Sources of inequality in annual earnings:
Lifetime inequality, life-cycle variation, transitory volatility

Moshe Justman[a] and Hadas Stiassnie[b]

Abstract

We estimate expected earnings trajectories and average lifetime earnings for males born between 1927-1987 from PSID data and use this to systematically represent inequality in annual earnings as the sum of inequality in lifetime earnings, dispersion from life-cycle variation in earnings, and unanticipated transitory volatility. Analyzing the increase in annual earnings inequality in the US between 1986 and 2018, we attribute 69% of the rise to increased volatility, 14% to increased lifetime inequality, and most of the rest to a large rise in the regressive incidence of earnings shocks partly offset by the rising share of younger high earners.

[a] Department of Economics, Ben Gurion University, justman@bgu.ac.il; corresponding author.
[b] Department of Economics, Ben Gurion University, hadas.stia@gmail.com

Sources of inequality in annual earnings:
Lifetime inequality, life-cycle variation, transitory volatility

1. Introduction

Measures of dispersion in annual earnings, used universally to track socio-economic inequality over time and compare it across populations, reflect not only the underlying dispersion of lifetime earnings but also the joint distribution of age or work experience in the labor force with different occupations and skill levels, each with its characteristic earnings trajectory, and short-term transitory volatility, or instability, much of which can be smoothed over time. Both volatility and the joint distribution of age and skill levels have seen dramatic changes over the last three decades, as the result of rapid technological innovation in the production and distribution of goods and services, the globalization of international trade and finance, and large waves of migration. Separating the effect of these changes on inequality in annual earnings allows us to assess the extent to which they reflect change in the underlying inequality of lifetime earnings and contributes to a better understanding of the nature of these recent changes in the joint distribution of age and occupations and in transitory volatility.

Short-term, transitory volatility that can be smoothed over time, thus inflating inequality in annual earnings beyond its underlying, "true" level, has been extensively studied. Shorrocks (1978) suggested looking at moving averages of earnings, to neutralize the effect of short-term volatility, and numerous studies have since implemented this and other related approaches (Burkhouser and Pupore, 1997, Gottschalk and Moffitt, 1994, 2009, among many others). Other studies estimate volatility directly from the short-term dispersion of relative change in earnings, which applied to the United States, find substantial earnings volatility, closely correlated with the inequality of annual earnings (Moffit and Zhang, 2020). These

studies attest to the substantial contribution of short-term volatility to inequality in annual earnings, though without explicitly quantifying its contribution.

The effect of changes in the age distribution of the population, and in the joint distribution of age with permanent earnings and transitory shocks on inequality in annual earnings has received much less attention. Observing labor force participants at different stages of their working lives necessarily induces substantial dispersion in annual earnings even where there is no fundamental inequality, and all individuals share the same life-cycle earnings trajectory. It follows that changes in the age distribution of the earning population due to demographic factors or changes in participation rates, or changes in the joint distribution of life-cycle earnings trajectories and age, can affect the distribution of annual earnings. While these effects are generally recognized and often in practice, usually by regressing earnings on a polynomial in age before analyzing their residuals (e.g., Gottschalk and Moffitt, 2009; Moffit and Zhang, 2020), these adjustments cannot quantify their full effect on the dispersion of annual earnings or identify the different channels through which they work. Among these channels are an accelerated absorption of higher skilled younger cohorts in the workforce temporarily compressing the distribution of annual earnings (illustrated by the first numerical example in Appendix B); cyclical variation in their rate of absorption (the second example); and an increase in the regressive incidence of transitory shocks (Bourguignon, 2010; Fields, 2010), as lower-paid, younger workers experience relatively larger negative earnings shocks during recessions.

This paper offers a comprehensive decomposition of year-by-year inequality in the annual earnings of male heads of household that quantifies these various effects.[1] We begin by

---

[1] This approach can be applied similarly to analyze inequality in their annual household income, and in women's earnings or income, each no less interesting in its own light.

estimating predicted life-cycle earnings trajectories and lifetime earnings for male heads of household born between 1927 and 1987 from data on their annual earnings, age, education and race, drawn from the United States Panel Study of Income Dynamics (PSID). We then use these estimates to represent log annual earnings as the sum of three terms: the log of average lifetime earnings; the log difference between predicted annual (age-specific) earnings and average lifetime earnings; and the log difference between actual and predicted annual earnings. The log variance of annual earnings, a measure of (relative) inequality in annual earnings, can then be represented as the sum of the variances of these three terms, measuring, respectively, inequality in estimated lifetime earnings, the life-cycle dispersion of earnings, and volatility, and twice their three pairwise covariances, which measure their interactions.[2] We then track annual inequality and its six components over time, from 1986 to 2018, and quantify the relative contribution of each factor to the change in annual inequality. This is the first such systematic decomposition of which we are aware.

The challenge of applying this approach to recent years is that it requires reliable estimates of life-cycle earnings trajectories and lifetime earnings for all individuals in a wide range of ages with positive earnings in a year—the same population for which annual equality is measured—many of whom are still far from retirement. Previous studies that estimated inequality in lifetime earnings did so historically, within cohorts or cohort groups near or past retirement, from long panels of individual earnings histories drawn from administrative data. These include Bjorklund's (1993) estimate of lifetime earnings inequality among Swedish cohorts born in 1924-36; Aaberge and Mogstad's (2015) estimate of inequality in lifetime earnings among the 1942-1944 Norwegian birth cohorts; and in the United States, Guvenen,

---

[2] We focus on the log variance of earnings because it allows this decomposition, comparing it where possible to other metrics, and acknowledging its limitations (Foster and Ok, 1999).

Kaplan, Song and Weidner's (2017) historical tracking of lifetime earnings inequality within single-year birth-cohorts from 1932 to 1958, using Social Security data from 1957-2013. This full-data approach cannot be applied to estimating inequality in lifetime earnings among active labor-force participants in a single calendar year, as it requires a much greater span of data than estimating lifetime inequality within a single-year birth cohort, and in any case could only be applied historically.[3]

The approach we employ here combines the incomplete earnings histories of male heads of households born between 1927 and 1987 with demographic and education variables to regress annual earnings on a cubic function of age and its interaction with education, race, and cohort-group, and on individual random effects.[4] This yields predicted annual earnings trajectories for each individual in our full sample, including younger cohorts far from retirement, from which we estimate their individual average lifetime earnings. To demonstrate the reliability of these projections we show that our estimates of average lifetime earnings closely match actual earnings averages for earlier cohorts for whom fuller earnings histories are available. and that our residual measure of volatility corresponds closely to short-panel measures. In addition, we show and that our findings are robust to the addition of future data, by estimating lifetime earnings in 1986-2008 using only data to 2008 and showing the close match between these restricted-data estimates and our full-data estimates for these years.

---

[3] For example, estimating lifetime earnings inequality among active individuals aged, say, 31-59 in a single year from complete earnings histories requires 29 earlier years of data, for the oldest cohort, and 29 later years for the youngest cohort, 59 years of earnings data in all.

[4] This follows the method developed in Justman and Krush (2013) and Justman and Stiassnie (2021). Coronado, Fullerton, and Glass (2000), Blomquist (1981), and Bowlus and Robin (2004) use similar methods to construct synthetic or simulated lifetime earnings profiles.

Previewing our results, we show that year-by-year inequality in estimated lifetime earnings among males aged 31-59 with positive earnings in the year was much stabler and its average level much lower than inequality in their annual earnings in each year. Thus, between 1986 and 2018, the log variance of lifetime earnings increased by 7%, from 0.202 to 0.216, while the log variance of annual earnings rose by 20%, from 0.493 to 0.592; increased inequality in lifetime earnings contributed only 14% of this 0.099 rise in the inequality of annual earnings. Other measures of inequality applied to annual and estimated lifetime earnings—the Gini coefficient and two Theil entropy indices—present a similar picture. Increased volatility contributed 69% of this increase in annual inequality, and most of the remainder was the net result of a large rise in the regressive incidence of earnings shocks partly offset by the rising share of younger cohorts with higher lifetime earnings compressing the earnings distribution. In the more recent doubling of the log variance of annual earnings, from a trough of 0.40 in 1998 to a peak of 0.82 in 2012, volatility again played a dominant role, accounting for 62% of this increase, while lifetime inequality accounted for only 6%, and the remainder was mostly due to an increase in the regressive incidence of earnings shocks. Our finding that increased volatility has played a dominant role in driving the rise in the inequality of annual earnings accords closely with previous studies; and our residual, long-panel measure of earnings volatility is closely correlated with earlier, short-panel volatility measures (Moffitt and Zhang, 2020). Our findings go beyond previous studies in explicitly estimating the changes in lifetime earnings inequality and in the joint distribution of age and earnings by calendar year, and quantifying the respective contributions of lifetime inequality, life-cycle variation, volatility, and their interactions to inequality in annual earnings.

In addition, we estimate lifetime earnings inequality within forty-one rolling ten-year cohort groups born between 1938 and 1987, which we compare to our calendar-year measures and to Guvenen et al.'s (2017) cohort-based estimates. Our calendar-year and cohort-group

time-series both indicate a small overall rise in lifetime inequality over their full timespans (which are not directly comparable), and both have similar average values. However, the cohort-group time series shows greater variation. The log variance of lifetime earnings within cohort groups increased by 50% over the first fifteen cohort-groups and then declined by almost as much over the next twenty-five. Comparing our cohort-specific estimates of inequality in lifetime earnings to those of Guvenen et al. (2017), where they overlap, and making necessary adjustments, we find that our results are generally consistent with theirs.

