

- **Observations:**
  - **high inequality** goes hand in hand with **slow transitions**
  - half life is $t_{1/2} = \ln(2)/\lambda \Rightarrow$ precise quantitative predictions
- **Rough idea:** λ = 2nd eigenvalue of "transition matrix" $\mathcal{A}^*$

Rough Idea of Proof ▸ back

- To understand, suppose x_{it} = finite-state Poisson process
  - $x_{it} \in \{x_1, ..., x_N\} \Rightarrow$ distribution = vector $p(t) \in \mathbb{R}^N$
  - dynamics
    $$\dot{p}(t) = \mathbf{A}^T p(t),$$
    where $\mathbf{A}$ = $N \times N$ (diagonolizable) transition matrix
  - Denote eigenvalues by $0 = |\lambda_1| < |\lambda_2| < ... < |\lambda_N|$ and corresponding eigenvectors by $(v_1, ..., v_N)$
  - **Theorem:** $p(t)$ converges to stationary dist. at rate $|\lambda_2|$
  - Proof sketch: decomposition
    $$p(0) = \sum_{i=1}^{N} c_i v_i \quad \Rightarrow \quad p(t) = \sum_{i=1}^{N} c_i e^{\lambda_i t} v_i$$
  - Example: symmetric two-state Poisson process with intensity ϕ
    $$\mathbf{A} = \begin{bmatrix} -\phi & \phi \\ \phi & -\phi \end{bmatrix}, \quad \Rightarrow \quad \lambda_1 = 0, \quad |\lambda_2| = 2\phi$$
    Intuitively, speed $|\lambda_2| \nearrow$ in switching intensity ϕ

Rough Idea of Proof

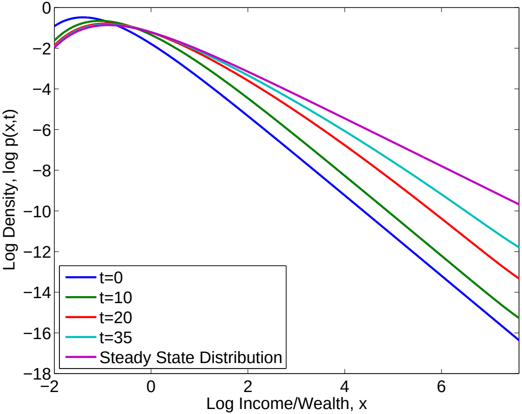
- Here: generalize this idea to continuous-state process
- Consider Kolmogorov Forward equation for x_{it}-process

$$p_t = \mathcal{A}^* p + \delta\delta_0, \quad \mathcal{A}^* p = -\mu p_x + \frac{\sigma^2}{2} p_{xx} - \delta p$$

- Exact generalization of finite-state $\dot{p}(t) = \mathbf{A}^T p(t)$
- **Proof** has two steps:
  1. realization that speed = second eigenvalue of operator $\mathcal{A}^*$
  2. analytic computation: $|\lambda_2| = \frac{1}{2}\frac{\mu^2}{\sigma^2}\mathbf{1}_{\{\mu<0\}} + \delta$

Transition in Upper Tail

- So far: **average** speed of convergence of whole distribution
- But care in particular about speed in **upper tail**
- Show: transition can be much **slower** in upper tail



Instructive Special Case: Steindl Model

- The special case $\sigma = 0, \mu > 0$ can be solved cleanly
  - x_t grows at rate μ, gets reset to $x_0 = 0$ at rate δ
  - stationary distribution $p(x) = \zeta e^{-\zeta x}, \zeta = \delta/\mu$
- Can show: for $t, x > 0$ density satisfies

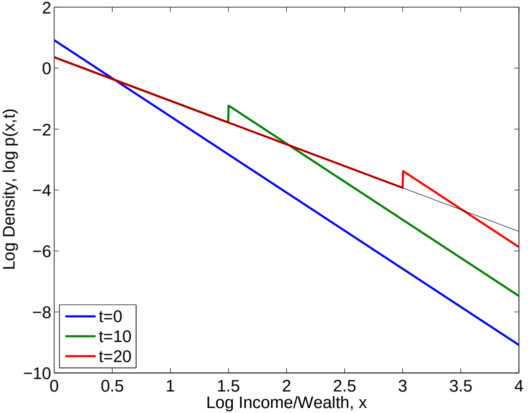
$$\frac{\partial p(x,t)}{\partial t} = -\mu \frac{\partial p(x,t)}{\partial x} - \delta p(x,t), \quad p(x,0) = \alpha e^{-\alpha x} \quad (*)$$

