The relationship between inequality in the annual earnings of US men and inequality in their lifetime earnings, 1986-2018

Moshe Justman[a] and Hadas Stiassnie[b]

Abstract

We answer the question, how much of the rise in inequality in the annual earnings of US men in recent decades reflects an increase in inequality in their lifetime earnings? Using PSID data, we estimate lifetime earnings trajectories and average lifetime earnings for US men aged 31-59 with positive earnings in individual years between 1986-2018, and compare relative inequality in their annual and lifetime earnings year by year. We find that the rise in the log variance of lifetime earnings accounted for only 14 percent of the rise in the log variance of annual earnings over this period. A decomposition of the log variance of annual earnings highlights the dominant factors driving its rise: increases in short-term volatility and in its regressivity, and changes in the distribution of earnings over the lifecycle.

Keywords: Earnings inequality, lifetime earnings, volatility, life-cycle variation.
JEL code: D31

[a] Department of Economics, Ben Gurion University of the Negev, justman@bgu.ac.il; corresponding author.

[b] Department of Economics, Ben Gurion University of the Negev, hadas.stia@gmail.com

Neither author has any financial or non-financial interests that are directly or indirectly related to the work submitted for publication. No specific funding was received for this study.

All data used in this study is freely available for public use.

The relationship between inequality in the annual earnings of US men and inequality in their lifetime earnings, 1986-2018

1 Introduction

Variation in the inequality of annual earnings generally reflects not only variation in the underlying inequality of lifetime earnings but also variation in the short-term volatility of earnings, and changes in the demographic structure of the workforce. Each of these potential sources of inequality in annual earnings carries different social and economic implications and suggests a different set of policy responses. Thus, where rising inequality in lifetime earnings may call for policies that lower barriers to investment in human capital for less advantaged populations, increased short-term volatility may indicate a need for better financial instruments to smooth unexpected short-term earnings shocks, while greater lifecycle variation in earnings may be addressed through credit solutions tailored to smoothing longer-term, anticipated earnings variation.

This broad understanding has motivated extensive research aimed at separating the permanent and transitory components of earnings (Friedman, 1957).[1] These include studies that use short overlapping panels to average out earnings within moving windows, to separate "within" and "between" variation (e.g., Shorrocks, 1978; Burkhauser and Poupore, 1997; Haider, 2001; Moffitt and Gottschalk, 2012), studies that apply errors in variables models to separate transitory shocks from permanent shocks (Carlton and Hall 1975; Gottschalk and Moffitt 1994; Moffitt and Gottschalk 2012), and studies that estimate short-term volatility from the dispersion of individual short-term variation in earnings (Shin and Solon, 2011; Carr and

[1] For reviews of this extensive literature, see Creedy (1999), Burkhouser and Couch (2009), Gottschalk and Moffitt (2009), Kopczuk et al. (2010), Meghir and Pistaferri (2011), Jännti and Jenkins (2013), Guvenen, Kaplan, Song and Weidner, (2017, 2022), Moffitt and Zhang (2020).

Wiemers, 2018; Moffitt, 2020). Many of these studies also take the preliminary step of first regressing earnings on a polynomial in age, to factor out the effect of changes in age composition.[2] These studies show large increases in earnings volatility in recent years but provide little direct evidence on how inequality in lifetime earnings has changed over time.

Other studies that estimate inequality in lifetime earnings have focused on lifetime inequality within birth cohorts, estimating lifetime earnings by averaging earnings histories drawn from long panels of earnings data and then measuring their dispersion.[3] Studies in this vein include Bjorklund (1993) on the 1924-1936 Swedish birth cohorts; Aaberge and Mogstad (2015) on the 1942-1944 Norwegian birth cohorts; Bönke, Corneo, and Lüthen (2015) on the 1935-49 West German birth cohorts; Haider's (2001) comparison of ten-year earnings totals for men aged 30-44 in 1969-1978 from Panel Study of Income Dynamics (PSID) data; and Leonesio and Del Bene's (2011) comparison of 12-year earnings distributions in 1981-1992 and 1993-2004 using US Social Security Administration (SSA) data. The most relevant of these for our current study is Guvenen et al.'s (2022) detailed analysis of 57 years of longitudinal SSA earnings data, which they use to average individual earnings between the ages of 25 and 55 for single-year US cohorts born between 1932-1958, and measure the level and distribution of lifetime earnings within each birth cohort of men and women. They find that inequality in

[2] Earlier studies by Paglin (1975) and Mookherjee and Shorrocks (1982) focused explicitly on the effect of changes in the age distribution on inequality in annual earnings.

[3] A different approach, used primarily to assess the lifetime implications of various policy reforms, applies strong stationarity assumptions to construct *synthetic* lifetime earnings trajectories from short-term data (Creedy, 1977; Blomquist 1981; Friesen and Miller 1983; Aaronson, 2002; Bowlus and Robin 2004). These strong assumptions limit what these studies can tell us about the evolution of lifetime inequality over time.

lifetime earnings increased significantly within each gender group over this period though overall lifetime inequality within cohorts remained flat as the gender gap in earnings narrowed.

Their focus on measuring lifetime inequality within birth cohorts—rather than within calendar years, as inequality in annual earnings is routinely measured—is dictated by their direct approach to measuring lifetime earnings and by the limits of the data Thus, when Guvenen et al. (2022) measure individual lifetime earnings by averaging annual earnings over 31 years, the 57 years of SSA earnings data available to them are not sufficient to describe the distribution of lifetime earnings within a single calendar year, let alone follow its evolution over time.[4] Cohort-based measures of inequality require far fewer years of data, but cannot be compared to common measures of inequality in annual earnings or to calendar-year measures of volatility, which invariably relate to a different population: economically active individuals in a calendar year within a given range of ages. Thus, they cannot shed light on the relationship between inequality in annual earnings, lifetime inequality, and short-term fluctuations.

This is the issue we address here, estimating lifetime inequality among the same population that annual earnings inequality is measured in each calendar year: individuals who fall within a set range of ages and are economically active in that year. We achieve this by taking estimating lifetime earnings from partial earnings histories combined with background data on age, race and education.[5] Using longitudinal PSID data, collected annually to 1996 and biannually thereafter, we estimate individual lifetime earnings trajectories for US men born

[4] Calculating the lifetime earnings of all individuals aged 25 to 55 in a *single* calendar year from full (31-year) earnings histories, as Guvenen et al. (2022) do, requires 61 years of earnings data: 30 earlier years for the oldest cohort and 30 later years for the youngest cohort.

[5] This approach builds on related methods used in Justman and Krush (2013) and Justman and Stiassnie (2021) to estimate intergenerational mobility in lifetime incomes.

between 1927 and 1987—a range of cohorts that includes not only cohorts observed throughout their working years but also older cohorts observed only from mid-life and younger cohorts still far from the end of their work-life—and use these trajectories to estimate individual lifetime earnings. This approach has the obvious disadvantage of using partial earnings histories drawn from survey data and substituting estimated annual earnings for missing observations. However, as we show, the lifetime earnings estimates it yields are closely correlated with simple averages calculated for individuals observed throughout their working years; and when we use these lifetime earnings estimates to construct cohort-based measures of lifetime inequality they behave similarly to Guvenen et al.'s (2022) measures where they overlap. Moreover, when we estimate inequality in lifetime earnings for earlier years using only data to 2008, we find a close match between these restricted-data estimates and our unrestricted estimates, which use earnings data to 2018, thus demonstrating the robustness of our findings to the addition of future data. In addition, our estimation yields measures of residual volatility that accord closely with traditional short-panel measures of volatility.

The advantage of this approach is that it allows us to compare inequality in the annual earnings of US men aged 31-59 with positive earnings in a given year to inequality in the lifetime earnings of the same population in each year, annually from 1986 to 1996, and biannually from 1996 to 2018. This is, to our knowledge, the first study to estimate inequality in lifetime earnings within calendar years, over several decades, and relate it to corresponding measures of inequality in annual earnings and volatility. Moreover, the individual earnings trajectories we estimate allow us to represent an individual's log annual earnings in a given year as the log sum of his average lifetime earnings, the predicted life-cycle deviation of his earnings in that year from his lifetime average, and an unpredicted earnings shock. This allows us to represent the log variance of the annual earnings of economically active individuals aged 31-59 in a given year as the sum of six terms: the log variances of their lifetime earnings, of

their life-cycle deviations, and of unpredicted shocks (our measure of residual volatility), plus twice their corresponding pairwise covariances. This allows us to systematically relate year-by-year changes in the inequality of annual earnings to changes in lifetime inequality, in life-cycle variation, and in short-term volatility, and in their interactions.

Previewing our main results, we find that the average level of inequality in lifetime earnings in 1986-2018 was substantially lower than inequality in annual earnings in that period and increased much less steeply. The log variance of annual earnings, a widely used measure of inequality, increased by 20.2 percent over this period, with an average value of 0.515. This proportional increase was three times greater than the 6.6 percent increase in the log variance of lifetime earnings, its average value in this period equal to 0.218. Similarly, the widely used Gini coefficient of annual earnings increased by 17.9 percent from 1986 to 2018, with an average level of 0.339, a proportional increase two and a half times greater than the 7.0 percent increase in the Gini coefficient of lifetime earnings, with an average value of 0.252.

Applying our decomposition to the log variance of annual earnings from 1986 to 2018, we find that rising volatility was the main factor driving the increase in the inequality of annual earnings over this period, amplified by an increase in the regressive incidence of these shocks. The increase in the inequality of lifetime earnings played a much smaller role, with the rise in the log variance of lifetime earnings equal to only 14 percent of the increase in the log variance of annual earnings. These increases were partially offset by a substantial decline in the covariance of life-cycle variation and lifetime earnings, possibly reflecting the disequilibrium effect of rising skill levels in younger cohorts compressing the earnings distribution. (In the Supplementary Appendix, part B, we illustrate this effect with a numerical example.) Recognizing the key drivers of observed increases in the inequality of annual earnings of US men in recent decades contributes to our understanding of past and future trends in earnings inequality and to formulating appropriate policy responses.

The structure of the paper is as follows. Section 2 estimates predicted annual earnings trajectories and average lifetime earnings. Section 3 compares inequality in lifetime earnings to inequality in annual earnings from 1986 to 2018. Section 4 decomposes the log variance of annual earnings, representing it as the sum of changes in the log variances of lifetime earnings, life-cycle variation, and transitory shocks, plus twice their three pairwise covariances. Section 5 presents evidence on the consistency of our measures with previous findings and on their robustness, and Section 6 concludes.

