



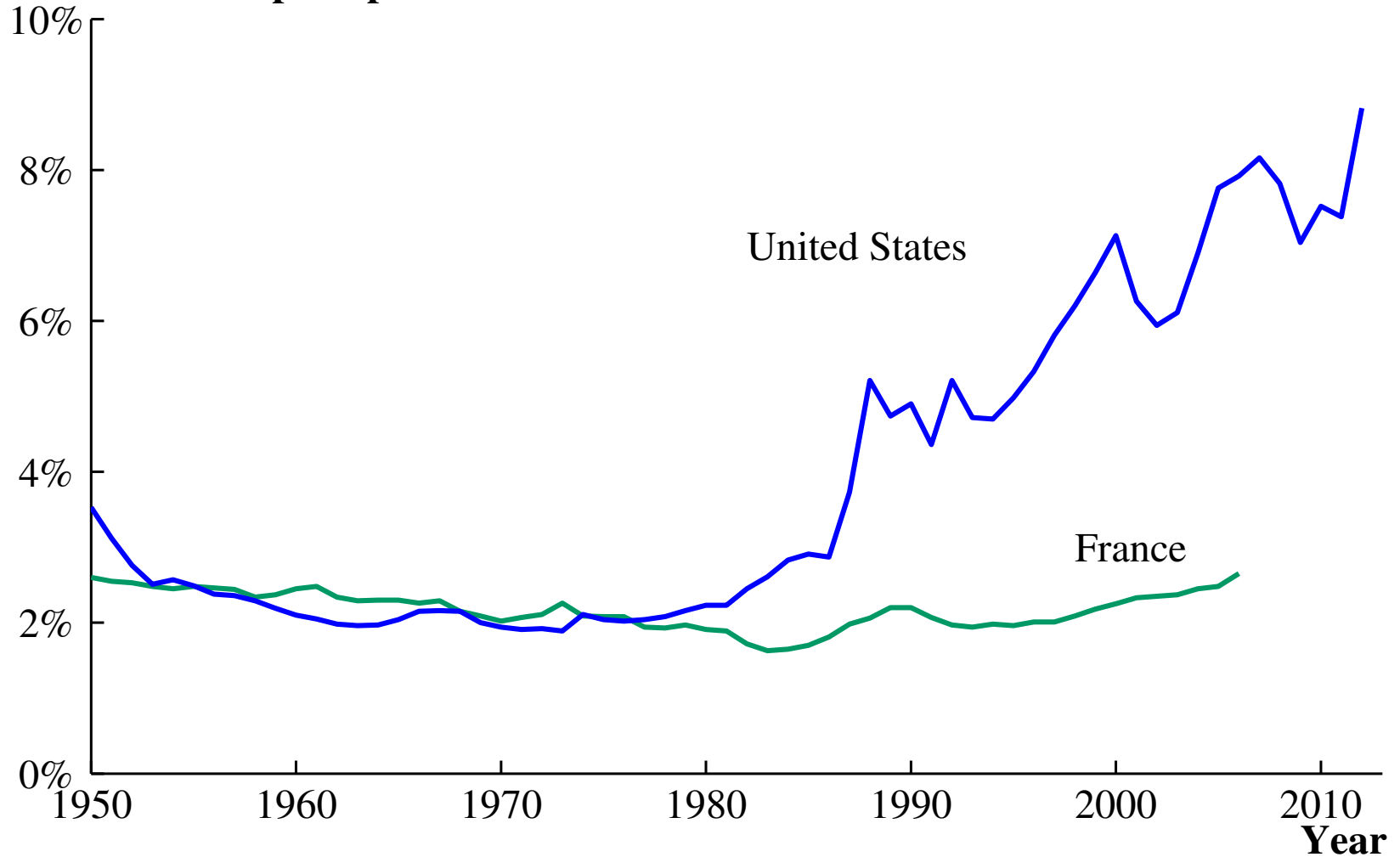
A Schumpeterian Model of Top Income Inequality

Chad Jones and Jihee Kim

SED – Warsaw 2015

Top Income Inequality in the United States and France

Income share of top 0.1 percent



Source: World Top Incomes Database (Alvaredo, Atkinson, Piketty, Saez)

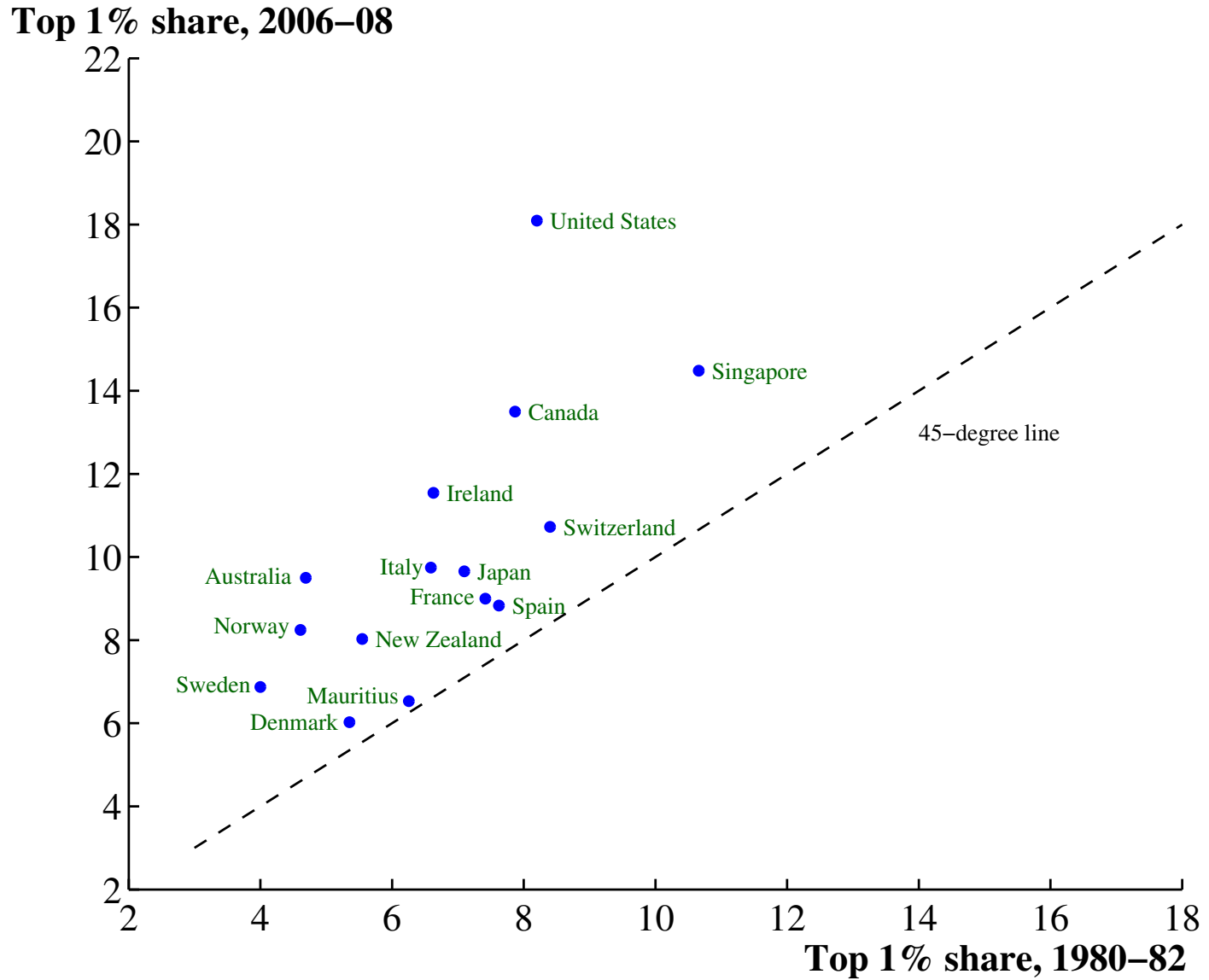
Related literature

- **Empirics:** Piketty and Saez (2003), Aghion et al (2015), Guvenen-Kaplan-Song (2015) and many more
- **Rent Seeking:** Piketty, Saez, and Stantcheva (2011) and Rothschild and Scheuer (2011)
- **Finance:** Philippon-Reshef (2009), Bell-Van Reenen (2010)
- **Not just finance:** Bakija-Cole-Heim (2010), Kaplan-Rauh
- **Pareto-generating mechanisms:** Gabaix (1999, 2009), Luttmer (2007, 2010), Reed (2001). GLLM (2015).
- **Use Pareto to get growth:** Kortum (1997), Lucas and Moll (2013), Perla and Tonetti (2013).
- **Pareto wealth distribution:** Benhabib-Bisin-Zhu (2011), Nirei (2009), Moll (2012), Piketty-Saez (2012), Piketty-Zucman (2014), Aoki-Nirei (2015)

