

Earnings Inequality Statistics from the LEHD

James R. Spletzer
U.S. Census Bureau

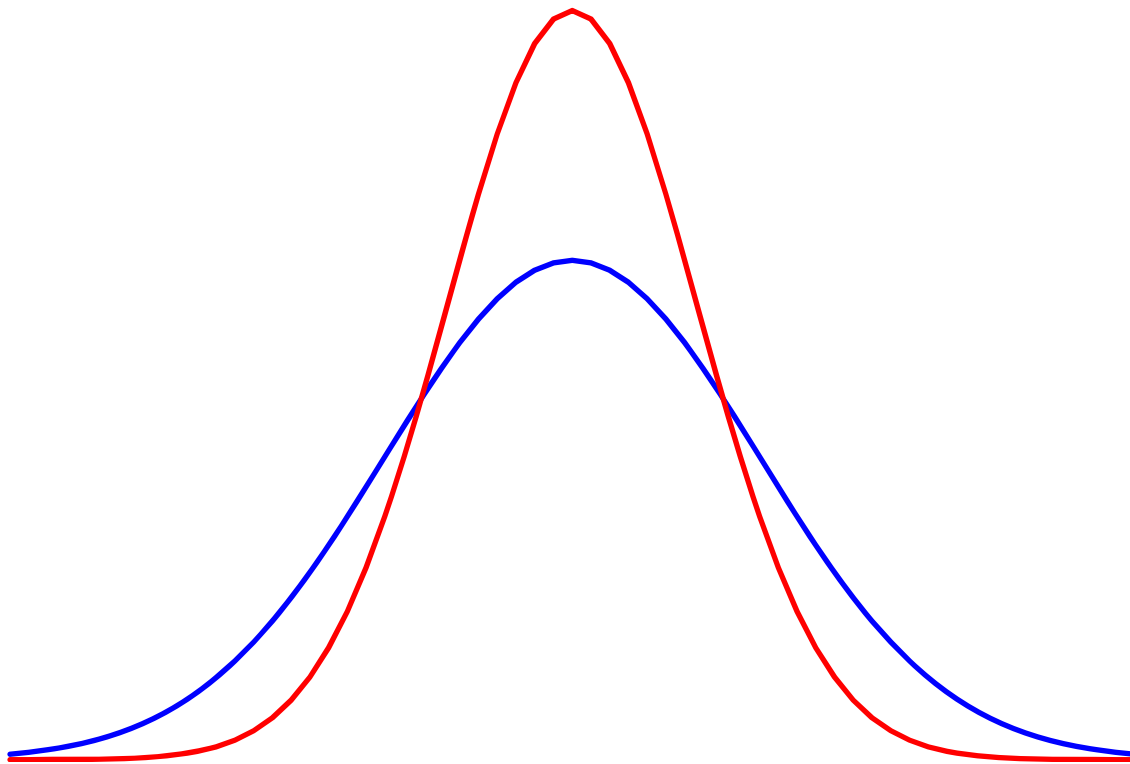
C2ER (Council for Community and Economic Research)
February 15, 2017

Disclaimer

Any opinions and conclusions expressed herein are those of the author and do not necessarily represent the views of the U.S. Census Bureau.

All results have been reviewed to ensure that no confidential information is disclosed.

Increasing Inequality

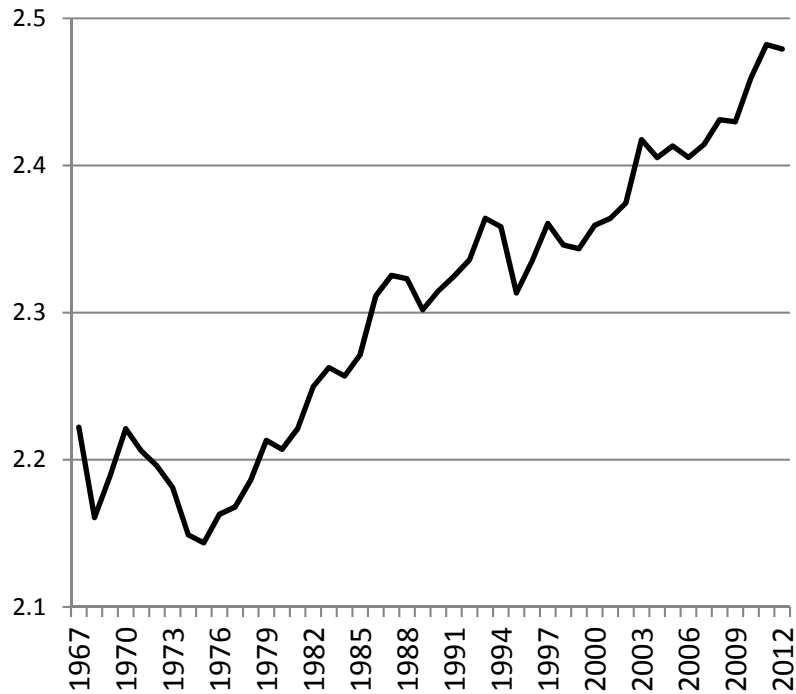


Increasing inequality refers to a widening distribution of earnings (from red to blue)

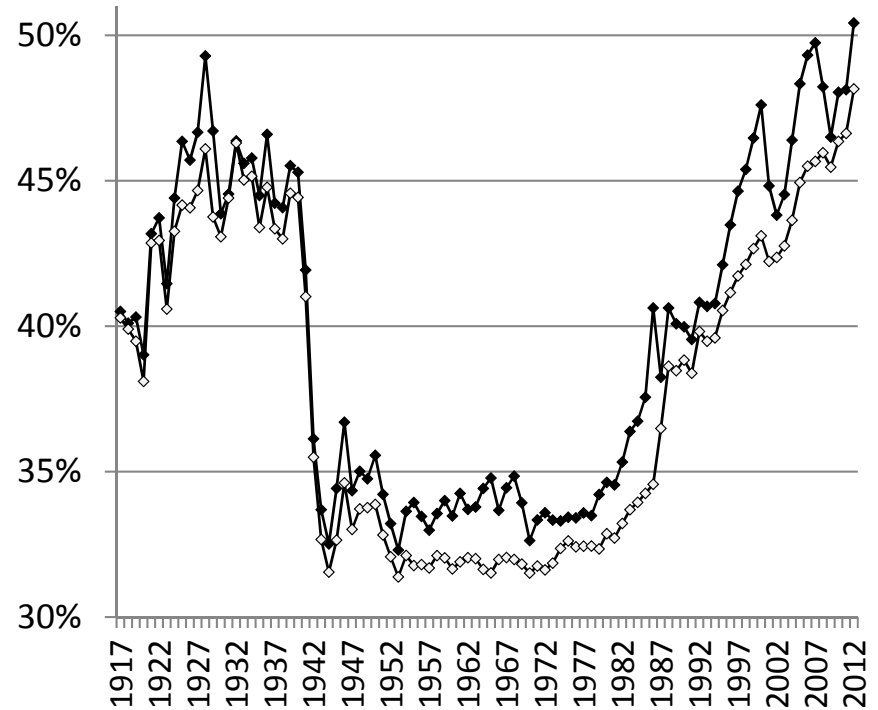
Increasing inequality is often measured using:

- the 90/10 ratio
- the variance
- the amount of mass in the tails

Increasing Inequality



— ln(90)-ln(10) Published CPS ASEC



◆ Top 10% Income Share (including capital gains)
 ◇ Top 10% Income Share (excluding capital gains)

Source: Saez (2013)

Stylized facts about increasing inequality

- 1) 90-10 ratio has been increasing since mid-to-late 1970s
 - Since the mid 1990s, 90-50 has been increasing & 50-10 has been flat
- 2) Top decile income share increasing since late 70s / early 80s
 - Much of this rise is in the top 1%
- 3) Almost everything we know empirically about inequality in the U.S. comes from CPS and IRS data
 - Time series of 90-10, 90-50, and 50-10 from CPS
 - Earnings shares of top 10%, top 5%, and top 1% from IRS

Why is inequality increasing?

