NBER WORKING PAPER SERIES

RECONCILING TRENDS IN VOLATILITY:
EVIDENCE FROM THE SIPP SURVEY AND ADMINISTRATIVE DATA

Michael D. Carr
Robert A. Moffitt
Emily E. Wiemers

Working Paper 27672
http://www.nber.org/papers/w27672

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
August 2020

This paper was prepared for presentation at the January, 2019 meetings of the Econometric Society in Atlanta. Carr and Wiemers gratefully acknowledge funding from the Russell Sage Foundation for this project (award #83-15-09). The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Reconciling Trends in Volatility: Evidence from the SIPP Survey and Administrative Data
Michael D. Carr, Robert A. Moffitt, and Emily E. Wiemers
NBER Working Paper No. 27672
August 2020
JEL No. C23,J3

ABSTRACT

As part of a set of papers using the same methods and sample selection criteria to estimate trends in male earnings volatility across a number of survey and administrative datasets, we conduct a new investigation of trends in male earnings volatility from the 1980s to 2014 using data from the Survey of Income and Program Participation (SIPP) survey and the SIPP Gold Standard File (SIPP GSF), which links the SIPP survey to administrative data on earnings. We find that the level of volatility is higher in the SIPP GSF than in the SIPP survey but that the trends are similar. Specifically, over the period where the datasets overlap between 1984 and 2012, volatility in the SIPP survey declines slightly while volatility in the SIPP GSF increases slightly but the differences are small in magnitude. Because the density of low earnings differs considerably across datasets, and volatility may vary across the earnings distribution, we estimate trends in volatility in the SIPP survey and SIPP GSF where we hold the earnings distribution fixed to resemble that in the Panel Study of Income Dynamics (PSID). We find that differences in the underlying earnings distribution explains almost all of the difference in the level of volatility between the SIPP survey and SIPP GSF and it somewhat reduces the small differences in trends.

Michael D. Carr
Department of Economics
University of Massachusetts Boston
100 Morrissey Blvd.
Boston, MA 02125
michael.carr@umb.edu

Robert A. Moffitt
Department of Economics
Johns Hopkins University
3400 North Charles Street
Baltimore, MD 21218
and IZA
and also NBER
moffitt@jhu.edu

Emily E. Wiemers
Department of Public Administration
and International Affairs
Syracuse University
Syracuse, NY 13244
eewiemer@syr.edu

1 Introduction

A substantial literature now exists studying trends in earnings instability over time in the United States. An early example in this literature is Gottschalk et al. (1994), which estimated an error components model of earnings in the Panel Study on Income Dynamics (PSID), and found that the transitory variance of earnings increased from the 1970s to the 1980s at the same time that permanent earnings inequality rose. The paper concluded that approximately half of the increase in earnings inequality over the period came from increasing transitory earnings variances and half to a widening of the distribution of permanent earnings. Extensions of this early work reached similar conclusions.[1]

An alternative to estimating an error components model of earnings on panel data is to use the distribution of short-run earnings changes as an estimate of the combined effect of transitory and permanent earnings variation (Shin and Solon, 2011). This latter measure is often referred to as earnings volatility and is the approach used in this paper. Estimates of volatility using the PSID imply similar conclusions about trends in instability as those reached by estimating transitory variances from error components models(Carr and Wiemers, 2018, Moffitt and Zhang, 2018a, Shin and Solon, 2011). While studies of volatility differ somewhat in the precise pattern of trends over time, most PSID studies have found that male earnings volatility stopped rising sometime in the 1980s and either declined or flattened out until the mid 1990s, but then increased again after the late 1990s through the Great Recession.

In recent years, volatility has been estimated using a wider variety of administrative and survey data sources (beyond the PSID) and a wider variety of methods.[2] Despite differences in methods and sample definitions across analyses, there is broad agreement across both survey and administrative data that volatility declined from the mid 1980s through 2000. Differences between studies emerge after 2000 and especially during the Great Recession. The level of

[1] See Moffitt and Zhang (2018b), Table 2, for a review of all published studies up to Spring 2018 and their findings.
[2] See Moffitt and Zhang (2018b), Table 3, for a list of studies and their findings up through Spring 2018.

volatility, however, is consistently higher in administrative data than survey data.

Among the analyses that use other sources of survey data to estimate volatility on individual earnings, Celik et al. (2012), which used survey data from the Survey of Income and Program Participation (SIPP) from 1984 to 2006, found that male earnings volatility declined over the entire period.[3] Ziliak, Hardy, and Bollinger (2011) used the panel component of the Current Population Survey (CPS) and found that volatility increased from the 1970s through the mid-1980s and stabilized through 2009. The stability in the 2000s in the CPS data is inconsistent with the PSID, where volatility rises during the 2000s (Carr and Wiemers, 2018, Moffitt and Zhang, 2018a).[4] Celik et al. (2012) also estimated volatility of male earnings in the CPS and found patterns similar to those of Ziliak, Hardy, and Bollinger (2011) except for an upturn in volatility in the mid-2000s.

A number of studies have used administrative data to examine trends in volatility in individual earnings. Sabelhaus and Song (2009, 2010) used administrative earnings data from the Social Security Master Earnings File (MEF) from 1980 to 2005 and found smoothly declining earnings volatility through the entire period on a pooled sample of men and women.[5] Guvenen, Ozkan, and Song (2014) also use earnings data from the MEF, and found slight declines in male earnings volatility between 1980 and 2011. Dahl, DeLeire, and Schwabish (2011, 2008) used both the SIPP matched to earnings data from the Detailed Earnings Records (DER) and the Continuous Work History Sample maintained by the Social Security Administration. The former contains the same universe of workers as the MEF, while the latter contains a 1 percent sample of issued Social Security numbers. They found declining earnings volatility from 1984 to 2005 when combining men and women, and slight declines in volatility for men when estimated separately.

[3]Bania and Leete (2009) found rising volatility in the SIPP from 1991 to 2003 but only examined intrayear volatility of household income.

[4]Ziliak, Hardy, and Bollinger (2011) replicates the methods from Shin and Solon (2011) and Moffitt and Gottschalk (2012) on their sample from the CPS and these methods show larger declines in earnings volatility for men than their preferred method.

[5]Dynan, Elmendorf, and Sichel (2012) found declining female earnings volatility in the PSID.

Carr and Wiemers (2018), using similar, though not identical, DER earnings data matched to the SIPP in the SIPP Gold Standard File (SIPP GSF) for 1980 through 2011, found rising volatility in the early 1980s, declining volatility from 1985 through 2000, and rising in the mid-2000s through the Great Recession. Celik et al. (2012) estimated volatility using administrative earnings in the Longitudinal Employment and Household Dynamics (LEHD) data, which are drawn from Unemployment Insurance records, and found no trend from 1992-2008 in the 12 states covered by the LEHD over the period. Finally, DeBacker et al. (2013), using tax records combined with W-2 data, also found no trend in gross male earnings volatility from 1987 to 2009.[6][7]

The differences in trends that emerge in the latter years between administrative data and survey data, respectively, are often taken to suggest that the survey data are in error, either because of respondent misreporting or other data quality problems in surveys such as nonresponse or attrition. But it is not clear whether the differences are a result of true discrepancies in trends across different sources of data or from the use of different samples and methods. Very few studies have attempted to determine whether differences are the result of different samples and methods or differences in the underlying data. The exceptions are Dahl, DeLeire, and Schwabish (2011), Celik et al. (2012), and Carr and Wiemers (2018). Dahl, DeLeire, and Schwabish (2011) compared estimates of volatility in household income across the SIPP survey, SIPP-matched administrative earnings, and administrative earnings alone. While they found that the data sources are capable of producing similar trends, they highlight that using imputed values in the SIPP produces an upward bias in estimated volatility after 1998. Celik et al. (2012) compares trends in volatility for individual earnings in the PSID, the CPS, the SIPP and the LEHD using the same sample selection criteria and methods across datasets. They found quite different trends across

[6] The authors also estimated an error components model and found no trend in the transitory variance.

[7] There are also studies that use administrative earnings data to estimate volatility at the household level. Hryshko, Juhn, and McCue (2017) used the SIPP GSF to examine trends in the transitory variance of earnings of married couples from 1980 to 2009 and also found that the variance declined through 2000 but then rose through 2009, though the sample differs considerably from Carr and Wiemers (2018). Dahl, DeLeire, and Schwabish (2011) use a slightly different version of the SIPP linked administrative data to estimate household level volatility for 1985 to 2005, and found small overall declines.

datasets even when using the same sample definitions and measures and were not able to identify a cause of the remaining difference. Carr and Wiemers (2018) compared trends in the PSID to those in the SIPP GSF using measures comparable to those typically used in the literature and similar samples on both datasets. They found that while the level of volatility was higher in administrative earnings, trends were similar, finding a notable increase in volatility from 2000 to 2011 which had not been present in other studies. But more work is needed to conduct comparable estimation across different data sets and to attempt to reconcile differences in findings, which is the primary goal of this paper and the others in this project.[8]

In this paper, we compare estimates of earnings volatility for working age men from SIPP survey data and the SIPP GSF, which links SIPP survey respondents to their administrative earnings records from the DER which is co-maintained by the Social Security Administration and the Internal Revenue Service. Both datasets draw their respective samples from the same universe of individuals – participants in the SIPP survey – and allow us to apply consistent sample definitions and methods across the two sources of earnings. We show that the trends are broadly similar but there are large level differences in volatility between the two data sources. We find that although volatility is consistently higher in the SIPP GSF, in both data sources we find a generally flat trend from the 1980s to 2014 with volatility increasing modestly in the SIPP GSF and declining modestly in the SIPP survey data. In neither data source do we find evidence of substantially declining volatility in our baseline estimates. We then investigate the extent to which changing assumptions about the treatment of imputed earnings in the SIPP survey data and the treatment of low earnings in both the SIPP survey data and the SIPP GSF affect the estimated trends in volatility. We show that including individuals with imputed earnings components in the SIPP has little impact on either the trend or level of volatility, as long as whole case imputations are excluded. Trimming a larger percent of low earners reduces the level of volatility

[8] There are three other studies in this project: McKinney and Abowd (2020), Moffitt and Zhang (2020), and Ziliak, Hokayem, and Bollinger (2020).

substantially, especially in the SIPP GSF, but does not affect the trend. Finally, we investigate the role of differences in the earnings distribution across datasets by comparing the trends in volatility in the SIPP and SIPP GSF when we reweight the earnings distribution in both datasets to resemble that of the PSID. We find that forcing the two SIPP datasets to have the same underlying earnings distribution virtually eliminates the differences in the level of volatility, and also reduces the small differences in trends that exist. We also compare these estimates of volatility, where the earnings distribution is weighted to resemble the PSID, to a baseline PSID volatility estimate. We show that forcing the SIPP GSF to more closely resemble the PSID lowers volatility, and also brings the trend through time more in line with PSID, though differences in trend are not eliminated, while in the SIPP survey data volatility falls but the trend is unaffected. This implies that, at the same level of inequality, volatility is higher in the PSID than in either the SIPP or the SIPP GSF and it rises faster in the period after the late 1990s.

After a brief initial section on our measures of volatility below, we turn to a detailed discussion of the SIPP survey and its complex earnings measures and of the SIPP GSF data and issues relating to the comparability of the two. We then present our main results obtained from a initial series of sample and specification decisions, followed by sensitivity tests to determine the robustness of our results to our initial decisions and to a number of other threats to the initial findings. Finally, we explore in some detail whether differences between the cross-sectional distributions of earnings in our two data sets can explain their differences in levels and trends of volatility, as well as comparing each of them to volatility trends in the PSID.