2 Estimating individual lifetime earnings trajectories and lifetime earnings

We estimate individual lifetime earnings trajectories from the following equation:

$$Y_{ict} = \Sigma_{c=1}^{2}\Sigma_{k=1}^{3}\{\alpha_{1kc}age_{it}^{k}cogr_{ic} + \Sigma_{h=1}^{6}\alpha_{2hkc}educ_{ih}age_{it}^{k}cogr_{ic} + \Sigma_{l=1}^{4}\alpha_{3lkc}race_{il}age_{it}^{k}cogr_{ic}\} + D_i + \varepsilon_{it} \quad (1)$$

It regresses Y_{ict}, the earnings of individual *i* in cohort-group *c* in year *t*, on the interaction of two cohort-group indicators, $\{cogr_{ic}\}$, with a cubic function of age, $\{age_{it}^{k}\}$, and with its interaction with six categories of schooling levels, indexed by h, $\{educ_{ih}\}$, and with four categories of race, indexed by *l,* $\{race_{il}\}$;[6] and on individual random effects, D_i. The α's are the regression coefficients and ε_{it} is an i.i.d. error term.

Our data source is the PSID, from 1969 to 2018, with data collected annually until 1996 and bi-annually thereafter. We focus on male heads of households born between 1927 (aged 59 in 1986, the first year of our analysis) and 1987 (aged 31 in 2018, the last year), and follow previous research in restricting our analysis to the representative national sample drawn from

[6] The cohort-groups are: 1927-1955, 1956-87. Schooling categories are: 8 years or less, 9-10, 11-12, 13-15, 16, 17 or more. Race categories are: white, African American, Latino, other.

the Survey Research Center (SRC), not including the over-sampling of low-income families (SEO) or Latino families.[7] Our earnings variable is "wages and salaries", which we adjust to 2018 prices using the personal consumption expenditures (PCE) index. Following Solon and Shin (2011) and subsequent studies, we trim the top and bottom one percent of positive earnings observations in each year. We require at least three untrimmed earnings observations between the ages of 25 and 59, as well as race and years of schooling. This leaves us with a sample of 3,371 male heads of households, with 45,939 annual earnings observations, an average of 13.6 observations per person. The number of individuals with positive earnings in each of the years we analyze between 1986 and 2018, ranges between 967 and 1,168 (we have far fewer annual observations before 1986). The average age at which earnings are observed is 39; about 90 percent of the sample are white and 8 percent are African American. (Table A1 in the Supplementary Appendix presents year-by-year descriptive statistics.)

Our estimation of equation (1) accounts for 0.679 of the variance in annual earnings, yielding estimates of 48 age-earnings profiles, one for each combination of schooling, race, and cohort group. Education strongly affects the shape of the age-earnings profile. Thus, for men with 16 years of schooling, peak earnings, at age 50, are about twice their level at age 30, while for men with 11-12 years of schooling, earnings at age 30 and 50 are similar. (Figure A1 in the Supplementary Appendix illustrates this graphically, showing four estimated age-earnings profiles: for 11-12 and 16 years of schooling, for each of the two cohort-groups.) We apply these estimation results to obtain predicted earnings for each individual in each year, which we denote $\hat{Y}_{it}$.

[7] We omit two earlier years of available data to maintain a consistent definition of earnings. Methodological changes in the PSID in the early 1990s may have affected its consistency around the change (Dynan, Elmendorf and Sichel, 2012). Our focus is mainly on later years.

We take predicted earnings at age 40, $\hat{Y}_{i40}$, as our proxy for lifetime earnings. The choice of age 40 follows the empirical work of Haider and Solon (2006), Aaberge and Mogstad (2015), and others, and we find that it closely approximates actual average lifetime earnings for individuals in our sample observed over the full life cycle. (Actual averages are a biased estimate of lifetime earnings for individuals observed only in their earlier years or only in their later years.) To illustrate this, we extract from our sample 958 men born between 1937 and 1957—and thus observed over three decades of their working life—each with at least ten annual untrimmed earnings observations and an average of 23 observations per person. We find a correlation of 0.984 between their actual average annual lifetime earnings and our proxy measure for this subgroup. (Figure A2 in the Supplementary Appendix presents a scatter diagram that illustrates this close correspondence).

3 Inequality in lifetime earnings and in annual earnings, compared

Using these estimates of individual lifetime earnings, we calculate inequality in lifetime earnings in each year, annually from 1986 to 1996 and biannually from 1996 to 2018, among all men in our sample aged 31 to 59 in a given year and with positive untrimmed earnings in that year; and compare it to inequality in annual earnings in that year among the same group of men. Figure 1 presents a graphic comparison of inequality in annual and lifetime earnings for two commonly used inequality measures, the variance of log earnings and the Gini coefficient. (Table A2 in the Supplementary Appendix presents numerical values for the four graphs.)

Figure 1 here

Both the log variance of earnings and the Gini coefficient indicate that inequality in lifetime earnings increased much less than inequality in annual earnings over this period and its average level was substantially lower. The log variance of lifetime earnings rose by 0.014 between 1986 and 2028, from 0.202 to 0.216, a proportional increase of 6.6 percent, where the log variance

of annual earnings rose by 0.099, from 0.493 to 0.592, a proportional increase of 20.2 percent, three times greater than the proportional increase in the log variance of lifetime earnings. Similarly, the Gini coefficient of lifetime earnings rose by 0.016 between 1986 and 2018, from 0.236 to 0.252, a proportional increase of 7.0 percent, where the Gini coefficient of annual earnings increased by 0.055 over this period, from 0.308 to 0.363, a proportional increase of 17.9 percent, two and a half times greater than the proportional increase in the Gini of lifetime earnings. Thus, for both measures, the proportional rise in the inequality of lifetime earnings was much smaller than the proportional rise in the inequality of annual earnings, and for both measures the average level of inequality in lifetime earnings was substantially lower than for annual earnings. The correlation between inequality in lifetime and annual earnings over this period is 0.755 for the log variance of earnings and 0.918 for the Gini coefficient.[8]

These patterns of variation in the inequality of lifetime and annual earnings are robust to changes in specifications. Using a dollar threshold to trim low earnings values—rather than trimming the bottom 1 percent—has no effect on our measures of inequality in lifetime earnings, and little effect on the Gini coefficient of annual earnings, though it somewhat dampens the rise in the log variance of annual earnings during the great recession (shown graphically in Figure A4 in the Supplementary Appendix,). None of our findings are sensitive to the top-trimming method used; or to using predicted earnings at age 45 as our proxy for lifetime earnings, instead of age 40; or to extending the range of ages over which we calculate earnings inequality in each year from 31-59 to 25-59.

[8] Figure A3 and Table A2 in the Supplementary Appendix show very similar patterns for two other measures of inequality, the Theil generalized entropy index for $\alpha = 0$ and $\alpha = 1$.

4 A decomposition of the log variance of annual earnings

We use the predicted earnings trajectories from our estimation, $\hat{Y}_{it}$, and our proxy for average annual earnings over the lifetime, $\hat{Y}_{i40}$, to decompose the log variance of earnings. Taking logs, we denote $y_{it} = \log(Y_{it})$, $\hat{y}_{it} = \log(\hat{Y}_{it})$ and $\hat{Y}_i = \log(\hat{Y}_{i4})$, and write log annual earnings, y_{it}, as the sum of three terms:

$$y_{it} = \hat{Y}_i + (\hat{y}_{it} - \hat{Y}_i) + (y_{it} - \hat{y}_{it}) \quad (2)$$

where $\hat{Y}_i$ is the log of individual *i*'s lifetime average earnings; $\hat{y}_{it} - \hat{Y}_i$, the log difference between individual *i*'s predicted earnings in year *t* and his lifetime average earnings, is his predicted lifetime variation in year *t*; and $y_{it} - \hat{y}_{it}$, the log difference between individual *i*'s actual and predicted earnings in year *t*, is an unexpected earnings shock. Inequality in annual earnings, measured as the log variance of annual earnings, is then the sum of six terms:

$$\begin{aligned} V(y_{it}) \;=\; & V(\hat{Y}_i) + V(\hat{y}_{it} - \hat{Y}_i) + V(y_{it} - \hat{y}_{it}) + \\ & 2Cov(\hat{Y}_i,\ \hat{y}_{it} - \hat{Y}_i) + 2Cov(\hat{Y}_i,\ y_{it} - \hat{y}_{it}) + 2Cov(\hat{y}_{it} - \hat{Y}_i,\ y_{it} - \hat{y}_{it}) \end{aligned} \quad (3)$$

The six terms on the right-hand side represent:

- Inequality in lifetime earnings, measured as the log variance of lifetime earnings, $V(\hat{Y}_i)$.
- The dispersion of life-cycle variation in earnings, measured as the variance of the log difference between predicted earnings and estimated lifetime earnings, $V(\hat{y}_{it} - \hat{Y}_i)$.
- Residual volatility, measured as the variance of the log difference between actual and predicted annual earnings, $V(y_{it} - \hat{y}_{it})$.
- The covariance of life-cycle variation in earnings with estimated lifetime earnings, $Cov(\hat{Y}_i,\ \hat{y}_{it} - \hat{Y}_i)$ (we expand on this below).
- The covariance of earnings shocks with lifetime earnings, $Cov(\hat{Y}_i,\ y_{it} - \hat{y}_{it})$, a measure of the regressive incidence of volatility.

- The covariance of transitory shocks with life-cycle variation, $Cov(\hat{y}_{it} - \hat{Y}_i, y_{it} - \hat{y}_{it},)$, which increases when younger individuals experience more negative transitory shocks (as younger individuals generally earn below their lifetime average earnings.)

Figure 2 and Table 1 here

The seven terms in equation (3) are plotted in Figure 2. (Numerical values for the seven graphs are given in Table A3 in the Supplementary Appendix). The log variance of annual earnings, $V(y_{it})$, represented by the dotted black line in Figure 2, fluctuates without trend until the turn of the millennium, then rises moderately during the dot-com recession and more sharply after the Great Recession, to a peak of 0.825 in 2012, before falling back to levels that remain slightly above its average level of 0.515. Table 1 presents, in its leftmost column, the average level of each of the seven variables in 1986-2018. The two largest components of the log variance of annual earnings are the log variance of lifetime earnings, $V(\hat{Y}_i)$, and residual volatility, $V(y_{it} - \hat{y}_{it})$, each with an average level equal to about 42 percent of the average level of the log variance of annual earnings. The second column of Table 1 shows the change in the value of each row variable from 1986 to 2018; the third column shows the ratio of this change to the increase in the log variance of annual earnings over this period; and the rightmost column shows the correlations of the row variable with the log variance of annual earnings.