The structure of the paper is as follows. Section 2 estimates predicted annual earnings and average lifetime earnings. Section 3 decomposes the log variance of annual earnings as the sum of the log variance of lifetime earnings, transitory shocks, and life-cycle variation, and their covariances, year by year. Section 4 applies short-panel measures of volatility to our data and compares them to our long panel measure of residual volatility. Section 5 estimates inequality in lifetime earnings within cohort groups and compares it to other estimates. Section 6 demonstrates the dynamic consistency of our findings, and Section 7 concludes.

2 Estimating earnings trajectories and average lifetime earnings

We begin by estimating the following Mincer-type earnings equation:

$$Y_{it} = \Sigma_{c=1}^{2}\{\alpha_{0c}cogr_{ic} + \Sigma_{k=1}^{3}\alpha_{1kc}age_{it}^{k}cogr_{ic} + \Sigma_{h=1}^{6}\Sigma_{k=1}^{3}\alpha_{2hkc}educ_{ih}age_{it}^{k}cogr_{ic}$$

$$+ \Sigma_{l=1}^{3}\Sigma_{k=1}^{3}\alpha_{3lkc}race_{il}age_{it}^{k}cogr_{ic}\} + D_i + \varepsilon_{it} \qquad (1)$$

which regresses Y_{ict}, the earnings of individual i belonging to cohort-group c in year t, on cohort-group constants $\{cogr_{ic}\}$; on a cubic function of age, $age_{it}^{k}, k = 1,2,3$, and its interaction with categorical variables indicating years of schooling $\{educ_{ih}\}$ and race $\{race_{il}\}$, and with cohort-group identifiers, which allows regression coefficients to vary by cohort-

group; and on individual random effects, D_i.[5] The α's are the regression coefficients and ε_{it} is an i.i.d. error term.[6] Estimation of (1) yields a predicted shape of the age-earnings profile for each combination of schooling, race, and birth-cohort group. Adding individual random effects we obtain predicted individual earnings in each year, $\hat{Y}_{it}$, and we take predicted earnings at age 40, $\hat{Y}_{i40}$, as our proxy for average lifetime earnings. We cannot use actual averages as our sample includes individuals observed only to mid-life. Using predicted earnings at age 45 yields very similar results. We denote: $y_{it} = \log(Y_{it}), \hat{y}_{it} = \log(\hat{Y}_{it}),$ and $\hat{Y}_i = \log(\hat{Y}_{i40})$.

Our data source is the PSID, using earnings data from 1969 to 2018, with data collected annually until 1996 and bi-annually thereafter. We follow previous research in restricting our analysis to male heads of households from the representative national sample drawn from the Survey Research Center (SRC), and do not include 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; and following Solon and Shin (2011) and subsequent studies, we trim the top and bottom one percent of positive earnings observations in each year. As we show in the following sections, using instead a linked dollar threshold to trim low earnings values dampens the rise in annual inequality and volatility during economic downturns but does not substantively affect our results. Earnings are trimmed

---

[5] Education categories are 8 or less years of schooling, 9-10, 11-12, 13-15, 16, 17 or more; race categories are white, Afro-American, Latino, other; cohort-groups are 1925-55, 1956-89.

[6] We make no assumptions on the structure of the error term, which is revealed in the decomposition of annual earnings inequality in Section 3.

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

at the top because the PSID changed its top-coding values over time; our findings are not sensitive to the top-trimming method used.

Our base sample comprises 3,371 male heads of households born between 1925 and 1989 (we use only the 1927-87 cohorts in analyzing inequality), with reported race and years of schooling, and with at least three untrimmed earnings observations between the ages of 25 and 59, yielding 45,939 annual earnings observations altogether, 13.6 per person. The average age at which earnings are observed is 39; average years of schooling rises from 13.9 in 1986 to 14.6 in 2018; about 90% percent of the sample are white and 8% are African American.

Equation (1) with individual random effects accounts for 0.679 of the overall variance in earnings observations. The two panels of Appendix Figure A1 show two age-earnings profiles derived from this estimation for each of the two cohort groups (the levels of the curves are determined by averaging individual random effects; coefficient estimates are available on request.) They highlight the effect of education on the shape of the age-earnings profile and the change in the shape of the curves over time. For both cohort groups, peak earnings of college graduates, at age 50, are about twice the level of their earnings at age 30, while individuals with 11-12 years of schooling experience less life-cycle variation.

To gauge how well $\hat{Y}_{i40}$, our lifetime earnings proxy, approximates actual lifetime earnings, we compare it in Figure 1 to actual average earnings for a subset of individuals in our sample with better data coverage: 958 males born between 1937 and 1957, and thus observed throughout their prime earning years, each with at least ten annual untrimmed earnings observations, an average of 23 per person. It shows a close correspondence between our estimates of anticipated average lifetime earnings and actual lifetime averages. In Section 6 we show that estimates using data to 2008 are robust to the addition of later data.

FIGURE 1 HERE

8

3 A decomposition of year-by-year inequality in annual earnings

Our primary measure of inequality in annual earnings by calendar year is the variance of log annual earnings, $V(y_{it})$, among male heads of households aged 31-59 with positive earnings in the year, after trimming the top and bottom one percent. We track it annually from 1986 to 1996 and then biannually to 2018, with an average of 1,046 positive earnings observations per year (Appendix Table A1 presents summary statistics; we have markedly fewer annual observations before 1986).

Shown as the topmost graph in Figure 2, the variance of log annual earnings fluctuates without trend until the turn of the millennium, reaching a trough of 0.402 in 1998, 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 value of 0.515 (numerical values in Appendix Table A2). Three alternative measures, the Gini coefficient and two Theil generalized entropy indices for $\alpha = 0$ and $\alpha = 1$ are shown graphically in Appendix Figure A2, and numerically in Appendix Table A3.[8] All follow a similar pattern, with a correlation of 0.95 between $V(y_{it})$ and the Theil index for $\alpha = 0$, and a correlation of 0.80 with both the Gini coefficient and the Theil index for $\alpha = 1$. Figure A2 also shows Moffitt and Zhang's (2020, Appendix Table 1) variance of log earnings series from1986 to 2016, similarly drawn from the PSID, which is nearly identical to ours.

FIGURE 2 HERE

---

[8] The two Theil indices are: $T_{\alpha=0} = (\frac{1}{N}) \sum_{i=1}^{N} \ln (\bar{Y}/Y_i)$ and $T_{\alpha=1} = (\frac{1}{N}) \sum_{i=1}^{N} (Y_i/\bar{Y}) \ln (\bar{Y}/Y_i)$

where $\bar{Y} = (\frac{1}{N}) \sum_{i=1}^{N} Y_i$ and Y_i denotes annual earnings.

Inequality in lifetime earnings, measured as the log variance of estimated lifetime earnings, $V(\hat{Y}_i)$, varies much less in this period, reaching its trough of 0.202 in 1986 and its peak of 0.240 in 2010, with a coefficient of variation over time, from 1986 to 2018, less than 5%. Alternative measures of lifetime earnings inequality, the Gini coefficient and our two Theil entropy measures, shown graphically in Appendix Figure A3 and numerically in Appendix Table A4, are similarly stable, with coefficients of variation over time all under 8%, and correlations of 0.88 and higher with $V(\hat{Y}_i)$. The dispersion of life-cycle variation in earnings, measured as $V(\hat{y}_{it} - \hat{Y}_i)$, is smaller in magnitude, ranging between 0.017 and 0.049. Volatility, measured as the log variance of residual transitory shocks, $V(y_{it} - \hat{y}_{it})$, shows much greater variation, from a trough of 0.139 in 1987 to a peak of 0.416 in 2012, with a correlation of 0.98 with annual earnings inequality, $V(y_{it})$, indicating the dominant role of volatility in driving changes in the variance of log annual earnings.

To quantify the relative contribution of the different factors shaping inequality in annual earnings, we write $y_{it} = \hat{Y}_i + (y_{it} - \hat{y}_{it}) + (\hat{y}_{it} - \hat{Y}_i)$, and use this to decompose the log variance of annual earnings, our measure of annual inequality, as the sum of six terms:

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

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

These six terms, plotted in Figure 2 (with numerical values in Appendix Table A2) represent:

- Inequality in lifetime earnings, measured as their log variance $V(\hat{Y}_i)$.

- Volatility, measured as the variance of the log difference between actual and predicted annual earnings, $V(y_{it} - \hat{y}_{it})$.

- The variance of the log difference between predicted earnings and estimated lifetime earnings, $V(\hat{y}_{it} - \hat{Y}_i)$, measuring the dispersion of life-cycle variation in earnings.

- The covariance of transitory shocks with life-cycle variation, $Cov(y_{it} - \hat{y}_{it}, \hat{y}_{it} - \hat{Y}_i)$, which increases when older cohorts with more education experience more positive transitory shocks (as predicted earnings increase with age, and life-cycle variation in earnings increases with education.)

- The covariance of transitory shocks with estimated lifetime earnings, $Cov(y_{it} - \hat{y}_{it}, \hat{Y}_i)$ a measure of the regressive incidence of volatility.

