Recent Trends in the Variability of Men's Earnings: Evidence from Administrative and Survey Data

Michael D. Carr* Emily E. Wiemers†

December 20, 2017

Abstract

Despite the rise in cross-sectional inequality since the late 1990s, there is little consensus on trends in earnings instability during this period. Even trends in simple measures of earnings volatility appear to differ substantially between administrative and survey data. Using consistent samples and methods in administrative earnings data matched to the Survey of Income and Program Participation (SIPP GSF) and survey data from the Panel Study of Income Dynamics (PSID), we examine earnings instability for men from 1978 through 2011. We use a variety of methods, from simple measures of earnings volatility to more complex error components models of earnings. In contrast to the apparent inconsistency in trends across administrative and survey data in the existing literature, we find similar increases in volatility in the SIPP GSF and the PSID. Analysis from the SIPP GSF suggests that the recent increase in inequality appears to be driven by increases in the variance of both the permanent and the transitory component of earnings. Results from models decomposing inequality in the PSID are not entirely consistent across methods, with one method pointing to very large increases in the variance of permanent earnings since the early 2000s and other methods pointing to increases in the variance of both the permanent and transitory component of earnings. We point to very rapid increases in inequality in the PSID during the 2000s as one reason why results from the PSID may be more sensitive to methodological choices.

*Department of Economics, University of Massachusetts Boston. Email: michael.carr@umb.edu

†Department of Economics, University of Massachusetts Boston. Email: emily.wiemers@umb.edu.

We gratefully acknowledge funding from the Russell Sage Foundation for this project (award #83-15-09).

1 Introduction

Increasing cross-sectional inequality must be the result of a widening distribution of permanent earnings, an increase in transitory earnings variability, or a combination of both. Because of the rapid rise in inequality in the United States beginning in the late 1970s and the explicit link between inequality and earnings instability, there is now a sizable literature devoted to measuring earnings instability and decomposing the variance of earnings into its permanent and transitory components. The early literature, which primariliy relies on data from the Panel Study of Income Dynamics (PSID), is in general agreement that earnings variability increased from the 1970s through the mid 1980s and declined into the early 1990s, with approximately half of the increase in earnings inequality attributable to increasing transitory earnings variances and half to a widening of the distribution of permanent earnings.[1] In more recent work that uses a more varied set of administrative and survey data sources, a lack of consensus has emerged about trends both in more recent years and during earlier years where there was previously a consensus.

We argue that the lack of consensus in trends in instability is driven largely by the introduction of new data sources, particularly administrative earnings data sources, and by sample selection and methodological choices. Despite the fact that comparisons of results across papers are frequently made, many analyses include simultaneous changes in data source, sample definitions, and methods which results in a set of estimates that are simply not comparable to each other. For example, Shin and Solon (2011) measure earnings volatility for working-age men using the PSID and find that earnings volatility increased between the 1970s and early 1980s, stabilized and decreased slightly through the 1990s, and increased again in the late 1990s and 2000s. In contrast, Sabelhaus and Song (2009, 2010) show declining earnings volatility from the 1980s through 2005 in administrative earnings data, but pool working-age men and women together. Ziliak, Hardy, and Bollinger (2011) use the panel component of the Current Population Survey and find that volatility

[1] See Gottschalk and Moffitt (2009), Gottschalk et al. (1994), Haider (2001), Moffitt and Gottschalk (2002, 2012), Shin and Solon (2011) for estimates using the PSID that span this period.

increased from the 1970s through the mid 1980s and stabilized thereafter, but use a sample of 16 to 60 year olds instead of the more customary sample of 25 to 59 year olds. The sample in Ziliak, Hardy, and Bollinger (2011) includes both men and women but trends are estimated separately by gender and confirm the declining volatility for women implied by Sabelhaus and Song (2010).[2] That is, as new data have been added to the literature on earnings volatility so have new sample definitions and methods.

The same issues make comparisons difficult in the literature that decomposes earnings inequality into its permanent and transitory components. Moffitt and Gottschalk (2012) show that, despite a surge in transitory earnings variances in the late 1990s and early 2000s in the PSID, total inequality is still roughly 50/50 transitory and permanent. But work decomposing inequality in administrative data show that the overwhelming majority of inequality is attributable to permanent earnings variances (Debacker et al., 2013, Kopczuk, Saez, and Song, 2010). While there are important methodological differences across papers which may contribute to differences in results, work using the PSID and the administrative data also decompose fundamentally different measures of earnings: Moffitt and Gottschalk (2012) decompose within-group earnings inequality (within age, education, and race) while Kopczuk, Saez, and Song (2010) and Debacker et al. (2013) decompose earnings inequality adjusted only for age. Given the inability to adjust for race and education in most administrative data, it is not surprising that the role of permanent earnings inequality appears bigger in these data.

In this paper, we provide estimates of volatility and error components earnings decomposition models using consistent methods and sample selection criteria on both administrative earnings data and the PSID. We make use of an underutilized dataset which links data from the Survey of Income and Program Participation (SIPP) to administrative earnings data. The data, known as the SIPP Gold Standard File (SPP GSF), draws its administrative earnings from the same sources as

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

other administrative data, but includes demographic and human capital data from the SIPP survey. Because the SIPP GSF contains information on race, gender, and educational attainment, we are the first to be able to apply methods to administrative data consistent with those typically applied to survey data – most notably, the ability to decompose *within group* inequality into its permanent and transitory components.

Relative to the existing literature, our analysis has two advantages. First, we measure earnings variability using simple measures of volatility and more complex error components models with consistent samples on two very different datasets. Insofar as we use both administrative and survey data, our paper follows Dahl, DeLeire, and Schwabish (2011), Celik et al. (2012), and Monti and Gathright (2013), but we extend these papers both in the time period under consideration and by estimating not only earnings volatility but also error components models of earnings which allow us to understand the reasons behind recent increases in cross-sectional inequality. Second, and perhaps even more importantly, we extend our analysis of earnings through 2011 (2012 in the PSID). Given the disagreements in the literature on the trends in earnings instability after the mid 1990s and the increases in inequality in recent years, this is an important contribution.

Across data sources, we find remarkable consistency in trends in earnings instability across the entire time period, including a substantial increase in volatility and other measures of transitory earnings variances in the 2000s. The results from the SIPP GSF also suggest that the recent increase in the transitory variance is not entirely cyclical, a feature that is hard to detect in the PSID due to its comparatively small cross-sectional sample sizes. However, we do find some discrepancies across methods using the PSID in the relative importance of increases in the variance of transitory earnings and a widening permanent earnings distribution to increases in inequality in recent years. These emerge only after 2000 in the PSID and seem to be linked to very rapidly rising inequality in the PSID during this period. We find no such discrepancies in the SIPP GSF where all of the methods show increases in inequality in the late 1990s and 2000s are driven by increases in the variance of both the permanent and transitory components of earnings.

The overall consistency in the trends of increasing earnings instability across administrative and survey data and across methods within each data source suggests that methodological and sample selection decisions are at the root of the discrepant results in the literature. The similarity of trends across the data sources is important because each data source has particular advantages. SIPP survey-linked administrative data has large cross-sectional sample sizes, no attrition, and no top coding, coupled with information on education, race, and gender allowing for the ability to analyze earnings instability across subgroups and across parts of the earnings distribution. The PSID includes information on labor supply characteristics over and above annual earnings allowing for an understanding of whether recent increases in earning instability are a result of changes in hours, unemployment, job tenure, or occupational structure. Each of these areas is important for understanding the causes and consequences of rising inequality. However, we exercise some caution in interpreting trends in the PSID after the 2000s where trends in inequality deserve further investigation.

2 Model of Earnings

Following Shin and Solon (2011), we begin by outlining a simple model of earnings that splits earnings into orthogonal permanent and transitory components:

$$y_{it} = \mu_i + \nu_{it} \tag{1}$$

where earnings of individual i in year t is the sum of permanent earnings (μ_i) and a transitory earnings shock (ν_{it}) which, in this simple model, is assumed to be independent of μ_i. The variance of earnings is then the sum of the variance of the permanent and transitory components of earnings:

$$\text{Var}(y_{it}) = \sigma^2_\mu + \sigma^2_{\nu t}. \tag{2}$$

As we outline in the introduction, and describe more fully in what follows, earnings volatility measures the variability of changes in the left hand side of Equation (1) differenced over short time periods, as given in Equation (3).

$$\text{Var}(y_{it} - y_{it-\tau}) = \sigma^2_{\nu t} + \sigma^2_{\nu t-\tau}. \tag{3}$$

where, for comparability between data sources, $\tau = 2$. Straightforwardly, if the variance of the transitory component is assumed to be constant over time $\sigma^2_{\nu t} + \sigma^2_{\nu t-\tau} = 2\sigma^2_{\nu}$. Volatility is a measure of the transitory variance of earnings. However, in the presence of time trends in the returns to permanent characteristics, time trends in the transitory earnings variance, long lasting permanent earnings shocks, and serial correlation in transitory variances, measures of earnings volatility and measures of the transitory variance from error components models of earnings will not yield the same results. In particular, measures of earnings volatility will include some of the variance in the permanent component of earnings. Shin and Solon (2011) argue that earnings volatility is still a useful measure because increases in the variance of the transitory component of earnings are likely to be accompanied by increases in earnings volatility and this measure is less dependent on specific parametric assumptions.

The simplest error components model separates the variance in the permanent and the transitory component of earnings in Equation (1) by using a random effects model to estimate the within and between person variances of medium-term earnings. Moffitt and Gottschalk (2012) call this method window averaging but it is similar in spirit to the method in Kopczuk, Saez, and Song (2010). This technique overstates the permanent component of earnings particularly in the presence of serial correlation in transitory shocks.

Over time, the literature has developed to model increasingly flexible specifications of earnings dynamics. Among the important features that have been captured are individual specific growth factors in permanent earnings, permanent earnings that evolve over the life cycle, serial correlation in transitory earnings, age-related heteroskedasticity in transitory earnings, and year-specific factor loadings for both permanent and transitory earnings (Baker and Solon, 2003, Doris, O'Neill, and Sweetman, 2011, Haider, 2001, Moffitt and Gottschalk, 2012).