Outline

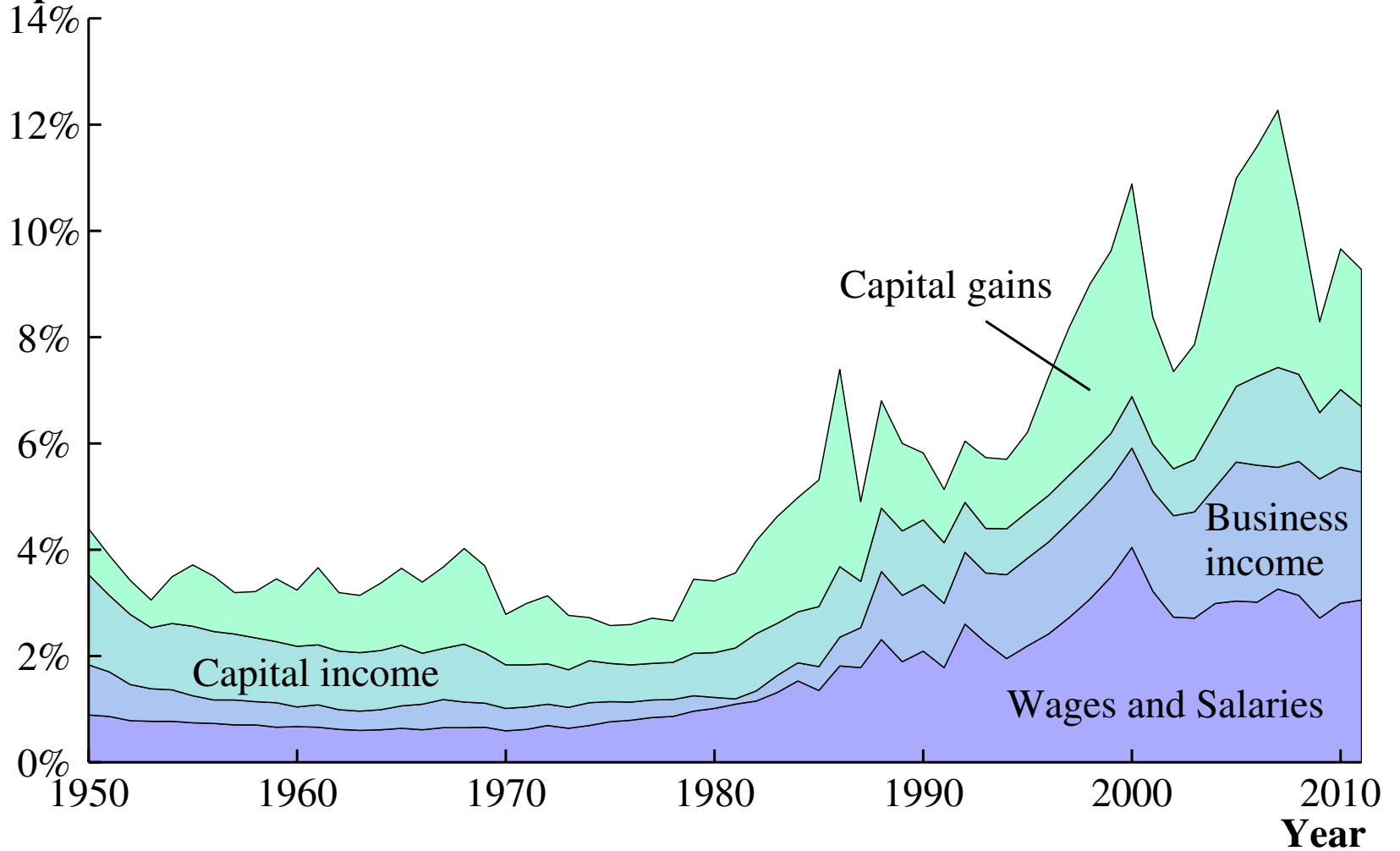
- Facts from World Top Incomes Database
- Simple model
- Full model
- Empirical work using IRS public use panel tax returns
- Numerical examples

Top Income Inequality around the World



The Composition of the Top 0.1 Percent Income Share

Top 0.1 percent income share



Pareto Distributions

$$\Pr [Y > y] = \left(\frac{y}{y_0} \right)^{-\xi}$$

- Let $\tilde{S}(p)$ = share of income going to the top p percentiles, and $\eta \equiv 1/\xi$ be a measure of **Pareto inequality**:

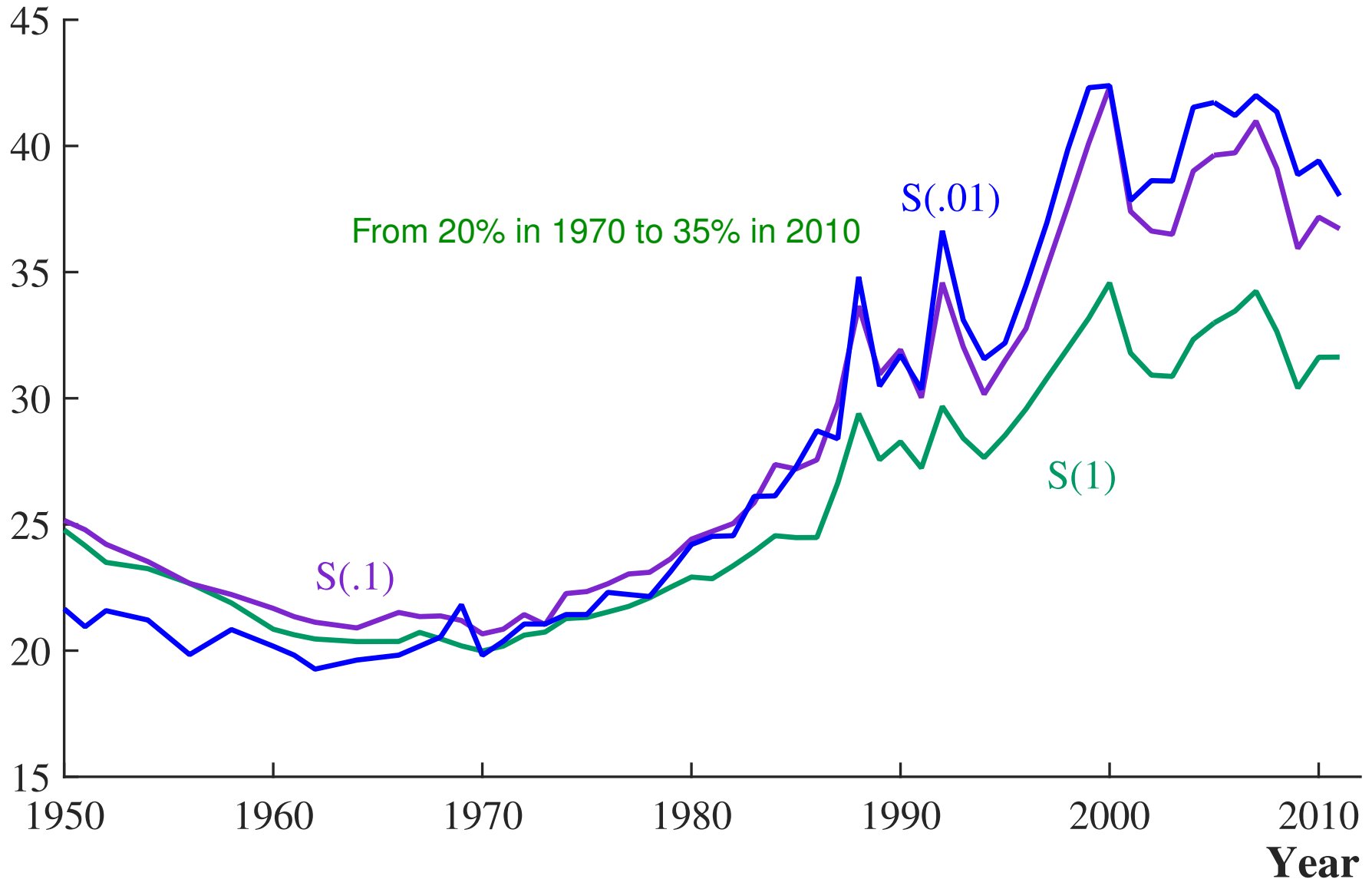
$$\tilde{S}(p) = \left(\frac{100}{p} \right)^{\eta-1}$$