2 Measures, Data, and Samples

2.1 Measures

We focus here on two simple measures of earnings volatility, both used extensively in the literature. The first is the variance of changes in log earnings over a short time horizon:

$$\text{Var}(y_{it} - y_{it-\tau}) \tag{1}$$

where y_{it} is log annual earnings of individual i at time t. In what follows, we show the results of estimating volatility net of life-cycle effects by reporting the variance of the residuals of a regression of $(y_{it} - y_{it-\tau})$ on a quadratic in age, estimated separately by year, as is standard in the literature starting with Shin and Solon (2011), with $\tau = 1$. We use the variance of earnings changes, rather than the standard deviation, because variances are easily decomposable.[9]

Our second measure is volatility estimated using the arc change, given in Equation 2 (Dahl, DeLeire, and Schwabish, 2011, 2008, Ziliak, Hardy, and Bollinger, 2011).

$$\text{Var}\left\{ \frac{y_{it} - y_{i,t-\tau}}{\frac{|y_{it}| + |y_{i,t-\tau}|}{2}} \right\} \tag{2}$$

An advantage of the arc change over the log difference method in Equation 1 is that it is bounded between -2 and 2, which reduces the impact of particularly large earnings changes on volatility. It also allows for the inclusion of individuals with zero earnings in one of two periods in a straightforward manner, though we exclude these individuals from our sample and work with samples of men with positive earnings in both periods, in line with the majority of the literature. As with Equation 1, we regress the arc change on a quadratic in age separately by year and calculate the variance of the changes from the residuals with $\tau = 1$. We examine the sensitivity of

[9] This measure of volatility is identical to the variance of transitory earnings if individual-specific permanent earnings are age and time invariant. Otherwise, the variance of shocks to the permanent component will be included.

our results to the use of age adjustments.

2.2 SIPP Survey

2.2.1 Data

The SIPP is a nationally representative sample of the civilian non-institutionalized population of the U.S. that began in 1984. There were 16 SIPP panels between 1984 and 2008, with each panel lasting between two and five years.[10] Within panels the SIPP is longitudinal, but each panel draws a new nationally representative sample of 14,000 to 52,000 households. SIPP panels after 1990 include a small oversample of low-income geographic areas that increases the number of households in or near poverty by 15% to 20% over what would be observed otherwise. During the 1980s and 1990s, most SIPP panels covered two years, with overlapping windows. For example, the 1984 panel started in 1983 and ended in 1986, while the 1985 panel started in 1985 and ended in 1987, but the samples in the 1984 and 1985 panels are different for overlapping years. Beginning with the 1993 panel, the overlapping samples were dropped, and panel lengths varied. The 1993 panel lasted three complete years. The 2008 panel provides four complete years of data, though it lasted for almost six years. We combine all waves of all panels, giving us data from 1984 to 2012, with a few missing years for reasons we discuss below.

SIPP interviews are conducted every four months and are retrospective, covering the previous four calendar months. Respondents report about income and program participation for each month of the previous four calendar months. SIPP households are separated into four rotation groups. At the beginning of the panel each rotation group starts in one month intervals. This design implies that sample sizes are smaller in the first and last four months of each panel as rotation groups rotate in and out of the sample.

[10]The SIPP was substantially redesigned in 2014 and we exclude the 2014 panel. The description that follows applies to SIPP panels through 2008.

2.2.2 Monthly Employment and Earnings Variables

Individuals are asked about their employment status in each month in the following sequence. They are first asked whether they worked at all during the previous four months. Then they are asked to report their employment status for each month. Individuals may have zero earnings in a month if they report either that they did not work at all during the four-month period or if they report being without a job for an entire month.

For respondents who report working, the SIPP asks respondents to report a maximum of five components of earnings for each month: wage and salary income from employment for up to two employers, business earnings from up to two businesses, and earnings from casual work. The SIPP also creates a separate total monthly earnings variable which is the sum of the five earnings components. From the monthly earnings measures, we construct a measure of annual earnings summed over all components for two successive years, allowing us to compute volatility measures.

2.2.3 Imputation

As in many other surveys, imputation of earnings because of item nonresponse (don't knows, refusals to answer, implausible values) is a non-trivial issue in the SIPP. In a report comparing imputation rates across surveys, Czaka and Denmead (2008) found that in the 2001 SIPP panel, 49.5% of those with wage and salary income in the SIPP had imputed values in at least one earnings component, the highest of any major survey. As a percent of total wage and salary income, 29% in the SIPP was the result of imputations (an even higher percent, 39.5%, of self-employment income in the SIPP is imputed) (Czaka and Denmead, 2008). Imputation rates have also changed over time in the SIPP. In the 1984 SIPP, only 8.8% of wage and salary income was imputed but this increased to 17.7% in 1993, 20.5% in 1997, and 24.9% in 2002 (National Research Council, 2009). Dahl, DeLeire, and Schwabish (2011) show that imputations in the

SIPP survey tend to bias estimates of volatility in household income upwards in the 1993, 1996, and 2001 panels, but less so in earlier panels. In our sample, imputation rates for wage and salary income fall in the 2004 and 2008 panels to about half the level in the 1996 and 2001 panels. Because of these high rates of nonresponse and imputation and because previous work using the SIPP survey has not always clearly articulated how imputations were handled, we devote considerable attention to the issue.

The first step is identifying which earnings values are imputed, which is complex in the SIPP. Imputations in the SIPP are flagged in two ways: with flags denoting whether an entire case has been imputed (meaning every variable except for those used for matching to a donor via a hot deck procedure) or whether just the labor force characteristics have been imputed, as well as flags indicating imputation of each of the five earnings components. It is necessary to use both types of imputations flags because individuals whose whole case is imputed or who have imputed labor force variables are given the imputation flags of the donor case for individual earnings components.

The method of flagging imputations has varied somewhat over time and we attempt to apply a consistent set of rules for excluding imputations. Most significant for our analysis is the way in which the survey deals with non-response to the question about whether an individual has had a job in the four-month reference period.[11] In SIPP panels prior to 1996, individuals who did not report whether they had a job in the reference period were treated as whole case imputations but, starting in the 1996 SIPP panel, only their job/business and labor force participation characteristics were imputed if not reporting that variable. Consistent treatment of labor force status over time requires that different imputation flags are used in different periods.[12] In contrast to the imputation flags on labor force status, the imputation flags for each earnings component

[11] The relevant questions are SC1000 in pre-1996 SIPP panels and EPDJBTHN in panels since 1996.

[12] The process of whole case imputation is referred to as "Type Z" and the process of partial imputation of labor force characteristics is referred to as "Little Type Z". The documentation is relatively sparse for the imputation processes in the SIPP prior to 1996 but the authors have used explanations in Pennell (2003) and discussion with Census Bureau staff as a basis for decisions about the treatment of imputations.

have been applied consistently over time. The SIPP-generated total earnings variable referred to above is imputed if any of the five components is missing or imputed, implying that the SIPP-provided sum of earnings need not equal the sum of earnings from its non-imputed components. Throughout the panels, there is no imputation flag on the SIPP-generated total monthly earnings variable.

Our approach to these issues is to minimize the use of imputed earnings data and to create a series of non-imputed earnings defined consistently over time. Imputations in the SIPP are implemented with a combination of the well-known hot deck method combined with other methods. Hot deck imputation has been shown to be problematic, in part because imputations are based on observables yet, as Bollinger et al. (2019) have shown for the CPS, nonresponse cannot be treated as missing at random. In particular, nonresponse is more likely among low and high earners conditional on observables. The evidence from the 2008 SIPP panel suggests that a similar pattern exists among earners in the SIPP (Chenevert, Klee, and Wilkin, 2016). This is the reason for minimizing the use of imputed earnings. However, excluding imputed variables may also bias estimates related to inequality (Bollinger et al. (2019)). In comparing SIPP-reported earnings and administrative earnings from the DER at the level of person-jobs, Abowd and Stinson (2013) show that conclusions about variance components of earnings are comparable between the SIPP-reported earnings and DER earnings except for the subsample of SIPP person-jobs where at least one year of SIPP earnings contained a Census Bureau imputation.[13] In these cases, the SIPP-reported earnings are less reliable than the DER. Because we are interested in comparing the variance of earnings changes in the SIPP survey and SIPP GSF, we draw on Abowd and Stinson (2013) and construct our sample using non-imputed earnings, which is consistent with the preferred estimates on the SIPP survey in both Celik et al. (2012) and Dahl, DeLeire, and Schwabish (2011). Further, as described next, we use the individual components of

[13] Abraham et al. (2013) document mismatches in employment status between administrative earnings histories and the CPS which seem to be concentrated among lower earnings.

earnings rather than the SIPP-constructed total earnings variable because the former has consistently-defined imputation flags over time and the latter has no imputation flags.

2.2.4 Creating Annual Earnings

To construct our baseline sample of men age 25 to 59 with positive earnings in two consecutive years, we first eliminate any observation with a whole-case imputation or for whom current work status is imputed. On the remaining sample, we then create monthly earnings by summing each of the five earnings components which is not imputed. Self-employment earnings are included among the components that we sum to be consistent with the earnings measure in the SIPP GSF (see below).[14] Excluding imputed earnings from the sum necessarily means our earnings total will be incomplete, so we also conduct a sensitivity test using the SIPP-constructed total monthly earnings, which contains imputed components. If an individual reports not working, they are assigned monthly earnings of zero. To construct two years of annual earnings for our volatility measure, we first select all men who have a valid non-imputed earnings (or a valid zero) in each month over a two-year period January 1 of the first year to December 31 of the second year, and then further select those who have at least one valid non-imputed earnings component in one month of each year. We then construct annual earnings by summing monthly earnings over the months in each calendar year.

2.2.5 Sample Loss from Rotation Groups and Attrition, and Final Sample

The construction of our baseline sample entails sample loss other than the exclusion of observations with only imputed earnings. One reason is the rotation group design of the SIPP combined with our calendar year earnings construction, and the other is the more familiar problem of attrition. The timing of the rotation groups affects the sample size we can use to

[14]Using only wage and salary earnings and including only male household heads – which is typical in the PSID – reduces the level of volatility in the SIPP survey but does not affect the trend.

calculate volatility because we require two full calendar years of earnings from January $t-1$ to December t. Appendix Table A1 lists the first and last month of each SIPP panel and the number of rotation groups used to estimate volatility in each pair of years. For example, the 1985 SIPP panel is used to construct earnings in 1984 and 1985, and all 4 rotation groups in the survey have some observations that can be used in that full two-year period. But for other panels, some rotation groups have to be omitted because their data overlap two calendar years and a full two-year calendar period cannot be constructed. For some panels and years, there are no rotation groups and no observations that can be observed fully, but because rotation groups are assigned randomly, we assume this sample loss does not create bias.

Loss from attrition, which means that some individuals are not observed from January $t-1$ to December t even though they could be, may create bias if it is correlated with individual volatility. Initial non-response rates vary across SIPP panels from about 9% to 13%. There is also sample loss within panels–usually cumulative attrition is over 20% by the fifth wave of a panel (National Research Council, 2009). The contact rules also change over time. After wave four of the 2001 panel, respondents were no longer dropped for eligibility after missing two consecutive interviews as they were in previous panels and waves. There are also idiosyncratic sample reductions. The 2004 panel was cut by 50% in wave 9 for budgetary reasons which reduces the sample sizes for 2006 and 2007. The so-called "Wave 1" bias, where attrition over a panel affects parameter estimates, is larger in the longer panels after 1996 (National Research Council, 2009).[15] We test the sensitivity of our results to the use of survey weights and inverse probability weights to correct for attrition and imputation, covered in more detail in Section 3.4.

Table 1 shows the magnitude of the various sample losses on the way to our final sample in column (F). The largest cut in sample size is in moving from Column (B) to Column (C), representing the loss from rotation group timing and attrition (and is largest in years when only a

[15] Seam bias, or the tendency for discrete changes to be too large or frequent in the first month of each wave, is unlikely to be a problem in this context because we are summing across waves.

subset of the rotation groups are available). However, there is also substantial attrition from imputation for those who do not have valid work questions (Column (D)) and from loss of those who do not have at least one non-imputed component of earnings for at least one month of the sample period (Column F). We also consider an additional sample, listed in column (E), which uses our baseline sample but instead of summing non-imputed components of earnings to calculate monthly earnings, we simply use the total earnings measure that the SIPP creates, which may be imputed. Note that the sample size is somewhat smaller due to earnings imputations of zero.