We outline a model which borrows elements from Haider (2001), Baker and Solon (2003), and Moffitt and Gottschalk (2012), proposed by Doris, O'Neill, and Sweetman (2011) in which log earnings for individual i in birth cohort a at time t, y_{iat} is given by:

$$\begin{aligned} y_{iat} &= q_a p_t(\alpha_i + \beta_i x_{it} + u_{it}) + s_a \lambda_t \nu_{it} \\ u_{it} &= u_{i,t-1} + \omega_{it} \\ \nu_{it} &= \rho \nu_{i,t-1} + \theta \varepsilon_{i,t-1} + \varepsilon_{it} \end{aligned} \tag{4}$$

where $E(\alpha_i) = E(\beta_i) = E(\omega_{it}) = E(\varepsilon_{it}) = 0$, σ^2_α is the variance of α_i, σ^2_β is the variance of β_i, $\sigma^2_{\alpha\beta}$ is the covariance between α_i and β_i, and σ^2_ε is variance of ε_{it}. The specification above implies that $\mu_i = \alpha_i + \beta_i x_{it} + u_{it}$ and incorporates individual specific age-earnings profiles in $\alpha_i + \beta_i x_{it}$ where x_{it} is the age of individual i at time t. Each individual has a different permanent growth rate of earnings which may be correlated with initial earnings. Permanent shocks to earnings arrive randomly and are modeled with a random walk in u_{it}. The variance of ω_{it} is σ^2_ω and $E(u_{i,t-1}, \omega_{it}) = 0$. The transitory component of earnings is characterized by an ARMA(1,1), which is standard in the literature (Baker and Solon, 2003, Haider, 2001, Moffitt and Gottschalk, 2012). We allow for calendar time shifts in the permanent and transitory component with p_t and λ_t, respectively, and cohort specific shifts with q_a and s_a. This model is functionally very similar to Haider (2001) and Moffitt and Gottschalk (2012), differing only by whether the cohort specific factors are estimated as parameters or defined as part of the variables, and that there are multiplicative year factor loadings on both permanent and transitory earnings. We note that the innovation of this paper is not in estimating a new earnings model but rather in using data that allow for us to estimate the same error components model of earnings on administrative and survey data.

3 Data

3.1 Administrative Data

The administrative data for this project come from the Survey of Income and Program Participation Gold Standard File (SIPP GSF).[3] The SIPP is a nationally representative sample of the civilian noninstitutionalized population of the U.S. that began in 1984. There have been 14 SIPP panels since 1984 and each panel lasts between two and four years. Within panels, the SIPP is longitudinal but each panel draws a new nationally representative sample of 14,000 to 52,000 households. The SIPP GSF links each individual in a SIPP household in the 1984, and 1990 – 2008 SIPP panels to their IRS and SSA earnings and benefits records.

The Census Bureau creates the data using a set of standardized extracts of SIPP survey variables common to multiple panels which are then linked with administrative earnings and benefits records.[4] Earnings histories come from the Detailed Earnings Records (DER), which are co-maintained by the SSA and the IRS and include FICA taxable and non-taxable earnings, self-employment earnings, and deferred earnings, the sum of which is not topcoded. If all earnings values are zero, then the individual had zero earnings for that year. The complete administrative SSA and IRS earnings history is linked to every individual that is ever surveyed in any of the included SIPP panels, but for earnings before 1978 only FICA taxable earnings are available. For example, if a 55 year old individual is surveyed in the 2000 panel, the SIPP GSF will include that individual's non-topcoded earnings from 1978 through 2000 *and* from 2001-2011.[5] In addition

[3]This analysis was first performed using the SIPP Synthetic Beta (SSB) which was funded by the U.S. Census Bureau and SSA, with additional funding from NSF Grants #0427889 and #0339191 using 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.

[4]Benedetto, Stinson, and Abowd (2013) describe the creation of the data in great detail.

[5]Their FICA taxable earnings back to 1951 are also available but we do not use the series of only FICA taxable earnings. Unlike the administrative data used in Kopczuk, Saez, and Song (2010), we do not have information on the

to the administrative earnings records, the SIPP GSF has basic demographic and human capital variables, marriage histories, fertility histories, as well as self-reported earnings from the SIPP survey. Though administrative earnings data are available outside of the SIPP panel window, variables collected in the SIPP panels cover only the years of the individual's SIPP panel.

Missing data in the SIPP GSF are multiply imputed. Missing data can arise either because the SIPP survey participant refused to answer a specific question or because the SIPP respondent could not be matched to administrative data.[6] The public use SIPP has missing observations that are imputed using a hot-deck method but missing observations in the SIPP GSF are imputed using a substantially more sophisticated multiple imputation method. The Census Bureau advises against excluding imputed observations and we follow this recommendation.[7] Because of the imputation process, the SIPP GSF consists of four implicates over which we average.

3.2 Survey Data

Survey data come from the Panel Study of Income Dynamics (PSID), the most commonly used dataset in the literature. The PSID is a household based panel survey first fielded in 1968 at which time it was representative of the population of households in the United States. The PSID has followed not just the original respondents but has added to its sample newly born (or adopted) children of those respondents. The PSID was conducted annually until 1997 and since then interviews have been done on a biennial basis. In each year respondents report their labor earnings in the calendar year prior to the interview and labor earnings are topcoded.

Critical to the study of volatility, and one of the primary motivations for using the PSID, is the

quarter in which the FICA earnings cap was reached and so we are unable to impute total earnings above the FICA earnings cap.

[6]The match rate 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. Beginning with the 2004 panel, the match rate increased to around 90% because the Census Bureau changed it's matching procedures removing the necessity to explicitly ask for social security numbers.

[7]See Abowd and Stinson (2013) for a discussion of imputations in the context of measurement error. Missing data must be imputed because patterns of missing observations can be used to link the SIPP GSF back to the regular SIPP survey, which the Census Bureau considers to be a violation of the privacy of the SIPP participants.

fact that the wave-to-wave response rate in the PSID is among the highest of any national survey in the world, with a rate of 95% - 98% in almost every wave since 1968 (Schoeni et al., 2013). Despite this fact, attrition over longer time horizons is meaningful (about 35% by 2009), and is correlated with both income and changes in life circumstances (Fitzgerald, Gottschalk, and Moffitt, 1998a) though the PSID remains remarkably representative when weighted properly (Fitzgerald, 2011, Fitzgerald, Gottschalk, and Moffitt, 1998b, McGonagle et al., 2012). We use PSID data through 2013, the latest year for which data are available. Because earnings are collected about the year prior to the survey year, in what follows we label the date of the PSID with the year to which earnings apply (i.e. 2013 is labelled 2012).

3.3 Samples and Measures

In both the SIPP GSF and PSID, we use a sample of men ages 25 to 59 with non-zero earnings. In the PSID we restrict our sample to household heads because these are the individuals for whom earnings are available throughout the survey. The headship restriction eliminates 8% of the PSID sample. We cannot make a similar sample restriction in the SIPP. In the PSID we exclude all of the oversamples (Survey of Economic Opportunity, Latino, and Immigrant). We also include only PSID sample members who are followed longitudinally – that is we include only the original 1968 sample members and their biological or adopted children.[8]

In both the SIPP GSF and in the PSID we measure the variability of labor earnings. In the PSID, we exclude the top and bottom 1% of positive earnings which has the dual advantage of reducing the impact of outlier earnings changes and eliminating topcoded earnings. For consistency, we follow the same procedure in the SIPP GSF though in these data earnings are not topcoded. To measure volatility, we age-adjust the changes in log earnings in each year using a quadratic in age.

[8]Our restriction of the PSID sample to include only household heads from the original 1968 sample and their biological or adopted children differs from Haider (2001), Shin and Solon (2011), and Moffitt and Gottschalk (2012) who include all PSID heads in their samples. We discuss the differences in samples and show that our results are robust to alternative sample choices in Appendix (A.2).

For our error components models, we measure earnings using the residuals from a regression of log earnings on a quadratic in age, four education categories (high school or less, some college, college, and advanced), and four race groups (white non-Hispanic, Black non-Hispanic, Hispanic, and Asian/Pacific Islander) and their interactions.

These restrictions result in a dataset with annual sample sizes between 95,125 and 155,591 in the SIPP GSF and between 1,038 and 2,551 in the PSID.

4 Methods

As outlined in Section (2), much of the literature that examines overall earnings volatility uses simple measures. We rely primarily on the standard deviation of two year changes in log earnings age-adjusted separately by year, or the square root of Equation (3), as used in Shin and Solon (2011). For the SIPP GSF through the entire period and the PSID through 1996 we use overlapping two-year changes (1980 - 1982, 1981 - 1983) and for the PSID after 1996 we use non-overlapping two-year changes (1996 - 1998, 1998 - 2000).

An alternative is to measure volatility using the standard deviation of the arc percent change in earnings (Dahl, DeLeire, and Schwabish, 2011, Ziliak, Hardy, and Bollinger, 2011). The arc-change is calculated as

$$\text{Volatility}_t = \sqrt{Var\left\{100 * \frac{y_{it} - y_{i,t-2}}{\frac{|y_{it}|+|y_{i,t-2}|}{2}}\right\}} \tag{5}$$

where we age-adjust the percent change in earnings in each year t. The main advantage of the arc-change is that individuals with zero earnings can be incorporated in a straightforward manner. The arc-change method also reduces the impact of outlier earnings changes by bounding percent changes between -200% and 200%. Though considering the role of individuals with zero earnings in earnings variability is not the focus of this paper, we include results using the arc-change measure with and without men with zero earnings to show that the overall trends are robust to an alternative method and to the inclusion of men with zero earnings.

We use three methods to decompose the variance in earnings into its permanent and transitory components. First, we follow Moffitt and Gottschalk (2012) and apply a window average method in which, for each year t, the log earnings residuals in the calendar year window $[t-w, t+w]$ are averaged for each individual i to obtain an estimate of the individual's permanent earnings. We use a 9-year window ($w = 4$). The difference between the log earnings residual for individual i in year t and individual $i's$ average residual, $y_{it} - \bar{y}_i$, is an estimate of the transitory earnings deviation in year t. We follow Gottschalk et al. (1994) and use a standard random effects decomposition to calculate the variance of the transitory component. The transitory variance from the window averaging method is a consistent estimate of the transitory variance under Equation (1) as long as the error structure of earnings follows the assumptions necessary for random effects variance estimations, namely, that μ_i and ν_{it} are independent of each other. However, in the presence of more complicated error structures of earnings, the variance from the window averaging method is not a consistent estimator of the transitory variance of earnings.

Second, we estimate Equation (4), which allows for a more complex error structure of earnings, using GMM. The moment conditions for the variances and covariances derived from Equation (4) take the following form:

$$\begin{aligned}\sigma^2_{at} =& \{q_a^2 p_t^2(\sigma^2_\alpha + \sigma^2_\beta x^2_{at} + 2\sigma_{\alpha\beta}x_{at} + \sigma^2_\omega x_{at})\} + \\ & \left\{ s_a^2\lambda_t^2 \left(\rho^{2t-2}\sigma^2_{\nu 1} + K\sum_{\omega=0}^{t-2}\rho^{2\omega} \right) \right\} \\ Cov(y_{at}, y_{a(t+s)}) =& q_a^2 p_t p_{t+s}\left\{\sigma^2_\alpha + \sigma^2_\beta X_{at}X_{a(t+s)} + \sigma_{\alpha\beta}\left(X_{at} + X_{a(t+s)}\right) + \sigma^2_\omega X_{at}\right\} + \\ & s_a^2\lambda_t\lambda_{t+s}\left(\rho^{2t+s-2}\sigma^2_{\nu 1} + \rho^2 K\sum_{\omega=0}^{t-1}\rho^{2\omega} + \rho^{s-1}\theta\sigma^2_\varepsilon\right)\end{aligned} \tag{6}$$

where $K = \sigma_\varepsilon(1 + \theta^2 + 2\rho\theta)$.