- If $\eta = 1/2$, then share to Top 1% is $100^{-1/2} \approx .10$
- If $\eta = 3/4$, then share to Top 1% is $100^{-1/4} \approx .32$
- **Fractal**: Let $S(a)$ = share of $10a$'s income going to top a :

$$S(a) = 10^{\eta-1}$$

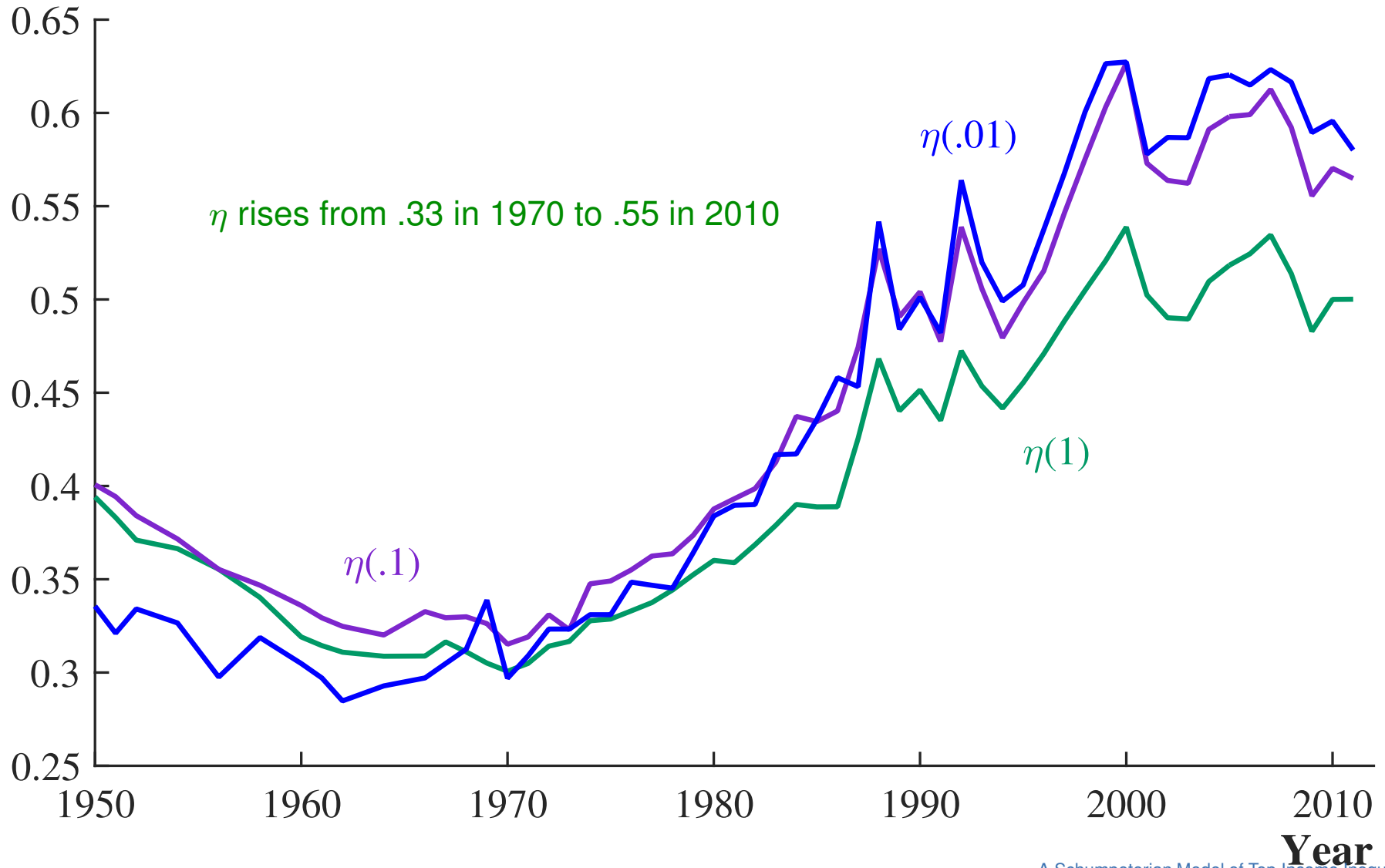
Fractal Inequality Shares in the United States

Fractal shares (percent)



The Power-Law Inequality Exponent η , United States

$1 + \log_{10}(\text{top share})$

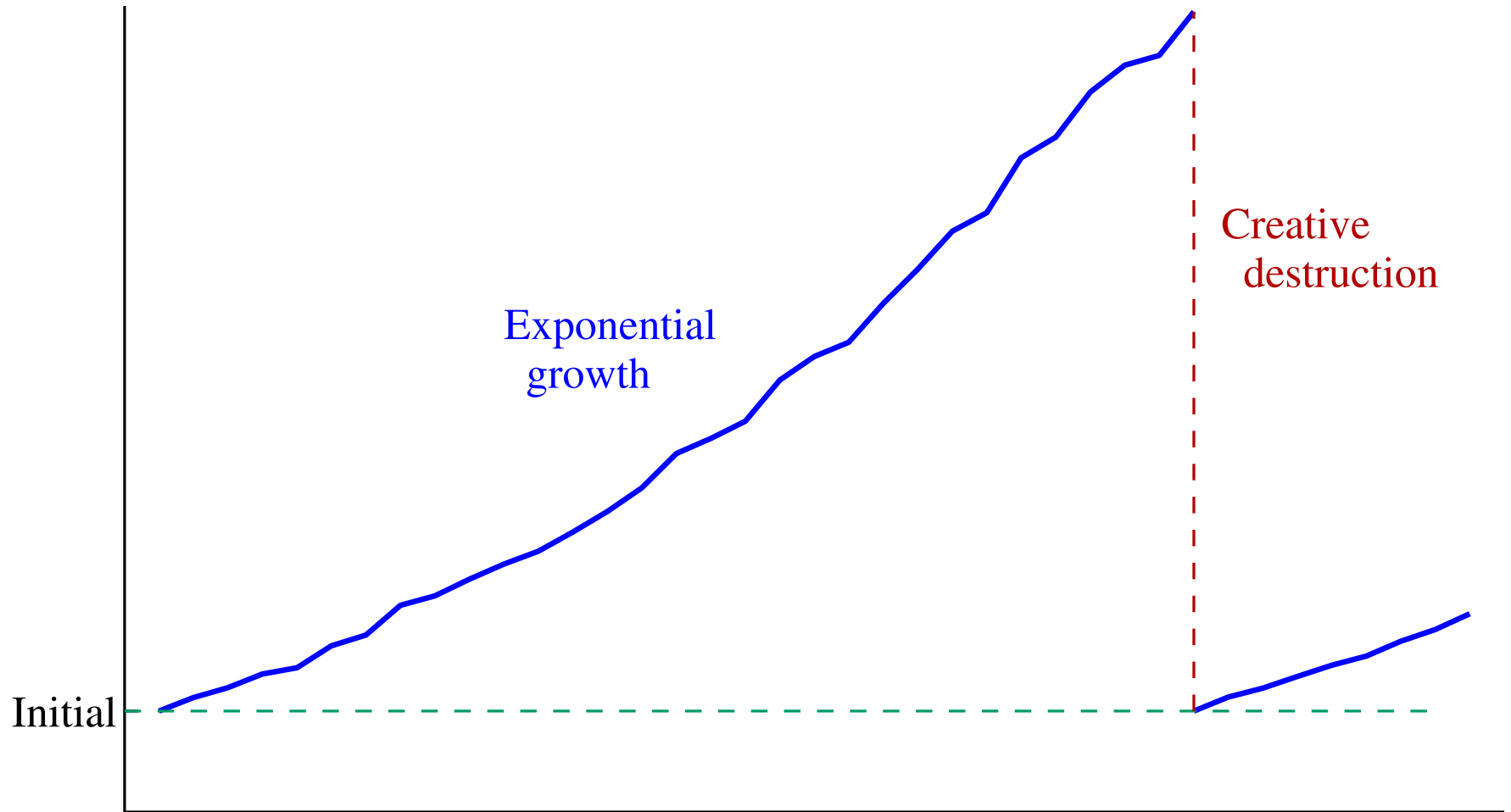




A Simple Model

Key Idea: Exponential growth w/ death \Rightarrow Pareto

Income



Time

Simple Model for Intuition

- Exponential growth often leads to a Pareto distribution.
- Entrepreneurs
 - New entrepreneur (“top earner”) earns y_0
 - Income after x years of experience:

$$y(x) = y_0 e^{\mu x}$$

- Poisson “replacement” process at rate δ
 - Stationary distribution of experience is exponential

$$\Pr [\text{Experience} > x] = e^{-\delta x}$$

What fraction of people have income greater than y ?

