Inequality Statistics from the LEHD

James R. Spletzer
Center for Economic Studies
US Census Bureau
James.R.Spletzer@census.gov

June 5, 2014

(The figures in this paper require a color printer)

Abstract

The growth of earnings inequality is at the forefront of current policy discussions, and has been referred to as "the defining challenge of our time." Most of what we know about increasing inequality comes from CPS and IRS statistics. This paper reviews the published CPS and IRS statistics, and compares them to similar statistics computed from the LEHD. The reason to create a time series of earnings distribution statistics from the LEHD is twofold: to confirm what we know from other datasets, and to use the large size of the LEHD linked employer-employer data to create detailed tabulations that are not possible with other data. Questions for the FESAC members are at the end of this paper.

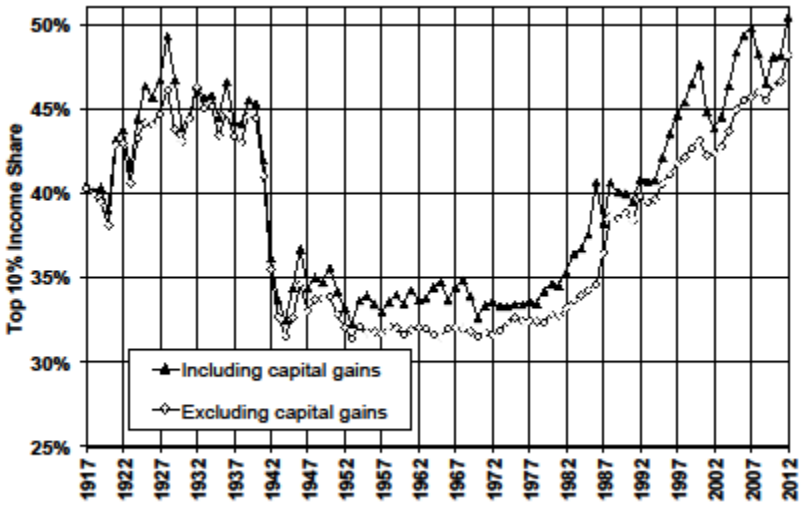
I. Introduction

Increasing earnings inequality refers to a widening of the earnings distribution over time. Documentation of increasing inequality can be traced back to several seminal articles in the early 1990s: Bound and Johnson (1992), Katz and Murphy (1992), Levy and Murnane (1992), and Juhn, Murphy, and Pierce (1993). Each of these articles analyzed microdata from the Current Population Survey (CPS).

There are several ways to measure the widening of the earnings distribution or the widening of the income distribution. Both the Bureau of Labor Statistics (BLS) and the Census Bureau publish the 90^{th} and the 10^{th} percentiles of the distribution, and the ratio of these two – the 90/10 ratio – is a commonly used measure of inequality. The BLS publishes the 90^{th} and the 10^{th} percentiles from the usual weekly earnings questions in the outgoing rotation groups (ORGs) of the CPS, and the Census Bureau publishes the 90^{th} and the 10^{th} percentiles from the income questions in the Annual Social and Economic Supplement (ASEC) of the CPS. These data sources are described in the next section of this paper. Published statistics from both the ORGs and the ASEC show an increase in the 90/10 ratio, starting in the mid-to-late 1970s or early 1980s, and continuing through the most recently published data.

Many researchers have used the publicly available CPS microdata to try to understand the sources of the increasing 90/10 ratio. Two explanations were given the most attention in the early literature. First, increasing inequality during the 1980s could be explained with a straightforward supply and demand explanation, where the demand for skilled workers (as measured by education or experience) rose faster than the supply of skilled workers. Second, and not mutually exclusive from the supply and demand explanation, is that institutions also appeared to play a role. The real value of the minimum wage fell during the 1980s, and the unionization rate also fell during this time period. Recent analysis of the sources of increasing inequality has focused on skill biased technical change (for example, Acemoglu and Autor, 2011), establishment-specific wage premiums (Card, Heining, and Kline, 2013), and the dramatic rise in the return to education (Autor, 2014).

During the past decade, motivated by the work of Piketty and Saez (2003), the academic and policy communities have focused on the increasing share of income accounted for by the top percentiles of the income distribution. The figure following this paragraph is from Saez (2013), who compiles publicly available data from the IRS Statistics of Income publications. This figure shows the pre-tax income share of the top decile, which in 2012 includes all families (tax paying units) with income exceeding $114,000. After remaining relatively constant at roughly 33 percent during the 1950s, 1960s, and 1970s, the top decile share began to increase in the late 1970s or early 1980s. In 2012, the top 10 percent of families account for 50 percent of all income. Saez goes on to show that almost all of this increase occurs within the top 1 percent of families (incomes exceeding $394,000 in 2012). Saez concludes that "those at the very top of the income distribution therefore play a central role in the evolution of U.S. inequality over the course of the twentieth century."



The Top Decile Income Share, 1917-2012, from Saez (2013)

Both measures of publicly available inequality statistics (the 90/10 ratio and the share of income going to the top percentiles) show that inequality has been increasing at a steady pace since the late 1970s or early 1980s. The goal of this short descriptive paper is to introduce another data source with time series information on both of these measures of inequality.

The Longitudinal Employer-Household Dynamics (LEHD) data is a longitudinally linked employer-employee dataset created by the U.S. Census Bureau. The source data underlying the LEHD are the universe of Unemployment Insurance wage records submitted by the states. Data from 1996 to the present are available from 20 states. From these data, we are able to create 90/10 ratios and top earnings shares similar to the CPS and the IRS statistics. The reasons to create similar inequality statistics from the LEHD are twofold: to provide another source of data for comparison and confirmation, and to utilize the large sample of the LEHD to provide statistics by detailed job characteristics or detailed demographic characteristics that are not possible with either the CPS or the IRS data.

This paper is structured as follows. Section II looks at what we have learned about the increasing 90/10 ratio from publicly available CPS statistics. Section III looks at what we have learned about increasing top percentile shares from the publicly available IRS, SSA, and CPS statistics. Section IV introduces the LEHD data and creates statistics that mimic the publicly available CPS and IRS statistics. Section V uses the large sample sizes of the LEHD to provide some inequality statistics not possible from the CPS or IRS statistics. Section VI concludes with several questions for the FESAC members.

II. Publicly Available 90/10 Statistics

Both the BLS and the Census Bureau publish the 90th and the 10th percentiles of the earnings or income distribution. In this section, the underlying data sources are described and

the 90/10 ratios based upon statistics downloaded from the statistical agency's websites are plotted. The data descriptions in sections IIA – IIC below are summarized in Table 1.[1]

IIA. Data Description: CPS ORG

Statistics for the 10th, 25th, 50th, 75th, and 90th percentiles of the weekly earnings distribution of full time wage and salary workers are available from the BLS website at http://www.bls.gov/webapps/legacy/cpswktab5.htm. These statistics are for the time period 2000 to the present, and earlier data are available by request from the CPS staff at BLS. These earlier data are 1979-1999 for annual statistics, and 1994-1999 data for quarterly statistics.

The publicly available CPS ORG statistics refer to the weekly earnings of individuals aged 16 and over working full time during the CPS reference week. These statistics refer only to the individual's main job, and thus any earnings from multiple jobs are not in scope. The earnings of self-employed individuals are also not in scope.

IIB. Data Description: CPS ASEC

Statistics for the 10th, 20th, 40th, 50th, 60th, 80th, 90th, and 95th percentiles of the annual household income distribution are available from the "P60 report" on the Census Bureau website (http://www.census.gov/prod/2013pubs/p60-245.pdf). The P60 report presents annual statistics for the time period 1967 to the present.

The publicly available CPS ASEC statistics refer to the income in the previous year for all persons aged 15 and over who currently reside in the household. There are 18 possible sources of income in the CPS ASEC statistics. The first source is earnings from all jobs held during the year, including self-employment earnings. The other 17 include income from sources such as unemployment compensation, social security, public assistance, interest, and dividends (the full list of 17 sources is given in Appendix A of the P60 report).

IIC. Comparing the CPS ORG and the CPS ASEC

As seen in Table 1, there are differences between the CPS ORG and the CPS ASEC measures of the 90th and the 10th percentile statistics. Most prominent are the differences in scope and definition: the CPS ORG measures the weekly earnings of full-time employed workers at their current main job, whereas the CPS ASEC measures the annual income of all household members at all jobs held during the past year. Note the differences in weekly versus annual, earnings versus income, full-time employed versus all persons, and main job versus all jobs. Regarding the earnings versus income difference, the CPS ORG measures earnings of wage and salary workers (which excludes the self-employed), whereas the CPS ASEC definition is more broad, in that it includes income from self-employment and also includes income from sources other than employment.

[1] Saez (2013) publishes the 90th percentile of the annual income of tax units. A modified Table 1 and Figure 1, which includes this IRS data series, are available upon request.

There is no clear prediction regarding how the 90th and the 10th percentiles from the CPS ORG and the CPS ASEC distributions should compare to each other. For example, including all household members and all jobs implies that the earnings in the CPS ASEC should be higher than what any one of these household members would report in the CPS ORG (assuming we could control for the different weekly versus annual frequency underlying the two data sources). On the other hand, including income from unemployment compensation and public assistance will add non-employed individuals to the distribution, and will likely add these individuals to the lower end of the distribution if the added sources of income are lower than average earnings of currently employed persons (assuming we could control for the different weekly versus annual frequency).

These differences in scope and definition will lead to differences in the CPS ORG and CPS ASEC statistics, but it needs to be noted that neither the CPS ORG nor the CPS ASEC statistics are preferred nor "more correct" for an analysis of increasing inequality. Analysts need to be aware that the two are different, and thus any empirical differences in the distribution statistics across the two data sources should not be surprising.

IID. Empirical Analysis of Publicly Available 90/10 Statistics

Figure 1 plots the 10th, 50th, and 90th percentiles of the earnings distribution from the CPS ORG and the household income distribution from the CPS ASEC. Each of the series in this Figure is from data downloaded from the statistical agency's websites (with 1979-1999 CPS ORG data obtained via special request). A few edits were made to the downloaded data. First, the CPS ORG data were transformed from nominal to 2012 real dollars using the CPI-U-RS; the CPS ASEC data are published in real 2012 CPI-U-RS dollars. Second, all series were transformed from levels to natural logs. And third, the natural log of 52 was added to the natural log of the CPS ORG weekly earnings data in order to graph the CPS ORG data on the same vertical scale as the CPS ASEC data.

Figure 1 shows that in any given year, the CPS ASEC household income distribution is wider than the CPS ORG earnings distribution. After adjusting the CPS ORG series by adding ln(52), the 10th percentile of the CPS ASEC series is lower than the 10th percentile of the CPS ORG series, while the 90th percentile of the CPS ASEC series is higher than the 90th percentile of the CPS ORG series. These differences are undoubtedly due to the discussion above – there are likely many non-employed persons in the ASEC with low non earning-related incomes, resulting in a lower level of the 10th percentile relative to the working full time concept of the ORG, and adding income across all jobs and all family members undoubtedly raises the level of the 90th percentile relative to the individual-based earnings data in the ORG.

Of key interest in Figure 1 are the trends in each of the series. The 10th percentile is essentially flat in both series during the 1979-2012 time period, falling by 0.034 log points in the CPS ORG, and rising by 0.011 log points in the CPS ASEC. The median (50th percentile) is modestly rising in both series during the 1979 to 2012 time period, increasing by 0.078 log points in the CPS ORG and by 0.050 log points in the CPS ASEC. Most dramatic is the rise in the 90th percentile in both series; the CPS ORG 90th percentile increases by 0.307 log points between 1979 and 2012, and the CPS ASEC 90th percentile increases by 0.277 log points

between 1979 and 2012.[2] This latter finding implies that the upper end of the earnings and income distributions is rising during the 1979-2012 time period, whereas the lower end of the distributions is relatively stagnant in inflation-adjusted terms.