We normalize the initial time and cohort factor loadings to be equal to one and estimate the parameters of the model using GMM, which minimizes the sum of squared deviations between the observed moments and moment conditions predicted by the model, using an identity weighting matrix and computing robust standard errors. Parameter estimates for the SIPP GSF and the PSID are in Appendix (A.3). The variances can change from year to year because the time-specific factor loadings change and because individuals age. Similar to Baker and Solon (2003) and Moffitt and Gottschalk (2012), in what follows we predict the permanent and transitory components of earnings for the 35 to 44 year old cohort in each year t. Following the literature, we use age-, education-, and race-adjusted residual log earnings (adjusted annually).

Finally, we follow Moffitt and Gottschalk (2012) and use an approximate nonparametric decomposition model that relies on the notion that trends in the autocovariance in earnings across a sufficiently long time horizon can be used to estimate the variance of the permanent component of earnings. Namely, if the time frame across which autocovariances are measured is sufficiently long so that transitory errors are no longer correlated, then the covariance of earnings in time t and time $t-\tau$ can be written as:

$$Cov(y_{iat}, y_{i,a-\tau,t-\tau}) = p_t p_{t-\tau} Cov(\mu_{ia}, \mu_{i,a-\tau}) \tag{7}$$

where τ represents lag-length.

Taking logs yields:

$$\log[Cov(y_{iat}, y_{i,a-\tau,t-\tau})] = \log(p_t) + \log(p_{t-\tau}) + \log[Cov(\mu_{ia}, \mu_{i,a-\tau})] \tag{8}$$

which can be estimated by OLS using a polynomial in age and τ as a nonparametric approximate of the long covariance in permanent earnings. The variance of the permanent component in year t is then estimated by evaluating (8) at $\tau = 0$. We follow Moffitt and Gottschalk (2012) and estimate this model using all lag lengths of 10 or over back to age 20.

5 Results

5.1 Volatility

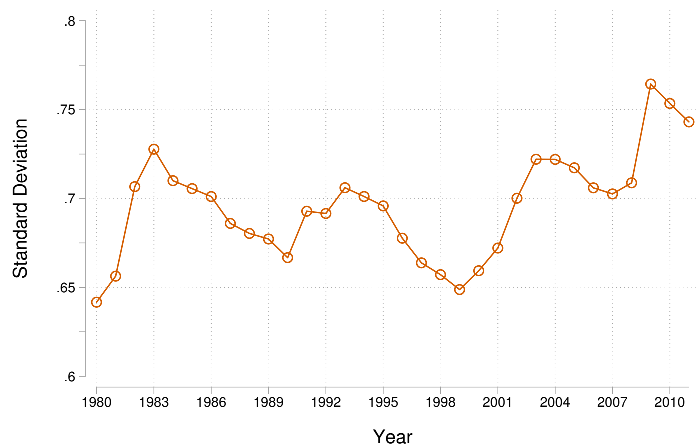
The results of estimating Equation (3) on both the SIPP GSF and the PSID are shown in Figure (1). The trends in volatility after 1980 in the two datasets are similar. The SIPP GSF and the PSID show a large increase in volatility during the early 1980s after which volatility declined until 1999, albeit with an increase in volatility during the recession in the early 1990s. In each series, volatility began to increase again in the late 1990s and continued its upward trend through 2011 (2012) in the SIPP GSF (PSID) when volatility in each case was higher than it was in the early 1980s. Volatility is about 40% higher in the SIPP GSF than in the PSID. We return to the level differences between the SIPP GSF and the PSID in Section (6).

Figure (2) shows volatility in the SIPP GSF, from Figure (1a), with the civilian unemployment rate. Consistent with Shin and Solon (2011), earnings volatility is highly cyclical. There are large increases in volatility during the recessions in the early 1980s, the early 1990s, the early 2000s, and especially during the Great Recession. However, over-and-above the cyclicality, the upward trend during the 2000s is unmistakable: volatility increased in the early 2000s during the recession and remained at this higher level after the recession ended. Volatility then increased sharply again during the Great Recession, though, at least in the SIPP GSF, it has declined during the recent recovery. Although cyclicality is apparent in the PSID, it is much less pronounced. Most of the previous literature on earnings volatility uses data through 2006 and, at that point, it was unclear whether the increase in volatility since the late 1990s was simply a consequence of the recession in the early 2000s (Moffitt and Gottschalk, 2012, Shin and Solon, 2011). Figure (2) shows that the recent increase in earnings volatility is not simply cyclical in nature.

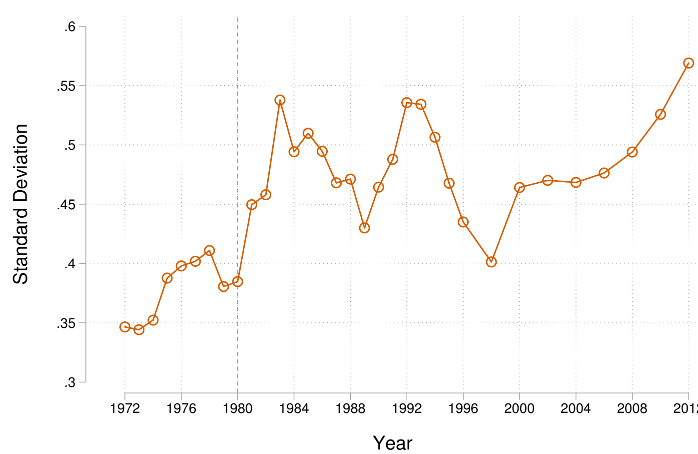
Indeed, one of the advantages of the SIPP GSF is its large samples and consistent series of annual earnings data which make cyclicality easier to distinguish from longer-term trends. The PSID not only has smaller sample sizes, but also contains some idiosyncrasies in its time series.

Figure 1: Earnings Volatility, SIPP GSF and PSID

(a) SIPP GSF

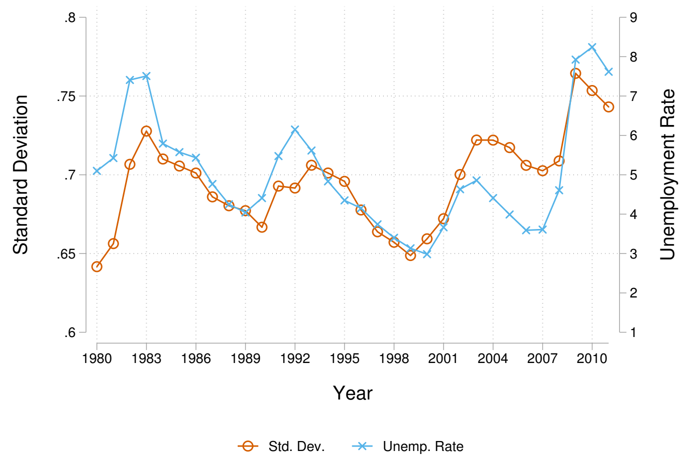


(b) PSID



Author's calculations using the SIPP GSF from 1978 to 2011 reported in Figure (1a) and using the PSID from 1970 to 2012 reported in Figure (1b). Volatility is the standard deviation of the age-adjusted two-year change in log earnings for a sample of men ages 25 to 59 with positive earnings, excluding the top and bottom 1% of annual earnings. Earnings changes age-adjusted separately by year using a quadratic in age. The PSID includes only sample men. The vertical line in 1980 in Figure (1b) represents the year in which the SIPP GSF data begins.

Figure 2: Earnings Volatility and Unemployment, SIPP GSF



Author's calculations using the SIPP GSF from 1980 to 2011. Volatility is the standard deviation of the age-adjusted two-year change in log earnings for a sample of men ages 25 to 59 with positive earnings, excluding the top and bottom 1% of annual earnings. Earnings changes age-adjusted separately by year using a quadratic in age. Unemployment rate is for individuals age 25 and over.

In particular, the PSID also does not allow for the calculation of two-year earnings changes in the years between biennial interviews after 1996 and, during the early 1990s, the PSID switched from in-person interviews to computer-assisted telephone interviews and from human to automated data editing which seems to have temporarily increased earnings volatility over and above that caused by the recession (Shin and Solon, 2011).

One of the benefits of our analysis is that samples and methods are held constant across datasets and we do not rely on previously published estimates of volatility using the PSID. However, the trends in volatility that we estimate using the PSID are very similar to Shin and Solon (2011), with the exception that we find a higher spike in volatility during the recession in the early 1990s. The difference between our estimates and those in Shin and Solon (2011) is attributable to the exclusion of non-sample individuals in the PSID in our sample. In Appendix (A.2), we discuss our sample choice in more depth and show the same figure including non-sample individuals for comparison to Shin and Solon (2011).

We also estimate volatility using the arc-change model outlined in Equation (5). Figure (3) shows the standard deviation of the arc percent change in earnings in the SIPP GSF and PSID, where Figure (3a) trims the top 1% of earnings when including individuals with zero earnings, and trims the top and bottom 1% when excluding individuals with zero earnings. Individuals with zero earnings in consecutive years are assigned a change of zero, as the change cannot be calculated using Equation (3). Overall, the trends in the arc-change are quite close to trends in volatility shown in Figure (1). Earnings instability increases in the early 1980s, declines through the late 1990s and then begins to increase after 1998. These trends are clear in both the SIPP GSF and the PSID. Most of the differences between the two datasets that we highlight above–cyclicality is more obvious in the SIPP GSF and the levels of volatility are higher in the SIPP GSF–hold using this alternative measure of earnings instability over a two-year period.

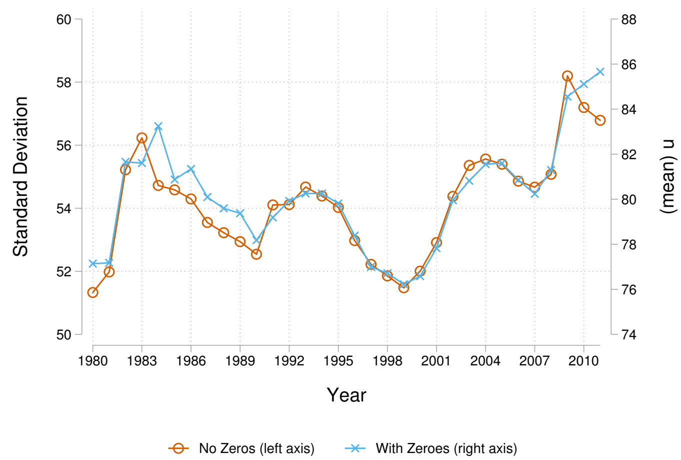
Figure (3) allows us to examine whether the trends that we estimate are robust to the inclusion of men with zero earnings. In the SIPP GSF, the trends excluding and including men with zero earnings are nearly identical though the level of volatility is higher when men with zero earnings are included. In the PSID, the trends in the 1970s through 1990s are also similar and volatility is higher when we include men with zero earnings. However, in the more recent period in the PSID, the upward trend in volatility is less pronounced and perhaps more cyclical when we include men with zero earnings.