- Equals fraction with at least $x(y)$ years of experience

$$x(y) = \frac{1}{\mu} \log \left(\frac{y}{y_0} \right)$$

- Therefore

$$\begin{aligned} \Pr [\text{Income} > y] &= \Pr [\text{Experience} > x(y)] \\ &= e^{-\delta x(y)} \\ &= \left(\frac{y}{y_0} \right)^{-\frac{\delta}{\mu}} \end{aligned}$$

- So power law inequality is given by

$$\eta_y = \frac{\mu}{\delta}$$

Intuition

- Why does the Pareto result emerge?
 - Log of income \propto experience (Exponential growth)
 - Experience \sim exponential (Poisson process)
 - Therefore log income is exponential
 - \Rightarrow Income \sim Pareto!
- A Pareto distribution emerges from exponential growth experienced for an exponentially distributed amount of time.

Full model: endogenize μ and δ and how they change



The Model

- Pareto distribution in partial eqm
- GE with exogenous research
- Full general equilibrium

Entrepreneur's Problem

Choose $\{e_t\}$ to maximize expected discounted utility:

$$U(c, \ell) = \log c + \beta \log \ell$$

$$c_t = \psi_t x_t$$

$$e_t + \ell_t + \tau = 1$$

$$dx_t = \mu(e_t)x_t dt + \sigma x_t dB_t$$

$$\mu(e) = \phi e$$

- x = idiosyncratic productivity of a variety
- ψ_t = determined in GE (grows)
- δ = endogenous creative destruction
- $\bar{\delta}$ = exogenous destruction

Entrepreneur's Problem – HJB Form

- The Bellman equation for the entrepreneur:

$$\rho V(x_t, t) = \max_{e_t} \log \psi_t + \log x_t + \beta \log(\Omega - e_t) + \frac{\mathbb{E}[dV(x_t, t)]}{dt} + (\delta + \bar{\delta})(V^w(t) - V(x_t, t))$$

where $\Omega \equiv 1 - \tau$

- Note: the “capital gain” term is

$$\frac{\mathbb{E}[dV(x_t, t)]}{dt} = \mu(e_t)x_t V_x(x_t, t) + \frac{1}{2}\sigma^2 x_t^2 V_{xx}(x_t, t) + V_t(x_t, t)$$

Solution for Entrepreneur's Problem

- Equilibrium effort is constant:

$$e^* = 1 - \tau - \frac{1}{\phi} \cdot \beta(\rho + \delta + \bar{\delta})$$

- Comparative statics:
 - $\uparrow \tau \Rightarrow \downarrow e^*$: higher “taxes”
 - $\uparrow \phi \Rightarrow \uparrow e^*$: better technology for converting effort into x
 - $\uparrow \delta$ or $\bar{\delta} \Rightarrow \downarrow e^*$: more destruction

Stationary Distribution of Entrepreneur's Income

- Unit measure of entrepreneurs / varieties
- Displaced in two ways
 - Exogenous misallocation ($\bar{\delta}$): new entrepreneur $\rightarrow x_0$.
 - Endogenous creative destruction (δ): inherit existing productivity x .
- Distribution $f(x, t)$ satisfies **Kolmogorov forward equation:**

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

- Stationary distribution $\lim_{t \rightarrow \infty} f(x, t) = f(x)$ solves
$$\frac{\partial f(x, t)}{\partial t} = 0$$

- Guess that $f(\cdot)$ takes the Pareto form $f(x) = Cx^{-\xi-1} \Rightarrow$

$$\xi^* = -\frac{\tilde{\mu}^*}{\sigma^2} + \sqrt{\left(\frac{\tilde{\mu}^*}{\sigma^2}\right)^2 + \frac{2\bar{\delta}}{\sigma^2}}$$

$$\tilde{\mu}^* \equiv \mu(e^*) - \frac{1}{2}\sigma^2 = \phi(1 - \tau) - \beta(\rho + \delta^* + \bar{\delta}) - \frac{1}{2}\sigma^2$$

- Power-law inequality is therefore given by

$$\eta^* = 1/\xi^*$$

Comparative Statics (given δ^*)

$$\eta^* = 1/\xi^*, \quad \xi^* = -\frac{\tilde{\mu}^*}{\sigma^2} + \sqrt{\left(\frac{\tilde{\mu}^*}{\sigma^2}\right)^2 + \frac{2\bar{\delta}}{\sigma^2}}$$

$$\tilde{\mu}^* = \phi(1 - \tau) - \beta(\rho + \delta^* + \bar{\delta}) - \frac{1}{2}\sigma^2$$

- Power-law inequality η^* increases if
 - $\uparrow \phi$: better technology for converting effort into x
 - $\downarrow \delta$ or $\bar{\delta}$: less destruction
 - $\downarrow \tau$: Lower “taxes”
 - $\downarrow \beta$: Lower utility weight on leisure

Luttmer and GLLM

- Problems with basic random growth model:
 - Luttmer (2010): Cannot produce “rockets” like Google or Uber
 - Gabaix, Lasry, Lions, and Moll (2015): Slow transition dynamics
- Solution from Luttmer/GLLM:
 - Introduce heterogeneous mean growth rates: e.g. “high” versus “low”
 - Here: $\phi_H > \phi_L$ with Poisson rate \bar{p} of transition ($H \rightarrow L$)

Pareto Inequality with Heterogeneous Growth Rates

$$\eta^* = 1/\xi_H, \quad \xi_H = -\frac{\tilde{\mu}_H^*}{\sigma^2} + \sqrt{\left(\frac{\tilde{\mu}_H^*}{\sigma^2}\right)^2 + \frac{2(\bar{\delta} + \bar{p})}{\sigma^2}}$$

$$\tilde{\mu}_H^* = \phi_H(1 - \tau) - \beta(\rho + \delta^* + \bar{\delta}) - \frac{1}{2}\sigma^2$$

- This adopts Gabaix, Lasry, Lions, and Moll (2015)
- Why it helps quantitatively:
 - ϕ_H : Fast growth allows for Google / Uber
 - \bar{p} : Rate at which high growth types transit to low growth types raises the speed of convergence = $\bar{\delta} + \bar{p}$.

Growth and Creative Destruction

Final output

$$Y = \left(\int_0^1 Y_i^\theta di \right)^{1/\theta}$$

Production of variety i

$$Y_i = \gamma^{n_t} x_i^\alpha L_i$$

Resource constraint

$$L_t + R_t + 1 = \bar{N}, \quad L_t \equiv \int_0^1 L_{it} di$$

Flow rate of innovation

$$\dot{n}_t = \lambda(1 - \bar{z})R_t$$

Creative destruction

$$\delta_t = \dot{n}_t$$

Equilibrium with Monopolistic Competition

- Suppose $R/\bar{L} = \bar{s}$ where $\bar{L} \equiv \bar{N} - 1$.
- Define $X \equiv \int_0^1 x_i di = \frac{x_0}{1-\eta}$. Markup is $1/\theta$.

Aggregate PF

$$Y_t = \gamma^{n_t} X^\alpha L$$

Wage for L

$$w_t = \theta \gamma^{n_t} X^\alpha$$

Profits for variety i

$$\pi_{it} = (1 - \theta) \gamma^{n_t} X^\alpha L \left(\frac{x_i}{X} \right) \propto w_t \left(\frac{x_i}{X} \right)$$

Definition of ψ_t

$$\psi_t = (1 - \theta) \gamma^{n_t} X^{\alpha-1} L$$

Note that $\uparrow \eta$ has a level effect on output and wages.

Growth and Inequality in the \bar{s} case

- Creative destruction and growth

$$\delta^* = \lambda R = \lambda(1 - \bar{z})\bar{s}\bar{L}$$

$$g_y^* = \dot{n} \log \gamma = \lambda(1 - \bar{z})\bar{s}\bar{L} \log \gamma$$

- Does rising top inequality always reflect positive changes?
 - No! $\uparrow \bar{s}$ (more research) or $\downarrow \bar{z}$ (less innovation blocking)
 - Raise growth and reduce inequality via \uparrow creative destruction.