- The covariance of life-cycle variation in earnings with estimated lifetime earnings, $Cov(\hat{y}_{it} - \hat{Y}_i, \hat{Y}_i)$, which varies inversely with the share of younger men with higher estimated lifetime earnings (illustrated by the first example in Appendix B).

TABLE 1 HERE

Table 1 presents, in column (1), the relative contributions of each of these six components to the increase of 0.099 in the log variance of annual earnings from 1986 to 2018. Column (2) presents their relative contribution to the much steeper trough-to-peak rise of 0.424, from 1998 to 2012; and column (3) shows the correlation of each component with inequality in annual earnings over time, from 1986 to 2018. We find that increased volatility accounted for 69% of the increase in annual earnings inequality between 1986 and 2018, with a correlation of 0.98 between the two, while the increase in lifetime inequality accounted for only 14% of this rise, with a correlation of 0.75. We attribute most of the remaining increase to the net effect of two strong countervailing trends. A large increase in the regressive incidence of transitory earnings shocks, measured as $Cov(y_{it} - \hat{y}_{it}, \hat{Y}_i)$, closely with annual earnings inequality, amplified the effect of volatility on inequality in annual earnings. This was partly offset by an increase in the share of younger cohorts with higher skills and higher estimated lifetime earnings leading to a decline in $Cov(\hat{y}_{it} - \hat{Y}_i, \hat{Y}_i)$, which compressed the distribution of annual earnings. (This latter effect is illustrated by the first numerical example in Appendix B). The

magnitude of change in both these covariances was substantially larger than the change in lifetime earnings inequality. Volatility had a similarly large contribution, 62%, to the much sharper rise in inequality from 1998 to 2012, as did the rise in the regressive incidence of earnings shocks, 34%, while the contribution of increased lifetime inequality was again comparatively small, 6%, though in absolute size, the increase in lifetime earnings inequality from 1998 to 2012, 0.027, was nearly twice its increase from 1986 to 2018, 0.014.

The three rightmost columns of Table 1 present the ratios of the average values of the six components of annual earnings inequality to average earnings inequality, over the entire period, 1986-2018, and over two sub-periods, 1986-1998 and 2000-2018. The ratio of lifetime earnings inequality to annual earnings inequality, 42% over the entire period, is greater in the earlier years, when it slightly exceeds volatility, and smaller in later years, highlighting the growing relative impact of earnings volatility on annual earnings inequality from another perspective. These ratios are all substantially smaller than the 60-65% ratio Bjorklund (1993) found in his analysis of Swedish tax data on 1924-36 birth cohorts, evidence of the greater earnings volatility in the US economy in this more recent period.

Extending the range of ages in which inequality is measured each year from 31-59 to 25-59 has virtually no effect on our measures of annual and lifetime inequality, and only slightly reduces the level of residual volatility, $V(y_{it} - \hat{y}_{it})$, while retaining its shape, as we show in Appendix Figure A4. Replacing our 1% threshold for trimming low earnings with a dollar threshold equal to thirteen weeks of full-time work at half the minimum wage (the threshold used by Guvenen, Ozkan and Song, 2014) dampens the recent increase in annual earnings inequality and volatility as it excludes a larger fraction of observations during economic downturns (Moffitt and Zhang, 2020), but leaves our substantive results unchanged. This is shown in Figure 3, which reproduces the decomposition in Figure 2 for earnings trimmed by this dollar threshold. The correlation of annual earnings inequality, $V(y_{it})$, with

volatility, $V(y_{it} - \hat{y}_{it})$, and with the regressive incidence of transitory shocks, $Cov(y_{it} - \hat{y}_{it}, \hat{Y}_i)$, remains very high, 0.94 and 0.93 respectively. Lifetime inequality, $V(\hat{Y}_i)$, and the variance of lifetime variation in earnings, $V(\hat{y}_{it} - \hat{Y}_i)$, are unaffected by the change in threshold. Appendix Table A5 recalculates Table 1, the decomposition of changes in the log variance of annual earnings over different time periods using this same dollar threshold, linked to the minimum wage. It presents a similar picture to Table 1.

<div align="center">FIGURE 3 HERE</div>

4   Measures of volatility from short overlapping panels

The pattern described by our long-panel measure of residual volatility is consistent with previous short-panel estimates of volatility, among them analyses of PSID data by Shin and Solon (2011), Moffitt and Gottschalk (2012), and Moffitt and Zhang (2018, 2020); Carr and Wiemer's (2018) comparative analysis of mobility measures drawn from the PSID and from survey-linked administrative earnings data; and Moffitt's (2020) summary of a comparative analysis of mobility measures drawn from four data sources.[9]

This is illustrated in Figure 4, which compares our long-panel measure of residual volatility, $V(y_{it} - \hat{y}_{it})$, to two widely-used short-panel measures which we apply to our data: 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, $2(Y_{it} - Y_{i,t-2})/(Y_{it} + Y_{i,t-2})$. Moffitt and Zhang (2018) refer to these as measures of "gross volatility". The three measures are presented in Figure 4

---

[9] For further references, see Burkhouser and Couch (2009), Gottschalk and Moffitt (2009), Kopczuk, Saez and Song (2010), Jäntti and Jenkins (2013), and Moffitt and Zhang (2020).

along with the variance of log annual earnings, biannually from 1986 to 2018 (numerical values in Appendix Table A6). The correlations of these short-panel measures—the variance of the two-year arc difference and of the log difference of earnings—with our long-panel measure of residual volatility, $V(y_{it} - \hat{y}_{it})$, are 0.88 and 0.92, with our measure falling mostly between the two short-panel measures, despite their very different definitions. The correlations of the short-panel measures with the variance of log annual earnings are 0.89 and 0.94, only slightly lower than the corresponding correlation of 0.98 for our measure of volatility.

FIGURE 4 HERE

Expanding the range of ages within which inequality is measured each year from 31-59 to 25-59 has almost no effect on the two-year arc difference and log difference of earnings, as Appendix Figure A5 shows for the years 1986-2012, with correlations of 0.97 for each of the two measures between the two age ranges. Replacing our one percent bottom trim of earnings with a dollar threshold equal to thirteen weeks of full-time work at half the minimum wage dampens the increase of both volatility measures after the Great Recession, as the dollar threshold excludes an increasing fraction of low earnings when economic conditions worsen. This lowers the correlation of our residual measure of volatility with the two short-panel measures to 0.70 and 0.74 (Appendix Figure A6). [10]

---

[10] Residualizing the differences by first regressing them on a polynomial in age yields nearly identical results. Alternative approaches more difficult to relate to the current approach, include "window averaging", which separates "within" and "between" earnings variances in windows of five years or more; and structured error-components models that separate transitory from persistent shocks (Gottschalk and Moffitt, 1994, 2009).

5 Inequality in lifetime earnings within cohort-groups

In this section, we derive measures of inequality in estimated lifetime earnings within forty-one rolling ten-year cohort-groups, the oldest born in 1938-47, the youngest in 1978-87. They measure inequality within a much narrower reference group of individuals, closer in age to one another, compared to our calendar-year measures of lifetime inequality, where the reference group is most of the currently active labor force. We compare these cohort-group measures to our calendar-year measures and to Guvenen et al.'s (2017) estimates of inequality in lifetime earnings within single-year birth cohorts born between 1938 and 1958.[11] In addition, we compare estimates of lifetime earnings inequality in the 1957-64 cohort-group drawn from our PSID data to corresponding estimates we derive from the 1979 National Labor Survey of Youth (NLSY79).

Our cohort-group estimates are presented in Figure 5, which plots the log variance of estimated lifetime earnings, their Gini coefficient, and their Theil generalized entropy indices for $\alpha = 0$ and $\alpha = 1$, within these forty-one rolling ten-year cohort-groups. (Appendix Tables A7 and A8 present descriptive statistics and numerical values of the inequality measures). All four measures follow a similar concave pattern with pairwise correlations of 0.94 and higher between them. Their average values are similar to our calendar-year measures of lifetime earnings inequality, and both sets of measures indicate a small rise in lifetime inequality over their full time-spans. However, the cohort-group measures exhibit a very different dynamic pattern, first rising by almost 50 percent to their peak values in 1953-62, and then falling back gradually to just above their initial values for the youngest, 1978-87 cohort group, with coefficients of variation over time roughly twice as large as their calendar-year counterparts.

FIGURE 5 HERE

---

[11] Our sample is too small to measure inequality within single-year birth cohorts.

Comparing our estimates to those of Guvenen et al. (2017), for our twelve earliest ten-year cohort-groups, from 1938-47 to 1949-58, which overlap Guvenen et al.'s (2017) range of birth cohorts, 1938-58. The two sets of estimates are not directly comparable, as Guvenen et al. (2017) estimate inequality within single-year cohorts from extensive Social Security data on commerce and industry workers, while we estimate inequality within ten-year cohort-groups with more varied employment. Nonetheless, both their series and ours show rising trends of similar magnitude. Their Figure 11(a) shows an increase of about 20% in the standard deviation of log lifetime earnings for males, from the 1938 to the 1958 birth cohorts where we find a 17% increase in the standard deviation of our log lifetime earnings estimates between the 1938-47 and 1949-58 cohort groups.[12] To further compare our results to theirs, we average their single-cohort means within each of our ten-year 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, shown in Figure 6, is 0.90, with both showing a similar, moderately rising trend, and their values higher than ours by 18% on average.

