Lecture 2

Key Facts on Income and Wealth Distribution

Distributional Macroeconomics

Part II of ECON 2149

Benjamin Moll

Harvard University, Spring 2018

A Budget Constraint to Organize our Thoughts

Want to think about

1. inequality of labor income
2. inequality of capital income
3. wealth inequality
4. consumption inequality
5. distribution of factor income (capital vs labor share)

A Budget Constraint to Organize our Thoughts

- N households indexed by $i = 1, ..., N$, discrete time $t = 0, 1, 2...$

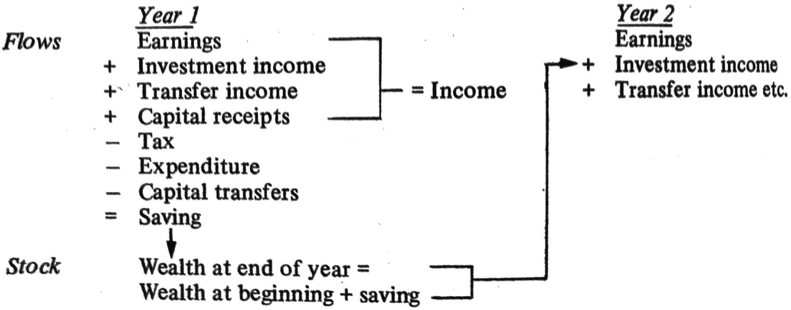
$$c_{it} + s_{it} = \underbrace{y_{it}^{\ell} + y_{it}^{k}}_{y_{it}}, \qquad a_{it+1} = s_{it} + a_{it}$$

$$\Rightarrow \quad a_{it+1} = \underbrace{y_{it}^{\ell} + y_{it}^{k}}_{y_{it}} + a_{it} - c_{it}$$

- y_{it}: total household income
- y_{it}^{ℓ}: labor income
- y_{it}^{k}: capital income
- c_{it}: consumption
- s_{it}: saving
- a_{it}: wealth

- Usual budget costraint = special case with $y_{it}^{\ell} = w_t \ell_{it}$, $y_{it}^{k} = r_t a_{it}$
- Power of above budget constraint: accounting identity
- Remark: nothing special about discrete time
  - could have also written $a_{i,t+1} = \int_0^1 s_{i,t+\tau} d\tau + a_{i,t}$
  - real world: continuous time, data sampled at discrete intervals

A Budget Constraint to Organize our Thoughts



Source: Atkinson (1975), “The Economics of Inequality”

Why useful?

- Aids clarity of thinking
- Consider following questions
  - when income inequality increases, do we expect wealth inequality to increase as well?
  - If so, will this happen simultaneously or with some lag?
- More later: personal vs factor income distribution
  - When will an increase in the capital share result in an increase in inequality?

Measuring Inequality

Measuring inequality

- Visualizing distributions: some key concepts you should know

    1. density
    2. cumulative distribution function
    3. quantile function
    4. Lorenz curve

- Some commonly used summary statistics (but always keep in mind: impossible to summarize distribution with one number)

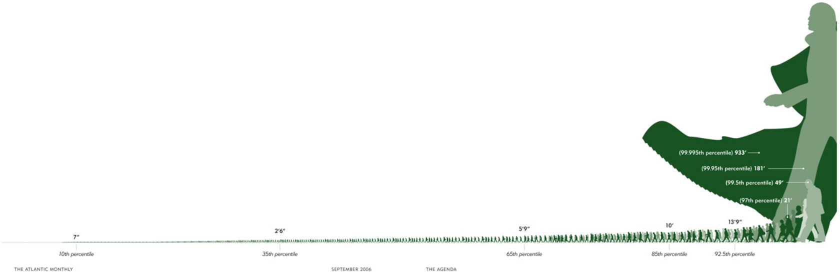
    1. 90-10 ratio, interquartile range and other percentile ratios
    2. top shares
    3. Gini coefficient

Quantile Function

- Quantile function = inverse of CDF

$$y(p) := F^{-1}(p), \quad F(y) := \Pr(y_{it} \leq y)$$

- Pen's parade:



Source: `http://www.theatlantic.com/magazine/archive/2006/09/the-height-of-inequality/305089/`

Lorenz Curve

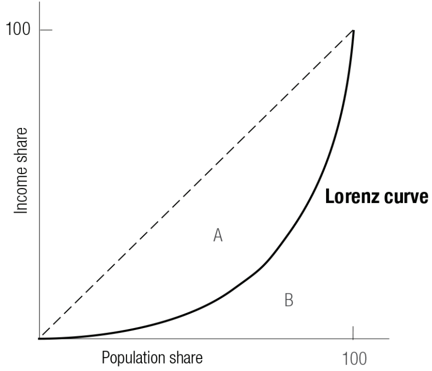


Figure 4. **Example of Lorenz curve for income.**

- $L(p)$:=share of total income going to bottom p%
- Relationship between Lorenz curve and quantile function

$$L'(p) = y(p)/\bar{y}$$

Atkinson's Theorem: Lorenz Dominance and Welfare

- Main message: if Lorenz curves for two distributions do not cross ("Lorenz dominance"), can rank them in terms of welfare
- Consider an income distribution F with density f
- For any u with $u' > 0$, $u'' < 0$, define welfare criterion

$$W(F) := \int_0^{\bar{y}} u(y) f(y) dy$$

- Theorem (Atkinson, 1970): Let F and G be two income dist'ns with equal means. Then F generates higher welfare than G if and only if the Lorenz curve for F lies everywhere above that for G:

$$W(F) \geq W(G) \quad \Leftrightarrow \quad L_F(p) \geq L_G(p) \quad \text{all } p \in [0, 1]$$