- Early literature (1992/1993) focused on skills and institutions
 - Inequality is within skill groups, as measured by education and experience
 - Institutions such as minimum wages and unions affect the lower half of the earnings distribution
- A second wave of literature (mid 2000s) focused on tasks {manual, routine, abstract} and the hollowing out of the earnings distribution to test the hypotheses of technological change and globalization
- The recent literature is focusing attention on the role of the firm and worker sorting across firms

Goals of this Presentation

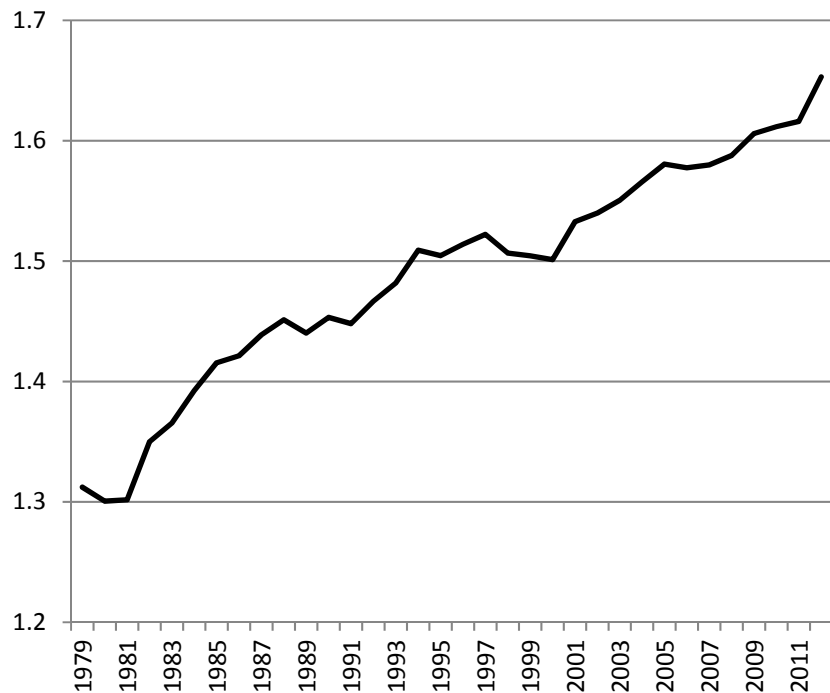
- ❖ Review publicly available earnings distribution statistics that inform us about increasing inequality
- ❖ Introduce another data source (LEHD) with time series information about the earnings distribution
- ❖ Discuss the value added of LEHD statistics
 - Comparison and confirmation
 - Utilize the large sample of the LEHD to provide inequality statistics by detailed demographic and job characteristics
 - Utilize the linked employee-employer aspect of the LEHD to analyze the role of the firm in increasing inequality

CPS {10,50,90} Earnings Percentiles

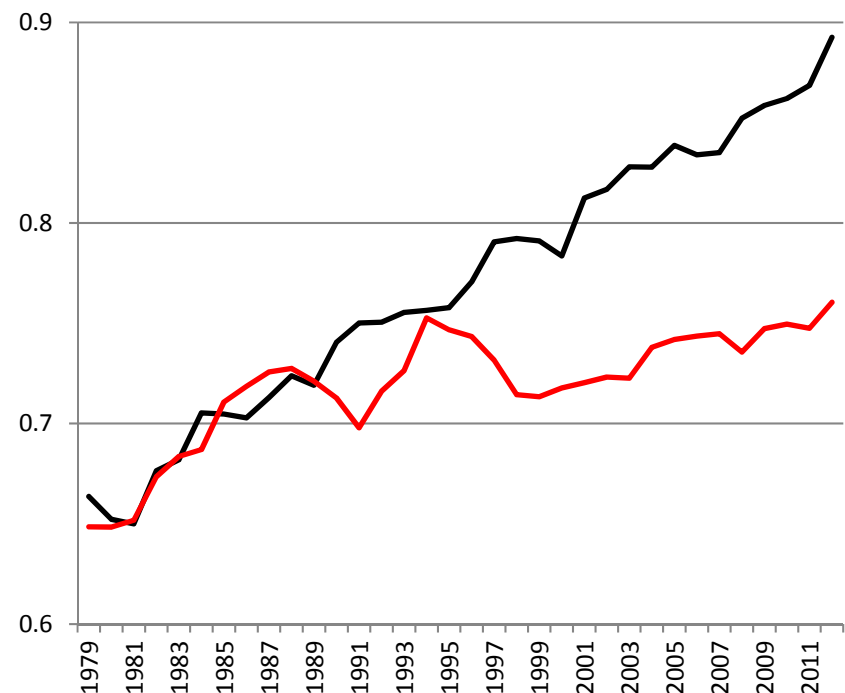
- ❖ Outgoing Rotation Group (ORG) data from the CPS
 - Usual weekly earnings of full time wage & salary workers at their main job
- ❖ 2000 – current tabulations available on BLS website
 - <http://www.bls.gov/webapps/legacy/cpswktab5.htm>
- ❖ Earlier tabulations available by request from BLS staff
 - Annual data beginning in 1979
 - Quarterly data beginning in 1994:Q1
- ❖ Two manipulations to published tabular data
 - Seasonally adjust the quarterly data
 - Convert to real (2012 CPI-U-RS) natural-logarithms

Increasing Inequality, 1979 - 2012

Published CPS-ORG, annual, real 2012 \$



— ln(90)-ln(10) Published CPS ORG



— ln(90)-ln(50) Published CPS ORG

— ln(50)-ln(10) Published CPS ORG

IRS Top Percentile Shares

❖ IRS, Annual since 1917

- Emmanuel Saez' website:

<http://elsa.berkeley.edu/~saez/TabFig2012prel.xls>

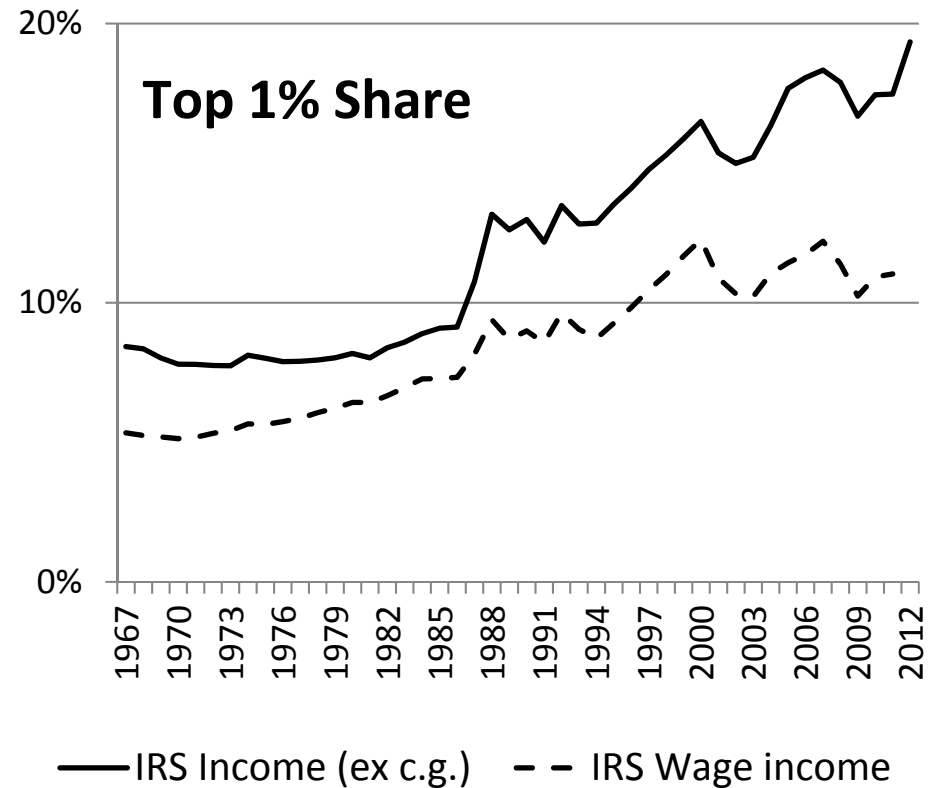
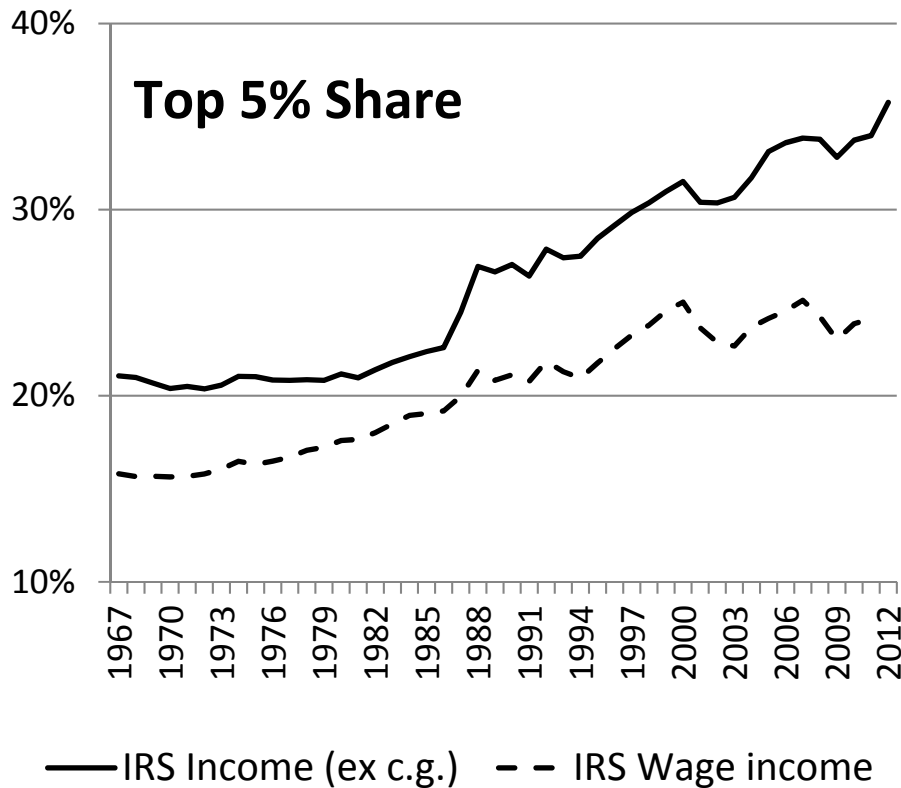
- Two different series:

(1) Annual income of tax units

(2) Annual salaries and wages of tax units

Increasing Inequality, 1967 - 2012

Published IRS tabulations

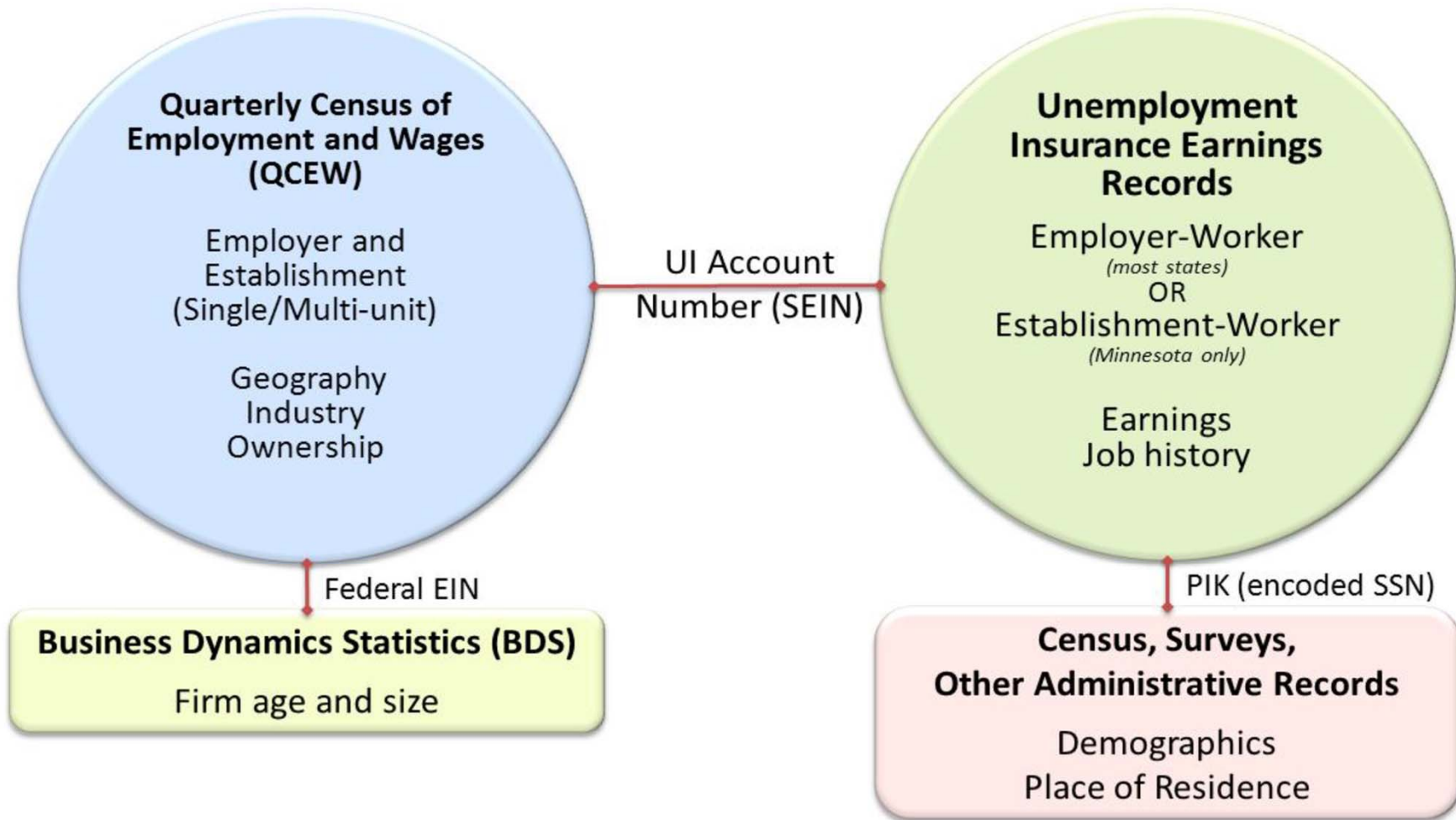


LEHD Data

- ❖ Longitudinal Employer-Household Dynamics
 - Longitudinally linked employer-employee microdata
 - Created at the U.S. Census Bureau
 - Microdata from the State UI administrative systems wage records and QCEW establishment data
 - Enhanced with demographics (age, gender, ...)
 - Enhanced with firm information (age, size)

- ❖ Different states have joined the LEHD at different times, and have provided different amounts of historical data
 - This presentation: 20 states with data from 1996:Q2 to 2012:Q2
 - These 20 states account for 48% of national employment

LEHD Data



LEHD Earnings (I)

- ❖ Earnings: all jobs or “full quarter” jobs?
 - Full quarter job is defined as the middle quarter of 3 consecutive quarters at the same employer, which allows us to assume that the person is working at the employer for the full quarter
 - Using all jobs results in a declining (not increasing) inequality
 - Why? A composition effect due to a declining number of short duration low paid jobs [see Hyatt & Spletzer, 2016]