FIGURE 6 HERE

We also compare measures of inequality in lifetime earnings within the 1957-64 cohort-group calculated from our PSID data to corresponding measures derived for a larger sample of males drawn from the NLSY79, applying the same method described here to the NLSY79 data. We first estimate equation (1) to obtain age-earnings profiles, which we find to be similar to our PSID-based profiles (examples shown in Appendix Figure A7), and then use predicted earnings at age 40 as our estimate of average lifetime earnings. Table 2 compares PSID and

---

[12] Relatedly, Haider (2001) using PSID data to 1991, found that inequality in 10-year earnings averages increased between the 1925-39 and 1938-52 cohort-groups.

NLSY79 estimates of four measures of inequality in lifetime earnings within the 1957-64 cohort group, and shows they are very similar. Appendix Figure A8 does the same for three measures of inequality in annual earnings within this cohort group, from 1988 to 2018, and finds both the levels and trends similar.

6 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 for younger cohorts from short earnings histories. As evidence of this, we estimates inequality in 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 this to our full-data estimates for this period, in two regards. First, Figure 7 directly compares our restricted-data estimates of lifetime earnings, using only data to 2008, to our full-data estimates, for 1,713 males born in 1938-77. There are a few outliers but the overall correlation is very high, 0.99. Then, Figure 8 compares our restricted-data estimates of year-by-year inequality in lifetime earnings from 1986 to 2008, measured as the log variance and as the Gini coefficient, to our full-data estimates for those years. The restricted-data estimates are slightly lower than the full-data estimates, with the differences increasing over time, but never exceeding 9 percent for the log variances, and 4 percent for the Gini coefficients. Correlations between the restricted-data and full-data estimates are 0.78 for the log variances and 0.92 for the Gini coefficients.

7 Conclusion

The dispersion of annual earnings, used universally to track socio-economic inequality over time and compare it across populations, reflects not only the underlying dispersion of lifetime earnings but also short-term volatility, much of which can be smoothed over time, and the joint distribution of age, or work experience, and occupations with different earnings trajectories in the labor force. This paper is concerned with separating the respective effect of these three factors on inequality in annual earnings.

To this end we first estimate the predicted life-cycle earnings trajectories and average lifetime earnings of male heads of household born between 1927 and 1987, from PSID data. We then use these estimates to present the inequality of annual earnings, measured as their log variance, as the sum of inequality in lifetime earnings, similarly measured; residual volatility, measured as the variance of the log difference between actual and predicted annual earnings; the dispersion of systematic life-cycle variation in earnings; and twice their pairwise covariances, year by year.

We then apply this framework to decompose inequality in annual earnings among male heads of households aged 31-59 with positive untrimmed earnings in each year, annually from 1986 to 1996 and then biannually to 2018. We find that while the log variance of annual earnings rose by 20% during this period, from 0.493 to 0.592, the log variance of estimated lifetime earnings was much smaller and behaved more stably, increasing by 7%, from 0.202 to 0.216, and thus accounting for only 14% of the increase the log variance of annual earnings. At the same time, volatility increased by 36% in this period, from 0.194 to 0.263, accounting for 69% of the rise in annual inequality, with which it is closely correlated over time, as previous studies have shown. We attribute the rest of the change in annual earnings inequality mostly to the net effect of two countervailing trends: a large increase in the regressive incidence

of transitory shocks—low earners are increasingly harder hit by recessions—partly offset by an increase in the share of young men in skilled occupations with higher lifetime earnings trajectories compressing the distribution of annual earnings.

In addition, we note that the stability of inequality in estimated lifetime earnings, which we find over the last three decades, is stability in *calendar-year* inequality among male heads of household actively participating in the workforce in each year. Inequality in lifetime earnings within cohort-groups, which we estimate for 41 rolling ten-year cohort-groups born between 1938 and 1987, follows a different dynamic pattern, rising initially by about 50% from the 1938-1947 to the 1953-1962 cohort groups, and then declining by almost as much to the 1978-1987 cohort-group. Thus, while both calendar-year and cohort-group measures of inequality in estimated lifetime earnings indicate a small rise in lifetime inequality overall, they behave very differently in the earlier and later periods.

References

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

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.

Bourguignon, F, 2010. Non-anonymous growth incidence curves, income mobility and social welfare dominance. *Journal of Economic Inequality* 9: 605-627.

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.

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.

Fields, G, 2010. Does income mobility equalize longer-term incomes? New measures of an old concept. *Journal of Economic Inequality* 8:409-427.

Foster, JE and E Ok, 1999. Lorenz dominance and the variance of logarithms. *Econometrica* 67 (4): 901-907.

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, 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.

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.

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.

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.

Tables and Figures

Table 1. *Decomposition of inequality in annual earnings, correlations with annual inequality, and average levels, 1986-2018 and sub-periods, 1% top and bottom trim*

	(1)	(2)	(3)	(4)	(5)	(6)
	1986-2018	1998-2012	1986-2018	1986-2018	1986-1998	2000-2018
	Change in $V(y_{it})$			*Average level of $V(y_{it})$*		
Inequality in annual earnings, $V(y_{it})$	0.099	0.424		0.515	0.452	0.591
	Contribution to change in $V(y_{it})$		*Correlation with $V(y_{it})$*	*Ratio of average level to the average level of $V(y_{it})$*		
Inequality in estimated lifetime earnings, $V(\hat{Y}_i)$	14%	6%	0.75	0.42	0.47	0.38
Residual volatility of earnings, $V(y_{it} - \hat{y}_{it})$	69%	62%	0.98	0.42	0.39	0.44
Dispersion of life-cycle variation, $V(\hat{y}_{it} - \hat{Y}_i)$	3%	0%	-0.01	0.05	0.07	0.04
Regressive incidence of shocks, $2 * Cov(y_{it} - \hat{y}_{it}, \hat{Y}_i)$	53%	34%	0.96	0.16	0.13	0.19
Covariance of volatility and life-cycle variation, $2 * Cov(y_{it} - \hat{y}_{it}, \hat{y}_{it} - \hat{Y}_i)$	-6%	1%	0.23	-0.04	-0.05	-0.02
Covariance of life-cycle variation and lifetime earnings, $2 * Cov(\hat{y}_{it} - \hat{Y}_i, \hat{Y}_i)$	-33%	-4%	-0.73	-0.02	0.00	-0.03

Notes: Male heads of households with at least three untrimmed earnings observations, aged 31-59 and with positive untrimmed earnings in the year, where trimming omits the top and bottom 1% of earnings in the year. Columns 1 and 2 quantify the decomposition of the log variance of annual earnings as the sum of variances and covariances, as specified in equation (2) for the full period, 1986-2018, and during the steep rise from 1998 to 2012; it follows from $y_{it} = \hat{Y}_i + (y_{it} - \hat{y}_{it}) + (\hat{y}_{it} - \hat{Y}_i)$, where y_{it} denotes log annual earnings and $\hat{y}_{it}$ and $\hat{Y}_i$ denote first-stage estimates of, respectively, predicted log annual earnings and log estimated average lifetime earnings. The top row shows the change in $V(y_{it})$ and the bottom six rows show the share of each component in this change; each column sums to 100%. Column 3 shows the correlation of each componenet with $V(y_{it})$ over the entire period. Columns 4, 5, and 6 show the ratio of average values over the full period and two sub-periods.

Table 2. *Measures of inequality in estimated lifetime earnings in the 1957-64 cohort-group calculated from NLSY79 data compared to our PSID estimates*

	N	Variance of log earnings	Gini coefficient	Mean to median ratio	Coefficient of variation
NLSY79	4,951	0.23	0.26	1.08	0.045
PSID	466	0.25	0.28	1.14	0.046