Endogenizing Research and Growth

Endogenizing $s = R/\bar{L}$

- Worker:

$$\rho V^w(t) = \log w_t + \frac{dV^W(t)}{dt}$$

- Researcher:

$$\begin{aligned} \rho V^R(t) = \log(\bar{m}w_t) + \frac{dV^R(t)}{dt} + \lambda (\mathbb{E}[V(x, t)] - V^R(t)) \\ + \bar{\delta}_R (V(x_0, t) - V^R(t)) \end{aligned}$$

- Equilibrium:

$$V^w(t) = V^R(t)$$

Stationary equilibrium solution

Drift of log x $\tilde{\mu}_H^* = \phi_H(1 - \tau) - \beta(\rho + \delta^* + \bar{\delta}) - \frac{1}{2}\sigma_H^2$

Pareto inequality $\eta^* = 1/\xi^*, \quad \xi^* = -\frac{\tilde{\mu}_H^*}{\sigma_H^2} + \sqrt{\left(\frac{\tilde{\mu}_H^*}{\sigma_H^2}\right)^2 + \frac{2(\bar{\delta} + \bar{p})}{\sigma_H^2}}$

Creative destruction $\delta^* = \lambda(1 - \bar{z})s^*\bar{L}$

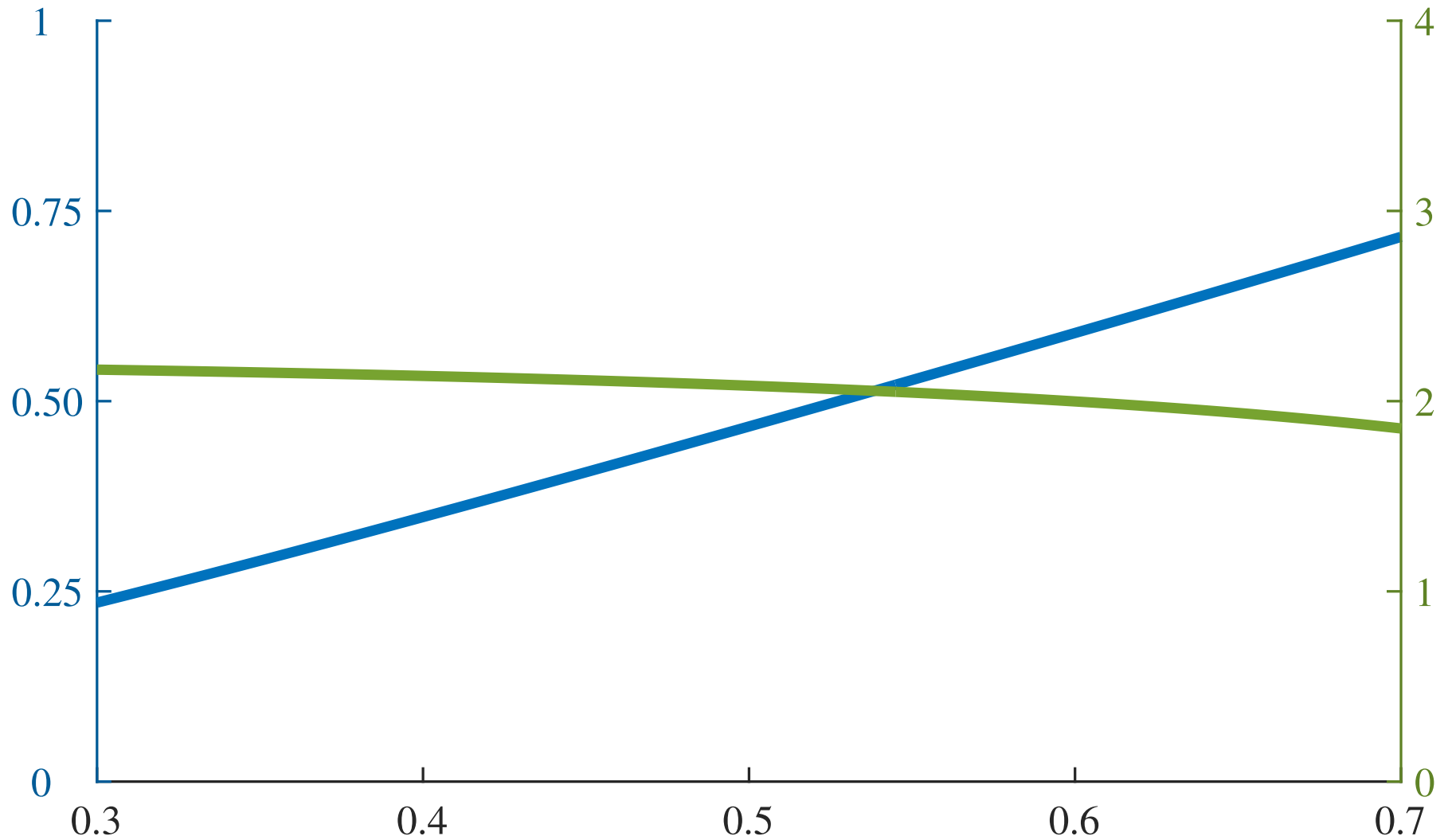
Growth $g^* = \delta^* \log \gamma$

Research allocation $V^w(s^*) = V^R(s^*)$

Varying the x-technology parameter ϕ

POWER LAW INEQUALITY

GROWTH RATE (PERCENT)



Why does $\uparrow \phi$ reduce growth?

- $\uparrow \phi \Rightarrow \uparrow e^* \Rightarrow \uparrow \mu^*$
- Two effects
 - **GE effect:** technological improvement \Rightarrow economy more productive so higher profits, but also higher wages
 - **Allocative effect:** raises Pareto inequality (η), so $\frac{x_i}{X}$ is more dispersed $\Rightarrow E \log \pi_i / w$ is lower. Risk averse agents undertake less research.
- Positive level effect raises both profits and wages. Riskier research \Rightarrow lower research and lower long-run growth.

Growth and Inequality

- Growth and inequality tend to move in **opposite** directions!
- Two reasons
 1. **Faster growth** \Rightarrow more creative destruction
 - Less time for inequality to grow
 - Entrepreneurs may work less hard to grow market
 2. **With greater inequality, research is riskier!**
 - Riskier research \Rightarrow less research \Rightarrow lower growth
- Transition dynamics \Rightarrow ambiguous effects on growth in medium run

Possible explanations: Rising U.S. Inequality

- Technology (e.g. WWW)
 - Entrepreneur's effort is more productive $\Rightarrow \uparrow \eta$
 - Worldwide phenomenon, not just U.S.
 - Ambiguous effects on U.S. growth (research is riskier!)
- Lower taxes on top incomes
 - Increase effort by entrepreneur's $\Rightarrow \uparrow \eta$

Possible explanations: Inequality in France

- Efficiency-reducing explanations
 - Delayed adoption of good technologies (WWW)
 - Increased misallocation (killing off entrepreneurs more quickly)
- Efficiency-enhancing explanations
 - Increased subsidies to research (more creative destruction)
 - Reduction in blocking of innovations (more creative destruction)