5.2 Window Averaging

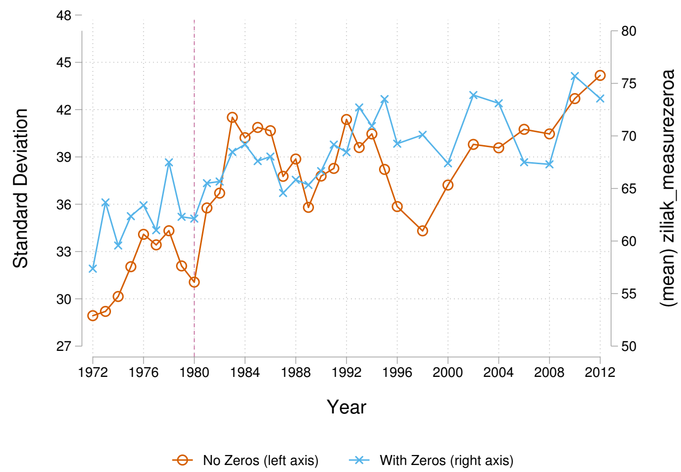
Figure (4) shows the trends in the transitory variance of earnings calculated using the window averaging method in the SIPP GSF and PSID. As with the volatility measures, the trends in the SIPP GSF and in the PSID are remarkably similar with increases in the transitory variance of earnings during the early 1980s, flat or declining transitory variances from the mid 1980s through the mid 1990s and an increase in the transitory variance starting in the late 1990s and continuing through the last year in each sample. Because the window averaging looks at annual deviations from rolling 9-year average earnings, the window averaging series are substantially less cyclical

Figure 3: Arc Percent Change in Earnings, SIPP GSF and PSID

(a) SIPP GSF



(b) PSID



Author's calculations using the SIPP GSF from 1978 to 2011 reported in Figure (3a) and using the PSID from 1970 to 2012 reported in Figure (3b). Volatility is the standard deviation of the age-adjusted two-year arc-change in earnings for a sample of men ages 25 to 59 with positive earnings. Earnings changes age-adjusted separately by year using a quadratic in age. The vertical line in 1980 in Figure (3b) represents the year in which the SIPP GSF data begins. When excluding zero earnings, the top and bottom 1% of earnings are trimmed. When including zero earnings, only the top 1% are trimmed.

than equivalent volatility series. As in Figure (1), the level of the transitory variance is now higher than at any other point in the sample and the increase in the transitory variance since the late 1990s is as large as the increase during the late 1970s and 1980s. Again, the levels in the SIPP GSF are much higher – nearly three times larger in the SIPP GSF than in the PSID. We return to the level differences in Section (6).

These series are most closely comparable to Moffitt and Gottschalk (2012)[9] who estimate trends in the transitory variance of earnings in the PSID through 2000. But, while Moffitt and Gottschalk (2012) [p.218] note that "the variance turns up ... for the year 2000 window ... this upturn is followed by a downturn in the years which follow," we find that the upturn in the early 2000s continued until 2008 (2007) in the PSID (SIPP GSF). Moreover, the trends in the transitory variance using the window averaging method are similar to those using the simpler volatility measure. Moffitt and Gottschalk (2012) could not have predicted the Great Recession which contributes to the recent upward trend in the transitory variance, but the upward trend in the transitory variance starting in the late 1990s does not appear to be purely cyclical.

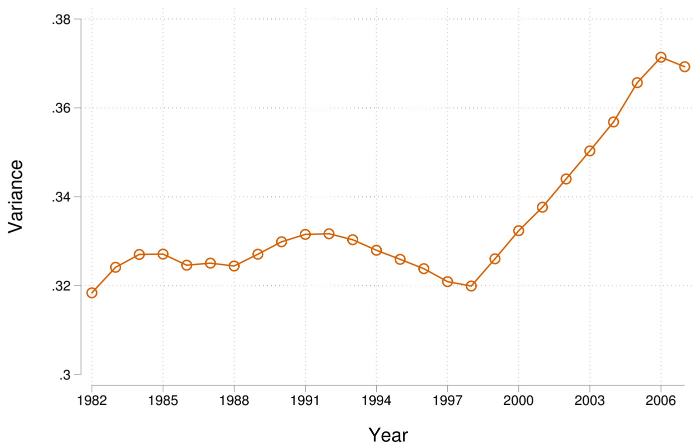
5.3 Error Components Model Results

The primary limitation of the window averaging method is the inability to separate long-lived transitory shocks from changes in permanent earnings. We now turn to the error components model of earnings based on the model presented in Equation (4), where the dependent variable is log earnings adjusted for education, age, and race. As is common in the literature, the estimated coefficients are used to graph annual predicted permanent, transitory, and total variances of earnings in Figure (5a) for the SIPP GSF, and in Figure (5b) for the PSID. The SIPP GSF covers only 1978 to 2011 while the PSID covers 1970 to 2012. For completeness, the coefficient estimates from one of the implicates in the SIPP GSF are reported in Table (A1) and for the PSID in Table (A2).

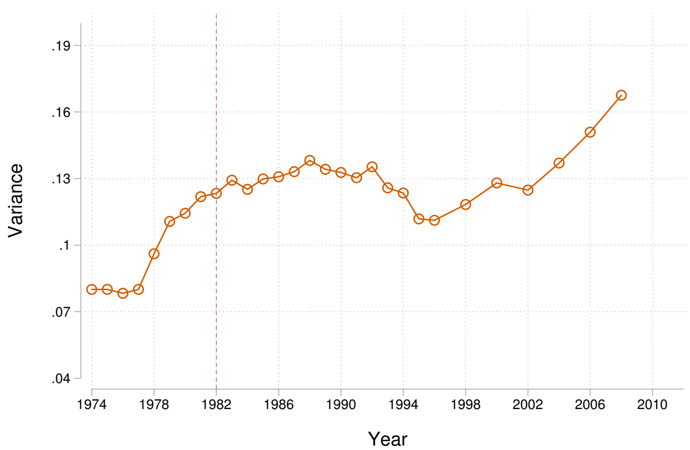
[9] Figure (4b) is nearly identical to Figure 8 in Moffitt and Gottschalk (2012) with the exception that we find a larger decline in the transitory variance of earnings during the 1990s. The differences are due to the exclusion of non-sample individuals. We show our estimates including non-sample men in Appendix (A.2).

Figure 4: Window Averaging

(a) SIPP GSF



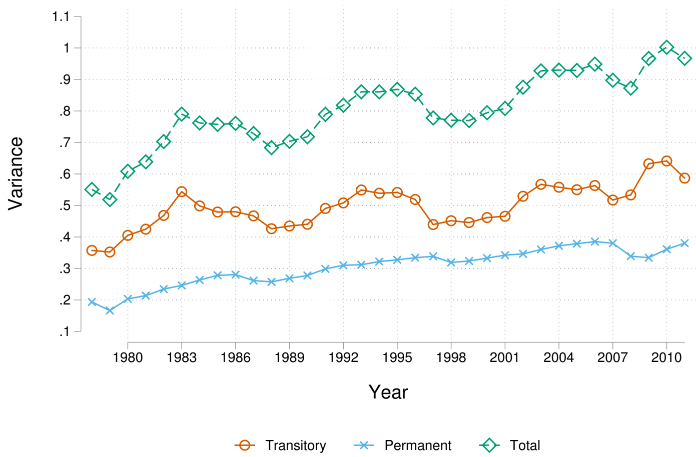
(b) PSID



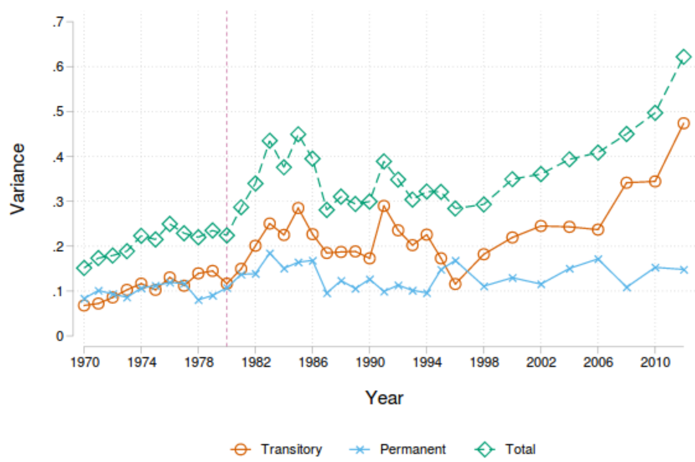
Author's calculations using the SIPP GSF from 1978 to 2011 in Figure (4a) and the PSID from 1970 to 2012 in Figure (4b). Sample is all men age 25 to 59 with positive earnings in the reference year, excluding the top and bottom 1% of earners. Decomposition based on the residual from a regression of log earnings on a quadratic in age interacted with race and education estimated separately by year. The PSID includes only sample men. The vertical line in 1980 in the PSID represents the year in which the SIPP GSF data begins.

Figure 5: Error Components Model

(a) SIPP GSF



(b) PSID



Author's calculations using the SIPP GSF from 1978 to 2011 in Figure (5a) and the PSID from 1970 to 2012 in Figure (5b). Sample is all men age 25 to 59 with positive earnings in the reference year, excluding the top and bottom 1% of earners. Decomposition based on the residual from a regression of log earnings on a quadratic in age interacted with race and education estimated separately by year. The PSID includes only sample men. The vertical line in 1980 in the PSID represents the year in which the SIPP GSF data begins.

Figures (5a) and (5b) show a increasing trend in the within group transitory variance of earnings in the SIPP GSF (PSID) since 1978 (1970). Consistent with the results in Figures (1) and (4), both the SIPP GSF and the PSID also show an increase in the transitory variance of earnings in the period since 1998. The increase is markedly larger in percent terms in the PSID and represents somewhat of a break from the relatively stable transitory variance of earnings in the 1990s, whereas in the SIPP GSF the transitory variance of earnings increases steadily over time. As we noted in our discussion of Figure (1), it appears that the changes in data collection and processing in the PSID in the early 1990s manifested in the form of higher levels of the transitory variance of earnings. However, it is also clear that one reason for the marked increase in the variance of the transitory component of earnings is the very rapid increase in overall inequality in the PSID after the mid-1990s. We will turn to this increase in inequality in our discussion of the approximate nonparametric decomposition. There are also increases in the variance in the permanent component of earnings in both models over time. In the SIPP GSF the variance of the permanent component of earnings increases from 0.32 in 1998 to 0.38 in 2011 while in the PSID the variance of the permanent component of earnings increases from 0.11 in 1998 to 0.15 in 2012.

