Reconciling the divergence in aggregate U.S. wage series*

Julien Champagne  André Kurmann
Bank of Canada    Federal Reserve Board

November 4, 2013

Abstract

This paper documents the gradual divergence in trend growth and business cycle volatility of two popular aggregate hourly wage series for the U.S. economy: average hourly compensation from the Labor Productivity and Cost (LPC) program and average hourly earnings from the Current Employment Statistics (CES). While the LPC wage increased by about 70% over the past four decades and became markedly more volatile starting in the 1980s, the CES wage grew by only about 20% over the same period and experienced a large drop in volatility post-1980. We establish that the divergence between the two aggregate hourly wage series is due to the different evolution of average labor earnings. Average hours worked, by contrast, evolve very similarly. We then use labor earnings data from the Current Population Survey (CPS), the National Income and Product Accounts (NIPAs), and Piketty and Saez (2003) in an attempt to reconcile the divergence between LPC and CES labor earnings. Our analysis indicates that differences in earnings concept and population coverage can account for a large part of the divergence. Our analysis also shows that earnings differences between the CPS and the LPC can be attributed almost entirely to earnings of high-income individuals and supplements such as employer contributions to pension and health plans, which are included in the LPC but not in the CPS. This result is interesting in its own right given the widespread use of micro earnings data from the CPS in cross-sectional studies.

*The views expressed in this paper do not necessarily represent the views of the Federal Reserve System or the Federal Open Market Committee. We thank seminar participants at the BLS, the 2013 Canadian Economic Association meetings, the Bank of Canada, and the Fall 2013 Midwest Macro meetings for comments; Jay Stewart for Stata codes and helpful discussions; John Schmitt, Jean Roth, and Barry Hirsch for support with the CPS data; Shawn Sprague for providing us with unpublished LPC data; and Christine Garnier for excellent research assistance.

Contact information: julienchampagne@bankofcanada.ca and andre.kurmann@frb.gov.
1 Introduction

The evolution of average hourly wages is a key indicator for economic analysis. In the U.S., two of the most popular and most readily available measures of average hourly wages for the non-farm business sector are average hourly compensation from the Labor Productivity and Costs (LPC) program and average hourly earnings from the Current Employment Statistics (CES).\footnote{Both data sources come from the Bureau of Labor Statistics (BLS).} This paper documents that over the past four decades, the two measures diverged substantially both in terms of trend growth and business cycle volatility. In particular:

1. While the LPC wage grew consistently over time and stands today about 70% higher in real terms than in 1970, the CES wage decreased by almost 10% between the mid-1970s and the mid-1990s and increased by only 20% total over the past four decades.

2. While the volatility of the LPC wage increased by 35% to 45% since the early 1980s, the volatility of the CES wage dropped by about 50%. Since the volatility of output declined by 40% to 50% since the early 1980s (i.e. the Great Moderation), the relative volatility of average hourly wages increased two- to threefold according to the LPC and but remained roughly unchanged according to the CES.

The main objective of the paper is to reconcile this divergence in trend and business cycle volatility of the LPC wage and the CES wage. Since each series is constructed by dividing an average labor earnings measure with an average hours measure, we start by decomposing the total divergence into differences coming from the earnings side and the hours side. We find that the divergence between the LPC wage and the CES wage – both in terms of trend growth and business cycle volatility – is driven by the different evolution of average labor earnings. Average hours worked, by contrast, evolve very similarly.

Next, we use data from a third source, the Current Population Survey (CPS), to examine potential reasons for the different evolution of average earnings from the LPC and the CES. Following Abraham, Spletzer and Stewart (1998) and Lemieux (2006), the CPS earnings series is constructed by combining information from the annual May supplements for 1973-1978 with information from the monthly outgoing rotation groups (ORG) from 1979 onward. The resulting CPS May/ORG extracts – referred to as ‘CPS data’ from hereon – represent a relatively long, consistent earnings series that, on the one hand, is based on a very similar earnings concept as the one used in the CES and, on the other hand, allows us to cover the same worker population as in the LPC.\footnote{Alternatively, we could have used earnings data from the March CPS supplements. As discussed below, the CPS May/ORG extracts have some advantages over the March CPS supplements for our purpose; but we plan}
data therefore allows us to separately quantify how much of the difference between LPC earnings and CES earnings is due to (i) differences in earnings concept; and (ii) differences in population coverage. Furthermore, the comparison of CPS earnings with LPC and CES earnings is interesting in its own right because the micro-data of the CPS is publicly available and its earnings series has been widely used in cross-sectional studies on U.S. labor market characteristics.\(^3\)

Analysis of the CPS data yields several interesting results. First, we document that the evolution of average earnings from the CPS falls in between the evolution of LPC earnings and CES earnings, both in terms of trend growth and changes in business cycle volatility. Given the shared characteristics of the CPS earnings data with both the LPC and the CES data, this result suggests that the divergence between LPC and CES earnings has indeed multiple sources.

Second, using information from the National Income and Product Accounts (NIPAs) about the wage and salaries portion of earnings in the LPC as well as labor income share data for high-earning individuals computed by Piketty and Saez (2003), we show that differences in earnings concept account for almost all of the differences between CPS and LPC earnings. In particular, LPC earnings include supplements such as employer contributions to pension and health plans as well as earnings of high-income individuals (including gains from exercising certain stock options) whereas CPS earnings do not. Once these components of earnings are controlled for, LPC earnings and CPS earnings evolve very similarly. The CPS data therefore provides a reliable measure of wages and salaries that is representative of a very large part of the U.S. workforce.

Third, based on information from the publicly available micro-data of the CPS, we find that differences in worker population coverage can account for a substantial part of the divergence in trend growth and volatility between CPS earnings and CES earnings. However, the sources of the remaining differences remain an open question. We conjecture that compositional changes in the CES due to a major sample expansion occurring between the early 1980s and the late 1990s represent one of the most plausible candidates.

We are not the first to document differences in average hourly wages for the U.S. economy. In particular, Abraham, Spletzer and Stewart (1998) analyze in detail the differences in trend growth of the three labor earnings series studied here. While our paper builds heavily on their analysis, we make three distinct contributions. First, we extend the sample analyzed by Abraham, Spletzer and Stewart (1998) by almost 20 years to show that the three wage series continued to diverge in the 1990s and the 2000s and that several of the reasons for the divergence discussed in their paper continue to be important. Second, Abraham, Spletzer and Stewart (1998) only consider the

\(^3\)See Bound and Johnson (1992); Katz and Murphy (1992); or Lemieux (2006) among many others.
divergence in trend growth whereas we also document the divergence in business cycle volatilities and show that the sources for this divergence are, to some extent, the same as the ones for the divergence in trends. Third, our analysis reveals that the difference in earnings concept can account for almost all of the divergence between LPC and CPS earnings, which is an important result given the widespread use of the CPS earnings data.

Our paper also builds on earlier work by Gali and Van Rens (2010) and Champagne and Kurmann (2013) who document the divergence in business cycle volatility of the CES wage relative to LPC and CPS wages. Compared to these two studies, we provide a detailed analysis for the reasons behind this divergence.4

Finally, while our findings are suggestive of important structural changes in the U.S. labor market, it is important to stress that the analysis of these changes, however important, is not the focus of the paper. Instead, the primary contribution of the paper is to provide a detailed account of the divergence in different popular aggregate hourly wage series so as to obtain better guidance on which wage series to use for the analysis of different aspects of the U.S. labor market.

The remainder of the paper proceeds as follows. Section 2 describes the different data series and documents the divergence in aggregate hourly wages. Section 3 examines different potential explanations. Section 4 concludes.

2 Divergent average wages: data and facts

We begin with a description of the different data sources (additional details are available in the appendix). Then we document the divergent evolution of the average hourly wage series in terms of trend, business cycle volatility and business cycle correlation. For ease of exposition, we directly report for average hourly wages, average earnings per worker and average hours per worker from the LPC, the CES and from the CPS May/ORG data.

2.1 Data

The Labor Productivity and Costs (LPC) database reports a variety of labor market variables for the non-farm business sector available quarterly starting in 1948. Its weekly earnings measure has

---

4There is also an extensive survey by Abraham and Haltiwanger (1995) on the correlation of hourly wages with the business cycle. Their focus is mostly on the sensitivity of results to the measurement of nominal wages, nominal prices, and cyclical conditions. Our focus, by contrast, is on the volatility of real hourly wages. In the interest of completeness, we also show some results for the correlation of our hourly wage series with the business cycle and how this correlation changed over time. Similar results are reported in Gali and Van Rens (2010) and Champagne and Kurmann (2013).
two components: ‘wages and salaries’; and ‘supplements’. The ‘wages and salaries’ component is based on the Quarterly Census of Employment and Wages (QCEW) – also known as the BLS’ ES-202 program – a mandatory employer-based program for all employees covered by unemployment insurance (UI) that spans about 98% of U.S. establishments and jobs. Wage and salary disbursements are very comprehensive and include executive compensation, commissions, tips, bonuses and gains from exercising non-qualified stock options. The ‘supplements’ component is based on estimates by the Bureau of Economic Analysis and consists of vacation pay, employer contributions to pension and health plans, and employer contributions for government social insurance. Average weekly hours in the LPC are based primarily on hours from the Current Employment Statistics (CES) survey (see below), supplemented by hours from the Current Population Survey (CPS) for workers not covered by the CES. The hourly wage series computed from these average earnings and hours measures (‘LPC wage’ for short) is very comprehensive, both in terms of earnings concept and population coverage.

The second source of earnings and hours comes from the Current Employment Statistics (CES), which is a monthly survey starting in 1964 and is administered on a voluntary basis. The sample was significantly expanded during the 1980s and 1990s and currently covers about 140,000 private-sector firms representing almost 500,000 establishments. The historical CES earnings only covers production and nonsupervisory workers and comprises regular wage and salary disbursements as well as overtime, commissions and bonuses but only if paid each pay period.\(^5\) Tips, irregular bonuses, gains from exercising stock options, and supplements are excluded. The hourly wage series computed from the CES measures of earnings and hours (‘CES wage’ for short) is therefore more restrictive than the LPC wage, both in terms of earnings concept and population coverage.

The third source of earnings and hours comes from the Current Population Survey (CPS), which is a monthly household survey of about 60,000 individuals that can be weighted to make them representative of the U.S. Census. Information on earnings and hours are available from different extracts of the CPS. As in Abraham, Spletzer and Stewart (1998), Lemieux (2006), and Champagne and Kurmann (2013), we use information from the annual CPS May supplements from 1973 to 1978 together with information from the monthly outgoing rotation groups (ORG) from 1979 onwards to construct annual series of earnings and hours for private-sector workers.\(^6\) CPS

\(^5\) Starting in 2006, the CES started collecting data for all workers. We use this information below.