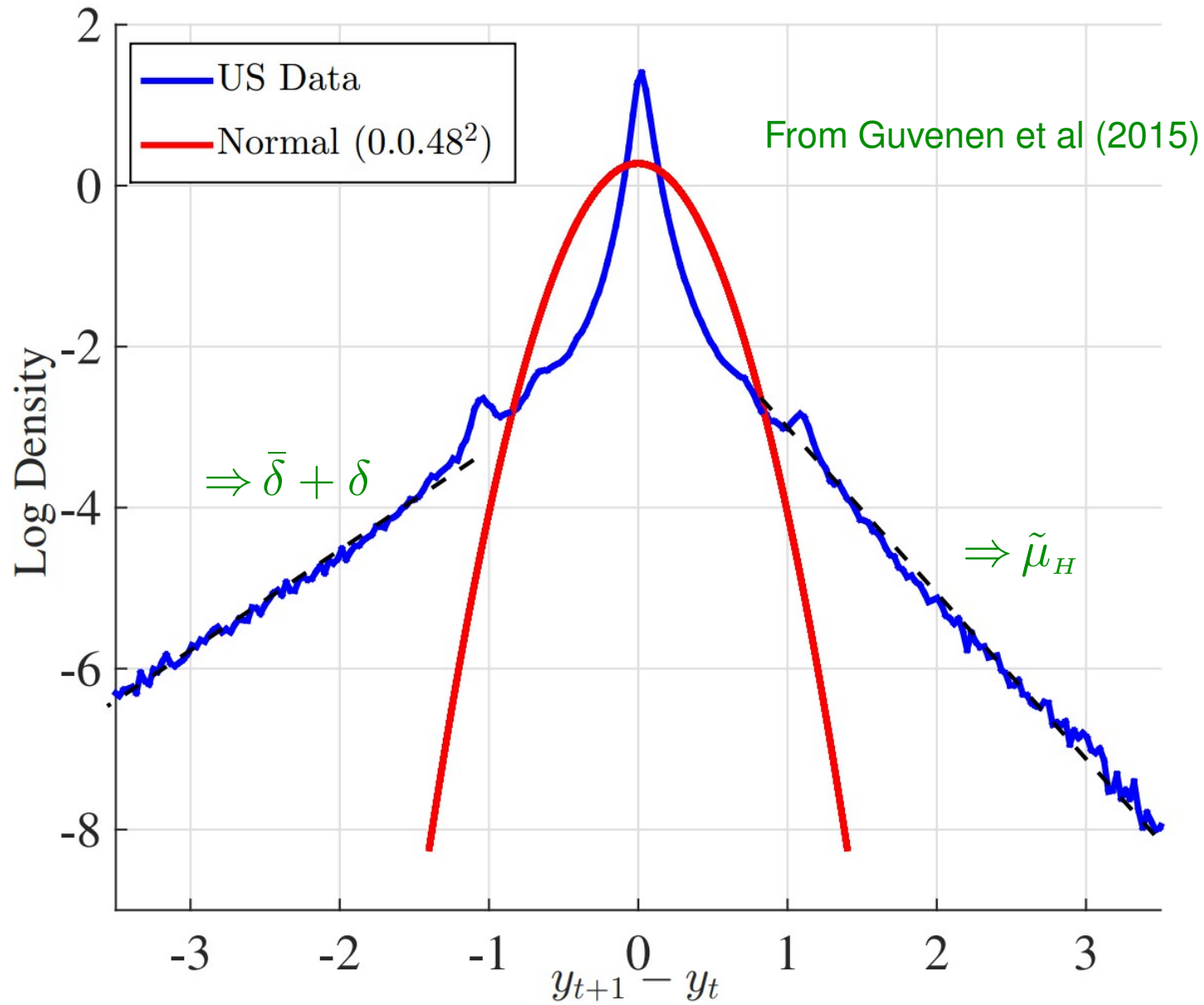


Micro Evidence

Overview

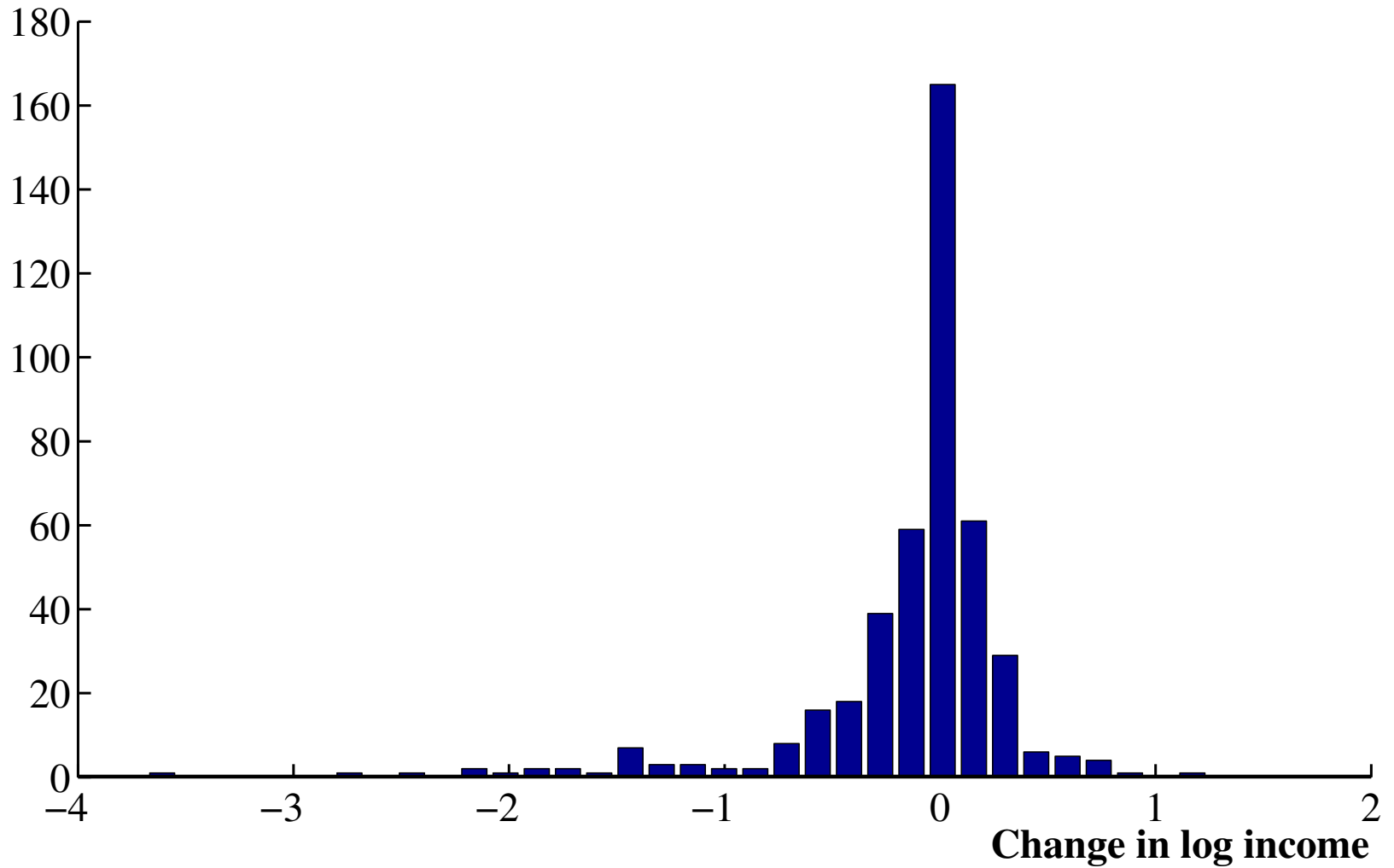
- Geometric random walk with drift = canonical DGP in the empirical literature on income dynamics.
 - Survey by Meghir and Pistaferri (2011)
- The distribution of growth rates for the Top 10% earners
 - Guvenen, Karahan, Ozkan, Song (2015) for 1995-96
 - IRS public use panel for 1979–1990 (small sample)

Pareto Tails of the Growth Rate Distribution



Growth Rates of Top 5% Incomes, 1988–1989

Number of observations



Results

Parameter	IRS 1979–81	IRS 1988–90	Guvenen et al. 1995–96
$\bar{\delta} + \delta$		0.07	...
σ_H		0.122	...
\bar{p}		0.767	...
$\tilde{\mu}_H$	0.244	0.303	0.435
Model: η^*	0.330	0.398	0.556
Data: η	0.33	0.48	0.55