This decomposition highlights the dominant role of volatility in driving the increase in the inequality annual earnings in these years, measured as their log variance. Of the six component terms, $V(y_{it} - \hat{y}_{it})$, the log variance of residual shocks, our measure of volatility, exhibited a rise equal to over two thirds of the increase in the log variance of annual earnings from 1986 to 2018, with a correlation of 0.981 with $V(y_{it})$. The term with the second largest change in this period is the regressive incidence of these shocks, measured as twice the covariance of unanticipated earnings shocks with average lifetime earnings, $2 * Cov(\hat{Y}_i, y_{it} - \hat{y}_{it})$. Its large

increase in 1986-2018—equal to over half the increase in the log variance of annual earnings—indicates that earnings shocks became substantially more regressive in this period: more positive for individuals with higher lifetime earnings and more negative for individuals with lower lifetime earnings. Its correlation with $V(y_{it})$ was similarly high, 0.961. These increases were partly offset by a decline in the covariance of life-cycle variation and lifetime earnings, $Cov(\hat{Y}_i, \hat{y}_{it} - \hat{Y}_i)$, which may have reflected a secular increase in the share of high-skilled men in younger cohorts causing an out-of-equilibrium compression of the distribution of annual earnings. (In part B of the Supplementary Appendix, we illustrate this effect with a numerical example.) The log variance of average lifetime earnings, $V(\hat{Y}_i)$, varied much less in this period, contributing 14 percent of the increase in the log variance of annual earnings in 1986-2018, with a correlation of 0.75 with $V(y_{it})$. The log variance of life cycle variation, $V(\hat{y}_{it} - \hat{Y}_i)$, and its correlation with unexpected shocks, $Cov(\hat{y}_{it} - \hat{Y}_i, y_{it} - \hat{y}_{it})$, changed very little in this period and contributed little to the increase in the inequality of annual earnings.

5 Consistency with previous studies and robustness to the addition of new data

In this section we first estimate lifetime earnings inequality within rolling ten-year cohort groups and find that they are generally consistent with Guvenen et al.'s (2017, 2022) cohort-based estimates where the time periods overlap. Comparing these cohort-group measures to our calendar-year measures of inequality in lifetime earnings highlights the very different dynamic patterns they exhibit. Next, we demonstrate the robustness of our findings to the addition of future data by re-estimating our measures of year-by-year inequality in lifetime earnings for the period 1986-2008 using only data to 2008. We show that the restricted-data estimates closely match our full-data estimates for this period. Finally, we show that our residual measure of volatility is similar in magnitude and correlates closely with previous estimates of volatility derived from two-year overlapping panels when applied to the same data.

5.1 Inequality in lifetime earnings within cohort-groups

Guvenen et al. (2022) describe the evolution over time of inequality in lifetime earnings within single-year birth cohorts born between 1932 and 1958, using administrative SSA data. To gauge the consistency of our findings with theirs, we measure inequality in our estimated lifetime earnings within forty-one rolling ten-year cohort-groups, the oldest born in 1938-47, the youngest in 1978-87.[9] Figure 3 plots the log variance and Gini coefficient of estimated lifetime earnings within these forty-one rolling ten-year cohort-groups. It highlights the very different dynamic patterns exhibited by these cohort-based measures of inequality in lifetime earnings compared to the calendar-year measures shown in Figure 1.

Figure 3 here

To compare our estimates to those of Guvenen et al. (2022) where they overlap, we compute the standard deviation of log lifetime earnings for the twelve ten-year cohort-groups from 1938-47 to 1949-58 and compare it to the standard deviation of log lifetime earnings within single-year cohorts of males born between 1938 and 1958 shown in Guvenen et al. (2022, Figure 8, Panel A, where birth-cohorts are denoted by the year they turn 25, from 1963 to 1983). We find that both series show rising trends of similar magnitude:[10] their series shows an increase of about 20 percent over this period where we find a 14 percent increase. To further compare our results to theirs, we average their single-cohort means within each of our ten-year

[9] Our sample is too small to measure inequality within single-year birth cohorts, as they do.

[10] We find these similarities despite substantial differences in populations and data sources. Guvenen et al. (2022) estimate inequality within single-year cohorts of commerce and industry workers from administrative SSA data where we estimate inequality within ten-year cohort-groups from PSID survey data.

cohort-groups, divide it by the median of their ten single-cohort medians, and compare this ratio to our corresponding mean to median ratios. The correlation between the two series is 0.90, with both series exhibiting a moderately rising trend (Figure A7 in the Supplementary Appendix shows this graphically).

5.2 Dynamic consistency: robustness to the addition of new data

The validity of our approach depends on the first-stage estimates accurately anticipating average lifetime earnings, especially for younger cohorts for whom we have only short earnings histories at younger ages. To gauge the accuracy of our first stage, we estimate lifetime earnings from 1986 to 2008 among males aged 31-59 with positive earnings in each year, using PSID data only to 2008, and compare the results to our full-data estimates of lifetime earnings for this period, for the 1,713 males in our sample born in 1938-77. We find a correlation of 0.99 between the two sets of estimates. (Figure A6 in the Supplementary Appendix presents a scatter diagram.) We then use these restricted-data estimates of lifetime earnings to calculate year-by-year inequality in lifetime earnings from 1986 to 2008, measured as the log variance and as the Gini coefficient, which we compare, in Figure 4, to our unrestricted estimates for those years. The general shape and magnitudes are very similar. The restricted-data estimates are slightly lower than the full-data estimates, more so for later years, but the relative difference between the restricted and unrestricted measures never exceeds 9 percent for the log variance, and 4 percent for the Gini coefficient. Correlations between the restricted-data and full-data estimates are 0.78 for the log variance and 0.92 for the Gini coefficient.

Figure 4 here

5.3 Consistency with estimates of volatility derived from overlapping two-year panels

The pattern described by our long-panel measure of residual volatility is consistent with two widely used short-panel measures of "gross volatility" (Moffitt and Zhang, 2018): the variance

of the two-year log difference of earnings, $V(y_{it} - y_{i,t-2})$ and the variance of the two-year arc-difference of earnings, $2(Y_{it} - Y_{i,t-2})/(Y_{it} + Y_{i,t-2})$.[11] We calculate these measures for our data and compare them to our long-panel measure of residual volatility, $V(y_{it} - \hat{y}_{it})$, biannually from 1986 to 2018. We find a correlation of 0.88 between the variance of the two-year log-difference of earnings and our long-panel measure of residual volatility, and a correlation of 0.92 between the variance of the arc-difference of earnings and our long-panel measure. The magnitudes are also similar, our measure falling mostly between the two short-panel measures. (Figure A5 in the Supplementary Appendix plots the three graphs, along with the log variance of annual earnings.) Residualizing the differences by first regressing them on a polynomial in age yields nearly identical results, as does expanding the range of ages within which inequality is measured each year, from 31-59 to 25-59.

6 Conclusion

This is the first study of which we are aware that estimates inequality in lifetime earnings among the same population for which inequality in annual earnings is routinely measured: all economically active men within a given age range in a given year. To this end we use PSID data to estimate individual lifetime earnings trajectories from incomplete earnings histories combined with data on education, race and birth year, and derive from them estimates of individual lifetime earnings for men born between 1927 and 1987. From these, we calculate measures of the dispersion of lifetime earnings among economically active US men aged 31-

[11] Shin and Solon (2011), Moffitt and Gottschalk (2012), and Moffitt and Zhang (2018, 2020) estimate these measures from PSID data. Our short-panel estimates are nearly identical to Moffitt and Zhang's (2020) estimates. Carr and Wiemer (2018) and Moffitt (2020) provide comparative analysis of mobility measures drawn from different data sources.

59 in each year, annually from 1986 to 1996 and biannually from 1996 to 2018, which we then compare to inequality in annual earnings within the same population in each year. Moreover, the earnings trajectories we estimate allow us to represent annual earnings in each year as the sum of average lifetime earnings, the expected life-cycle deviation from the lifetime average in that year, and an unexpected earnings shock; and to represent the log variance of annual earnings in each year, as the sum of the log variances of lifetime earnings, life-cycle variation in earnings, and earnings volatility, plus twice their pairwise covariances.

We find that inequality in lifetime earnings increased much less, proportionally, than inequality in annual earnings in this period. The log variance of lifetime earnings increased by 6.6 percent where the log variance of annual earnings increased by 20.2 percent, and the Gini coefficient of lifetime earnings increased by only 7.0 percent where the Gini coefficient of annual earnings increased by 17.9 percent. In addition, we find that the average level of inequality in lifetime earnings was substantially lower than the average level of inequality in annual earnings during this period by either measure.

Applying our decomposition to quantify the contribution of different factors to the rise in the log variance of annual earnings between 1986 and 2018, we find that increased volatility and greater regressivity of earnings shocks contributed most of the rise, with the small increase in the inequality of lifetime earnings contributing only a small fraction. These positive effects were partly offset by a decline in the covariance of life-cycle variation and lifetime earnings, possibly reflecting a secular increase in the share of high-skilled men in younger cohorts, which compressed the distribution of annual earnings.

Comparing our results to previous studies, we use our estimates of lifetime earnings to compute inequality within forty-one overlapping ten-year cohort groups and compare them to Guvenen et al.'s (2022) findings on inequality in lifetime earnings within cohort groups, which

used complete earnings histories drawn from administrative data. We find that for overlapping cohorts, the two studies show similar trends, which are very different from the dynamic patterns exhibited by our calendar-year estimates of inequality in lifetime earnings. In addition, to test the robustness of our findings to the addition of new data, we estimate lifetime earnings and inequality to 2008 using only data up to 2008 and compare these estimates to our unrestricted estimates, using data to 2008, for these years. We find a close correspondence between the restricted and unrestricted estimates. Finally, we find that our residual measure of volatility is closely correlated with, and of similar magnitude to earlier short-panel measures of volatility.

A systematic and timely understanding of the sources of observed inequality in annual earnings can help address its social and economic implications more effectively.

Acknowledgments: We thank Richard Burkhouser, Danny Cohen-Zada, Leila Gautham, Mark Gradstein, Stephen Jenkins, Jacques Silber, and participants in meetings of the Society for the Study of Economic Inequality and the International Association for Research on Income and Wealth, and in seminars at Ben Gurion University and Bar Ilan University, for their helpful comments and suggestions.

The Panel Study of Income Dynamics notes that collection of its data, which we used in the current study, was partly supported by the National Institutes of Health under grant number R01 HD069609 and R01 AG040213, and the National Science Foundation under award numbers SES 1157698 and 1623684.

References

Aaberge, R and M Mogstad, 2015. Inequality in current and lifetime income. *Social Choice and Welfare* 44: 217–230.

Aaronson, S, 2002. The Rise in Lifetime Earnings Inequality among Men. Unpublished.

Bjorklund, A, 1993. A comparison between actual distributions of annual and lifetime income: Sweden 1951-89. *Review of Income and Wealth* 39: 377–386.

Blomquist, S, 1981. A Comparison of Distributions of Annual and Lifetime Income: Sweden around 1970. *Review of Income and Wealth* 27: 243–264.

Bönke, T, G Corneo, and H Lüthen. 2015. Lifetime Earnings Inequality in Germany. *Journal of Labor Economics* 33 (1): 171–208.

Bowlus, A and J-M Robin, 2004. Twenty years of rising inequality in U.S. lifetime labour income values. *Review of Economic Studies* 71: 709–742.

Burkhauser, R and J Poupore, 1997. A cross-national comparison of permanent inequality in the United States and Germany. *Review of Economics and Statistics* 79(1):10-17.

Burkhauser, R and K Couch, 2009. Intragenerational inequality and intertemporal mobility, in W Salverda, B Nolan and T Smeeding, (Eds.), *Oxford Handbook of Economic Inequality*. Oxford University Press, pp. 522–548.