\(^6\) An interviewed individual appears in the CPS for two periods of four consecutive months, separated by eight months during which the individual is left out of the survey. Before 1979, earnings questions were asked only in March and May of each year (the March and May supplements). Thereafter, an earnings question is asked each month for individuals who are at the end of a four-month rotation (the ORG extracts). This information is collected each year by the NBER into a single merged ORG file, available on the NBER website. For the years between 1973
earnings are comprised of wages and salaries, including overtime, tips and commissions (OTC) and bonuses if earned and paid in each period.\textsuperscript{7} Irregular bonuses, gains from exercising stock options and supplemental benefits are excluded. Furthermore, earnings are topcoded. For hourly-paid workers, the CPS topcodes earnings at $99.99 per hour, a threshold rarely crossed. For salaried workers, the CPS topcodes weekly earnings at $999 until 1989; $1923 between 1989 and 1997; and $2884 from 1998 onwards. For certain years, this puts a substantial share of workers above the topcode. To reduce the risk of breaks in the earnings’ trend or spurious volatility induced by irregular topcode adjustments, we multiply topcoded earnings by a constant factor of 1.3 before averaging across individuals. While this constant-factor adjustment is standard in the labor literature, we also experiment with more sophisticated topcode adjustments (see the discussion in Section 3 and the appendix for details). Consistent with the LPC wage and the CES wage, the ‘CPS wage’ is computed as average weekly earnings divided by average weekly hours. Compared to the two other data sources, the CPS wage is similar to the CES wage in terms of its earnings concept but provides coverage representative of the non-farm business workforce of the US economy, as the LPC wage.\textsuperscript{8} We exploit this ‘in-between’ characteristic of the CPS wage relative to the LPC wage and the CES wage for much of our analysis.

All three earnings measures and therefore all three hourly wage series are deflated using the Personal Consumption Expenditure (PCE) index from the National Income and Products Accounts (NIPAs).\textsuperscript{9}

\textsuperscript{7}CPS respondents are first asked if they are salaried or paid by the hour. In the first case they report earnings on a weekly basis, in the latter case on an hourly basis. Salaried workers report all regular earnings including overtime, tips and commissions (OTC) if earned and paid in each period. Hourly-paid workers are asked to only report their regular hourly wage rate. In 1994, a new, separate question about OTC earnings was added for hourly-paid workers. As in Abraham, Spletzer and Stewart (1998), we do not use this additional OTC information for hourly workers for the baseline results. We use this information, however, in Section 3.

\textsuperscript{8}As opposed to CPS earnings, CES earnings do not include tips. According to Abraham, Stewart and Spletzer (1998), however, tips represented only a very small part of total earnings. It is therefore unlikely that tips play a major role in explaining the divergence of the CES wage.

\textsuperscript{9}As Abraham and Haltiwanger (1995) point out, there is controversy about the price index that should be used to deflate wages. While the choice of deflator may be important for the determination of the real wage level at any given point in time, our conclusions about the divergence in the three series is not affected by the use of alternative deflators.
2.2 Trends

Figure 1 plots the evolution of the three average real hourly wage series (all in natural logs).

Three observations stand out. First, in the early 1970s, the LPC wage is already about 35% higher than the CES wage and the CPS wage. Second, the LPC wage grows at a substantially higher rate over the sample, ending up 86% and 67% higher than the CES wage and the CPS wage, respectively, in 2009. Third, while the CPS wage grows consistently throughout the sample, the CES wage experiences a prolonged decline between the mid 1970s and the early 1990s.

2.3 Business cycle volatilities

To compute business cycle statistics, we take logarithms of the different hourly wage series and extract the business cycle component using the Hodrick-Prescott (H-P) filter. Then, we compute standard deviations of each series for the pre-1984 period and the post-1984 period. The break in

---

10The H-P filter constant is set to 1600 for quarterly data and 6.25 for annual data as recommended by Ravn and Uhlig (2002). Results are robust to alternative filtering methods.
1984 is motivated by the Great Moderation literature that estimates a significant change in output volatility around 1984 (e.g. McConnell and Perez-Quiros, 2000).

The first panel of Table 1 reports standard deviations for quarterly series of the LPC wage and the CES wage for the subsamples 1964:1-1983:4 and 1984:1-2011:4, with standard errors provided in brackets.\textsuperscript{11} The second panel of Table 1 reports the same standard deviations using annualized data for the samples 1973-1983 and 1984-2011 together with standard deviations for the CPS wage. Both tables also show the corresponding standard deviation of non-farm business chain-weighted GDP as a benchmark and report the ratio of the standard deviation of the different wage series to the standard deviation of GDP (denoted relative standard deviation).

<table>
<thead>
<tr>
<th>Standard Deviation</th>
<th>Relative Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pre-84</td>
</tr>
<tr>
<td>Quarterly data</td>
<td></td>
</tr>
<tr>
<td>Output</td>
<td>2.73</td>
</tr>
<tr>
<td></td>
<td>(0.31)</td>
</tr>
<tr>
<td>LPC wage</td>
<td>0.68</td>
</tr>
<tr>
<td></td>
<td>(0.08)</td>
</tr>
<tr>
<td>CES wage</td>
<td>1.13</td>
</tr>
<tr>
<td></td>
<td>(0.20)</td>
</tr>
<tr>
<td>Annual data</td>
<td></td>
</tr>
<tr>
<td>Output</td>
<td>2.91</td>
</tr>
<tr>
<td></td>
<td>(0.19)</td>
</tr>
<tr>
<td>LPC wage</td>
<td>0.64</td>
</tr>
<tr>
<td></td>
<td>(0.08)</td>
</tr>
<tr>
<td>CPS wage</td>
<td>0.64</td>
</tr>
<tr>
<td></td>
<td>(0.09)</td>
</tr>
<tr>
<td>CES wage</td>
<td>1.01</td>
</tr>
<tr>
<td></td>
<td>(0.15)</td>
</tr>
</tbody>
</table>


Table 1. Business cycle volatilities.

There is a clear divergence in business cycle volatility for the three hourly wage series. While the volatility of the LPC wage increases by 35% to 45% between the pre-84 period and the post-84 period, the volatility of the CPS wage increases only by about 20% and the volatility of the CES wage drops by about 50%. Since the volatility of output drops by 40% to 50% between the two periods (i.e. the Great Moderation), the relative volatility of hourly wages increases two- to threefold according to the LPC and the CPS but remains roughly unchanged according to the CES. Furthermore, the LPC wage and the CPS wage are equally volatile in the pre-84 period, whereas the volatility of the CES wage is substantially higher during that period.

\textsuperscript{11}Standard errors are computed via the delta method from GMM-based estimates. See the appendix for details.
2.4 Correlations

Table 2 reports the correlation coefficients of the three H-P filtered wage series, both with respect to non-farm business GDP and total non-farm business hours from the LPC. As before, the first panel shows results for quarterly data for the samples 1964:1-1983:4 and 1984:1-2011:4; and the second panel shows results for annual data for the samples 1973-1983 and 1984-2011.

<table>
<thead>
<tr>
<th></th>
<th>Correlations w/ GDPnfb</th>
<th>Correlations w/ Hours</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pre-84</td>
<td>Post-84</td>
</tr>
<tr>
<td><strong>Quarterly data</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LPC wage</td>
<td>0.35</td>
<td>-0.22</td>
</tr>
<tr>
<td></td>
<td>(0.18)</td>
<td>(0.10)</td>
</tr>
<tr>
<td>CES wage</td>
<td>0.60</td>
<td>-0.34</td>
</tr>
<tr>
<td></td>
<td>(0.12)</td>
<td>(0.18)</td>
</tr>
<tr>
<td><strong>Annual data</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LPC wage</td>
<td>0.39</td>
<td>-0.25</td>
</tr>
<tr>
<td></td>
<td>(0.30)</td>
<td>(0.13)</td>
</tr>
<tr>
<td>CES wage</td>
<td>0.66</td>
<td>-0.34</td>
</tr>
<tr>
<td></td>
<td>(0.22)</td>
<td>(0.23)</td>
</tr>
<tr>
<td>CPS wage</td>
<td>0.17</td>
<td>-0.46</td>
</tr>
<tr>
<td></td>
<td>(0.27)</td>
<td>(0.23)</td>
</tr>
</tbody>
</table>

Notes: Total sample extends from 1964:1 to 2011:4 for quarterly data; from 1973 to 2011 for annual data. Nonfarm business sector. Cyclical indicator are: (1) real chained-$ nonfarm business GDP (NIPA s) per capita, and (2) hours per capita from LPC. Standard errors appear in parentheses below estimates.

Table 2. Business cycle correlations.

While there are noteworthy differences in the pre-84 period (e.g. the correlation of the CES wage with both output and hours is markedly higher than for the other two wage series), all three wage series experience a sizable drop in correlation into negative territory for the post-84 period.

3 Potential explanations

Since each of the average hourly wage series is constructed by dividing an earnings measure with an hours worked measure, we start by decomposing the divergence in trend and business cycle volatility into differences coming from the earnings side and the hours side. Second, we consider three specific sources of divergence: (i) differences in earnings concepts; (ii) differences in population coverage; and (iii) measurement issues.

There are, of course, other differences between the wage series. However, as our analysis reveals, the three candidate sources we consider are likely to account for a large part of the divergence.
3.1 Earnings versus hours

Each of the average hourly wage series is constructed by dividing an earnings measure by an hours measure; i.e.

\[ w_{it} = W_{it}/H_{it}, \]

where \( w_{it} \) denotes the average hourly wage from data source \( i \) at time \( t \); \( W_{it} \) the corresponding average (weekly) earnings measure; and \( H_{it} \) the corresponding average (weekly) hours measure. Hence, the divergence in the different average hourly wage series must come from different evolutions in either earnings, hours, or both.

Figures 2 and 3 plot the evolution of log average weekly earnings and log average weekly hours used in the computation of the three hourly wage series.

![Figure 2. Log real weekly earnings (2005 dollars).](image)

Figure 2 shows that similarly to average hourly wages, there is already a level difference in 1973 between weekly earnings from the LPC and the two other weekly earnings measures. Thereafter, weekly earnings from both the LPC and the CPS grow consistently although the average growth rate of LPC weekly earnings is higher. By contrast, weekly earnings from the CES fall substantially...
between the mid-1970s and the early 1990s before recovering to their early 1970s level by 2010.

Figure 3 shows that while the LPC and the CES measure of weekly hours both decrease in similar fashion over time, the CPS measure of weekly hours fluctuates around an approximately constant level.

To quantify the importance of these differences for the divergence in the three hourly wage series, we use growth accounting techniques. First, we decompose the log difference of the average hourly wage from data source $i$ between 1973 and 2011 into the log differences of the corresponding weekly earnings and the weekly hours measures. Then, we subtract the same decomposition for the log difference of the average hourly wage from data source $j$ to obtain the percent contributions of differences in weekly earnings growth and weekly hours growth for the difference in average hourly wage growth; i.e.

$$
\Delta \log w_i - \Delta \log w_j = (\Delta \log W_i - \Delta \log W_j) - (\Delta \log H_i - \Delta \log H_j),
$$

(1)

where $\Delta \log w_i$ denotes the log difference in the average hourly wage from data source $i$ between 1973 and 2011 and so forth. Figure 4 reports the results.
Figure 4. Accounting for the divergence in average hourly wage growth.

Figure 4 shows that the divergence in average hourly wage growth between the LPC and CES is entirely due to the difference in earnings growth. By contrast, only about two thirds of the considerably smaller divergence in average hourly wage growth between the LPC and the CPS is due to smaller weekly earnings growth in the CPS. The remaining third of the divergence in average hourly wage growth is due to the fact that LPC weekly hours decreased consistently over time whereas CPS weekly hours remained approximately constant.