Three numerical examples

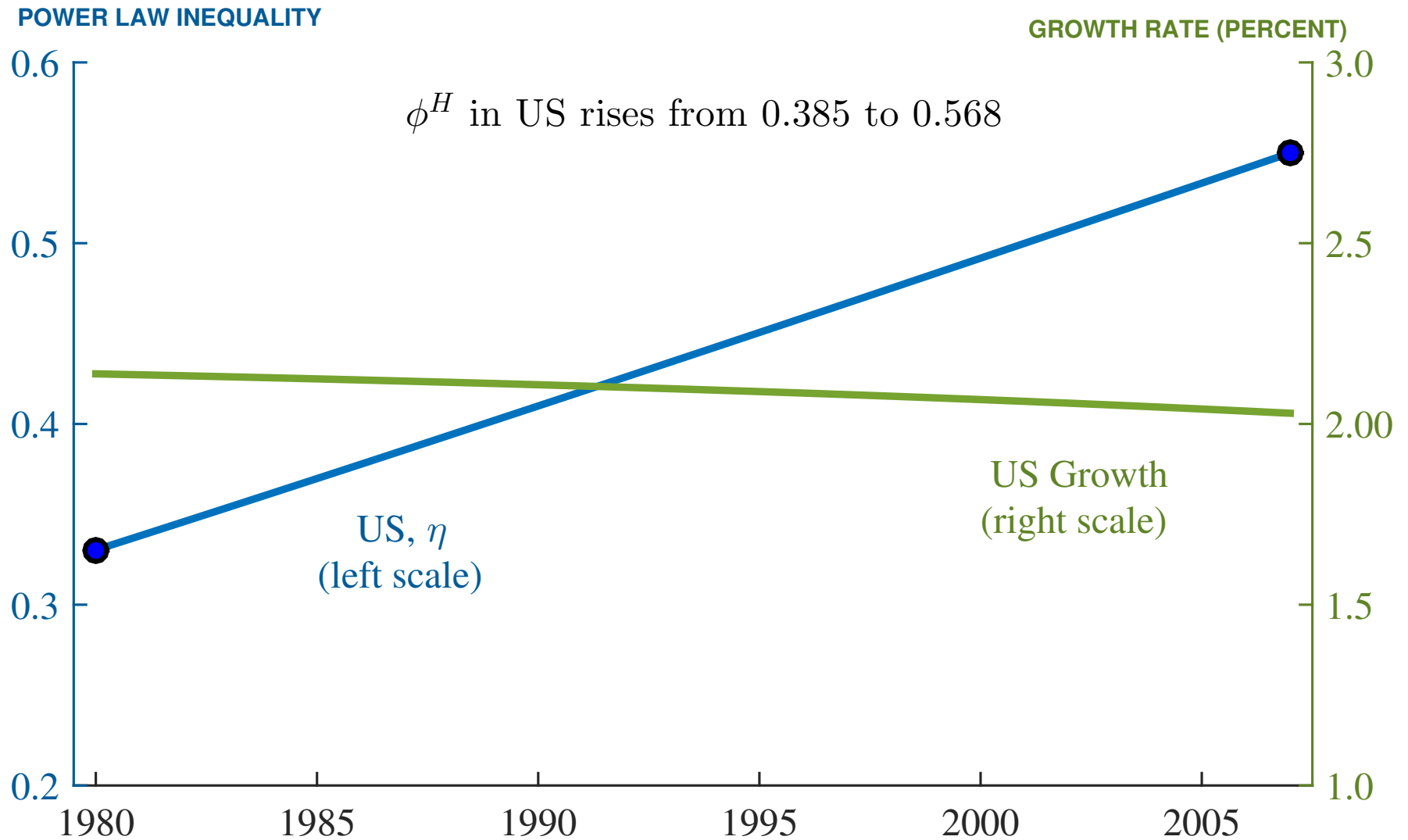
Three numerical examples

- The examples
 1. Match U.S. inequality 1980–2007 (ϕ)
 2. Match inequality in France (\bar{z}, \bar{p})
 3. Match U.S. and French data using taxes (τ)
- Why these are just examples
 - Identification problem: observe μ but not structural parameters, e.g. ϕ and τ
 - Sequence of steady states, not transition dynamics

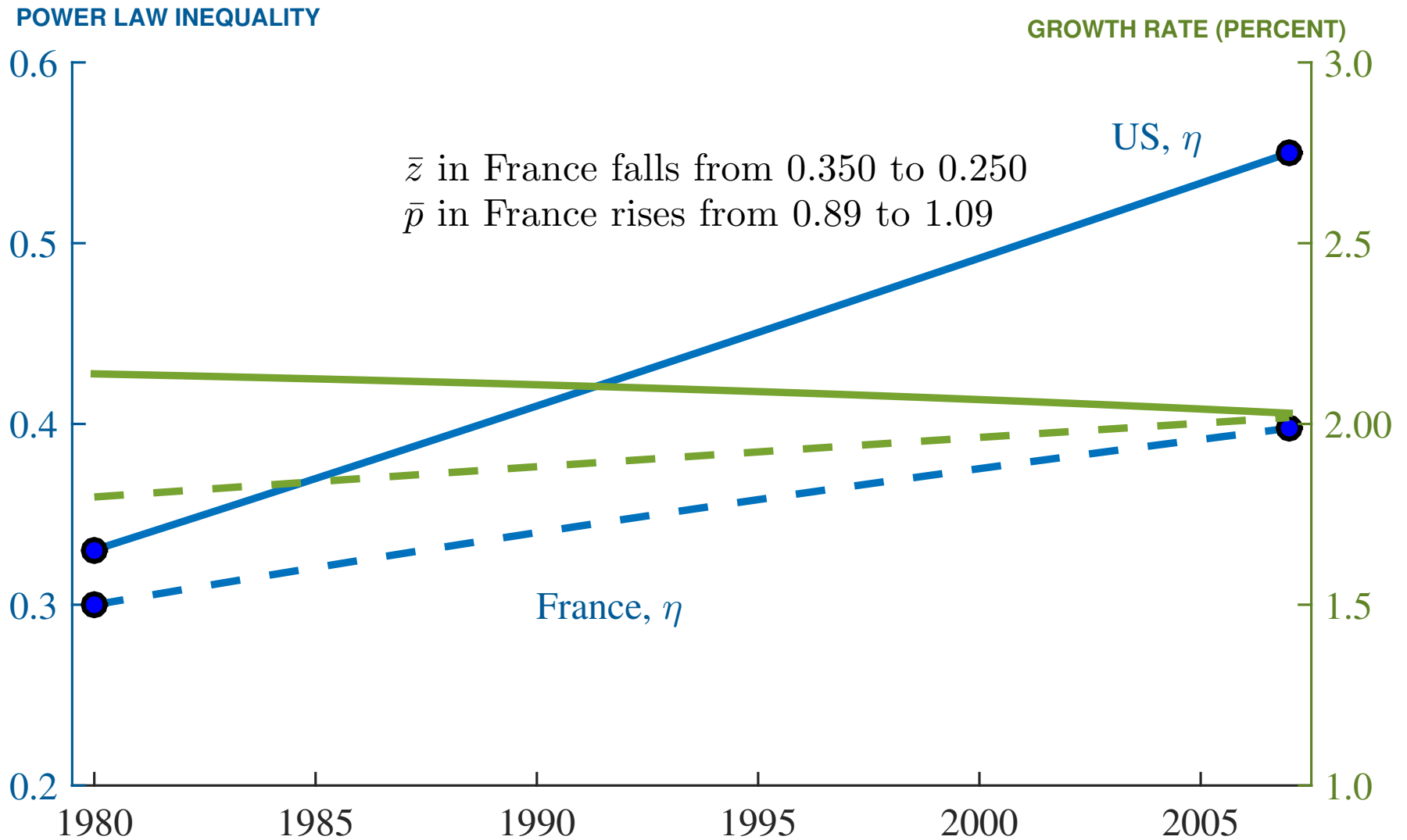
Parameters

- Parameters consistent with IRS panel:
 - $\phi \approx 0.5 \Rightarrow \tilde{\mu}_H \approx .3$
 - $\sigma_H = \sigma_L = .122$
 - $\bar{p} = 0.767$
 - $\bar{q} = .504$ – 2.5% of top earners are high growth
- Other parameter values
 - Match U.S. growth of 2% per year and Pareto inequality in 1980
 - $\bar{\delta} = 0.04$ and $\gamma = 1.4 \Rightarrow \delta + \bar{\delta} \approx 0.10$
 - $\rho = 0.03, \bar{L} = 15, \tau = 0, \theta = 2/3, \beta = 1, \lambda = 0.027, \bar{m} = 0.5, \bar{z} = 0.20$