- Easy to extend to unequal means – Shorrocks (Ecma, 1983)
- Proof in two steps
  1. Lorenz dominance $\Leftrightarrow$ 2nd-order stochastic dominance
  2. 2nd-order stochastic dominance $\Leftrightarrow$ welfare ranking

Step 1 of proof: Lorenz dominance $\Leftrightarrow$ SOSD

Lemma 1: Let F and G be two income distributions with equal means. Then $L_F(p) \geq L_G(p)$, all $p \in [0,1] \Leftrightarrow \int_0^y [F(x) - G(x)]dx \leq 0$ for all y

Proof of Lemma 1 ($\Rightarrow$ part, see Atkinson (1970) for $\Leftarrow$ part):

- Denote mean by μ, pth quantile by $y_F(p)$, i.e. $F(y_F(p)) = p$. Have

$$L_F(p) := \frac{1}{\mu} \int_0^{y_F(p)} yf(y)dy$$

- Integrate by parts $\mu L_F(p) = y_F(p)p - \int_0^{y_F(p)} F(y)dy$
- Compare L_F and L_G at point p – WOLG assume $y_F(p) \leq y_G(p)$

$$\mu[L_F(p) - L_G(p)] = [y_F(p) - y_G(p)]p - \left[\int_0^{y_F(p)} F(y)dy - \int_0^{y_G(p)} G(y)dy\right]$$

$$= -\int_0^{y_G(p)} [F(y) - G(y)]dy + \left[\int_{y_F(p)}^{y_G(p)} F(y)dy - (y_G(p) - y_F(p))F(y_F(p))\right]$$

- Mean value theorem: $\int_{y_F(p)}^{y_G(p)} F(y)dy = (y_G(p) - y_F(p))F(\hat{y})$ for some $\hat{y} \in [y_F(p), y_G(p)] \Rightarrow$ 2nd term $\geq 0 \Rightarrow \mu[L_F(p) - L_G(p)] \geq 0$

Step 2 of proof: SOSD $\Leftrightarrow$ welfare ranking

Lemma 2: Let F and G be two income distributions. Then $W(F) \geq W(G) \Leftrightarrow \int_0^y [F(x) - G(x)]dx \leq 0$ for all $y \in [0, \bar{y}]$

Proof of Lemma 2 ($\Leftarrow$ part, see risk aversion literature for $\Rightarrow$ part):

$$\begin{aligned} W(F) - W(G) &= \int_0^{\bar{y}} u(y)f(y)dy - \int_0^{\bar{y}} u(y)g(y)dy \\ &= \int_0^{\bar{y}} u'(y)[G(y) - F(y)]dy \\ &= -\int_0^{\bar{y}} u''(y)S(y)dy + u'(\bar{y})S(\bar{y}) \end{aligned}$$

$$\text{where} \quad S(y) := -\int_0^{y} [F(x) - G(x)]dx$$

- From 2nd-order stochastic dominance $S(y) \geq 0$ for all y
- Further $u' > 0$, $u'' < 0$ for all y by assumption
- Hence $W(F) - W(G) \geq 0$

Publicly Available Data Sources for U.S.

- Survey of Consumer Finances (SCF)

  `http://www.federalreserve.gov/econresdata/scf/scfindex.htm`

- Panel Study of Income Dynamics (PSID)

  `https://psidonline.isr.umich.edu/`

- Consumer Expenditure Survey (CEX)

  `http://www.bls.gov/cex/`

- Current Population Survey (CPS)

  `http://www.census.gov/programs-surveys/cps.html`

- IRS public use tax model data (Piketty-Saez), through NBER

  `http://www.nber.org/taxsim-notes.html`, `http://users.nber.org/~taxsim/gdb/`

- for features, pros and cons of these see Gianluca Violante's lecture notes "Micro Data: A Helicopter Tour" `http://www.econ.nyu.edu/user/violante/NYUTeaching/QM/Fall15/Lectures/Lecture2_Data.pdf`

Other countries or other variables

- World Wealth and Income Database (Piketty-Saez top shares)
  http://www.wid.world/
- ECB Household Finance and Consumption Survey (HFCS)
  https://www.ecb.europa.eu/pub/economic-research/research-networks/html/researcher_hfcn.en.html
- Luxembourg Income Study Database
  http://www.lisdatacenter.org/our-data/lis-database/
- IPUMS International (household-level micro data from around the world): https://international.ipums.org/international/
- Execucomp (Executive Compensation)
  https://wrds-web.wharton.upenn.edu/wrds/ds/execcomp/exec.cfm
  http://www.anderson.ucla.edu/rosenfeld-library/databases/business-databases-by-name/execucomp
- Billionaire Characteristics Database
  http://www.iie.com/publications/interstitial.cfm?ResearchID=2917

Administrative Data

- If you want to work in this area, may want to try to get your hands on some administrative data
  - large samples, long panels, whole population, not top coded
  - though other issues, e.g. if tax data $\Rightarrow$ attempts at tax evasion
- U.S.: hard to get access
  - IRS: see papers by Chetty, Saez, Hendren, Stantcheva,...
  - SSA: see papers by Jae Song & co
  - exception with easy access is IRS public use tax model data
- May want to go to other countries (world $\neq$ just U.S.!)
  - Norway has a wealth tax and Denmark, Sweden used to $\Rightarrow$ have administrative wealth data in addition to income data

Income Inequality in U.S.

Income Concepts, Individuals vs Households

Earnings of Person 1

+

Earnings of Person 2

+

Income from capital

+

Private transfers

= Household market income

+

State transfers

=

Household gross income

−

Direct taxes

=

Household disposable income

divided by number of equivalent adults = Household equivalised disposable income

Disposable income + Value of public services

=

Household extended income

FIGURE 1.5: Guide to household income

Source: Atkinson (2015), "Inequality: What Can Be Done?"

