

Lectures 6 and 7

Theories of Top Inequality

Distributional Dynamics and Differential Operators

ECO 521: Advanced Macroeconomics I

Benjamin Moll

Outline

1. Gabaix (2009) “Power Laws in Economics and Finance”
2. Literature on inequality and random growth
3. Gabaix-Lasry-Lions-Moll (2016) “The Dynamics of Inequality”
 - tools: differential operators as transition matrices
 - will be extremely useful for analysis, computation of fully-fledged heterogeneous agent models later on

Power Laws

- **Definition 1:** S follows a **power law** (PL) if there exist $k, \zeta > 0$ s.t.

$$\Pr(S > x) = kx^{-\zeta}, \quad \text{all } x$$

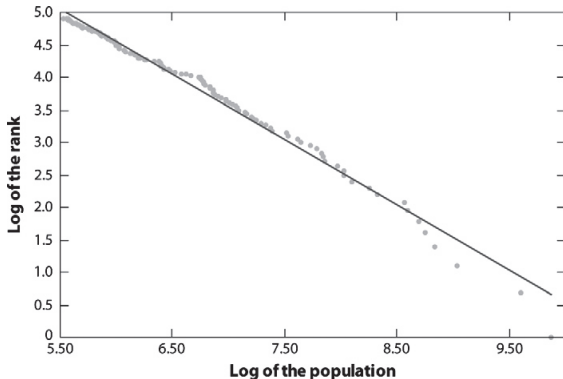
- S follows a PL $\Leftrightarrow S$ has a Pareto distribution
- **Definition 2:** S follows an **asymptotic power law** if there exist $k, \zeta > 0$ s.t.

$$\Pr(S > x) \sim kx^{-\zeta} \quad \text{as } x \rightarrow \infty$$

- Note: for any f, g $f(x) \sim g(x)$ means $\lim_{x \rightarrow \infty} f(x)/g(x) = 1$
- Surprisingly many variables follow power laws, at least in tail

City Size

- Order cities in US by size (NY as first, LA as second, etc)
- Graph $\ln \text{Rank}$ ($\ln \text{Rank}_{NY} = \ln 1$, $\ln \text{Rank}_{LA} = \ln 2$) vs. $\ln \text{Size}$
- Basically plot log quantiles $\ln \Pr(S > x)$ against $\ln x$



City Size

- **Surprise 1:** straight line, i.e. city size follows a PL

$$\Pr(S > x) = kx^{-\zeta}$$

- **Surprise 2:** slope of line ≈ -1 , regression:

$$\ln \text{Rank} = 10.53 - 1.005 \ln \text{Size}$$

i.e. city size follows a PL with exponent $\zeta \approx 1$

$$\Pr(S > x) = kx^{-1}.$$

- A power law with exponent $\zeta = 1$ is called “Zipf’s law”
- Two natural questions:
 1. Why does city size follow a power law?
 2. Why on earth is $\zeta \approx 1$ rather than any other number?

Where Do Power Laws Come from?

- Gabaix's answer: random growth
- Economy with continuum of cities
- S_t^i : size of city i at time t

$$S_{t+1}^i = \gamma_{t+1}^i S_t^i, \quad \gamma_{t+1}^i \sim f(\gamma) \quad (\text{RG})$$

- S_t^i follows random growth process $\Leftrightarrow \log S_t^i$ follows random walk
- Gabaix shows: (RG) + stabilizing force (e.g. minimum size) \Rightarrow power law. Use “Champernowne's equation”
- Easier: continuous time approach

Random Growth Process in Continuous Time

- Consider random growth process over time intervals of length Δt

$$S_{t+\Delta t}^i = \gamma_{t+\Delta t}^i S_t^i$$

- Assume in addition that $\gamma_{t+\Delta t}^i$ takes the particular form

$$\gamma_{t+\Delta t}^i = 1 + g\Delta t + \nu \varepsilon_t^i \sqrt{\Delta t}, \quad \varepsilon_t^i \sim \mathcal{N}(0, 1)$$

- Substituting in

$$S_{t+\Delta t}^i - S_t^i = (g\Delta t + \nu \varepsilon_t^i \sqrt{\Delta t}) S_t^i$$

- Or as $\Delta t \rightarrow 0$

$$dS_t^i = gS_t^i dt + \nu S_t^i dW_t^i$$

i.e. a geometric Brownian motion!

Stationary Distribution

- Assumption: city size follows random growth process

$$dS_t^i = gS_t^i dt + \nu S_t^i dW_t^i$$

- Does this have a stationary distribution? No! In fact

$$\log S_t^i \sim \mathcal{N}((g - \nu^2/2)t, \nu^2 t)$$

⇒ distribution explodes.