Also consistent with our results in Figures (1) and (4), the transitory variance and the overall variance in the SIPP GSF is substantially higher than in the PSID. Figure (5) also shows that the level of the variance of the permanent component of earnings is substantially higher in the SIPP GSF than in the PSID. The higher levels of the predicted overall variance of earnings in the SIPP GSF implies that predicted inequality in the administrative data is substantially higher than in the PSID. We return to these level differences in Section (6).

Because our methods more closely replicate those used in Moffitt and Gottschalk (2012) it is perhaps unsurprising that our qualitative results from the PSID closely resemble theirs. Although the model used here attributes somewhat more of the total variance to transitory variance, the broad trends and the contribution to changes in both the transitory and permanent variances are quite similar across the two models, both of which are estimated on the PSID. The notable exception to this is what happens between 2000 and 2004, where Moffitt and Gottschalk (2012) find an increase in the permanent variance and a decrease in the transitory, while we find an increase in the transitory and a decrease in the permanent. Because Moffitt and Gottschalk (2012) stop their analysis in 2004 we do not know whether this the beginning of a trend and thus a divergence between the two slightly different models, or a cyclical pattern in the transitory component, either way the fit of the two models prior to 2004 is quite similar. Further, the fact that this period contains a recession, a rise in transitory variances is to be expected.

Debacker et al. (2013) provide the only other error components model of earnings estimated using administrative data. Debacker et al. (2013), using a sample of male tax filers, finds that the variance of the permanent component is about five times bigger than the transitory component, and that the entire increase in the variance of earnings in the 1990s is due to increases in the variance of the permanent component. However, the model estimated in Debacker et al. (2013) is not consistent with those developed in Baker and Solon (2003), Haider (2001) or Moffitt and Gottschalk (2012) or the model used here. In particular, Debacker et al. (2013) model estimates an MA(2) in the transitory component of earnings, while we follow Moffitt and Gottschalk (2012) and estimate an ARMA(1,1) in the transitory component. This is potentially important as the estimate of the AR component tends to be large (e.g. it is 0.846 in Moffitt and Gottschalk (2012)) and would lead to an overstatement of the variance of the permanent component. Second, Debacker et al. (2013) does not allow variances to vary by both age and birth cohort. As Baker and Solon (2003) make clear, year-to-year changes in the variance of permanent and transitory earnings change for three reasons: (1) different birth cohorts have different average variances, (2) variances change over the life-cycle, and (3) the variances change net of differences across birth cohorts and age. The failure to decompose every year at the same age, and to use the birth cohort in each year associated with the age at which the decomposition is performed, means that the trends through time confuse changes in the variances through time with life-cycle effects and birth cohort effects. Finally, the data in Debacker et al. (2013) do not contain information on educational attainment or race and so they are unable to use residuals from a regression of log earnings on age, education, and race which necessarily increases the contribution of the permanent component of earnings to the total variance.

This latter point, the inability to decompose earnings net of returns to education, is quite important. In results shown in Figure (A3) in Appendix (A.4) we show estimates of the same error components model depicted in Figure (5a) on earnings that are age-adjusted but not adjusted for education or race. Consistent with Haider (2001), we find that the estimated share of cross-sectional inequality that is due to inequality in permanent earnings is much higher when not controlling for education and race. This is not surprising given the rise in returns to schooling for men, especially during the 1980s to mid 1990s (Autor, Katz, and Kearney, 2008). One important contribution of this work is that we are able to closely follow the literature and estimate an error components model of earnings using administrative earnings data *and* controls for race and education.

5.4 Approximate Nonparametric Decomposition Results

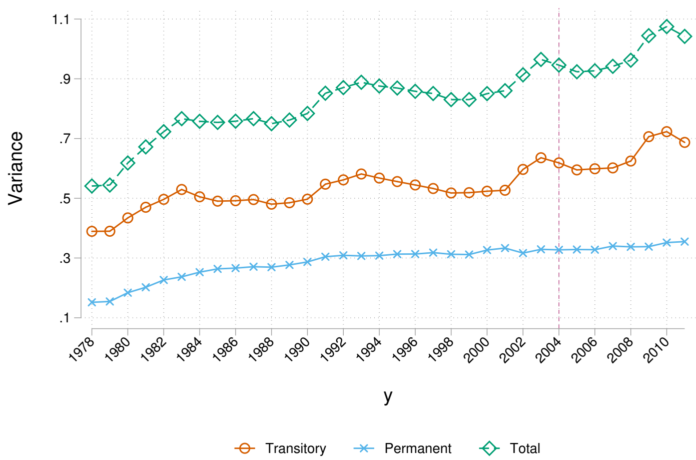
To complement the estimates from the error components model, we provide one additional method of estimating both the permanent and transitory components of earnings that is less reliant on parametric assumptions. Figure (6) shows the results from the approximate nonparametric decomposition outlined in Section (4) with Figure (6a) showing the results from the SIPP GSF and Figure (6b) shows the results from the PSID.

Figure (6a) shows increases in the variance of both the permanent and transitory component of earnings over the entire period and in the period since the mid-1990s. This model is broadly consistent with Figure (5a) and suggests that the increases in the transitory variance since the mid-1990s implied by the increases in volatility and the increases in the transitory variance estimated with the window average method were accompanied by increases in the variance of the permanent component of earnings.

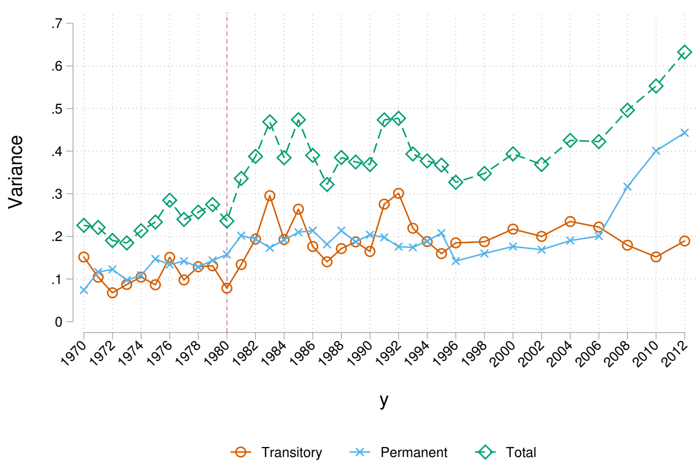
However, Figure (6b) using the PSID is not consistent with Figure (5b). Both figures show

Figure 6: Approximate Nonparametric Decomposition

(a) SIPP GSF



(b) PSID



Author's calculations using the SIPP GSF from 1978 to 2011 in Figure (6a) and the PSID from 1970 to 2012 in Figure (6b). Sample is all men age 25 to 59 with positive earnings in the reference year, excluding the top and bottom 1% of earners. Decomposition based on the residual from a regression of log earnings on a quadratic in age interacted with race and education estimated separately by year. The PSID includes only sample men. The vertical line in 1980 in the PSID represents the year in which the SIPP GSF data begins.

a very rapid increase in the overall variance of earnings (overall inequality). However, the approximate nonparametric method attributes nearly all of the increase to increasing dispersion in the permanent component of earnings while the more complete error components model attributes nearly all of the increase in inequality to increases in the variance of the transitory component of earnings. We do not know exactly what is behind the differences in the results from the two models in the PSID but we note that these differences begin to appear in the years after 2006 when inequality growth expands rapidly in the PSID while expanding more slowly in the SIPP GSF.

6 Accounting for higher volatility in the SIPP

Though in our analysis, we highlight the similarity of the increase in earnings instability since the late 1990s in the PSID and the SIPP GSF, the level of earnings instability and the level of earnings inequality is substantially higher in the SIPP GSF than in the PSID when we apply a top and bottom 1% trim to each sample. Using similar data, Sabelhaus and Song (2009) show that including earnings at the very bottom of the distribution – in their case earnings under the eligibility criteria for a year towards Social Security earnings – increases volatility substantially. They find that trimming the bottom of the earnings distribution to exclude individuals with earnings below the Social Security threshold reduces volatility to approximately the same level as we find in the PSID, though the trend through time remains different. In this section, we explore whether the level differences that we find between the SIPP GSF and the PSID become smaller when we apply a more aggressive trim to the SIPP GSF at the bottom of the earnings distribution.

Table (1) shows the 1st, 5th, and 99th percentiles of the earnings distribution in the SIPP GSF and in the PSID for selected years, in constant 2011 dollars. The very bottom of the earnings distributions in the PSID and the SIPP GSF are quite different. Consistent with other analyses of inequality, in each dataset, earnings at the low percentiles decrease and at the high percentiles increase over time. However for each year, the 1st percentile in the PSID is substantially higher

Table 1: Earnings Percentiles by Year for Full Sample

	SIPP GSF			PSID	
	P1	P5	P99	P1	P99
1978	1268.59	8192.16	262375.19	4884.15	228889.73
1980	998.40	6040.76	216913.59	3900.87	195043.44
1990	665.40	4368.86	248090.62	3710.04	241539.22
2000	576.80	4489.15	385890.94	4025.23	306658.66
2010	322.26	2854.63	397690.00	1526.72	331667.59

Author's calculations using the SIPP GSF and PSID from 1978 to 2011. Sample is all men age 25 to 59 with positive earnings in the reference year.