Carr, M and E Wiemers, 2018. New evidence on earnings volatility in survey and administrative data, *AEA Papers and Proceedings* 108: 287-291.

Coronado, J, D Fullerton and T Glass, 2000. The Progressivity of Social Security, National Bureau of Economic Research, WP 7520.

Creedy, J, 1977. The distribution of lifetime earnings. *Oxford Economic Papers* 29: 412-429.

Creedy, J, 1999. Lifetime versus annual income distribution. In *Handbook of Income Inequality Measurement* (pp. 513-534). Springer, Dordrecht.

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

Friedman, M, 1957. *A Theory of the Consumption Function.* Princeton University Press.

Friesen, PH and Miller, D, 1983. Annual inequality and lifetime inequality. *The Quarterly Journal of Economics*, 98(1), 139-155.

Gottschalk, P and R Moffitt, 1994. The growth of earnings instability in the U.S. labor market. *Brookings Papers on Economic Activity* 1994: 217–72.

—, 2009. The rising instability of U.S. earnings. *Journal of Economic Perspectives* 23: 3–24.

Guvenen, F, G Kaplan, J Song, and J Weidner, 2017. Lifetime incomes in the United States over six decades, National Bureau of Economic Research, Working Paper No. 23371.

Guvenen, F, G Kaplan, J Song, and J Weidner, 2022. Lifetime earnings in the United States over six decades. *American Economic Journal: Applied Economics* 14: 446-479.

Guvenen, F, S Ozkan, and J Song, 2014. The nature of countercyclical income risk. *Journal of Political Economy* 122(3): 621-60.

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

Haider, S and G Solon, 2006. Life-Cycle Variation in the Association between Current and Lifetime Earnings. *American Economic Review* 96 (4): 1308-1320.

Jäntti, M and S Jenkins, 2013. Income mobility. In *Handbook of Income Distribution*, eds. Anthony B. Atkinson and Francois Bourguignon. North Holland.

Justman, M and A Krush, 2013. Less equal and less mobile: Evidence of a decline in intergenerational income mobility in the United States, Melbourne Institute of Applied Economic and Social Research WP 43/13, University of Melbourne.

Justman, M and H Stiassnie, 2021. Intergenerational income mobility in lifetime incomes. *Review of Income and Wealth* 67 (4): 928-949.

Kopczuk, W, E Saez, and J Song, 2010. Earnings Inequality and Mobility in the United States: Evidence from Social Security Data since 1937. *Quarterly Journal of Economics* 125:91-128.

Leonesio, Michael V., and Lisa Del Bene. 2011. "The Distribution of Annual and Long-Run US Earnings:1981–2004." *Social Security Bulletin* 71 (1): 17–33.

Meghir C, and L Pistaferri, 2011. Earnings, consumption, and lifecycle choices. In: Ashenfelter O, Card D (eds) *Handbook of Labor Economics*. Elsevier, Amsterdam.

Moffitt, R, 2020. Reconciling trends in US male earnings volatility: Results from a four data set project. National Bureau of Economic Research, Working Paper No. 23371.

Moffitt, R and P Gottschalk, 2012. Trends in the transitory variance of male earnings: Methods and evidence. *Journal of Human Resources* 47 (1): 204–36.

Moffitt, R and S Zhang, 2018. The PSID and income volatility: Its record of seminal research and some new findings. Annals of the American Association of Political Science, 680.

Moffitt, R and S Zhang, 2020. Estimating Trends in Male Earnings Volatility with the Panel Study of Income Dynamics. NBER WP 27674.

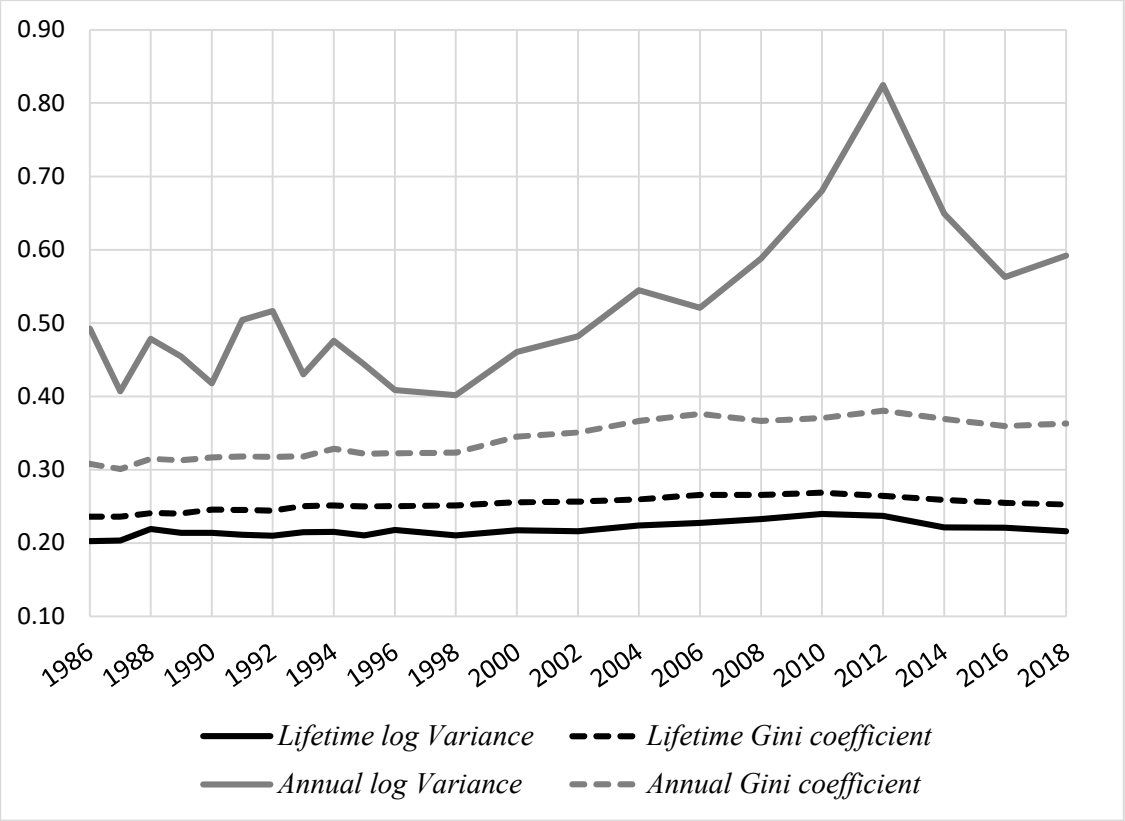
Mookherjee D and A Shorrocks, 1982. A Decomposition Analysis of the Trend in UK Income Inequality. *Economic Journal* 92: 886-902.

Panel Study of Income Dynamics, public use dataset. Produced and distributed by the Survey Research Center, Institute for Social Research, University of Michigan, Ann Arbor.

Shin, D and G Solon, 2011. Trends in men's earnings volatility: What does the Panel Study of Income Dynamics show? *Journal of Public Economics* 95: 973–82.

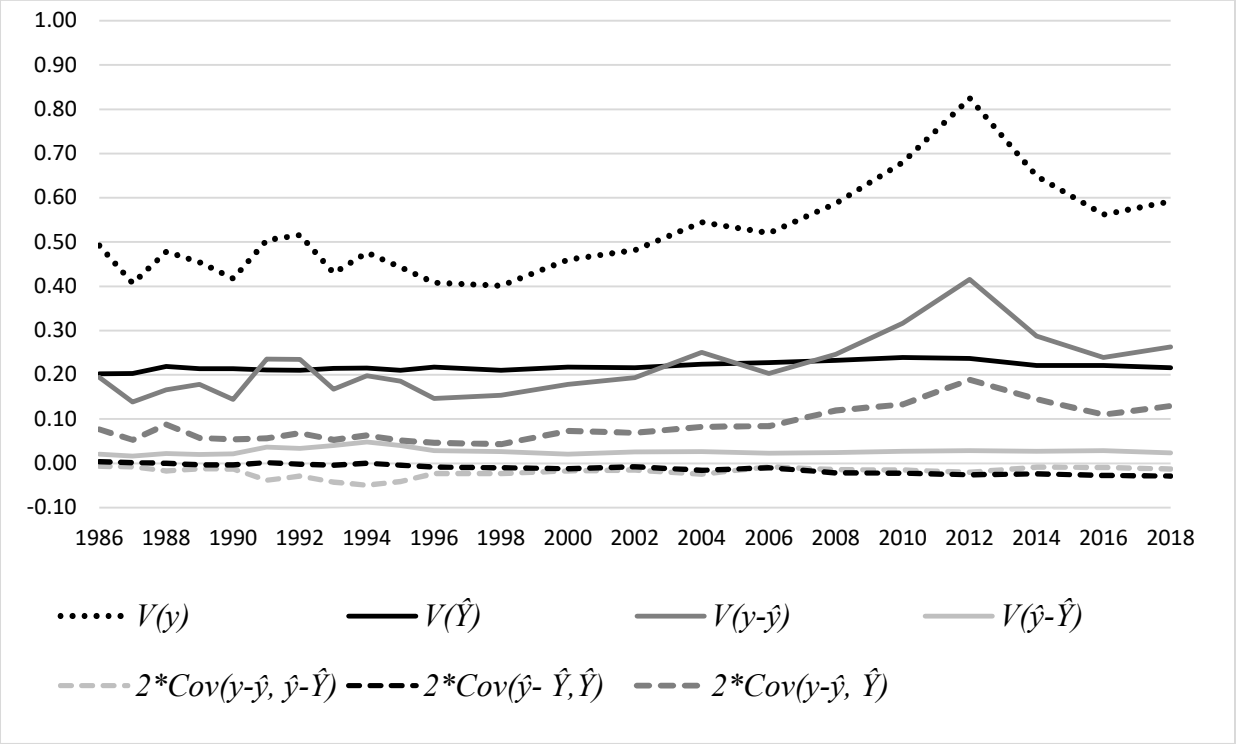
Shorrocks, A, 1978, Income inequality and income mobility, *Journal of Economic Theory,* 19: 376-393.