U.S. Income Distribution

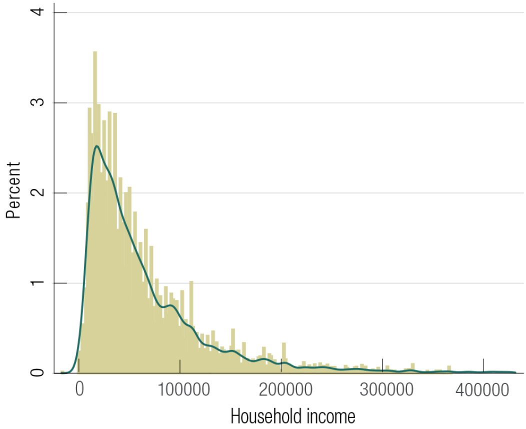


Figure 6. **Histogram of the 2013 income distribution (2013 USD).**

Source: Kuhn and Rios-Rull (2016)

U.S. Income Distribution

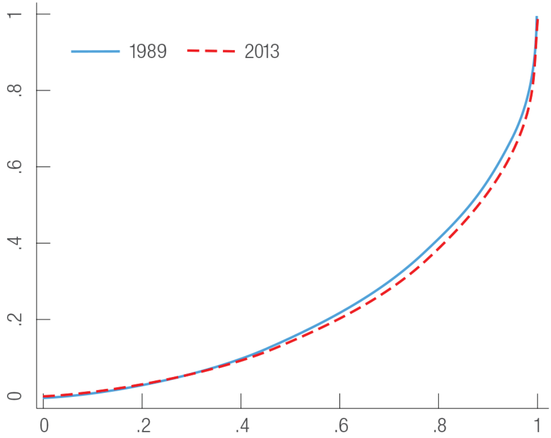


Figure 5. **Lorenz curves of income in 1989 and 2013.**

Source: Kuhn and Rios-Rull (2016)

Evolution of Household Income Distribution in U.S.

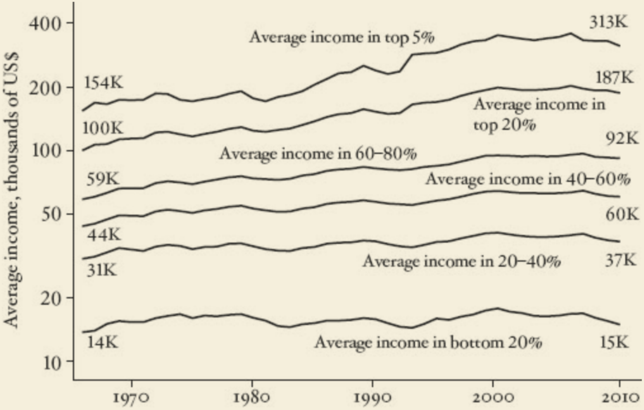
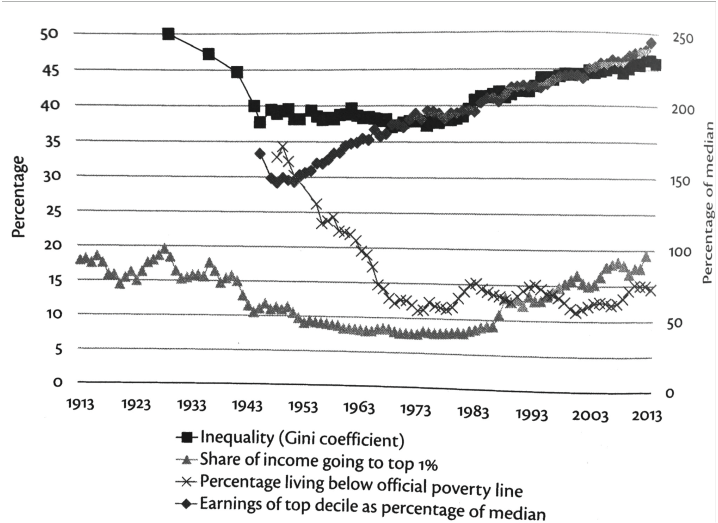


FIGURE 3 The distribution of family income in the United States.

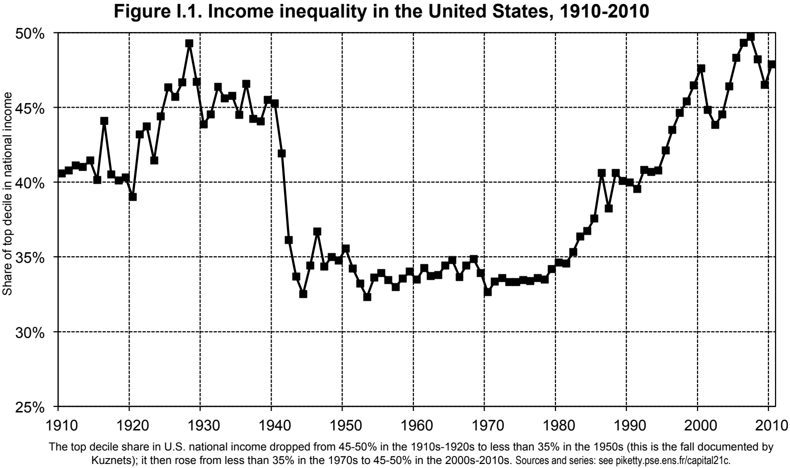
Source: Deaton (2015), “The Great Escape”

Evolution of Household Income Distribution in U.S.



Source: Atkinson (2015), “Inequality: What Can Be Done?”

Evolution of Top 10% Income Share in U.S.



Source: `http://piketty.pse.ens.fr/en/capital21c2`

Evolution of Household Income Distribution in U.S.