- Gabaix insight: random growth process + **stabilizing force** does have a stationary distribution and that's a PL
 - Note: Gabaix uses “friction” rather than “stabilizing force”
 - use the latter because “friction” already means something else
- Simplest possible stabilizing force: $g < 0$ and minimum size S_{\min}
 - if process goes below S_{\min} it is brought back to S_{\min} (“reflecting barrier”)

Stationary Distribution

- Use Kolmogorov Forward Equation
- Recall: stationary distribution satisfies

$$0 = -\frac{d}{dx}[\mu(x)f(x)] + \frac{1}{2}\frac{d^2}{dx^2}[\sigma^2(x)f(x)]$$

- Here geometric Brownian motion: $\mu(x) = gx$, $\sigma^2(x) = \nu^2x^2$

$$0 = -\frac{d}{dx}[gxf(x)] + \frac{1}{2}\frac{d^2}{dx^2}[\nu^2x^2f(x)]$$

Stationary Distribution

- **Claim:** solution is a Pareto distribution, $f(x) = S_{\min}^{\zeta} x^{-\zeta-1}$
- **Proof:** Guess $f(x) = Cx^{-\zeta-1}$ and verify

$$\begin{aligned} 0 &= -\frac{d}{dx}[gx Cx^{-\zeta-1}] + \frac{1}{2} \frac{d^2}{dx^2}[\nu^2 x^2 Cx^{-\zeta-1}] \\ &= Cx^{-\zeta-1} \left[g\zeta + \frac{\nu^2}{2} (\zeta - 1)\zeta \right] \end{aligned}$$

- This is a quadratic equation with two roots $\zeta = 0$ and

$$\zeta = 1 - \frac{2g}{\nu^2}$$

- For mean to exist, need $\zeta > 1 \Rightarrow$ impose $g < 0$
- Remains to pin down C . We need

$$1 = \int_{S_{\min}}^{\infty} f(x) dx = \int_{S_{\min}}^{\infty} Cx^{-\zeta-1} dx \quad \Rightarrow \quad C = S_{\min}^{\zeta} \cdot \square$$

Tail inequality and Zipf's Law

- “Tail inequality” (fatness of tail)

$$\eta := \frac{1}{\zeta} = \frac{1}{1 - 2g/\nu^2}$$

is increasing in g and ν^2 (recall $g < 0$)

- Why would Zipf's Law ($\zeta = 1$) hold? We have that

$$\bar{S} = \int_{S_{\min}}^{\infty} xf(x)dx = \frac{\zeta}{\zeta - 1} S_{\min}$$

$$\Rightarrow \zeta = \frac{1}{1 - S_{\min}/\bar{S}} \rightarrow 1 \quad \text{as} \quad S_{\min}/\bar{S} \rightarrow 0.$$

- Zipf's law obtains as stabilizing force becomes small

Alternative Stabilizing Force: Death

- No minimum size
- Instead: die at Poisson rate δ , get reborn at S_*
- Can show: correct way of extending KFE (for $x \neq S_*$) is

$$\frac{\partial f(x, t)}{\partial t} = -\delta f(x, t) - \frac{\partial}{\partial x} [\mu(x) f(x, t)] + \frac{1}{2} \frac{\partial^2}{\partial x^2} [\sigma^2(x) f(x, t)]$$

- Stationary $f(x)$ satisfies (recall $\mu(x) = gx, \sigma^2(x) = \nu^2 x^2$)

$$0 = -\delta f(x) - \frac{d}{dx} [gx f(x, t)] + \frac{1}{2} \frac{d^2}{dx^2} [\nu^2 x^2 f(x)] \quad (\text{KFE}')$$

Alternative Stabilizing Force: Death

- To solve (KFE'), guess $f(x) = Cx^{-\zeta-1}$

$$0 = -\delta + \zeta g + \frac{\nu^2}{2}\zeta(\zeta - 1)$$

- Two roots: $\zeta_+ > 0$ and $\zeta_- < 0$. General solution to (KFE'):

$$\Rightarrow f(x) = C_- x^{-\zeta_- - 1} + C_+ x^{-\zeta_+ - 1} \quad \text{for } x \neq S_*$$

- Need solution to be integrable

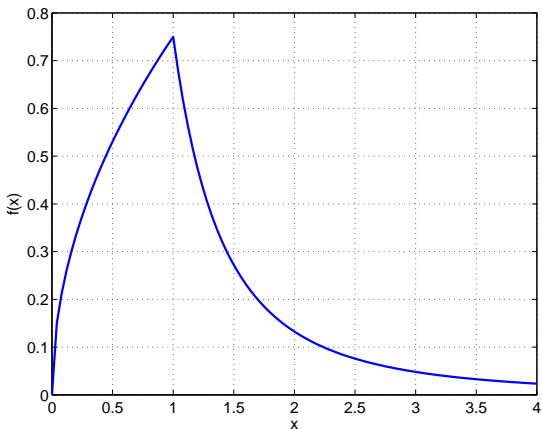
$$\int_0^\infty f(x) dx = f(S_*) + \int_0^{S_*} f(x) dx + \int_{S_*}^\infty f(x) dx < \infty$$

- Hence $C_- = 0$ for $x > S_*$, otherwise $f(x)$ explodes as $x \rightarrow \infty$
- And $C_+ = 0$ for $x < S_*$, otherwise $f(x)$ explodes as $x \rightarrow 0$

Alternative Stabilizing Force: Death

- Solution is a **Double Pareto** distribution:

$$f(x) = \begin{cases} C(x/S_*)^{-\zeta_- - 1} & \text{for } x < S_* \\ C(x/S_*)^{-\zeta_+ - 1} & \text{for } x > S_* \end{cases}$$



Generalizations and Other Stabilizing Forces

- See Appendix D of “The Dynamics of Inequality” for a pretty exhaustive list

- death and rebirth with $S_t^i \sim \psi(S)$
- additive term

$$dS_t^i = ydt + gS_t^i dt + vS_t^i dW_t^i, \quad g < 0, y > 0$$

-
- In general, distribution **will not** be exactly Pareto or exactly double-Pareto
- But often, under quite weak assumptions, it will still follow **asymptotic power law**, i.e.

