# Reconciling Trends in U.S. Male Earnings Volatility:

## Results from Survey and Administrative Data


Robert Moffitt
Johns Hopkins University

John Abowd
U.S. Bureau of the Census
and Cornell University

Christopher Bollinger
University of Kentucky

Michael Carr
University of Massachusetts Boston

Charles Hokayem
U.S. Bureau of the Census

Kevin McKinney
U.S. Bureau of the Census

Emily Wiemers
Syracuse University

Sisi Zhang
Jinan University

James Ziliak
University of Kentucky


June 16, 2022

Abstract

Reconciling Trends in U.S. Male Earnings Volatility:
Results from Survey and Administrative Data

There is a large literature on earnings and income volatility in labor economics, household finance, and macroeconomics. One strand of that literature has studied whether individual earnings volatility has risen or fallen in the U.S. over the last several decades. There are strong disagreements in the empirical literature on this important question, with some studies showing upward trends, some showing downward trends, and some showing no trends. Some studies have suggested that the differences are the result of using flawed survey data instead of more accurate administrative data. This paper summarizes the results of a project attempting to reconcile these findings with four different data sets and six different data series--three survey and three administrative data series, including two which match survey respondent data to their administrative data. Using common specifications, measures of volatility, and other treatments of the data, four of the six data series show a lack of any significant long-term trend in male earnings volatility over the last 20-to-30+ years when differences across the data sets are properly accounted for. A fifth data series (the PSID) shows a positive net trend but small in magnitude. A sixth, administrative, data set, available only since 1998, shows no net trend 1998-2011 and only a small decline thereafter. Many of the remaining differences across data series can be explained by differences in their cross-sectional distribution of earnings, particularly differences in the size of the lower tail. We conclude that the data sets we have analyzed, which include many of the most important available, show little evidence of any significant trend in male earnings volatility since the mid-1980s.

The literature on labor market volatility is vast and touches on multiple areas of macroeconomics, household finance, labor economics, and overlaps between them. The classic study of permanent vs transitory components of income and their implications for consumption, saving, and the marginal propensity to consume is just one example (Friedman, 1957; Hall and Mishkin, 1982). On the micro level, this literature has spilled over into household finance, with its concern with liquidity constraints, ability to deal with income shocks, possible inadequacy of assets to deal with such shocks, and consequent inability to smooth consumption sufficiently (Carroll, 1997; Gourinchas and Parker, 2002; Blundell et al., 2008; Ganong and Noel, 2019). In labor economics, a literature going back to the 1960s and 1970s on sectoral models of the labor market, with one sector characterized by high wages and stable jobs and another characterized by low wages and unstable jobs, has reemerged in recent discussions of technological change and the decline of union and manufacturing jobs, since the latter are generally more stable than average (Taubman and Wachter, 1986; Katz and Autor, 1999). The impact of income uncertainty on investments in human capital, both educational and on-the-job, and on investments in children at young ages, has generated yet another discussion in labor economics (Levhari and Weiss, 1974; Cunha et al., 2005; Carneiro and Ginja, 2016).

An important empirical branch of this literature concerns whether volatility has changed over time in the U.S. A priori, whether labor market volatility should be expected to have risen or fallen differs by perspective. On the one hand, the just-mentioned literature on structural change in the U.S. labor market suggests that earnings instability might have increased, at least for workers with medium or lower skills. Katz and Autor (1999), for example, in their review

of the early literature on increasing earnings inequality, make the connection between rising earnings inequality and rising instability directly. Haider (2001) also explicitly draws a connection between growing earnings inequality and earnings instability. On the other hand, a prominent hypothesis in macroeconomics is that the 1980s ushered in a period known as the Great Moderation, reflected in declining levels of aggregate volatility (McConnell and Perez-Quiros, 2000). While there is no necessary connection between aggregate volatility and volatility at the micro level (as noted by Davis and Kahn, 2008, and Dynan et al., 2012), some macroeconomists argue that a lack of decline in individual earnings volatility matching the aggregate volatility decline is intuitively difficult to explain (Sabelhaus and Song, 2010).

The project which this Overview summarizes represents an effort to bring several data sets to bear on the question of whether U.S. earnings volatility at the micro level has risen, fallen, or remained constant over the last several decades. It is motivated in large part by the disparate findings on this question which have appeared when different data sets have been used. The workhorse data set for estimating trends in individual earnings volatility in the U.S. has been the Panel Study of Income Dynamics (PSID), a longitudinal survey that has been ongoing since 1968 (and is hence the longest-running general-purpose socioeconomic panel in the world), which has attempted to maintain reasonable population representativeness and which asks extensive questions on labor market activity. The use of the PSID for the study of male earnings volatility began with Gottschalk and Moffitt (1994), who found male earnings volatility to have increased from 1970 to 1987, with the largest increase occurring among the less educated. About a dozen PSID studies subsequent to the Gottschalk-Moffitt study have also found increases in male earnings volatility over time (Dynarski and Gruber (1997), Haider (2001), Hyslop (2011), Keys (2008), Heathcote et al. (2010), Shin and Solon (2011), Dynan et al.

(2012), Moffitt and Gottschalk (2012), Jensen and Shore (2015), and Carr and Wiemers (2018); see Moffitt and Zhang (2018) for a review). However, as we will discuss in detail in the first section of the paper below, findings have often differed in other data sets. While some differing findings have been found in other survey data sets (e.g., the Survey of Income and Program Participation and, partially, the Current Population Survey), the largest differences have emerged from studies using administrative data from Social Security, tax, or Unemployment Insurance records, which often find no trend in earnings volatility or even a decline. We review these studies in detail in the first section of the paper below. The difference in trends found in administrative data, which are often presumed to be more accurate than survey data, suggests that the PSID may be biased by reporting error, attrition bias, or some other issue.

The project brings four different subprojects and six different data series to bear on this question, each using common samples to the maximum extent feasible, common definitions of volatility, and common other treatments of the data (trimming of outliers, treatment of nonworkers, imputations, and others). One paper reexamines the oft-used PSID, but adds to previous work by extending the time period up through 2016--which turns out to be important--and by conducting a number of analyses of bias that might come from attrition and other threats to representativeness. A second paper uses data from the Current Population Survey (CPS), using earnings reports of individuals who are in the survey twice over a two-year period. But this paper also links the CPS sample to Social Security earnings files, permitting a direct comparison of survey reports and administrative data reports for the same individuals. The third paper uses the Survey of Income and Program Participation (SIPP), a well-known Census Bureau survey intended to be representative of the population and which consists of a rolling series of 2-to-5-year panels, thereby permitting estimates of the volatility of year-to-year changes

in earnings. But, like the second paper, this third paper also employs a data set of Social Security earnings data matched to the SIPP survey respondents, again providing the opportunity to compare trends in earnings volatility between the two types of data. The fourth paper uses only administrative data drawn from Unemployment Insurance (UI) wage records collected by employers and reported to state governments. The well-known file is called the Longitudinal Employer-Household Dynamics (LEHD) data set and has the near universe of earnings of UI-covered workers in the states that have provided their data. The importance of the LEHD is not only its vast sample size compared to the survey-based data sets, but its different sampling frame—namely, the near-total universe of US workers and not just those who agreed to participate in a survey.