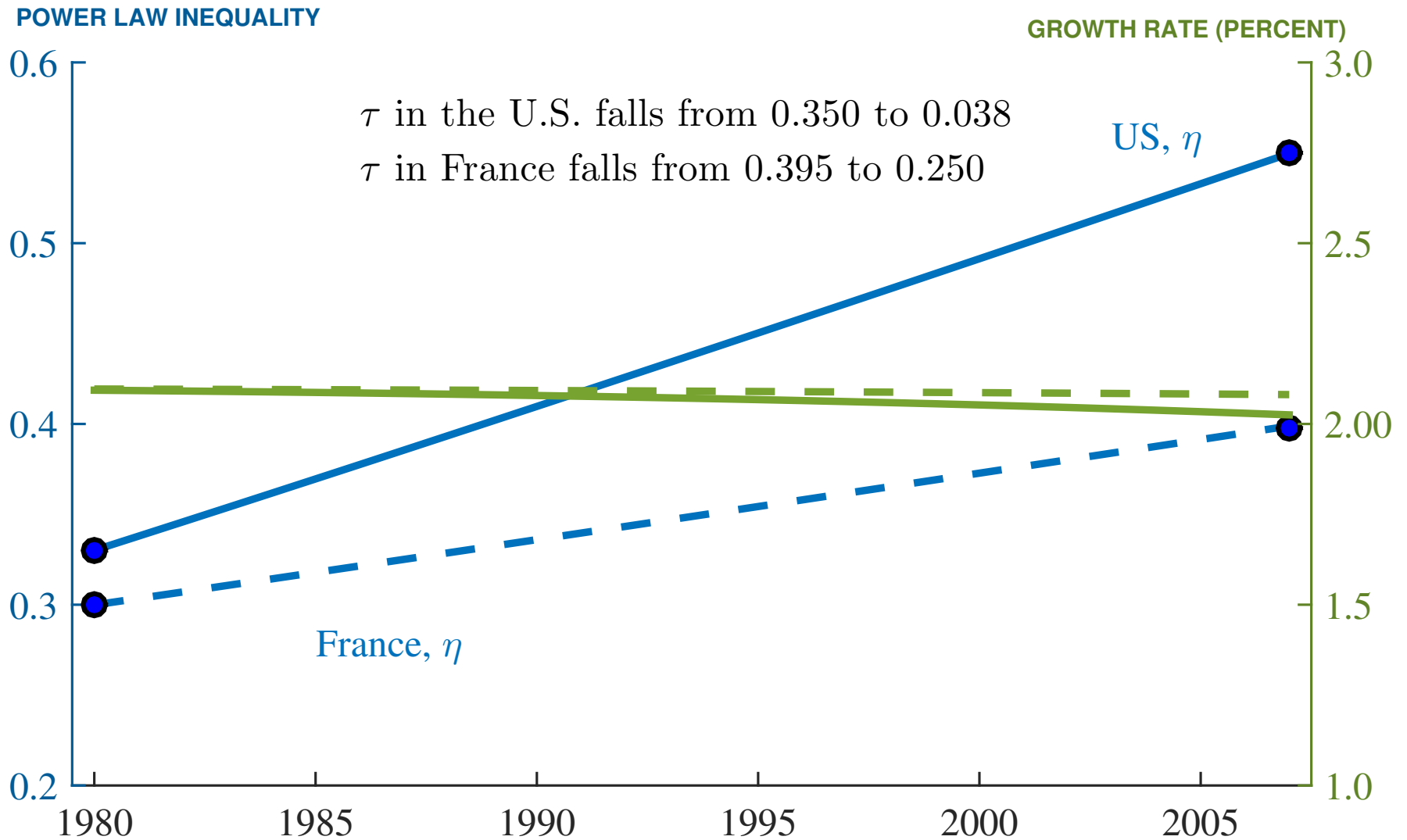
Numerical Example: Matching U.S. Inequality



Numerical Example: U.S. and France



Numerical Example: Taxes and Inequality



Conclusions: Understanding top income inequality

- Information technology / WWW:
 - Entrepreneurial effort is more productive: $\uparrow\phi \Rightarrow \uparrow\eta$
 - Worldwide phenomenon (?)
- Why else might inequality rise by less in France?
 - Less innovation blocking / more research: raises creative destruction
 - Regulations limiting rapid growth: $\uparrow\bar{p}$ and $\downarrow\phi$

Theory suggests rich connections between:
models of top inequality \leftrightarrow micro data on income dynamics

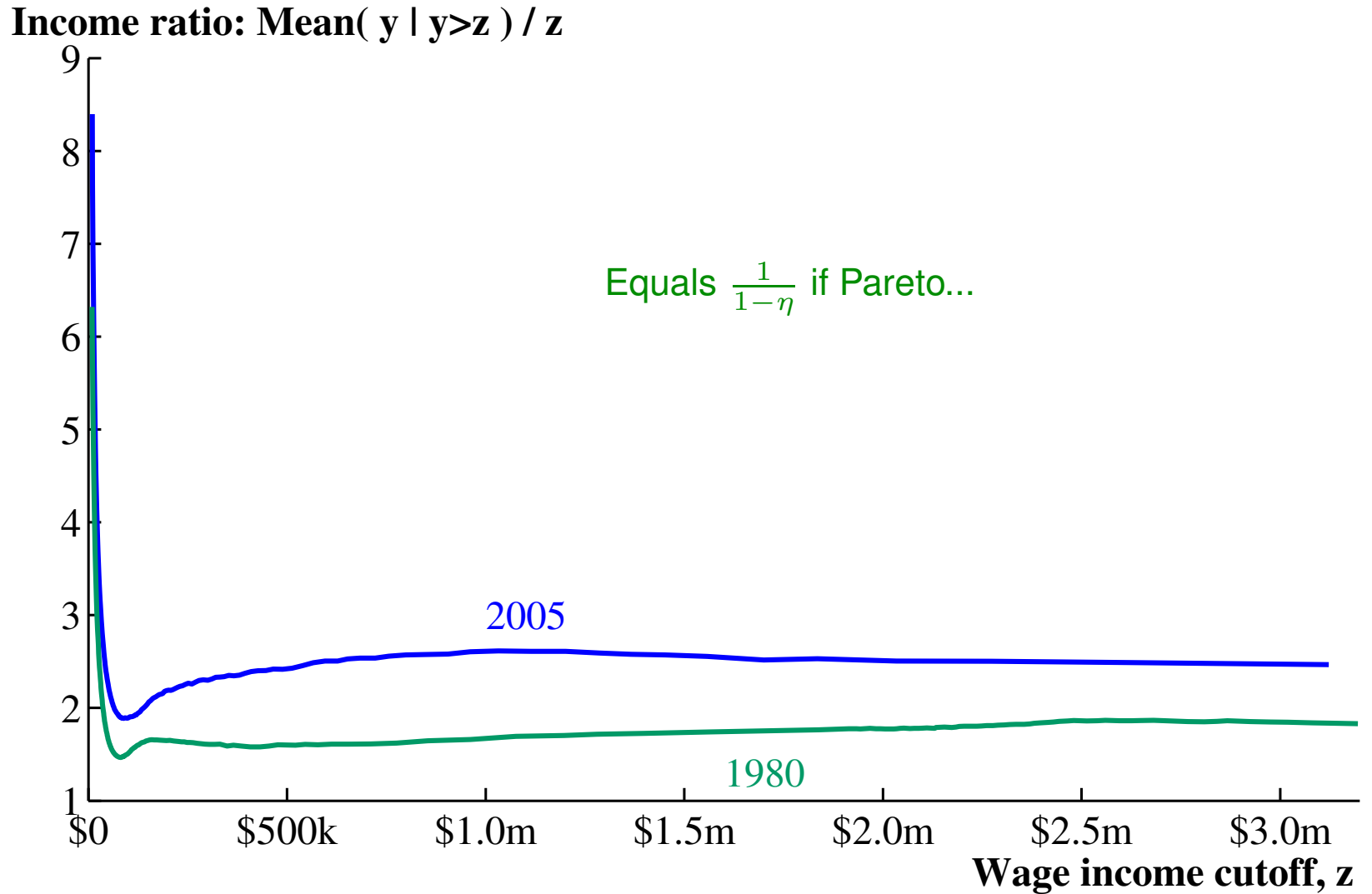


Extra Slides

Overview

- Atkinson / Piketty / Saez stylized facts on top income inequality
 - Rising sharply in US since 1980
 - More stable in France and Japan
 - Why?

The Pareto Nature of Labor Income



Skill-Biased Technical Change?

- Let $x_i = \text{skill}$ and $\bar{w} = \text{wage per unit skill}$

$$y_i = \bar{w}x_i^\alpha$$

- If $\Pr [x_i > x] = x^{-1/\eta_x}$, then

$$\Pr [y_i > y] = \left(\frac{y}{\bar{w}}\right)^{-1/\eta_y} \quad \text{where } \eta_y = \alpha\eta_x$$

- That is y_i is Pareto with inequality parameter η_y
 - SBTC ($\uparrow \bar{w}$) shifts distribution right but η_y unchanged.
 - $\uparrow \alpha$ would raise Pareto inequality...
 - This paper: **why** is $x \sim \text{Pareto}$, and **why** $\uparrow \alpha$

Why is experience exponentially distributed?

- Let $F(x, t)$ denote the distribution of experience at time t
- How does it evolve over discrete interval Δt ?

$$F(x, t + \Delta t) - F(x, t) = \underbrace{\delta \Delta t (1 - F(x, t))}_{\text{inflow from above } x} - \underbrace{[F(x, t) - F(x - \Delta x, t)]}_{\text{outflow as top folks age}}$$

- Dividing both sides by $\Delta t = \Delta x$ and taking the limit

$$\frac{\partial F(x, t)}{\partial t} = \delta(1 - F(x, t)) - \frac{\partial F(x, t)}{\partial x}$$

- Stationary: $F(x)$ such that $\frac{\partial F(x, t)}{\partial t} = 0$. Integrating gives the exponential solution.

How the model works

- $\uparrow \phi$ raises top inequality, but leaves the growth rate of the economy unchanged.
 - Surprising: a “linear differential equation” for x .
- Key: the distribution of x is stationary!
- Higher ϕ has a positive level effect through higher inequality, raising everyone’s wage.
 - But growth comes via research, not through x ...

Lucas at “micro” level, Romer/AH at “macro” level