The top panel of Figure 2 presents the 90/10 ratios computed from the publicly available CPS data. Since the 10th percentile and 90th percentile series are expressed in natural logs, the 90/10 ratio is merely the vertical distance between the 90th and the 10th percentile series in Figure 1. And one final technical note – in order to easily put the CPS ORG and the CPS ASEC 90/10 ratios on the same graph, the 90/10 ratios are presented in the top panel of Figure 2 as an index, with 1997=100.[3]

The data in the top panel of Figure 2 show that the 90/10 ratios in both the CPS ORG and the CPS ASEC have been increasing during the last several decades. The CPS ASEC 90/10 ratio was essentially flat from 1967 to the mid-to-late 1970s, and has been increasing since. The CPS ORG 90/10 series, which starts in 1979, has been increasing since 1981. The increasing 90/10 ratio is similar in the CPS ORG and the CPS ASEC during the latter half of the 1990s and through most of the 2000s. Although the two series diverge in 2012, both the CPS ORG and the CPS ASEC have increased by 5 to 6 percent during the 1997 to 2011 time period.

An often used decomposition of the 90/10 ratio is based on the following identity: ln(90/10) = ln(90/50) + ln(50/10). This decomposition shows how much of the increasing 90/10 ratio is due to growth in the upper part of the distribution (the growth between the 50th and the 90th percentiles) and how much is due to growth in the lower part of the distribution (the growth between the 10th and the 50th percentiles). The 90/50 and the 50/10 ratios are presented in the bottom panel of Figure 2.

The bottom panel of Figure 2 shows that the 90/50 ratio has been increasing steadily through the entire time period covered by each series. From 1979 to 2011, both the CPS ORG and the CPS ASEC 90/50 ratios have increased from roughly 84 to 109 (using the 1997=100 scale). Both data series agree that the 50/10 ratio has been relatively flat since the mid-1980s (acknowledging that the two series differ on whether the 50/10 ratio was flat or rising during the early 1980s).

In summary, the statistics available from the publicly available CPS ORG and CPS ASEC both show that inequality, as measured by the ratio of the 90th percentile to the 10th percentile of the distribution, has been increasing since the late 1970s or early 1980s. Furthermore, looking at the 90/50 and 50/10 ratios, both series agree that the increase in inequality since the mid 1980s has been in the upper half of the distribution, with little if any growth in the lower half of the distribution.

---

[2] Converting these .307 and .277 log point increases into percentage increases in the levels emphasizes how dramatic these growth rates are [the conversion is done via exp(.307)-1 and exp(.277)-1]. The CPS ORG 90th percentile rose 35.9 percent between 1979 and 2012 (from 71,732 to 97,500) and the CPS ASEC 90th percentile rose 31.9 percent between 1979 and 2012 (from 110,648 to 146,000).

[3] 1997 is chosen since this is the start of the LEHD time series used later in this paper.

III. Publicly Available Top Percentile Share Statistics

Analysis of the 90/10 ratio from the two CPS series provides time series evidence on the rise in inequality for much of the distribution. However, analysis of the 90/10 ratio does not provide any evidence of what is happening in the tails of the distribution – specifically, what is happening below the 10^{th} percentile or what is happening above the 90^{th} percentile. The time series evidence from the IRS statistics, compiled and published by Saez (2013), provides evidence on the substantial changes occurring amongst tax units with the top incomes. In this section, three different sources of data with information about the shares of top earners are presented. Information on the data sources, the scope, and the definition of earnings or income is given in Table 2.

IIIA. Data Description: CPS ASEC

In addition to the statistics for the various percentiles of the household income distribution, the Census Bureau's P60 report also contains statistics on the share of household income for the top five percent of the distribution. These statistics are annual, for the time period 1967 to the present. Relevant details of the scope and the definition of household income are discussed in the earlier section IIB.

IIIB. Data Description: IRS (Saez)

As mentioned in the introduction, Saez (2013) compiles publicly available data from the IRS Statistics of Income publications and presents them in easily accessible excel spreadsheets (available at http://elsa.berkeley.edu/~saez/TabFig2012prel.xls). The statistics that have received the most attention are those that document the shares of the top 10 percent, the top 5 percent, and the top 1 percent of the income distribution. These statistics refer to the annual income of tax units, where a tax unit is defined as a married couple living together or a single adult. The definition of income is the IRS concept of gross income, which includes all income items reported on tax returns and before all deductions (for further details, see Piketty and Saez, 2003).

Saez (2013) presents top income shares with and without capital gains. All figures in this paper use the series without capital gains (similar to the CPS ASEC measure of top income shares). Furthermore, Saez presents separate series for gross income and a series based only on the salaries and wages component of income. Salaries and wages is a subset of gross income, excluding small business and farm income, partnership and fiduciary income, dividends, interest, rents, royalties, and other small items reported as other income. The graphs in this paper will use both the comprehensive measure of income (excluding capital gains) and the salaries and wage measure of income (referred to as "wage income" in the graphs).

IIIC. Data Description: SSA (Saez)

Table B5 of the spreadsheet in Saez (2013) presents shares of the top earnings of individuals from the Social Security Administration (SSA) data. Although it is unclear if the earnings concepts between the IRS and the SSA data are similar or different, comparing the IRS

and the SSA statistics should provide evidence on how the top shares differ across tax units in the IRS data and individuals in the SSA data.

IIID. Empirical Analysis of Publicly Available Top Share Statistics

The top panel of Figure 3 presents the top 5% income shares from four data sources, and the bottom panel of Figure 3 presents the top 1% income shares from three data sources (the CPS ASEC does not publish the top 1% share).[4] It is immediately obvious that there are two discrete jumps in these graphs that require explanation. The IRS income series increased from 22.6 percent in 1986 to 27.0 percent in 1988, which Piketty and Saez (2003) attribute to income shifting between the corporate sector and the individual sector following the Tax Reform Act of 1986. The CPS ASEC series increased from 18.6 percent in 1992 to 21.0 percent in 1993; the most likely explanation for this increase is given in footnote 8 of Table A2 of the P60 report, which describes the transition of the ASEC from paper and pencil to computer-assisted interviewing, as well as revised (upward for most) topcodes for various income measures.

The four series in the top panel of Figure 3 differ somewhat on the amount of income (or earnings) accounted for by the top 5 percent. In 2011, the IRS income measure, which excludes capital gains, states that 34 percent of annual income is held by the top 5 percent of tax units. The equivalent statistics are 28.2 percent for the SSA measure of individual earnings, 24.2 percent for the IRS salaries and wage measure, and 22.3 percent for the CPS ASEC household income measure. Although some variation in the measures should not be surprising, given the differences in scope and the definition of income, it is important to note that all four publicly available data sources agree that the top 5 percent account for more than 20 percent of all income or earnings.

The main conclusion evident in the top panel of Figure 3 is that the share of income (or earnings) accounted for by the top 5 percent increased between 1993 and 2011.[5] The IRS income measure, which excludes capital gains, increased by 6.6 percentage points, from 27.4 percent to 34.0 percent. The SSA measure of individual earnings increased by 3.9 percentage points (from 24.3 to 28.2), the IRS wage income measure increased by 2.9 percentage points (from 21.3 to 24.2), and the CPS ASEC measure of household income increased by 1.3 percentage points (from 21.0 to 22.3). While there is some disagreement on the magnitude of the increase, all four series agree that the share of income (or earnings) accounted for by the top 5 percent is increasing during the 1993 to 2011 time period.

The bottom panel of Figure 3 shows the share of income (or earnings) accounted for by the top 1 percent of the distribution. In terms of levels, each series agrees that between 11 and 18 percent of income (or earnings) in 2011 is held by the top 1 percent of the distribution. Furthermore, each series agrees that the growth of this top 1 percent share increased by more than 20 percent between 1990 and 2011. This growth is cyclical, with obvious declines in the 2001 and the 2007-2009 recessions. And finally, the top and bottom panel of Figure 3 can be

[4] A modified Figure 3, which includes a graph for the top 10% income share from the two IRS series and the SSA series, is available upon request.

[5] 1993 is chosen as the starting point of the comparison for two reasons: to include the SSA series which starts in 1990, and to avoid the 1993 discontinuity in the CPS ASEC series.

directly compared, since the top 1 percent is a strict subset of the top 5 percent. In terms of increasing inequality, more than half of the growth of the top 5 percent share between 1993 and 2011 occurs in the top 1 percent share.[6]

In summary, the publicly available statistics from the CPS ASEC, the IRS, and the SSA all show increases in the share of income (or earnings) held by the top 5 percent of the distribution between 1993 and 2011. The IRS and SSA measures show that more than half of this growth is in the top 1 percent of the distribution. This growth in inequality in the top shares is not measured by the 90/10 ratios available from the publicly available CPS, and thus these top share statistics provide further measures of increasing inequality above and beyond the measured growth in inequality shown in the 90/10 ratios.

IV. Inequality Statistics from the LEHD Data

This section describes the LEHD data, describes the computation of 90/10 ratios and top earnings shares that are comparable to the publicly available CPS ORG, the CPS ASEC, and the IRS statistics, and compares the LEHD statistics to the published CPS and IRS statistics.

IVA. Data Description: LEHD

The Longitudinal Employer Household Dynamics (LEHD) is a longitudinally linked employer-employee dataset created by the U.S. Census Bureau as part of the Local Employment Dynamics federal-state partnership. The data are derived from state-submitted Unemployment Insurance (UI) wage records and the Quarterly Census of Employment and Wages (QCEW) data. Every quarter, employers who are subject to state UI laws – approximately 98% of all private sector employers, plus state and local governments – are required to submit to the states information on their workers (the wage records) and their workplaces (the QCEW). The wage records and the QCEW data submitted by the states to the U.S. Census Bureau are enhanced with census and survey microdata in order to incorporate information about worker demographics (age, gender, race and ethnicity, and education) and the firm (firm age and firm size). Abowd et al. (2009) provide a thorough description of the source data and the methodology underlying the LEHD data and one of its main public use data products, the QWI. A job in the LEHD is defined as the presence of an individual-employer match, and earnings is defined as the amount earned from that job during the quarter.

Because states have joined the LEHD program at different times and have provided various amounts of historical data upon joining the LEHD program, the length of the time series of LEHD data varies by state. The data used here are private sector data from the 20 states that have data available from 1996:Q2 through 2012:Q1.[7] These 20 states account for 48 percent of national employment.

---

[6] Specifically, the top 5 percent share of the IRS income measure (excluding capital gains) increased by 6.6 percentage points, the top 1 percent share increased by 4.7 percentage points, and thus 71 percent of the growth in the top 5 percent is due to growth in the top 1 percent. Similar statistics for the IRS wage income measure are 2.9, 2.0, and 69 percent. Similar statistics for the SSA measure are 3.9, 2.2, and 58 percent.

[7] CA, CO, CT, HI, ID, IL, KS, LA, MD, MN, MO, MT, NC, NJ, NM, OR, RI, TX, WA, WI.

IVB. 90/10 Ratios from the LEHD, Quarterly

Table 3 presents details about the earnings distribution statistics created from the LEHD. One key thing to note is that both quarterly and annual 10th, 50th, and 90th percentile statistics are created from the LEHD. The quarterly statistics are for comparison to the published CPS ORG statistics, and the annual statistics are for comparison to the CPS ASEC statistics.[8]

The LEHD quarterly statistics that are compared to the CPS ORG quarterly statistics refer to the earnings of individuals in their main full-quarter job. A full quarter job is defined as the middle quarter of an employer-employee match that lasts for three consecutive quarters. This is done as a best comparison to the CPS ORG statistics, which refer to the earnings of individuals working full time. The 20 state LEHD does not have any measures of labor supply within the quarter (such as weeks worked or hours worked), and as such, restricting to jobs where the individual worked at the same employer in both the previous quarter and the next quarter assumes that the individual is working full time in terms of the weeks dimension (we can not control for the hours dimension of "full time" in the 20 state LEHD data). Furthermore, the LEHD quarterly statistics used here restrict to the main job, where main job is defined as the higher paying of two jobs if the individual is holding two full-quarter jobs.