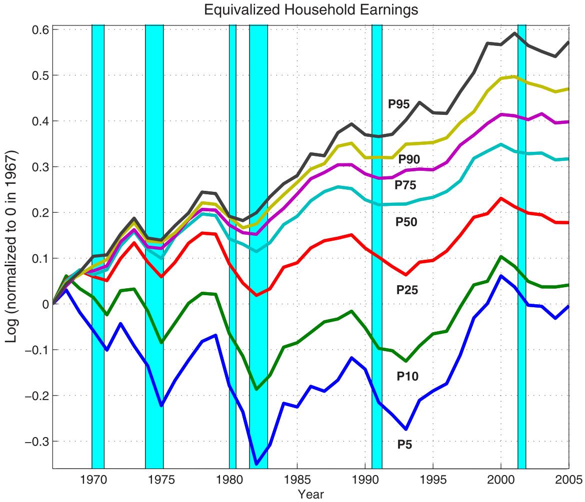


Fig. 9. Percentiles of the household earnings distribution (CPS). Shaded areas are NBER recessions.

Source: Heathcote-Perri-Violante (2010), "Unequal We Stand..."

Other Countries

See `https://ourworldindata.org/incomes-across-the-distribution/`

Inequality in the tails: back to the roots...

- ... more precisely 1896 and

- In 1896, Vilfredo Pareto examined income and wealth distribution across Europe
  - published “Cours d’économie politique”, for whole book see `http://www.institutcoppet.org/2012/05/08/cours-deconomie-politique-1896-de-vilfredo-pareto/`
  - relevant part `http://www.princeton.edu/~moll/pareto.pdf`

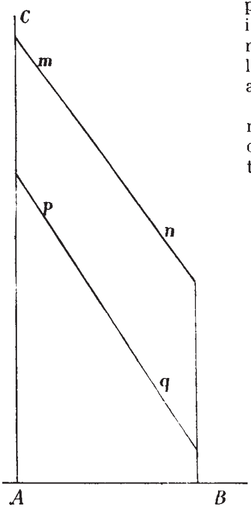


Fig. 47.

poser en ligne droite [1]. Disons immédiatement que nous allons retrouver cette tendance dans les nombreux exemples que nous aurons encore à examiner.

Un autre fait, tout aussi, et même plus remarquable, c'est que les courbes de la répartition des revenus, en Angleterre

Schedule D — Année 1893-94.

x £	N GREAT BRITAIN	N IRELAND
150	400 648	17 717
200	234 185	9 365
300	121 996	4 592
400	74 041	2 684
500	54 419	1 898
600	42 072	1 428
700	34 269	1 104
800	29 311	940
900	25 033	771
1000	22 896	684
2000	9 880	271
3000	6 069	142
4000	4 161	88
5000	3 081	68
10000	1 104	22

et en Irlande, présentent un parallélisme à peu près complet. Ce fait est à rapprocher d'un autre, que nous allons bientôt constater : les inclinaisons des lignes mn, pq obtenues pour dif-

(958) 1 C'est à dire que la courbe réelle est interpolée par une droite

Power Laws

- Pareto (1896): upper-tail distribution of number of people with an income or wealth X greater than a large x is proportional to $1/x^{\zeta}$ for some $\zeta > 0$
$$\Pr(X > x) = kx^{-\zeta}$$
- Definition 1: x follows a power law (PL) if there exist $k, \zeta > 0$ s.t.
$$\Pr(X > x) = kx^{-\zeta}, \quad \text{all } x$$
- x follows a PL $\Leftrightarrow$ x has a Pareto distribution
- Definition 2: x follows an asymptotic power law if there exist $k, \zeta > 0$ s.t.
$$\Pr(X > x) \sim kx^{-\zeta} \quad \text{as} \quad x \to \infty$$
- Note: for any f, g $f(x) \sim g(x)$ means $\lim_{x\to\infty} f(x)/g(x) = 1$
- Surprisingly many variables follow power laws, at least in tail
  - see Gabaix (2009), "Power Laws in Economics and Finance," very nice, very accessible

Power Laws

- Another way of saying same thing: top inequality is fractal
  - ... top 0.01% is M times richer than top 0.1%,... is M times richer than top 1%,... is M times richer than top 10%,...
  - to see this, note that top p percentile x_p satisfies

$$kx_p^{-\zeta} = p/100 \quad \Rightarrow \quad \frac{x_{0.01}}{x_{0.1}} = \frac{x_{0.1}}{x_1} = ... = 10^{1/\zeta}$$

  - average income/wealth above pth percentile is

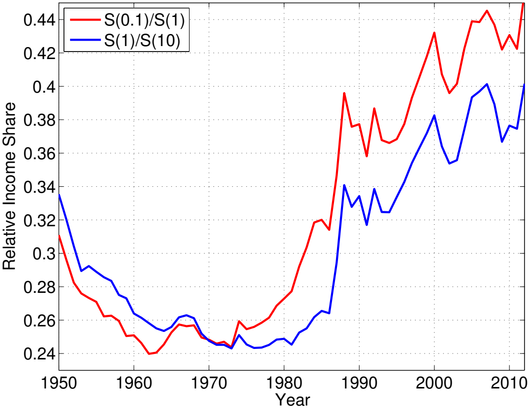
$$\bar{x}_p = \mathbb{E}[x|x \geq x_p] = \frac{\int_{x_p}^{\infty} x\zeta k x^{-\zeta-1} dx}{kx_p^{-\zeta}} = \frac{\zeta}{\zeta - 1} x_p \quad \Rightarrow$$

$$\frac{\bar{x}_{0.01}}{\bar{x}_{0.1}} = \frac{\bar{x}_{0.1}}{\bar{x}_1} = ... = 10^{1/\zeta}$$

- Related result: if x has a Pareto distribution, then share of x going to top p percent is

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

The income distribution's tail has gotten fatter



- $\frac{S(0.1)}{S(1)}$ = fraction of top 1% share going to top 0.1%
- $\frac{S(1)}{S(10)}$ = analogous, find top inequality $\eta = 1/\zeta$ from

$$\frac{S(p/10)}{S(p)} = 10^{\eta-1} \quad \Rightarrow \quad \eta = 1 + \log_{10} \frac{S(p/10)}{S(p)}$$

Wealth Inequality in U.S.