The four papers each focus on trends in male earnings volatility over the years for which they have data. The PSID goes back to 1970, the SIPP goes back to the early 1980s, the matched CPS data we have goes back to 1996, and the LEHD goes back to 1998. There is thus full overlap after 1998 and partial overlap in many earlier years. A summary of the analyses and findings of each paper is given in this Overview, and the four individual papers which follow in this issue of the journal provide additional detail on each subproject.

The major findings of the analysis are that, when treated on a comparable basis, there is considerable agreement in volatility trends across the data sets, although less in levels. In terms of levels, the LEHD has the highest volatility and, for the two data sources where linked administrative and survey data are available, the administrative data show higher levels of volatility than the survey data. In terms of trends, we summarize results for three successive time periods separately. First, while none of the data sets reach back as far in time as the PSID, two that reach back to the 1970s or early 1980s (the SIPP survey and administrative data) are

consistent with the PSID by showing some upward trends in volatility over that period. But, second, both the PSID and those two data sets also show no average trend from the mid-1980s to the late 1990s. Third, after the late 1990s, all six data series are available and all show countercyclical volatility patterns, rising just before and in the early phases of the Great Recession and falling afterwards. However, with the exception of the PSID and the LEHD, all data sets show no net trend from the late 1990s to the mid- or late 2010s after the Recession was over. The PSID shows a positive net trend, but smaller than that which occurred in the 1970s-1980s, and the LEHD shows a small negative net trend after 2011. We conclude that the evidence shows no strong overall trend in volatility among working men in the U.S. since the mid-1980s, i.e., over the last 30 years.

While we conduct a number of sensitivity tests to the robustness of these findings, by far the most important test we conduct relates to the size of the left tail of the cross-sectional earnings distribution in the different data series. The LEHD has a much larger left tail than the other data sets, and the two data sources that have matched survey and administrative data show larger left tails in the latter than in the former. Other studies have found similar patterns and have often ascribed it to underreporting of short employment spells, which have low earnings, in survey data. We show that when the cross-sectional earnings of the data sets are required to have the same distribution (that of the PSID), the levels of volatility in all the data sets are much closer to each other. Equally important, this exercise converts the small negative trend in the LEHD to a small positive trend and makes the trends in the other administrative data sets over that same period positive instead of zero. The changes are a result of different trends over time in the left tail of the earnings distribution in the PSID and the other data sets. We conclude that

the small differences in trends across the data series, particularly in the late periods of our data, are mostly a result of differences in trends in the left tail of earnings.

The next section of the paper reviews the conflicting findings in past work in more detail. The different data sets used in the project appear in the following section, after which the methods and results across the data sets are summarized and attempts at reconciling their differences are reported. A final section summarizes the findings and draws lessons for future work.

## I. Past Work: Additional Detail

As noted in the Introduction, there have been over a dozen studies of male earnings volatility using the PSID (Moffitt and Zhang, 2018, has a detailed table of those published prior to 2018, describing their samples and results, including some studies of female earnings volatility). They do not always align perfectly with each other in methodology, and many estimate error components models instead of the gross volatility models studied here, instead using the transitory variance as the measure of volatility. The studies differ by what years of data were available at that time the studies were conducted. Almost all studies show rising volatility from the 1970s to the 1980s, and either no trend or a downward trend through the 1990s (ignoring cyclical movements). Those studies which had data into the 2000s (e.g., Shin and Solon (2011) and Dynan et al. (2012)) show rising volatility since that time. Carr and Wiemers (2018) had data through 2012 and showed that PSID volatility continued to rise during the Great Recession. Moffitt and Zhang (2018) had data through 2014 and showed that volatility started to decline after the Recession.

The results from other survey data sets are mostly, but not always, roughly consistent with the PSID. Using across-wave matched CPS observations, Celik et al. (2012) also found increases in volatility from the 1970s to the 1980s, stability through the 2000s, then a resumption of an increase, like the PSID. But Ziliak et al. (2011), also using CPS matched data, found the same early trends as Celik et al. but more of a stable trend through the late 2000s, unlike Celik et al. and the PSID. Koo (2016) found similar trends with CPS matched data. Dahl et al. (2008, 2011) examined volatility in household income in the SIPP survey and found no trend over time unless imputed income is included. Celik et al. also estimated trends in male earnings volatility in the SIPP, finding an actual decline after 1984.

Starker differences with the PSID are often found in studies using administrative data. On the one hand, Carr and Wiemers (2018), using Social Security earnings records of SIPP respondents, found patterns similar to the PSID—rising through the early 1980s, declining through 2000, and then rising through the mid-2000s. But Guvenen et al. (2014), using Social Security earnings data not linked to a survey, found slight declines in male volatility from 1980 to 2011. Bloom et al. (2017), also using Social Security earnings records, showed separate volatility trends for men and women and found those for men to decline from 1978 to 2013. Dahl et al. (2008) use the Continuous Work History Sample (CWHS) from Social Security records and find declining male earnings volatility between 1985 and 2003, while Dahl et al. (2011) show declining earnings volatility for all workers in both the CWHS and SIPP-SSA linked data (their linked sample differed from that of Carr and Wiemers (2018)). And using administrative data from Unemployment Insurance records, Celik et al. found no trend in volatility from 1992-2008 for the 12 states available in the data, while DeBacker et al. (2013), using tax records, found no trend from 1987 to 2009. There are also studies showing declining volatility but which

combine men and women (Sabelhaus and Song (2009, 2010); Braxton et al. (2021), but these are non-comparable because female earnings volatility shows a decline in both the CPS and the PSID (Ziliak et al., 2011; Dynan et al., 2012). Thus the largest discrepancies with the PSID are from those studies using administrative data, with some studies using Social Security data finding declining volatility and those using UI or tax data finding stable rather than increasing volatility.

There has been little work on reconciling the discrepancies in trends across the data sets. Dahl et al. (2011) suggested that the differences in volatility trends in SIPP survey data and matched Social Security earnings data might be related to imputed earnings values in the SIPP (we carefully examine this issue with the SIPP survey data). Celik et al. presented the diverse findings from different data sets, but found no explanation for the differences (their study is the closest to ours but differs in a number of respects). Carr and Wiemers (2018) compared volatility trends in the PSID to those in SIPP-based Social Security earnings data—but not matched to the PSID--and found them to be approximately the same, similar to what we find.

## II. The Data Sets

The six data series used in this project are shown in Table 1: the PSID, the CPS survey, CPS linked Social Security earnings records, the SIPP survey, SIPP-linked Social Security earnings records, and UI earnings from the LEHD. The PSID has been analyzed many times before, so the primary purpose of including it is only to provide a baseline estimate using the same sample definitions, measures of volatility, and other analysis features as those in the other five data series.