The 10th, 50th, and 90th percentiles from the LEHD-quarterly and the published CPS ORG statistics are plotted in Figure 4. All statistics are quarterly, from 1996 to 2012. All series are expressed in natural logs, are in real 2012 dollars, and are seasonally adjusted. The natural log of 13 was added to the natural log of the CPS ORG weekly earnings data in order to graph the CPS ORG data on the same vertical scale as the LEHD quarterly earnings statistics.

The immediate conclusion from Figure 4 is the similarity between the 50th and the 90th percentiles of the CPS ORG and the LEHD statistics. They are almost identical in both level and trend. However, the 10th percentiles are similar in trend but different in level. Transforming the natural logs into levels and dividing by 13, the 10th percentile of the (real) CPS ORG weekly earnings measure is $358 in 2012:Q2, compared to $187 for the LEHD 10th percentile. These findings imply that the left tail of the LEHD quarterly earnings of individuals in full-quarter jobs is quite different than the left tail of the CPS ORG weekly earnings of individuals working full time. Two plausible explanations for this difference are that multiplying the CPS ORG weekly earnings level by 13 is a poor method to compare weekly earnings to quarterly earnings for the lower part of the earnings distribution, or there are some part-time (hours<35) workers with full-quarter jobs in the LEHD.

Given the similarity in trends between the CPS ORG and the LEHD-quarterly earnings percentiles, it should not be surprising that the 90/10, the 90/50, and the 50/10 ratios have the same trends in both data series. The 90/10 ratio, normalized at 1997=100, is presented in the top

[8] A technical note about the 10th, 50th, and 90th percentile statistics computed from the LEHD microdata is warranted. The LEHD program at the Census Bureau does not allow statistics to be reported that might identify a specific individual. As such, percentile statistics such as the median (50th percentile) can not be reported, since this refers to the earnings of a specific individual in the dataset. To protect confidentiality, all medians reported in this paper refer to a modeled estimate, which is created as the average over all individuals between the 49.5th percentile and the 50.5st percentile. The 10th and the 90th percentiles are created similarly.

panel of Figure 5. Both the published CPS ORG statistics and the LEHD-quarterly statistics show that the 90/10 ratio of inequality increased by eight percent between 1997 and the first quarter of 2012. The 90/50 and the 50/10 ratios are presented in the bottom panel of Figure 5 (normalized at 1997=100). Both the published CPS ORG statistics and the LEHD-quarterly statistics show that the 90/50 ratio of inequality increased by 13 to 15 percent between 1997 and 2012, whereas the 50/10 ratio of inequality only increased by 2 to 4 percent during the same timeframe.

In summary, the LEHD measures of the 50th and the 90th quarterly earnings percentiles of individuals in full-quarter main jobs are very similar to the published weekly earnings statistics of individuals working full-time in the CPS ORG earnings data (adjusted for 13 weeks worked in the quarter). Furthermore, the time series of the 90/10, 90/50, and 50/10 ratios in the LEHD quarterly data show similar trends of increasing inequality compared to the published CPS ORG data.

IVC. 90/10 Ratios from the LEHD, Annual

The comparison of the LEHD and the CPS ORG just described is based on quarterly statistics. Comparing the LEHD and the CPS ASEC requires annual statistics. The LEHD annual statistics used in this paper sum the quarterly earnings of all jobs held by an individual (in 20 states) during four quarters. The 10th, 50th, and 90th percentiles from the LEHD-annual and the published CPS ASEC statistics are plotted in Figure 6. All series are expressed in natural logs and are in real 2012 dollars.

Unlike the comparison of the LEHD-quarterly and the CPS ORG statistics, the LEHD-annual and the CPS ASEC statistics are not very similar. The 10th, 50th, and 90th percentile statistics from the CPS ASEC are all higher than the corresponding statistics from the LEHD-annual, with the difference most apparent for the 10th percentile. These differences should not be surprising, since the LEHD measures the annual earnings of individuals, whereas the CPS ASEC measures the annual income of households. Transforming the natural logs into levels, the CPS ASEC 10th percentile is $12,251 in 2011, compared to $2,123 for the LEHD 10th percentile. The 50th percentile in the CPS ASEC is $51,100, compared to $24,490 in the LEHD annual data. The 90th percentile in the CPS ASEC is $146,611, compared to $89,626 in the LEHD annual data. In terms of levels, the distribution of the annual income of households in the published CPS ASEC is much higher at several key percentiles than is the distribution of the annual earnings of individuals in the LEHD annual data.

The top panel of Figure 7 plots the 90/10 ratios from the CPS ASEC and the LEHD-annual (normalized to 1997=100). Whereas the CPS ASEC 90/10 ratio is rising during the 1997 to 2011 time period, the LEHD 90/10 ratio is flat. This flat LEHD 90/10 ratio is unexpected given what we know from the literature and from the published statistics analyzed earlier in this paper. One reason why the LEHD annual 90/10 ratio is flat during the 1997 to 2011 time period becomes apparent when looking at the bottom panel of Figure 7. The 90/50 ratios from the published CPS ASEC statistics and from the LEHD-annual track each other somewhat closely: the CPS ASEC 90/50 ratio rose by 8.6 percent from 1997 to 2011, and the 90/50 ratio from the LEHD-annual rose by 6.5 percent during the same timeframe. The 50/10 ratios from the two

series are both flat between 1997 and 2006, and then diverge after 2006. From 2006 to 2011, the CPS ASEC 50/10 rose by 2.7 percent, whereas the LEHD-annual 50/10 fell by 3.0 percent. Explaining this divergence will require more research.

IVD. Top Percentile Share Statistics from the LEHD

The LEHD annual earnings series that was used in the previous subsection is also used to compute top percentile earnings shares. To restate, this is an annual measure which sums the earnings from all jobs (in 20 states) that an individual holds during the four quarters of the year.[9] Table 4 documents how this LEHD series compares to the published CPS ASEC and the IRS top percentile earnings shares in terms of scope and definition. There are two main things to note in this comparison. First, the LEHD is individual based whereas the CPS ASEC and IRS top percentile shares refer to households and tax units, respectively. Second, the LEHD measure is based on earnings, whereas the CPS ASEC measure is based on income; the IRS data compiled by Saez (2013) have two series – one for income and one for wage income.

The top panel of Figure 8 plots the top five percent earnings and income shares from various series, and the bottom panel plots the top one percent earnings and income shares from various series. In the top panel, the LEHD top 5 percent share is below the IRS income measure (excluding capital gains), and is above the SSA measure, the IRS wage income measure, and the CPS ASEC measure. The IRS income measure states that 34.0 percent of all income (excluding capital gains) is held by the top five percent of tax units in 2011; the LEHD states that 29.9 percent of all earnings is held by the top five percent of individuals in 2011. Statistics for the SSA, IRS wage income, and CPS ASEC are 28.2 percent, 24.2 percent, and 22.3 percent, respectively. All five series report a sizable amount of income or earnings is held by the top five percent.

In terms of growth, the LEHD top five percent share is growing by 7.1 percent during the 1997 to 2011 time period. This is less than the 13.8 percent reported in the IRS income data and the 11.4 percent reported in the SSA data, but greater than the 3.8 percent reported in the IRS wage data and the 2.8 percent reported in the published CPS ASEC data. In terms of correlations, the time series patterns across many of the datasets are similar. Based upon the 1997 to 2011 time period, the LEHD top 5 percent share and the IRS wage income top 5 percent share have a correlation of .885; the LEHD's correlation with the other published data series are .811 with the SSA series, .691 with the IRS income series (excluding capital gains), and .260 with the CPS ASEC series.

The top one percent share of earnings (or income) in the various series is reported in the bottom panel of Figure 8. The LEHD series matches well with the IRS and the SSA series in terms of both levels and time series. The time series correlation of the LEHD series with the IRS wage income series is .950, is .881 with the SSA earnings series, and is .561 with the IRS income (excluding capital gains) series.

[9] To be consistent with the computation of 90/10 ratios in the LEHD, the top 5 percent and the top 1 percent earnings shares in the LEHD are defined from a modeled estimate of the 95th and the 99th percentiles.

V. Detailed Inequality Statistics

This section presents extensions of the LEHD statistics from sub-sections IVB and IVD above. The simple methodology is to use the large sample sizes and the linked employee-employer structure of the LEHD in order to present inequality statistics by demographic and firm characteristics.

VA. 90/10 Ratios by Firm Size

The LEHD inequality statistics in sub-section IVB were 90/10 ratios for individuals in their main full-quarter job. These statistics were found to be quite similar to the CPS ORG statistics. In this sub-section, we present these LEHD 90/10 ratios by select firm characteristics, which is something not possible using the CPS ORG statistics.

Figure 9 presents the 10^{th}, 50^{th}, and 90^{th} earnings percentiles by four categories of firm size: firms with less than 20 employees, firms with 20 to 99 employees, firms with 100 to 999 employees, and firms with 1000 or more employees.[10] The weighted average across firm size categories for each percentile in Figure 9 is the LEHD series in Figure 4. For each of the percentiles in Figure 4 – the 10^{th}, the 50^{th}, and the 90^{th} – earnings are monotonically increasing in firm size: at each of these points in the earnings distribution, small firms pay less (on average) than large firms.

The 90/10 ratios by firm size are presented in Figure 10a. To set the context, recall from Figure 5 that the LEHD 90/10 ratio rose by eight percent between 1997 and 2012. We learn from Figure 10a that this inequality growth varies substantially by firm size. Individuals whose main full quarter job is in small firms with less than 20 employees exhibit little inequality growth – the 90/10 ratio only rises by two percent between 1997 and 2012. On the other hand, the 90/10 ratio increases by over twelve percent for individuals whose main full quarter job is in large firms with more than 1000 employees.

These differences in inequality growth by firm size are evident in the top half of the earnings distribution. The top panel of Figure 10b shows that the 90-50 ratio grows by 21 percent in large firms, which is much higher than the 5 percent growth in small firms. The bottom panel of Figure 10b shows that the 50-10 ratio is relatively flat over time in all firm sizes.

VB. The Role of the Firm

These 90/10 and 90/50 ratios by firm size show that the firm plays a role in the magnitude of increasing earnings inequality. In this subsection, this role of the firm is examined more formally.

For the purposes of this subsection only, we switch from 90/10 ratios and examine variances, since we can decompose the variance into the within and across firm components. Letting "i" index individuals and "f" index firms, the variance of earnings is:

[10] Firm information is linked into the LEHD from the Census Bureau's Business Dynamics Statistics (BDS), and refers to the size of the national firm. See Haltiwanger, Hyatt, McEntarfer, Sousa, and Tibbets (2014) for details.

$$V(\$) = \left(\frac{1}{N}\right)\sum_{i=1}^{N}\left(\$_{if} - \bar{\$}\right)^2$$

$$= \left(\frac{1}{N}\right)\sum_{i=1}^{N}\left(\$_{if} - \overline{\$_f}\right)^2 + \left(\frac{1}{N}\right)\sum_{i=1}^{N}\left(\overline{\$_f} - \bar{\$}\right)^2$$

$$= \left[\frac{\sum_{f=1}^{F} N_f V_f(\$)}{\sum_{f=1}^{F} N_f}\right] + \left[\frac{\sum_{f=1}^{F} N_f\left(\overline{\$_f} - \bar{\$}\right)^2}{\sum_{f=1}^{F} N_f}\right]$$

$$= \begin{bmatrix} Within\ firm \\ Variance \end{bmatrix} + \begin{bmatrix} Across\ firm \\ Variance \end{bmatrix}$$

The analysis here builds on an existing literature.[11] This literature tells us that about half of cross-sectional earnings variance is across businesses, and between half and all of the growth in earnings variance is across businesses.

Using the LEHD quarterly earnings of individuals in their main full-quarter job, the variance of earnings is given in the top panel of Figure 11. The earnings data underlying this graph has been winsorized at the 99th percentile of the state-year-quarter distribution to minimize the effects of outliers, and the quarterly time series of variance has been seasonally adjusted. The variance increases during the 1996 to 2012 period, which is interpreted as increasing earnings inequality.