A first thing to note

- Data for wealth considerably murkier than for income
- Particularly true for top wealth inequality
  - excellent summary by Kopczuk (2015), “What Do We Know About Evolution of Top Wealth Shares in the United States?”
- Main thing that’s clear: wealth more unequally distributed than income
- Pen’s parade for wealth: `https://www.youtube.com/watch?v=QPKKQnijnsM`

Households Hold Many Different Assets and Liabilities

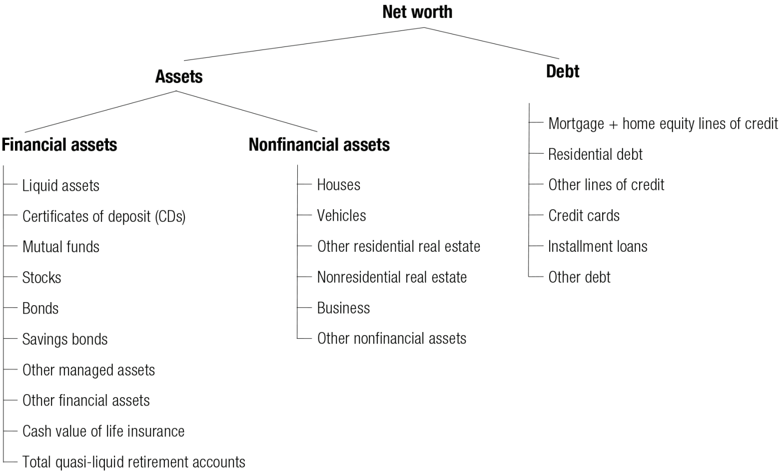
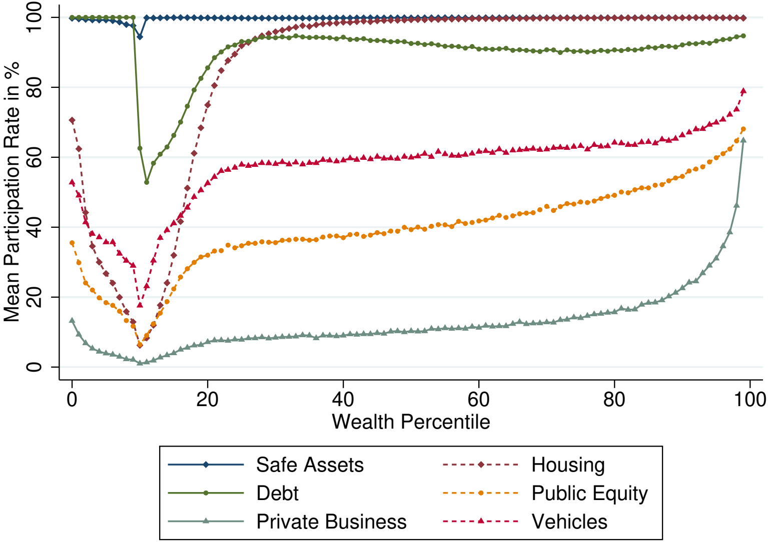


Figure 7. **SCF household portfolio.**

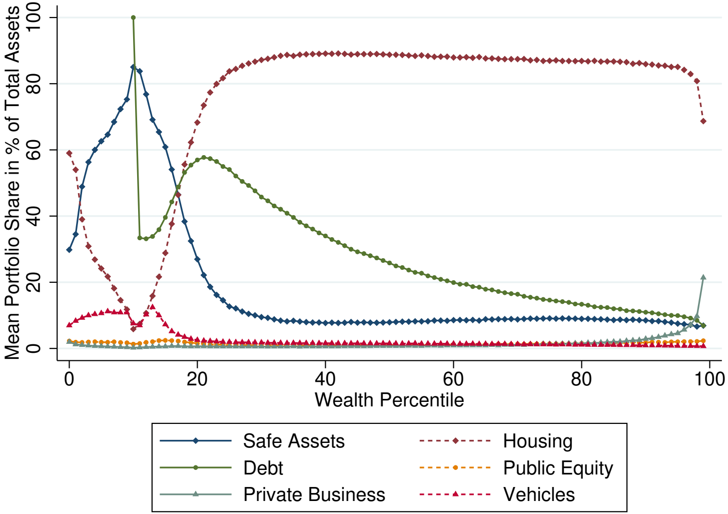
Source: Kuhn and Rios-Rull (2016)