[Table 1 about here.]

2.3 SIPP Gold Standard File

2.3.1 Data

We compare estimates of volatility calculated using annual earnings in the SIPP survey to analogous estimates using the SIPP Gold Standard File (SIPP GSF). The SIPP GSF contains all individuals in a SIPP household in the 1984 and in the 1990-2008 SIPP panels plus a link, if one was found, to earnings values in the DER. Links to any and all DER earnings from 1978 through 2014 are linked to each individual in any of the just-mentioned SIPP panels.[16] [17] This match is both prospective and retrospective so, for example, an individual appearing in the 1996 SIPP panel would be matched retrospectively to their administrative earnings records prior to 1996 and prospectively to their administrative earnings records from 1996 on.

[16]This analysis was first performed using the SIPP Synthetic Beta (SSB) on the Synthetic Data Server housed at Cornell University which is funded by NSF Grant #SES-1042181. These data are public use and may be accessed by researchers outside secure Census facilities. For more information, visit https://www.census.gov/programs-surveys/sipp/methodology/sipp-synthetic-beta-data-product.html. Final results for this paper were obtained from a validation analysis conducted by Census Bureau staff using the SIPP Completed Gold Standard Files and the programs written by this author and originally run on the SSB. The validation analysis does not imply endorsement by the Census Bureau of any methods, results, opinions, or views presented in this paper.

[17]The SIPP GSF file is not directly linked to our SIPP survey files, because the Census Bureau does not provide the person IDs that would permit such a match.

The match rate between survey and administrative data for most panels is quite high. In the 1980's and 1990's panels, the match rate hovers around 80%. In 2001, the match rate dropped to 47% because many SIPP participants refused to provide social security numbers for matching. Beginning with the 2004 panel, the match rate increased to around 90% because the Census Bureau changed its matching procedures removing the necessity to explicitly ask for social security numbers. Aggregate annual match rates for men age 25 to 59 decline slightly over time from about 80 to 70 percent with a cumulative match rate of 74% across the entire period. This is in part because the earnings distribution in a given year pools together individuals from all SIPP panels, though a low match rate in one panel has a minimal impact on the share of individuals in a given year who are matched. In our analysis, we use only the sample of individuals who could be matched to their administrative earnings records.

Earnings histories in the SIPP GSF come from the DER, which are co-maintained by the SSA and the IRS. The DER represents the same universe of earners as the MEF, but contains a limited set of earnings measures. The measure of earnings that we use represents total earnings from all FICA-covered and non-FICA covered jobs with a W-2 or Schedule C (self-employment) filing. W-2 earnings are the sum of amounts from Box 1 (Total Wages, Tips, and Bonuses) and Box 12 (earnings deferred to a 401(k) type account). Earnings are not top coded after 1978.

As with the SIPP survey, we use a sample of men age 25 to 59 with positive earnings in two consecutive years. We further limit the sample to only those individuals who could be matched to administrative earnings histories. On average, the baseline sample in the SIPP GSF before trimming the earnings distribution has 103,106 observations, ranging from 73,000 to 117,000. As with the SIPP survey, sample sizes tend to increase over time.

2.4 Comparability between the Survey and Administrative Samples

For our analysis, we define samples that are appropriate for each respective dataset. This means that, although both samples come from SIPP survey participants, the SIPP survey and SIPP GSF samples will generally be different in any given pair of years where volatility can be estimated in both datasets. Additionally, the level of earnings will differ between survey and administrative data for individuals who are in both samples, as discussed above. The variables available publicly to researchers in the SIPP GSF make it impossible to estimate volatility in both survey and administrative earnings on identical samples using earnings measures that are consistent through time in the SIPP survey.

In any given year where volatility can be calculated in both the SIPP GSF and the SIPP survey, who is included in the two samples used to estimate volatility may differ for at least three reasons. First, since SIPP GSF earnings records for a given year come from individuals from different panels of the survey, and the SIPP survey data come from survey respondents only in that year, a sample for any given calendar year is not drawn from the same sampling frame and population. We examine the importance of this in our analysis with a sensitivity test that uses only SIPP GSF earnings data from those in the SIPP survey in that year. Second, there is the traditional misreporting problem arising from individuals with only under-the-table earnings and hence appear in the survey data but not the administrative data, and from individuals who report zero earnings in the survey data but have positive DER earnings. Both will result in different individuals in the two samples, in addition to differences in the level of earnings for individuals who are in both samples. This is complicated in our case because the SIPP survey has all of the imputation and non-response issues described above and we include in our survey sample only a subsample of all observations. Third, only individuals matched to administrative earnings histories are included in the SIPP GSF, while no analogous constraint exists for the survey. The inability to match all SIPP respondents to the DER implies that there may be individuals who are included in SIPP survey data whose records could not be linked to administrative data sources. The SIPP survey is also not included as part of the SIPP GSF between 1985 and 1988.

Although the SIPP GSF is constructed by linking SIPP survey data to the DER, we treat the SIPP GSF and the SIPP survey as if they are separate datasets, and define samples that are appropriate for each respective dataset and compare them. However, several other papers have conducted detailed comparisons of individual earnings between the two data sources in actual linked samples so that the same individuals are in both samples. In the 2008 panel, Chenevert, Klee, and Wilkin (2016) find that for nearly 70% of the sample, earnings in the SIPP survey are within $5000 of earnings in the DER and in about 18% of the sample, the earnings in the two data sources differ by more than $10,000. In most cases where the two do not match, the DER earnings are higher than the SIPP survey earnings. Differences between the SIPP survey earnings and those in the DER are larger when earnings are reported by proxy or contain imputations (Chenevert, Klee, and Wilkin, 2016). Cristia and Schwabish (2009) draw similar conclusions in the 1996 panel. Using a similar strategy of SIPP survey respondents linked to their DER earnings, Gottschalk and Huynh (2010) use the 1996 SIPP panel and show that earnings inequality is lower in the SIPP survey data than in the SIPP linked to the DER. Additionally, Abowd and Stinson (2013) use a different strategy and compare earnings at the person-job level in the 1990 - 1996 SIPP panels. They show that the correlations in earnings at the person-job level between adjacent years is lower in the SIPP survey than in the DER. The variance of earnings at the person-job level is not systematically higher in the DER than in the SIPP. Nearly all papers conclude that including imputed earnings in the SIPP survey yields earnings that are less comparable with the administrative earnings from the DER. As discussed above, this finding is reinforced by Dahl, DeLeire, and Schwabish (2011), who compare volatility in household earnings between SIPP survey earnings and DER earnings using an identical set of households, again finding that the trend in volatility between 1985 and 2005 is similar in the two datasets as long as imputations are excluded from the survey data.

Table 2 compares the demographic characteristics in our SIPP GSF and SIPP survey samples, taken over all person-year observations. The first column (All) includes all men on the SIPP GSF file age 25 to 59 who participated in any SIPP panel, the second column (Matched) includes the subset of those men who can be matched to their administrative earnings histories including those with zero earnings, the third column (Volatility Sample: GSF) includes the subset of matched men who have positive earnings in two consecutive years and hence for whom we estimate volatility in the SIPP GSF, and the last column (Volatility Sample: SIPP) contains the baseline SIPP survey sample described above. Comparing the subsamples from the SIPP GSF with the whole sample, the matched sample is slightly better educated and slightly more likely to be white than the full sample, but have the same average age. The volatility sample in the SIPP GSF is again somewhat better educated, somewhat more likely to be white, and about the same age. The volatility sample in the SIPP survey has higher educational attainment than any of the other samples and is more likely to be white. However, the age differences are small.

[Table 2 about here.]

Beyond balance in demographic characteristics, there is the concern about whether the two samples yield similar earnings distributions given that volatility and inequality are related to each other, a point we return to below. Specifically, a central issue for estimates of volatility is that earnings growth rates are highly sensitive to the effect of small absolute earnings changes at the bottom of the earnings distribution. Table 3 shows selected earnings percentiles for selected years in the SIPP GSF and the SIPP survey. In the two SIPP samples, the density of low earnings is higher in the administrative data. Though median earnings and the top half of the earnings distribution are similar across the two datasets, there is a much longer left tail in the earnings distribution in the administrative data.[18] Additionally the first percentile of earnings in the administrative data drops substantially in real terms over the period. In the SIPP survey data, the earnings distribution has a shorter left tail–indicated by the consistently higher 1st, 5th and 10th percentiles–but is also less stable at the very bottom. The falling 10th and 5th percentiles in the SIPP GSF implies a growing share of individuals at the bottom of the earnings distribution in the administrative data. Note also that the cross-sectional samples sizes of the SIPP survey are

[18]The long left tail of the earnings distribution is also present in the LEHD (Abowd, McKinney, and Zhao, 2018, Juhn and McCue, 2010).

considerably smaller than the SIPP GSF.

[Table 3 about here.]

The most straightforward way to address the issue of low earnings is to trim the earnings distribution prior to calculating earnings changes. We trim on percentile points of the earnings distribution in each dataset in each year. In Section 3.2 we show results under three different trimming methods: trimming the top and bottom 1% of the earnings distribution, trimming the top and bottom 5% of the distribution, and doing no trimming. We refer to these at the 1% trim, 5% trim, and untrimmed, respectively. In Section 4 we directly consider the role of differences in the earnings distribution for our estimates of volatility by reweighting the earnings distributions in both the SIPP survey data and the SIPP GSF to an external source of data, the PSID.

3 Results

In our main results, we focus on differences in the levels and trends in volatility in the SIPP GSF and SIPP survey data. We first show our baseline results for trends using our preferred sample definitions and methods. We then conduct a number of sensitivity tests to those results by varying the method of trimming, the method of age adjusting, the inclusion of imputed earnings in the survey data, the use of weights to adjust for attrition bias in the SIPP survey, and the restriction of the SIPP GSF to include only respondents from the SIPP panel corresponding to each calendar year.

3.1 Baseline Trends

Figure 1a shows trends in volatility in the SIPP GSF and SIPP survey using the arc change and log change methods as given in Equations 1 and 2, respectively. In Figure 1a, earnings are trimmed at the top and bottom 1% of the earnings distribution in each year.

In terms of levels, volatility in the SIPP GSF is roughly 2 times larger than that in the SIPP survey data. In both datasets volatility is larger in log changes than in arc changes, and there is a larger relative difference between the SIPP GSF and the SIPP survey when volatility is measured in log changes. The larger difference in the level of volatility between the SIPP GSF and the SIPP survey in log changes in part reflects the higher weight that the log change measure places on larger earnings changes, and what must be a higher density of large changes in the SIPP GSF. The higher density of large earnings changes could be the result of the higher density of low earnings in the SIPP GSF, a point we return to below.

In terms of trends, the major conclusion from Figure 1a is that neither the administrative nor survey data show any discernible trends over the entire period. While the SIPP GSF volatility shows a mild U-shaped pattern, at least if one begins in 1984 rather than 1980–declining through about 2000 and then rising through the Recession peak in 2010–the final values in 2014 are almost identical to their initial values in 1980. For the SIPP survey, there appears to be a mild downward trend when the arc change measure is used but less when the log change measure is used. But even the former decline is very mild. We will argue that the lack of a significant long-term trend in either the survey or administrative data is robust to all sensitivity tests we conduct and therefore that the two types of data are consistent with one another.