- **Result:** the solution to $(*)$ is

$$p(x,t) = \zeta e^{-\zeta x} \mathbf{1}_{\{x \le \mu t\}} + \alpha e^{-\alpha x + (\alpha - \zeta) t} \mathbf{1}_{\{x > \mu t\}}$$

where $\mathbf{1}_{\{\cdot\}}$ = indicator function

Instructive Special Case: Steindl Model



Observations:

1. transition is slower in upper tail: it takes time $\tau(x) = x/\mu$ for the local PL exponent to converge to its steady state value ζ
2. initially, tail exhibits parallel shift

Transition in Tail: General Case

- Distribution $p(x, t)$ satisfies a Kolomogorov Forward Equation

$$p_t = -\mu p_x + \frac{\sigma^2}{2} p_{xx} - \delta p + \delta \delta_0 \qquad (*)$$

- Can solve this, but not particularly instructive
- Instead, use so-called "Laplace transform" of p

$$\widehat{p}\,(\xi, t) := \int_{-\infty}^{\infty} e^{-\xi x} p\,(x, t)\, dx = \mathbb{E}\left[e^{-\xi x}\right]$$

- $\widehat{p}$ has natural interpretation: $-\xi$th moment of income/wealth $w_{it} = e^{x_{it}}$
  - e.g. $\widehat{p}(-2, t) = \mathbb{E}[w_{it}^2]$

Transition in Upper Tail

- **Proposition:** The Laplace transform of p, $\widehat{p}$ satisfies

$$\widehat{p}(\xi, t) = \widehat{p}_\infty(\xi) + (\widehat{p}_0(\xi) - \widehat{p}_\infty(\xi))\, e^{-\lambda(\xi)t}$$

  with moment-specific speed of convergence

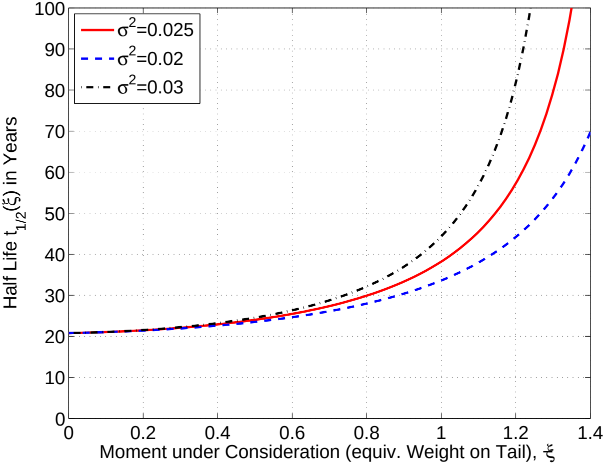
$$\lambda(\xi) = \mu\xi - \frac{\sigma^2}{2}\xi^2 + \delta$$

- Hence, for $\xi < 0$, the higher the moment $-\xi$, the slower the convergence (for high enough $|\xi| < \zeta$)

- Key step: Laplace transform transforms PDE $(*)$ into ODE

$$\frac{\partial \widehat{p}(\xi, t)}{\partial t} = -\xi\mu\widehat{p}(\xi, t) + \xi^2\frac{\sigma^2}{2}\widehat{p}(\xi, t) - \delta\widehat{p}(\xi, t) + \delta$$

Transition in Upper Tail



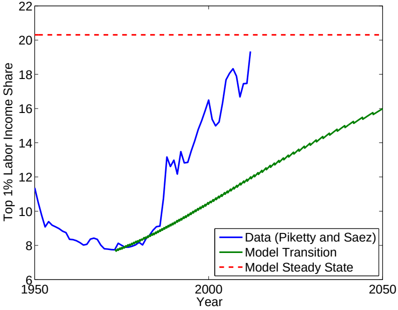
Dynamics of Income Inequality

- Recall process for log wages

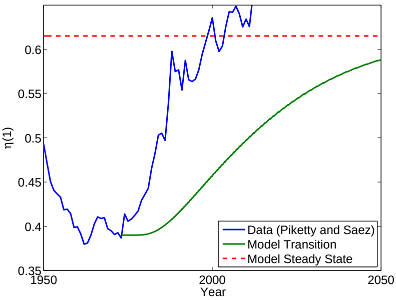
$$d\log w_{it} = \mu dt + \sigma dZ_{it} \qquad + \text{death at rate } \delta$$