For all data sets, only men 25-59 in each year are included. Regarding sampling frames, five of them are based on survey sampling frames of the non-institutional population and may be subject to nonresponse bias from those declining to participate (Groves et al., 2002; National Research Council, 2013). In addition, the PSID is only representative of the 1968 US population (at best, ignoring attrition and other issues) because it does not include post-1968 immigrants, as all the other data series do (see the CPS and PSID papers for some investigation of this issue). The CPS only includes those who were at the same address in the surveys one-year apart (because the Census Bureau just returns to the same address), but the CPS paper finds this restriction to affect only the level of volatility, not its trend (attriters have higher levels of volatility). The Social Security earnings records matched to the CPS and SIPP surveys are drawn from the same source—the Detailed Earnings Records (DER)—and necessarily exclude those who do are not linked across data sets, which may include undocumented workers and some who work off the books. The LEHD draws its administrative earnings histories from Unemployment Insurance records, and therefore excludes individuals not properly using Social Security numbers as well as those who work off the books, and excludes those who are not covered by the UI system. These issues are discussed in the LEHD paper. Most of these differences are unalterable and introduce inevitable noncomparability to some unknown degree.

Regarding the samples, as is well known, the PSID collects sufficient earnings information only on household heads and their spouses. Both the CPS and SIPP surveys have headship information but the SIPP administrative data do not and the LEHD does not because family composition is not collected in UI records, which are reported by employers. Thus the possible importance of headship can be explored in only three of the data series. As for the definition of earnings, the PSID, SIPP survey, and CPS files have wage and salary earnings and self-

employment income separately, but the SIPP administrative data and the LEHD contain both and they cannot be separated. Again, some examination of this issue can be conducted but not across all data sets. Estimates of volatility including and excluding nonworkers are conducted in the project but only where possible, for the LEHD data have no nonworkers (in the LEHD, this means the absence of a positive earnings record). The LEHD sample size is vastly larger than that in any of the other data sets, and the PSID sample size is the smallest.

In this Overview paper, we refer to a number of sensitivity tests conducted to gauge the importance of these cross-data differences. However, the details of those tests appear only in the individual papers.

## III. Volatility Measures

All analyses use simple and transparent summary measures of gross earnings volatility, calculating the earnings change from one year to a subsequent year, either one or two years later, depending on the data set. The measure we report in this Overview paper uses the variance of what is called the arc percent change, which is simply the percent change in earnings relative to the average in the two years (a measure commonly used in macroeconomic studies, such as Davis et al., 2006). Another common measure is the variance of log earnings differences, but this measure is more sensitive to the tails. But the individual papers report results for that measure as well and find little difference in estimated trends. No attempt is made to decompose the variance of earnings changes into permanent and transitory components; this is left for future work. Also, while the results shown in this Overview paper are based on volatility estimated with earnings directly, the individual papers also show results for volatility calculated using residuals from a regression of either the change in log earnings or the arc percent change on age

and age squared, also common in the literature. No difference in volatility trends is found when using such residuals. The baseline specification also trims the top and bottom one percent of the cross-sectional earnings distribution in each year to remove outliers, but sensitivity tests to trimming are conducted and, in fact, differences in the sizes of the tails of the earnings distribution in different data sets play a role in the analyses and explain some differences in volatility levels and trends, as summarized below and analyzed in detail in the papers. All papers work with a sample of men who worked in both years, but results are also obtained when men with zero earnings in one year are included and which therefore capture volatility in the movement into and out of employment. Results including nonworkers are summarized in Online Appendix A.

## IV. Results

### A. Baseline

Figure 1 shows our baseline results, using the samples and earnings variables listed in Table 1, for men working both periods. The PSID shows patterns mostly consistent with prior work, with rising volatility from the 1970s to the mid-1980s, then following a stable trend around significant fluctuations through about 2002, then rising in the period leading up to and including the Great Recession, and then falling post-Recession from 2012 to 2016. The last four years of PSID data are new to this project and show that volatility has declined back to its pre-Recession level in 2006, which was somewhat above its level in the mid-1980s.

The series for the other five data sets are often different from the PSID in level but not always in trend. The SIPP administrative data series, which starts in 1980, is higher in level than the PSID but follows a similar slight decline from 1982 through about 1999, but with

fluctuations over that period much milder in magnitude than for the PSID (perhaps the result of a larger sample size—see Table 1). It then also rises before the Great Recession and falls afterwards, although again not always of the same magnitude and at exactly the same time points as the PSID. The SIPP survey data have a lower level of volatility than the PSID (and much lower than the SIPP administrative data series) but has an approximately similar pattern in the first half of the period—a rise then fall from 1985 to 1999, but occasionally moving in opposite directions (e.g., 1988-1990). But the main difference with the SIPP survey is that it rises much less before and during the Great Recession than the PSID and the SIPP administrative data series.

The two CPS series shown—one for survey data and one for administrative data—are computed only on non-imputed observations (see the CPS paper). The series are lower in level than the PSID but their separate levels are quite close to one another, an important finding suggesting that any response error in the CPS survey is small enough to be ignored, at least for the purpose of earnings volatility measurement. Both CPS series only begin in 1996 and, over that period, both rise with the Recession and then fall afterwards, returning to their original 1996 levels by 2015, implying no net trend. This differs from the PSID, which is still somewhat higher by that date than it was in 1996. Finally, the LEHD, which also only starts in 1998, has the highest volatility level of all the series. In terms of trends, it has two countercyclical spikes which leaves its value in 2011 the same as its initial value in 1998, but declines from 2011 to 2016 and ends up below its 1998 level. Most other series also declined after 2011 but not to values below their 1998 levels.

Overall, the differences in volatility levels across the series are greater than their differences in trends. As we will describe below, we are able to greatly narrow the differences in levels

across the data sets. As for trends, the series are consistent with each other in many major respects. Prior to 1998, there is rough consistency across the three data sets we have for that period—essentially, no average trend in volatility from the mid-1980s to the 1998-2002 period. After 1998, the two CPS series, the SIPP administrative data series, and the PSID all show increases prior to and including the Great Recession, followed by declines. The magnitude of the upward trend is greatest for the PSID. However, the SIPP survey and the LEHD show little or no trend, and we will have some hypotheses for the differences with these two data sets below.