Figure 1. *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*

Number of individulas: 958
Correlation: .986, Rank correlation: .984
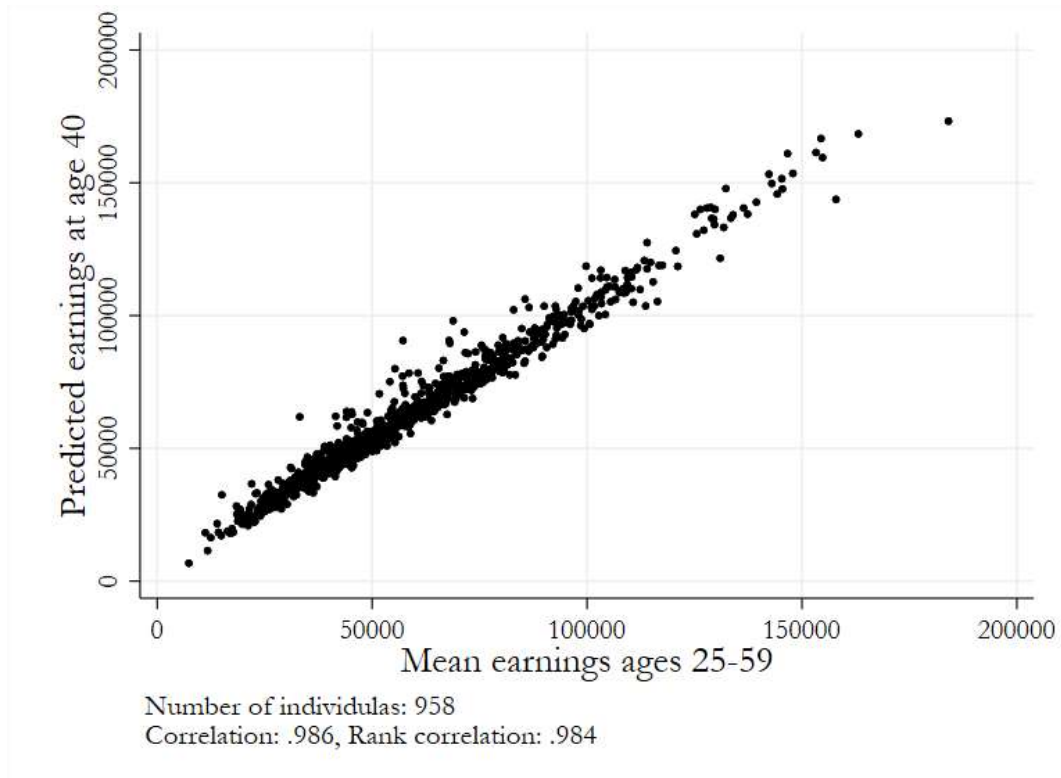
Note: Predicted earnings at age 40 is our estimate of average lifetime earnings, from a regression of annual earnings on a cubic polynomial in age and its interaction with education, race and cohort group indicators, and on individual random effects.

**Figure 2.** *Decomposition of the yearly variance of log annual earnings, 1986-2018: variances and pairwise covariances of log lifetime earnings, life-cycle variation, and residual volatility; 1% top and bottom trim*
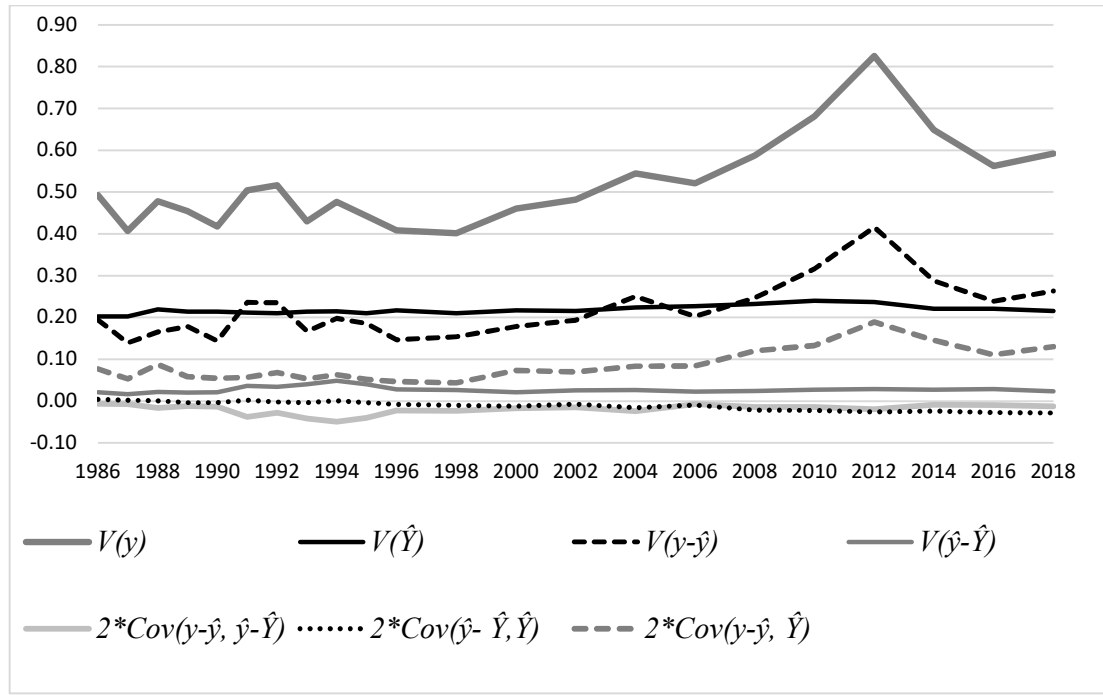
Note: Male heads of households with at least three untrimmed earnings observations after trimming the top and bottom 1% in each year; aged 31-59 and with positive untrimmed earnings in the year; annually to 1996 and biannually thereafter. The variable y denotes log annual earnings and $\hat{y}$ and $\hat{Y}$ denote first-stage estimates of, respectively, predicted log annual earnings and log estimated average lifetime earnings. The decomposition follows from $y_{it} = \hat{Y}_i + (y_{it} - \hat{y}_{it}) + (\hat{y}_{it} - \hat{Y}_i)$, as specified in equation (2).

Figure 3. *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*
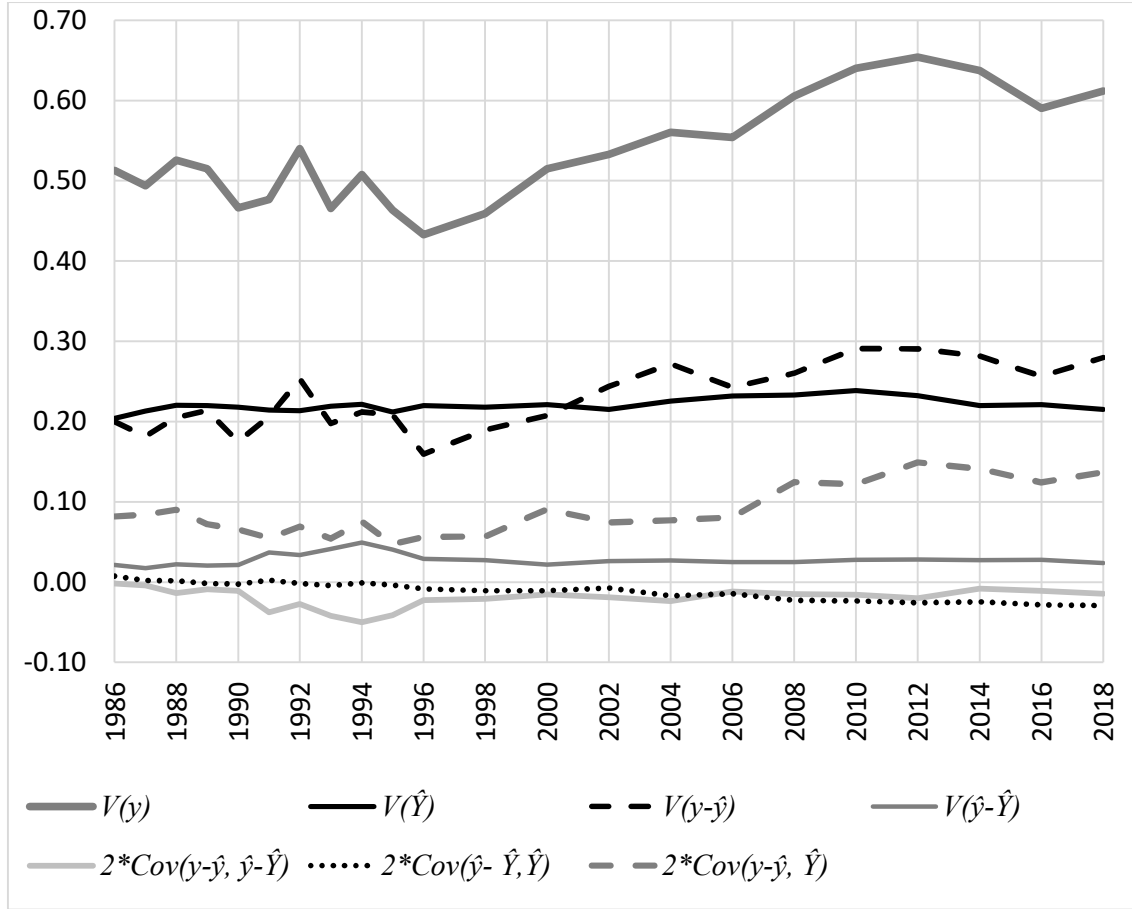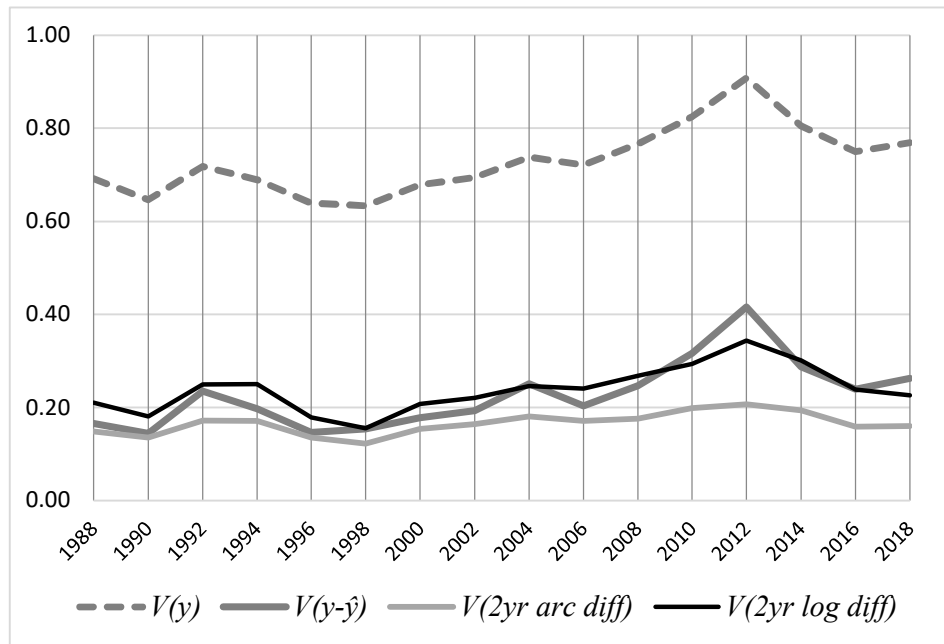