The trends as well as fluctuations over time are more easily seen in Figure 1b, which normalizes each series to its respective pooled mean, thereby removing level differences. The series show a similar time trend in volatility for the two data sets, with the mild U-shaped pattern in the SIPP GSF also apparent. Both datasets show declining volatility into the late 1990s, and increasing volatility after, with both ending up about where they began. Relative fluctuations in the survey measures are much larger than in the SIPP GSF, particularly the increase in volatility in the early 1990s and the decline into 1999. While both datasets show a local peak in 1993 and trough in 1999, volatility in the SIPP survey declined about 70% over this time while in the SIPP GSF it only declined about 10%. Table 3 shows that the larger relative changes seen in the SIPP survey in the early 1990s coincide with a period of instability in the 1st and 5th percentiles of the earnings distribution. We speculate that changes in interview mode in from in-person to telephone in the early 1990s, and a redesign of the SIPP survey in 1996, may be responsible for these swings. R. Moffitt and S. Zhang suggest that changes in interview mode in the PSID may have contributed to higher levels of volatility in that dataset during the same period as well (Moffitt and Zhang, 2020).

[Figure 1 about here.]

The increase in volatility during the Great Recession is similar in the two series. Volatility increased by about 30% in both datasets.[19] In both datasets, volatility fell after 2010. In the SIPP GSF volatility fell to its pre-recession levels by 2014. The SIPP survey data ends in 2012 when volatility remained above its pre-recession level. Over the period between 1985 and 2012 when the series in SIPP GSF and SIPP survey data overlap, volatility increased modestly in the SIPP GSF and fell modestly in the SIPP survey data. In the SIPP GSF, volatility is at its pooled mean in 1985 and 10% above its pooled mean in 2012. In the SIPP survey, volatility is also at its pooled mean in 1985 and 10% below its pooled mean in 2012.

Because the trends in volatility measured in log changes and arc changes are similar, in what follows we show the sensitivity tests for for volatility in arc changes. Results in log changes are available on request.

3.2 Trimming Low Earnings

Our first sensitivity test is to our baseline 1% trimming method. The differential trends in the administrative and survey data could be affected by differences in the density of observations in the tails, especially if tail observations have different volatility trends. We showed in Table 3 that

[19] We are missing two data points (2007 - 2008 and 2008 - 2009) in the SIPP survey data because of the dating of the 2004 and 2008 panels.

trends in very low earnings in the SIPP survey and SIPP GSF are quite different. This could affect trends in volatility as, again, low earnings can have an outsized impact on earnings changes measured in percents. Figures 2 and 3 show the effect of different methods of trimming low earnings on the baseline estimates of volatility in the SIPP GSF and the SIPP survey data, respectively. The figures show volatility in the untrimmed earnings distribution, with the 1% trim (baseline estimates), and with the 5% trim.

Figure 2 shows that trimming the SIPP GSF reduces volatility. Volatility in the untrimmed earnings distribution measured in arc changes fluctuates around 0.23. Trimming the top and bottom 1% of the earnings distribution reduces volatility measured in arc changes by about 17%. Trimming the top and bottom 5% reduces volatility by about 50% measured in arc changes compared to untrimmed. Thus, about 50% of volatility in the SIPP GSF in any given year comes from individuals with earnings below the 5th percentile in either t or $t-1$, or both. The normalized series show that trimming high and low earnings does little to alter the trends in volatility.

[Figure 2 about here.]

Figure 3 repeats this trimming exercise for the SIPP survey data. Again, volatility declines with each successive trim. Volatility measured in arc changes declines by 35% when trimming the top and bottom 1% of earnings, and by 38% when trimming the top and bottom 5% of earnings, compared to untrimmed volatility.The normalized series confirm that trimming has minimal impact on the trend. The exception is a decline in volatility between 2005 and 2007 when earnings are trimmed at the top and bottom 5% that is not apparent in the other trims.Trimming high and low earnings has a similar impact in the two datasets: levels decline by similar magnitudes but trends are unchanged.

[Figure 3 about here.]

Carr and Wiemers (2020) have pointed out that an important issue in all trimming procedures is that trimming on percentile points is preferred to trimming on levels of earnings. Trimming on levels of earnings is problematic if the tails are changing over time as, for example, would happen if earnings inequality is increasing in part because the lower tail of earnings is declining in real value. In that case, a trim in fixed real dollars would exclude an increasing fraction of the population, and would alter trends in volatility if the lower earners have a different level of volatility than other workers. Figure 4 is taken from Carr and Wiemers (2020) and shows the effect on volatility in the SIPP GSF of trimming earnings at the bottom 1% and bottom 5% along with the trimming methods used in Kopczuk, Saez, and Song (2010), which excluded observations below a fixed real dollar amount ($3770); in DeBacker et al. (2013) and Guvenen, Ozkan, and Song (2014) which used a nominal dollar trim tied to the level of the federal minimum wage in each year (Min Wage); and in Sabelhaus and Song (2009, 2010) which used a nominal dollar earnings level for trimming tied to the Social Security minimum earnings qualification level in each year (SSA).

[Figure 4 about here.]

While the untrimmed, 1%, and 5% trim results show no trend, as we have found here, the SSA trimming method shows a marked decline in volatility while the $3770 and Min Wage trims show a slightly negative trend. The downward trends in the SSA, the $3770, and the minimum wage trim are consistent with the other estimates of volatility using administrative data sources that use these trims. The decline in volatility with these trims occurs because the lower tail of the earnings distribution has higher volatility than the rest of the distribution and, over time, and these methods trim an increasing fraction of low earnings. The trend in the untrimmed data series between 1985 and 2003 is consistent with Dahl, DeLeire, and Schwabish (2008), who find small declines in volatility in untrimmed earnings for men in the Continuous Work History Sample, and nearly identical levels and trends in SIPP data linked to the DER. While Sabelhaus and Song

(2010) also find declining volatility in an untrimmed earnings distribution, they pool men and women together for this estimate, and Dahl, DeLeire, and Schwabish (2008) finds larger declines for women than men over most of this time period.

3.3 Age-Adjusting

We also test the sensitivity of our results to the choice to measure volatility using age-adjusted earnings changes. Figure 1a shows trends in volatility in age-adjusted earnings changes which is standard in the literature. Figure 5 reports those age-adjusted baseline results and results using non-age adjusted earnings changes. In neither series are the differences in trends visible–the series overlap almost exactly. Thus the results are not sensitive to whether earnings changes are age adjusted. This necessarily implies that neither the age distribution nor the coefficients on the age variables in the first-stage regression are changing much over time.

[Figure 5 about here.]

3.4 Imputations in the SIPP Survey

As we described at length in our discussion of survey imputations, our annual earnings variable in each year sums only non-imputed earnings in months when the individual is working. But some men have imputed earnings components as well as non-imputed earnings, so we conduct a sensitivity test by computing volatility in our baseline sample of individuals but using the SIPP-generated total monthly earnings variable as the measure of earnings. Figure 6 shows the results, demonstrating that the level of volatility is essentially identical whether we use our non-imputed monthly earnings or imputed monthly earnings. The similarities in trends are confirmed when each series is normalized to its respective mean. Implicitly, this suggests that the year-to-year volatility of imputed components is the same as that for the non-imputed for the men in our sample and that they both move similarly over time. We emphasize that the results do depend on using imputation flags that are consistent over time and excluding whole-case imputations, as we discussed in detail above. Dahl, DeLeire, and Schwabish (2011) show that using whole case imputes imparts an upward bias on estimates of household volatility.[20]

[Figure 6 about here.]

3.5 Aligning the SIPP GSF and the Survey Data

As we noted in our description of the SIPP GSF administrative earnings data, we pool earnings data, using, in each year, respondents who participated in any SIPP panel during, before, or after that year. While the sampling frame for each SIPP panel is a well-defined representative frame for the U.S. population in that period, it is unclear how to define the sampling frame in a period which draws from the U.S. population in different years. To test for any problems introduced by this approach, we estimate volatility using SIPP GSF observations only on the subsample of men age 25 to 59 who participated in a given SIPP panel in the same year. For example, if we consider earnings changes between 1984 and 1985, we compare our current SIPP GSF estimates, which use observations on men from all SIPP panels, to that estimated only for men who participated in the 1984 SIPP panel who are age 25 to 59 in 1984 and 1985. Figure 7 shows the results of this exercise, showing volatility in age-adjusted arc changes with a 1% trim, where the SIPP panel year estimate is weighted using the initial SIPP survey weight of individuals in the panel.

[Figure 7 about here.]

While the trends are broadly similar, some differences emerge. Between 1985 and 1999, the level of volatility is quite similar in the full SIPP GSF sample and in the sample restricted to panel

[20]We also note that our trends reported for the SIPP survey are different than that reported in Celik et al. (2012), regardless of which method, trim, or earnings measure we use. Celik et al. (2012) find a downward trend in volatility in the SIPP, with a decline of about 20%. We are unable to define a sample that we believe is both consistent through time and handles monthly earnings consistently over the calendar year that replicates this trend.

respondents. Between 1999 and 2007, volatility in the full sample increases more than the panel specific sample. By 2009, the levels of volatility converge again. It appears as if trends in volatility in the 2001 and 2004 panels are slightly different than the full SIPP GSF sample. Because of slightly higher levels of volatility in 1985 and slightly lower volatility in 2013 when using only those individuals present in a SIPP panel, the overall trend in the SIPP GSF declines slightly in the panel specific sample. Nevertheless, the trends are broadly similar across the two estimates.

3.6 Attrition in the SIPP Survey

As we discussed previously, attrition in the SIPP survey is non-trivial. It is well known that attrition that is correlated with the outcome variable being examined can generate bias in estimates from the non-attrition sample alone. In the SIPP survey, the selection procedures requiring that a full two-year sequence of non-imputed earnings be available likely selects for individuals with more stable labor supply and earnings (Bollinger et al., 2019, Fitzgerald, Gottschalk, and Moffitt, 1998), which may bias volatility downward. The changes in the survey in 1996 also represent a challenge to creating a consistent time series that is population representative. We address this issue by testing the sensitivity of our results to to the use of sample weights provided by the SIPP survey and those we construct ourselves which correct more directly for the correlation between attrition and the level of earnings.

The SIPP constructs weights for each individual in the sample which are based on the usual differential selection probabilities from the geographic units used to draw the sample, for clustering in the sampling design, and for differential initial nonresponse probabilities. They also are constructed to account for differential attrition and have a poststratification design to meet outside totals for the distribution of the population.[21] In Figure 8 we test the sensitivity of our results from the SIPP survey to the use of these sample weights. It shows volatility in our

[21] https://www.census.gov/programs-surveys/sipp/methodology/weighting.html

preferred specification, the age-adjusted arc change using the 1% trim in the SIPP survey data, unweighted and weighted. We weight the SIPP survey data using the calendar year weights for year t. These weights are assigned to everyone who is observed at a control date (usually Jan 1 of the calendar year) and for every month of the calendar year for which they are in scope for the survey. Figure 8 shows that the trend results are nearly identical in the weighted and unweighted data.

[Figure 8 about here.]

Although the survey weights are designed to adjust for attrition from the SIPP survey, they do so by accounting for differences in the distribution of household characteristics not including the level of earnings. Fitzgerald, Gottschalk, and Moffitt (1998) showed that attrition probabilities in the PSID are significantly correlated with lagged earnings, so we develop our own weights and test for earnings-related attrition bias in this way. In this common method, typically referred to as inverse probability weighting and often used for attrition adjustments (Wooldridge, 2010), we estimate a first stage probit for attrition propensities as a function of lagged earnings and then use these to reweight the volatility estimates based on the respondent sample. Specifically, we estimate the probability of remaining in the sample in year t, conditional on age and earnings in $t-1$, as given in Equation 3.

$$Pr(Remain = 1) = \beta_0 + \beta_1 age_{i,t-t} + \beta_2 age^2_{i,t-1} + \beta_3 earn_{i,t-1} + u_{it} \tag{3}$$

We then use the inverse of the predicted probability of remaining in the sample, multiplied by the SIPP sample weights, to weight estimates of volatility in the SIPP. The combined weight accounts for both attrition from the SIPP survey as a whole and for selective attrition based on age and earnings from our sample.