We devote a brief additional discussion to the post-1998 trends to explore these differences a bit further. There is always an issue in comparing growth rates of any short aggregate time series that has significant fluctuations without a formal statistical model because calculations of those rates can be highly sensitive to the chosen starting point. For example, earnings volatility growth in the PSID is much faster than in other data sets from 1998, but 1998 was a low point that clearly deviated negatively from its trend (see the PSID paper for a possible explanation for this deviation). Both SIPP series have a dip in 1998 and the two CPS series have a dip in 2000 which suggests not using them as a starting points. While not claiming any particular formal justification for our procedure, we address this issue by measuring volatility growth after 2002 relative to the 1992-2002 average for the PSID and the SIPP administrative data, and relative to the 1998-2002 average for the other data sets. These particular intervals roughly average over one complete cycle for each data set, as can be seen from Figure 1. The results, appearing in Appendix Figure 1, show, first, that four of the data sets have remarkably similar growth and decline patterns after 2002 relative to their initial averages, and end up with a net zero growth by the end. A fifth, the PSID, is an outlier and is mainly distinguished by a continued high level

after 2012 and a lack of decline in that post-Great-Recession period compared to the pattern exhibited by the other data sets, resulting in a net positive growth by the end (but the PSID is not much different than the others prior to 2012). The LEHD is the other outlier, showing very little cyclical growth and a stronger decline post-Recession than in the other data sets and ending with a small negative growth after about 2011. We conclude that average volatility growth rates in the 2000s and partway into the 2010s were quite similar for most of the data sets, with some exceptions. Combined with our finding of very little volatility trend in the three data sets covering the period from the mid-1980s to the 1998-2002 period, we also conclude that there is little evidence for significant trends upward or downward over the last 20-30 years.

B. Explaining the Differences

The difference across the six data series which we find to be most important in explaining their level and trend differences is related to differences in their cross-distributional distributions of earnings. Before we present those results, we briefly summarize the large number of other hypotheses we have explored which have little or no explanatory power for differences in volatility trends. The details of these investigations are in the individual papers; here we just summarize the findings.

For example, the restriction of the PSID to household heads is explored by estimating volatility trends in the CPS and SIPP survey, whose results above include non-heads, on heads only. Both data sets show lower levels of volatility for heads than for non-heads but trends are unaffected (Carr et al., 2022, Appendix Figure A.5; Ziliak et al., 2022, Appendix Figure S.9). For the differences in volatility trends for wage and salary earnings, both the SIPP and CPS permit the estimation of volatility excluding the self-employed, as required in the PSID. The SIPP

data show a lower level of volatility for wage and salary earners but no change in trend, while the CPS has similar findings (Carr et al., 2022, Appendix Figure A.6; Ziliak et al., 2022, Appendix Figure S.9). For the PSID exclusion of immigrants, the CPS paper finds no volatility trend differences for immigrants and non-immigrants (Ziliak et al., 2022, Appendix Figure S.8). And regarding our baseline regression specification, volatility trends are little affected by using regression residuals from log earnings equations instead of log earnings itself, or by using the log earnings difference instead of the arc percent change (Moffitt and Zhang, 2022, Appendix Figure 2; Carr et al., 2022, Appendix Figures A.1 and A.2; Ziliak et al., 2022, Appendix Figure S.3 and Figure S.6; McKinney and Abowd, 2022, Figures 1 and Appendix Figure B2). In addition, our trimming at the 1st and 99th percentile points of the cross-sectional earnings distributions has no effect on trends; no trimming at all produces more fluctuations in our estimated trends and trimming greater proportions of the tails produces fewer fluctuations, but in neither case are trends affected (Moffitt and Zhang, 2022, Appendix Figure 7; Carr et al., 2022, Appendix Figure A.3 and A.4); Ziliak et al., 2022, Appendix Figures S.4 and S.5; McKinney and Abowd, 2022, Figure 1).

The papers in the project note that it is important to trim at percentile points and not to use real dollar trims, as employed in some prior work using administrative data (Kopczuk et al. (2010), Guvenen et al. (2014), Bloom et al. (2017); Sabelhaus and Song (2009, 2010)). As noted by Carr and Wiemers (2021), using real dollar trims creates bias if either the tails of the distribution are changing in real dollar terms or if the trends in volatility are different in the tails. Online Appendix B to this Overview reports the results of our analysis of this issue with our different data series and shows that, in some data sets, real dollar trims sometimes reduce the upward trends in volatility and, in some cases, change an upward trend to a negative trend.

Attrition is more of a potential threat because it only occurs in survey data sets. Bias in trends can occur if those who are missing from a survey have differences in volatility, and attrition in our survey data sets range from 20 to 44 percent. For one of our data sets—the CPS—we have administrative data for those who attrite. CPS attrition takes place when a family interviewed in the first year has moved or is otherwise unavailable in the second year. The CPS analysis shows that those missing in the second year have much higher levels of volatility than those who are not missing but that trends are unaffected when they are included (Ziliak et al., 2022, Figure 4). In addition, in all three survey data sets--the CPS, SIPP, and PSID—we use standard inverse probability weighting to test for attrition bias by estimating the probability of attrition as a function of observables and then reweighting the volatility calculation on the non-attriter sample with the inverse of the predicted probabilities (Wooldridge, 2010). This eliminates bias under the selection-on-observables assumption. The results show virtually unchanged trends in volatility after this adjustment (Moffitt and Zhang, 2022, Appendix Figure 10; Carr et al., 2022, Appendix Figure A8; Ziliak et al., 2022, Figure 5). The results for the PSID, in fact, show a stronger upward trend in volatility after reweighting, consistent with prior studies of the PSID indicating that higher-volatility individuals are more likely to drop out in the first place (Fitzgerald et al., 1998).

In addition to attrition, a fraction of respondents in all surveys have missing data on specific variables ("item nonresponse") because of do-not-know responses and refusals-to-answer, from implausible values indicating response error, and for other reasons. Surveys typically impute new values for those missing responses and the statistical properties of those imputations have attracted a great deal of discussion in the literature (e.g., Andridge and Little, 2010). While imputation rates for earnings in the PSID are very low—3 to 4 percent—they are

very high in the CPS and SIPP, with recent rates of about 45 and 49 percent of all observations, respectively, and rising over time. Further, for both the CPS and SIPP, studies linking survey data to administrative earnings reports show that missingness is decidedly non-ignorable, with nonresponse higher in the tails of the earnings distribution (Bollinger et al., 2019; Chenevert et al., 2015). The non-ignorability is important because imputation methods for the CPS and SIPP made by the U.S. Census Bureau only use observables to impute values for the missing observations. For our project, while more sophisticated approaches to the problem are possible, we take simple approaches for the CPS and the SIPP. The CPS analysis shows that volatility levels and trends are very different when using the administrative earnings data for those whose earnings are imputed. Specifically, using imputed survey earnings show much higher levels of volatility as well as a rising trend (Ziliak et al., 2022, Figure 3). But, as noted previously, both the level and trends in volatility are virtually identical in survey and administrative data when using only non-imputed survey earnings.

The SIPP analysis cannot link administrative data directly to the survey data but it is apparent that imputation is a serious problem in the survey data. Methods of imputation and their coding have changed over time, which makes a truly consistent measure of imputation not possible with the SIPP survey. Carr et al. (2021, Figure 2) shows that the more serious problem is the presence of so-called "whole imputes," where the entire observation is imputed. Including those observations (which were excluded for the calculations shown in Figure 1) greatly raises the level of volatility as well as changing the flat trend to a positive one.