- $\sigma^2 = \text{Var(permanent earnings)}$
- **Literature:** σ has increased over last forty years
  - documented by Kopczuk, Saez and Song (2010), DeBacker et al. (2013), Heathcote, Perri and Violante (2010) using PSID
  - but Guvenen, Ozkan and Song (2014): σ flat/decreasing in SSA data
- **Can increase in σ explain increase in top income inequality?**

Dynamics of Income Inequality: Model vs. Data



(a) Top 1% Labor Income Share

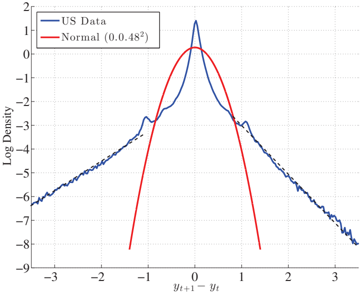


(b) Pareto Exponent

- Experiment $\sigma^2 \uparrow$ from 0.01 in 1973 to 0.025 in 2014 (Heathcote, Perri and Violante, 2010)
- Note: PL exponent $\eta = 1 + \log_{10} \frac{S(0.1)}{S(1)}$ (from $\frac{S(0.1)}{S(1)} = 10^{\eta - 1}$)

Jumps Don't Help Either

- Standard random growth model: income innovations are log-normally distributed
- Recent research: not a good description of the data, e.g. Guvenen-Karahan-Ozkan-Song:



- Natural question: can jumps generate fast transitions?
- Answer: no! While useful descriptively, jumps do not increase the speed of convergence

Jumps Don't Help Either

- Extend income process to

$$dx_{it} = \mu dt + \sigma dZ_{it} + \text{jumps with intensity } \phi \text{ drawn from } f$$

- **Proposition:** With jumps, speed of convergence is

$$\lambda(\xi, \phi) := \xi\mu - \xi^2\frac{\sigma^2}{2} + \delta - \phi(\widehat{f}(\xi) - 1)$$
$$\widehat{f}(\xi) := \int_{-\infty}^{\infty} e^{-\xi g} f(g) dg,$$

Jumps have no effect whatsoever on average speed of convergence

$$\lambda = \delta$$

and they slow down the speed of convergence in the tail

$$\xi < 0 \quad \Rightarrow \quad \lambda(\xi, \phi) \text{ decreasing in } \phi$$

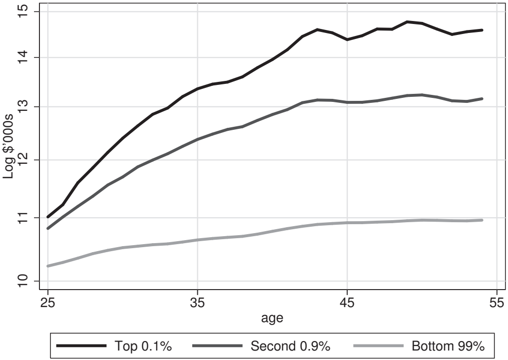
OK, so what drives top inequality then?

Two candidates:

1. heterogeneity in mean growth rates
2. deviations from Gibrat's law, e.g. due to changes in skill prices

Heterogeneity in Mean Growth Rates

(A) Mean earnings by age



- Guvenen, Kaplan and Song (2014): between age 25 and 35
  - earnings of top 0.1% of lifetime inc. grow by ≈ 25% each year
  - and only ≈ 3% per year for bottom 99%

Heterogeneity in Mean Growth Rates

- Two regimes: H and L

$$dx_{it} = \mu_H dt + \sigma_H dZ_{it}$$
$$dx_{it} = \mu_L dt + \sigma_L dZ_{it}$$