$$\Pr(S > x) \sim kx^{-\zeta} \quad \text{as } x \rightarrow \infty$$

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), Panousi (2012), Benhabib, Bisin, Zhu (2012, 2014), Piketty and Zucman (2014), Piketty and Saez (2014), **Piketty (2015)**, Benhabib and Bisin (2016) is nice survey
- **Dynamics** of income and wealth distribution
 - Aoki and Nirei (2014), Gabaix, Lasry, Lions and Moll (2016), Hubmer, Krusell, Smith (2016)

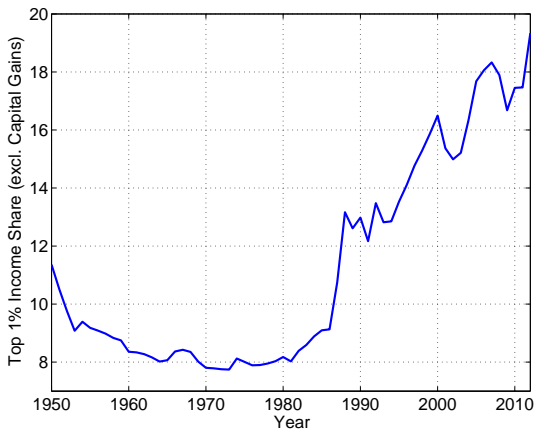
Literature: Inequality and Random Growth

From Piketty “About Capital in the Twenty-First Century” (AEA P&P, 2015)

- “Technically, one can indeed show that if shocks take a **multiplicative** form, then the inequality of wealth converges toward a distribution that has a Pareto shape for top wealth holders [...], and that the inverted Pareto coefficient (an indicator of top end inequality) is a steeply rising function of the gap $r - g$.”
- Idea: $\mu(x) = (r - g - \text{constant})x$
- In book this point unfortunately gets lost in discussion about how $r - g$ affects capital share
 - factor income vs personal income distribution
 - no general connection between capital share and inequality (see end of Lecture 5)

The Dynamics of Inequality

Question

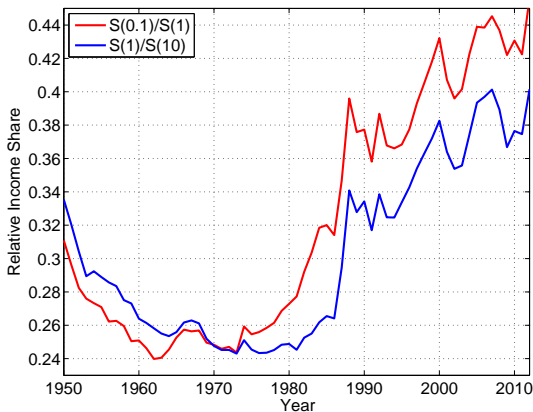


- In U.S. past 40 years have seen rapid rise in top income inequality
- Why?

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”



- $\frac{S(0.1)}{S(1)}$ = fraction of top 1% share going to top 0.1%
- $\frac{S(1)}{S(10)}$ = analogous

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 income inequality?
 - e.g. falling income taxes? superstar effects?
- **What we do:**
 - study **transition dynamics** of cross-sectional income distribution in theories with random growth mechanism
 - contrast with data, **rule out** some theories, **rule in** others
- **Today:** income inequality. **Paper:** also wealth inequality.

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 $\text{Var}(\text{permanent earnings})$~~
 - rise of “superstar” entrepreneurs or managers

A Random Growth Theory of Income Dynamics

- 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

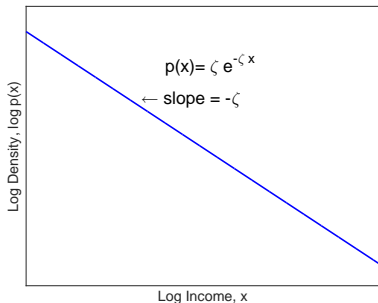
$$d \log w_{it} = \mu dt + \sigma dZ_{it}$$

- **growth rate** of wage w_{it} is **stochastic**
- μ, σ depend on model parameters
- see Appendix C: log-utility + constant returns (same trick as AK-RBC model in Lecture 4)

Stationary Income Distribution

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

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



- Convenient to work with log income $x_t = \log w_t$

$$\Pr(\tilde{w} > w) \sim C w^{-\zeta} \quad \Leftrightarrow \quad \Pr(\tilde{x} > x) \sim C e^{-\zeta x}$$

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

Stationary Income Distribution

- Have $x_{it} = \log w_{it}$ follows

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

- Need additional “stabilizing force” to ensure existence of stat. dist.
 - income application: death/retirement at rate δ
 - alternative: reflecting barrier
- Distribution $p(x, t)$ satisfies ($\psi(x)$ = distribution of entry wages)

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

- With reflecting barrier at $x = 0$, have boundary condition

$$0 = -\mu p(0, t) + \frac{\sigma^2}{2} p_x(0, t)$$

Derivation: $\int_0^\infty p(x, t) dx = 1$ for all t , and hence from (*)

$$0 = \int_0^\infty p_t dx = \left[-\mu p + \frac{\sigma^2}{2} p_x \right]_0^\infty$$

Stationary Income Distribution

- Stationary Distribution $p_\infty(x)$ satisfies

$$0 = -\mu p_x + \frac{\sigma^2}{2} p_{xx} - \delta p + \delta \psi$$

- Find solution via guess-and-verify: plug in $p(x) = C e^{-\zeta x}$

$$0 = \mu \zeta + \frac{\sigma^2}{2} \zeta^2 - \delta + \delta \frac{\psi(x)}{C e^{-\zeta x}}$$