Differences in Cross-sectional Earnings Distributions. It has been noted in a number of prior studies that administrative data on earnings from Social Security and UI earnings records appear to have larger left tails of the earnings distribution than survey data sets (Kornfeld and

Bloom, 1999; Juhn and McCue, 2010; Celik et al., 2012; Spletzer, 2014; Abraham et al., 2013; Abowd et al., 2018). The most common hypothesis for this difference is that survey respondents often fail to report earnings from short, part-year jobs, especially if asked about earnings over a previous 12-month period (although there are also some earnings reports in survey data that do not show up in administrative data (Juhn and McCue, 2010; Abraham et al., 2013; Abowd and Stinson, 2013)). This difference is strongly exhibited in our six data series, as shown by the cumulative distribution functions in Figure 2, which shows the LEHD to have the largest left tail of earnings. The left tail of the SIPP administrative data is larger than that of the SIPP survey data and the left tail of the CPS administrative data is larger than that of the CPS survey data. The PSID has the smallest left tail. For example, about 25 percent of the LEHD observations have earnings less than $20,000 per year while only about 5 percent of PSID observations do, a large difference. These differences will almost surely cause the levels of volatility to be larger in administrative data because the literature has shown volatility to be higher at lower earnings levels than at higher ones. Figure 1 shows volatility levels indeed to be higher in administrative data than in survey data. But it can affect trends in volatility as well as levels if either those trends differ between the lower tail and the rest of the earnings distribution or if the size of the left tail is changing differently in the different data sets.

We explore this issue by adjusting all data series to the same cross-sectional distribution. Given the key role of the PSID in this literature, we benchmark the other five to it, which effectively means downweighting the left tails of the other five data sets' distributions. We do so by first assembling the minimum and maximum value of the PSID cross-sectional earnings distribution in each year (which are the 1$^{st}$ and 99$^{th}$ percentile points of the untrimmed distribution) and then computing ventile percentile points in each year between those year-by-

year minima and maxima. For each of the other five data series, we discard observations below the PSID minimum and above the PSID maximum in each year, and then use the PSID ventile points to compute a weighted average volatility, weighting the observations in each ventile range by .05. The results will reveal whether any differences in volatility levels and trends across the data series arise from differences in their cross-sectional earnings distributions rather than differences in volatility trends conditional on location in the cross-sectional distribution (at least within the PSID range).

Figure 3 shows the benchmarked volatility for the five series other than the PSID, as well as the PSID for comparison. The benchmarking has a dramatic impact on the data series differences, as can be seen by a comparison to Figure 1. The levels of volatility for the three administrative data sets (CPS administrative, SIPP administrative, and LEHD) are greatly lowered because of the trimming and/or downweighting of their large left tails. Now the levels of all data sets except the PSID are relatively tightly concentrated in a narrow range in the 1998-2015 period (they are all still quite a bit below the PSID level). The trends for the five also show similar time patterns in their overlapping years, rising from the mid-1990s through the peak of the Great Recession and declining thereafter, as in Figure 1, except that now the LEHD trend is a small net positive from the first year to the last instead of a small net negative. The changed trend in LEHD volatility arises because the PSID real 1$^{st}$ percentile point declines over time while the lower percentile points in the LEHD do not, leading to the inclusion of an increasing fraction of (high volatility) low LEHD earners over time (McKinney and Abowd, 2022, Figure 3 and Appendix Figure G1). Thus the underlying source of the change is a difference in trends in the left tails of the earnings distributions.

Finally, to compare trends after the 1990s, we can benchmark all six data sets to their initial average as discussed above but using the reweighted data series instead of the unweighted series. As shown in Appendix Figure 2, the LEHD series ends with a net small positive growth in volatility but which is close to zero, and the CPS survey follows an almost identical pattern. The SIPP survey follows a pattern similar to these two, over the years it is available. The CPS and SIPP administrative series do not fall as much—because of the downweighting of their left tails—and now follow trends almost identical to that of the PSID. We conclude that the majority of the large differences in volatility levels across the data series, and much of the small differences in trend, are explained by the differences in their cross-sectional earnings distributions.

## V. Summary and Conclusions

The project which this Overview summarizes has been narrowly focused on the sole question of how the gross earnings volatility of prime-age men has evolved over the last 40-50 years in the US and whether differing findings on this question across different data sets can be reconciled. The central finding is that, when put on a comparable basis, male earnings volatility in six survey and administrative data sets shows no sign of a major net increase or decrease since the late 1980s or early 1990s, although experiencing significant countercyclicality. There is some evidence that volatility increased from the 1970s to the 1980s but only from a subset of our data series which go back that far. Our findings should be regarded as a significant contribution to our understanding of the evolution of the US labor market.

One take-away from our study is that all data sets have their strengths and weaknesses, and acknowledging those characteristics and investigating how they may affect the results of a

particular study should be an important goal for work in many areas of economics. Reconciliation studies which attempt to resolve the differing findings from different data sets are relatively rare, possibly because they are messy and difficult. But our project demonstrates that progress can be made when careful and detailed attention is paid.

The narrowness of the exercise we have conducted also means that many interesting questions have not been addressed. Extensions to other demographic groups by age, gender, and marital status would be of interest. For age, the relationship between volatility and retirement has been little studied. For gender, while there has been a bit of attention to women's volatility (see Moffitt and Zhang (2018) for references and see the CPS paper in this volume), it is an understudied topic. On marital status, while the concept of the family as an insurance mechanism has received considerable attention, very little of that literature adopts the kind of dynamic framework that an incorporation of volatility requires. Decomposing male earnings volatility into its components of hours worked and wage rates would be of interest, a topic long studied in connection with business cycles (where volatility is greater in hours than in wages), trends in their separate volatility has not received much attention. More work has been done on trends in U.S. job mobility and its possible decline, but more models of the joint determination of earnings and job mobility would be of interest. Error components models exploring the dynamic structure of earnings evolution could yield additional insights on time-series trends in all those components, building on the already extensive literature on decomposition of dynamic permanent and transitory components. And an important finding of our work is that earnings volatility is quite different in different parts of the earnings distribution, particularly at low earnings, which suggests that future work address that source of heterogeneity as well (see Hardy and Ziliak, 2014, for one such past study).

Acknowledgements and Disclosures

The authors would like to thank Joe Altonji, the participants in a Summer 2019 meeting sponsored by the Washington Center for Equitable Growth, the participants of the Summer 2020 NBER Conference in Research and Wealth workshop, the participants of the June 2021 Summer meetings of the Econometric Society, and the participants of a seminar at the Chicago Federal Reserve Bank for comments. This draft also incorporates suggestions made by the JBES editorial staff and referees on individual papers. No funding was received for the work on this paper. The authors report there are no competing interests to declare.