The results of this exercise are shown in Figure 9. The results are nearly identical with and without weights, indicating that volatility in the baseline sample is not sensitive to selective attrition based on earnings between $t-1$ and t. While we cannot rule out the existence of attrition correlated with unobservables beyond household characteristics and lagged earnings, the lack of any effect based on those observables suggests that any bias may not be large.

[Figure 9 about here.]

4 Adjusting for Differences in the Cross-Sectional Distribution of Earnings

In Table 3, we showed that the SIPP GSF and SIPP survey data have different cross-sectional earnings distributions, particularly at the bottom of the earnings distribution. If the level of volatility is different in different portions of the cross-sectional distribution, this could generate differences in the levels of average volatility in the two data sets. If the density of low earnings is trending differently in different datasets, this could affect the trend in volatility. Additionally, if volatility is trending differently for individuals at different points in the cross-sectional distributional, trends could also be affected. We examine this issue in this section by reweighting the SIPP survey and SIPP GSF to the observed cross-sectional earnings distribution of the PSID. This exercise allows us not only to benchmark our two SIPP data sets to each other, but also to compare our results to those from other papers in this project which are also benchmarking their cross-sectional earnings distributions to the PSID.[22]

Figure 10 shows log earnings ratios for selected percentile points relative to the median for the SIPP survey, SIPP GSF, and the PSID.[23] Figure 10 confirms that the biggest differences in the earnings distributions between the three datasets are at the bottom of the earnings distribution. The P50/P10 in the SIPP GSF is about the same level as the P50/P5 in the SIPP survey, and 30%

[22]This paper is part of a joint project estimating volatility in the PSID, the CPS, and the LEHD, as noted earlier. The papers reporting results for the latter two are also reweighting to the PSID, to establish a common set of estimates.

[23]Estimates for the PSID were provided to us by R. Moffitt and S. Zhang in their companion paper in this project.

to 50% higher than the P50/P5 in the PSID. In addition, the density of low earnings in both the SIPP survey and the PSID trends down slightly between the 1980s and late 1990s, while it is flat or increasing everywhere in the SIPP GSF. The levels and trends in the top half of the earnings distribution are more similar across datasets, though the increase in inequality is higher at the top of the earnings distribution in the SIPP GSF than in the SIPP survey data or the PSID. Log earnings ratios in the SIPP survey data also show clear breaks between panels after 1996 where inequality increases between the last year of one panel and the first year of the next but then declines over the life of a panel.

[Figure 10 about here.]

In order to separate differences in inequality from differences in volatility, we reweight the cross-sectional earnings distributions in the SIPP survey and SIPP GSF, respectively, to the distribution in the PSID. To do so, we first trim the SIPP survey and SIPP GSF at the 1st and 99th percentiles of the PSID in each year and then divide the remaining cross-sectional earnings distribution into ventiles determined by the ventile cut points in the PSID. We then estimate weighted volatility in arc changes using as weights the inverse share of observations in each earnings group determined by the ventile cutpoints in the PSID. We use two different sets of weights, one in which the weights are determined by the PSID annual cross section in each year t (Perm 2), and one in which we fix the weights based on the year 2000 PSID distribution and apply the same weights to all years (Perm 3). We repeat this exercise with age-adjusted earnings and age-adjusted earnings changes in Perm 5 and Perm 6. Given differences in the density of low earnings, this exercise reduces the weight on low earnings in the SIPP GSF relative to the SIPP survey, and reduces the weight on low earnings for both of them relative to the PSID. Complete details of this process are outlined in Appendix B.[24]

[24]The ventile cut points were provided by R. Moffitt and S. Zhang, as were the volatility estimates from the PSID shown in Figure 11

Figure 11 shows the level of volatility and the normalized trend in the SIPP GSF and SIPP survey datasets. Figures 11a and 11b show the level and normalized trends of volatility in arc changes without age adjustments after reweighting to the PSID and include the PSID trend as a reference. Figures 11c and 11d show all of the permutations for the SIPP GSF and SIPP survey, respectively. Reweighting the SIPP survey and SIPP GSF to the PSID virtually eliminates the level difference in volatility between the two, as demonstrated by the similarity in level between Perm 2 on the SIPP GSF and SIPP, respectively. It also reduces the small differences in trend between the two. Overall, reweighting cuts SIPP GSF volatility in half while reducing the level in the SIPP survey data by about 20%. Reweighted volatility in the SIPP GSF is also 20% to 40% lower than the PSID. This reweighting exercise shows that the differences in the cross-sectional earnings distributions between the SIPP survey and SIPP GSF account for the entire difference in the level of volatility between the two. It also shows that, after accounting for differences in the distribution of cross-sectional earnings, the level of volatility is lower in the SIPP GSF than in the PSID.

[Figure 11 about here.]

Figure 11b, which displays mean-normalized trends, shows that overall trends in both the observed and reweighted series remain similar. The only notable change is the more exaggerated U-shape seen in Perm 2 of the SIPP GSF. This U-shape comes from the fact that between the mid 1980s and the late 1990s the density of low earnings in the PSID trends down slightly while it trends steadily up in the SIPP GSF. This is confirmed in Figure 11c, which shows that fixing the earnings distribution in 2000 (Perm 3) produces a flatter trend relative to the slightly U-shaped pattern found when allowing the distribution to change from year to year, as in Perm 2. Figures 11c and 11d also confirm that overall conclusions are not sensitive to age adjusting, as demonstrated by the similarities between Perms 2 and 5, and Perms 3 and 6, respectively.

4.1 Summary

In this paper we compare estimates of earnings volatility from SIPP survey data and the SIPP GSF, which links SIPP survey respondents to their respective administrative earnings records. We show that the trends in the SIPP survey and the SIPP GSF are similar. Volatility increases modestly in the SIPP GSF and declines modestly in the SIPP survey data between the mid-1980s and the early 2010s. Levels in the SIPP GSF are considerably higher than in the SIPP survey data. The results are robust to age-adjusting, more aggressively trimming low earnings, the use of imputations, and different methods of handling population representativeness. We also show that differences in the underlying earnings distribution explain the differences in the trend and level of volatility in the SIPP GSF and SIPP survey data. When we force the earnings distributions in the two datasets to be the same, we find that the differences in levels of volatility largely disappear, and the minor differences in trends that exist are reduced.

This paper is part of a larger project designed to consider estimates of earnings volatility across multiple datasets. As part of this project, we also compare volatility in the SIPP GSF and SIPP survey data to that in PSID. We compare volatility using the same sample selection criteria and methods and find that there remain differences in both the level and trend in volatility. Volatility in the SIPP GSF is higher than the PSID, while volatility in the SIPP survey is lower, and both shower smaller relative increases between 1999 and 2010. Even when the earnings distributions are forced to be the same across the three datasets, there remain differences in both level and trend in volatility between the SIPP and the PSID. We find that, for both the SIPP GSF and the SIPP, survey volatility falls to a lower level than the PSID when weighting the cross-sectional earning distribution to the PSID, implying that volatility is higher in the PSID at a fixed level of inequality.

More broadly, the results of this paper and the other papers in this group have important implications for the broader literature on volatility. Using 3 different survey data sets and 3 different administrative data sets, we find broad agreement that volatility has shown no major trend upward or downward in the past 35 years. Thus our paper, along with the others, finds no difference in the evidence obtained from survey and administration data. Another finding that has broad applicability is that, in some cases, differences in the underlying cross-sectional earnings distribution in different data sets plays a role in explaining differences in both levels and trends in volatility. The substantial differences across data sources in the level of inequality, across methods used to estimate volatility in the weight placed on different portions of the earnings distribution, and across sample definitions that may themselves alter the level and trend in inequality, could contribute to some of the disagreement that has developed in this literature.

References

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.

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):1451–1467.

Abraham, Katherine G., John C. Haltiwanger, Kristin Sandusky, and James Speltzer. 2013. "Exploring differences in employment between household and establishment data." *Journal of Labor Economics* 31:129–172.

Bania, Neil and Laura Leete. 2009. "Monthly Household Income Volatility in the U.S., 1991/1992 vs. 2002/2003." *Economics Bulletin* 29 (3):2100–2112.

Bollinger, Christopher R., Barry T. Hirsch, Charles M. Hokayem, and James P. Ziliak. 2019. "Trouble in the Tails? What We Know about Earnings Nonresponse 30 Years after Lillard, Smith, and Welch." *Journal of Political Economy* 127 (5):2143–2185.

Carr, Michael D. and Emily E. Wiemers. 2018. "New Evidence on Earnings Volatility in Survey and Administrative Data." *American Economic Review Papers and Proceedings* 108.

———. 2020. "The Role of Low Earnings in Differing Trends in Earnings Volatility." *Washington Center for Equitable Growth Working Paper* .

Celik, Sule, Chinhui Juhn, Kristin McCue, and Jesse Thompson. 2012. "Recent trends in earnings volatility: evidence from survey and administrative data." *BE Journal of Economic Analysis & Policy* 12 (2):1–26.

Chenevert, Rebecca L., Mark A. Klee, and Kelly R. Wilkin. 2016. “Do Imputed Earnings Earn Their Keep? Evaluating SIPP Earnings and Nonresponse with Administrative Records.” *WORKING PAPER NUMBER SEHSD-WP2016-18, SIPP-WP-275, Census Bureau* :1–74.

Cristia, Julian and Jonathan A. Schwabish. 2009. “Measurement error in the SIPP: Evidence from Administrative Matched Records.” *Journal of Economic and Social Measurement* 34 (1):1–17.

Czaka, John L. and Gabrielle Denmead. 2008. “Income Data for Policy Analysis: A Comparative Assessment of Eight Surveys.” Final Report to the U.S. Department of Health and Human Services, Mathematica Policy Research.

Dahl, Molly, Thomas DeLeire, and Jonathan 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.

Dahl, Molly, Thomas DeLeire, and Jonathan A Schwabish. 2008. “Trends in Earnings Variability over the Past 20 Years.” Tech. rep., Congression Budget Office.

DeBacker, Jason, Bradley Heim, Vasia Panousi, and Ivan Vidangos. 2013. “Rising Inequality: Transitory or Permanent? New Evidence from a Panel of Tax Returns.” *Brookings Papers on Economic Activity* :67–122.

Dynan, Karen, Douglas Elmendorf, and Daniel Sichel. 2012. “The Evolution of Household Income Volatility.” *B.E. Journal of Economic Analysis and Policy* 12 (2).

Fitzgerald, John, Peter Gottschalk, and Robert Moffitt. 1998. “An Analysis of Sample Attrition in Panel Data: The Michigan Panel Study of Income Dynamics.” *Journal of Human Resources* 33 (2):251–299.

Gottschalk, Peter and Minh Huynh. 2010. “Are Earnings Inequality and Mobility Overstated?

The Impact of Nonclassical Measurement Error." *The Review of Economics and Statistics* 92 (2):302–315. URL `https://doi.org/10.1162/rest.2010.11232`.

Gottschalk, Peter, Robert Moffitt, Lawrence F Katz, and William T Dickens. 1994. "The Growth of Earnings Instability in the US Labor Market." *Brookings Papers on Economic Activity* 1994 (2):217–272.

Guvenen, Fatih, Serdar Ozkan, and Jae Song. 2014. "The Nature of Countercyclical Income Risk." *Journal of Polical Economy* 22 (3):621–660.

Hryshko, Dmytro, Chinhui Juhn, and Kristen McCue. 2017. "Trends in Earnings Inequality and Earnings Instability Among U.S. Couples: How Important is Assortative Matching?" *Labour Economics* 48:168–82.

Juhn, Chinhui and Kristin McCue. 2010. "Comparing Measures of Earnings Instability Based on Survey and Administrative Reports." *U.S. Census Bureau Center for Economic Studies Paper No. CES-WP-10-15* :1–32.