- Assume $\lim_{x \rightarrow \infty} \psi(x)/e^{-\zeta x} = 0 \Rightarrow$ last term drops for large x & ζ solves

$$0 = \mu \zeta + \frac{\sigma^2}{2} \zeta^2 - \delta$$

with positive root

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

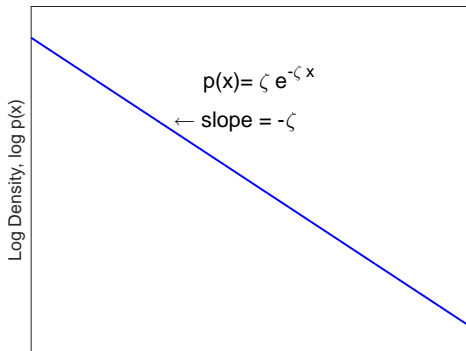
- Tail inequality $\eta = 1/\zeta$ increasing in μ, σ , decreasing in δ

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

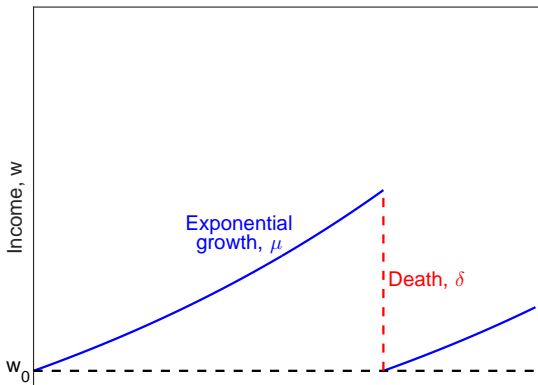
Transitions: The Thought Experiment

- Suppose economy is in Pareto steady state
- At $t = 0$, $\sigma \uparrow$. Know: in long-run \rightarrow higher top inequality



Instructive Special Case: $\sigma = 0$ ("Steindl Model")

- In special case $\sigma = 0$, can solve full transition dynamics
 - w_t grows at rate μ , gets reset to $w_0 = 1$ at rate δ
 - stationary distribution $f(w) = \zeta w^{-\zeta}$, $\zeta = \delta/\mu$
 - stationary distribution of $x_t = \log w_t$: $p(x) = \zeta e^{-\zeta x}$, $\zeta = \delta/\mu$
 - at $t = 0$ $\mu \uparrow$ Know from $\zeta = \delta/\mu$ in long-run top inequality \uparrow



Transition in Steindl Model

- Denote
 - old steady state distribution: $p_0(x) = \alpha e^{-\alpha x}$
 - new steady state distribution: $p_\infty(x) = \zeta e^{-\zeta x}$
- Can show: for $t, x > 0$ density satisfies

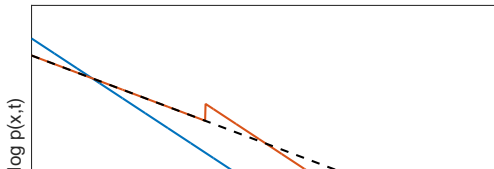
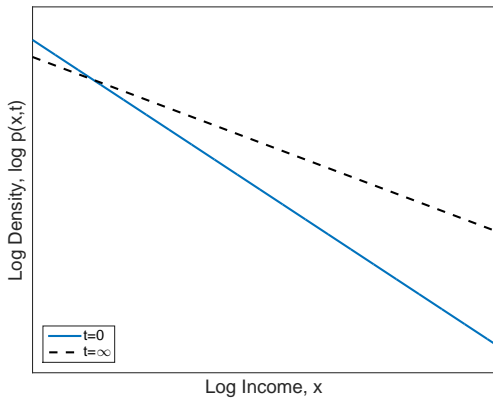
$$\frac{\partial p(x, t)}{\partial t} = -\mu \frac{\partial p(x, t)}{\partial x} - \delta p(x, t), \quad p(x, 0) = \alpha e^{-\alpha x} \quad (*)$$

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

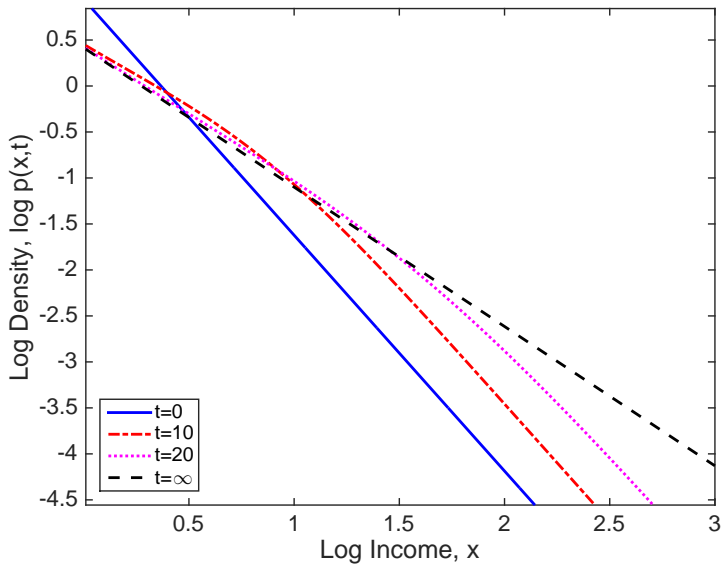
$$p(x, t) = \zeta e^{-\zeta x} \mathbf{1}_{\{x \leq \mu t\}} + \alpha e^{-\alpha x + (\alpha - \zeta)t} \mathbf{1}_{\{x > \mu t\}}$$