**References**


Abowd, John M., and Martha H. Stinson. 2013. Estimating Measurement Error in Annual Job Earnings: A Comparison of Survey and Administrative Data. *Review of Economics and Statistics* 95 (5):1451–67.

Abowd, John M.; Kevin L. McKinney; and Nellie L. Zhao. 2018. "Earnings Inequality and Mobility Trends in the United States: Nationally Representative Estimates from Longitudinally Linked Employer-Employee Data." Journal of Labor Economics 36(S1): S183-S300.

Abraham, Katherine G., John C. Haltiwanger, Kristin Sandusky, and James Spletzer. 2013. Exploring Differences in Employment between Household and Establishment Data. *Journal of Labor Economics* 31 (S1): 129-72.

Andridge, Rebecca R. and Roderick J. A. Little. 2010. A Review of Hot-Deck Imputation for Survey Non-Response. *International Statistical Review* 78(1): 40-64.

Bania, Neil, and Laura Leete. 2009. Monthly Household Income Volatility in the U.S., 1991/92 vs. 2002/03. *Economics Bulletin* 29 (3): 2100–12.

Bloom, Nicholas; Faith Guvenen; Luigi Pistaferri; John Sabelhaus; Sergio Selgado; and Jae Song. 2017. The Great Micro Moderation. Working Paper.

Blundell, Richard; Luigi Pistaferri; and Ian Preston. 2008. Consumption Inequality and Partial Insurance. *American Economic Review* 98(5): 1887-1921.

Bollinger, Christopher; Barry T. Hirsch; Charles M. Hokayem; and James P. Ziliak. 2019. Trouble in the Tails? What We Know about Earnings Nonresponse Thirty Years after Lillard, Smith, and Welch. *Journal of Political Economy* 127(5): 2143-85.

Braxton, J. Carter; Kyle Herkenhoff; Jonathan Rothbaum; and Lawrence Schmidt. 2021. Changing Income Risk across the Skill Distribution: Evidence from a Generalized Kalman Filter. Working Paper 29567. Cambridge: National Bureau of Economic Research.

Carneiro, Pedro and Rita Ginja. 2016. Partial Insurance and Investments in Children. *Economic Journal* 126(October): F66-F95.

Carr, Michael D. and Emily E. Wiemers. 2018. New Evidence on Earnings Volatility in Survey and Administrative Data. *American Economic Association Papers and Proceedings* 108 (May): 287-91.

Carr, Michael D. and Emily E. Wiermers. 2021. The Role of Low Earnings in Differing Trends in Earnings Volatility. *Economics Letters* 199(C).

Carr. Michael D.; Robert A. Moffitt; and Emily E. Wiemers. 2022. Reconciling Trends in Male Earnings Volatility: Evidence from the SIPP Survey and Administrative Data. *Journal of Business and Economic Statistics*, this issue.

Carroll, Christopher D. 1997. Buffer-Stock Saving and the Life Cycle/Permanent Income Hypothesis. *Quarterly Journal of Economics* 112(1): 1-55.

Celik, Sule, Chinhui Juhn, Kristin McCue, and Jesse Thompson. 2012. Recent Trends in Earnings Volatility: Evidence from Survey and Administrative data. B.E. *Journal of Economic Analysis and Policy* 12 (2).

Chenevert, Rebecca L.; Mark A. Klee; and Kelly R. Wilkin. 2015. Do Imputed Earnings Earn Their Keep? Evaluating SIPP Earnings and Nonresponse with Administrative Records. Washington: U.S. Census Bureau.

Cunha, Flavio; James Heckman; and Salvador Navarro. 2005. Separating Uncertainty from Heterogeneity in Life Cycle Earnings. *Oxford Economic Papers* 57(2): 191-261.

Dahl, Molly, Thomas DeLeire, and Jonathan A. Schwabish. 2008. Trends in Earnings Variability over the Last 20 Years. Washington: Congressional Budget Office.

Dahl, Molly, Thomas DeLeire, and Jonathan A. Schwabish. 2011. Estimates of Year-to-Year Volatility in Earnings and in Household Incomes from Administrative, Survey, and Matched Data. *Journal of Human Resources* 46 (4): 750–74.

Davis, Steven J.; John Haltiwanger; Ron Jarmin; and Javier Miranda. 2006. Volatility and Dispersion in Firm Growth Rates: Public Traded Vs Privately Held Firms. *NBER Macroeconomics Annual* 21:107-158.

Davis, Steven J. and James A. Kahn. 2008. Interpreting the Great Moderation: Changes in the Volatility of Economic Activity at the Macro and Micro Levels. *Journal of Economic Perspectives* 22(4): 155-80.

DeBacker, Jason Matthew, Bradley T. Heim, Vasia Panousi, Shanthi Ramnath, and Ivan Vidangos. 2013. Rising Inequality: Transitory or Permanent? New Evidence from a U.S. Panel of Household Income 1987-2006. *Brookings Papers on Economic Activity* 2013 (4): 67-142.

Dynan, Karen, Douglas Elmendorf, and Daniel Sichel. 2012. The Evolution of Household Income Volatility. *The B.E. Journal of Economic Analysis & Policy* 12 (2): 1-42.

Dynarski, Susan and Jonathan Gruber. 1997. Can Families Smooth Variable Earnings? *Brookings Papers on Economic Activity* 1997 (1):229–303.

Fitzgerald, John, Peter Gottschalk, and Robert A. Moffitt. 1998. An Analysis of Sample Attrition in Panel Data: The Michigan Panel Study of Income Dynamics. *Journal of Human Resources* 33 (2): 251-99.

Friedman, Milton. 1957. The Permanent Income Hypothesis. In *A Theory of the Consumption Function*, 20-37. Elizabeth: Princeton University Press.

Ganong, Peter and Pascal Noel. 2019. Consumer Spending During Unemployment: Positive and Normal Implications. *American Economic Review* 109(7): 2383-2424.

Gottschalk, Peter, and Robert Moffitt. 1994. The Growth of Earnings Instability in the U. S. Labor Market. *Brookings Papers on Economic Activity* 1994 (2): 217–72.

Gourinchas, Pierre-Olivier and Jonathan A. Parker. 2002. Consumption over the Life Cycle." *Econometrica* 70(1): 47-89.

Groves, Robert M.; Don A. Dillman; John L. Eltinge; and Roderick J.A. Little, eds. 2002. Survey Nonresponse. New York: Wiley.

Guvenen, Fatih, Serdar Ozkan, and Jae Song. 2014. The Nature of Countercyclical Income Risk. *Journal of Political Economy* 122(3): 621-60.

Haider, Steven. 2001. Earnings Instability and Earnings Inequality of Males in the United States, 1967-1991. *Journal of Labor Economics* 19 (October): 799-836.

Hall, Robert E, and Frederic S Mishkin. 1982. The Sensitivity of Consumption to Transitory Income: Estimates from Panel Data on Households. *Econometrica* 50 (2): 461–81.