The results in Figure 4 confirm the previous findings by Abraham, Spletzer and Stewart (1998) for a substantially longer sample. The close association between LPC and CES hours should not come as a surprise since LPC hours are primarily constructed from CES hours. The divergence between CES (respectively LPC) hours and CPS hours is also relatively well-known and has been investigated in a recent paper by Frazis and Stewart (2010). We return to discussing their findings below.

The decomposition of average hourly wages into weekly earnings and weekly hours can also be used to analyze the divergence in business cycle volatility. Specifically, the variance of average
hourly wage growth from data source \( i \) can be expressed as

\[
\sigma_{wi}^2 = \sigma_{W_i}^2 + \sigma_{H_i}^2 - 2 \rho_{W_i,H_i}(\sigma_{W_i} \sigma_{H_i}),
\]

where \( \sigma_{wi}^2 \equiv \text{Var}(\Delta \log w_{it}) \); \( \sigma_{H_i}^2 \equiv \text{Var}(\Delta \log H_{it}) \); and \( \rho_{W_i,H_i} \equiv \text{Corr}(\Delta \log W_{it}, \Delta \log H_{it}) \). By subtracting this decomposition for some subsample \( a \) from the decomposition of some other subsample \( b \) (i.e. between the pre-84 period and the post-84 period), we obtain

\[
\sigma_{wi}^2(b) - \sigma_{wi}^2(a) = \left[ \sigma_{W_i}^2(b) - \sigma_{W_i}^2(a) \right] + \left[ \sigma_{H_i}^2(b) - \sigma_{H_i}^2(a) \right] - 2 \left[ \rho_{W_i,H_i}(b)\sigma_{W_i}(b)\sigma_{H_i}(b) - \rho_{W_i,H_i}(a)\sigma_{W_i}(a)\sigma_{H_i}(a) \right] .
\]

By manipulating this expression further to decompose the multiplicative parts, we end up with

\[
\sigma_{wi}^2(b) - \sigma_{wi}^2(a) = \left[ \sigma_{W_i}^2(b) - \sigma_{W_i}^2(a) \right] + \left[ \sigma_{H_i}^2(b) - \sigma_{H_i}^2(a) \right] - \left\{ \frac{\rho_{W_i,H_i}\sigma_{H_i}(b)}{2} \right\} \left[ \sigma_{W_i}(b) - \sigma_{W_i}(a) \right] + \left\{ \frac{\rho_{W_i,H_i}\sigma_{W_i}(b)}{2} \right\} \left[ \sigma_{H_i}(b) - \sigma_{H_i}(a) \right] .
\]

Hence, the change in variance of average hourly wage growth is accounted for by changes in either the volatility of earnings growth; the volatility of hours growth; or the correlation between earnings and hours growth. This allows us to quantify the sources of the divergence in business cycle volatility between the different average hourly wage series.

As a preliminary to this volatility accounting exercise, Table 3 shows the post-84 to pre-84 changes in volatilities and correlations of the three weekly earnings and weekly hours measures,
together with the corresponding changes in the hourly wage volatilities reported in Table 1.

<table>
<thead>
<tr>
<th></th>
<th>LPC</th>
<th>CES</th>
<th>CPS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1973-1984</td>
<td>1984-2011</td>
<td>post 84 - pre 84</td>
</tr>
<tr>
<td>Std(hourly wage)</td>
<td>0.64</td>
<td>0.85</td>
<td>0.21</td>
</tr>
<tr>
<td>Std(weekly earnings)</td>
<td>0.90</td>
<td>0.80</td>
<td>-0.10</td>
</tr>
<tr>
<td>Std(weekly hours)</td>
<td>0.41</td>
<td>0.45</td>
<td>0.04</td>
</tr>
<tr>
<td>corr(weekly earnings, weekly hours)</td>
<td>0.78</td>
<td>0.17</td>
<td>-0.61</td>
</tr>
<tr>
<td></td>
<td>1973-1984</td>
<td>1984-2011</td>
<td>post 84 - pre 84</td>
</tr>
<tr>
<td>Std(hourly wage)</td>
<td>1.01</td>
<td>0.50</td>
<td>-0.51</td>
</tr>
<tr>
<td>Std(weekly earnings)</td>
<td>1.29</td>
<td>0.49</td>
<td>-0.79</td>
</tr>
<tr>
<td>Std(weekly hours)</td>
<td>0.38</td>
<td>0.36</td>
<td>-0.02</td>
</tr>
<tr>
<td>corr(weekly earnings, weekly hours)</td>
<td>0.79</td>
<td>0.35</td>
<td>-0.44</td>
</tr>
<tr>
<td></td>
<td>1973-1984</td>
<td>1984-2011</td>
<td>post 84 - pre 84</td>
</tr>
<tr>
<td>Std(hourly wage)</td>
<td>0.64</td>
<td>0.76</td>
<td>0.12</td>
</tr>
<tr>
<td>Std(weekly earnings)</td>
<td>0.76</td>
<td>0.69</td>
<td>-0.07</td>
</tr>
<tr>
<td>Std(weekly hours)</td>
<td>0.41</td>
<td>0.29</td>
<td>-0.13</td>
</tr>
<tr>
<td>corr(weekly earnings, weekly hours)</td>
<td>0.56</td>
<td>-0.04</td>
<td>-0.59</td>
</tr>
</tbody>
</table>

*Note: Annual data, H-P filtered. Standard deviations are multiplied by 100.

Table 3: Changes in standard deviations and correlations of weekly earnings and weekly hours.

Three observations stand out. First, the volatility of weekly earnings from both the LPC and the CPS experience a small decline although, interestingly, the volatility of weekly earnings from the CPS declines by a slightly smaller amount. In comparison, the volatility of CES weekly earnings drops about 8 times as much. Second, the volatility of weekly hours remains approximately constant in the LPC and the CES but decreases slightly in the CPS. Third, the correlation of weekly earnings with weekly hours drops in all three data sets but the drop is larger for the LPC and the CPS. These observations imply, maybe somewhat surprisingly, that the increase in volatility of the average hourly wage in the LPC and the CPS is entirely due to the drop in correlation between weekly earnings and weekly hours (since a drop in correlation affects hourly wage volatility positively). In turn, the large drop in volatility of the CES wage is driven by the large drop in volatility of weekly earnings.

Figure 5 displays the results of the volatility accounting exercise based on the numbers in Table

---

12 The small decline in volatility of weekly earnings from the LPC and the CPS is consistent with recent findings from micro-data that for most individuals, the volatility of labor income has remained approximately constant (e.g. Dynan et al., 2008; Jensen and Shore, 2008).

13 Since LPC hours and CES hours are very highly correlated in both subsamples (0.99 and 0.98, respectively), the larger drop in correlation between earnings and hours in the LPC is entirely due to the different cyclical properties of earnings.
Figure 5. Accounting for the divergence in business cycle volatility of average hourly wages.

The figure provides quantitative confirmation for the above observation that the large drop in correlation between earnings and hours is behind the increase in hourly wage volatility in the LPC and the CPS. The increase in the volatility of the LPC wage is larger than the increase in the volatility of the CPS wage because in the LPC, the drop in correlation between earnings and hours is attributed a larger weight in the above variance decomposition (due to larger average earnings and hours volatilities in the LPC). In turn, the large drop in the volatility of the CES wage is primarily due to the large fall in earnings volatility and the somewhat smaller drop in correlation between earnings and hours.

We take away two main lessons from the trend and volatility decomposition exercises:

1. The divergence in both trend and volatility between the LPC wage and the CES wage is entirely driven by the divergence in trend and volatility of weekly earnings. Weekly hours from the LPC and CES behave, by construction, very similarly.

2. The divergence in trend growth of the LPC wage relative to the CPS wage is due to the smaller growth of weekly earnings in the CPS and, to a lesser extent, the difference in the

\[ \Delta \sigma^2(w) = \Delta \sigma^2(W) - \Delta \sigma^2(H) \]

\[ \Delta \text{corr}(w,H) = \Delta \text{corr}(W,H) \]

Note: The figure decomposes the change in the variance of the hourly wage (blue) into changes in the variance of earnings (red), the variance of hours (green), and the correlation between earnings and hours (purple) between 1973-1983 and 1984-2011 for the LPC, CES, and CPS datasets. The accompanying table reports the differences in changes between the difference components.

<table>
<thead>
<tr>
<th>Change in Variance</th>
<th>LPC-CES</th>
<th>%</th>
<th>LPC-CPS</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>[ \Delta \sigma^2(w_{LPC}) - \Delta \sigma^2(w) ]</td>
<td>1.090</td>
<td>100.00%</td>
<td>0.1453</td>
<td>100.00%</td>
</tr>
<tr>
<td>[ \Delta \sigma^2(W_{LPC}) - \Delta \sigma^2(W) ]</td>
<td>0.941</td>
<td>86.34%</td>
<td>-0.0411</td>
<td>-28.31%</td>
</tr>
<tr>
<td>[ \Delta \sigma^2(H_{LPC}) - \Delta \sigma^2(H) ]</td>
<td>-0.002</td>
<td>-0.21%</td>
<td>0.0426</td>
<td>29.34%</td>
</tr>
<tr>
<td>[ \Delta \text{corr}(w_{LPC},H_{LPC}) - \Delta \text{corr}(w,H) ]</td>
<td>0.151</td>
<td>13.87%</td>
<td>0.1438</td>
<td>98.97%</td>
</tr>
</tbody>
</table>

14The volatility accounting formula is derived for first-differenced data whereas the results in Table 3 pertain to H-P filtered data. This introduces an approximation error that is, however, only of minor quantitative importance.
evolution of weekly hours. The larger increase in volatility of the LPC wage relative to the CPS wage is mainly due to a smaller contribution of the drop in correlation between earnings and hours in the CPS.

Based on these results, differences in earnings behavior become the main focus of our attempt to reconcile the divergence in average hourly wages.

3.2 Differences in earnings concepts

As described in Section 2, LPC earnings are based on a very broad concept that includes executive compensation, tips, bonuses and gains from executing non-qualified stock options; as well as supplements such as vacation pay and employer contributions to pension and health plans. By contrast, CPS and CES earnings only include compensation that is earned and paid each period; and completely exclude supplements.

Given the similarity in earnings concepts between the CPS and the CES, we focus on the comparison between LPC earnings and CPS earnings. First, we try to compare the wages and salaries component of the LPC to the one from the CPS. Second, we use information in the CPS on overtime, tips and commissions (OTC) to assess the importance of a particular type ‘regular’ bonus payments for hourly-paid workers. Third, we use wage share information on top income earners from Piketty and Saez (2003, updated to 2010) to quantify the role played by earnings of high-income individuals in the LPC.

3.2.1 Wages and salaries

To compare the wages and salaries component of LPC earnings to the one from the CPS, we need to strip out supplements components from LPC earnings. Unfortunately, this cannot be done directly in the LPC dataset because it does not contain separate information on these two components of compensation. Separate information on the two components is, however, provided by the income tables in the National Income and Product Accounts (NIPAs), from which the LPC program computes its earnings series. At the same time, the publicly available NIPA data do not contain all information to reconstruct the non-farm business coverage employed by the LPC.\textsuperscript{15} We therefore consider a ‘private non-agriculture’ coverage of earnings that is straightforward to compute from the NIPAs.