Norway: Participation Rates by Asset Class



Note: safe assets = deposits + bonds + informal loans

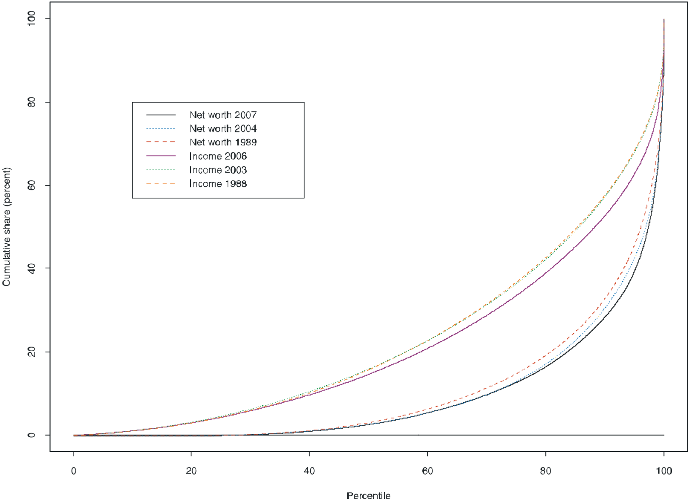
Norway: Portfolio Shares by Asset Class



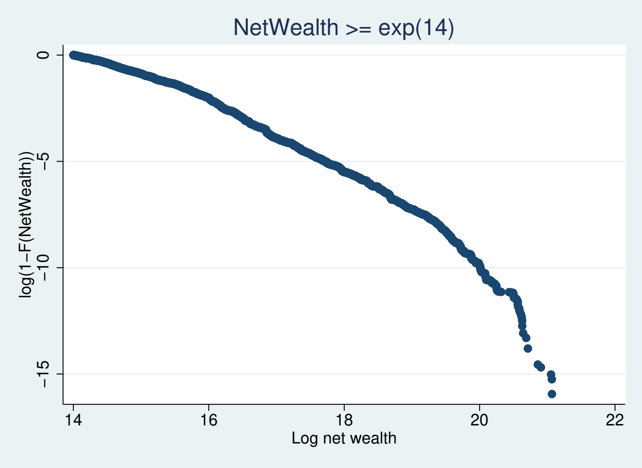
Notes: 15th pctile = 0 net worth. Safe assets = deposits + bonds + informal loans. Wealthy often hold stocks through private holding companies.

Wealth Lorenz Curve (Kennickell, 2009)

Figure A1: Lorenz curves for 1988, 2003 and 2006 total family income and 1989, 2004 and 2007 net worth.



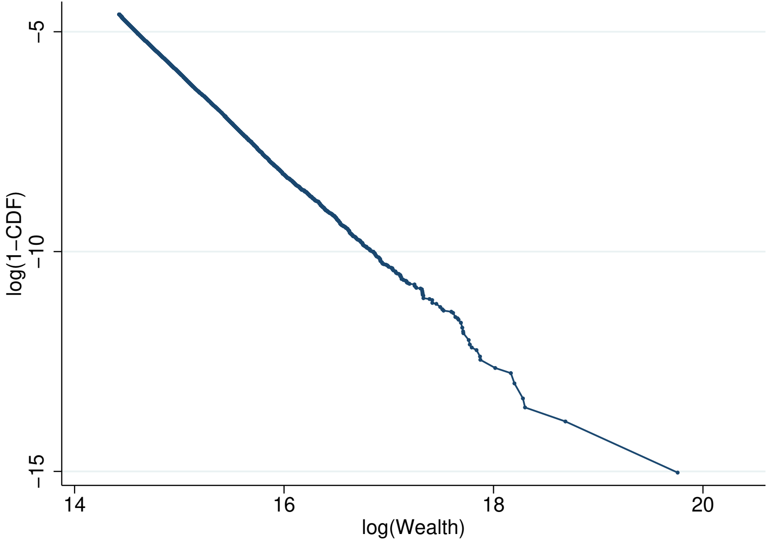
Pareto Tail of Wealth Distribution in SCF



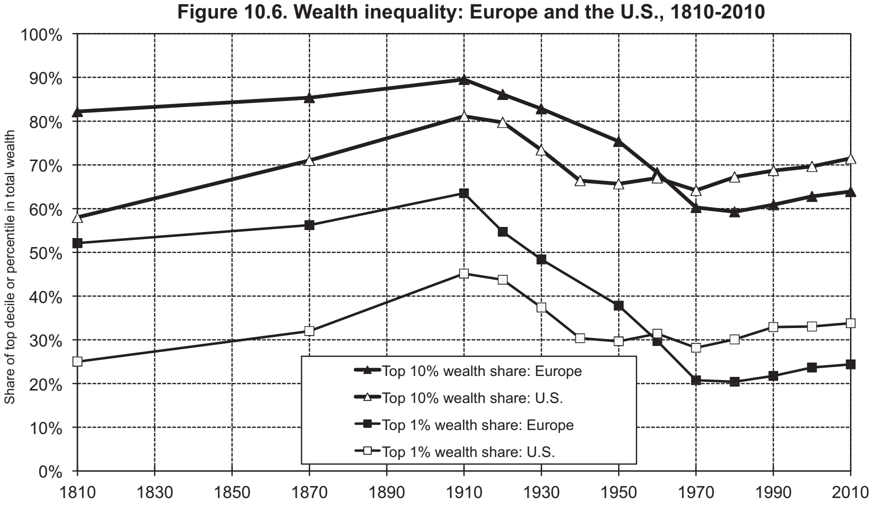
- Source: own calculations using SCF
- Nice article on power laws and random growth (Lectures 5 and 6)

  `http://nautil.us/issue/44/luck/investing-is-more-luck-than-talent`

Pareto Tail of Wealth Distribution in Norwegian Data

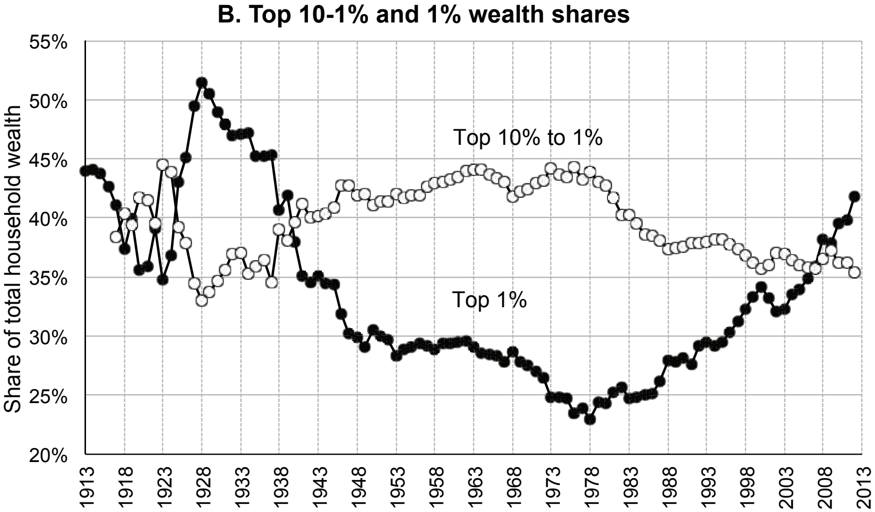


Piketty’s most interesting figure



Until the mid 20th century, wealth inequality was higher in Europe than in the United States.
Sources and series: see piketty.pse.ens.fr/capital21c.