Note: Male heads of households with at least three untrimmed earnings observations after trimming earnings below a threshold equal to thirteen weeks of full-time work at half the minimum wage, and the top 1% of earnings in each year, aged 31-59 and with positive untrimmed earnings in the year; annually to 1996 and biannually thereafter. The variable y denotes log annual earnings and $\hat{y}$ and $\hat{Y}$ denote first-stage estimates of, respectively, predicted log annual earnings and log estimated average lifetime earnings. The decomposition follows from $y_{it} = \hat{Y}_i + (y_{it} - \hat{y}_{it}) + (\hat{y}_{it} - \hat{Y}_i)$, as specified in equation (2).

Figure 4. *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%. 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 $y_{it} - y_{i,t-2}$, where $y_{it} = \log Y_{it}$.

Figure 5. *Four measures of inequality in estimated lifetime earnings within 10-year cohort groups, 1938-87*
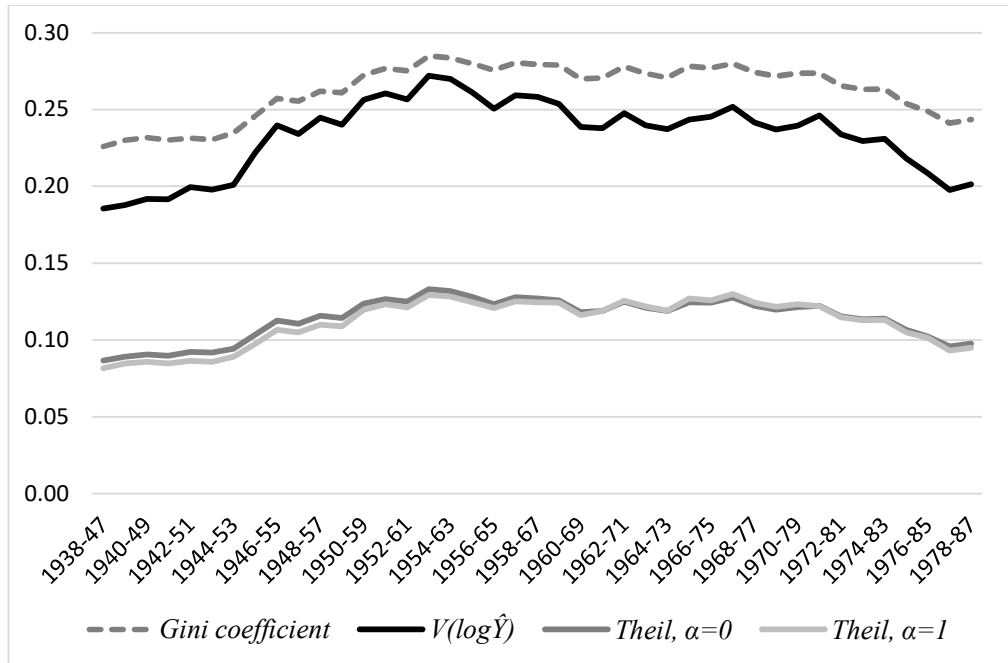
Note: Male heads of households with at least three untrimmed earnings observations after trimming the top and bottom 1% in each year. Average lifetime earnings estimated as predicted earnings at age 40, from a regression of annual earnings on a cubic polynomial in age and its interaction with education, race, and cohort group indicators, and on individual random effects. "Theil, $\alpha = 0$" and "Theil, $\alpha = 1$" are the Theil generalized entropy indices for $\alpha = 0$ and $\alpha = 1$ specified in footnote 8.

Figure 6. *Mean to median ratios of lifetime earnings within ten-year cohort groups, 1938-47 to 1949-58, current estimates compared to Guvenen et al. (2017)*
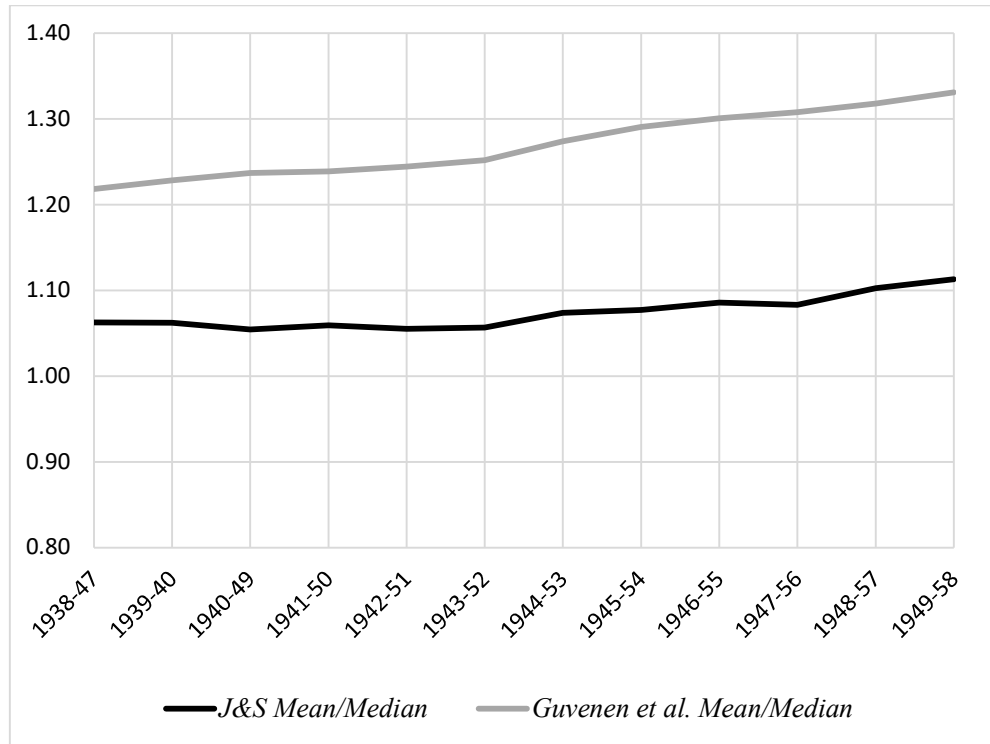
Note: "Guvenen et al. Mean/Median" denotes cohort-group mean to median ratios we constructed from Guvenen et al.'s (2017, Table A1) annual values by taking the within-group averages of their yearly means and the median of their yearly medians. They identify cohorts by the year they turn 25; we identify cohorts by their birth year. "J&S Mean/Median" denotes our mean to median ratios within each ten-year cohort group.

Figure 7. *Estimated lifetime earnings for males born in 1938-77 using only data up to 2008, plotted against our full data estimates (2018 PCE prices)*
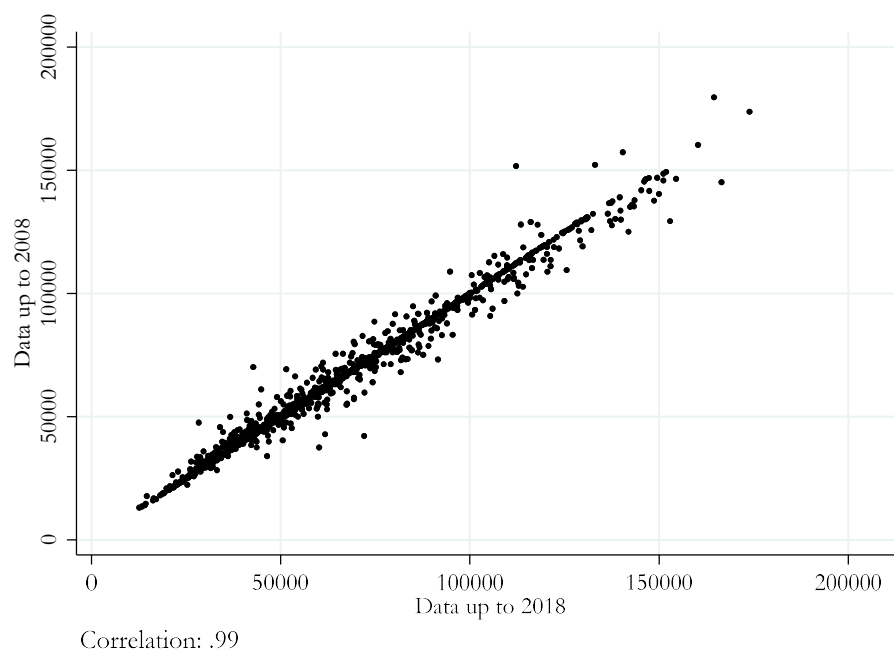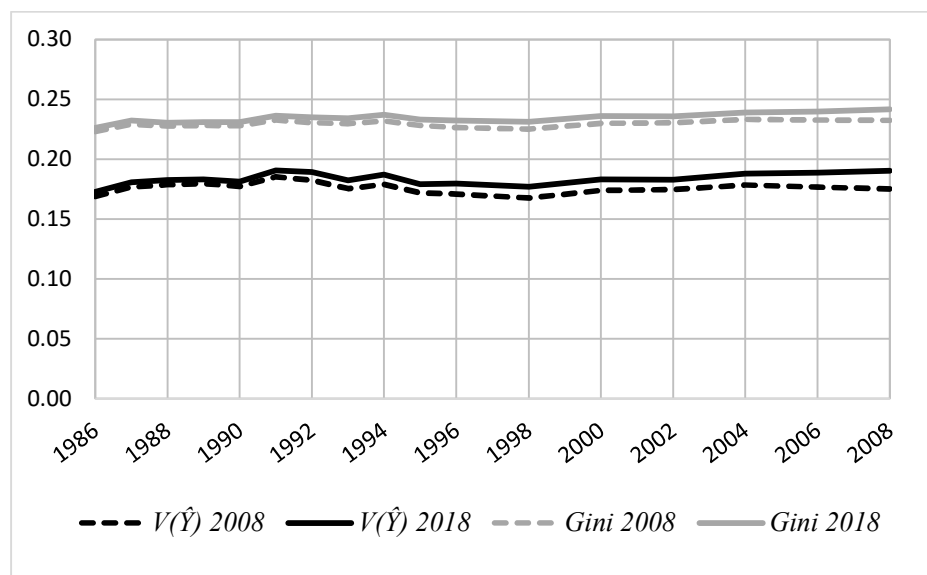