\textsuperscript{15}As explained in more detail in the appendix, the LPC non-farm business data excludes farms, households and non-profits and general government; but includes agricultural services, forestry and fishing, government services and imputed data for self-employed. Several of these added and subtracted components are unavailable publicly over the entire sample period.
Figure 6 plots the resulting NIPA series of average weekly earnings both for total compensation including supplements (labeled ‘NIPA total compensation’) and the wages and salaries component (labeled ‘NIPA wages&salaries’). The figure also contains, for comparison, the LPC (non-farm business) series of weekly earnings from above; and two weekly private non-agricultural earnings series computed from the CPS. The series labeled ‘CPS private non-agricultural’ is the equivalent of the CPS wage from above. The series labeled ‘CPS non-agricultural with OTC’ includes OTC payments and is discussed further below.

As shown by the top two lines, there is some difference in total compensation (i.e. earnings including supplements) between non-farm business sector covered by the LPC and the private non-agricultural sector as computed from the NIPAs. However, this difference is small compared to the substantial and widening gap between NIPA private non-agricultural earnings based on total compensation and earnings based on the wages and salaries component. This gap, which started to appear in the 1960s illustrates that the inclusion of supplements are an important reason for the level difference and the divergence in trends between LPC earnings and CPS earnings (and therefore CES earnings).

Table 4 displays the business cycle volatility of the same earnings series for the pre-84 and
post-84 subsamples.

<table>
<thead>
<tr>
<th></th>
<th>Standard Deviation</th>
<th>Relative Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pre-84</td>
<td>Post-84</td>
</tr>
<tr>
<td>HP-Filter</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Output (nfb)</td>
<td>2.91</td>
<td>1.42</td>
</tr>
<tr>
<td>(0.19)</td>
<td>(0.22)</td>
<td></td>
</tr>
<tr>
<td>LPC total compensation (nfb)</td>
<td>0.90</td>
<td>0.81</td>
</tr>
<tr>
<td>(0.14)</td>
<td>(0.12)</td>
<td></td>
</tr>
<tr>
<td>NIPA total compensation (private non-agri)</td>
<td>0.89</td>
<td>0.71</td>
</tr>
<tr>
<td>(0.13)</td>
<td>(0.10)</td>
<td></td>
</tr>
<tr>
<td>NIPA wages &amp; salaries (private non-agri)</td>
<td>0.76</td>
<td>0.84</td>
</tr>
<tr>
<td>(0.11)</td>
<td>(0.10)</td>
<td></td>
</tr>
<tr>
<td>CPS (private non-agri)</td>
<td>0.76</td>
<td>0.69</td>
</tr>
<tr>
<td>(0.13)</td>
<td>(0.12)</td>
<td></td>
</tr>
<tr>
<td>CPS with OTC (private non-agri)</td>
<td>0.84</td>
<td>0.79</td>
</tr>
<tr>
<td>(0.14)</td>
<td>(0.12)</td>
<td></td>
</tr>
</tbody>
</table>


Table 4. Comparison of aggregate weekly earnings measures.

The decline in volatility of total compensation in the post-84 period is slightly larger for the NIPAs private non-agricultural aggregate than for the LPC non-farm business aggregate. This difference is due to small differences in the evolution of volatility for segments of the population that are included in the LPC (such as self-employed) but excluded from the NIPAs aggregate and vice versa.

More interesting is the increase in volatility of the ‘wages and salaries’ component computed from the NIPAs. This result obtains because supplements, which are excluded in this ‘wages and salaries’ series but included in total compensation, fell in volatility in line with the business cycle. In other words, once supplements are stripped out from NIPA earnings (or, equivalently, LPC) earnings, there is thus also a divergence in earnings volatility between NIPA earnings and CPS earnings (which does not contain supplements).

3.2.2 Overtime, tips and commissions

As described in Section 2, salaried workers in the CPS report all earnings including OTC if earned and paid each period. Hourly-paid workers, by contrast, were historically asked to only report their regular hourly wage rate, thus excluding OTC. Since hourly-paid workers represent almost 60% of the workforce (and around 50% of the total wage bill), the CPS earnings measure is likely to underestimate the true level of regular earnings. Furthermore, since overtime work increased substantially during the 1980s and 1990s (see Kuhn and Lozano, 2008) and is likely to be more volatile than regular pay, the CPS earnings measure is also likely to underestimate the true trend and volatility of earnings over time.
To obtain a measure of earnings including OTC for hourly workers in the CPS, we proceed as follows. Prior to 1994, we use the greater of weekly earnings and wage rate * hours to obtain a measure of earnings that includes at least some of the OTC received by hourly-paid workers (see the appendix for details and a discussion on why this measure is incomplete). Starting in 1994, the CPS introduced a separate question about OTC earnings for hourly-paid workers, which we add to the usual wage rate * hours series.

The last line in Figure 6 and Table 4 report the results, labeled 'CPS with OTC (private non-agricultural)’. Adding OTC for hourly-paid workers indeed leads to a systematic level increase with a discrete jump in 1994 when the OTC question was introduced. Including OTC thus covers at least part of the gap between CPS earnings and NIPA / LPC earnings. As Table 4 indicates, adding OTC also leads to a smaller drop in the volatility of CPS earnings for the post-1984 subsample. Further analysis reveals, however, that this is mostly the result of a discontinuity resulting from the introduction of additional OTC earnings question in 1994. Once we control for this discontinuity, the smaller drop in volatility disappears.

3.2.3 Labor earnings of high-income individuals

While the CPS contains all information to construct a representative average weekly earnings series of private non-agricultural workforce, respectively the entire U.S. workforce, publicly available earnings data is topcoded for high-income individuals (see the description in Section 2), thus downweighing their contribution to average weekly earnings. But even if earnings for these high-income individuals was not topcoded, the CPS data would still miss a substantial portion of high-income individuals’ compensation due to the restrictive earnings concept in the CPS. In contrast, the earnings concept in the LPC (and NIPAs) is much more comprehensive because the QCEW, its basis for earnings, includes irregular cash bonuses and gains from exercising non-qualified stock options. This may affect the evolution of earnings in non-trivial ways.\footnote{Irregular bonuses, but not gains from exercises stock options, are included in the CPS March supplements. We plan to quantify the importance of these bonuses in a subsequent version of the paper.}

This difference between CPS and LPC / NIPAs earnings are particularly interesting to investigate because recent evidence shows that the share of total labor income by high-earning individuals has increased importantly over the past decades. Most prominently, based on tabulations from the Internal Revenue Service (IRS), Piketty and Saez (2003, updated to 2010) document that the top 1% individuals of the income distribution saw their share of total economy-wide income increase from a stable 8% between the 1950s to the mid-1990s to 23.5% in 2007. This remarkable growth is due mostly to the growing inequality in labor income and implies earnings of high-income individ-

\footnote{Irregular bonuses, but not gains from exercises stock options, are included in the CPS March supplements. We plan to quantify the importance of these bonuses in a subsequent version of the paper.}
uals play an increasingly important role for the trend of average weekly earnings as recorded in the LPC / NIPAs. Likewise, if earnings of high-income individuals have become more volatile, then this could explain why the volatility of the LPC hourly wage increased so much in the post-84 period.\textsuperscript{17}

To assess the role played by earnings of high-income individuals, we use information on top wage income shares from Piketty and Saez to calculate a separate series of average weekly earnings for the top 5\% earners and the remaining 95\% in each year. We then compare the two series to average weekly earnings for the corresponding top 5\% earners in the CPS (of which a fraction have topcoded earnings) and the remaining 95\%.\textsuperscript{18} Since the earnings concept of the IRS data used by Piketty and Saez is very similar to the one employed in the QCEW, the comparison allows us to consider the role played by high-income individuals and their irregular earnings that are not taken into account in the CPS. At the same time, since IRS data does not allow a distinction into different sectors, the Piketty-Saez data is only available for ‘all economy’. We therefore recompute the CPS

\textsuperscript{17}This is particularly relevant for stock options because they are likely to be exercised in upturns when their value is higher than their fair-market value at the time they were granted (i.e. the time when they should have been recorded as compensation). See Mehran and Tracy (2001) who argue that the growth of stock options in the 1990s and their inclusion in compensation at the time of exercise has biased the evolution of compensation upwards. The authors also conjecture that increased use of stock options may render compensation more variable.

\textsuperscript{18}We use a 5\%-95\% split simply because in the CPS data, the fraction of individuals with topcoded earnings never exceeds 5\%. 

earnings series for an ‘all economy’ equivalent. Figure 7 shows the results.

Unsurprisingly, there is a large and widening difference in average weekly earnings between the top 5% earners and the remaining 95%. What is striking, however, is that average weekly earnings for the 95% computed from Piketty and Saez (2003, 2011), labeled ‘P-S 0-95’, evolves almost identically to the one computed from the CPS, labeled ‘CPS 0-95’. In contrast, for the top 5% earners, weekly earnings computed from Piketty-Saez grow at a substantially higher pace than weekly earnings for the 5% top earners in the CPS.

To further quantify the importance of this difference in weekly earnings for high-income individuals, we take Piketty and Saez’ (2003, 2011) weekly earnings information for the top-income groups (i.e. top 0.01%, 0.1%-0.01%, 0.5%-0.1%,...to 1%-5%) and extrapolate new values for topcoded CPS earnings for each year from 1973 to 2009 (the last year for which the Piketty-Saez data is currently available). Based on this extrapolation, we then compute a new ‘topcode corrected’ weekly earnings measure for the CPS and analyze to what extent this measure tracks the evolution of weekly earnings from the NIPA wages and salaries portion. The specifics of the procedure are described in
We notice that weekly earnings from the topcode corrected measure of the CPS closes most of the gap with the NIPA wages and salaries series. If, in addition, we adjust for the small difference in weekly earnings for the remaining 95% earners between the CPS and the Piketty-Saez measure, labeled ’CPS topcode & 0-95 corrected’, we basically close the gap.

Table 5 reports on the business cycle volatility of weekly earnings of the different series in Figures 7 and 8. For comparison, the table also shows the NIPA wages and salaries series for ‘all economy’.
The first panel confirms that weekly earnings for the top 5% earners are substantially more volatile than for the remaining 95% and have, according to the Piketty-Saez data, greatly increased in volatility for the post-84 period. The second panel shows that correcting the CPS earnings series with the Piketty-Saez data reduces the drop in volatility and, in one instance, slightly increases it.\footnote{Also note from this panel that the volatility of the 'all economy' aggregate of NIPA wages and salaries falls substantially while it increased for the 'private non-agricultural' aggregate. By contrast, for the CPS, there is no corresponding fall in earnings volatility of the 'all economy' aggregate relative to the 'private non-agricultural' aggregate. We plan to analyze this issue further in subsequent versions of the paper.}

We conclude from this investigation that differences in earnings concept between the QCEW and the CPS explain the vast majority of the divergence in weekly earnings between the two data sets and therefore also account for at least part of the divergence in trend and volatility between the LPC wage and the CES wage.