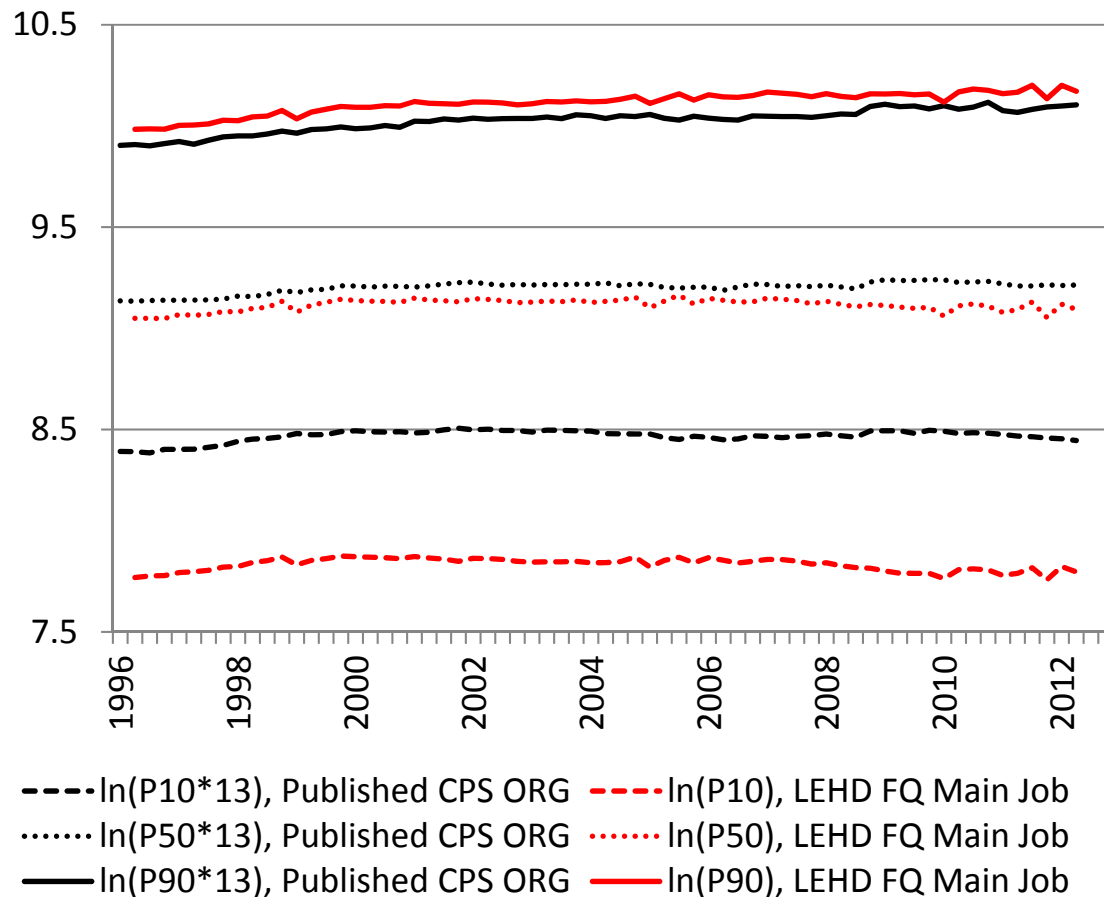
- ❖ Attempt to mimic the CPS-ORG earnings concept
 - Use quarterly earnings of individuals in their main full quarter job
 - Consistent with most of the inequality literature, which imposes a labor supply restriction (such as full time workers in the CPS)

LEHD Earnings (II)

- ❖ Our 20-state LEHD data
 - 4.0 billion jobs, 1996:Q2 – 2012:Q2
 - 2.6 billion full-quarter jobs 1996:Q2 – 2012:Q2
 - approximately 39½ million FQ jobs each quarter (65 quarters)
- ❖ Only 2 manipulations to LEHD full-quarter earnings data
 - Winsorize earnings at the 99.5 % of state-year-quarter distribution
 - Convert to real (2012 CPI-U-RS) natural-log earnings
- ❖ All quarterly time series are seasonally adjusted

CPS-ORG (*13) and LEHD

{10,50,90} earnings percentiles, quarterly SA, real 2012 \$



Very similar 50th & 90th percentiles

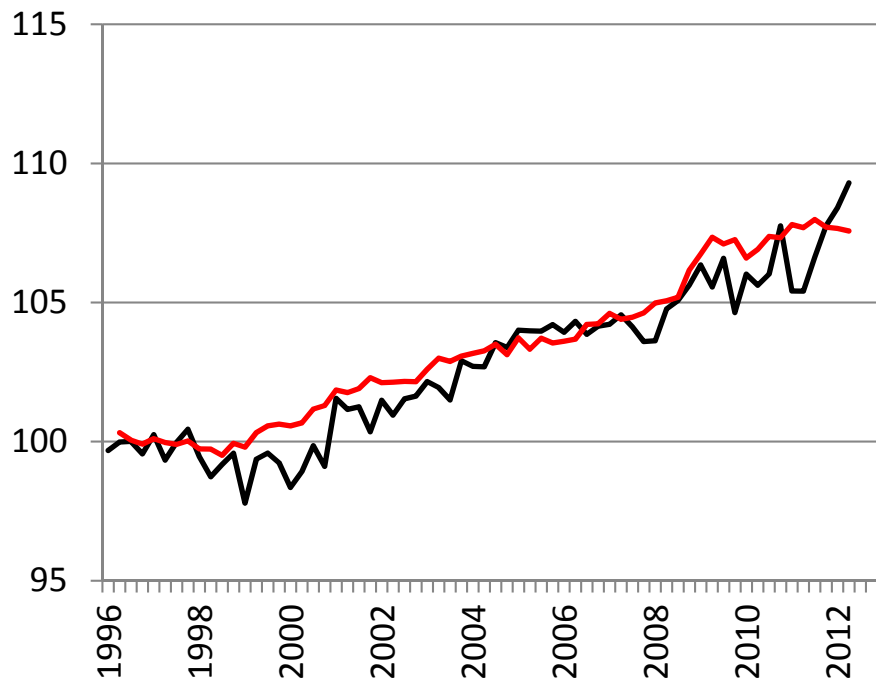
But different 10th percentiles:

- perhaps “*13” is a poor method of transforming CPS weekly earnings to quarterly earnings

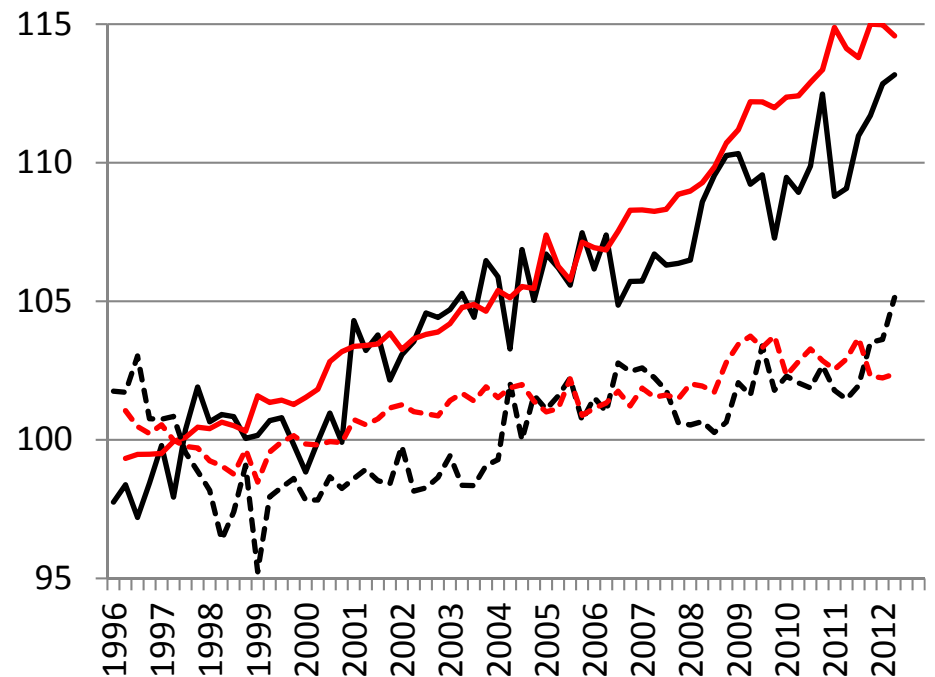
- perhaps there are part-time (<35 hours) workers in the LEHD