The interesting graph is the bottom panel of Figure 11. On average, 50.3 percent of cross-sectional earnings variance in this LEHD quarterly data is across firms.[12] This 50 percent estimate is consistent with the existing literature. In this LEHD data, 93.5 percent of the growth of earnings variance is across firms. Intuitively, increasing inequality in the United States during the 1996 to 2012 time period is driven by an increasing variance of average wages across firms rather then by an increasing variance across workers within firms. This obviously reflects sorting of workers into firms (Abowd, Kramarz, and Margolis, 1999, Card, Heining, and Kline, 2013), yet quantifying the exact type of sorting requires further research.

VC. Top Percentile Shares by Age and Gender

The LEHD inequality statistics in sub-section IVD were top five percent and top one percent earnings shares (computed on an annual basis for all jobs held by an individual). The data presented in Figure 8 showed that the LEHD statistics are quite similar to the publicly available SSA statistics, and lie midway between the IRS income (excluding capital gains) and the IRS wage income statistics. In this sub-section, we present these LEHD top earnings shares by select demographic characteristics of the worker, which is something (to the best of my knowledge) not publicly available from the SSA or the IRS statistics.

[11] A non-exhaustive literature review is Groshen (1991), Davis and Haltiwanger (1991), Dunne, Foster, Haltiwanger, and Troske (2004), and Handwerker and Spletzer (2014).

[12] A technical note warrants mentioning. The definition of firm underlying this graph is the State UI account, rather than the national EIN firm identifier. Quantifying the effects of different definitions of firms will be done in future drafts of this paper.

The column of the top panel of Table 5 labeled "All Workers" presents the age composition of all workers, averaged over all years 1997 through 2011. We see that 31.8 percent of workers in the LEHD data are less than 30 years old, and 11.2 percent are between 30 and 34 years old. The right column of the top panel of Table 5 presents the age composition of workers who are in the top five percent of the earnings distribution. Not surprisingly, there are very few young workers in the top five percent of the earnings distribution: only 2.7 percent of these workers are less than 30, and 8.8 percent are between 30 and 34 years old. Relative to all workers in the population, workers in the top five percent of the earnings distribution are much more likely to be between the ages of 40 and 54. The data in Table 5 show that 53.4 percent of the workers in the top five percent are between the ages of 40 and 54, relative to 30.9 percent in the working population.

The bottom panel of Table 5 presents statistics for the gender composition of workers, averaged over all years 1997 through 2011. In the population of all workers, 53.1 percent of workers are male. However, males are 79.0 percent of workers in the top five percent of the earnings distribution.

VI. Questions for the FESAC Members

As is customary with papers written for the Federal Economic Statistics Advisory Committee, this paper concludes with several questions for the committee members:

1) Are the inequality statistics currently in the public domain sufficient for analysts and policymakers, or should the statistical agencies publish similar 90/10 ratios and top five percent earnings distribution statistics from other data sources?

2) Do you have any thoughts about the similarities and differences in inequality statistics that exist from different data sources? Differences undoubtedly reflect different concepts and definitions across data sources, but differences might also signal issues that the statistical agencies should research and understand.

3) One goal of publishing inequality statistics is to help understand why inequality is increasing. As such, are there any specific statistics from the LEHD that you would find useful?

References

Abowd, John M., Francis Kramarz, and David Margolis. 1999. "High Wage Workers and High Wage Firms." *Econometrica*, pp. 251-334.

Abowd, John M., Bryce E. Stephens, Lars Vilhuber, Fredrik Andersson, Kevin L. McKinney, Marc Roemer, and Simon Woodcock. 2009. "The LEHD Infrastructure Files and the Creation of the Quarterly Workforce Indicators." In *Producer Dynamics*, edited by Timothy Dunne, J. Bradford Jensen, and Mark J. Roberts, pp. 149-230, Chicago: University of Chicago Press.

Acemoglu, Daron and David Autor. 2011. "Skills, Tasks, and Technologies: Implications for Employment and Earnings." In *Handbook of Labor Economics*, edited by Orley Ashenfelter and David Card, pp. 1043-1171, London: Elsevier.

Autor, David H. 2014. "Skills, education, and the rise of earnings inequality among the 'other 99 percent.'" *Science*, vol. 344 (6186), pp. 843-851.

Bound, John and George Johnson. 1992. "Changes in the Structure of Wages in the 1980's: An Evaluation of Alternative Explanations." *The American Economic Review*, vol. 82(3), pp. 371-392.

Card, David, Jörg Heining, and Patrick Kline. 2013. "Workplace Heterogeneity and the Rise of West German Wage Inequality." *The Quarterly Journal of Economics*, vol. 128(3), pp. 967-1015.

Davis, Steve J. and John Haltiwanger. 1991. "Wage Dispersion Between and Within U.S. Manufacturing Plants." *Brookings Papers on Economic Activity*, pp. 115-200.

Dunne, Timothy, Lucia Foster, John Haltiwanger, and Kenneth R. Troske. 2004. "Wage and Productivity Dispersion in United States Manufacturing: The Role of Computer Investment." *Journal of Labor Economics*, Vol. 22, No. 2, pp. 397-429.

Groshen, Erica L. 1991. "Sources of Intra-Industry Wage Dispersion: How Much do Employers Matter?" *The Quarterly Journal of Economics*, Vol. 106, No. 3, pp. 869-884.

Haltiwanger, John, Henry Hyatt, Erika McEntarfer, Liliana Sousa, and Stephen Tibbets. 2014. "Firm Age and Size in the Longitudinal Employer-Household Dynamics Data." Center for Economic Studies Working Paper #14-16, U.S. Census Bureau, http://ideas.repec.org/p/cen/wpaper/14-16.html.

Handwerker, Elizabeth Weber and James R. Spletzer. 2014. "Occupational Concentration, Wages, and Growing Wage Inequality." Paper presented at the annual AEA meetings, Philadelphia PA.

Juhn, Chinhui, Kevin M. Murphy, and Brooks Pierce. 1993. "Wage Inequality and the Rise in Returns to Skill." *Journal of Political Economy*, vol. 101 (3), pp. 410-442.

Katz, Lawrence F. and Kevin M. Murphy. 1992. "Changes in Relative Wages, 1963-1987: Supply and Demand Factors." *The Quarterly Journal of Economics*, vol. 107(1), pp. 35-78.

Levy, Frank and Richard J. Murnane. 1992. "U.S. Earnings Levels and Earnings Inequality: A Review of Recent Trends and Proposed Explanations." *Journal of Economic Literature*, vol. 30(3), pp. 1333-1381.

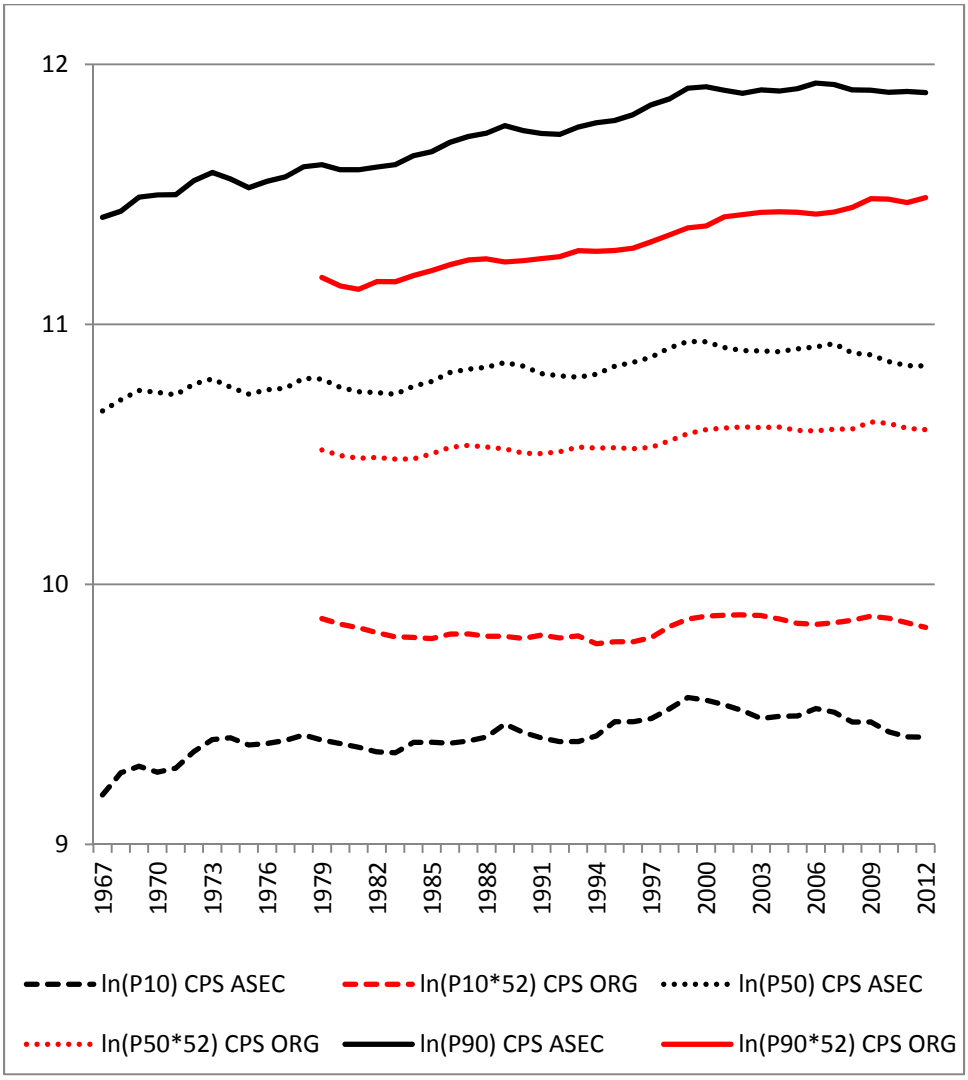
Piketty, Thomas and Emmanuel Saez. 2003. "Income Inequality in the United States, 1913-1998." *The Quarterly Journal of Economics*, vol. 118(1), pp. 1-39.

Saez, Emmanuel. 2013. "Striking it Richer: The Evolution of Top Incomes in the United States (Updated with 2012 preliminary estimates)." Manuscript, University of California Berkeley, http://elsa.berkeley.edu/users/saez/saez-UStopincomes-2012.pdf.

Table 1: 10^{th}, 50^{th}, and 90^{th} Percentiles
Publicly Available CPS ORG and CPS ASEC