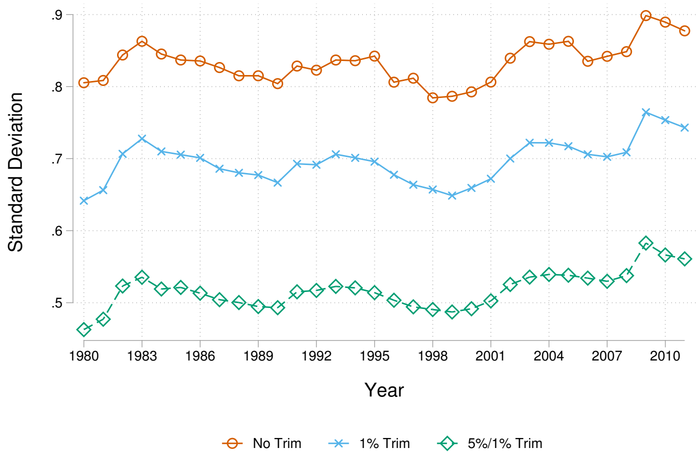
than the 1st percentile in the SIPP GSF, though generally smaller than the 5th percentile in the SIPP GSF. Table (1) highlights the concern that the low 1% threshold in the SIPP GSF means that small absolute changes in earnings at the very bottom of the earnings distribution may explain the higher level of volatility in the SIPP GSF compared with the PSID. In what follows, we show estimates of earnings volatility using untrimmed SIPP GSF data, the 1% top and bottom trim that we have applied throughout the paper, and a bottom 5% and top 1% trim of earnings. We show that while the levels of earnings instability are sensitive to the choice of trim, the trends are not.[10]

Figure (7a) shows the results of estimating earnings volatility in the SIPP GSF using untrimmed data, using our preferred 1% top and bottom trim, and using the more aggressive bottom 5% trim. Figure (7a) shows that earnings volatility is sensitive to trimming earnings at the bottom. Untrimmed, the variance of log earnings changes fluctuates between 0.8 and 0.9, a 1% trim of the top and the bottom decreases volatility to 0.65 to 0.78, and a further trim of the bottom to 5% decreases volatility to between 0.48 and 0.6. Volatility in the PSID varies between 0.35 and 0.54, or roughly 11% less that the most aggressive trim in any given year. Trimming earnings at the

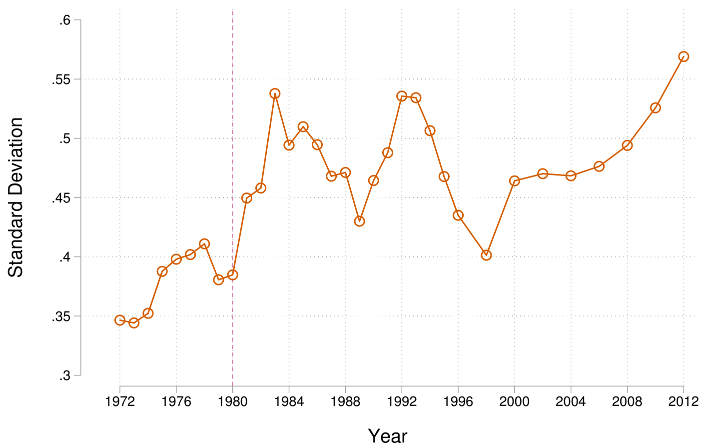
[10]Two other ways we could have trimmed include a fixed threshold in real earnings (such as the eligibility criteria for a year toward Social Security earnings) or applying the first percentile from the PSID to the SIPP GSF. The latter is not appealing as the 1st percentile in the PSID is quite variable; sometimes it is above the 5th percentile in the SIPP GSF and sometimes it is below. The fixed threshold has the advantage of being fixed, so the impact of low earnings on volatility does not change through time due to the introduction of a lower earnings threshold, but has the disadvantage of excluding more individuals in later years than early years as the earnings distribution changes. However, this is a common choice in the administrative data and we show some results with a fixed trim in Appendix (A.5).

Figure 7: Effect of Trimming on Volatility

(a) SIPP GSF



(b) PSID



Author's calculations using the SIPP GSF from 1978 to 2011 reported in Figure (7a) and using the PSID from 1970 to 2012 reported in Figure (7b). Volatility is the standard deviation of the age-adjusted two-year change in log earnings for a sample of men ages 25 to 59 with positive earnings. Earnings changes age-adjusted separately by year using a quadratic in age. The PSID includes only sample men. The vertical line in 1980 in Figure (7b) represents the year in which the SIPP GSF data begins. The "1% trim" trims top and bottom 1% of earnings. The "5%/1% Trim" trims bottom 5% and top 1%. Earnings are trimmed prior to age-adjusting.

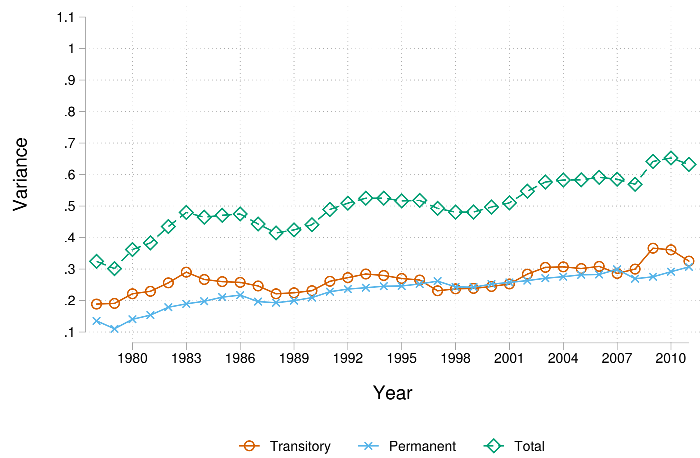
bottom of the earnings distribution brings earnings volatility in the SIPP GSF much closer to that seen in the PSID, but the trend in the SIPP GSF remains unchanged.

Trimming earnings more aggressively at the bottom of the earnings distribution reduces volatility which should correspond to a reduction in the variance of the transitory component of earnings from the error components model. Figure (8) show the results of estimating our error components model in the SIPP GSF with the more aggressive bottom 5% trim. As in Figure (5a), both the transitory and permanent variance of earnings increase over time but the relative contribution of the transitory variance is much smaller – the level of the total variance and of the contribution of the transitory component of earnings to total inequality declines. While the results from the error components model using 1% trim at the top and bottom of the earnings distribution suggests that over 70% of inequality comes from transitory earnings variability, the results using the more aggressive trim suggest that the contribution of the permanent and transitory components to the total variance of earnings is about equal. This is consistent with the notion that small changes in earnings at low levels of earnings substantially increase the level of the variance of the transitory component of earnings.

Table (1) also provides evidence on why the variance of the transitory component of earnings increases so rapidly after 2000 in the PSID in Figure (5b). Between 2000 and 2010, the first percentile of earnings in the PSID drops from $4025 to $1526 - a 60% decline. We know that the transitory component of earnings is higher among individuals with low levels of earnings, if more such individuals are included in the PSID in latter years, the transitory component of earnings should rise. One reason the approximate nonparametric method may attribute this increase to the permanent component of earnings is that it relies on individuals for whom long autocovariances are observable in the PSID to estimate trends in the permanent variance of earnings. These individuals may have lower earnings volatility than younger individuals, individuals with more sporadic labor force attachment, or individuals who attrit from the PSID over long windows of time.

Which is the most appropriate trim remains an open question. There are three primary moti-

Figure 8: Error Components Model, SIPP GSF 5%/1% Trim of Earnings



Author's calculations using the SIPP GSF from 1978 to 2011. Sample is all men age 25 to 59 with positive earnings in the reference year. Decomposition based on the residual from a regression of log earnings on a quadratic in age interacted with race and education estimated separately by year.

vations for trimming in the PSID: (1) earnings are top-coded, (2) the density of low earnings is thin, and (3) there may be measurement error particularly in unusually low earnings. In the SIPP GSF, none of these concerns apply. Cross-sectional sample sizes are large enough that there are no serious earnings density problems, and there is no top-coding nor measurement error. On the other hand, even the 5% threshold represents marginal labor force attachment at best, and allowing those individuals to play a substantial role in the level of volatility and cross-sectional inequality may not be appropriate.

7 Conclusion

The sharp rise in inequality during the late 1970s and early 1980s was the result of both a widening of the distribution of permanent earnings and an increase in transitory earnings fluctuations (Haider, 2001, Moffitt and Gottschalk, 2012). After a period of relative stability during the 1990s, cross-sectional inequality began to increase rapidly again in the late 1990s and yet, there is still little agreement on whether earnings volatility increased or on the relative importance of transitory shocks and permanent earnings inequality for overall inequality during this period. The difficulty in reconciling recent trends across the literature is a result of important differences in methods, samples, and datasets across papers.

This project is the first to use an administrative panel dataset and a survey dataset, with consistent methods and sample selection criteria, to estimate models of earnings volatility and error components models of earnings. In contrast to the recent literature, we find similar trends in the survey and administrative data. Across both administrative and survey data we find that earnings volatility has increased rapidly since the late 1990s. This increase in volatility is suggestive evidence of an increase in the variance of the transitory component of earnings. In the SIPP GSF, three different decomposition techniques – an informal window averaging method, a formal error components model of earnings, and a nonparametric method using long autocovariances – confirm that there was an increase in the transitory variance of earnings which was accompanied by a widening distribution of permanent earnings. In the PSID, the three decomposition techniques suggest increases in both the transitory and permanent component of earnings but different methods put different weight on the two components. We suspect that very rapidly increasing inequality in the PSID after 2006, combined with somewhat small sample sizes, especially when relying on individuals observed over many years in the PSID, all contribute to the sensitivity of results estimated with different techniques in the last few years. Our analysis shows that the differences in the trends that emerge between survey and administrative data are mainly the result of differences in sample and methods, and not a result of differences in underlying trends in the two data sources. However, we exercise caution when interpreting results from the PSID over the last few years. The very rapid rise in inequality in the PSID after 2000 is something that should be explored in more depth.

Showing that recent trends in earnings variability are similar in administrative data and in the PSID is useful because it sets the stage to exploit the benefits of each dataset in understanding the causes and consequences of increasing earnings instability. The pronounced increase in both volatility and transitory earnings variability raises important questions about the distribution of earnings risk in an environment of growing inequality. Rising inequality may be less of a concern if it is accompanied by an increasing share of transitory earnings variation because it implies that inequality in lifetime earnings may be lower than inequality in annual earnings, that changes in earnings rank may be more likely even as inequality has increased, and that consumption may be largely insured against these earnings fluctuations. These assertion are only true, however, if individuals have sufficient means to smooth possibly long-lived transitory shocks and if the transitory shocks generally cancel out over time. If, for example, some individuals generally experience negative transitory shocks while some experience generally positive shocks, then rising transitory variances may result in a divergence of lifetime earnings and a divergence in movements of rank for individuals starting their career at the bottom and top of the earnings distribution. Perhaps even more importantly, Blundell, Pistaferri, and Preston (2008) show empirically that while consumption is largely insured against transitory earnings changes, poor families do experience changes in consumption due to transitory earnings instability. The large sample sizes and the survey information on educational attainment in the SIPP GSF allows for future work on these important issues. Despite these advantages, the SIPP GSF does not allow for an analysis of the underlying components of earnings instability such as hours, wages, or weeks worked. However, the PSID does allow for such an analysis which is bolstered by understanding that underlying trends in earnings instability in the PSID are matched in administrative data.

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):1451–1467.

Autor, David H., Lawrence F. Katz, and Melissa S. Kearney. 2008. "Trends in U.S. Wage Inequality: Revising the Revisionists." *Review of Economics and Statistics* 90 (2):300–323.

Baker, Michael and Gary Solon. 2003. "Earnings Dynamics and Inequality Among Canadian Men, 1976-1992." *Journal of Labor Economics* 21 (2):289–321.

Benedetto, Gary, Martha H Stinson, and John M Abowd. 2013. "The Creation and Use of the SIPP Synthetic Beta." US Census Bureau.

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

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.

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):750–774.