- Assumptions
  - $\mu_H > \mu_L$
  - fraction θ enter labor force in H-regime
  - switch from H to L at rate ψ, $L =$ absorbing state
  - retire at rate δ
- See Luttmer (2011) for similar model of firm dynamics
- **Proposition:** The dynamics of $\hat{p}(x,t) = \mathbb{E}[e^{-\xi x}]$ satisfy

$$\hat{p}(\xi,t) - \hat{p}_\infty(\xi) = c_H(\xi) e^{-\lambda_H(\xi)t} + c_L(\xi) e^{-\lambda_L(\xi)t}$$

$$\lambda_H(\xi) := \xi\mu_H - \xi^2 \frac{\sigma_H^2}{2} + \psi + \delta \gg \lambda_L(\xi)$$

and $c_L(\xi), c_H(\xi) =$ constants

"Superstar shocks" to skill prices

- Second candidate for fast transitions: $x_{it} = \log w_{it}$ satisfies

$$x_{it} = \chi_t y_{it}$$
$$dy_{it} = \mu dt + \sigma dZ_{it} \qquad (*)$$

  i.e. $w_{it} = (e^{y_{it}})^{\chi_t}$ and $\chi_t =$ stochastic process $\neq 1$

- Note: implies deviations from Gibrat's law

$$dx_{it} = \mu dt + x_{it} dS_t + \sigma dZ_{it}, \quad S_t := \log \chi_t \neq 0$$

- Call χ_t (equiv. S_t) **"superstar shocks"**

- **Proposition:** The process $(*)$ has an **infinitely fast** speed of adjustment: $\lambda = \infty$. Indeed

$$\zeta_t^x = \zeta^y / \chi_t \qquad \text{or} \qquad \eta_t^x = \chi_t \eta^y$$

  where ζ_t^x, ζ^y are the PL exponents of incomes x_{it} and y_{it}.

- **Intuition:** if power χ_t jumps up, top inequality jumps up

A Microfoundation for "Superstar Shocks"

- χ_t term can be microfounded with changing skill prices in assignment models (Sattinger, 1979; Rosen, 1981)
- Here adopt Gabaix and Landier (2008)
  - continuum of firms of different size $S \sim \text{Pareto}(1/\alpha_t)$.
  - continuum of managers with different talent T, distribution

    $$T(n) = T_{\max} - \frac{B}{\beta} n^{\beta_t}$$

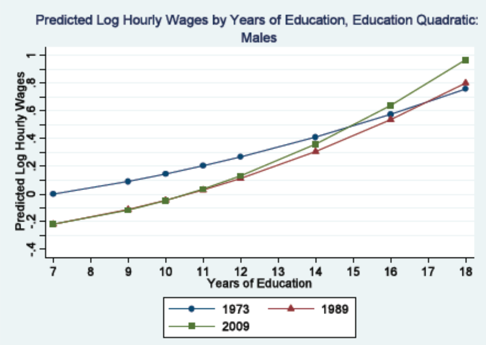
    where $n :=$ rank/quantile of manager talent
  - Match generates firm value: constant $\times TS^{\gamma_t}$
- Can show: $w(n) = e^{a_t} n^{-\chi_t}$ $(= e^{a_t + \chi_t y_{it}}, y_{it} = -\log n_{it})$

$$\chi_t = \alpha_t \gamma_t - \beta_t$$

- Increase in χ_t due to
  - β_t, γ_t: (perceived) importance of talent in production, e.g. due to ICT (Garicano & Rossi-Hansberg, 2006)
- Other assignment models (e.g. with rent-seeking, inefficiencies) would yield similar microfoundation

Empirical Evidence on "Superstar shocks"

1. Acemoglu and Autor (2011): "convexification" of skill prices



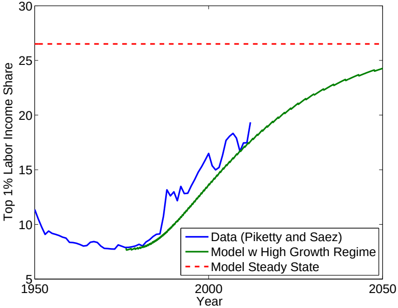
2. Recall

$$dx_{it} = \mu dt + x_{it} dS_t + \sigma dZ_{it}, \quad S_t := \log \chi_t \neq 0$$

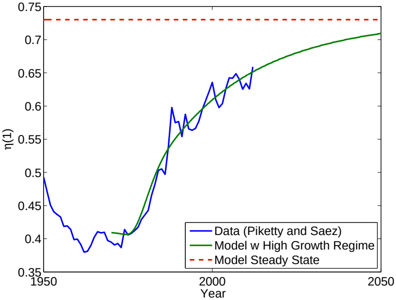
Parker and Vissing-Jorgenson (2009) and Guvenen (2014) find evidence of S_t shocks at business cycle frequencies