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

Transition in Steindl Model



General Case



General Case

- Recall Kolmogorov Forward equation for $p(x, t)$

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

- Question: at what speed does $p(x, t)$ converge to $p_\infty(x)$?
- need a “distance measure”
- Use L^1 norm:

$$\|p(x, t) - p_\infty(x)\| := \int_{-\infty}^{\infty} |p(x, t) - p_\infty(x)| dx$$

- measures **average** distance between p and p_∞
- Later: more general distance measures

General Case: Average Speed of Convergence

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

$$\lambda := - \lim_{t \rightarrow \infty} \frac{1}{t} \log \|p(x, t) - p_\infty(x)\| \quad (*)$$

- without reflecting barrier

$$\lambda = \delta$$

- with reflecting barrier

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

- Interpretation of (*): exponential convergence at rate λ

$$\|p(x, t) - p_\infty(x)\| \sim k e^{-\lambda t} \quad \text{as } t \rightarrow \infty$$

- Half life is $t_{1/2} = \ln(2)/\lambda \Rightarrow$ precise quantitative predictions

Before proving this, let's take a step back...

- ... and take a somewhat different perspective on the Kolmogorov Forward equation
 - exploit heavily analogy to finite-state processes
- This will also be extremely useful for **computations**
- Let's focus on case with reflecting barrier at $x = 0$ and $\delta = 0$
- Kolmogorov Forward equation is

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

with boundary condition

$$0 = -\mu p(0, t) + \frac{\sigma^2}{2} p_x(0, t)$$

Key: operator in KFE = transpose of transition matrix

- Just for a moment, suppose x_{jt} = finite-state Poisson process
- $x_{jt} \in \{x_1, \dots, x_N\} \Rightarrow$ distribution = vector $\mathbf{p}(t) \in \mathbb{R}^N$
- Dynamics of distribution

$$\dot{\mathbf{p}}(t) = \mathbf{A}^T \mathbf{p}(t),$$

where $\mathbf{A} = N \times N$ transition matrix

- Key idea: KFE is exact analogue with continuous state
- Can write in terms of differential operator \mathcal{A}^*

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

with boundary condition $0 = -\mu p(0) + \frac{\sigma^2}{2} p_x(0)$

- \mathcal{A}^* analogue of transpose of transition matrix \mathbf{A}^T

This can be made more precise...

- **Definition:** the inner product of two functions v and p is $\langle v, p \rangle = \int_0^\infty v(x)p(x)dx$ (analogue of $\mathbf{v} \cdot \mathbf{p} = \sum_{i=1}^N v_i p_i$)
- **Definition:** the **adjoint** of an operator \mathcal{A} is the operator \mathcal{A}^* satisfying

$$\langle \mathcal{A}v, p \rangle = \langle v, \mathcal{A}^*p \rangle$$

Note: adjoint = analogue of matrix transpose $\mathbf{A}\mathbf{v} \cdot \mathbf{p} = \mathbf{v} \cdot \mathbf{A}^T\mathbf{p}$

- **Definition:** An operator \mathcal{B} is **self-adjoint** if $\mathcal{B}^* = \mathcal{B}$
- **Definition:** the **infinitesimal generator** of a Brownian motion is the operator \mathcal{A} defined by

$$\mathcal{A}v = \mu v_x + \frac{\sigma^2}{2} v_{xx}$$

with boundary condition $v_x(0) = 0$

- same operator shows up in HJB equations, e.g.

$$\rho v = u + \mu v_x + \frac{\sigma^2}{2} v_{xx}, \quad u = \text{period return}$$

- will call it “HJB operator”, plays role of transition matrix

\mathcal{A}^* is adjoint of \mathcal{A} (& vice versa)

- Result: \mathcal{A}^* in the Kolmogorov Forward equation is the adjoint of \mathcal{A}
- Proof:

$$\begin{aligned}\langle v, \mathcal{A}^* p \rangle &= \int_0^\infty v \left(-\mu p_x + \frac{\sigma^2}{2} p_{xx} \right) dx \\ &= \left[-v\mu p + \frac{\sigma^2}{2} v p_x \right]_0^\infty - \int_0^\infty \left(-\mu v_x p + \frac{\sigma^2}{2} v_x p_x \right) dx \\ &= \left[-v\mu p + \frac{\sigma^2}{2} v p_x - \frac{\sigma^2}{2} v_x p \right]_0^\infty + \int_0^\infty \left(\mu v_x p + \frac{\sigma^2}{2} v_{xx} p \right) dx \\ &= v(0) \left(\mu p(0) - \frac{\sigma^2}{2} p_x(0) \right) + \frac{\sigma^2}{2} v_x(0) p(0) + \langle \mathcal{A} v, p \rangle \\ &= \langle \mathcal{A} v, p \rangle.\end{aligned}$$

- key step is to use integration by parts and boundary conditions

Carries over to **any** diffusion process

- ... with x -dependent μ and σ
- “HJB operator” (infinitesimal generator)

$$\mathcal{A}v = \mu(x) \frac{\partial v}{\partial x} + \frac{\sigma^2(x)}{2} \frac{\partial^2 v}{\partial x^2}$$

with appropriate boundary conditions

- “Kolmogorov Forward operator”

$$\mathcal{A}^*p = -\frac{\partial}{\partial x}(\mu(x)p) + \frac{1}{2} \frac{\partial^2}{\partial x^2}(\sigma^2(x)p)$$