Increasing Inequality, 1996 - 2012

CPS-ORG (*13) & LEHD, quarterly SA, real 2012 \$



— $\ln(P90^*13) - \ln(P10^*13)$, Published CPS ORG, 1997=100
 — $\ln(P90) - \ln(P10)$, LEHD FQ Main Job, 1997=100

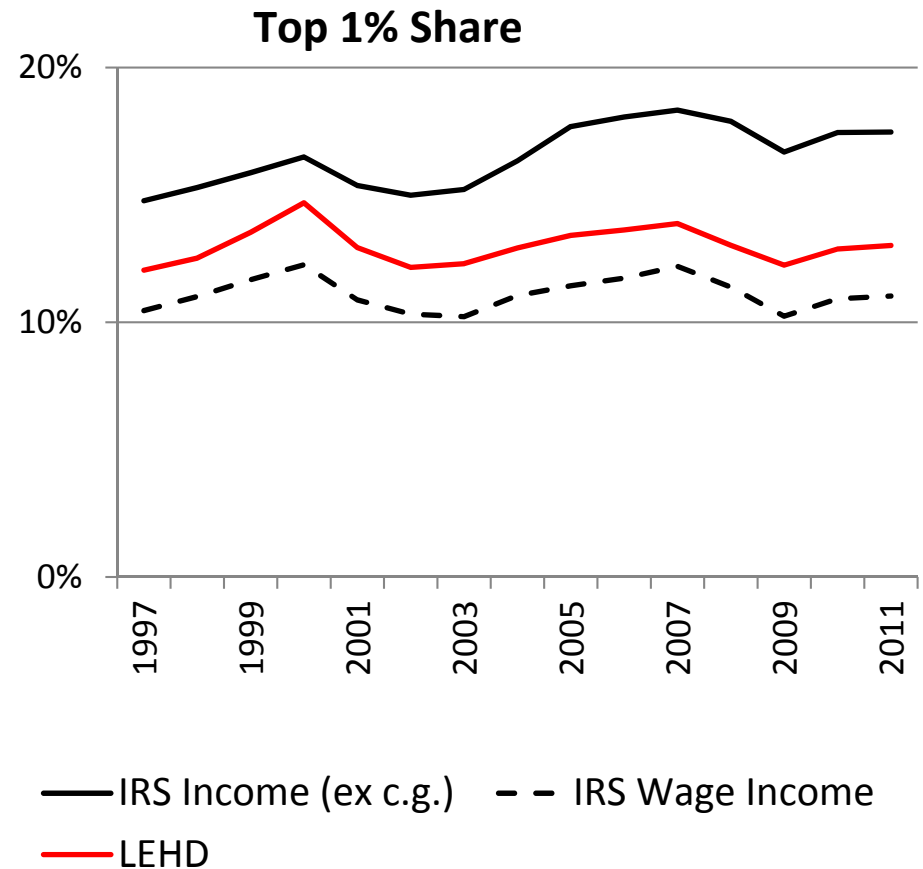
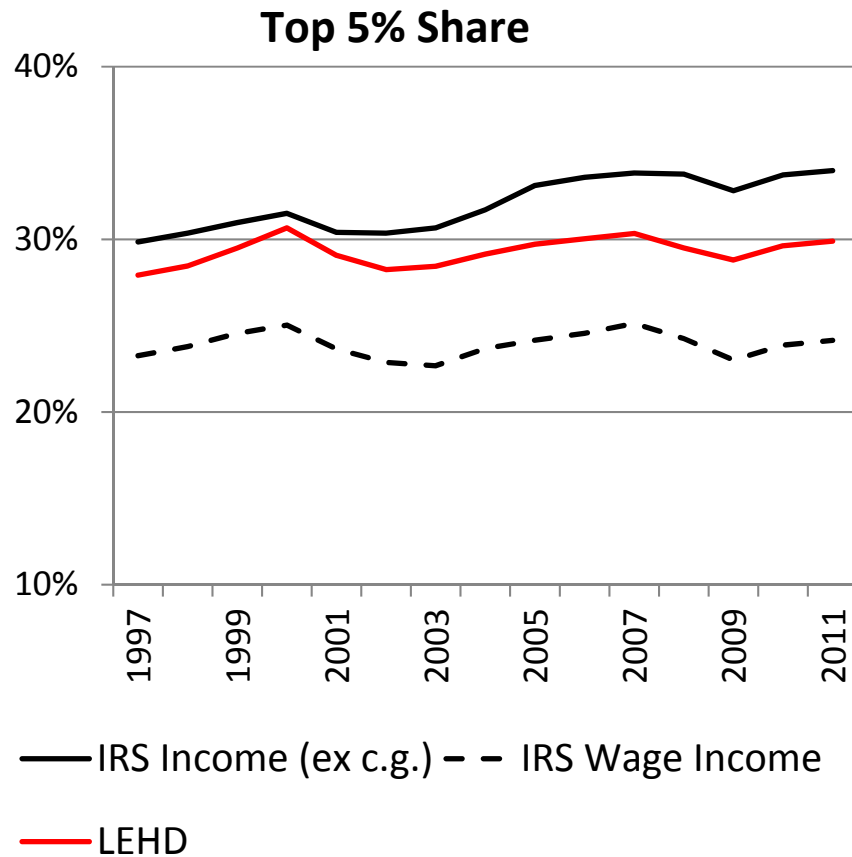


— $\ln(P90^*13) - \ln(P50^*13)$, Published CPS ORG, 1997=100
 — $\ln(P90) - \ln(P50)$, LEHD FQ Main Job, 1997=100
 - - - $\ln(P50^*13) - \ln(P10^*13)$, Published CPS ORG, 1997=100
 - - - $\ln(P50) - \ln(P10)$, LEHD FQ Main Job, 1997=100

LEHD Top Percentile Shares

- ❖ Attempt to mimic the IRS earnings concept
 - Annual earnings of individuals from all jobs during the year
 - 943 million individual-year observations, 1997 – 2011
 - approximately 63 million persons each year (15 years)

IRS, SSA, CPS-ASEC, and LEHD



Summary of Comparison

- ❖ Acknowledge differences in scope and definitions
- ❖ Comparing LEHD {10, 50, 90} with CPS ORG:
 - 50th & 90th percentiles almost identical, 10th different
 - 90/10, 90/50 and 50/10 trends very similar
- ❖ Comparing LEHD top % shares with IRS income:
 - Levels similar, time series correlations are $>.8$
- ❖ Now turn to the value added of LEHD statistics

LEHD Top Percentile Shares

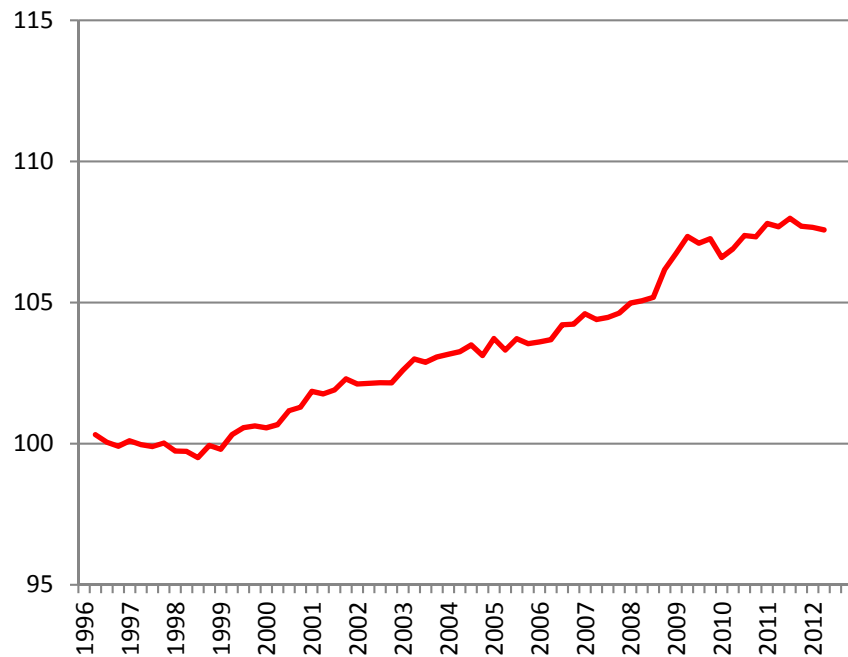
What are the age and gender distributions of workers in the top 5%?