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 (1):91–128.

McKinney, Kevin L. and John M. Abowd. 2020. "Male Earnings Volatility in LEHD Before, During, and After the Great Recession." Working Paper.

Moffitt, Robert and Sisi Zhang. 2018a. "Income Volatility and the PSID: Past Research and New Results." *American Economic Association Papers and Proceedings* 108:277–80.

———. 2018b. "The PSID and Income Volatility: Its Record of Seminal Research and Some New Findings." *The Annals of the American Academy of Political and Social Science* 680 (1):48–81.

Moffitt, Robert A and Peter Gottschalk. 2012. "Trends in the Transitory Variance of Male Earnings Methods and Evidence." *Journal of Human Resources* 47 (1):204–236.

Moffitt, Robert A. and Sisi Zhang. 2020. "Estimating Trends in Male Earnings Volatility Using the Panel Study of Income Dynamics." Working Paper. `http://www.econ2.jhu.edu/people/Moffitt/moffitt-zhangdraft2020-7-7%20final.pdf` (accessed 7/8/2020).

National Research Council. 2009. *Reengineering the Survey of Income and Program Participation*. Washington, D.C.: The National Academies Press.

Pennell, Stephen G. 2003. "Cross-Sectional Imputation and Longitudinal Editing Procedures in the Survey of Income and Program Participation." *U.S. Department of Commerce, Census Bureau* :1–73.

Sabelhaus, John and Jae Song. 2009. "Earnings Volatility Across Groups and Time." *National Tax Journal* 2 (62):347–364.

———. 2010. "The Great Moderation in Micro Labor Earnings." *Journal of Monetary Economics* 57.

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):973–982.

Wooldridge, Jeffrey M. 2010. *Econometric Analysis of Cross Section and Panel Data, 2e*. Boston, MA: MIT Press.

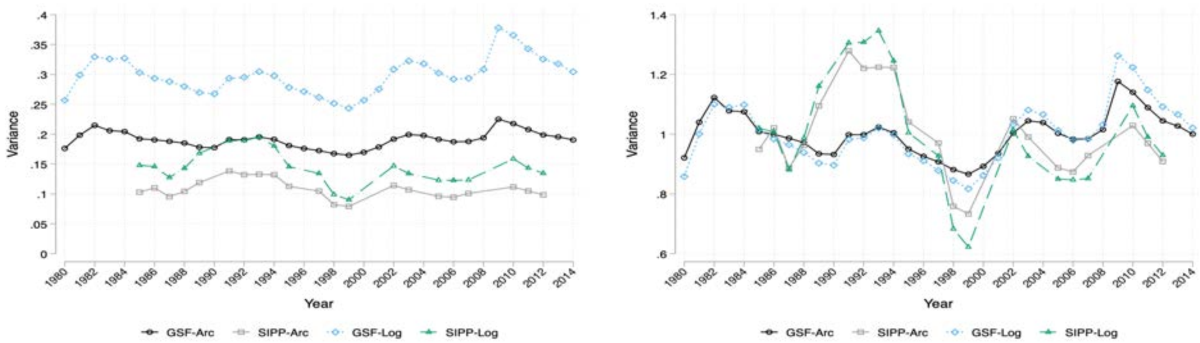
Ziliak, James P, Bradley Hardy, and Christopher Bollinger. 2011. "Earnings Volatility in America: Evidence from Matched CPS." *Labour Economics* 18 (6):742–754.

Ziliak, James P., Charles Hokayem, and Christoper R. Bollinger. 2020. "Trends in Earnings Volatility using Linked Administrative and Survey Data." Working Paper.

http://ukcpr.org/sites/ukcpr/files/research-pdfs/DP2020-01.pdf (accessed 7/8/2020).

Figure 1: Volatility, Age Adjusted, 1% Trim Sample

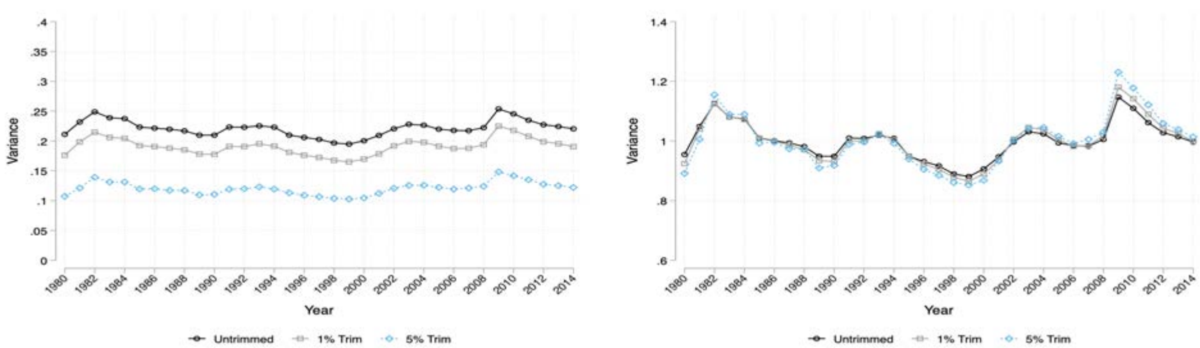
(a) Levels (b) Mean Normalized



Notes: Authors' calculations based on SIPP GSF and SIPP. Sample includes men age 25 to 59 with positive annual earnings in two consecutive years. Earnings are trimmed at the bottom and top 1% of the full distribution of earnings. Earnings changes are adjusted using a quadratic in age, separately by year.

Figure 2: Volatility, Age Adjusted, SIPP GSF, Selected Trims

(a) Arc Change (b) Normalized Arc Change

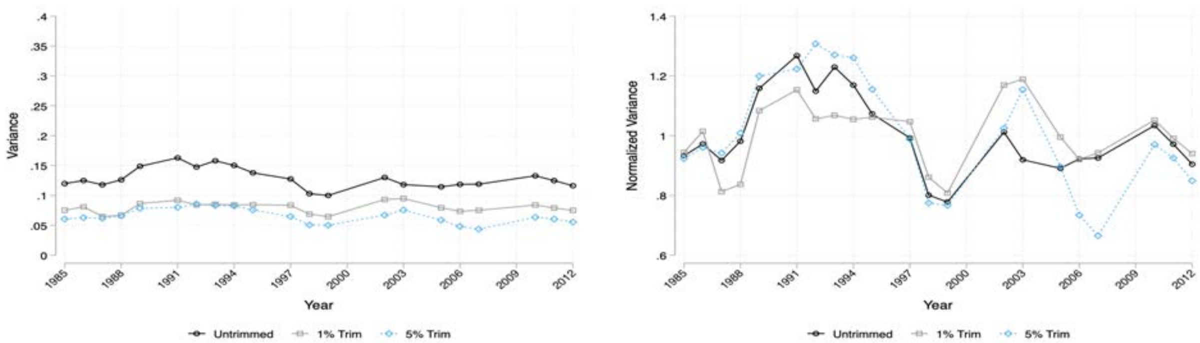


Notes: Authors' calculations using SIPP GSF. Sample includes men age 25 to 59 with positive annual earnings in two consecutive years. Earnings changes are age-adjusted using a quadratic in age, separately by year.

Figure 3: Volatility, Age Adjusted, SIPP Survey, Selected Trims

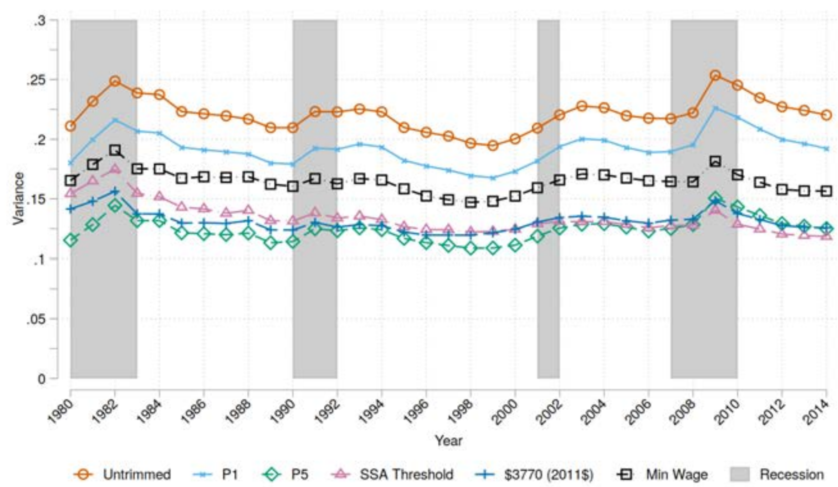
(a) Arc Change

(b) Normalized Arc Change



Notes: Authors' calculations using SIPP survey. Sample includes men age 25 to 59 with positive annual earnings in two consecutive years. Earnings changes are age-adjusted using a quadratic in age, separately by year.

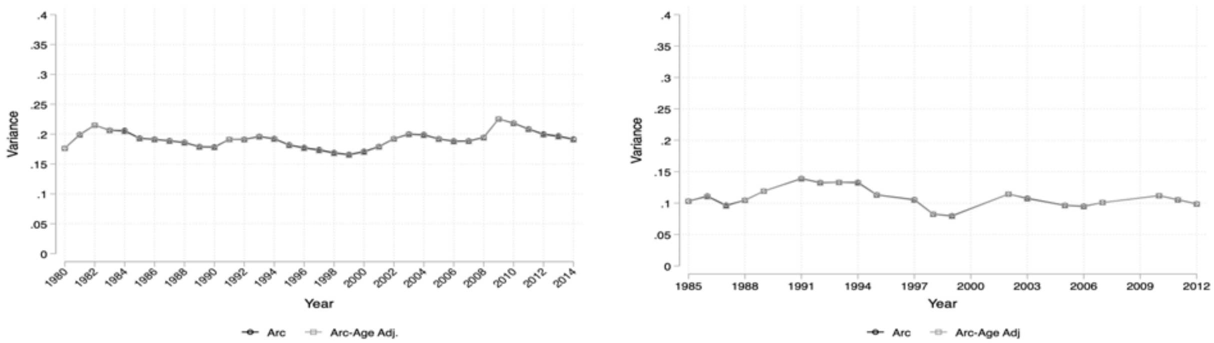
Figure 4: Volatility, Age Adjusted, SIPP GSF, Alternative Trimming Methods



Notes: Carr and Wiemers (2020). Sample is men age 25 to 59 with positive earnings in t and $t-1$. Trimming methods are: no trim (Untrimmed), bottom 1% (1%), bottom 5% (5%), SSA earnings threshold (SSA), annual federal minimum wage (Min Wage), and fixed $3,770 in 2011 dollars ($3770).