Figure 1. *The log variance and Gini coefficient of annual and lifetime earnings in calendar years, annually from 1986 to 1996 and biannually from 1996 to 2018.*



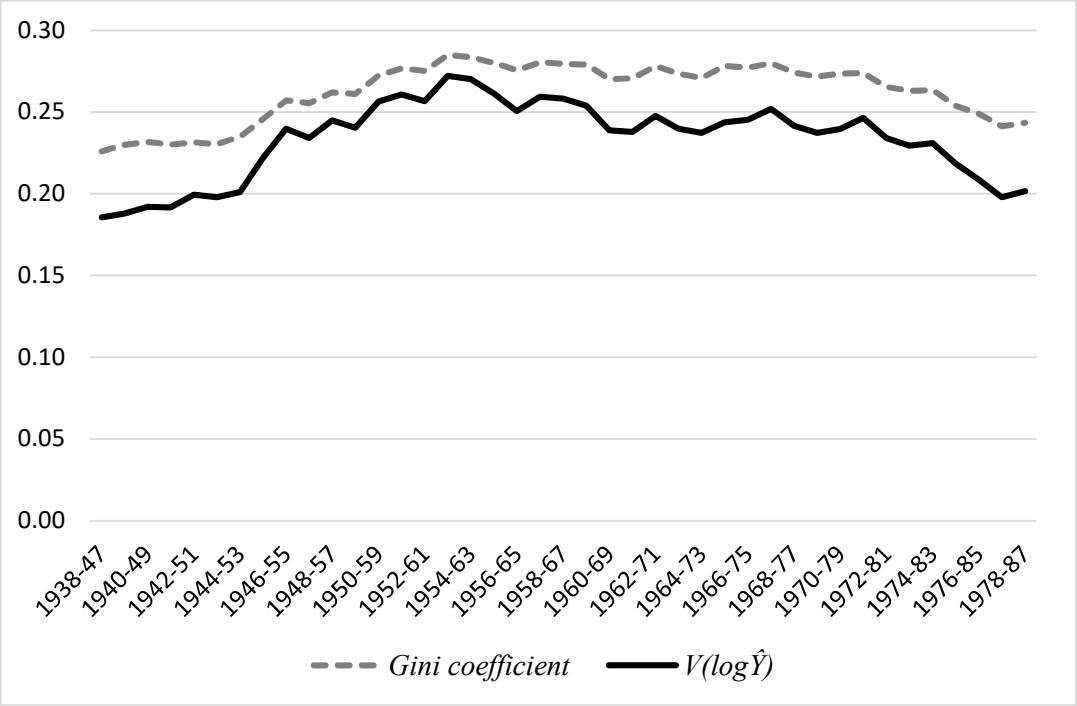
Note: Our data set comprises male heads of households with at least three untrimmed earnings observations between the ages 25-59 after trimming the top and bottom 1 percent in each year. Measures of annual inequality in a given year measure inequality in annual earnings among individuals in our data set aged 31-59 and with positive untrimmed earnings in that year. Measures of lifetime inequality in a given year measure inequality in estimated lifetime earnings among the same group of individuals. Average lifetime earnings are predicted earnings at age 40, estimated from a regression of annual earnings on a cubic polynomial in age and its interaction with education, race, and cohort group indicators, and on individual random effects.

Figure 2. *Decomposition of the variance of log annual earnings as the sum of the variances and pairwise covariances of log lifetime earnings, life-cycle variation, and residual volatility, annually from 1986 to 1996 and then biannually to 2018*



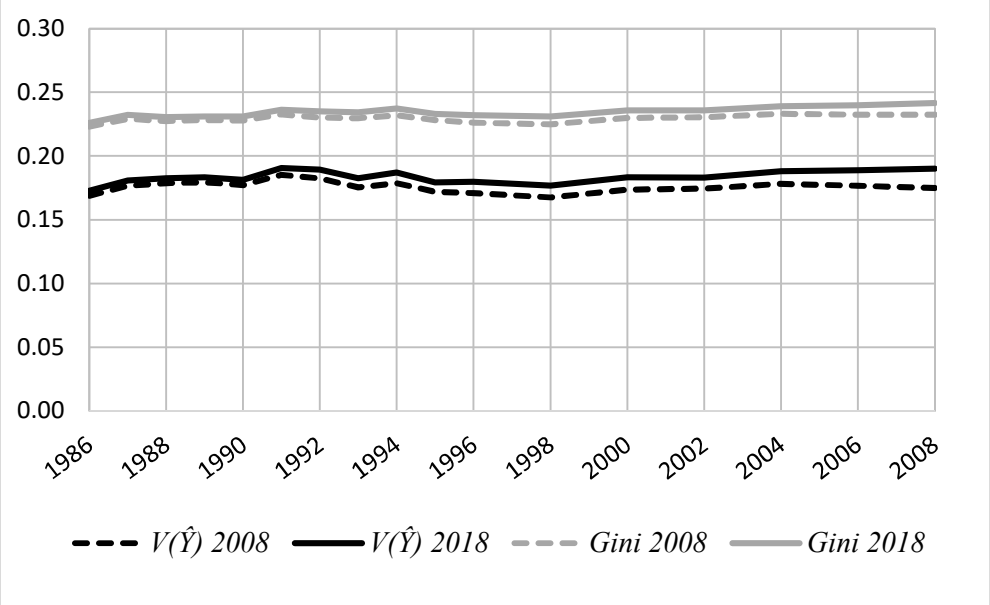
Note: Our data set comprises male heads of households with at least three untrimmed earnings observations between the ages 25-59 after trimming the top and bottom 1 percent in each year. All measures in a given year calculated among individuals in our data set aged 31-59 and with positive untrimmed earnings in that year. The variable y_{it} denotes the log annual earnings of individual i in year y; $\hat{y}_{it}$ denotes individual i's log predicted annual earnings year t; and $\hat{Y}_i$ denotes individual i's log estimated average lifetime earnings. The decomposition shown here follows from: $y_{it} = \hat{Y}_i + (\hat{y}_{it} - \hat{Y}_i) + (y_{it} - \hat{y}_{it})$.

Figure 3. *The Gini coefficient and the log variance of lifetime earnings within rolling 10-year cohort groups, 1938-87*



Note: Our dataset and lifetime earnings estimates are as described in the note to Figure 1. The Gini coefficients and log variance of lifetime earnings are calculated over all individuals in our data set in a given ten-year birth-cohort group, from 1938-47 to 1978-86.

Figure 4. *Two measures of year-by-year inequality in estimated lifetime earnings, 1986-2008, using only data to 2008, and using all data to 2018, annually from 1986 to 1996 and then biannually to 2008.*



Note: The solid lines are the lifetime Gini coefficient and log variance estimated from the full data set, shown in Figure 1. The dashed lines are the lifetime Gini coefficient and log variance estimated in the same way from data to 2008.

Table 1. *Decomposition of inequality in annual earnings: relative average values, ratio of change over the full period, and correlation with annual inequality, 1986-2018*

	Average value of row variable, 1986-2018	*Change in row variable, 1986-2018*	*Ratio of change in row variable to change in* $V(y_{it})$	*Correlation of row variable with* $V(y_{it})$
Log variance of annual earnings, $V(y_{it})$	0.515	0.099	1.000	1.000
Log variance of estimated lifetime earnings, $V(\hat{Y}_i)$	0.218	0.013	0.135	0.755
Log variance of life-cycle variation, $V(\hat{y}_{it} - \hat{Y}_i)$	0.028	0.003	0.031	-0.011
Log variance of residual shocks (volatility), $V(y_{it} - \hat{y}_{it})$	0.215	0.069	0.692	0.981
Covariance of life-cycle variation and lifetime earnings, $2 * Cov(\hat{y}_{it} - \hat{Y}_i, \hat{Y}_i)$	-0.010	-0.032	-0.326	-0.728
Regressive incidence of shocks, $2 * Cov(y_{it} - \hat{y}_{it}, \hat{Y}_i)$	0.084	-0.053	0.533	0.961
Covariance of volatility and life-cycle variation, $2 * Cov(y_{it} - \hat{y}_{it}, \hat{y}_{it} - \hat{Y}_i)$	-0.020	-0.006	-0.064	0.225

Notes: Table rows refer to the seven graphs shown in Figure 2. The variable y_{it} *denotes the log annual earnings of individual i in year y;* $\hat{y}_{it}$ *denotes individual i's log predicted annual earnings year t; and* $\hat{Y}_i$ *denotes individual i's log estimated average lifetime earnings. The decomposition shown here follows from:* $y_{it} = \hat{Y}_i + (\hat{y}_{it} - \hat{Y}_i) + (y_{it} - \hat{y}_{it})$. *All measures in a given year calculated among individuals in our data set aged 31-59 and with positive untrimmed earnings in that year.*

Appendix, Part A: Supplementary tables and figures

Table A1. Descriptive statistics

		Age			Years of schooling				Mean	Median
Year	N	25-34	35-54	55-59	0-12	13-15	16+	White	earnings	earnings
1986	967	22%	65%	13%	40%	24%	36%	93%	60,617	56,821
1987	969	21%	67%	12%	39%	25%	36%	93%	62,097	57,443
1988	976	21%	69%	11%	39%	25%	36%	93%	61,714	55,841
1989	980	20%	70%	10%	38%	25%	37%	92%	60,745	53,693
1990	979	18%	73%	9%	36%	25%	39%	92%	61,818	54,671
1991	1008	19%	72%	9%	37%	25%	38%	91%	59,881	52,902
1992	1011	18%	74%	9%	37%	25%	38%	92%	62,602	56,359
1993	1067	18%	73%	9%	36%	25%	39%	93%	63,364	54,988
1994	1080	19%	73%	8%	36%	25%	38%	92%	63,162	53,867
1995	1070	16%	76%	8%	36%	24%	40%	92%	63,824	54,259
1996	1064	16%	76%	8%	34%	25%	41%	93%	65,465	56,078
1998	1060	13%	77%	10%	34%	24%	42%	93%	70,549	60,440
2000	1093	14%	75%	12%	34%	25%	42%	92%	73,210	60,874
2002	1091	15%	72%	13%	33%	26%	41%	92%	71,044	57,610
2004	1168	17%	69%	14%	30%	27%	42%	91%	73,119	59,581
2006	1126	17%	68%	15%	30%	29%	41%	91%	76,376	60,690
2008	1145	18%	66%	15%	31%	28%	41%	91%	72,088	58,613
2010	1081	20%	64%	16%	30%	28%	42%	91%	70,246	58,810
2012	1084	21%	64%	15%	29%	28%	43%	91%	70,925	58,449
2014	1025	21%	65%	14%	29%	28%	43%	90%	71,702	57,893
2016	1016	19%	67%	14%	28%	28%	44%	90%	74,530	62,373
2018	979	18%	69%	13%	25%	29%	45%	90%	76,463	63,000

See notes to Figure 1.