Debacker, Jason, Bradley Heim, Vasia Panousi, Shanthi Ramnath, and Ivan Vidangos. 2013. "Rising Inequality: Transitory or Persistent? New Evidence from a Panel of U.S. Tax Returns." *Brookings Papers on Economic Activity* Spring.

Doris, Aedín, Donal O'Neill, and Olive Sweetman. 2011. "GMM Estimation of the Covariance Structure of Longitudinal Data on Earnings." *Stata Journal* 11 (3):1–21.

Fitzgerald, John. 2011. "Attrition in Models of Intergenerational Links Using the PSID with Extensions to Healtha and Sibling Models." *BE Journal of Economic Analysis & Policy* 11 (3):Article 2.

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

———. 1998b. "The Impact of Attrition in the Panel Study of Income Dynamics on Intergenerational Analysis." *Journal of Human Resources* 33 (2):300–344.

Gottschalk, Peter and Robert Moffitt. 2009. "The Rising Instability of U.S. Earnings." *Journal of Economic Perspectives* 23 (4):3–24.

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.

Haider, Steven J. 2001. "Earnings instability and earnings inequality of males in the United States: 1967–1991." *Journal of Labor Economics* 19 (4):799–836.

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.

McGonagle, Katherine, Robert Schoeni, Narayan Sastry, and Vicki Freedman. 2012. "The Panel Study of Income Dynamics: Overview, Recent Renovations, and Potential for Life Course Research." *Longitudinal and Life Course Studies* 3:268–284.

Moffitt, Robert A and Peter Gottschalk. 2002. "Trends in the Transitory Variance of Earnings in the United States." *Economic Journal* 112 (478):C68–C73.

———. 2012. "Trends in the Transitory Variance of Male Earnings Methods and Evidence." *Journal of Human Resources* 47 (1):204–236.

Monti, Holly and Graton Gathright. 2013. "Measuring Earnings Instability Using Survey and Administrative Data." Working Paper, US Census Bureau.

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.

Schoeni, Robert, Frank Stafford, Katherine McGonagle, and Patricia Andreski. 2013. "Response Rates in National Panel Surveys." *Annals of the American Academy of Political and Social Science* 645 (1):60–87.

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.

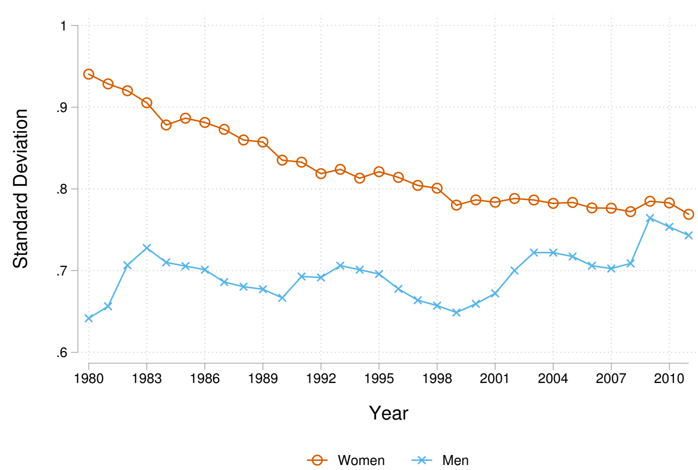
Ziliak, James P, Bradley Hardy, and Christopher Bollinger. 2011. "Earnings Volatility in America: Evidence from Matched CPS." *Labour Economics* 18 (6):742–754.

A Appendix

A.1 Volatility for Men and Women

The results of estimating Equation (3) separately for men and women are shown in Figure (9). While our analysis focuses on men, Sabelhaus and Song (2009, 2010) pool men and women together in their analysis of administrative earnings data. Recall that they find steadily declining earnings volatility from the early 1980s through the 2000s. Figure (9) shows that volatility for men and women, respectively, differs both in level and in trend. While volatility is higher overall for women, it falls considerably over this time period. Volatility for men, on the other hand, is more cyclical and ultimately ends higher than where it began. Clearly, the finding in Sabelhaus and Song (2009, 2010) that volatility is steadily falling through this time period is driven by the choice to pool men and women. If we were to pool men and women, our data would also show moderate declines in earnings volatility.

Figure 9: Earnings Volatility by Gender, SIPP GSF

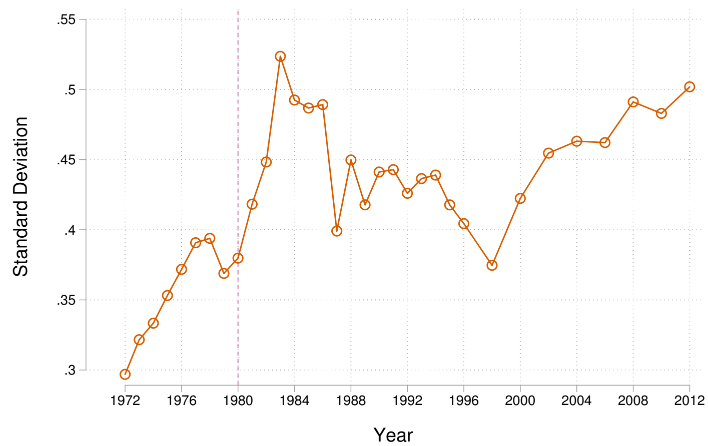


Author's calculations using the SIPP GSF from 1978 to 2011. Volatility is the standard deviation of the age-adjusted two-year change in log earnings for a sample of individuals ages 25 to 59 with positive earnings, excluding the top and bottom 1% of annual earnings. Earnings changes age-adjusted separately by year using a quadratic in age.

A.2 PSID including Non-Sample Men

Our sample includes only male heads of household who are followed longitudinally in the PSID. An alternative sampling choice would be to include all male heads of household regardless of whether they are followed longitudinally. To understand the difference between these samples, we briefly outline the relevant PSID "rules." First, the PSID only follows respondents who lived in a household in the original 1968 sample or who are the biological or adopted children of these original 1968 sample members. Their spouses or cohabiting partners are only respondents in the PSID as long as they coreside with a PSID sample member who is followed longitudinally. Second, the PSID preferences men when assigning who is the head of household. A household head can only be female if she is unmarried or she has lived with her cohabiting partner for a year or less. These two rules combined suggest that including all male heads in the PSID sample biases the sample toward men who are either married or in stable cohabiting relationships. In contrast, including only male heads who are followed longitudinally does not introduce such a bias. Below we show estimates of earnings volatility and the window averaging model using all male heads to show consistency with Shin and Solon (2011) and Moffitt and Gottschalk (2012).

Figure A1: Earnings Volatility, PSID including Non-Sample Men



Author's calculations using the PSID from 1970 to 2012. Volatility is the standard deviation of the age-adjusted two-year change in log earnings for a sample of men ages 25 to 59 with positive earnings, excluding the top and bottom 1% of annual earnings. Earnings changes age-adjusted separately by year using a quadratic in age.

Figure A2: Window Averaging, PSID including Non-Sample Men

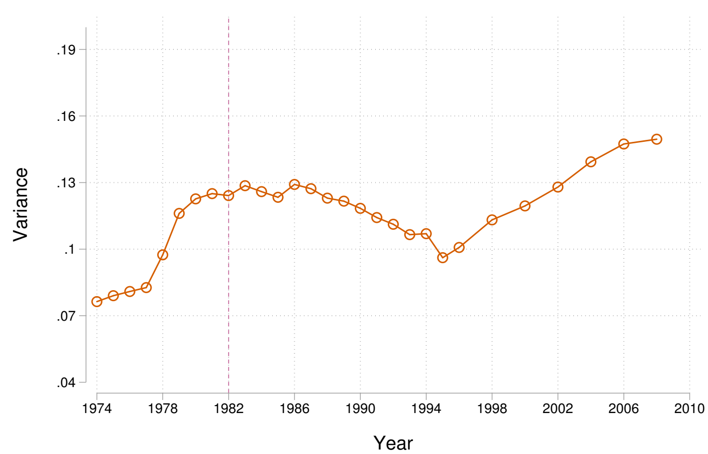


Figure (A2) uses the PSID from 1970 to 2012. Sample is all men age 25 to 59 with positive earnings in the reference year, excluding the top and bottom 1% of earners. Decomposition based on the residual from a regression of log earnings on a quadratic in age interacted with race and education estimated separately by year.

A.3 Parameter Estimates for Equation (4), SIPP GSF and PSID

Table A1: GSF Estimate of Earnings Dynamics Model

	Coef.	Std. Err.
ρ	0.7125	[0.0000]
σ^2_α	0.5896	[0.0017]
σ^2_β	0.0001	[0.0000]
σ^2_ϵ	0.3490	[0.0001]
$\sigma^2_{\nu_1}$	0.3707	[0.0000]
σ^2_ω	0.0057	[0.0000]
$\sigma_{\alpha\beta}$	-0.0104	[0.0000]
θ	-0.3579	[0.0000]
l_{1979}	0.9538	[0.0001]
l_{1980}	1.0046	[0.0002]
l_{1981}	1.0152	[0.0002]
l_{1982}	1.0620	[0.0002]
l_{1983}	1.1447	[0.0003]
l_{1984}	1.0886	[0.0002]
l_{1985}	1.0718	[0.0002]
l_{1986}	1.0706	[0.0002]
l_{1987}	1.0786	[0.0002]
l_{1988}	1.0554	[0.0002]
l_{1989}	1.0708	[0.0002]
l_{1990}	1.0760	[0.0002]
l_{1991}	1.1330	[0.0002]
l_{1992}	1.1533	[0.0002]

l_{1993}	1.2029	[0.0002]
l_{1994}	1.1934	[0.0002]
l_{1995}	1.1867	[0.0002]
l_{1996}	1.1621	[0.0002]
l_{1997}	1.0978	[0.0002]
l_{1998}	1.0962	[0.0002]
l_{1999}	1.0846	[0.0002]
l_{2000}	1.1205	[0.0002]
l_{2001}	1.1153	[0.0002]
l_{2002}	1.1853	[0.0002]
l_{2003}	1.2192	[0.0003]
l_{2004}	1.2107	[0.0003]
l_{2005}	1.2168	[0.0003]
l_{2006}	1.2252	[0.0003]
l_{2007}	1.1679	[0.0003]
l_{2008}	1.1782	[0.0003]
l_{2009}	1.2845	[0.0003]
l_{2010}	1.3080	[0.0003]
l_{2011}	1.2330	[0.0003]
p_{1979}	1.0089	[0.0001]
p_{1980}	1.0877	[0.0002]
p_{1981}	1.1218	[0.0002]
p_{1982}	1.1894	[0.0002]
p_{1983}	1.2325	[0.0003]
p_{1984}	1.2888	[0.0003]
p_{1985}	1.3336	[0.0004]

p_{1986}	1.3510	[0.0005]
p_{1987}	1.3660	[0.0005]
p_{1988}	1.3634	[0.0005]
p_{1989}	1.4033	[0.0006]
p_{1990}	1.4389	[0.0007]
p_{1991}	1.5018	[0.0008]
p_{1992}	1.5469	[0.0009]
p_{1993}	1.5582	[0.0010]
p_{1994}	1.5992	[0.0011]
p_{1995}	1.6191	[0.0013]
p_{1996}	1.6370	[0.0014]
p_{1997}	1.7237	[0.0017]
p_{1998}	1.7415	[0.0019]
p_{1999}	1.7690	[0.0021]
p_{2000}	1.8083	[0.0023]
p_{2001}	1.8457	[0.0026]
p_{2002}	1.8663	[0.0028]
p_{2003}	1.9052	[0.0031]
p_{2004}	1.9580	[0.0035]
p_{2005}	1.9795	[0.0039]
p_{2006}	2.0043	[0.0043]
p_{2007}	2.1283	[0.0052]
p_{2008}	2.1651	[0.0057]
p_{2009}	2.1628	[0.0060]
p_{2010}	2.2755	[0.0071]
p_{2011}	2.3395	[0.0080]

q_2	0.9298	[0.0002]
q_3	0.8369	[0.0004]
q_4	0.7143	[0.0006]
q_5	0.5647	[0.0006]
s_2	0.9534	[0.0000]
s_3	0.9102	[0.0000]
s_4	0.9105	[0.0000]
s_5	0.9541	[0.0001]