Age	All Workers	Workers in the top 5%
<30	31.8%	2.7%
30-34	11.2%	8.8%
35-39	11.4%	15.1%
40-44	11.4%	18.5%
45-49	10.6%	18.7%
50-54	8.9%	16.2%
55-59	6.6%	11.3%
60-64	4.1%	5.8%
>65	4.1%	2.9%

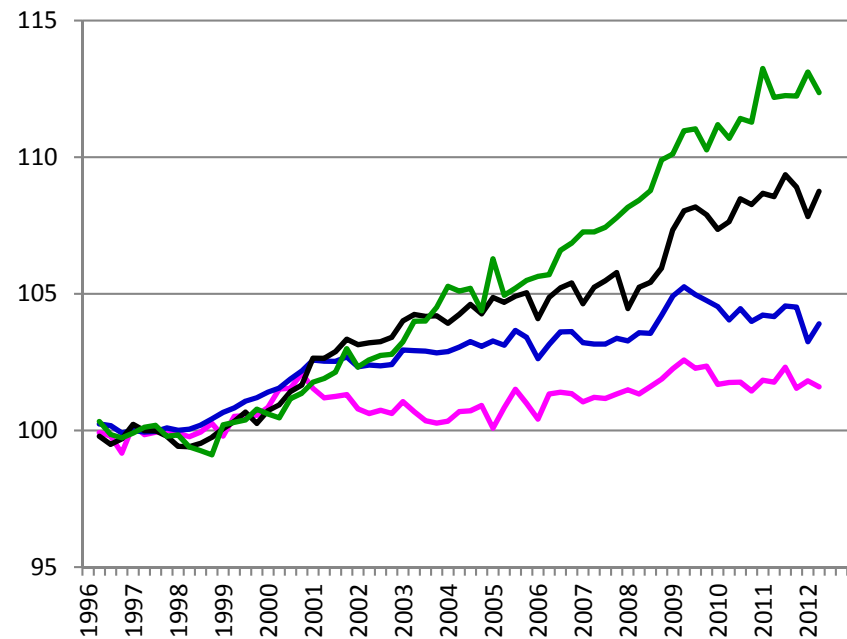
Gender	All Workers	Workers in the top 5%
Male	53.1%	79.0%
Female	46.9%	21.0%

Does the Firm Matter?

LEHD, quarterly SA, real 2012 \$, by firm size



— $\ln(P90) - \ln(P10)$, LEHD FQ Main Job, 1997=100

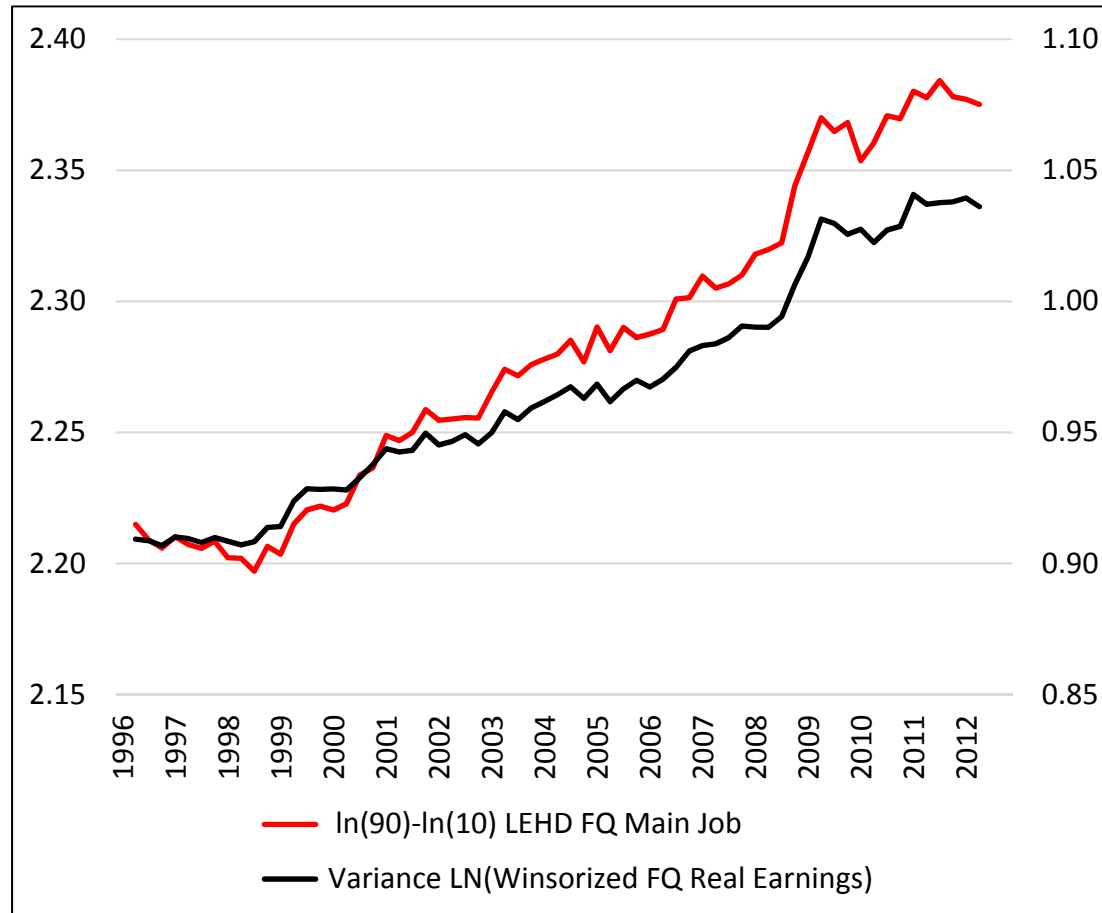


- $\ln(P90) - \ln(P10)$ Firm Size <20, 1997=100
- $\ln(P90) - \ln(P10)$ Firm Size 20-99, 1997=100
- $\ln(P90) - \ln(P10)$ Firm Size 100-999, 1997=100
- $\ln(P90) - \ln(P10)$ Firm Size 1000+, 1997=100

The Role of the Firm

- We know that quite a bit of cross-sectional earnings variance is across establishments (Groshen 1991 QJE, and many more)
- We also know that a sizeable amount of the *growth* in earnings variance is across establishments or firms:
 - Early literature
 - Davis & Haltiwanger (1991 Brookings), LRD [48%]
 - Dunne, Foster, Haltiwanger, & Troske (2004 JoLE), LRD [90%]
 - Recent literature
 - Card, Heining, & Kline (2013)
 - Barth, Bryson, Davis, & Freeman (2016), LEHD [68%]
 - Handwerker & Spletzer (2016), OES [73%]
 - Song, Price, Guvenen, Bloom, & von Wachter (2015), SSA [101%]

Variance of LEHD Earnings



We switch from 90-10 ratios to variances

Why? To use the simple decomposition:

Total Variance =

Variance within firms +

Variance across firms

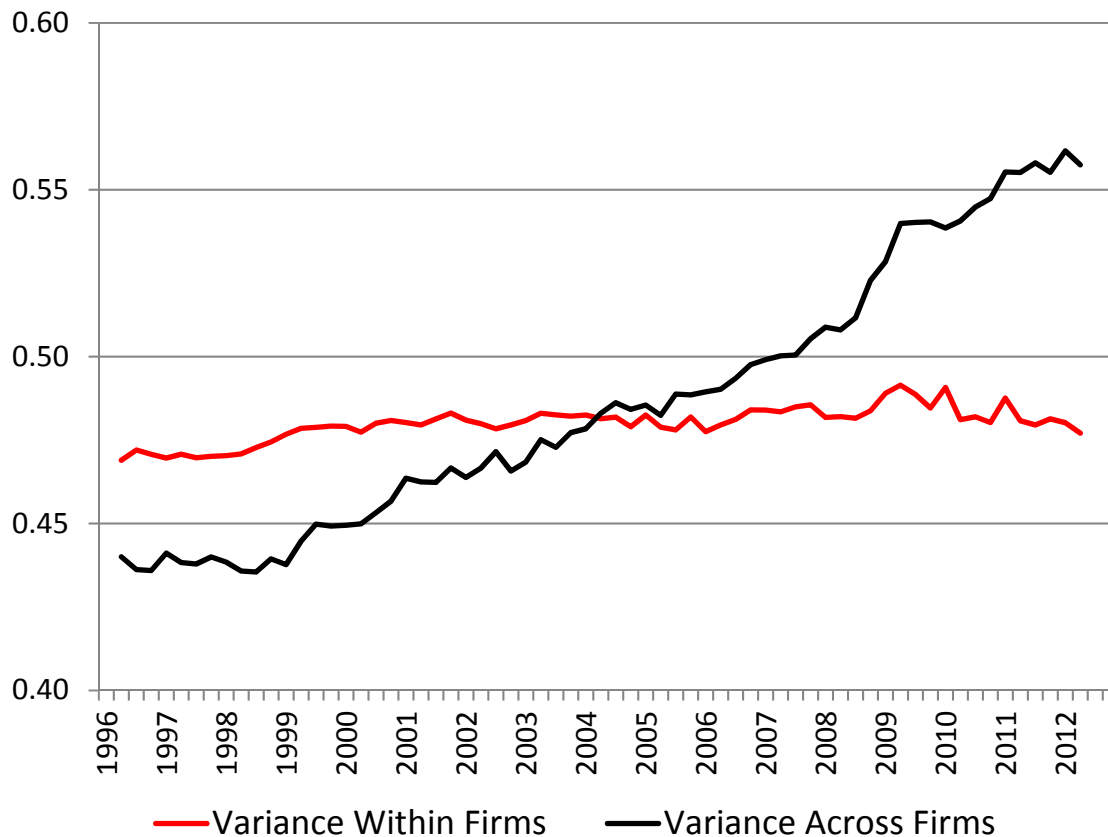
Variance Decomposition

Let “i” index individuals and “f” index firms

The variance of wages is:

$$\begin{aligned} V(w) &= \left(\frac{1}{N}\right) \sum_{i=1}^N (w_i - \bar{w})^2 \\ &= \left(\frac{1}{N}\right) \sum_{i=1}^N (w_{if} - \bar{w})^2 \\ &= \left(\frac{1}{N}\right) \sum_{i=1}^N (w_{if} - \bar{w}_f)^2 + \left(\frac{1}{N}\right) \sum_{i=1}^N (\bar{w}_f - \bar{w})^2 \\ &= \left[\frac{\sum_{f=1}^F N_f V_f(w)}{\sum_{f=1}^F N_f} \right] + \left[\frac{\sum_{f=1}^F N_f (\bar{w}_f - \bar{w})^2}{\sum_{f=1}^F N_f} \right] \\ &= [\text{Variance within firms}] + [\text{Variance across firms}] \end{aligned}$$

Variance Decomposition: LEHD



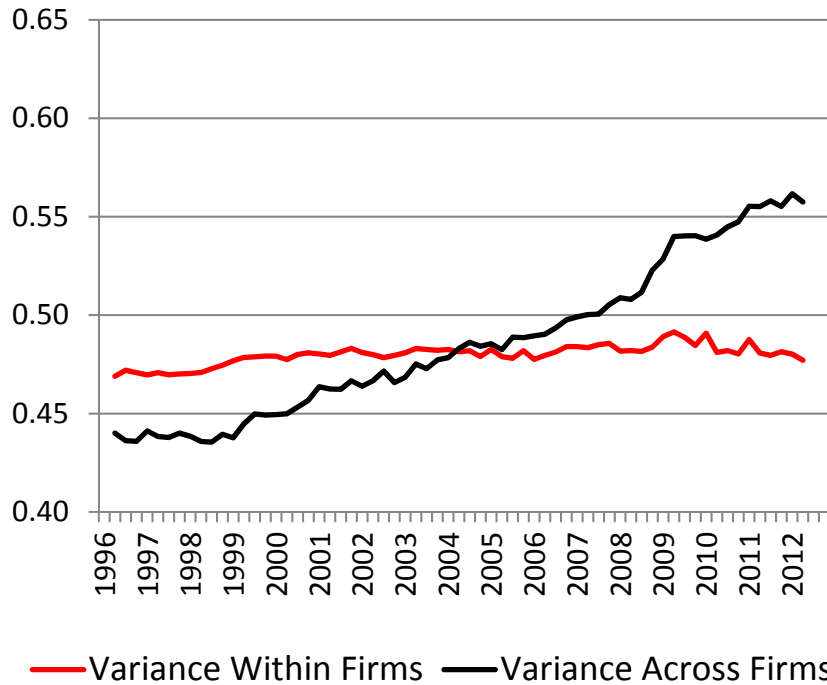
On average, 50.3% of cross-sectional earnings variance is across firms

93.5% of the growth in earnings variance is across firms

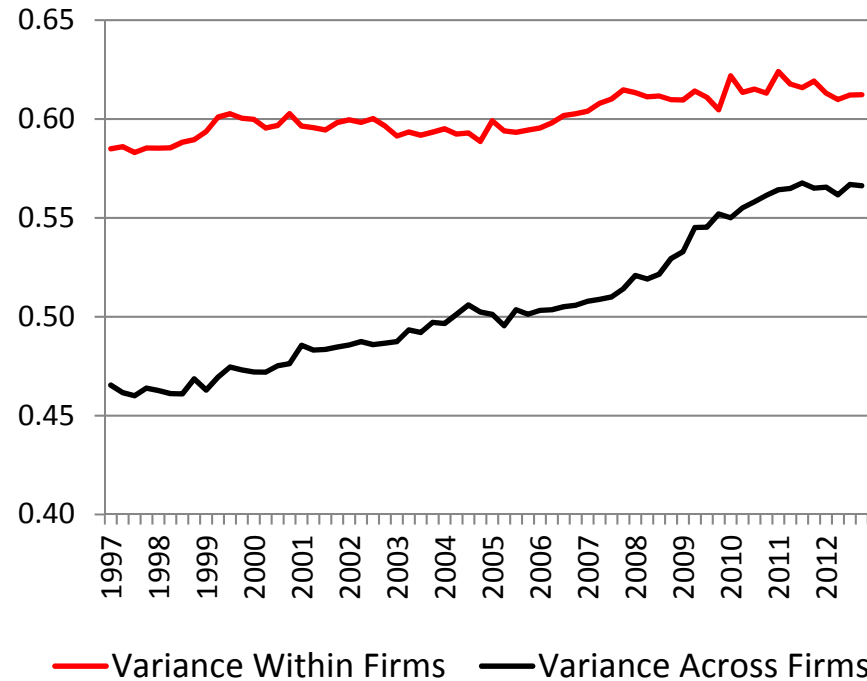
Firm is defined as the State UI account number