Hardy, Bradley, and James P. Ziliak. 2014. Decomposing trends in income volatility: The 'wild ride' at the top and bottom. *Economic Inquiry* 52 (1): 459–76.

Heathcote, Jonathan; Kjetil Storesletten; and Giovanni L. Violante. 2010. The Macroeconomic Implications of Rising Wage Inequality in the United States. *Journal of Political Economy* 118 (August): 681-722.

Hyslop, Dean R. 2001. Rising U.S. Earnings Inequality and Family Labor Supply: The Covariance Structure of Intrafamily Earnings. *American Economic Review* 91 (4): 755–77

Jensen, Shane T, and Stephen H Shore. 2015. Changes in the Distribution of Earnings Volatility. *Journal of Human Resources* 50 (3): 811–36.

Juhn, Chinhui and Kristin McCue. 2010. Comparing Measures of Earnings Instability Based on Survey and Administrative Reports. *US Bureau Census for Economics Studies Paper* CES-WP-10-15.

Katz, Lawrence F. and David H. Autor, 1999. Changes in the Wage Structure and Earnings Inequality. In *Handbook of Labor Economics, Vol. 3A*, eds. O. Ashenfelter and D. Card. Amsterdam and New York: Elsevier North-Holland.

Keys, B. 2008. Trends in Income and Consumption Volatility, 1970-2000. *In Income Volatility and Food Assistance in the United States*, ed. D. Jolliffe and J. Ziliak. Kalamazoo, Mich: W.E. Upjohn Institute.

Koo, K.H. 2016. The Evolution of Earnings Volatility During and After the Great Recession. *Industrial Relations* 55: 705-732.

Kopczuk, Wojciech, Emmanuel Saez, and Jae Song. 2010. Earnings Inequality and Mobility in the United States: Evidence from Social Security Data since 1937. *Quarterly Journal of Economics* 125 (Feb): 91-128

Kornfeld, Robert and Howard S. Bloom. 1999. Measuring Program Impacts on Earnings and Employment: Do Unemployment Insurance Wage Reports from Employers Agree with Surveys of Individuals? *Journal of Labor Economics* 17: 168-197.

Levhari, David and Yoram Weiss. 1974. The Effect of Risk on the Investment in Human Capital. *American Economic Review* 64(6): 950-963.

McConnell, Margaret M. and Gabriel Perez-Quiros. 2000. Output Fluctuations in the United States: What Has Changed since the Early 1980s? *American Economic Review* 90(5): 1464-76.

McKinney, Kevin L. and John M. Abowd. 2022. Male Earnings Volatility in LEHD Before, During, and After the Great Recession. *Journal of Business and Economic Statistics*, this issue.

Moffitt, Robert A. and Peter Gottschalk. 2012. Trends in the Transitory Variance of Male Earnings: Methods and Evidence. *Journal of Human Resources* 47(1): 204-36.

Moffitt, Robert A. and Sisi Zhang. 2018. Income Volatility and the PSID: Past Research and New Results. *American Economic Association Papers and Proceedings* 108 (May): 277-80.

Moffitt, Robert A. and Sisi Zhang. 2022. Estimating Trends in Male Earnings Volatility with the Panel Study of Income Dynamics. *Journal of Business and Economic Statistics*, this issue.

National Research Council. 2013. Nonresponse in Social Science Surveys. Washington: National Academy Press.

Sabelhaus, John and Jae Song. 2009. Earnings Volatility Across Groups and Time. *National Tax Journal* 62 (June): 347-364.

Sabelhaus, John, and Jae Song. 2010. The Great Moderation in Micro Labor Earnings. *Journal of Monetary Economics* 57 (4): 391–403.

Shin, Donggyun, and Gary Solon. 2011. Trends in Men's Earnings Volatility: What Does the Panel Study of Income Dynamics Show? *Journal of Public Economics* 95 (7–8). Elsevier B.V.:973–82.

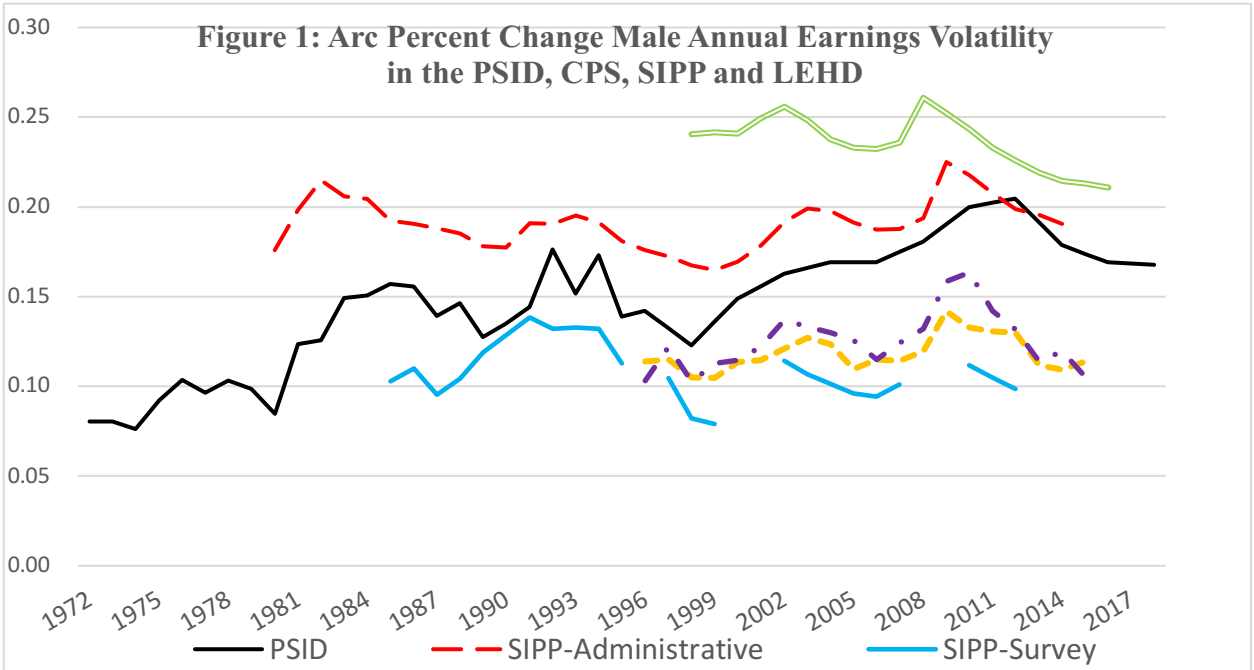
Spletzer, James R. 2014. Inequality Statistics from the LEHD. Working Paper. Washington: U.S. Bureau of the Census.

Taubman, Paul and Michael Wachter. 1986. Segmented Labor Markets. In *Handbook of Labor Economics*, Vol.2, eds. Orley Ashenfelter and Richard Layard. New York and Amsterdam: Elsevier North-Holland.