Table A2: PSID Estimate of Earnings Dynamics Model

	Coef.	Std. Err.
ρ	0.6970	[0.0019]
σ^2_α	0.5745	[0.0276]
σ^2_β	0.0002	[0.0000]
σ^2_ϵ	0.1109	[0.0005]
$\sigma^2_{\nu_1}$	0.0680	[0.0001]
σ^2_ω	0.0017	[0.0000]
$\sigma_{\alpha\beta}$	−0.0103	[0.0000]
θ	−0.3419	[0.0014]
l_{1971}	0.8337	[0.0119]
l_{1972}	1.5868	[0.0219]
l_{1973}	0.8465	[0.0122]
l_{1974}	0.8880	[0.0126]
l_{1975}	0.8274	[0.0122]
l_{1976}	0.9286	[0.0133]
l_{1977}	0.8590	[0.0119]
l_{1978}	0.9580	[0.0116]
l_{1979}	0.9762	[0.0123]
l_{1980}	0.8758	[0.0120]
l_{1981}	0.9913	[0.0163]
l_{1982}	1.1495	[0.0169]
l_{1983}	1.1248	[0.0200]
l_{1984}	1.0664	[0.0187]
l_{1985}	1.2000	[0.0206]

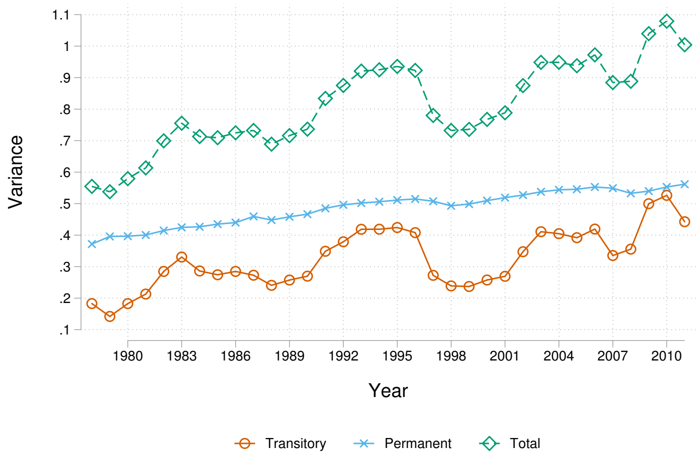
l_{1986}	1.0703	[0.0161]
l_{1987}	0.9665	[0.0129]
l_{1988}	0.9724	[0.0158]
l_{1989}	0.9755	[0.0127]
l_{1990}	0.9348	[0.0137]
l_{1991}	1.2100	[0.0179]
l_{1992}	1.0904	[0.0177]
l_{1993}	1.0457	[0.0162]
l_{1994}	1.1046	[0.0157]
l_{1995}	0.9671	[0.0157]
l_{1996}	0.7909	[0.0143]
l_{1998}	0.9919	[0.0135]
l_{2000}	1.0892	[0.0169]
l_{2002}	1.1509	[0.0176]
l_{2004}	1.2246	[0.0210]
l_{2006}	1.2091	[0.0219]
l_{2008}	1.4513	[0.0285]
l_{2010}	1.4576	[0.0314]
l_{2012}	1.7098	[0.0433]
p_{1971}	1.1400	[0.0073]
p_{1972}	0.7391	[0.0067]
p_{1973}	1.1266	[0.0109]
p_{1974}	1.2978	[0.0125]
p_{1975}	1.3787	[0.0171]
p_{1976}	1.4701	[0.0260]
p_{1977}	1.5055	[0.0310]

p_{1978}	1.4371	[0.0300]
p_{1979}	1.5720	[0.0369]
p_{1980}	1.7697	[0.0504]
p_{1981}	2.0608	[0.0736]
p_{1982}	2.1313	[0.0811]
p_{1983}	2.5223	[0.1222]
p_{1984}	2.3327	[0.1208]
p_{1985}	2.4850	[0.1321]
p_{1986}	2.5556	[0.1530]
p_{1987}	2.4283	[0.1485]
p_{1988}	2.8470	[0.2158]
p_{1989}	2.7135	[0.2165]
p_{1990}	3.0516	[0.2922]
p_{1991}	2.7725	[0.2600]
p_{1992}	3.0317	[0.3245]
p_{1993}	2.9309	[0.3268]
p_{1994}	2.9027	[0.3278]
p_{1995}	3.6508	[0.5700]
p_{1996}	3.9330	[0.7675]
p_{1998}	3.5957	[0.6441]
p_{2000}	4.0124	[0.9034]
p_{2002}	3.9068	[0.9292]
p_{2004}	4.5901	[1.3918]
p_{2006}	5.0419	[1.8499]
p_{2008}	4.8579	[1.7150]
p_{2010}	5.9625	[2.7718]

p_{2012}	6.0718	[2.8718]
q_2	0.6154	[0.0045]
q_3	0.4245	[0.0053]
q_4	0.3282	[0.0062]
q_5	0.1845	[0.0037]
s_2	1.0498	[0.0028]
s_3	1.1973	[0.0045]
s_4	1.1575	[0.0060]
s_5	1.0835	[0.0076]

A.4 Age-adjusted Error Components Model in the SIPP GSF

Figure A3: Error Components Model, SIPP GSF 1%/1% Trim of Age-adjusted Earnings



Author's calculations using the SIPP GSF from 1978 to 2011. Sample is all men age 25 to 59 with positive earnings in the reference year. Decomposition based on the residual from a regression of log earnings on a quadratic in age estimated separately by year.

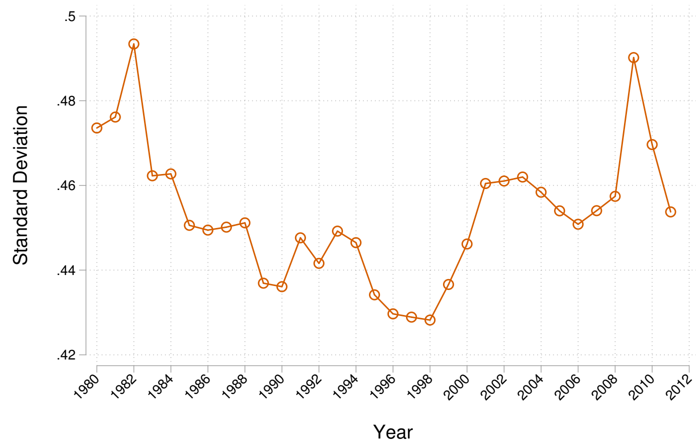
A.5 Fixed Trim in SIPP GSF

Figure (A4a) shows the volatility of one-year changes in age-, education-, and race-adjusted earnings and the results from the error components model using a trim fixed at the minimum threshold of one-fourth of a full-year full-time minimum wage in 2013 ($3770) indexed to inflation. There is no trim at the top of the earnings distribution. This is similar to the trim used in Debacker et al. (2013), Kopczuk, Saez, and Song (2010), Sabelhaus and Song (2010).

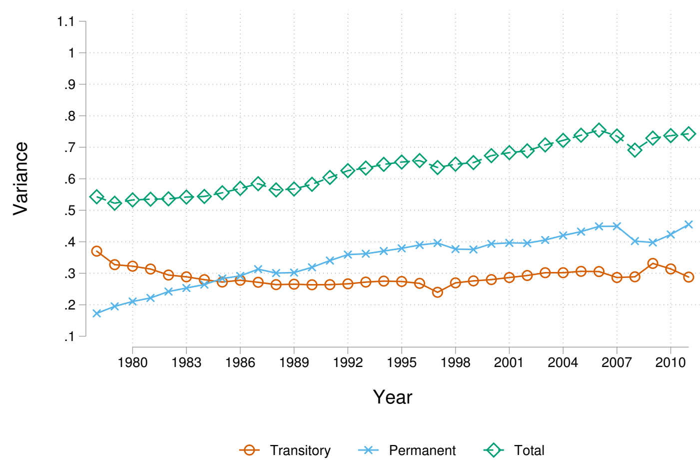
The fixed trim substantially changes the shape of the trend in volatility and in the transitory component of earnings from a steady upward trend to a more u-shaped trend. In the error components model, it also shows that at the beginning of the period, the variance of the transitory component of earnings was larger than that of the permanent component of earning but that the variance permanent component of earnings is higher than the transitory component by the end of the period. The change in shape occurs because when using a fixed trim, the level of the transitory variance rises at the beginning of the period because in the 1980s, the fixed trim occurs the 1st and 5th percentile of the earnings distribution. By the end of the period, the fixed trim is well above the 5th percentile. While this trim holds the impact of small changes in earnings at the bottom of the earnings distribution fixed over time, it does not account for some of the rise in inequality which occurred near the bottom of the earnings distribution. But, even with a fixed trim at the bottom of the earnings distribution, since the late 1990s, the variance of both the permanent and the transitory component of earnings have been increasing.

Figure A4: Volatility and Error Components Model with Fixed Trim, SIPP GSF

(a) SIPP GSF Volatility



(b) SIPP GSF ECM



Author's calculations using the SIPP GSF from 1978 to 2011. Figure (A4a) shows standard deviation of the age-adjusted one-year change in log earnings for a sample of men ages 25 to 59 with positive earnings. Earnings changes are age-adjusted separately by year using a quadratic in age. Figure (A4b) shows the decomposition based on the residual from a regression of log earnings on a quadratic in of age interacted with race and education estimated separately by year. Earnings are trimmed at $3770 2013$ prior to adjusting.