Variance Decomposition: LEHD

Two different definitions of firm

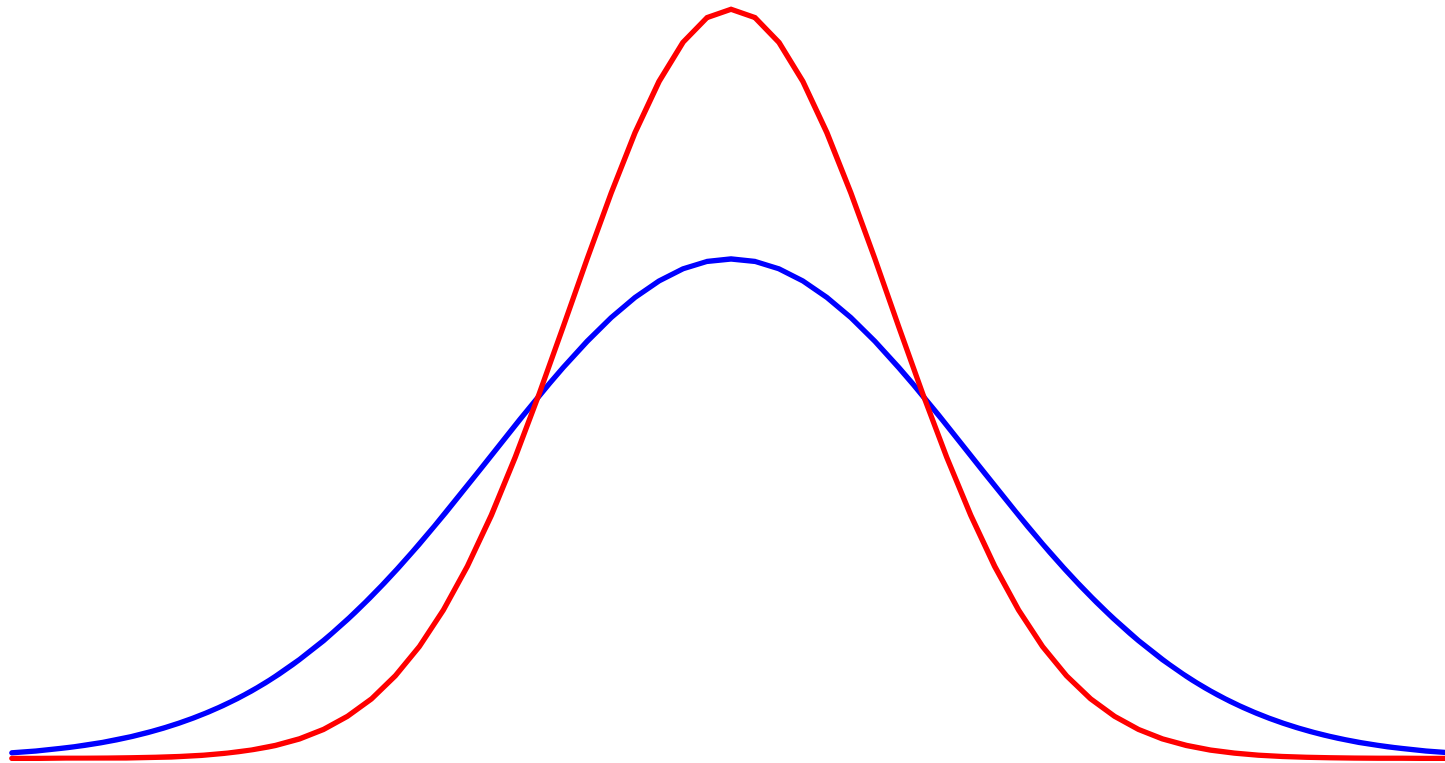


94% of variance growth is across firms defined by State UI Numbers (SEINs)



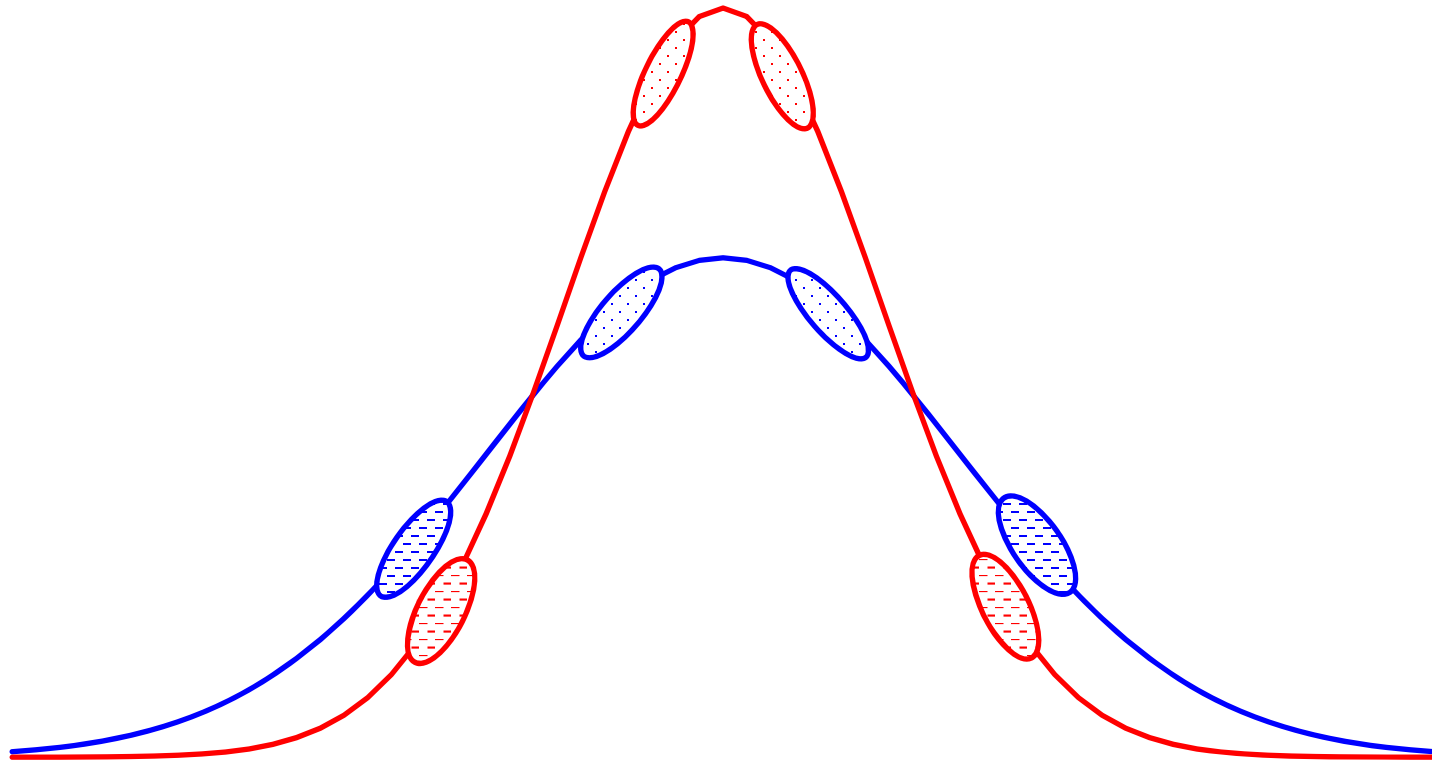
78% of variance growth is across firms defined by National Enterprise

Variance Growth



Variance Growth

Example: all growth is across firms, none within firms



Why is the Firm Important?

Card, Heining, & Kline (2013 QJE):

- Increasing inequality in Germany at the low end of the distribution
 - sharp decline in % employment covered by collective agreements
 - plants, particularly births, opting out of traditional collective bargaining system and paying low wages

U.S. is different: increasing inequality is at the high end

- Three leading proposed mechanisms
 - worker-firm sorting across firms
 - rent sharing
 - firm earnings differentials might be industry differentials?

Summary

- 1) LEHD is an under-utilized data source that has information on the earnings distribution
 - Research tabulations from the LEHD closely mimic inequality tabulations from the CPS-ORG and the IRS
 - The LEHD's large sample size allows for publication by detailed demographic and job characteristics
 - The LEHD's employer-employee links allow for research into how the firm influences increasing inequality
- 2) Next steps are trying to understand why inequality has been rising during the past several decades

Comments appreciated

Jim Spletzer

James.R.Spletzer@census.gov

301-763-4069