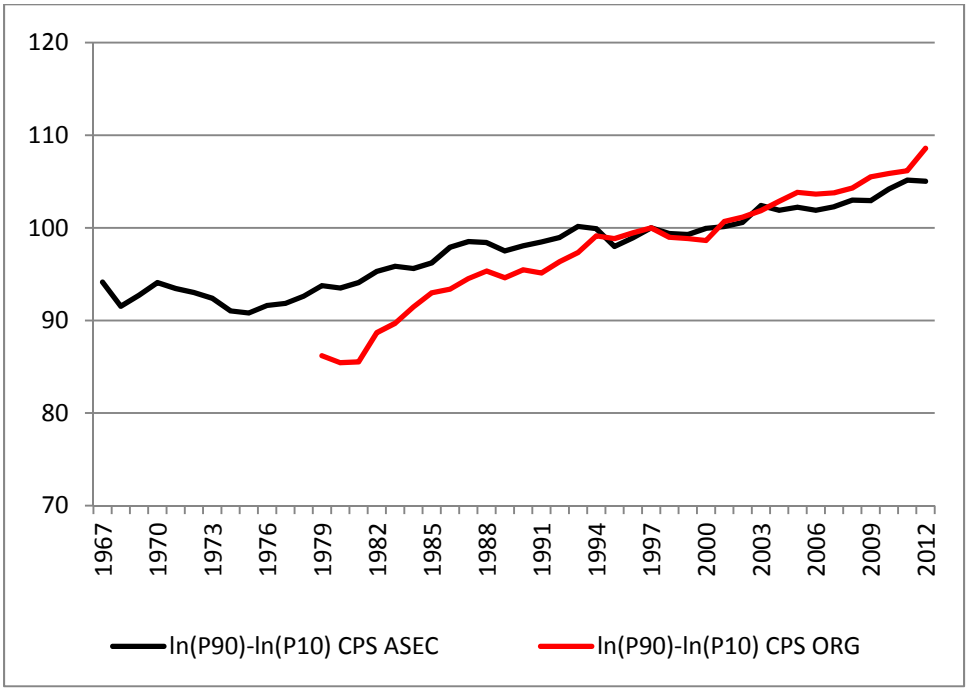
	CPS ORG	CPS ASEC
Data Available	10^{th}, 50^{th}, & 90^{th} percentiles Annual, 1979 – current Quarterly, 1994 – current also available, but not used here comparing to ASEC	10^{th}, 50^{th}, & 90^{th} percentiles Annual, 1967 - current
Source	Current Population Survey (CPS) Outgoing Rotation Groups (ORG) 2000 - current data from http://www.bls.gov/webapps/legacy/cpswktab5.htm, earlier data available upon request	Current Population Survey (CPS) Annual Social and Economic Supplement (ASEC) Table A-2 of http://www.census.gov/prod/2013pubs/p60-245.pdf
Scope and definition (simplified)	Weekly earnings of individuals working full-time	Annual income of households
Scope	Usual weekly earnings of full-time wage and salary workers 16 years and over (at their main job) Quarterly earnings are tabulated from three months of interviews; annual earnings are tabulated from twelve months of interviews	Previous year income (from all jobs) for all persons aged 15 and over currently residing in the household
Measure of earnings and income	Data represent earnings before taxes and other deductions and include any overtime pay, commissions, or tips usually received Earnings of the self-employed are excluded, regardless of whether their businesses are incorporated.	Income is earnings from wages and salaries, net income from self-employment, plus income from 17 other sources such as social security, pension income, public assistance, interest, dividends, rents, … Capital gains are not included in income. Wage and salary earnings includes commissions, tips, and cash bonuses earned from all jobs held during the previous calendar year, before payments for personal income taxes, social security, etc.

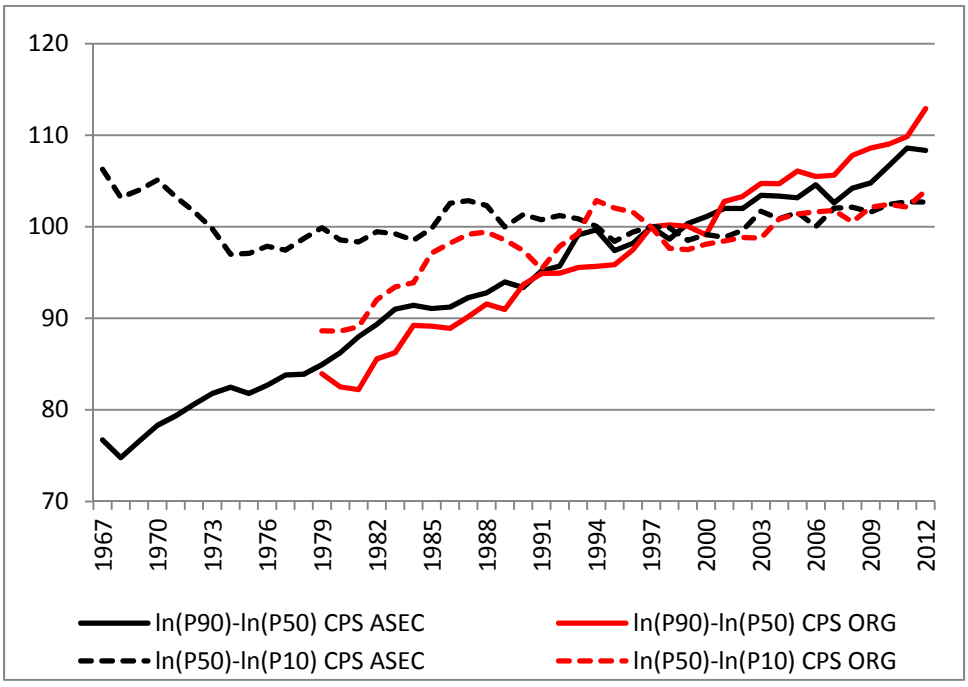
Figure 1: 10th, 50th, and 90th Percentiles, 1967-2012
Publicly Available CPS ORG and CPS ASEC



Notes: CPS ASEC and CPS ORG are defined in Table 1.
All series are in real 2012 dollars.

Figure 2: 90/10, 90-50, and 50-10 Ratios (1997=100), 1967-2012
Publicly Available CPS ASEC and CPS ORG



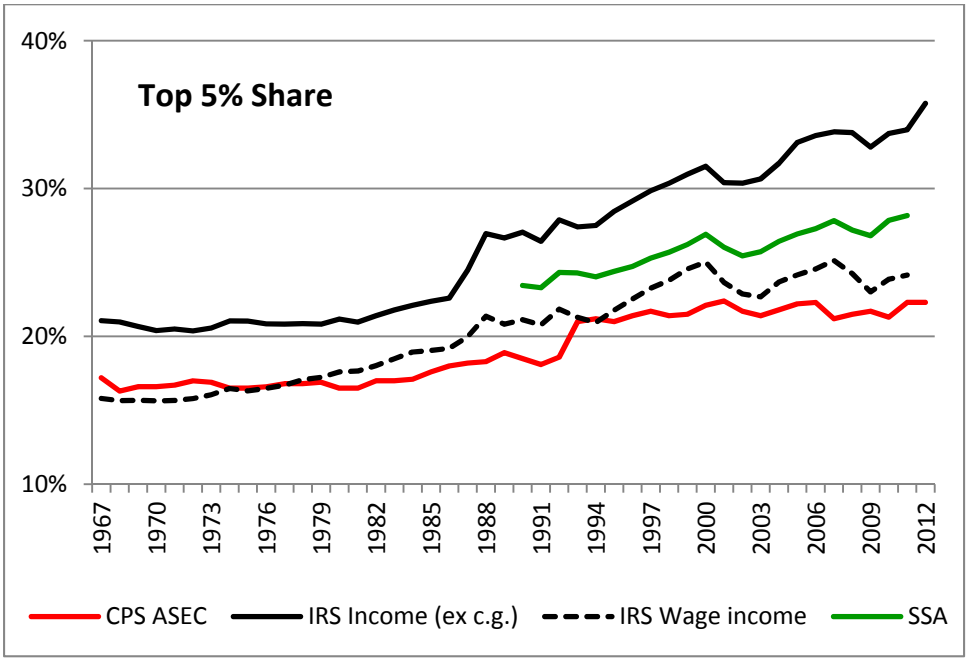


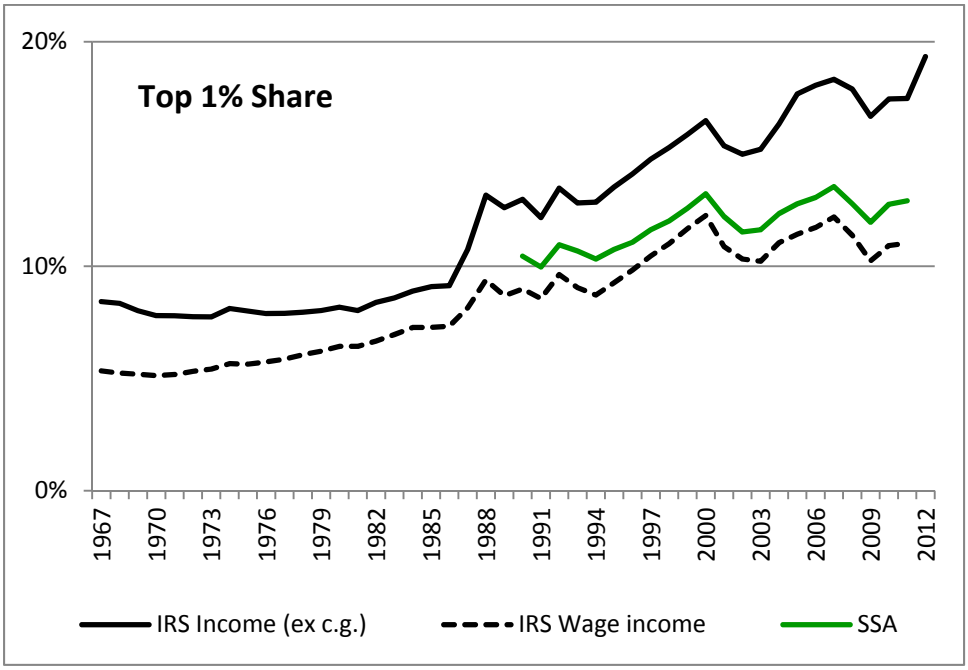
Notes: CPS ASEC and CPS ORG are defined in Table 1.
All series are in real 2012 dollars.

Table 2: Top 10%, 5%, and 1% Shares
Publicly Available CPS ASEC, IRS (Saez), and SSA (Saez)

	CPS ASEC	IRS (Saez)	SSA (Saez)
Data Available	Top 5% share Annual, 1967 - current	Top 10%, 5%, and 1% share Annual, 1917 - current for income, 1927 – current for wage income	Top 10%, 5%, and 1% share Annual, 1990 – current
Source	Current Population Survey (CPS) Annual Social and Economic Supplement (ASEC) Table A-2 of http://www.census.gov/prod/2013pubs/p60-245.pdf	ITS tax return statistics, compiled and updated annually by Emmanuel Saez Tables A1 and B2 of http://elsa.berkeley.edu/~saez/TabFig2012prel.xls	SSA tabulations of annual individual wage earnings, compiled and updated annually by Emmanuel Saez Table B5 of http://elsa.berkeley.edu/~saez/TabFig2012prel.xls
Scope and definition (simplified)	Annual income of households	Annual income (or annual salaries and wages) of tax units	Annual earnings of individuals
Scope	Previous year income (from all jobs) for all persons aged 15 and over currently residing in the household	Annual gross income reported by tax units, before deductions (from all jobs) A tax unit is defined as a family (such as a couple with dependents, or a head of household with dependents, or a single person). A family can be smaller than a household.	
Measure of earnings and income	Income is earnings from wages and salaries, net income from self-employment, plus income from 17 other sources such as social security, pension income, public assistance, interest, dividends, rents, … Capital gains are not included in income. Wage and salary earnings includes commissions, tips, and cash bonuses earned from all jobs held during the previous calendar year, before payments for personal income taxes, social security, etc.	Annual gross income includes salaries and wages, small business and farm income, partnership and fiduciary income, dividends, interest, rents, royalties, and other small items reported as other income. All data reported in this paper exclude realized capital gains. Separate series are presented for income (as defined above) and the salaries and wages component of income	Wage income is defined as W2 wage income (wages, salaries, and tips) inclusive of elective retirement contributions (such as 401(k) contributions). Wage income also includes bonuses, and profits from exercised stock options.