### 3.3 Differences in population coverage

As described in Section 2, the LPC and the CES cover different segments of the U.S. workforce. Specifically weekly earnings in the LPC is based on the QCEW and includes labor income of the near totality of workers in the non-farm business sector. By contrast, the CES historical sample covers only earnings of production and non-supervisory workers. In 2006, the CES started collecting earnings and hours information for all workers in the sampled establishments. Comparing earnings for this 'all worker' sample to earnings from the sample of production and non-supervisory workers should therefore provide information on the role played by the difference in population coverage.
Figure 9 shows the result of this comparison, together with average earnings from the non-farm business population in the CPS (which, as discussed above, is based on a very similar earnings concept as the CES).

The figure reveals that average weekly earnings for the 'all worker' sample of the CES are about 20% higher than for the production and non-supervisory sample and almost match the level of average earnings in the CPS. The difference in population coverage therefore explains most of the gap between CES earnings and CPS earnings at the end of the sample (respectively LPC earnings once the difference in earnings concept is taken into account). What needs to be explored is whether the difference in population coverage also explains the divergence in trends and volatility over time between CES earnings and CPS earnings.

To assess this possibility, we employ the strategy used in Abraham, Spletzer and Stewart (1998) and create a weekly earnings measure from the CPS data intended to replicate the population coverage in the CES. In a first instance, we use the official definition of production and non-supervisory workers from the BLS. As can be seen from Figure 10, the resulting series, labeled 'CES replication 1', fails to generate the trend and business cycle dynamics of weekly earnings in the CES. As opposed to CES earnings, from CES replication 1 are already substantially below CPS earnings.

---

The sample for this exercise stops in 2002 because occupations definitions in the CPS changed in 2003, making
earnings in the early 1970s and grow at approximately the same pace as CPS earnings thereafter. This result confirms, for a substantially longer sample, the findings reported by Abraham, Spletzer and Stewart (1998).

Figure 10. Replicating log average real earnings from the CES with CPS data.

Plewes (1982) and Abraham, Spletzer and Stewart (1998) argue, however, that establishments in the CES often mistakenly interpret production and non-supervisory workers as employees paid by the hour and other employees that are non-exempt under the Fair Labor Standards Act. Hence, by restricting the CPS sample to the definition of production and non-supervisory workers in the BLS, we may not necessarily capture what establishments in the CES sample report. In a second instance, we use an alternative definition of production and non-supervisory workers constructed by Abraham, Spletzer and Stewart (1998) based on their assessment of what CES establishments report. As Figure 10 shows, the resulting series, labeled 'CES replication 2', implies an even larger gap in earnings relative to the CES in the early 1970s. Thereafter, however, earnings from CES replication track earnings from the CES more closely, including a significant decline in earnings between the late 1970s and the early 1990s. Again, this result broadly confirms the findings reported in Abraham, Spletzer and Stewart (1998) for a longer sample.

We thank Jay Stewart for kindly providing us with the Stata codes used in Abraham et al. (1998). Our second CES replication therefore uses exactly the same definitions they use in their paper.
It is also interesting to assess the extent to which the two CES replications are capable of generating the large fall in earnings volatility in the CES from the pre-84 to the post-84 subsample. Table 6 reports the results for both 1st-differenced and H-P filtered data.

<table>
<thead>
<tr>
<th>Changes in average earnings volatility</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard Deviation</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>1st-difference</td>
</tr>
<tr>
<td>CPS</td>
</tr>
<tr>
<td>CES replication 1</td>
</tr>
<tr>
<td>CES replication 2</td>
</tr>
<tr>
<td>CES</td>
</tr>
<tr>
<td>HP-filter</td>
</tr>
<tr>
<td>CPS</td>
</tr>
<tr>
<td>CES replication 1</td>
</tr>
<tr>
<td>CES replication 2</td>
</tr>
<tr>
<td>CES</td>
</tr>
</tbody>
</table>


Table 6. Replicating average real earnings volatility from the CES with CPS data.

Earnings volatility from CES replication 1 drops somewhat more than earnings volatility from the CPS but not nearly enough to replicate the change in earnings volatility from the CES.\(^{22}\) In contrast, CES replication 2 comes close to matching the higher volatility of actual CES weekly earnings in the pre-84 period and, depending on the filtering method used, accounts for 79\% to 91\% of the decline in volatility of the CES weekly earnings in the post-84 period.\(^{23}\) This further suggests that differences in population coverage can explain at least part of the divergence in trend and business cycle volatility of earnings in the CES.

Naturally, the same issues of representativeness may explain the different evolution of weekly hours from the CES and the CPS. In a recent paper, Frazis and Stewart (2010) investigate this possibility. They find that both CES replication 1 and 2 with the CPS sample decreases average hours by 1.3 to 1.7 hours, which basically closes the initial gap between CES and CPS hours. However, neither of the replications can account for the downward trend in CES hours. Nevertheless, it is interesting to assess the extent to which applying the two replication to both weekly earnings

\(^{22}\)Notice that the CPS earnings volatility numbers for the post-1984 period are slightly different than in Table 3 because the sample here stops in 2002 instead of 2011.

\(^{23}\)Because first-differencing cuts out a part of the fluctuations that are typically associated with the business cycle, we think that 79\% is a more accurate and conservative estimate of how the CES replication 2 accounts for the decline in CES weekly earnings after 1984.
and weekly hours allows us to explain the divergence in trend and business cycle between the CES hourly wage and the CPS hourly wage (and therefore, the LPC hourly wage). Figure 11 and Table 7 display the results.

Figure 11. Replicating log average real hourly wages from the CES with CPS data.

In terms of trends, CES replication 2 comes surprisingly close to the CES hourly wage although
there remains a large level difference between the two.

<table>
<thead>
<tr>
<th></th>
<th>Standard Deviation</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pre-84</td>
<td>Post-84</td>
<td>Relative</td>
</tr>
<tr>
<td><strong>1st-difference</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CPS</td>
<td>1.21</td>
<td>1.38</td>
<td>1.14</td>
</tr>
<tr>
<td>CES replication 1</td>
<td>1.18</td>
<td>1.34</td>
<td>1.14</td>
</tr>
<tr>
<td>CES replication 2</td>
<td>1.70</td>
<td>1.27</td>
<td>0.75</td>
</tr>
<tr>
<td>CES</td>
<td>1.49</td>
<td>1.17</td>
<td>0.79</td>
</tr>
<tr>
<td><strong>HP-filter</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CPS</td>
<td>0.64</td>
<td>0.64</td>
<td>1.00</td>
</tr>
<tr>
<td>CES replication 1</td>
<td>0.69</td>
<td>0.59</td>
<td>0.87</td>
</tr>
<tr>
<td>CES replication 2</td>
<td>0.99</td>
<td>0.63</td>
<td>0.64</td>
</tr>
<tr>
<td>CES</td>
<td>1.01</td>
<td>0.35</td>
<td>0.34</td>
</tr>
</tbody>
</table>


Table 7. Replicating average real hourly wage volatility from the CES with CPS data.

In terms of business cycle volatility, results are broadly similar to the replication exercise for average earnings in Table 6. CES replication 1 generates hourly wage volatilities that remain close to the CPS hourly wage volatility. For first-differenced data, CES replication 2 accounts for the main share of the fall in volatility of CES hourly wages. For H-P filtered data, CES replication 2 accounts for about 55% of the fall in volatility of CES hourly wages relative to the fall in volatility of CPS hourly wages, which is somewhat less than was the case for weekly earnings but still substantial. Also, for H-P filtered data, CES replication 2 comes very close to generating the higher pre-84 hourly wage volatility in the CES.

Overall, the CES replication exercises with CPS data suggest that the segment of workers for which establishments in the CES sample have traditionally reported earnings is not representative of the non-farm business sector workforce and that this lack of representativeness plays a major role in the divergence of the CES wage from the other wage series. In addition, the difference between CES replication 1 and CES replication 2 suggest that the historical earnings and hours series from the CES do not even cover the subset of the workforce they are supposed to represent. This makes the use of historical CES earnings and hours series problematic.
3.4 Measurement issues

This last section describes a particular set of issues that affect the CES earnings and hours measures but not the LPC and CPS measures, and how these issues may explain the remaining part of the divergence in the CES earnings and hours measures.24

First, the CES sample underwent a substantial expansion from about 160,000 to 400,000 establishments between 1980 and 2006. This expansion is likely to have led to spurious changes in the CES average earnings measure. Specifically, based on UI microfiles, Plewes (1982) documents the following characteristics of the CES sample for the early 1980s:

- Whereas 72% of all private-sector employment came from services-providing industries, this proportion in the CES sample was only 30.4%.

- The service establishments that the CES captured were on average much larger and older than the ones in the UI records.25 Since the proportion of small and young establishments in service-providing industries was higher than in goods-producing industries, this implied that the CES sample contained few small and young establishments.

The sample expansion of the CES that started in the early 1980s led to a better representation of the service-providing industries and, more generally, of small and young establishments. Because small and young establishments in service industries hire on average less skilled workers for whom hourly wages are lower and have become less volatile (see Champagne and Kurmann, 2013), the observations by Plewes (1982) imply that the resulting shift in sample composition may have led to spurious changes in the CES average earnings measures – both in terms of trend and business cycle volatility.26 It would be interesting to assess this conjecture using QCEW and CES earnings data that distinguishes establishments by industry and age. Such an exercise hinges on the availability of a micro-data that extends sufficiently far back in time.27

---

24 This is not to say that the CPS does not suffer from measurement issues. For example, as claimed by some researchers, hours in the CPS may be overreported (although Frazis and Stewart (2010) argue that once all the necessary adjustments for sample representation and reporting period are made, overreporting in the CPS is not significant). The point here is, however, that we are not aware of systematic changes in measurements issues for the CPS that could lead to a bias in the evolution of the CPS wage similar to, we argue, is the case for the CES earnings and hours.

25 An important reason for this underrepresentation is, according to Plewes (1982), the inability / unwillingness of small establishments to provide information for the CES survey (BLS 790 questionnaire), which is voluntary.

26 Note that employment numbers in the CES are benchmarked once a year to UI records (the source of the QCEW). Earnings and hours are, however, not benchmarked and therefore do not undergo a regular bias correction.

27 Furthermore, since average earnings in the CES are a weighted sum of earnings across industries, the sample
4 Conclusion

The evolution of average hourly wages is a key indicator for economic analysis. In the U.S., the Labor Productivity and Costs (LPC) program and the Current Employment Statistics (CES) provide the two most popular and most readily available measures of average hourly wages for the non-farm business sector. In this paper, we document that over the past four decades, the two measures diverged substantially both in terms of trend growth and business cycle volatility. Particularly, while the LPC wage is today about 70% higher in real terms than in 1970, the CES wage decreased by almost 10% between the mid-1970s and the mid-1990s and increased by only 20% total over the past four decades. Furthermore, while the volatility of the LPC wage increased by 35% to 45% since the early 1980s, the volatility of the CES wage dropped by about 50%.

We try to reconcile this divergence in trend and business cycle volatility of the LPC and the CES average hourly wages by first decomposing the total divergence into differences coming from the earnings side and the hours side. We find that the divergence between the LPC wage and the CES wage – both in terms of trend growth and business cycle volatility – is driven by the different evolution of average labor earnings. Average hours worked, by contrast, evolve very similarly. As a result, we turn the focus on earnings to explain the divergence between the two hourly wage series.