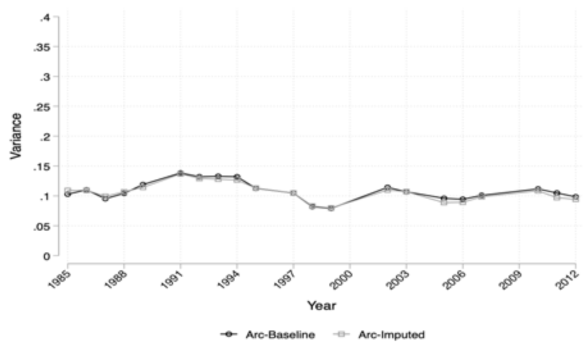
Figure 5: Volatility, Age and Not Age Adjusted, 1% Trim Sample

(a) SIPP GSF   (b) SIPP



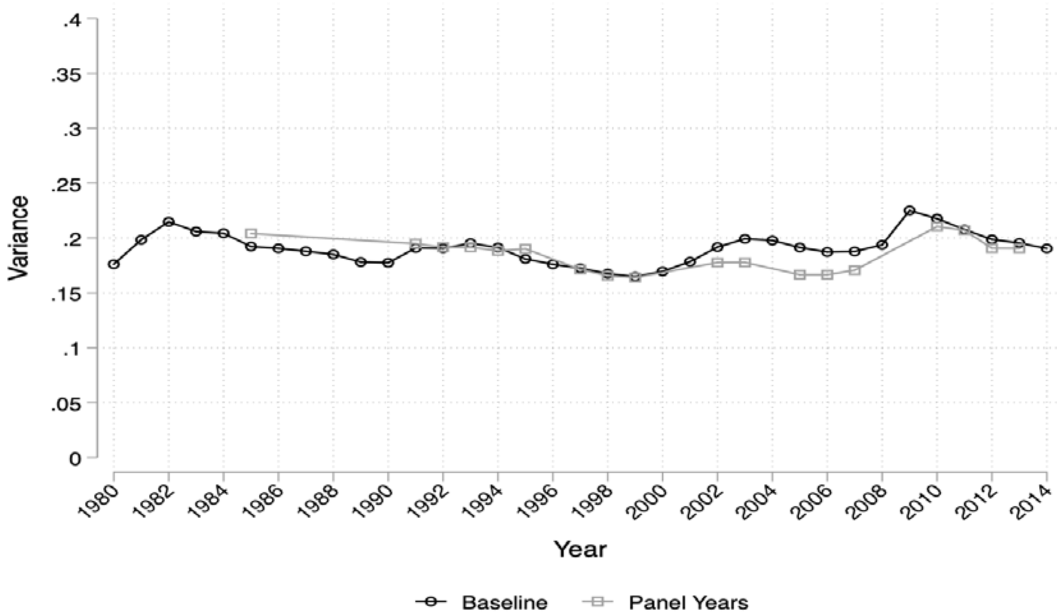
Notes: Authors' calculations using SIPP GSF and SIPP. Sample includes men age 25 to 59 with positive annual earnings in two consecutive years. Earnings are trimmed at the bottom and top 1% of the full distribution of earnings for men age 25 to 59 with positive earnings. Earnings changes are age-adjusted using a quadratic in age, separately by year.

Figure 6: Volatility, Age Adjusted, With and Without Imputed Earnings Components, 1% Trim Sample, SIPP Survey



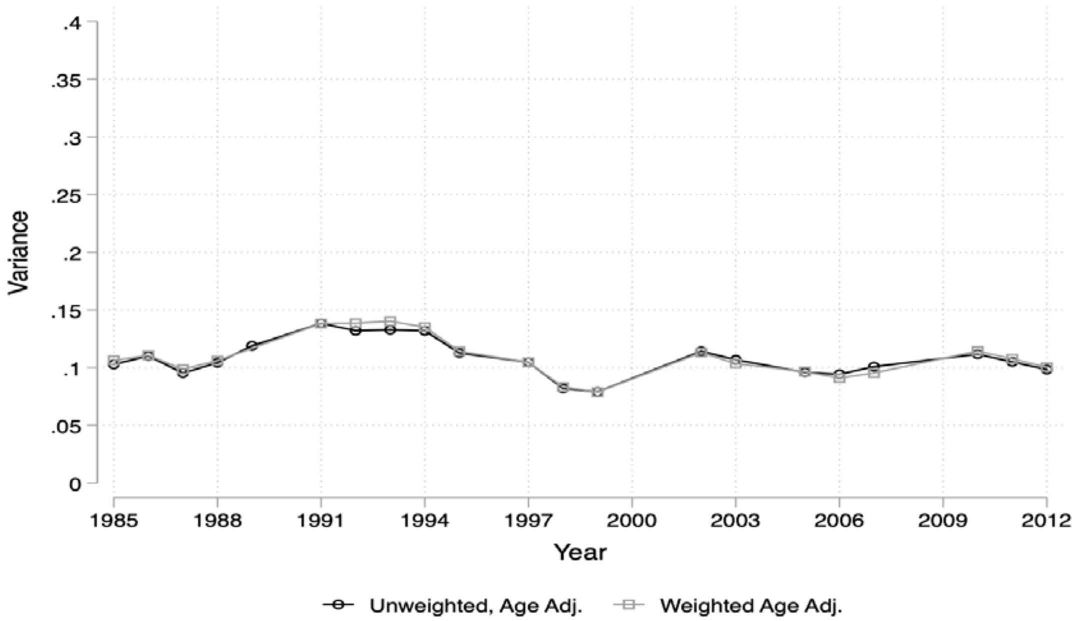
Notes: Authors' calculations based on SIPP. Earnings are trimmed at the bottom and top 1% of the full distribution of men age 25 to 59 with positive earnings. Non-imputed sample uses the sum of non-imputed components of earnings for individuals wit at least one non-imputed component. Imputed earnings using SIPP imputed total earnings for individuals with at least one non-imputed earnings component.

Figure 7: Volatility, Age Adjusted, Full Sample and Sample Present in Survey Panel, 1% Trim Sample, SIPP GSF



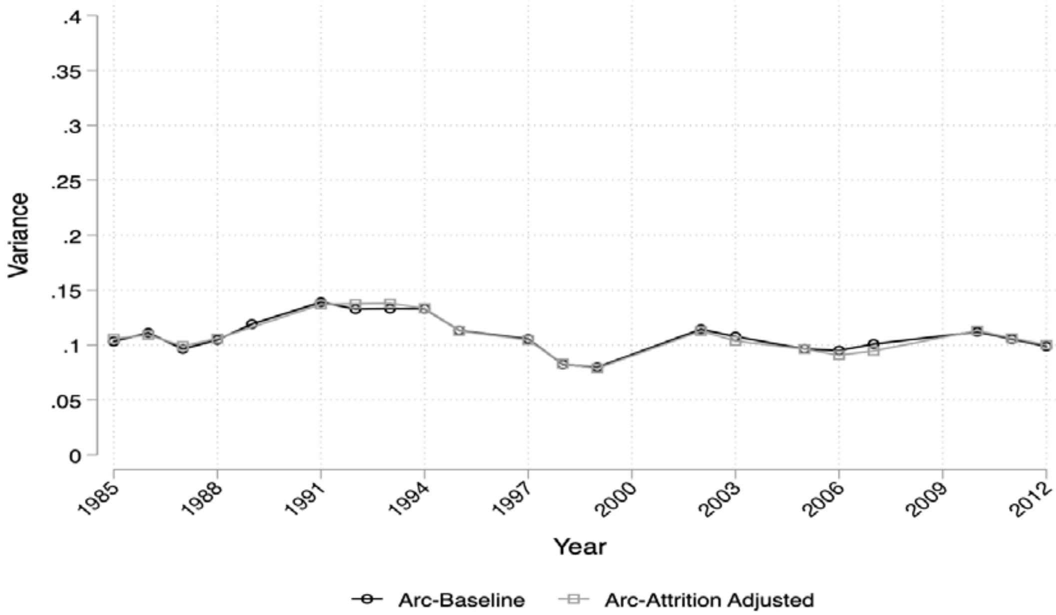
Notes: Authors' calculations using SIPP GSF. Earnings are trimmed at the bottom and top 1% of the full distribution of mean age 25 to 59 with positive earnings. Full sample pools all panels together in each calendar year. Survey Panel sample limits the full sample to include on SIPP panel years, and only individuals who participate in the panel. Only administrative earnings are used.

Figure 8: Volatility, Age Adjusted, With and Without Survey Weights, 1% Trim Sample, SIPP Survey



Notes: Authors calculations using SIPP. Earnings are trimmed at the bottom and top 1% of the full distribution of men age 25 to 59 with positive earnings. Weighted with calendar year weights.

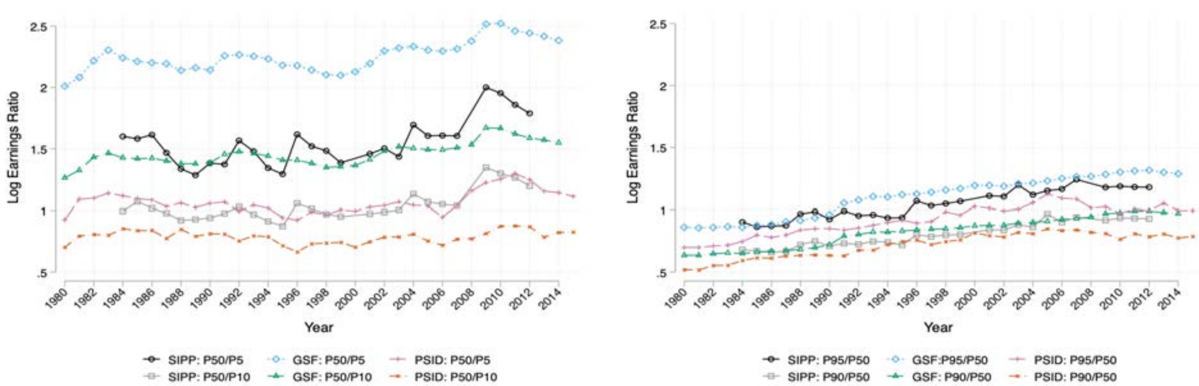
Figure 9: Volatility, Age Adjusted, With and Without Inverse Probability Weights, 1% Trim Sample, SIPP Survey



Notes: Authors calculations using SIPP. Earnings are trimmed at the bottom and top 1% of the full distribution of men age 25 to 59 with positive earnings. Weighted with calendar year weights multiplied by inverse probability of remaining in the sample between $t-1$ and t.

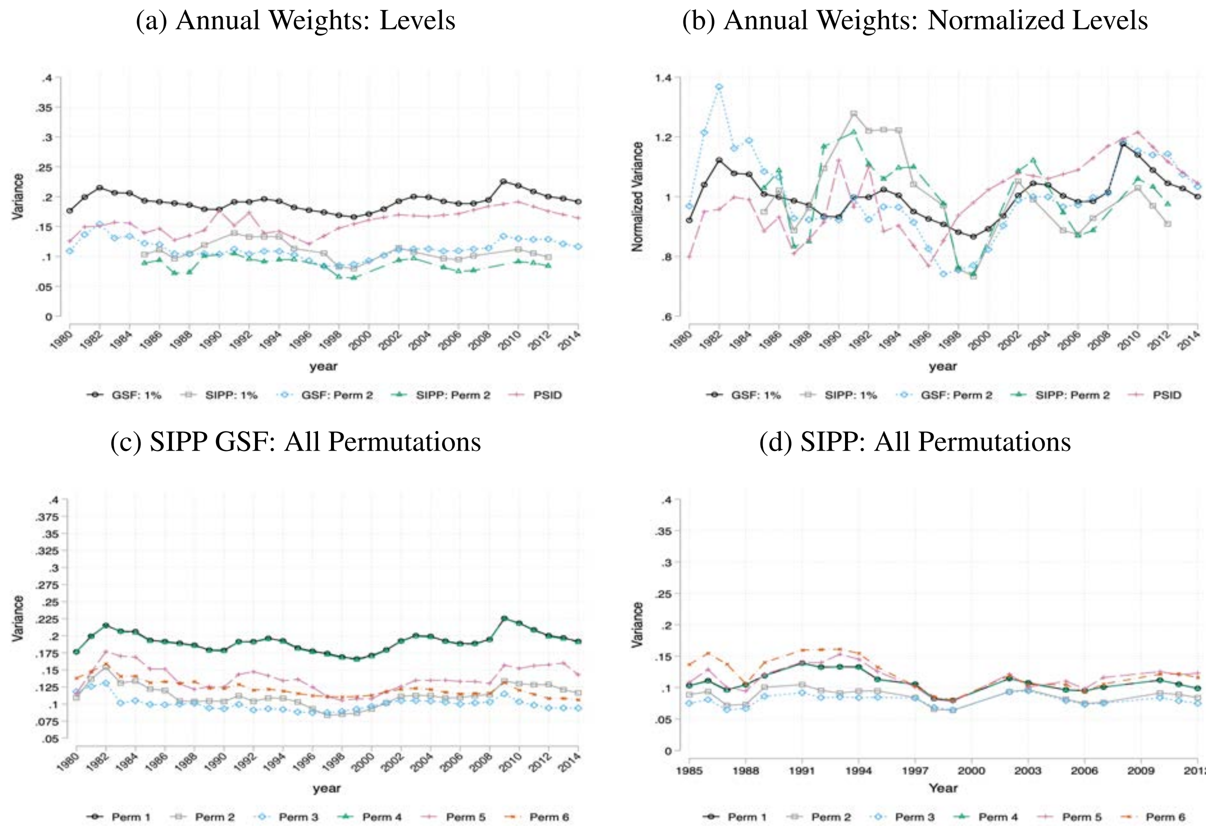
Figure 10: Cross-Sectional Log Earnings Differences Between Percentile Points

(a) Bottom End (b) Top End



Notes: Earnings are trimmed at the bottom and top 1% of the full distribution of men age 25 to 59 with positive earnings in all datasets. Total labor earnings are reported in all datasets.