Correlation: .99

Figure 8. *Two measures of year-by-year inequality in estimated lifetime earnings, 1986-2008, using only data to 2008, and using all data to 2018*

Appendix A: Additional Tables and Figures

Table A1. *Year-by-year summary statistics, males aged 31-59 with positive untrimmed earnings in the year, top and bottom 1% trimmed annually; 2018 PCE prices.*

Year	N	Age group			Years of schooling			White	Mean earnings	Median earnings
		31-34	35-54	55-59	0-12	13-15	16+			
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	1,008	19%	72%	9%	37%	25%	38%	91%	59,881	52,902
1992	1,011	18%	74%	9%	37%	25%	38%	92%	62,602	56,359
1993	1,067	18%	73%	9%	36%	25%	39%	93%	63,364	54,988
1994	1,080	19%	73%	8%	36%	25%	38%	92%	63,162	53,867
1995	1,070	16%	76%	8%	36%	24%	40%	92%	63,824	54,259
1996	1,064	16%	76%	8%	34%	25%	41%	93%	65,465	56,078
1998	1,060	13%	77%	10%	34%	24%	42%	93%	70,549	60,440
2000	1,093	14%	75%	12%	34%	25%	42%	92%	73,210	60,874
2002	1,091	15%	72%	13%	33%	26%	41%	92%	71,044	57,610
2004	1,168	17%	69%	14%	30%	27%	42%	91%	73,119	59,581
2006	1,126	17%	68%	15%	30%	29%	41%	91%	76,376	60,690
2008	1,145	18%	66%	15%	31%	28%	41%	91%	72,088	58,613
2010	1,081	20%	64%	16%	30%	28%	42%	91%	70,246	58,810
2012	1,084	21%	64%	15%	29%	28%	43%	91%	70,925	58,449
2014	1,025	21%	65%	14%	29%	28%	43%	90%	71,702	57,893
2016	1,016	19%	67%	14%	28%	28%	44%	90%	74,530	62,373
2018	979	18%	69%	13%	25%	29%	45%	90%	76,463	63,000

Table A2. *The variance of log annual earnings and its decomposition, 1986-2018*

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

Note: Males aged 31-59 with positive untrimmed earnings in the year; top and bottom 1% trimmed each year; 2018 PCE prices. The decomposition follows from: y = Ŷ + (ŷ − Ŷ) + (y − ŷ), where y is log annual earnings, Ŷ is log estimated lifetime earnings and ŷ is predicted annual earnings, as specified in equation (2):

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

Table A3: *Measures of inequality in annual earnings, 1986-2018*

Year	N	V(y)	Gini coefficient	Theil entropy index $\alpha=0$	Theil entropy index $\alpha=1$
1986	967	0.493	0.308	0.194	0.157
1987	969	0.407	0.301	0.172	0.148
1988	976	0.479	0.315	0.196	0.164
1989	980	0.455	0.313	0.190	0.162
1990	979	0.418	0.317	0.185	0.167
1991	1,008	0.504	0.318	0.201	0.168
1992	1,011	0.516	0.318	0.205	0.168
1993	1,067	0.430	0.318	0.187	0.168
1994	1,080	0.476	0.328	0.203	0.180
1995	1,070	0.443	0.322	0.192	0.172
1996	1,064	0.408	0.322	0.184	0.169
1998	1,060	0.402	0.323	0.185	0.171
2000	1,093	0.461	0.345	0.212	0.199
2002	1,091	0.482	0.351	0.220	0.204
2004	1,168	0.545	0.367	0.245	0.226
2006	1,126	0.521	0.376	0.248	0.241
2008	1,145	0.588	0.367	0.253	0.226
2010	1,081	0.680	0.370	0.274	0.232
2012	1,084	0.825	0.381	0.306	0.246
2014	1,025	0.649	0.369	0.266	0.229
2016	1,016	0.562	0.360	0.242	0.216
2018	979	0.592	0.363	0.252	0.223

Note: Males aged 31-59 with positive untrimmed earnings in the year; top and bottom 1% trimmed each year; 2018 PCE prices. V(y) is the variance of log annual earnings. The Gini coefficient values are substantially smaller than published Gini coefficients in household income from the Current Population Survey, reflecing the greater homogeneity of our sample, which includes only male heads of household aged 31-59, and possibly due to attrition in the PSID affecting extreme values disproportionately. The two Theil entropy indices are specified in footnote 8.

Table A4: *Year-by-year inequality in estimated lifetime earnings, 1986-2018*

Year	N	V(Ŷ)	Gini coefficient	Theil entropy index $\alpha=0$	Theil entropy index $\alpha=1$
1986	967	0.20	0.24	0.09	0.09
1987	969	0.20	0.24	0.09	0.09
1988	976	0.22	0.24	0.10	0.09
1989	980	0.21	0.24	0.10	0.09
1990	979	0.21	0.25	0.10	0.10
1991	1,008	0.21	0.25	0.10	0.10
1992	1,011	0.21	0.24	0.10	0.10
1993	1,067	0.21	0.25	0.10	0.10
1994	1,080	0.22	0.25	0.10	0.10
1995	1,070	0.21	0.25	0.10	0.10
1996	1,064	0.22	0.25	0.10	0.10
1998	1,060	0.21	0.25	0.10	0.10
2000	1,093	0.22	0.26	0.11	0.10
2002	1,091	0.22	0.26	0.11	0.11
2004	1,168	0.22	0.26	0.11	0.11
2006	1,126	0.23	0.27	0.11	0.11
2008	1,145	0.23	0.27	0.12	0.11
2010	1,081	0.24	0.27	0.12	0.12
2012	1,084	0.24	0.26	0.12	0.11
2014	1,025	0.22	0.26	0.11	0.11
2016	1,016	0.22	0.26	0.11	0.10
2018	979	0.22	0.25	0.10	0.10

Note: Males aged 31-59 with positive untrimmed earnings in the year; 2018 PCE prices. Estimated lifetime earnings are predicted earnings at age 40 derived from our first stage regression. The two Theil entropy indices are specified in footnote 8.

Table A5. *Decomposition of inequality in annual earnings, correlations with annual inequality, and average levels, 1986-2018 and sub-periods, dollar threshold linked to minimum wage*

	(1)	(2)	(3)	(4)	(5)	(6)
	1986-2018	1998-2012	1986-2018	1986-2018	1986-1998	2000-2018
	Change in $V(y_{it})$			*Average level of* $V(y_{it})$		
Inequality in annual earnings, $V(y_{it})$	0.099	0.424		0.515	0.452	0.591
	Contribution to change in $V(y_{it})$		*Correlation with* $V(y_{it})$	*Ratio of average value to average* $V(y_{it})$		
Inequality in estimated lifetime earnings, $V(\hat{Y}_i)$	11%	7%	0.58	0.41	0.44	0.38
Residual volatility of earnings, $V(y_{it} - \hat{y}_{it})$	81%	52%	0.94	0.43	0.41	0.44
Dispersion of life-cycle variation, $V(\hat{y}_{it} - \hat{Y}_i)$	2%	0%	-0.21	0.05	0.06	0.04
Regressive incidence of shocks, $2 * Cov(y_{it} - \hat{y}_{it}, \hat{Y}_i)$	55%	47%	0.93	0.16	0.14	0.19
Covariance of volatility and life-cycle variation $2 * Cov(y_{it} - \hat{y}_{it}, \hat{y}_{it} - \hat{Y}_i)$	-13%	0%	0.35	-0.04	-0.05	-0.03
Covariance of life-cycle variation and lifetime earnings $2 * Cov(\hat{y}_{it} - \hat{Y}_i, \hat{Y}_i)$	-37%	-8%	-0.78	-0.02	0.00	-0.03

Notes: Male heads of households with at least three untrimmed earnings observations, aged 31-59 and with positive untrimmed earnings in the year, where trimming omits earnings below a threshold equal to thirteen weeks of full-time work at half the minimum wage, and the top 1% of earnings in the year. Columns 1 and 2 quantify the decomposition of the log variance of annual earnings as the sum of variances and covariances, as specified in equation (2) for the full period, 1986-2018, and during the steep rise from 1998 to 2012; it follows from $y_{it} = \hat{Y}_i + (y_{it} - \hat{y}_{it}) + (\hat{y}_{it} - \hat{Y}_i)$, where y_{it} denotes log annual earnings and $\hat{y}_{it}$ and $\hat{Y}_i$ denote first-stage estimates of, respectively, predicted log annual earnings and log estimated average lifetime earnings. The top row shows the change in $V(y_{it})$ and the bottom six rows show the share of each component in this change; each column sums to 100%. Column 3 shows the correlation of each componenet with $V(y_{it})$ over the entire period. Columns 4, 5, and 6 show the ratio of average values over the full period and two sub-periods.