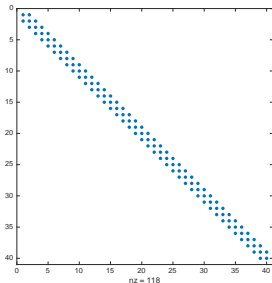
with appropriate boundary conditions

- **Result:** \mathcal{A}^* is adjoint of \mathcal{A}
- Proof: integration by parts just like previous slide

Computation of Kolmogorov Forward Equations

- That operator in KFE = transpose of transition matrix is very useful for computations
- Use finite difference method $p_i^n = p(x_i, t^n)$
- Key: already know how to discretize \mathcal{A}
- recall from Lectures 3 and 4 that discretize HJB equation as

$$\rho v = u + \mu v_x + \frac{\sigma^2}{2} v_{xx} \quad \text{as} \quad \rho \mathbf{v} = \mathbf{u} + \mathbf{A} \mathbf{v}$$



Computation of Kolmogorov Forward Equations

- By same logic: **correct discretization of \mathcal{A}^* is \mathbf{A}^\top**
- Discretize

$$p_t = -\mu p_x + \frac{\sigma^2}{2} p_{xx} \quad \text{or} \quad p_t = \mathcal{A}^* p \quad (\text{KFE})$$

as (explicit scheme)

$$\frac{\mathbf{p}^{n+1} - \mathbf{p}^n}{\Delta t} = \mathbf{A}^\top \mathbf{p}^n$$

or slightly better (implicit scheme)

$$\frac{\mathbf{p}^{n+1} - \mathbf{p}^n}{\Delta t} = \mathbf{A}^\top \mathbf{p}^{n+1} \quad \Rightarrow \quad \mathbf{p}^{n+1} = (\mathbf{I} - \Delta t \mathbf{A}^\top)^{-1} \mathbf{p}^n$$

- can also obtain these finite-difference schemes directly from (KFE), i.e. without using “operator in KFE = transpose of transition matrix”
 - Section 2 in http://www.princeton.edu/~moll/HACTproject/HACT_Numerical_Appendix.pdf
- but if have already computed \mathbf{A} for HJB equation, no need to do discretization again – get (KFE) for free!

Back to the proof of average-speed proposition

- To gain intuition, suppose again finite-state process $\mathbf{p}(t) \in \mathbb{R}^N$ with

$$\dot{\mathbf{p}}(t) = \mathbf{A}^T \mathbf{p}(t)$$

- assume \mathbf{A} is diagonalizable
- denote eigenvalues by $0 = |\lambda_1| < |\lambda_2| < \dots < |\lambda_N|$
- corresponding eigenvectors by $(\mathbf{v}_1, \dots, \mathbf{v}_N)$
- **Theorem:** $\mathbf{p}(t)$ converges to \mathbf{p}_∞ at rate $|\lambda_2|$ (“spectral gap”)
- Proof sketch: decomposition

$$\mathbf{p}(0) = \sum_{i=1}^N c_i \mathbf{v}_i \quad \Rightarrow \quad \mathbf{p}(t) = \sum_{i=1}^N c_i e^{\lambda_i t} \mathbf{v}_i$$

- 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

Proof of proposition (reflecting barrier, $\delta = 0$)

- Generalize this idea to continuous-state process
- Analyze Kolmogorov Forward equation

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

in same exact way as $\dot{\mathbf{p}}(t) = \mathbf{A}^T \mathbf{p}(t)$

- **Proof** has two steps:
 1. realization that speed = second eigenvalue (spectral gap) of operator \mathcal{A}^*
 2. analytic computation: spectral gap given by

$$|\lambda_2| = \frac{1}{2} \frac{\mu^2}{\sigma^2}$$

Analytic Computation of Spectral Gap

- Discrete eigenvalue problem

$$\mathbf{A}\mathbf{v} = \lambda\mathbf{v}$$

- Continuous eigenvalue problem

$$\mathcal{A}\varphi = \lambda\varphi$$

or

$$\mu\varphi'(x) + \frac{\sigma^2}{2}\varphi''(x) = \lambda\varphi(x)$$

with boundary condition $\varphi'(0) = 0$.

- In principle, could analyze that one directly, **but...**

Analytic Computation of Spectral Gap

- **Definition:** an operator \mathcal{B} is *self-adjoint* if $\mathcal{B}^* = \mathcal{B}$
- **Result:** all eigenvalues of a self-adjoint operator are real
- want to analyze eigenvalues of \mathcal{A}
 - but problem: \mathcal{A} is not self-adjoint
 - eigenvalues could have imaginary parts
- Solution: construct **self-adjoint transformation \mathcal{B}** of \mathcal{A} as follows
 1. Consider stationary distribution p_∞ satisfying

$$0 = \mathcal{A}^* p \quad \Rightarrow \quad p_\infty = e^{(2\mu/\sigma^2)x}$$

2. Consider $u = v p_\infty^{1/2} = v e^{(\mu/\sigma^2)x}$. Can show u satisfies

$$u_t = \mathcal{B}u, \quad \mathcal{B}u := \frac{\sigma^2}{2} u_{xx} - \frac{1}{2} \frac{\mu^2}{\sigma^2} u$$

with boundary condition $u_x(0) = \frac{\mu}{\sigma^2} u(0)$.