Table A2. Year-by-year inequality in annual and lifetime earnings, four measures

Year	*N*	*Log variance*		*Gini coefficient*		*Theil, α=0*		*Theil, α=1*	
		Annual	*Lifetime*	*Annual*	*Lifetime*	*Annual*	*Lifetime*	*Annual*	*Lifetime*
1986	967	0.493	0.202	0.308	0.236	0.194	0.095	0.157	0.089
1987	969	0.407	0.203	0.301	0.236	0.172	0.095	0.148	0.089
1988	976	0.479	0.219	0.315	0.241	0.196	0.101	0.164	0.093
1989	980	0.455	0.214	0.313	0.241	0.190	0.099	0.162	0.092
1990	979	0.418	0.214	0.317	0.246	0.185	0.101	0.167	0.097
1991	1,008	0.504	0.211	0.318	0.245	0.201	0.101	0.168	0.096
1992	1,011	0.516	0.210	0.318	0.244	0.205	0.100	0.168	0.095
1993	1,067	0.430	0.215	0.318	0.250	0.187	0.104	0.168	0.100
1994	1,080	0.476	0.215	0.328	0.251	0.203	0.104	0.180	0.101
1995	1,070	0.443	0.210	0.322	0.250	0.192	0.102	0.172	0.100
1996	1,064	0.408	0.218	0.322	0.251	0.184	0.104	0.169	0.100
1998	1,060	0.402	0.210	0.323	0.251	0.185	0.103	0.171	0.101
2000	1,093	0.461	0.217	0.345	0.256	0.212	0.107	0.199	0.105
2002	1,091	0.482	0.216	0.351	0.256	0.220	0.107	0.204	0.106
2004	1,168	0.545	0.224	0.367	0.260	0.245	0.110	0.226	0.108
2006	1,126	0.521	0.227	0.376	0.266	0.248	0.114	0.241	0.114
2008	1,145	0.588	0.233	0.367	0.266	0.253	0.115	0.226	0.114
2010	1,081	0.680	0.240	0.370	0.269	0.274	0.118	0.232	0.117
2012	1,084	0.825	0.237	0.381	0.264	0.306	0.115	0.246	0.112
2014	1,025	0.649	0.221	0.369	0.259	0.266	0.109	0.229	0.107
2016	1,016	0.562	0.221	0.360	0.255	0.242	0.104	0.216	0.104
2018	979	0.592	0.216	0.363	0.253	0.252	0.102	0.223	0.102

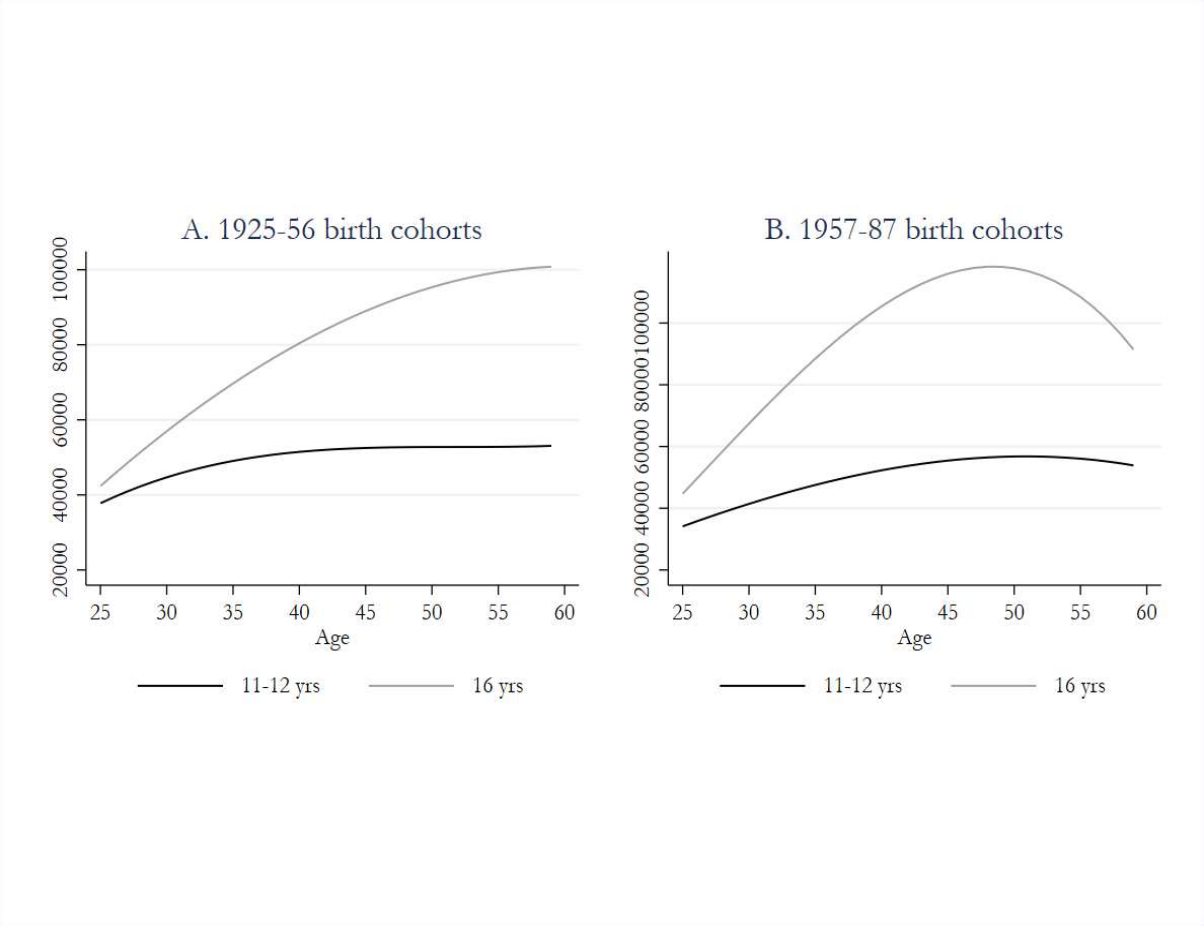
Note: Numerical values for the graphs shown in Figure 2 and Figure A3.

Table A3. Decomposition of the log variance of annual earnings

		Annual earnings	Lifetime earnings	Short-term shocks	Lifecycle variation	Lifetime and shocks	Lifetime and lifecycle	Lifetime and shocks
N	*Year*	*V(y)*	*V(Ŷ)*	*V(y-ŷ)*	*V(ŷ-Ŷ)*	*2*Cov(y-ŷ, ŷ-Ŷ)*	*2*Cov(ŷ-Ŷ,Ŷ)*	*2*Cov(y-ŷ, Ŷ)*
967	1986	0.493	0.202	0.194	0.021	-0.006	0.004	0.077
969	1987	0.407	0.203	0.139	0.017	-0.007	0.002	0.053
976	1988	0.479	0.219	0.166	0.022	-0.017	0.000	0.088
980	1989	0.455	0.214	0.178	0.020	-0.012	-0.003	0.058
979	1990	0.418	0.214	0.144	0.021	-0.013	-0.003	0.055
1,008	1991	0.504	0.211	0.236	0.036	-0.038	0.002	0.057
1,011	1992	0.516	0.210	0.235	0.034	-0.028	-0.002	0.068
1,067	1993	0.430	0.215	0.167	0.041	-0.042	-0.004	0.053
1,080	1994	0.476	0.215	0.198	0.049	-0.049	0.000	0.063
1,070	1995	0.443	0.210	0.186	0.040	-0.041	-0.004	0.051
1,064	1996	0.408	0.218	0.147	0.028	-0.023	-0.008	0.047
1,060	1998	0.402	0.210	0.154	0.027	-0.023	-0.010	0.044
1,093	2000	0.461	0.217	0.179	0.021	-0.017	-0.012	0.073
1,091	2002	0.482	0.216	0.193	0.026	-0.015	-0.007	0.069
1,168	2004	0.545	0.224	0.251	0.027	-0.024	-0.016	0.083
1,126	2006	0.521	0.227	0.203	0.023	-0.008	-0.010	0.084
1,145	2008	0.588	0.233	0.247	0.024	-0.014	-0.021	0.120
1,081	2010	0.680	0.240	0.317	0.027	-0.014	-0.022	0.133
1,084	2012	0.825	0.237	0.416	0.028	-0.020	-0.026	0.189
1,025	2014	0.649	0.221	0.288	0.027	-0.008	-0.024	0.145
1,016	2016	0.562	0.221	0.239	0.029	-0.009	-0.027	0.110
979	2018	0.592	0.216	0.263	0.024	-0.013	-0.028	0.130

Note: Numerical values for the graphs shown in Figure 2.

Figure A1. Four estimated age-earnings profiles, by years of schooling and cohort group



Note: Graphs estimated from equation (1), averaging individual random effects.

Figure A2. *Predicted earnings at age 40 against average actual earnings, for a subsample of males born in 1937-57 with 10 or more untrimmed earnings observations (2018 PCE prices)*