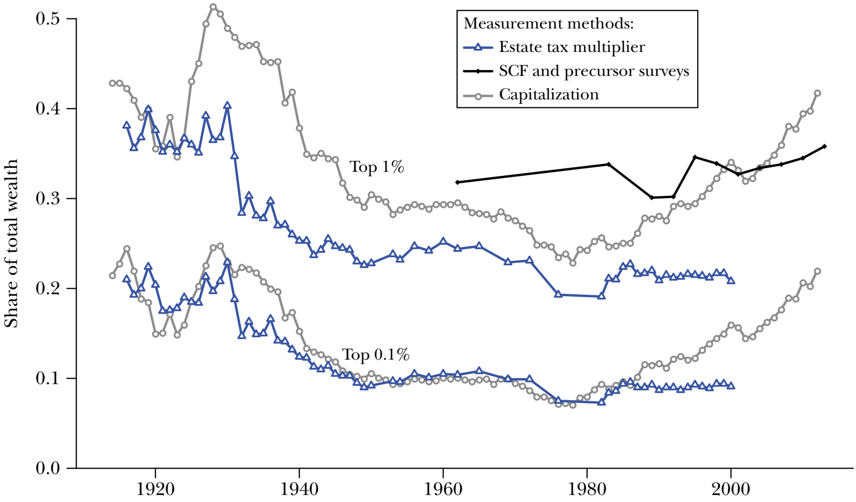
Saez-Zucman: it's even more extreme



Kopczuk: it's not so clear

Figure 1

Top 0.1% and Top 1% Wealth Shares



Capitalization Method

- First use: Robert Giffen (1913), next Charles Stewart (1939)
  - `http://www.nber.org/chapters/c9522.pdf`
  - interesting discussion by Milton Friedman
- Used by Saez and Zucman (2016)
- Idea of capitalization method
  - observe $y_{it}^k = r_{it} a_{it}$
  - estimate $\hat{a}_{it} = y_{it}^k / \bar{r}_t = a_{it} \times r_{it} / \bar{r}_t$
- Potential problem: $r_{it} \neq \bar{r}$, systematically with a_{it}
  - see Fagereng, Guiso, Malacrino and Pistaferri (2016)

Estate Multiplier Method

Due to Mallet (1908) `http://piketty.pse.ens.fr/files/Mallet1908.pdf`

- split population into groups $g = 1, ..., G$
    - e.g. percentiles 1 to 100 of the population
    - N_g = no of people in group g
    - p_g = mortality rate in group g
    - D_g = no of deaths in group g
- This equation holds by definition:

$$D_g = p_g N_g$$

- Similarly, denoting W_g= total wealth in group g, E_g = total estates

$$E_g = p_g W_g$$

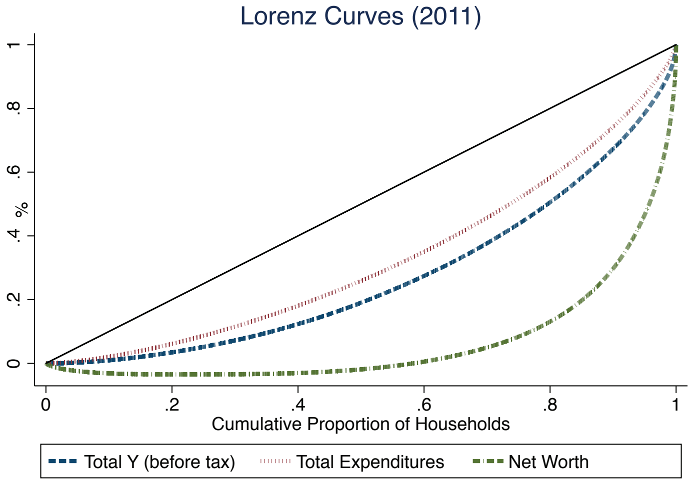
- Therefore, given data on p_g and E_g, can calculate

$$W_g = E_g / p_g$$

or $W_g = m_g E_g$ where $m_g = 1/p_g$ is the "estate multiplier"

“3D Inequality”: Consumption, Income and Wealth

“3D Inequality”: Consumption, Income and Wealth



- Wealth inequality > income inequality > consumption inequality
- Source: own calculations using PSID

"3D Inequality": Consumption, Income and Wealth

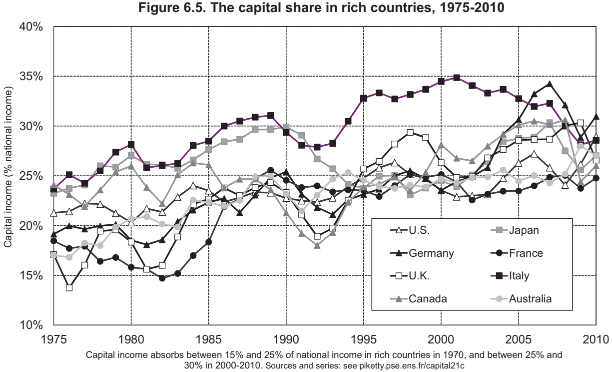
Table 2: PSID Households across the net worth distribution: 2006