We use data from a third source, the Current Population Survey (CPS), to examine potential reasons for the different evolution of average earnings from the LPC and the CES. Earnings data in the CPS is based on a very similar earnings concept as the one used in the CES and at the same time allows us to cover the same worker population as in the LPC. We show that the evolution of average earnings from the CPS falls in between the evolution of LPC earnings and CES earnings, both in terms of trend growth and changes in business cycle volatility. Moreover, using additional information about the wage and salaries portion of earnings in the LPC as well as labor income share data for high-earning individuals computed by Piketty and Saez (2003), we show that differences expansion is likely have decreased the measurement error in some of the industries, thus reducing the volatility of average earnings. To see this, think of measured earnings across establishments in some industry $i$, $\widehat{W}_{i,t}$, as the sum of "true" earnings, $W_{i,t}$, plus an uncorrelated error term $\varepsilon_{i,t}$

$$\widehat{W}_{i,t} = W_{i,t} + \varepsilon_{i,t}.$$  

The variance of measured earnings is therefore

$$Var(\widehat{W}_{i,t}) = Var(W_{i,t}) + Var(\varepsilon_{i,t}).$$

Under the hypothesis that the errors are i.i.d. across establishments of industry $i$, an increase in the number of establishments per industry leads to a decrease in the variance of errors and therefore a decrease in the variance of measured CES earnings.
in earnings concept account for almost all of the differences between CPS and LPC earnings. The CPS data therefore provides a representative measure of wages and salaries for large part of the U.S. workforce. Finally, we use occupational and industry information in the CPS to show that differences in worker population coverage can account for a substantial part of the divergence in trend growth and volatility between CPS earnings and CES earnings. However, the sources of the remaining differences remain an open question. We conjecture that compositional changes in the CES due to a major sample expansion occurring between the early 1980s and the late 1990s represent one of the most plausible candidates.

Even though our findings are suggestive of important structural changes in the U.S. labor market, it is not the focus of the paper. Instead, the primary contribution of the paper is to provide a detailed account of the divergence in different popular aggregate hourly wage series so as to obtain better guidance on which wage series to use when analysing different aspects of the U.S. labor market.

References


A Data Description

Here we describe in details the different variables used throughout the paper. We also provide data sources and series’ IDs.

A.1 Macro Variables

The different macro variables used throughout the paper are:

- **Output**: Gross Domestic Product, Non-farm business, Chained-$2005. From the NIPA tables of the Bureau of Economic Analysis (BEA). Series ID: A358RX1. We divide this series by the U.S. population (see below) to get an hours per capita measure.

- **Price deflator**: The main series we use is the Personal Consumption Expenditure (PCE) deflator, from the NIPA tables of the BEA; index, 2005=100. Series ID: A002RD3.

- **Population**: Non-civilian population, 16 years old and over; from the Bureau of Labor Statistics’ (BLS) Labor Productivity and Costs (LPC) program. Series ID: LNU00000000.

A.2 Labor Productivity and Costs (LPC)

The Major Productivity and Costs program of the BLS produces labor productivity and costs (LPC) measures for the private-sector U.S. economy. Below we list the variables we use from the LPC dataset. All of them are available quarterly (seasonally adjusted) and annually.

- **Compensation**: Total compensation from the LPC dataset is comprised of a ‘wages and salaries’ component, and a ‘supplements’ component. The ‘wages and salaries’ component is based on earnings data from the Quarterly Census of Employment and Wages (QCEW), previously known as the BLS ES-202 program. The QCEW is "...a cooperative program involving the Bureau of Labor Statistics (BLS) of the U.S. Department of Labor and the State Employment Security Agencies (SESAs)...[and] produces a complete tabulation of employment and wage information for workers covered by State unemployment insurance (UI) laws and Federal workers covered by the Unemployment Compensation for Federal Employees (UCFE) program". This represents about 98 percent of all U.S. jobs. The definition of labor earnings in the QCEW are very comprehensive. Specifically: "Wage and salary

---

28 The proportion of wages and salaries in total compensation has been trending downwards in a constant way through time, from around 91% of total compensation in the mid-1960s to 80% in 2010.
disbursements consist of the monetary remuneration of employees (including the salaries of corporate officers, commissions, tips, bonuses, and severance pay); employee gains from exercising nonqualified stock options; distributions from nonqualified deferred compensation plans; and an imputation for pay-in-kind (such as the meals furnished to the employees of restaurants).” See http://www.bea.gov/regional/pdf/spi2005/Complete_Methodology.pdf for more information.

The 'supplements' components consists of employer contributions for employee pension and insurance funds and employer contributions for government social insurance. To derive total compensation for the non-farm business sector, the LPC substracts compensation of employees working in public administration offices, in the farm sector, and in non-profit institutions and private households. Moreover, the LPC imputes earnings of self-employed individuals using comparable data from workers in the CPS.

The total compensation measure we use from LPC is series ID: PRS85006063, which is in levels and not publicly available (the LPC website of the BLS only publishes the corresponding index series).

- **Hours**: Total hours in the LPC database mainly comes from the Current Establishment Survey (CES) for production and nonsupervisory workers (see CES description below), supplemented by other sources to estimate hours of workers not covered by the CES. For example, LPC computes an estimate of average weekly hours for nonproduction and supervisory workers by applying a CPS-based ratio of [nonproduction & supervisory workers] / [production & non-supervisory workers] to CES production & nonsupervisory worker average weekly hours. The total hours measure we use is LPC series ID: PRS84006033. This series is in levels and not publicly available, as for the total compensation series.

- **Total employment**: We use LPC’s employment series PRS85006013, which, as for compensation and hours above, is in levels.

---

29The estimates for the 'supplements' portion of total compensation come from various sources, such as the IRS, the Medical Expenditure Panel Survey, or the American Council on Life Insurance. The estimates are compiled by the Bureau of Economic Analysis (BEA).

30Note that workers employed in 'general government' are not included in the non-farm business measure, while workers in 'government enterprises' are.

31We thank Shawn Sprague for supplying us with the level series from LPC.

32Further note that when reporting statistics for total hours (e.g. in Table 2, where we use total hours as a cyclical indicator), we use LPC series ID: PRS84006033 and divide it by the U.S. population to get an hours per capita measure.
• **Average weekly earnings:** We divide by 52 the ratio of total compensation to total employment.

• **Average weekly hours:** We divide by 52 the ratio of total hours to total employment.

• **Average hourly wage:** We compute average hourly earnings by dividing average weekly earnings with average weekly hours.

### A.3 The Current Population Survey (CPS)

The Current Population Survey (CPS) is a monthly survey of about 60,000 households. It collects a variety of information on households’ demographics and employment.\(^{33}\) Since we analyze (mainly) earnings and hours in this paper, we need information on both from the CPS. However, earnings and hours questions are not asked to all CPS respondents each month. Specifically, an interviewed individual appears in the CPS for two periods of four consecutive months, separated by eight months during which the individual is left out of the survey. Between 1973 and 1978, the CPS asked all the respondents in the sample about weekly earnings and weekly hours once a year only. This data was collected in May in what is called the 'May supplements'. Starting in 1979, weekly earnings and hours questions are asked each month to the individuals who are at the end of a four-month rotation – the 'Outgoing Rotation Group' (ORG). Hence, from 1979 onward, one fourth of the CPS sample is asked about earnings and hours each month.\(^{34}\)

Following Abraham et al. (1998) and Lemieux (2006), we use the earnings and hours information from the CPS May supplements and the ORG extracts to create an annual series of weighted average weekly earnings and weighted average weekly hours from 1973 onwards. The individual weights used in this calculation are provided by the CPS to make the resulting sample representative of the U.S. workforce.

Since sectoral coverage of the LPC and the CES series differs slightly and the LPC coverage cannot be replicated directly, we use a private non-agricultural coverage for the CPS that resembles the NIPA and CES coverage (see below); i.e. we remove from the CPS May / ORGs extracts all unemployed; self-employed; individuals under 16 years of age; all government, agricultural and private household workers; as well as former armed force personnel. For 1973-1978, the May supplements yield an average of 30,406 individual observations per year. For 1979 onward, the combination of 12 monthly ORG files yields an average of 139,230 individual observations per year.

---

\(^{33}\)For more documentation on the CPS and in particular the May / ORG extracts, see Schmitt (2003); and Roth and Feenberg (2007).

\(^{34}\)In March of each year, the CPS also asks all individuals in the sample about their annual labor earnings. Extending our earnings analysis using the CPS March earnings data remains to be done.
Lastly, note that the actual CPS ORG extracts (1979-2011) we use are from the Center for Economic Policy Research (CEPR). These extracts are based on the 'Merged Outgoing Rotation Groups' files ('MORGs', i.e. the ORGs, merged in annual files) compiled by the National Bureau of Economic Research (NBER). We use the CEPR data because the CEPR modifies the NBER MORGs to make them more user-friendly. But the greatest advantage with the CEPR data is that they provide detailed documentation on the modifications and additions they make to the NBER's MORG files. We use this documentation to replicate the CEPR’s adjustments for the NBER MORG extracts for the CPS May Supplement files so as to have consistent variables throughout the whole sample (1973-onwards).

- **Compensation:** Workers in the CPS May/ORG extracts report earnings in two different ways, depending on whether they are salaried or paid by the hour. Salaried workers report usual weekly earnings, defined as compensation normally received, including bonuses, overtime, tips and commissions (OTC) if paid and earned each period but excluding payments in kind, stock options, any other form of irregular bonuses, and any supplements to wage earnings. Hourly-paid workers report their usual hourly wage rate, which is not supposed to take into account OTC or any form of irregular pay, and are also asked their usual weekly earnings, as asked to salaried workers. Hence, CPS earnings contain some fraction of bonuses and OTC if paid and earned each period but no irregular form of compensation.

To create consistent average hourly (and weekly) earnings series from this data, two issues need to be addressed. The first issue concerns topcoding of high earnings; the second issue concerns the computation of treatment of OTC earnings for hourly-paid workers.

Topcoding concerns the fact that the CPS limits (i.e. topcodes) publicly available data of individuals with high earning to a maximum value that varies over time and depends on whether a worker reports weekly earnings or the hourly wage rate. For the latter, the CPS topcodes the hourly rate at $99.99, a threshold rarely crossed. For the former (i.e. salaried workers), the CPS topcodes weekly earnings at $999 until 1989; $1923 between 1989 and 1997; and $2884 from 1998 onward. For certain years, this puts a substantial share of workers above the topcode, which may lead to earnings discontinuities around topcode changes. To reduce
this risk of discontinuities, we multiply topcoded weekly earnings by a factor of 1.3 before averaging across individuals. While this constant-factor adjustment is standard in the labor literature (e.g. Abraham et al., 1998; Lemieux, 2006), it does not completely eliminate the possibility of discontinuities from topcode changes. Alternatively, one can use more sophisticated adjustment methods that estimate mean earnings of individuals above the topcode from the cross-sectional distribution of earnings below the topcode. The most popular among these methods is based on the Pareto distribution which, for certain years, has been shown to provide a better approximation of actual earnings in confidential CPS samples. In the paper, we provide a new method to account for topcoding, by using data from Piketty and Saez (2003) on the top income earners in the U.S. See the next section of this Appendix for more details.