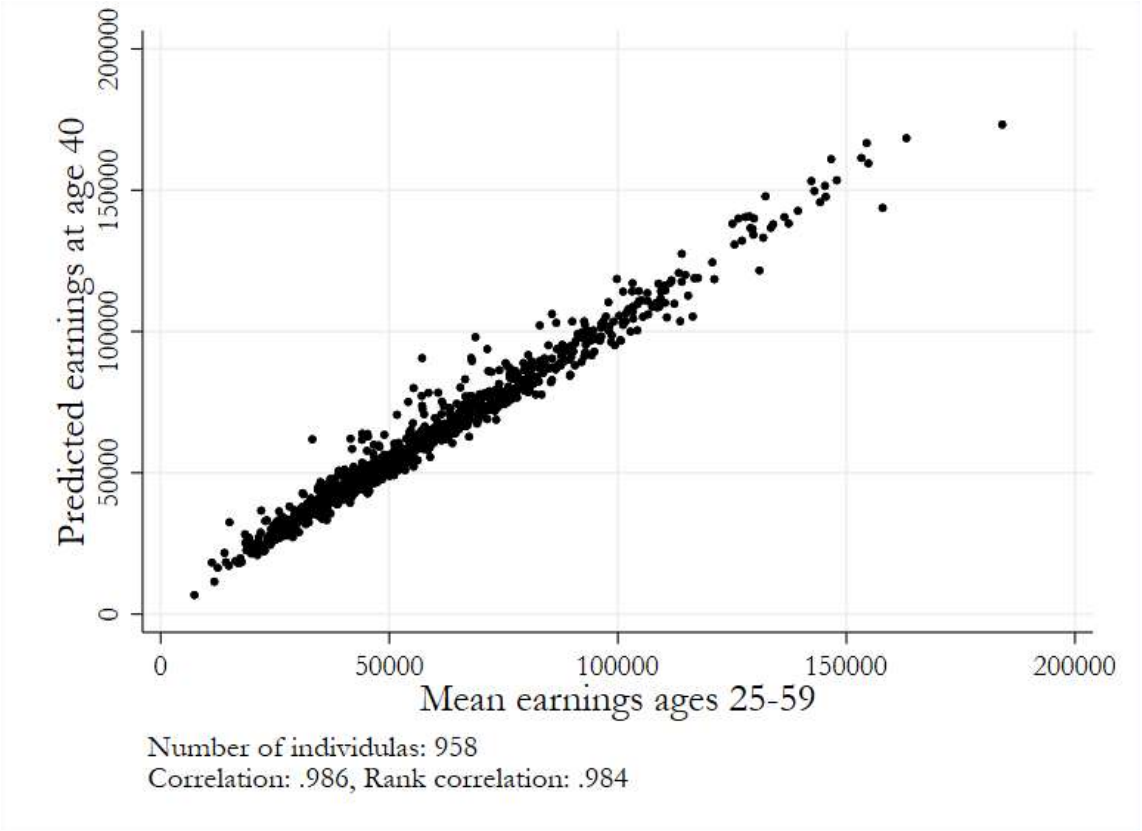
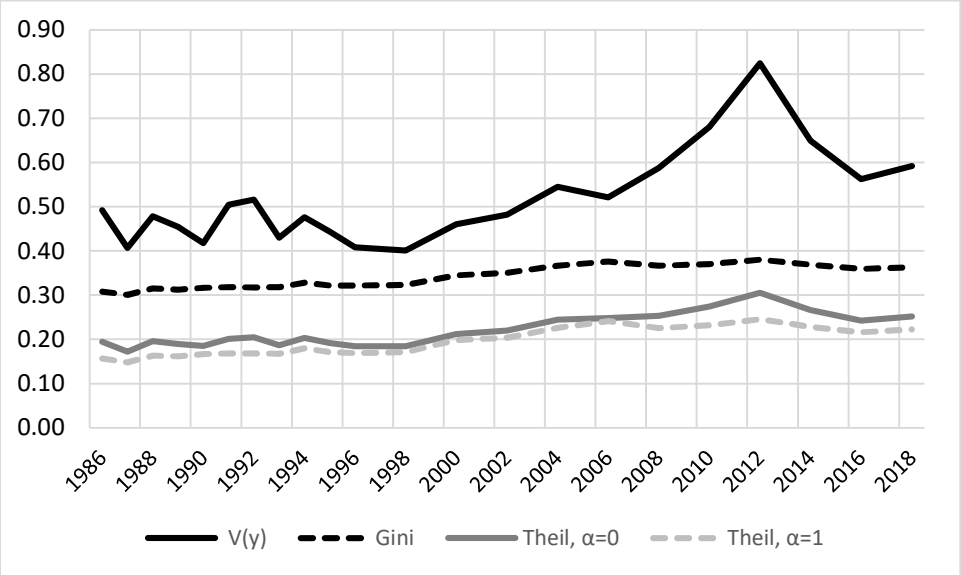
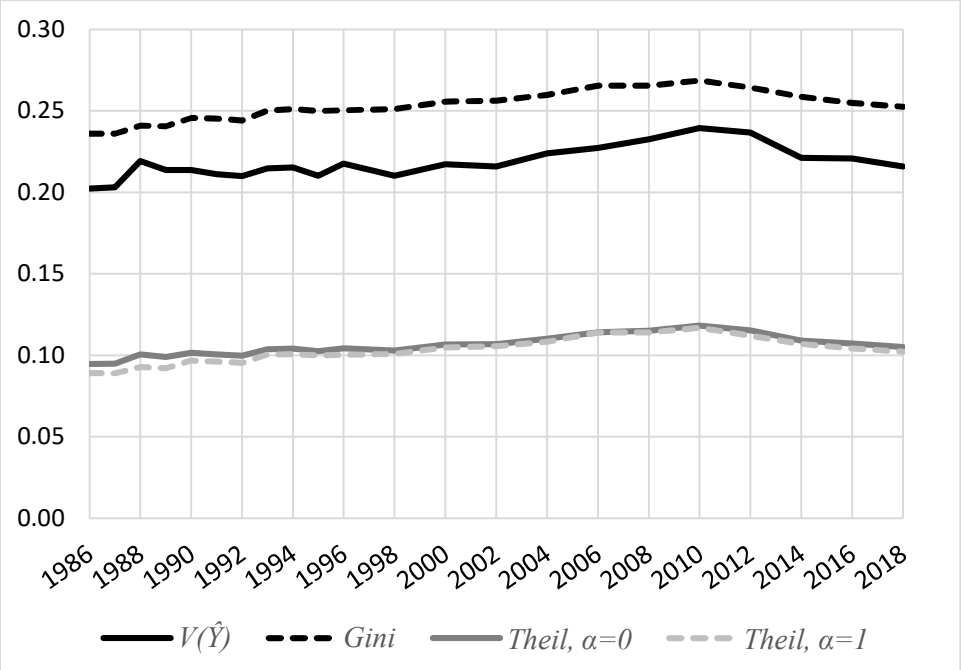


Figure A3. Four measures of inequality in annual and lifetime earnings

Annual earnings

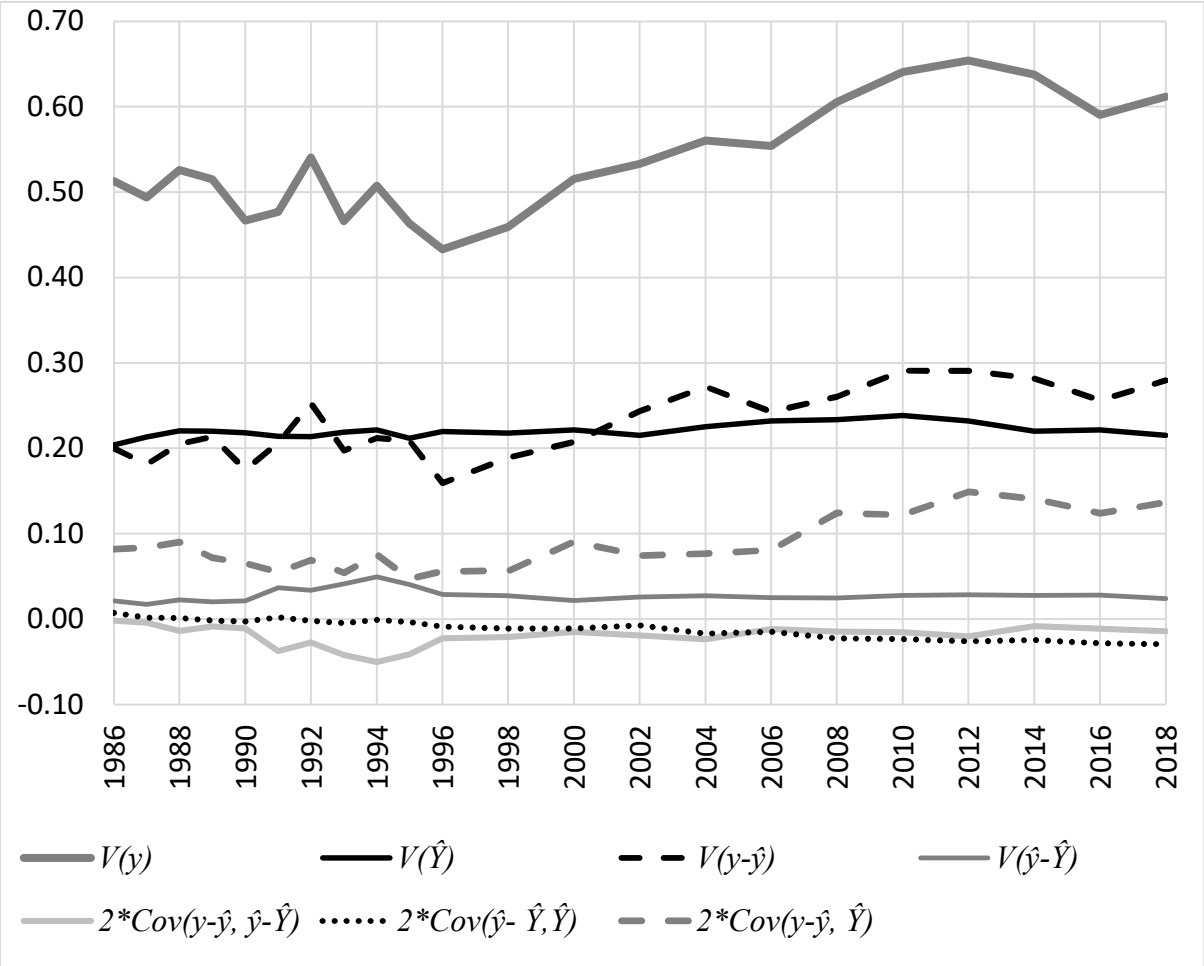


Lifetime earnings



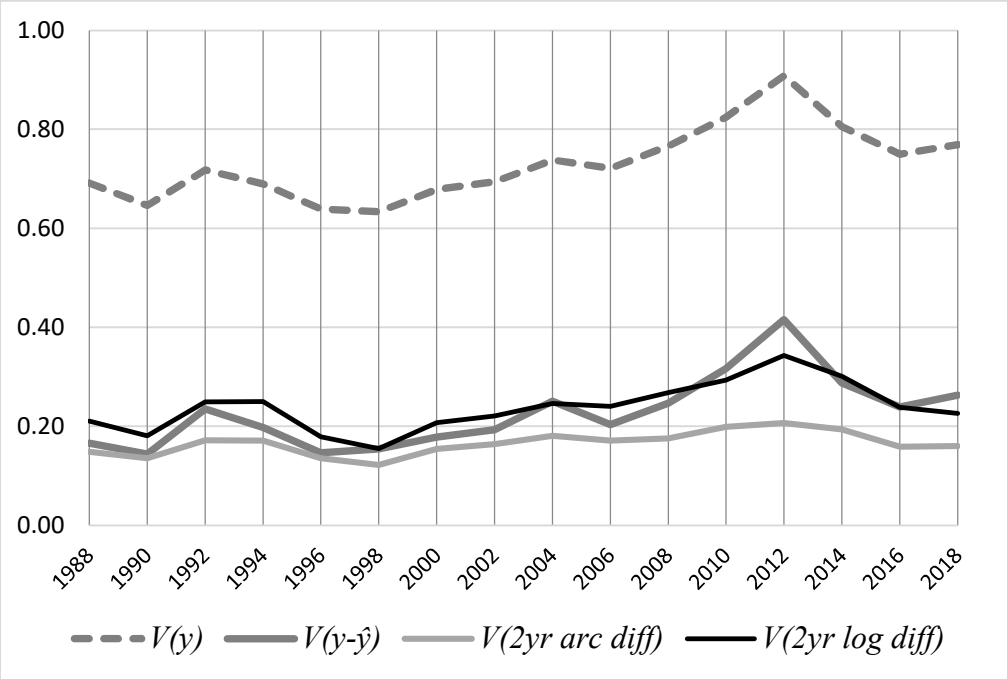
Note: The eight graphs include the four graphs shown in Figure 1 and all eight graphs refer to the same population as described in the note to Figure 1.

Figure A4. *Decomposition of the yearly variance of log annual earnings, 1986-2018: variances and pairwise covariances of log lifetime earnings, life-cycle variation, and residual volatility; dollar threshold linked to minimum wage*



Note: This figure recalculates Figure 2 except that earnings are bottom-trimmed below a threshold equal to thirteen weeks of full-time work at half the minimum wage, instead of trimming the bottom 1 percent in each year. All measures in a given year calculated among individuals in our data set aged 31-59 and with positive untrimmed earnings in that year. The variable y_{it} denotes the log annual earnings of individual i in year y; $\hat{y}_{it}$ denotes individual i's log predicted annual earnings year t; and $\hat{Y}_i$ denotes individual i's log estimated average lifetime earnings. The decomposition follows from: $y_{it} = \hat{Y}_i + (\hat{y}_{it} - \hat{Y}_i) + (y_{it} - \hat{y}_{it})$.