	% Share of:			% Expend. Rate		Head's	
NW Q	Earn.	Disp Y	Expend.	Earn.	Disp Y	Age	Edu (yrs)
*Q*1	9.8	8.7	11.3	95.1	90.0	39.2	12
*Q*2	12.9	11.2	12.4	79.3	76.4	40.3	12
*Q*3	18.0	16.7	16.8	77.5	69.8	42.3	12.4
*Q*4	22.3	22.1	22.4	82.3	69.6	46.2	12.7
*Q*5	37.0	41.2	37.2	83.0	62.5	48.8	13.9
	Correlation with net worth						
	0.26	0.42	0.20				

Source: Krueger, Mitman and Perri (2016)

Personal Income Distribution vs Factor Income Distribution

Factor Shares and Inequality



- Developed countries: sizeable increase in capital share

  (Elsby-Hobijn-Sahin, Karabarbounis-Neiman, Piketty-Zucman, Rognlie)

- Usual argument: "capital is back" ⇒ income inequality will increase/already has

- Logic: capital income more concentrated than labor income

Factor Shares and Inequality

- Nicest discussion I've seen: James Meade (1964) "Efficiency, Equality and the Ownership of Property", Section II

  `http://www.princeton.edu/~moll/meade.pdf`

- Succinct summary in 2006 Economic Report of President: "Wealth is much more unequally distributed than labor income. As a result, the extent to which aggregate income is divided between returns to labor and returns to wealth (capital income) matters for aggregate inequality. When the labor share of income falls, the offsetting increase in capital income (returns to wealth) is distributed especially unequally, increasing overall inequality."

Factor Shares and Inequality

- David Ricardo (1821): "The produce of the earth – all that is derived from its surface by the united application of labour, machinery, and capital, is divided among three classes of the community; namely, the proprietor of the land, the owner of the stock or capital necessary for its cultivation, and the labourers by whose industry it is cultivated. [...] To determine the laws which regulate this distribution, is the principal problem in Political Economy"
- What is the relationship between capital (or other factor) share and inequality?
- Use our organizing framework to think about this

Relationship between capital share and inequality?

- Consider following question: when does an increase in capital share coincide with increase in income inequality?
- Use extension of Meade's analysis (1964, Section II). Also see Atkinson & Bourguignon (2000, Section 1).
- Recall total income $y_i = y_i^k + y_i^\ell$.
- Assume continuum of households $i \in [0, 1]$ and order households such that y_i is increasing with i
- Define aggregates

$$Y := \int_0^1 y_i di, \quad Y^\ell := \int_0^1 y_i^\ell di, \quad Y^k := \int_0^1 y_i^k di$$

- Capital share is

$$\alpha := Y^k / Y$$

Relationship between capital share and inequality?

- As measure of inequality take share of income held by top p% (equiv Lorenz curve)

$$S(p) = \frac{1}{Y}\int_{i(p)}^{1} y_i di, \quad i(p) := \text{p'th percentile household}$$

- Question: when α increases, what happens to $S(p)$?
- Easy to see that $\frac{y_i}{Y} = \alpha \frac{y_i^k}{Y^k} + (1-\alpha)\frac{y_i^\ell}{Y^\ell}$. Hence

$$S(p) = \alpha \hat{S}^k(p) + (1-\alpha)\hat{S}^\ell(p)$$

$$\hat{S}^k(p) := \frac{1}{Y^k}\int_{i(p)}^{1} y_i^k di$$

  i.e. share of capital income going to top p percent of total income, and similarly for $\hat{S}^\ell(p)$

- Same formula as Meade's: $i_1 = p_1(1-q) + \ell_1 q$ (see his Section II)

Meade's 1964 Analysis

- Recall formula for top p% income share:
$$S(p) = \alpha \hat{S}^k(p) + (1-\alpha)\hat{S}^\ell(p)$$
- When α increases, does $S(p)$ increase for all p?
- Meade: in data $\hat{S}^k(p) > \hat{S}^\ell(p)$, hence $\alpha \uparrow \Rightarrow S(p) \uparrow$ for all p
- But note implicit assumption: $\hat{S}^k(p)$ and $\hat{S}^\ell(p)$ are constant for all p when $\alpha \uparrow$. How likely is this?
- Would happen only if y_i^k/Y^k and y_i^ℓ/Y^ℓ constant for all i
    - everyone's y_i^k scales up exactly proportionately with Y^k
    - everyone's y_i^ℓ scales down exactly proportionately with Y^ℓ
- Example: "capitalist-worker economy" in which bottom of distribution has only labor income, top has only capital income
$$y_i^k = 0, \; y_i^\ell = Y^\ell/\theta \quad \text{for } i \leq \theta, \qquad y_i^k = Y^k/(1-\theta), \; y_i^\ell = 0 \quad \text{for } i > \theta$$
- If only interested in (say) top 10% share: slightly weaker conditions

More Sophisticated Analysis

- More likely that whatever factor causes $Y^k \uparrow$ affects some individuals' y_i^k proportionately more than others. Then

$$\frac{\partial S(p)}{\partial \alpha} = \underbrace{\hat{S}^k(p) - \hat{S}^\ell(p)}_{\text{due to between-factor distribution}} + \underbrace{\alpha \frac{\partial \hat{S}^k(p)}{\partial \alpha} + (1-\alpha)\frac{\partial \hat{S}^\ell(p)}{\partial \alpha}}_{\text{due to changes in within-factor distribution}}$$

- Crucial question: sign and size of second term?
- In principle, 2nd term can be $+$ or $-$, may outweigh 1st term $(+)$ in which case Meade's analysis is misleading
- Two authors questioning relation between capital share & inequality
  - Blinder (1975): "the division of national income between labor and capital has only a tenuous relation to the size distribution"
  - Krugman (2016) `http://krugman.blogs.nytimes.com/2016/01/08/economists-and-inequality/`