The second issue with creating a consistent average hourly wage series from CPS data concerns the treatment of OTC earnings for hourly-paid workers. Prior to 1994, hourly-paid workers were simply asked to report their hourly wage rate as well as their weekly earnings. With the redesign of the CPS survey in 1994, the CPS introduced an additional question on weekly OTC earnings for hourly-paid workers but not for salaried workers. The consequence of this additional question is a more accurate measurement of OTC earnings for hourly-paid workers starting in 1994 (see below for more discussion). Due to this discontinuity in earnings reporting, the main statistics in the paper using CPS earnings data simply use the hourly wage rate times the usual hours worked as the measure of weekly earnings for hourly-paid workers. See below for more details and the treatment of OTC earnings.

- **Hours**: Hours in the CPS May / ORGs are recorded as the usual number of hours per week

---

39 Furthermore, the 1994 CPS redesign also affected the hours reported by individuals in the ORGs. See hours bullet point below for more information.
40 For example, before 1994, hourly-paid workers provided their hourly wage (not including OTC earnings), and then were asked to provide their usual weekly earnings (supposedly including OTC earnings). Our calculations (see paper) show that workers often did not include OTC in weekly earnings and, as a result, the average difference between weekly earnings for hourly-paid workers vs. weekly earnings not including it (i.e. hourly wage rate times weekly hours) is small. Starting in 1994, the new, more precise question about OTC earnings made the hourly-paid individuals respond more precisely about their weekly OTC earnings. Confirmation of this issue comes from Polivka (2000) who concludes: "Prior to 1994, workers identified as paid by the hour were simply asked to report their hourly rate, the number of hours they worked and then a weekly amount in addition. The repetitive process of asking these questions irked some respondents provoking statements such as, "Well, figure it out yourself." (Polivka and Rothgeb, 1993)."

38
worked on the main job. As for compensation above, the CPS redesign in 1994 created a small consistency problem with hours; from 1994 onward, the redesigned CPS allowed respondents to indicate that their "hours vary". As a result, starting in 1994, no response for usual weekly hours is recorded for these individuals. As Schmitt (2003) notes: "a sizeable share of workers (typically, 6-7%) chose to report that their hours vary. Since the distribution of hourly earnings for these workers may differ systematically from that of workers whose hours generally do not vary, simply excluding the group of workers whose hours vary may reduce comparability of wage series across the 1994 redesign." The CEPR CPS ORG extracts we use in this paper impute weekly hours for these individuals whose hours are missing. As a result, we use the CEPR-generated variable 'uhoursi', which is a variable created by the CEPR using the variable 'uhourse' in the NBER CPS ORG extracts ('uhourse' reports the usual weekly earnings for each individual CPS respondent in the NBER ORGs). 'uhoursi' equals 'uhourse' for the 1979-1993 period; from 1994-on, it equals 'uhourse', unless a CPS respondent answered that his hours vary. In that case, 'uhoursi' equals the hours imputation made by the CEPR.\footnote{See Schmitt (2003) for more details on the imputation procedure.}

- **Average weekly earnings:** To compute average weekly earnings, we proceed differently for salaried and hourly-paid workers. For salaried workers, we simply use the weekly earnings reported under the variable 'earnwke' in the CPS ORGs extracts for the whole sample. For hourly-paid workers, our main results compute weekly earnings as hourly wage rate times weekly hours, thus omitting OTC earnings.\footnote{As mentioned above, to compute average weekly earnings in the CPS sample, we use a weighted average of individual weekly earnings, where the weights are individual weights provided in the CPS May / ORGs extracts.}

In the second half of the paper, we account for OTC earnings of hourly-paid workers as follows. Before 1994, weekly earnings of hourly-paid workers are computed as $\max(wage\_rate \times weekly\_hours, weekly\_earnings)$; from 1994 onward, weekly earnings of hourly-paid workers are computed as $\max(wage\_rate \times weekly\_hours + OTC, weekly\_earnings)$, where the term 'weekly\_earnings' in the brackets refers to the usual weekly earnings variable, labeled 'earnwke' in the CPS ORGs.\footnote{Note that between 1979 and 1988, the response of hourly-paid workers for the weekly earnings question has been recorded under another variable in the CPS ORGs, labeled "uearnwke" See Feenberg and Roth (2007).}

- **Average weekly hours:** To compute average weekly hours, we use a weighted average of the variable 'uhoursi' described above. As for average weekly earnings, the weights used are individual weights provided in the CPS.
• **Average hourly wage**: We compute average hourly wages by dividing (weighted) average weekly earnings with (weighted) average weekly hours.

### A.4 Current Employment Statistics

The CES is a monthly survey of employment, wages and hours in the private non-agricultural establishments. The CES grew from about 166,000 to about 330,000 establishments between 1980 and 1993; and then to over 400,000 establishments in 2006. Today, the CES covers about 141,000 firms representing approximately 486,000 individual worksites. While the CES reports data for all employees as far back as 1939, it only reports earnings and hours from 1964 onwards and only for production workers in the goods-producing sector and nonsupervisory workers in the service-providing sector.\(^{44}\)

- **Compensation**: Chapter 2 of the BLS Handbook of Methods states that: "*Aggregate payrolls include pay before deductions for Social Security, unemployment insurance, group insurance, withholding tax, salary reduction plans, bonds, and union dues. The payroll figures also include overtime pay, shift premiums, and payments for holidays, vacations, sick leave, and other leave made directly by the employer to employees for the pay period reported. Payrolls exclude bonuses, commissions, and other lump-sum payments (unless earned and paid regularly each pay period or month), or other pay not earned in the pay period (such as retroactive pay). Tips and the value of free rent, fuel, meals, or other payments in kind are not included."

- **Hours**: Chapter 2 of the BLS Handbook of Methods states that: "*Total hours during the pay period include all hours worked (including overtime hours), hours paid for standby or reporting time, and equivalent hours for which employees received pay directly from the employer for sick leave, holidays, vacations, and other leave. Overtime and other premium pay hours are not converted to straight-time equivalent hours."

- **Average weekly earnings**: We downloaded the average weekly earnings series from the CES, series ID: CES0500000030. It is computed (in the CES) as the weekly average of total earnings divided by total employment.

- **Average weekly hours**: We downloaded the average weekly hours series from the CES, series ID: CES0500000007. It is computed (in the CES) as the weekly average of total hours divided by total employment.

\(^{44}\)Note that since March 2006, the CES also publishes series of weekly earnings and hours that cover all employees in the non-farm business sector.
- **Average hourly wage**: The average hourly wage is computed as the ratio of average weekly earnings to average weekly hours.

### A.5 National Income and Product Accounts (NIPAs)

The National Income and Product Accounts, produced by the Bureau of Economic Analysis (BEA), provide, among several other macroeconomic variables, detailed information on compensation of workers at the national and industry levels. Contrary to LPC, NIPAs provide compensation of employees series for the whole economy (not restricted to the non-farm business sector); moreover, the NIPA tables provide details on the ‘wage and salaries’ and ‘supplements’ parts of total compensation (while LPC only provides information on total compensation).

To calculate non-farm business series for the variables defined below, we take total private figures from the NIPA tables and subtract total agriculture, which includes farms, agricultural services, forestry, and fishing. Note that this definition is similar to the CES’ non-farm business sector, as well as our CPS non-farm business definition; however, it slightly differs from LPC’s non-farm business definition (see LPC subsection above). The reason for these differences is that from the publicly-available NIPA tables, we do not have the information to replicate LPC’s non-farm business sector definition. Further note that NIPA private non-agricultural series do not include an imputation of self-employed workers’ earnings, hours, and employment.

- **Compensation**: Total compensation in the NIPAs is computed as the sum of ‘wages and salaries’ and ‘supplements to wages and salaries’. As in LPC, the ‘wages and salaries’ portion of total compensation is based on earnings data from the Quarterly Census of Employment and Wages (QCEW), and the ‘supplements to wages and salaries’ is computed by the BEA.
  - The total compensation measure we use from the NIPAs is taken from Table 6.2, series ID: A033RC0.
  - The ‘wages and salaries’ portion of total compensation, available from the NIPAs, is taken from Table 6.3 ID: A034RC0.

- **Total employment**: Total employment series comes from NIPA Table 6.4, series ID: A4201C0. It includes all full- and part-time workers. The BEA source for employment is the QCEW.\(^{45}\)

- **Hours**: Total hours are taken from NIPA Table 6.9, series ID: B4701C0. Total hours include all hours worked by full- or part-time employees. It is computed (by the BEA) as total NIPA...

\(^{45}\)The CES actually benchmarks its employment estimates each year to Census data (i.e. QCEW).
employment times average weekly hours. The source of average weekly hours is LPC and do not include self-employed workers.

- **Average weekly compensation and wages & salaries:** We divide by 52 the ratio of total compensation to total employment for average weekly compensation, and divide by 52 the ratio of total wages and salaries to total employment for average weekly wages and salaries.

- **Average weekly hours:** We divide by 52 the ratio of total hours to total employment.

- **Average hourly compensation and wages & salaries:** We divide average weekly compensation with average weekly hours to get average hourly compensation, and divide average weekly wages and salaries by average weekly hours to get average hourly wages & salaries.

**B Topcode adjustments: Income data from Piketty-Saez**

We use the dataset on income inequality constructed by Piketty and Saez (P-S, thereafter) from IRS data first release with their seminal 2003 QJE paper and updated to 2010 since then. This dataset is available on Saez’s website at: http://elsa.berkeley.edu/~saez/. P-S provide an analysis of inequality at two levels: 1) at the income level (with and without capital gains) and 2) at the wages and salaries level. Since our work focus on the wage portion of income, we will use their data on "wage inequality" instead of income inequality, because the earnings concept from the P-S wage inequality data is very similar to the earnings concept in the QCEW (detailed above).

In their wage inequality dataset, P-S provide wage shares (of total wages and salaries), average wages and salaries for the top-10% decile and numerous fractiles within the top-10%, as well as wages and salaries threshold values for these fractiles. Below, we will use these average (and thresholds) wages and salaries to estimate mean earnings for topcoded observations in the CPS data.

**B.1 Using P-S data to estimate means above the topcode in the CPS**

We use the information in the P-S dataset on average and threshold wages and salaries values for various fractiles within the top-10% of income earners. Then, using the proportion of CPS respondents with topcoded earnings each year in the CPS May and ORGs, we can impute a value for weekly earnings to these respondents using the P-S data. Here are the detailed steps we follow to compute the assigned weekly earnings (to topcoded earnings in the CPS May / ORGs) in each year:
1. Gather, for the years 1973 to 2009, average (nominal) annual earnings for the top-5%, top-1%, top-0.5%, top-0.1%; and gather threshold values for the 95th, the 99th, the 99.5th, and the 99.9th percentiles from P-S dataset.\textsuperscript{46}

2. Convert these annual earnings in weekly earnings (divide annual earnings by 52).

3. Gather, from the CPS May / ORGs, the densities (%) of workers with topcoded weekly earnings in each year.\textsuperscript{47}

4. Compute the values to assign (from the P-S dataset) in each year to the topcode earnings observations in the CPS May / ORGs.