Figure 3: Top 5% and 1% Shares, 1967-2012
Publicly Available CPS ASEC, IRS (Saez), and SSA (Saez)



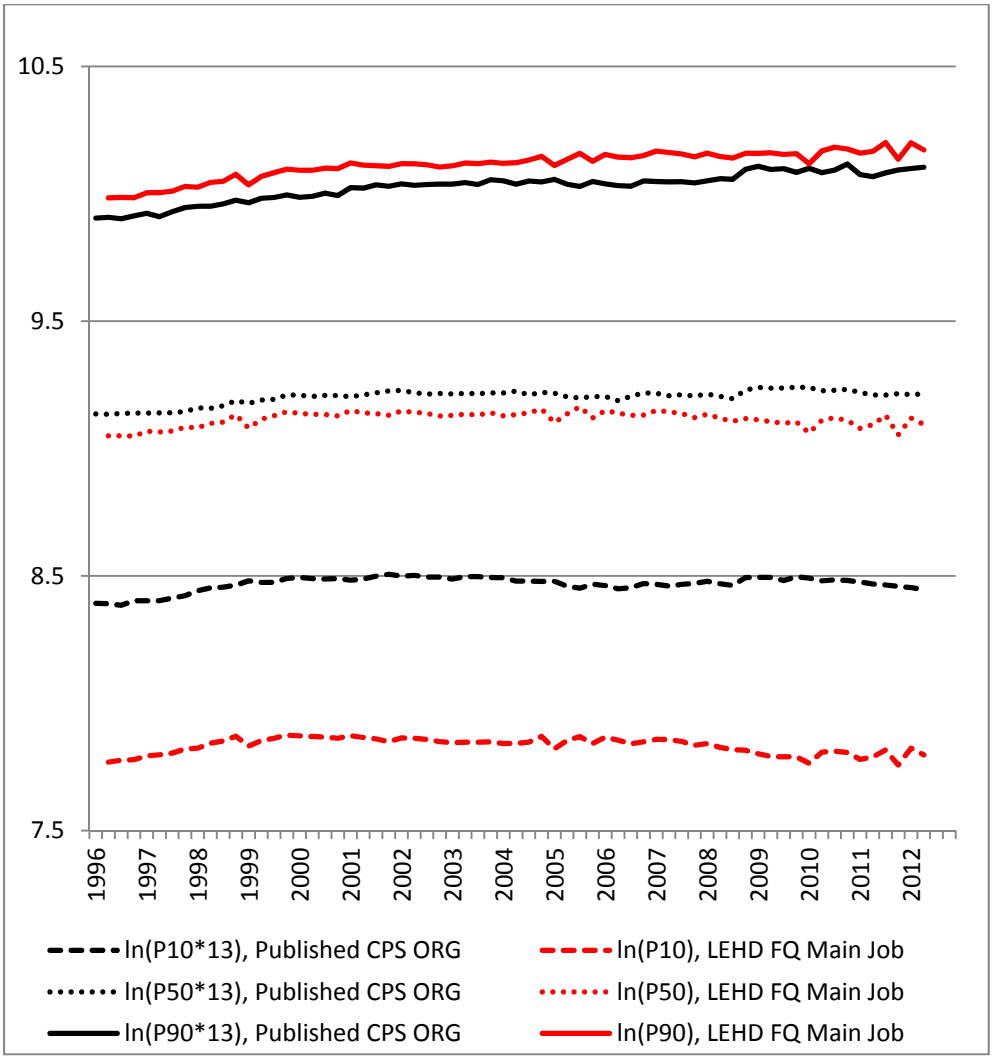


Notes: CPS ASEC, IRS, and SSA series are defined in Table 2.
All series are in real 2012 dollars.

Table 3: 10^{th}, 50^{th}, and 90^{th} Percentiles
CPS ORG, CPS ASEC, and LEHD

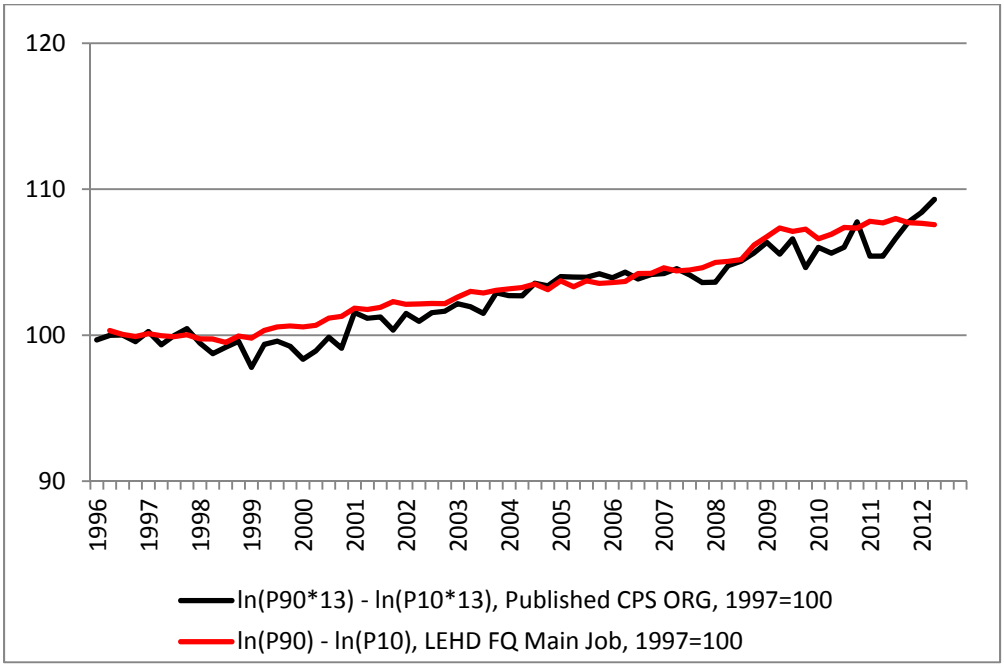
	CPS ORG	CPS ASEC	LEHD
Data Available	10^{th}, 50^{th}, & 90^{th} percentiles Quarterly, 1994 - current	10^{th}, 50^{th}, & 90^{th} percentiles Annual, 1967 - current	10^{th}, 50^{th}, & 90^{th} percentiles Quarterly, 1996:2 – current Annual, 1997 – current
Source	Current Population Survey (CPS) Outgoing Rotation Groups (ORG) 1994 - current data from http://www.bls.gov/webapps/legacy/cpswktab5.htm, earlier data available upon request	Current Population Survey (CPS) Annual Social and Economic Supplement (ASEC) Table A-2 of http://www.census.gov/prod/2013pubs/p60-245.pdf	Longitudinal Employer Household Dynamics (LEHD) Tabulations from confidential microdata
Scope and definition (simplified)	Weekly earnings of individuals working full-time	Annual income of households	Quarterly earnings of individuals in full-quarter jobs Annual earnings of individuals
Scope	Usual weekly earnings of full-time wage and salary workers 16 years and over (at their main job) Quarterly earnings are tabulated from three months of interviews; annual earnings are tabulated from twelve months of interviews	Previous year income (from all jobs) for all persons aged 15 and over currently residing in the household	Quarterly earnings of workers 16 years and over in their full-quarter main job. Full-quarter jobs are those where the worker is at the same employer in both the previous and the following quarter Annual earnings for persons aged 15 and over (from all jobs) 20 states, private sector
Measure of earnings and income	Data represent earnings before taxes and other deductions and include any overtime pay, commissions, or tips usually received Earnings of the self-employed are excluded, regardless of whether their businesses are incorporated.	Income is earnings from wages and salaries, net income from self-employment, plus income from 17 other sources such as social security, pension income, public assistance, interest, dividends, rents, … Capital gains are not included in income. Wage and salary earnings includes commissions, tips, and cash bonuses earned from all jobs held during the previous calendar year, before payments for personal income taxes, social security, etc.	Earnings include gross wages and salaries, bonuses, stock options, tips and other gratuities, and the value of meals and lodging, where supplied.

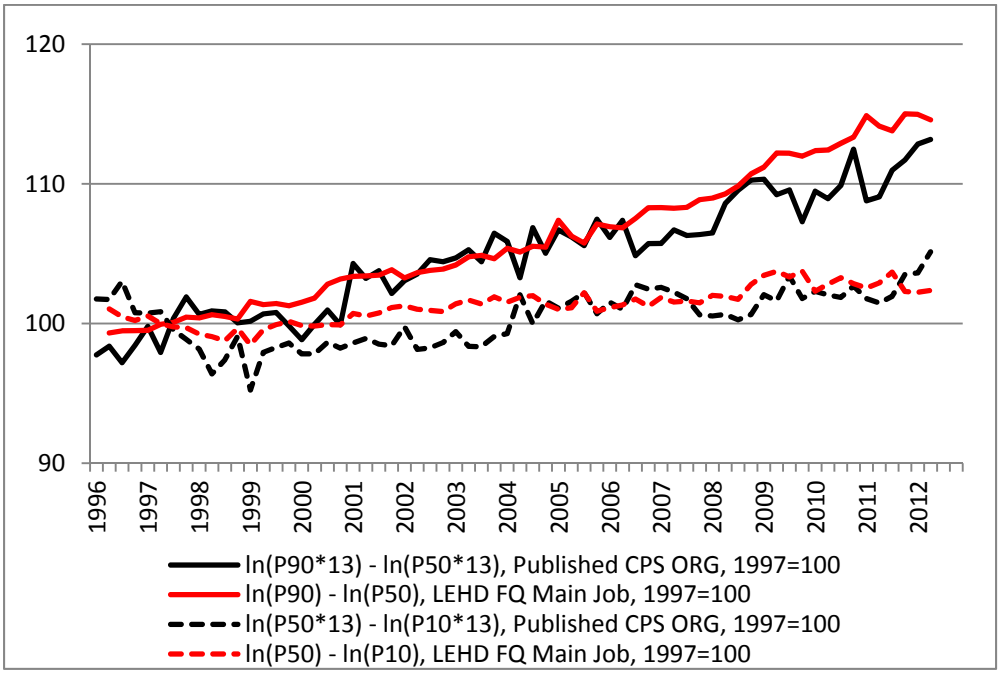
Figure 4: 10th, 50th, and 90th Percentiles, 1996-2012
CPS ORG and LEHD, Quarterly



Notes: CPS ORG and LEHD are defined in Table 3.
All series are in real 2012 dollars. All series are seasonally adjusted.

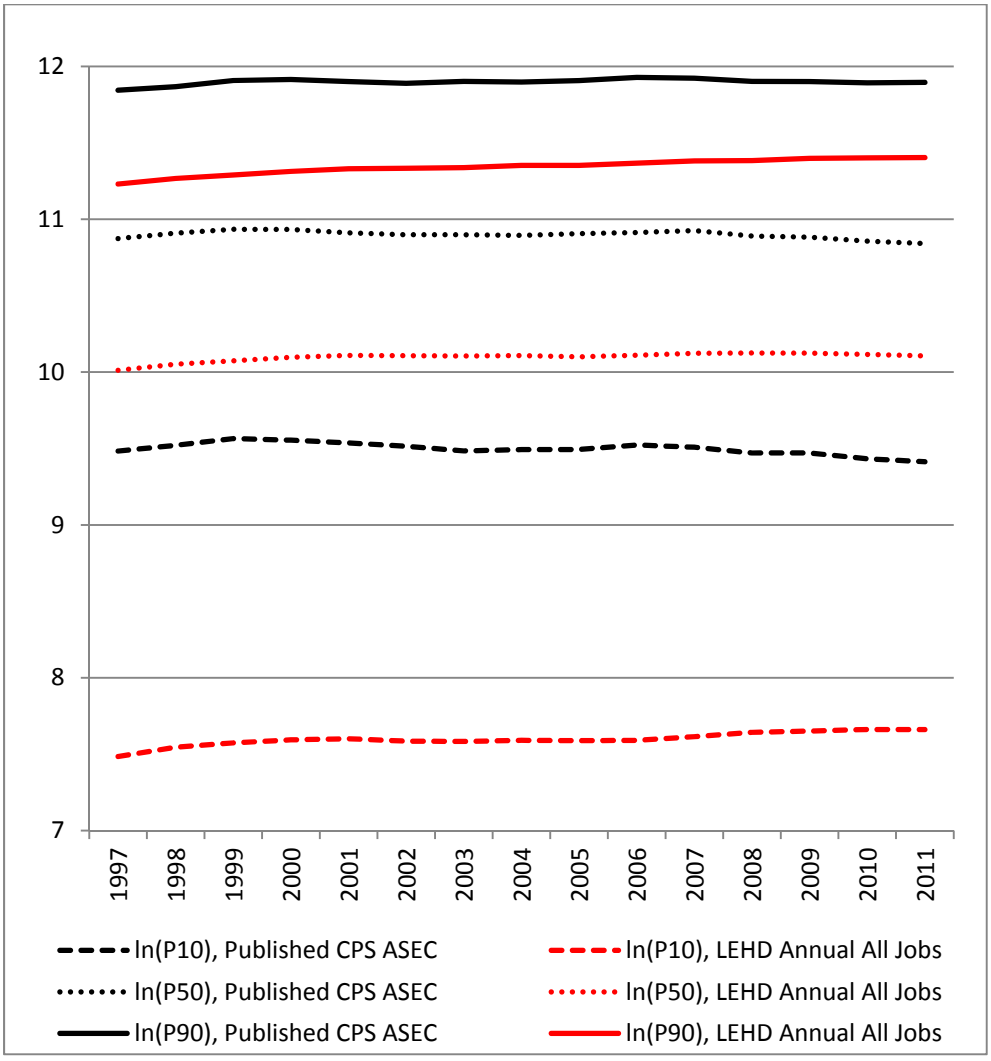
Figure 5: 90/10, 90-50, and 50-10 Ratios (1997=100), 1996-2012
CPS ORG and LEHD, Quarterly