Figure A5. *Two short-panel measures of volatility; $V(y - \hat{y})$, our residual measure of volatility; and annual earnings inequality, $V(y)$; 1986-2018, biannually*



Note: Male heads of households aged 31-59 with positive untrimmed earnings in the year, after trimming the top and bottom 1%. Y_{it} is individual i's earnings in year t and $y_{it} = \log Y_{it}$. The two-year arc-difference of earnings is $2(Y_{it} - Y_{i,t-2})/(Y_{it} + Y_{i,t-2})$, and the two-year log difference is $y_{it} - y_{i,t-2}$.

Figure A6. *Estimated lifetime earnings for males born in 1938-77 using only data up to 2008, plotted against our full data estimates (2018 PCE prices)*

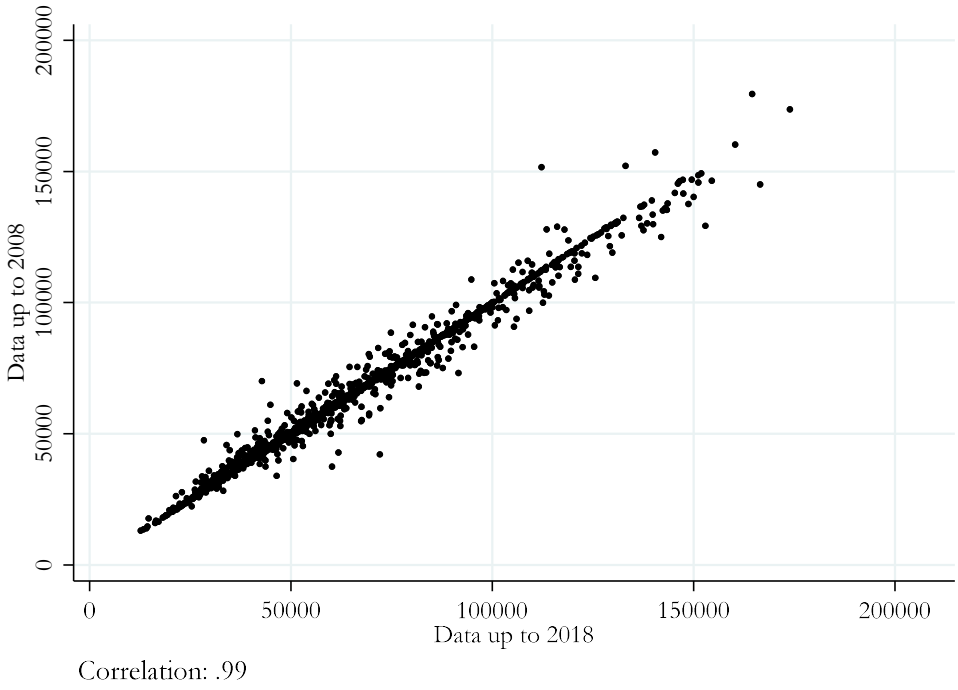
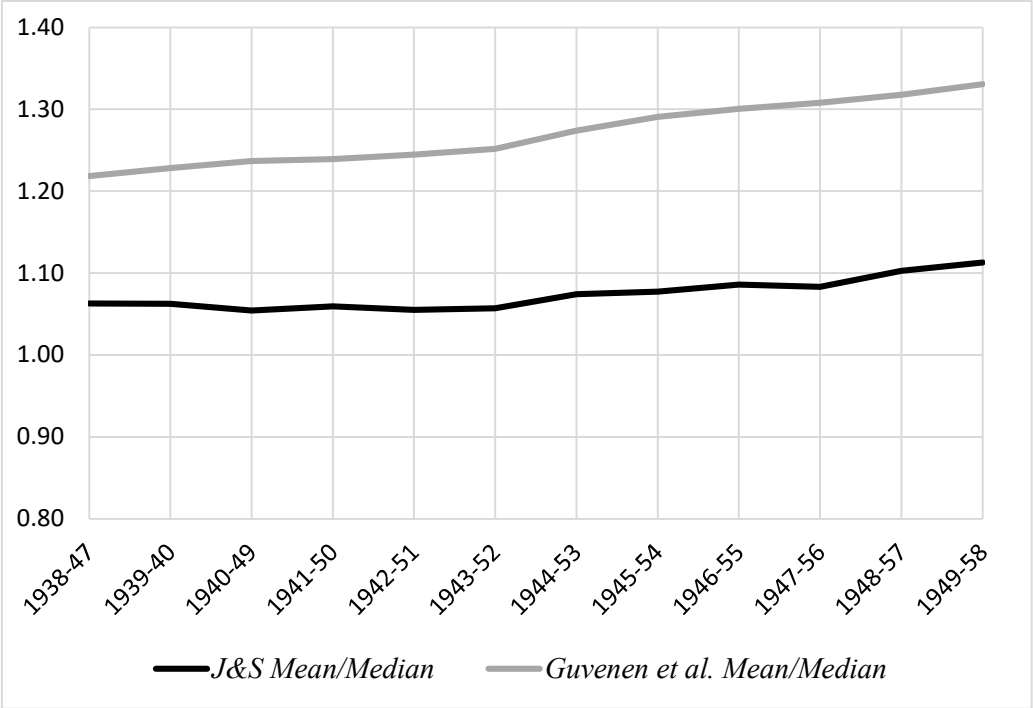


Figure A7. *Mean to median ratios of lifetime earnings within ten-year cohort groups, 1938-47 to 1949-58, our estimates compared to Guvenen et al. (2017)*



Note: "Guvenen et al. Mean/Median" denotes cohort-group mean to median ratios we constructed from Guvenen et al.'s (2017, Table A1) annual values by taking the within-group averages of their yearly means and the median of their yearly medians. They identify cohorts by the year they turn 25; we identify cohorts by their birth year. "J&S Mean/Median" denotes our mean to median ratios within each ten-year cohort group.

Appendix, Part B

An example of how the diffusion over time of a change in the skill level of the work force can temporarily compress the distribution of earnings, reducing inequality in annual earnings in the short-term much more than it reduces inequality in lifetime earnings.

The following numerical example illustrates how gradual structural change in the age distribution of lifetime earnings trajectories, possibly driven by technological progress, can temporarily affect inequality in annual earnings through its effect on the covariance of lifetime earnings with life-cycle variation in earnings. Consequently, though this structural change has an identical effect on current and lifetime inequality in the long term, it causes wide swings in current inequality divorced from its long-term effect. Substantial skill deepening, measured by years of education, characterized the US labor force in recent decades.

Consider an overlapping generations economy with a fixed population of eight individuals, each living two periods, that we follow over three periods. Half the population in the first (and each successive) period is young (Y) and half is old (O). Each person is either low skilled and employed in a lower-skilled, lower earnings occupation (L), or high skilled and employed in a higher-skilled, higher-earning occupation (H). Each occupation has its fixed earnings trajectory. Current earnings are: $Y_{LY} = 20$, $Y_{HO} = 60$, $Y_{HY} = 40$, and $Y_{HO} = 120$; and average lifetime earnings are $Y_L = 40$ and $Y_H = 80$. There are no transitory shocks and both earnings trajectories have the same ratio of old to young earnings.

Table B1 presents annual and average lifetime earnings by generation in each period. In the first period, half of each generation is low-skilled, and the other is high-skilled. Technological change increases demand for high-skilled workers, which is satisfied gradually. In the second period, the first-period older generation exits the economy, the younger generation becomes older, and a new younger generation enters the economy with a higher skill ratio than the older generation: of the four new entrants one is low-skilled and three are

high-skilled. In the third period, the new younger generation again comprises one low-skilled and three high-skilled workers, so now the two active generations have the same occupational structure, completing the transition to a new, higher-skilled steady state.

Table B1: *The effect of gradual skill deepening on current and average lifetime earnings*

		Young				*Old*			
Period 1	*Annual earnings*	20	20	40	40	60	60	120	120
	Avg lifetime earnings	40	40	80	80	40	40	80	80
Period 2	*Annual earnings*	20	40	40	40	60	60	120	120
	Avg lifetime earnings	40	80	80	80	40	40	80	80
Period 3	*Annual earnings*	20	40	40	40	60	120	120	120
	Avg lifetime earnings	40	80	80	80	40	80	80	80

Taking logarithms, we write: $y_{it} = \hat{Y}_i + (y_{it} - \hat{Y}_i)$, where t = 1,2,3 is a period index and we use this to decompose the log variance of current earnings as the sum of three terms: $V(y_{it}) = V(\hat{Y}_i) + V(y_{it} - \hat{Y}_i) + 2Cov(y_{it} - \hat{Y}_i, \hat{Y}_i)$, as in equation (2) in the text. The top three rows of Table B2 present this decomposition year by year. Change in the log variance of earnings in "the long run", from period 1 to 3, is the same for current and lifetime earnings, a decline of 0.03. However, the interim sees large swings in current inequality, first falling then rising, reflecting changes in the covariance of lifecycle change and lifetime earnings. The disproportionate entry of higher-skilled young generation workers in period 2, increasing the covariance of lifetime earnings with life-cycle variation in earnings, temporarily compresses the earnings distribution, sharply lowering inequality in current earnings. Then, in period 3, the recurring influx of higher-skilled younger generation restores the generational balance in skill distribution, the covariance of lifetime earnings with life-cycle variation in earnings bounces back to its first-period level of zero, while lifetime inequality declines, causing inequality in current earnings to *rise* to a level slightly lower than its level in the first period. Fig. B1 shows

this graphically. Lifetime inequality declines moderately and steadily while current inequality fluctuates strongly, driven by fluctuation in the age distribution of skills.

Table B2: *Decomposition of change in the log variance of current earnings over time*

	Log variance of current earnings, $V(y_{it})$	*Log variance of lifetime earnings* $V(\hat{Y}_i)$	*Log variance of life-cycle change* $V(y_{it} - \hat{Y}_i)$	*Covariance of life-cycle change and lifetime earnings* $2Cov(y_{it} - \hat{Y}_i, \hat{Y}_i)$
Period 1 value	0.422	0.120	0.302	0
Period 2 value	0.319	0.113	0.302	-0.095
Period 3 value	0.392	0.090	0.302	0
Change, period 1 to 2	-0.103	-0.008	0	-0.095
Change, period 2 to 3	0.073	-0.023	0	0.095
Change, period 1 to 3	-0.030	-0.030	0	0
Share of change 1 to 2		7.3%	0	92.7%
Share of change 2 to 3		-31.0%	0	131.0%
Share of change 1 to 3		100.0%	0	0

Figure B1: *Changes in the inequality of annual earnings,* $V(y_{it})$*, in lifetime inequality* $V(\hat{Y}_i)$*, and in twice the covariance of life-cycle variation in earnings and lifetime earnings,* $2Cov(y_{it} - \hat{Y}_i, \hat{Y}_i)$*, when the share of high-skilled workers increases gradually*