- To see that \mathcal{B} is self-adjoint: $\langle \mathcal{B}u, p \rangle = \langle u, \mathcal{B}p \rangle$ using same steps as before (integration by parts)

Eigenvalues of \mathcal{B}

The first eigenvalue of \mathcal{B} is $\lambda_1 = 0$ and the second eigenvalue is $\lambda_2 = -\frac{1}{2}\frac{\mu^2}{\sigma^2}$. All remaining eigenvalues satisfy $|\lambda| > |\lambda_2|$

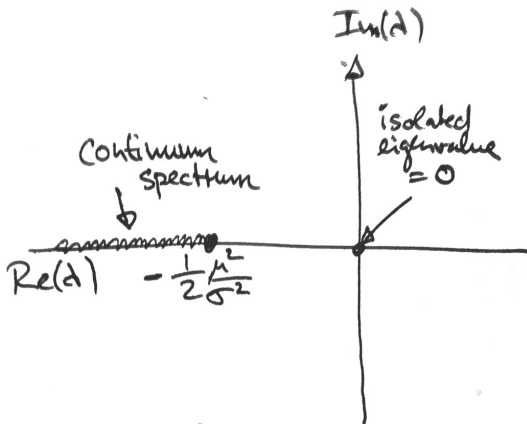


Figure: Spectrum of \mathcal{B} in complex plane

Proof of Lemma

- Consider eigenvalue problem

$$\mathcal{B}\varphi = \lambda\varphi$$

$$\frac{\sigma^2}{2}\varphi''(x) - \frac{1}{2}\frac{\mu^2}{\sigma^2}\varphi(x) = \lambda\varphi(x) \quad (\text{E})$$

with boundary condition $\varphi'(0) = \frac{\mu}{\sigma^2}\varphi(0)$

- Can show: for $\lambda \in \left(-\frac{1}{2}\frac{\mu^2}{\sigma^2}, 0\right)$ all solutions to (E) satisfying boundary condition explode as $|x| \rightarrow \infty$. See appendix of paper.
- Intuition why rate of convergence of \mathcal{B} is $\frac{1}{2}\frac{\mu^2}{\sigma^2}$

- recall $\mathcal{B}u := \frac{\sigma^2}{2}u_{xx} - \frac{1}{2}\frac{\mu^2}{\sigma^2}u$

- consider case $\sigma \approx 0$: $\frac{1}{2}\frac{\mu^2}{\sigma^2}$ term large relative to $\frac{\sigma^2}{2}$

$$u_t = \mathcal{B}u \approx -\frac{1}{2}\frac{\mu^2}{\sigma^2}u \quad \Rightarrow \quad u(x, t) \approx u_0(x)e^{-\frac{1}{2}\frac{\mu^2}{\sigma^2}t}$$

i.e. operator \mathcal{B} features exponential decay at rate $\frac{1}{2}\frac{\mu^2}{\sigma^2}$

Transition in Upper Tail

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

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

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

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

- \hat{p} has natural interpretation: $-\xi$ th moment of income/wealth
 $w_{it} = e^{x_{it}}$
 - e.g. $\hat{p}(-2, t) = \mathbb{E}[w_{it}^2]$
- only works in case without reflecting barrier/lower bound

Transition in Upper Tail

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

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

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

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

Can the model explain the fast rise in inequality?

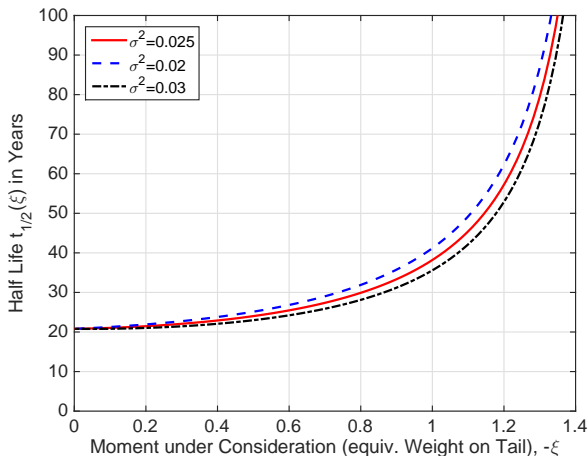
- Recall process for log wages

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

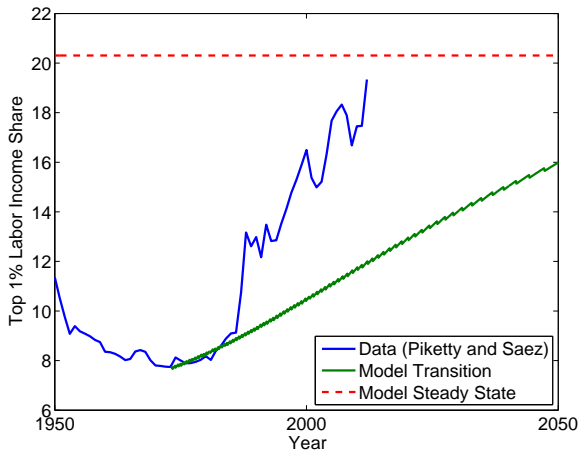
- $\sigma^2 = \text{Var}(\text{permanent earnings})$
- **Literature:** σ has increased over last forty years
 - Kopczuk, Saez and Song (2010), DeBacker et al. (2013), Heathcote, Perri and Violante (2010) using PSID
 - but Guvenen, Ozkan and Song (2014): σ flat/decreasing in SSA data
- **Can increase in σ explain increase in top income inequality?**
 - experiment: $\sigma^2 \uparrow$ from 0.01 in 1973 to 0.025 in 2014 (Heathcote-Perri-Violante)

Putting the Theory to Work

- Recall formula $\lambda(\xi) = \mu\xi - \frac{\sigma^2}{2}\xi^2 + \delta$
- Compute half-life $t_{1/2}(\xi) = \log 2/\lambda(\xi)$



Transition following Increase in σ^2 from 0.01 to 0.025



OK, so what drives top inequality then?