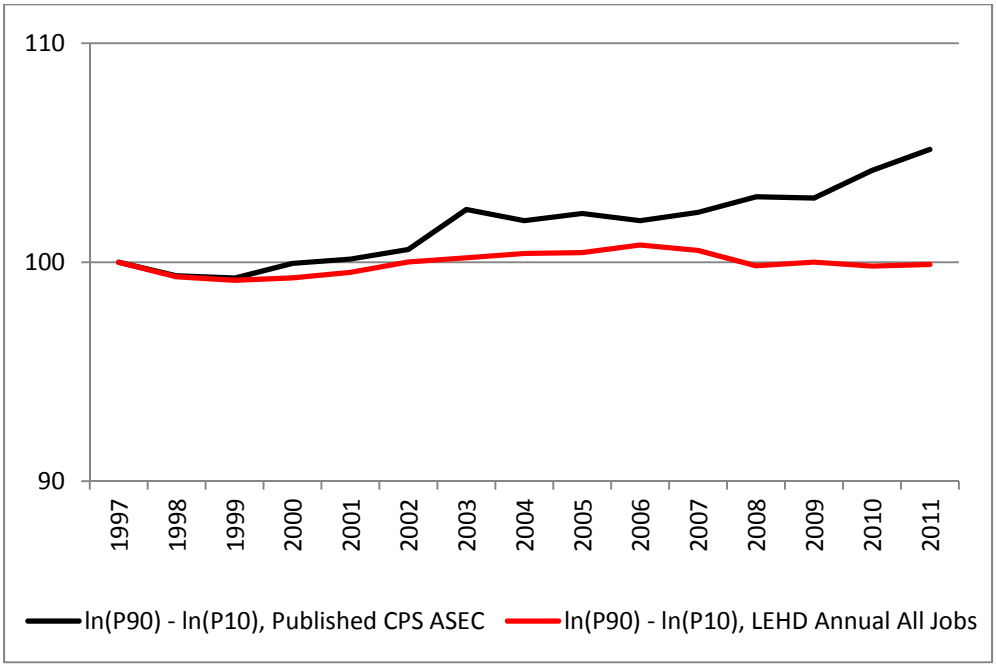
Notes: CPS ORG and LEHD are defined in Table 3.
All series are in real 2012 dollars. All series are seasonally adjusted.

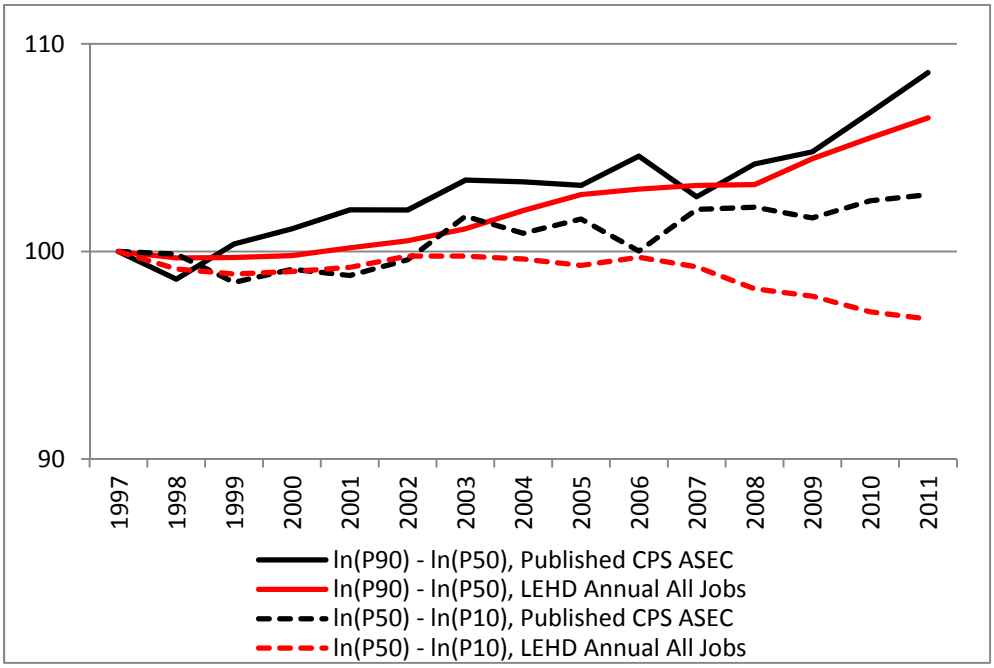
Figure 6: 10th, 50th, and 90th Percentiles, 1997-2011
CPS ASEC and LEHD, Annual



Notes: CPS ASEC and LEHD are defined in Table 3.
All series are in real 2012 dollars.

Figure 7: 90/10, 90-50, and 50-10 Ratios (1997=100), 1997-2011
CPS ASEC and LEHD, Annual





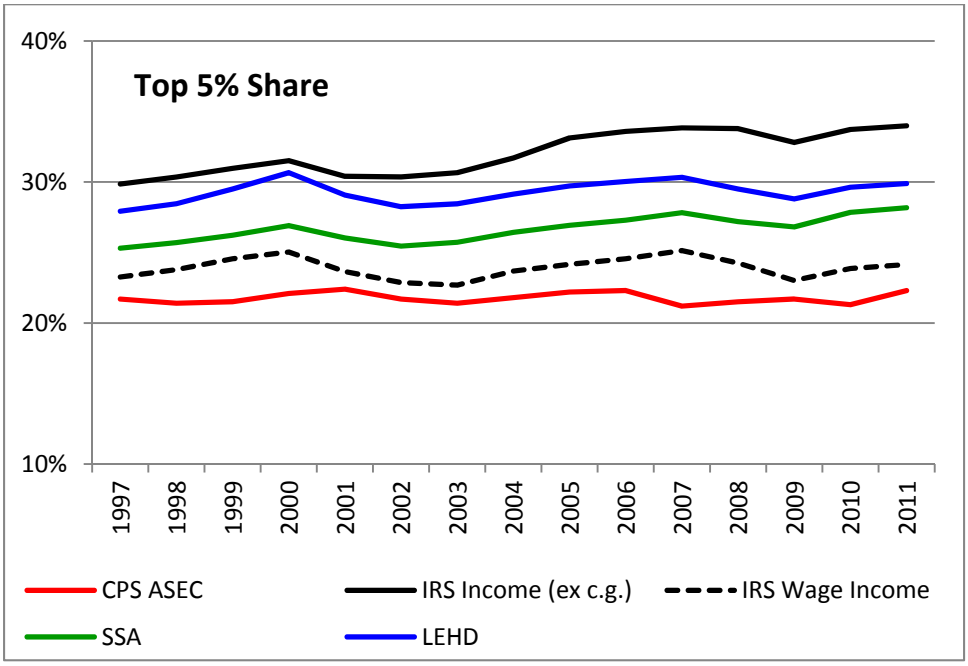
Notes: CPS ASEC and LEHD are defined in Table 3.
All series are in real 2012 dollars.

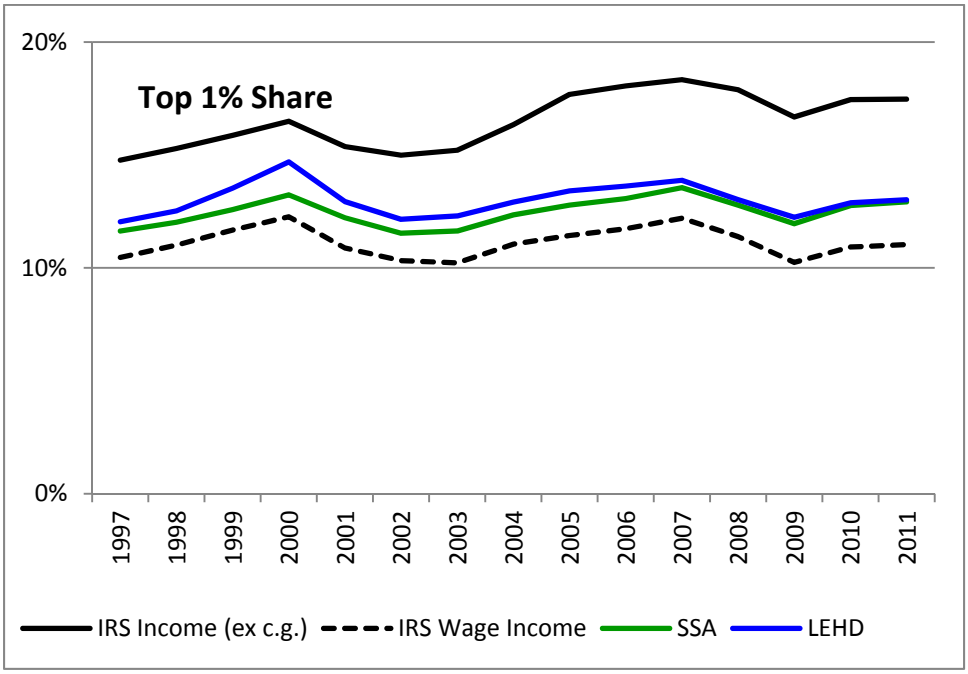
Table 4: Top 10%, 5%, and 1% Shares
CPS ASEC, IRS (Saez), and LEHD

	CPS ASEC	IRS (Saez)	LEHD
Data Available	Top 5% share Annual, 1967 - current	Top 10%, 5%, and 1% share Annual, 1917 - current for income, 1927 – current for wage income	Top 10%, 5%, and 1% share Annual, 1997 – current
Source	Current Population Survey (CPS) Annual Social and Economic Supplement (ASEC) Table A-2 of http://www.census.gov/prod/2013pubs/p60-245.pdf	ITS tax return statistics, compiled and updated annually by Emmanuel Saez Tables A1 and B2 of http://elsa.berkeley.edu/~saez/TabFig2012prel.xls	Longitudinal Employer Household Dynamics (LEHD) Tabulations from confidential microdata
Scope and definition (simplified)	Annual income of households	Annual income (or annual salaries and wages) of tax units	Annual earnings of individuals
Scope	Previous year income (from all jobs) for all persons aged 15 and over currently residing in the household	Annual gross income reported by tax units, before deductions. A tax unit is defined as a family (such as a couple with dependents, or a head of household with dependents, or a single person). A family can be smaller than a household.	Annual earnings for persons aged 15 and over (from all jobs) 20 states, private sector
Measure of earnings and income	Income is earnings from wages and salaries, net income from self-employment, plus income from 17 other sources such as social security, pension income, public assistance, interest, dividends, rents, … Capital gains are not included in income. Wage and salary earnings includes commissions, tips, and cash bonuses earned from all jobs held during the previous calendar year, before payments for personal income taxes, social security, etc.	Annual gross income includes salaries and wages, small business and farm income, partnership and fiduciary income, dividends, interest, rents, royalties, and other small items reported as other income. All data reported in this paper exclude realized capital gains. Separate series are presented for income (as defined above) and the salaries and wages component of income	Earnings include gross wages and salaries, bonuses, stock options, tips and other gratuities, and the value of meals and lodging, where supplied.

Notes: See Table 2 for a description of the top 5% and 1% earnings shares in the SSA data.