Steps (1) to (3) are straightforward, but step (4) is more complicated. There are two main reasons why the assignation of P-S values to topcoded observations is not simple; first, the P-S average and thresholds earnings values do not correspond exactly to the densities of observations with topcoded earnings in the CPS. As a result, we need a procedure that uses some average and/or threshold fractiles values from P-S and compute an approximative earnings value to assign to CPS topcoded observations.

Second, as mentioned above, the topcode value changes through time in the CPS; as a result, the densities of observations at the topcode change throughout the sample, especially when the topcode value changes. For example, in 1988, the proportion of observations with topcoded weekly earnings (for the whole economy\textsuperscript{48}) was 4.17%, while in 1989 is drops to 0.45%. The same pattern is observed between 1997 and 1998, the other time the topcode value changes in the CPS: in 1997, the % of topcoded earnings is 1.50%, while in 1998 it is 0.60%. These changes in densities when the topcode values change are important because they can guide us in imputing reasonable values to the topcoded observations. For example, we assume that it is highly improbable that in 1988 (or before), more than 0.45% of individuals made above $1923/week, since in 1989, only 0.45% of individuals made more than $1923/week. Consequently, even though 4.17% of individuals had topcoded earnings (topcoded at $999/week) in 1988, we assume that no more than 0.45% is assigned

\textsuperscript{46}Ideally, we would like to use more precise values (i.e. that coincide exactly with the CPS densities of observations at the topcode described below) but P-S provide only these average and thresholds values for earnings.

\textsuperscript{47}Note that for simplicity, we only consider salaried workers as potential workers with topcoded weekly earnings. The reason behind this is that the topcode level for the hourly wage is $99/hour, a threshold almost never crossed throughout the sample.

\textsuperscript{48}As specified in the paper, the whole economy sector is defined as all workers less private households and military workers (who are not asked the earnings questions).
a value higher than $1923/week.\footnote{Of course, this assumption eliminates the possibility of large swings in high incomes due to business cycles that would change the density of people with topcoded earnings in the CPS. The reason we make this assumption is that when we do not take into account these sharp drops in densities when the topcode value changes, we obtain unrealistically large decreases in CPS average wages in the years after the topcode value changes, i.e. a year where the density is very high (e.g. 1988), vs. 1989, when the topcode value changes.} Let us now turn to the actual assignment procedure from the P-S dataset.

- For years where the density (%) of topcoded earnings is lower than 1%, we use wage information in P-S for fractiles within the top-1% to impute values to these topcoded observations. For example, in 1975, the density (%) of topcoded earnings was 0.21%. For this year, we assign to 0.1% of observations the top-0.1% average weekly earnings value in P-S (labelled "$P(99.9 - 100)$", for the average weekly earnings of the top-0.1%), and to the remaining topcoded values (i.e. 0.21%-0.1% = 0.11%) we assign the P-S average earnings value $P(99.5 - 99.9)$, i.e. the average earnings for individuals with wages between the 99.5 and 99.9 fractiles, since we do not have the exact average earnings value from P-S for those 0.11% observations. The detailed formula to compute the assigned weekly earnings in 1975 is thus:

$$\text{assigned wkly earnings} = \frac{0.1}{0.21} \times P(99.9 - 100) + \frac{(0.21 - 0.1)}{0.21} \times P(99.5 - 99.9)$$

- For years where the density (%) of topcoded earnings is higher than 1%, we use average earnings value for the top-1% (i.e. $P(99 - 100)$), and a weighted average of the $P95$ and $P99$ percentiles thresholds for the rest of topcoded observations (again, we proceed accordingly because we only have details on $P(95 - 99)$, $P95$, and $P99$ between the 95th and 99th percentiles in the P-S data). Take year 2005 as an example (where % topcoded earnings is 1.20%); the detailed formula to compute the assigned weekly earnings (in 2005) from P-S is thus:

$$\text{assigned wkly earnings} = \frac{1}{1.20} \times P(99 - 100) + \frac{(1.20 - 1)}{1.20} \times \left[ \frac{0.20}{4} \times P95 + \frac{(4 - 0.20)}{4} \times P99 \right]$$

where, as above, $PXX$ corresponds to the weekly earnings threshold for the $XXth$ percentile.

- Finally, to be consistent with our assumption above\footnote{Recall that topcode values in the CPS change two times throughout the sample, in 1989 and in 1998. We assume it is improbable that in 1988 (or before), more than 0.45% of individuals made above $1923/week, since in 1989, only 0.45% of individuals made more than $1923/week. The same analogy applies from 1989 to 1997: in these years, we assume that no more than 0.60% of individuals made above $2884/week, since this is the proportion of individuals with topcoded earnings (at $2884/week) in 1998.}, between 1979-88 and 1989-97, if the density (of topcoded observations) in any year is higher than the density in the year the
topcode value subsequently changes (i.e. in 1989 and in 1998), we do not assign an earnings value from P-S that is higher than the "new" topcode value in the CPS (in the year it changes). To illustrate this "rule", take an example: in 1988, the topcode density is 4.17%. We thus use the procedure described above to assign an earnings value to the top-0.45% (i.e. the density in 1989 after the topcode value changes in the CPS from $999/week to $1923/week); for the rest of the topcoded observations (i.e. 4.17% - 0.45% = 3.72%), we follow the same procedure as above unless the assigned topcode values exceeds $1923/week, the new CPS topcode value in 1989. In that case, we simply use a 1.3 multiplicative factor (times the CPS topcode value in 1988, i.e. $999*1.3). Therefore, the detailed formula to compute the assigned weekly earnings in 1998 is:

$$\text{assigned wkly earnings} = \frac{0.45}{4.17} \left( \frac{0.1}{0.45} \cdot P(99.9 - 100) + \frac{(0.45 - 0.1)}{0.45} \cdot P(99.5 - 99.9) \right) + \frac{(4.17 - 0.45)}{4.17} \cdot 1.3 \cdot 999$$

i.e. all observations above 0.45% were assigned $999/week time 1.3 since the assigned P-S value found was higher than the CPS topcode value in 1989 (i.e. $1923/week). This would mean that more than 0.45% of workers in 1988 would have earned weekly earnings above than $1923, even though only 0.45% of CPS workers had topcoded earnings in 1989 at $1923/week.

By using a 1.3 multiplicative factor, we rule out this possibility.

C Computation of Standard Errors

Standard errors and relative standard errors in the text are obtained using the delta method from GMM-based estimates. In the first stage, define

$$f(x_{it}, \mu) = \begin{bmatrix} x_{1t} - \mu_1 \\ \vdots \\ x_{Nt} - \mu_N \\ x_{1t}x_{1t} - \mu_{11} \\ \vdots \\ x_{Nt}x_{Nt} - \mu_{NN} \end{bmatrix},$$

where \(x_{it}\) are the time series of interest for \(t = 1, \ldots, T; \mu_i = E(x_{it})\) for \(i = 1, \ldots, N;\) and \(\mu_{ij} = E(x_{it}x_{jt})\) for \(i, j = 1, \ldots, N.\) The GMM estimator sets \(\hat{\mu}\) such that \(\frac{1}{T} \sum_{t=1}^{T} f(x_{it}, \mu) = 0.\) The asymptotic distribution of the GMM estimator is given by

$$\sqrt{T}(\hat{\mu} - \mu) \rightarrow N \left(0, \{D'S^{-1}D\}^{-1}\right),$$
where
\[ D = E \left[ \left( \frac{\partial f(x_t, \mu)}{\partial \mu'} \right)' \right] \]
is the Jacobian matrix (N x N since our GMM procedure is just-identified), and where
\[ S = \sum_{j=-\infty}^{\infty} E \left[ f(x_t, \mu)f(x_t, \mu)' \right]. \]

Next, compute the covariance matrix for the standard errors
\[ COV(\mu) = \frac{\{D'S^{-1}D\}^{-1}}{T}. \]

To construct a sample analog of S, we use the Newey-West estimate of S:
\[ \hat{S}_T = \sum_{j=-k}^{k} \left\{ \left( \frac{k - |j|}{k} \right) \frac{1}{T} \sum_{t=1}^{T} f(x_t, \hat{\mu})f(x_{t-j}, \hat{\mu})' \right\}. \]

Now, our moments of interest are standard deviations and relative standard deviations (non-linear functions of the moments found above), so we use the delta Method to estimate the standard errors of these standard deviations and relative standard deviations. For example, consider the standard deviation of a random variable \( x_{it} \):
\[ \sigma_x = (E(x_t^2) - E(x_t)^2)^{1/2}. \]

Here we interpret \( \sigma_x \) as a function of the population moments \( E(x_t) \) and \( E(x_t^2) \). Moreover, define
\[ \mu = \begin{bmatrix} E(x_t) & E(x_t^2) \end{bmatrix} \equiv \begin{bmatrix} \mu_x & \mu_x \end{bmatrix}, \]

and thus
\[ \sigma_x(\mu) = (\mu_{xx} - \mu_x^2)^{1/2} = X(\mu_x, \mu_{xx}). \]

The delta method states that
\[ \sqrt{T}(\hat{X} - X) \rightarrow N \left( 0, \frac{\partial X}{\partial \mu} \left\{ D'S^{-1}D \right\}^{-1} \frac{\partial X'}{\partial \mu} \right). \]

Since \( D = \frac{\partial f}{\partial \mu} = -I \), where \( I \) denotes the identity matrix of appropriate dimension, \( \{D'S^{-1}D\}^{-1} \) reduces to \( S \). Furthermore, we can compute the derivative of \( X \) with respect to \( \mu \)
\[ \frac{\partial X}{\partial \mu} = \left[ \begin{array}{c} \frac{\partial \sigma_x}{\partial \mu_x} \\ \frac{\partial \sigma_x}{\partial \mu_{xx}} \end{array} \right] = \left[ \begin{array}{c} -\mu_x/\sigma_x \\ 1/(2\sigma_x) \end{array} \right]. \]
With these in hand, along with the estimate of $S$, $\hat{S}_T$, we can compute the standard error of $\sigma_x$. We use the same procedure to find the standard errors for relative standard deviations (e.g. the ratio of the standard deviations of wages and output), where the derivative of $X$ with respect to $\mu$ is:

$$
\frac{\partial X}{\partial \mu} = \frac{\partial \{\sigma_x/\sigma_y\}}{\partial \mu} = \begin{bmatrix}
\frac{\partial \sigma_x}{\partial \mu_x} \\
\frac{\partial \sigma_y}{\partial \mu_y} \\
\frac{\partial \sigma_x}{\partial \mu_{xx}} \\
\frac{\partial \sigma_y}{\partial \mu_{yy}} \\
\frac{\partial \mu_x}{\partial \mu_{xx}} \\
\frac{\partial \mu_y}{\partial \mu_{yy}} \\
\end{bmatrix} = \begin{bmatrix}
\frac{-\mu_x}{\sigma_x \sigma_y} \\
\frac{\mu_y X}{\sigma^2} \\
\frac{1}{2 \sigma_x \sigma_y} \\
\frac{-1}{2 \sigma^2} \\
\end{bmatrix}.
$$