Revisiting the Rise in Income Inequality

- Casual evidence: very rapid income growth rates since 1980s (Bill Gates, Mark Zuckerberg)
- Jones and Kim (2015): in IRS/SSA data, average growth rate in upper tail of the growth rate distribution ↑ since late 1970s
- Experiment in model with het. growth rates: in 1973 growth rate of H-types ↑ by 8%



(a) Top 1% Labor Income Share



(b) Pareto Exponent

Wealth Inequality and Capital Taxes

- A simple model of top wealth inequality based on Piketty and Zucman (2015, HID), Piketty (2015, AERPP),...

$$dw_{it} = [y + (r - g - \theta)w_{it}]dt + \sigma w_{it} dZ_{it}$$
$$r = (1 - \tau)\tilde{r}, \quad \sigma = (1 - \tau)\tilde{\sigma}$$

  - y: labor income
  - $R_{it}dt = rdt + \sigma dZ_{it}$: after-tax return on wealth
  - τ: capital tax rate
  - g: economy-wide growth rate
  - θ: MPC out of wealth

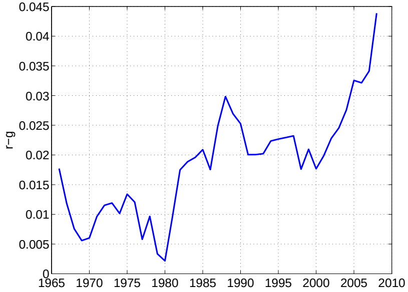
- Stationary top inequality

$$\eta = \frac{1}{\zeta} = \frac{\sigma^2/2}{\sigma^2/2 - (r - g - \theta)}$$

- **Can $r - g$ explain observed dynamics of wealth inequality?**

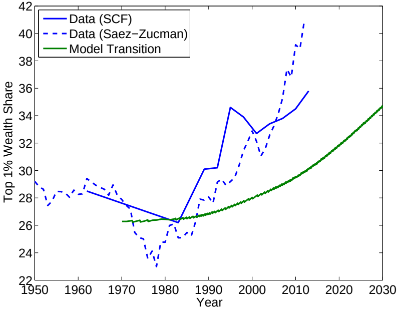
Wealth Inequality and Capital Taxes

- Compute $r_t - g_t = \tilde{r}_t(1 - \tau_t) - g_t$ with ▸ details
  - $\tilde{r}_t$ from Piketty and Zucman (2014)
  - τ_t = capital tax rates from Auerbach and Hassett (2015)
  - g_t = smoothed growth rate from PWT

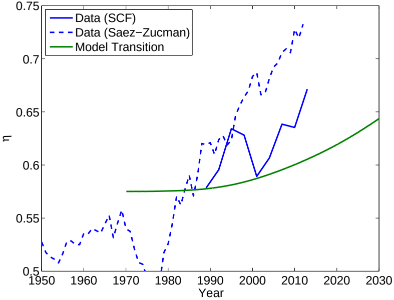


- $\sigma = 0.3$ = upper end of estimates from literature
- θ calibrated to match inequality in 1978

Dynamics of Wealth Inequality



(a) Top 1% Wealth Share



(b) Power Law Exponent

Note: PL exponent $\eta = 1 + \log_{10} \frac{S(0.1)}{S(1)}$ (from $\frac{S(0.1)}{S(1)} = 10^{\eta - 1}$)

OK, so what drives top wealth inequality then?

- Rise in **rate of returns** of super wealthy relative to wealthy (top 0.01 vs. top 1%)
  - better investment advice?
  - better at taking advantage of "tax loopholes"?
  - Kacperczyk, Nosal and Stevens (2015) provide some evidence
- Rise in **saving rates** of super wealthy relative to wealthy
  - Saez and Zucman (2014) provide some evidence

Conclusion

- Transition dynamics of standard random growth models **too slow** relative to those observed in the data
- Two parsimonious deviations that generate **fast transitions**
  1. heterogeneity in mean growth rates
  2. "superstar shocks" to skill prices
- Rise in top **income** inequality due to
  - ~~simple tax stories, stories about Var(permanent earnings)~~
  - **rise in superstar growth (and churn) in two-regime world**
  - **"superstar shocks" to skill prices**
- Rise in top **wealth** inequality due to
  - ~~increase in $r - g$ due to falling capital taxes~~
  - **rise in saving rates/RoRs of super wealthy**