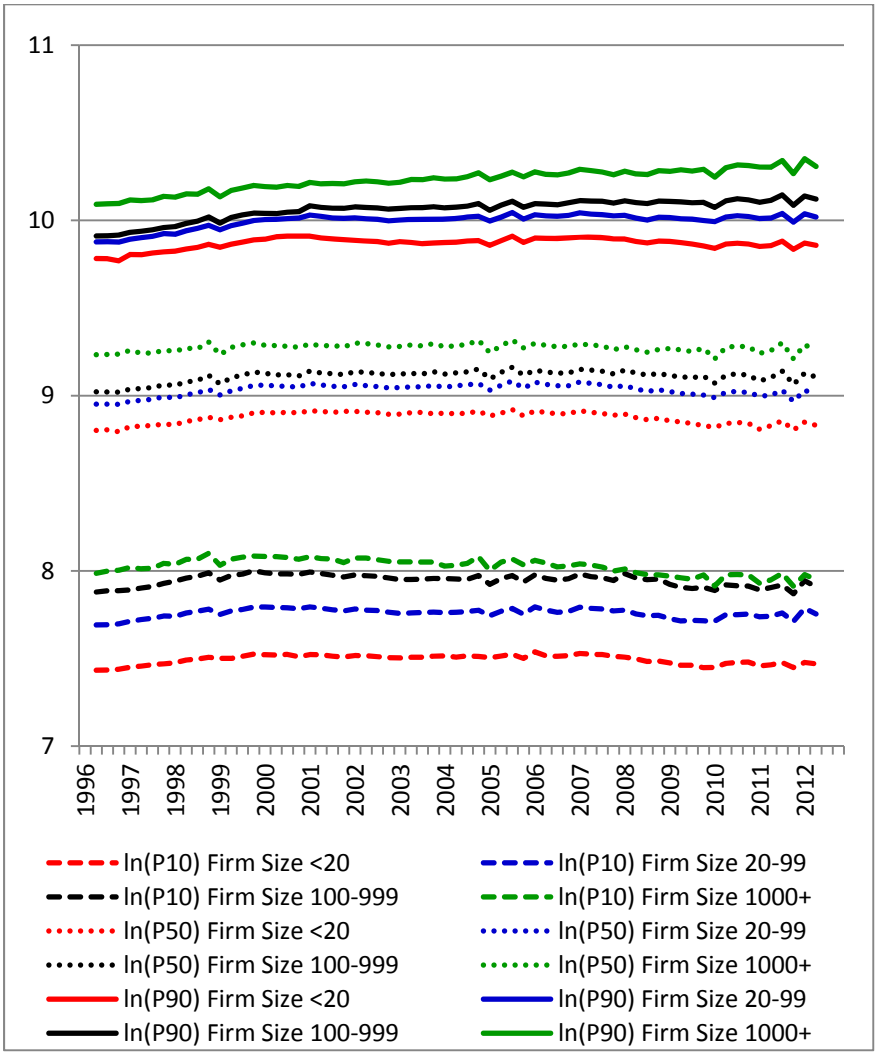
Figure 8: Top 5% and 1% Shares, 1997-2011
CPS ASEC, IRS (Saez), SSA (Saez), and LEHD





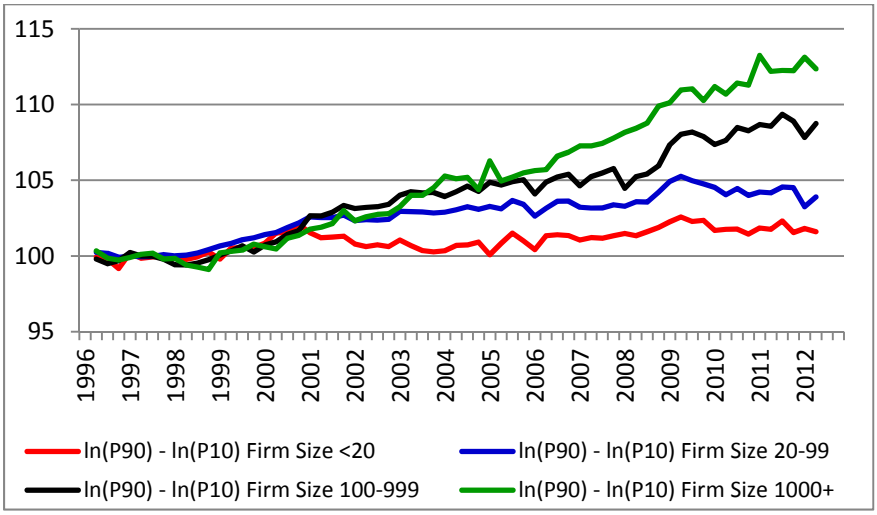
Notes: CPS ASEC, IRS, SSA, and LEHD are defined in Table 4.
All series are in real 2012 dollars.

Figure 9: 10th, 50th, and 90th Percentiles, 1996-2012, by Firm Size LEHD, Quarterly



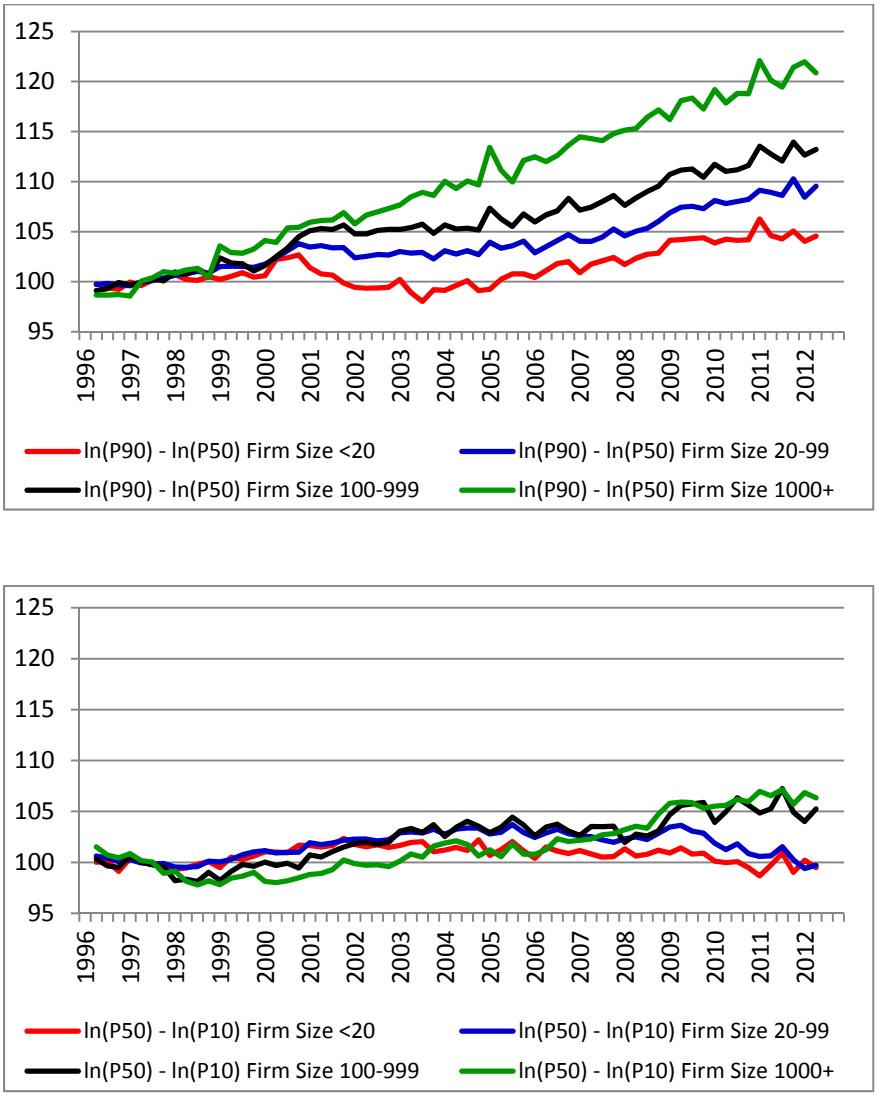
Notes: LEHD is defined in Table 3.
All series are in real 2012 dollars. All series are seasonally adjusted.

Figure 10a: 90/10 Ratios (1997=100), 1996-2012, by Firm Size LEHD, Quarterly



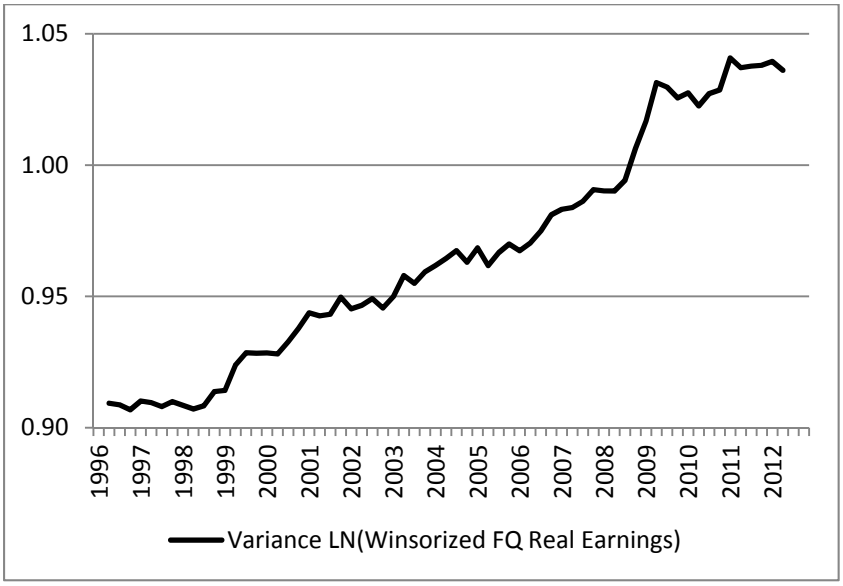
Notes: LEHD is defined in Table 3.
All series are in real 2012 dollars. All series are seasonally adjusted.

Figure 10b: 90-50 and 50-10 Ratios (1997=100), 1996-2012, by Firm Size LEHD, Quarterly



Notes: LEHD is defined in Table 3.
All series are in real 2012 dollars. All series are seasonally adjusted.

Figure 11: Variance Decomposition, 1996-2012, Within and Across Firms LEHD, Quarterly



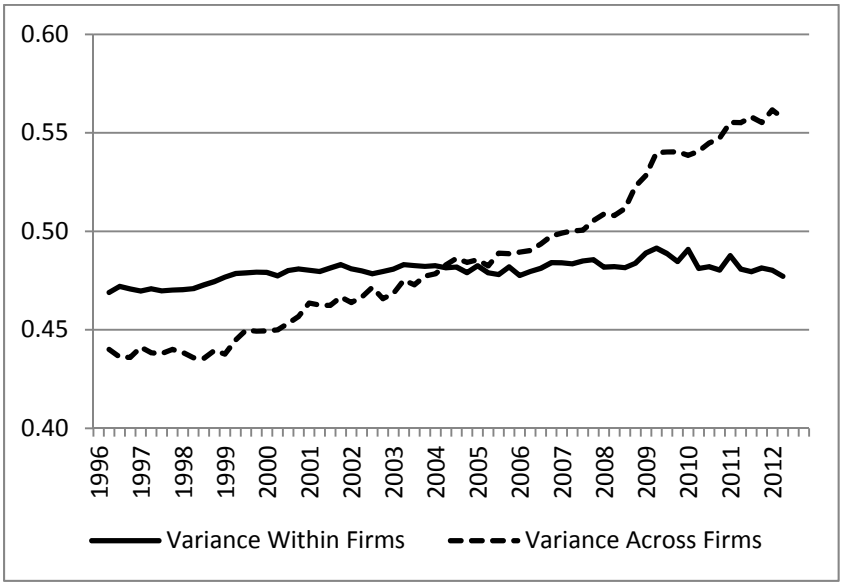


Table 5: Age and Gender Composition, All Workers and Workers in Top 5% Share LEHD Annual, 1997-2011

Age	All Workers	Workers in Top 5% Earnings Share
<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%
Total	100.0%	100.0%

Gender	All Workers	Workers in Top 5% Earnings Share
Male	53.1%	79.0%
Female	46.9%	21.0%
Total	100.0%	100.0%

Notes: LEHD is defined in Table 4.
All series are in real 2012 dollars.