Figure 11: Volatility, 1% Trim Sample, Weighted to PSID



Notes: Authors' calculations based on SIPP GSF. PSID estimate from Moffitt and Zhang (2020). PSID Earnings are trimmed at the bottom and top 1% of the full distribution of men age 25 to 59 with positive earnings. SIPP GSF earnings are trimmed at the PSID top and botttom 1%, and at the min/max of two-year average earnings, separately by year. SIPP GSF weighted using inverse share of individuals in each PSID ventile. Perms 2 and 3 are not age-adjusted, Perms 5 and 6 are. Perms 2 and 5 use annual weights, and Perms 3 and 6 use year 2000 weights.

Table 1: SIPP Survey Sample N

Year	N of Men 25-59 Jan of t-1 (B)	N of (B) + Jan t-1 - Dec t (C)	N of (C) + Not missing labor force variables Jan t-1 - Dec t (D)	N of (D) + One non-imputed component per month or not working + Pos. imputed earnings t-1 & t (E)	N of (D) + One non-imputed component per month or not working + Pos. non-imputed earnings t-1 & t (F)
1985	11657	5664	4792	3363	3377
1986	7764	3938	3478	2430	2450
1987	6642	4112	3594	2495	2535
1988	6636	5550	4888	3427	3467
1989	6861	1391	1231	885	899
1991	12678	9466	8881	6777	6845
1992	8491	6408	6019	4417	4480
1993	11443	8739	8122	5944	5944
1994	11648	8677	8058	5709	5785
1995	11042	2140	2011	1415	1430
1997	10791	7091	6089	3623	3623
1998	19603	13167	11224	6658	6658
1999	17691	9786	8356	4954	4954
2002	20835	11654	9764	4977	4977
2003	15943	2824	2408	1309	1309
2005	25188	15286	13332	8883	8883
2006	21868	6488	5813	3901	3901
2007	20581	1518	1375	902	902
2010	21739	12560	10797	6986	6986
2011	19856	11525	9837	6256	6256
2012	18281	10774	9295	5862	5862

Notes: Authors calculations from SIPP survey.

Table 2: Demographic Characteristics in the SIPP GSF and Survey Samples

	All	Matched	Volatility Samples GSF	Volatility Samples SIPP
<High School	0.185	0.165	0.141	0.101
High School	0.305	0.302	0.305	0.310
Some College	0.264	0.273	0.283	0.283
College	0.155	0.164	0.172	0.185
College+	0.090	0.096	0.100	0.121
White	0.725	0.750	0.771	0.834
Black	0.116	0.108	0.099	0.064
Other	0.053	0.050	0.046	0.044
Hispanic	0.107	0.092	0.083	0.058
Age	40.380	40.720	40.150	39.404

Notes: Authors' calculations based on SIPP GSF and SIPP. Full sample is all men age 25 to 59. Matched sample is all men age 25 to 59 who can be matched to administrative records. Volatility Sample GSF is all men age 25 to 59 with positive earnings in the SIPP GSF. Volatility Sample SIPP is all mean age 25 to 59 with at least one non-imputed earnings component and with positive earnings in the SIPP survey data.

Table 3: Selected Earnings Percentile Points: SIPP GSF and SIPP Survey Data

Year	P1	P5	P10	P25	P50	P75	P90	P95	P99	N
SIPP										
1984	1,398	8,379	14,909	26,307	39,411	55,455	76,738	94,568	155,647	6,131
1985	1,210	9,006	14,266	25,660	39,160	54,787	75,078	90,238	150,950	2,957
1986	423	7,100	13,891	25,561	40,675	57,543	77,742	95,017	151,605	3,065
1987	1,197	8,721	15,042	26,151	40,733	57,280	77,512	95,766	150,045	4,159
1988	987	8,393	14,956	25,477	39,161	57,402	80,902	103,185	156,895	4,322
1990	944	7,492	13,444	23,669	37,827	55,669	77,797	98,054	149,043	8,424
1991	914	7,356	13,151	23,774	37,498	55,789	78,231	101,987	145,904	5,428
1992	1,423	6,885	12,834	23,253	37,303	54,870	78,024	98,031	140,953	7,346
1993	584	5,994	12,586	23,783	38,014	56,030	79,601	99,240	138,914	7,325
1996	1,461	8,268	14,906	24,750	39,483	58,912	86,289	110,306	236,896	5,298
2001	1,608	9,696	16,503	27,005	42,538	64,624	95,877	123,628	303,540	7,346
2004	1,888	8,820	15,535	27,088	43,764	67,592	100,150	129,402	309,243	12,522
2009	1,373	6,448	12,540	24,410	42,076	67,951	103,258	132,221	368,604	10,418
GSF										
1980	951	5,693	12,020	25,770	42,550	60,070	80,420	100,700	165,900	82,000
1984	633	4,219	9,517	22,790	39,670	58,300	76,030	93,870	173,200	91,000
1985	683	4,359	9,628	22,910	39,810	58,830	77,100	95,670	178,800	95,000
1986	701	4,446	9,670	22,920	40,170	59,590	78,420	97,240	180,900	97,500
1987	690	4,469	9,860	22,840	40,070	59,770	78,580	99,140	195,400	101,000
1988	747	4,690	10,030	22,850	39,850	59,740	78,900	99,590	202,600	104,000
1990	705	4,523	9,618	22,020	38,510	58,480	79,170	100,900	202,900	109,000
1991	594	3,909	8,719	20,890	37,390	57,860	82,520	107,200	204,900	110,000
1992	581	3,896	8,578	20,860	37,600	58,620	83,960	110,400	214,300	112,000
1993	520	3,903	8,624	20,650	37,220	58,510	84,790	112,400	218,800	114,000
1996	606	4,332	9,364	21,690	38,260	60,130	88,610	117,900	265,600	119,000
2001	605	4,780	10,450	24,740	42,940	67,550	103,200	141,900	318,100	124,000
2004	516	4,142	9,503	23,940	42,770	68,610	105,300	143,900	328,000	122,000
2009	387	3,325	7,749	21,340	41,190	68,650	108,600	148,600	337,500	118,000
2014	476	3,831	8,813	22,060	41,530	69,590	109,700	150,600	359,800	115,000

Notes: Authors calculations on SIPP survey and SIPP GSF. Sample is men age 25 to 59 with positive earnings in year t. SIPP is limited to men with non-imputed earnings. Constant 2010 dollars deflated using the PCE.

A SIPP Rotation Groups

Table A1: SIPP Rotation Groups In Each Volatility Year

SIPP Panel	First Ref. Month	Last Ref. Month	Volatility Year t	# Rot Groups in Estimates
1984	Jun-83	Jun-86	1985	4
1985	Oct-84	Jul-87	1986	4
1986	Oct-85	Mar-88	1987	4
1987	Oct-86	May-89	1988	4
1988	Oct-87	Dec-89	1989	1
1989	Jan-89	Dec-89	1990	0
1990	Oct-89	Aug-92	1991	4
1991	Oct-90	Aug-93	1992	4
1992	Oct-91	Aug-94	1993	4
1993	Oct-92	Dec-95	1994	4
			1995	1
1996	Dec-95	Feb-00	1996	0
			1997	2
			1998	4
			1999	3
			2000	0
2001	Oct-00	Dec-03	2001	0
			2002	4
			2003	1
2004	Oct-03	Dec-07	2004	0
			2005	4
			2006	4
			2007	1
2008	May-08	Dec-13	2008	0
			2009	0
			2010	4
			2011	4
			2012	4

Notes: Dates determined from SIPP survey complete technical documentation for each panel.

B Reweighting Methods

B.1 Reweighting Earnings Distribution

The method used to reweight inequality in the SIPP and the SIPP GSF to the PSID in earnings is straightforward, and largely described in the text. We first define a sample in the PSID. This sample consists of men, age 25 to 59, with positive earnings between the 1st and 99th percentiles of the positive earnings distribution in t and $t-1$. We further limit the sample to men who are heads of household and who are followed longitudinally by the PSID. Using this sample, we estimate two-year average earnings over t and $t-1$, and estimate the 19 cutpoints associated with earnings ventiles.

In both the SIPP survey and the SIPP GSF, we begin with the baseline samples described above. We trim the earnings distribution of this sample at the 1st and 99th percentiles of the PSID, forcing annual earnings to have the same range in all datasets. We then estimate two-year average earnings on this sample, and again trim two-year average earnings on the minimum and maximum of the PSID for each year t. This results in a sample that has the same range of annual earnings in t and $t-1$, and the same range of two-year average earnings.

For each year t, we divide the sample in the SIPP survey and SIPP GSF, respectively, into 20 groups based on the 19 earnings ventile cutpoints from the PSID. For each year t, we define a weight given as $0.05/(n_j/N)$ where n_j is the sample size of each earnings bin, $j = \{1, 2, ..., 20\}$, and N is the full sample size. That is, we weight using the inverse share of observations in each earnings bin. The result is a weighted variance of earnings where each of the 20 earnings bins contributes an equal share to the total variance.

B.2 Reweighting Age-Adjusted Earnings Distribution

For age-adjusted volatility, the process is similar, but we use the age-adjusted earnings distribution. This introduces additional complications. We begin by defining the same sample in the PSID as described above. We then age-adjust log earnings using a quadratic in age, as given in Equation 4.

$$y_i = \beta_0 + \beta_1 age_i + \beta_2 age_i^2 + u_i \tag{4}$$

This results in age-adjusted log earnings. To estimate age-adjusted earnings, we use exponentiate predicted earnings for each individual i to estimate $\hat{U}_i$, or residual earnings for individual i. Following the procedure outlined in Wooldridge (2010), we estimate predicted log earnings ($\hat{y}_i$) from Equation 4. Residual earnings are then

$$\begin{aligned} \hat{U}_i &= r_i - \overline{r}_i \\ r_i &= (Y_i - \overline{e^{\hat{u}_i}} e^{\hat{y}_i}) \end{aligned} \tag{5}$$

where $e^{\hat{y}_i}$ is exponentiated predicted log earnings for individual i, $\overline{e^{\hat{u}_i}}$ is the exponentiated residual from Equation 4 averaged across all i individuals, and $\overline{r}_i$ is the average residual averaged across all i individuals. The result is age-adjusted residual earnings ($\hat{U}_i$), that are forced to be mean zero for all datasets due to the mean subtraction. We then average this measure of earnings over years t and $t-1$ to create the ventile cutpoints in the PSID, that is, we estimate $(\hat{U}_{it} + \hat{U}_{i,t-1})/2$ and use this quantity to define the earnings ventiles. We do the same procedure in the SIPP survey and the SIPP GSF.

What follows is identical to the non-age-adjusted version, except that we weight age-adjusted earnings changes using the age-adjusted earnings distribution instead of weighting earnings changes using the earnings distribution.