Wooldridge, Jeffrey M. 2010. *Econometric Analysis of Cross Section and Panel Data*. Cambridge: MIT Press.

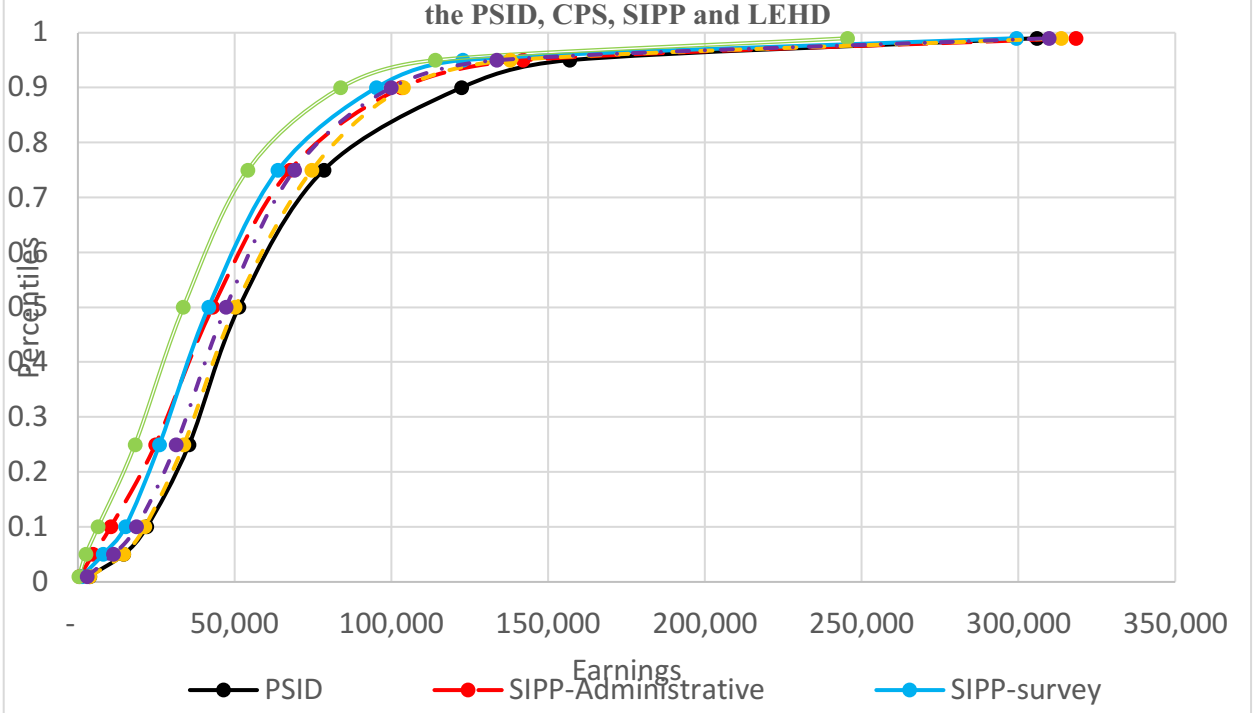
Ziliak, James P., Bradley Hardy, and Christopher Bollinger. 2011. Earnings Volatility in America: Evidence from Matched CPS. *Labour Economics* 18(6): 742-54.

Ziliak, James P.; Charles Hokayem; and Christopher R. Bollinger. 2022. Trends in Earnings Volatility using Linked Administrative and Survey Data. *Journal of Business and Economic Statistics*, this issue.

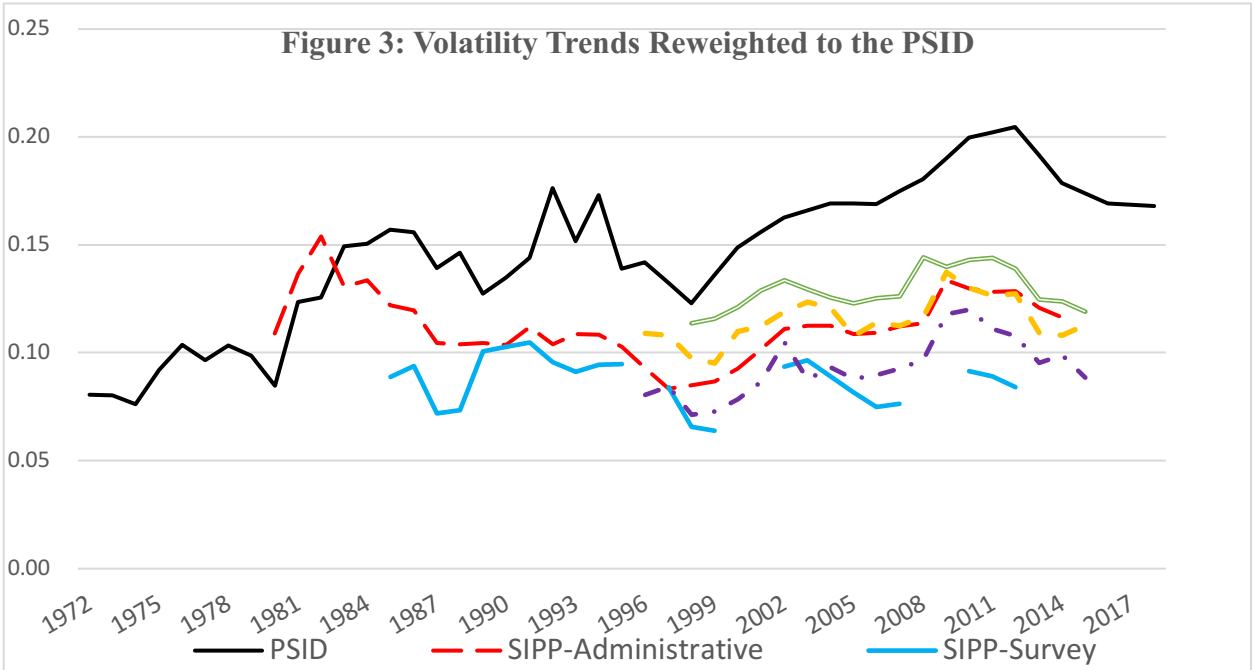



*Notes*: Sample consists of men with positive earnings in both time periods with a 1 percent top and bottom trim of the cross-sectional earnings distribution in each year. Moffitt and Zhang (2021, Appendix Figure 2). CPS: Ziliak et al. (2021, Appendix Figure S.6). SIPP: Carr et al.

**Figure 2: Cross-Sectional Cumulative Distribution Function of Earnings in the PSID, CPS, SIPP and LEHD**




*Notes*: The PSID, CPS-Survey, CPS-Administrative use 2000 data, and SIPP-Administrative,




*Notes*: Sample consists of men with positive earnings in both time periods with a 1 percent top and bottom trim of the cross-sectional earnings distribution in each year. Cross-sectional distribution of all data sets except the PSID reweighted to match the PSID cross-sectional