Two candidates:

1. “type dependence”: heterogeneity in mean growth rates
2. “scale dependence”: “superstar shocks” to skill prices

Both are violations of Gibrat’s law

Type Dependence

- Casual evidence: very rapid income growth rates since 1980s (Bill Gates, Mark Zuckerberg)
- Two regimes: H and L with $\mu_H > \mu_L$

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

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

- Assumptions
 - drop from H to L at rate ψ
 - retire at rate δ
- See Luttmer (2011) for similar model of firm dynamics
- **Proposition:** Speed of transition determined by

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

Scale Dependence

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

$$\begin{aligned}x_{it} &= \chi_t y_{it} \\ dy_{it} &= \mu dt + \sigma dZ_{it}\end{aligned}\tag{*}$$

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

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

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

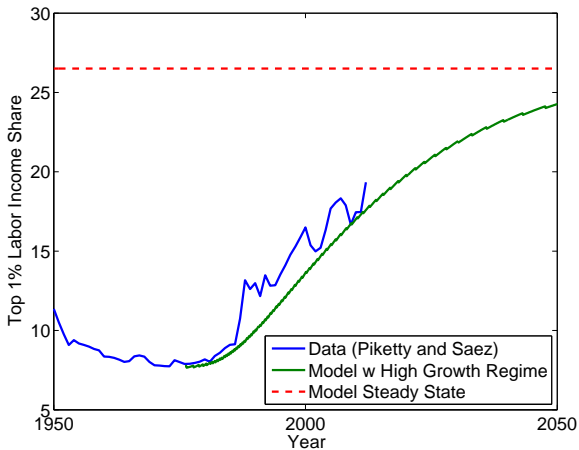
- Match generates firm value: constant $\times T S \gamma_t$
- Can show: $w(n) = e^{a_t} n^{-\chi_t}$ ($= e^{a_t + \chi_t y_{it}}$, $y_{it} = -\log n_{it}$)

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

- 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)

Revisiting the Rise in Income Inequality

- Jones and Kim (2015): in IRS/SSA data, $\mu_H \uparrow$ since 1970s
- Experiment: in 1973 $\mu_H \uparrow$ by 8%



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 $\text{Var}(\text{permanent earnings})$~~
 - rise in superstar growth (and churn) in two-regime world
 - “superstar shocks” to skill prices
- See paper for wealth inequality results

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

Tools Summary

- Differential operators as transition matrices
- At fundamental level, **everything same** whether discrete/continuous time/space
 - nothing special about continuous t
 - nothing special about continuous x
 - all results from discrete time/space carry over to infinite-dimensional (i.e. continuous) case
 - but computational advantages (e.g. sparsity) – next lecture
- Analogies
 - function $p \Leftrightarrow$ vector \mathbf{p}
 - (linear) operator $\mathcal{A} \Leftrightarrow$ matrix \mathbf{A}
 - adjoint $\mathcal{A}^* \Leftrightarrow$ transpose \mathbf{A}^T

Open Questions

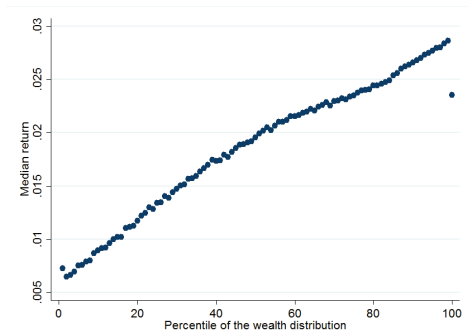
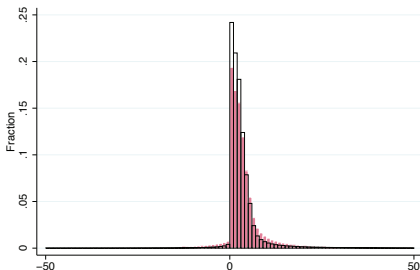
- “What fraction” of top inequality is **efficient** in the sense of people getting paid marginal product? What fraction due to rent-seeking?
- What are the **underlying economic forces** that drove the increase in top inequality?
 - technical change?
 - globalization?
 - superstars?
 - rent-seeking?
 - particular sectors/occupations?
- Evidence for scale- and type-dependence?
 - for wealth: Fagereng, Guiso, Malacrino and Pistaferri (2016), “Heterogeneity and Persistence in Returns to Wealth”
 - what about income?

Fagereng-Guiso-Malacrino-Pistaferri

- Using Norwegian administrative data (Norway has wealth tax), document massive heterogeneity in returns to wealth
 - range of over 500 basis points between 10th and 90th pctile
 - returns positively correlated with wealth

Distribution of returns on wealth

) Full sample



- Interesting open question: can a process for returns to wealth like the one documented by FGMP quantitatively generate fast dynamics in top wealth inequality?