Table A6. *Our long panel measure of volatility in annual earnings, and two overlapping 2-year short-panel measures, males aged 31-59 and with untrimmed earnings in the year, top and bottom 1% trimmed, 1986-2018, biannually*

Year	$V(y_{it})$	$V(y_{it} - \hat{y}_{it})$	$V(y_{it} - y_{i,t-2})$	$V\left(\dfrac{Y_{it} - Y_{i,t-2}}{(Y_{it} + Y_{i,t-2})/2}\right)$
1986	0.69	0.17	0.21	0.15
1988	0.65	0.14	0.18	0.14
1990	0.72	0.23	0.25	0.17
1992	0.69	0.20	0.25	0.17
1994	0.64	0.15	0.18	0.14
1996	0.63	0.15	0.16	0.12
1998	0.68	0.18	0.21	0.15
2000	0.69	0.19	0.22	0.16
2002	0.74	0.25	0.25	0.18
2004	0.72	0.20	0.24	0.17
2006	0.77	0.25	0.27	0.18
2008	0.82	0.32	0.29	0.20
2010	0.91	0.42	0.34	0.21
2012	0.81	0.29	0.30	0.19
2014	0.75	0.24	0.24	0.16
2016	0.77	0.26	0.23	0.16
2018	0.69	0.17	0.21	0.15

Note: Y_{it} denotes annual earnings of individual i in year t, y_{it} is log annual earnings, and $\hat{y}_{it}$ is log predicted annual earnings. $V(y_{it} - \hat{y}_{it})$ is the variance of the log difference between actual and predicted earnings, our residual measure of volatility. $V(y_{it} - y_{i,t-2})$ is the variance of the two-year log difference in earnings, and $V\left(\frac{Y_{it} - Y_{i,t-2}}{(Y_{it} + Y_{i,t-2})/2}\right)$ is the variance of the arc-difference in earnings, two short-panel measures of volatility.

Table A7. *Descriptive statistics for ten-year cohort groups, 1938-87;*
the four disjoint cohort-groups are shaded

Birth cohort	N	Observations	Average observed annual earnings	Average estimated average lifetime earnings	Years of schooling
1938-47	494	4,025	67,650	60,341	13.8
1939-40	523	4,569	68,803	60,824	14.0
1940-49	546	4,918	68,460	60,998	14.1
1941-50	567	5,325	69,505	62,026	14.1
1942-51	588	5,691	68,592	61,571	14.2
1943-52	606	6,046	68,343	61,604	14.2
1944-53	634	6,649	68,051	61,711	14.2
1945-54	650	7,131	67,422	61,206	14.2
1946-55	671	7,597	67,881	61,600	14.2
1947-56	676	7,754	66,891	61,487	14.1
1948-57	661	7,787	66,515	61,249	14.0
1949-58	645	7,677	66,128	61,525	13.9
1950-59	654	7,848	66,728	62,231	13.9
1951-60	665	8,047	66,489	62,400	13.9
1952-61	649	7,896	67,546	63,937	13.9
1953-62	645	7,693	68,062	65,070	13.9
1954-63	633	7,268	67,352	65,078	13.9
1955-64	606	6,794	68,080	66,545	14.0
1956-65	578	6,170	67,688	67,124	13.9
1957-66	548	5,698	68,185	67,556	14.0
1958-67	538	5,294	69,093	69,127	14.0
1959-68	521	4,904	68,294	68,872	14.0
1960-69	513	4,595	68,119	69,553	14.0
1961-70	496	4,072	69,275	71,242	14.1
1962-71	498	3,841	68,180	70,892	14.1
1963-72	494	3,622	67,691	71,333	14.1
1964-73	490	3,418	69,422	72,530	14.2
1965-74	501	3,300	70,177	73,275	14.1
1966-75	505	3,212	69,221	73,040	14.2
1967-76	525	3,203	70,472	75,120	14.2
1968-77	537	3,113	70,507	75,256	14.2
1969-78	560	3,102	70,584	76,151	14.3
1970-79	576	3,004	70,547	76,630	14.4
1971-80	575	2,839	69,452	76,286	14.4
1972-81	593	2,734	69,707	77,360	14.5
1973-82	584	2,503	69,452	77,258	14.5
1974-83	599	2,352	68,664	77,666	14.5
1975-84	602	2,173	67,603	78,089	14.5
1976-85	597	1,980	68,268	78,714	14.6
1977-86	591	1,762	67,239	78,106	14.6
1978-87	576	1,503	66,407	78,439	14.7

Note: Earlier cohorts are observed mostly at later ages when annual earnings are mostly higher than average lifetime earnings, while the opposite holds for later cohorts.

Table A8. *Inequality in estimated lifetime earnings within rolling ten-year cohort groups, 1938-87; the five disjoint cohort-groups are shaded*

Birth cohorts	N	Var $(\log \hat{Y})$	Theil, $\alpha = 0$	Theil, $\alpha = 1$	Gini coefficient	Mean to median ratio
1938-47	494	0.186	0.09	0.08	0.23	1.06
1939-40	523	0.188	0.09	0.08	0.23	1.06
1940-49	546	0.192	0.09	0.09	0.23	1.05
1941-50	567	0.192	0.09	0.08	0.23	1.06
1942-51	588	0.199	0.09	0.09	0.23	1.05
1943-52	606	0.198	0.09	0.09	0.23	1.06
1944-53	634	0.201	0.09	0.09	0.24	1.07
1945-54	650	0.222	0.10	0.10	0.25	1.08
1946-55	671	0.240	0.11	0.11	0.26	1.09
1947-56	676	0.234	0.11	0.10	0.26	1.08
1948-57	661	0.245	0.12	0.11	0.26	1.10
1949-58	645	0.240	0.11	0.11	0.26	1.11
1950-59	654	0.256	0.12	0.12	0.27	1.14
1951-60	665	0.261	0.13	0.12	0.28	1.15
1952-61	649	0.257	0.12	0.12	0.28	1.13
1953-62	645	0.272	0.13	0.13	0.29	1.16
1954-63	633	0.270	0.13	0.13	0.28	1.14
1955-64	606	0.261	0.13	0.12	0.28	1.13
1956-65	578	0.251	0.12	0.12	0.28	1.13
1957-66	548	0.259	0.13	0.13	0.28	1.13
1958-67	538	0.258	0.13	0.12	0.28	1.15
1959-68	521	0.254	0.13	0.12	0.28	1.14
1960-69	513	0.239	0.12	0.12	0.27	1.12
1961-70	496	0.238	0.12	0.12	0.27	1.14
1962-71	498	0.248	0.12	0.13	0.28	1.14
1963-72	494	0.240	0.12	0.12	0.27	1.14
1964-73	490	0.237	0.12	0.12	0.27	1.14
1965-74	501	0.244	0.12	0.13	0.28	1.15
1966-75	505	0.245	0.12	0.13	0.28	1.15
1967-76	525	0.252	0.13	0.13	0.28	1.16
1968-77	537	0.242	0.12	0.12	0.27	1.15
1969-78	560	0.237	0.12	0.12	0.27	1.14
1970-79	576	0.240	0.12	0.12	0.27	1.14
1971-80	575	0.246	0.12	0.12	0.27	1.14
1972-81	593	0.234	0.12	0.11	0.27	1.14
1973-82	584	0.229	0.11	0.11	0.26	1.14
1974-83	599	0.231	0.11	0.11	0.26	1.13
1975-84	602	0.218	0.11	0.10	0.25	1.12
1976-85	597	0.209	0.10	0.10	0.25	1.11
1977-86	591	0.198	0.10	0.09	0.24	1.09
1978-87	576	0.201	0.10	0.09	0.24	1.09

Note: Var (log Ŷ) is the log variance of estimated lifetime earnings. The two Theil entropy indices are specified in footnote 8.

Figure A1. *PSID age-earnings profiles, by years of schooling, white males born in 1925-56 and 1957-87; derived from a regression of earnings on a cubic polynomial in age and its interaction with education, race, and cohort-group (equation 1), and individual random effects. Heights of the graphs are averages of the random effects for white males, by schooling and cohort.*
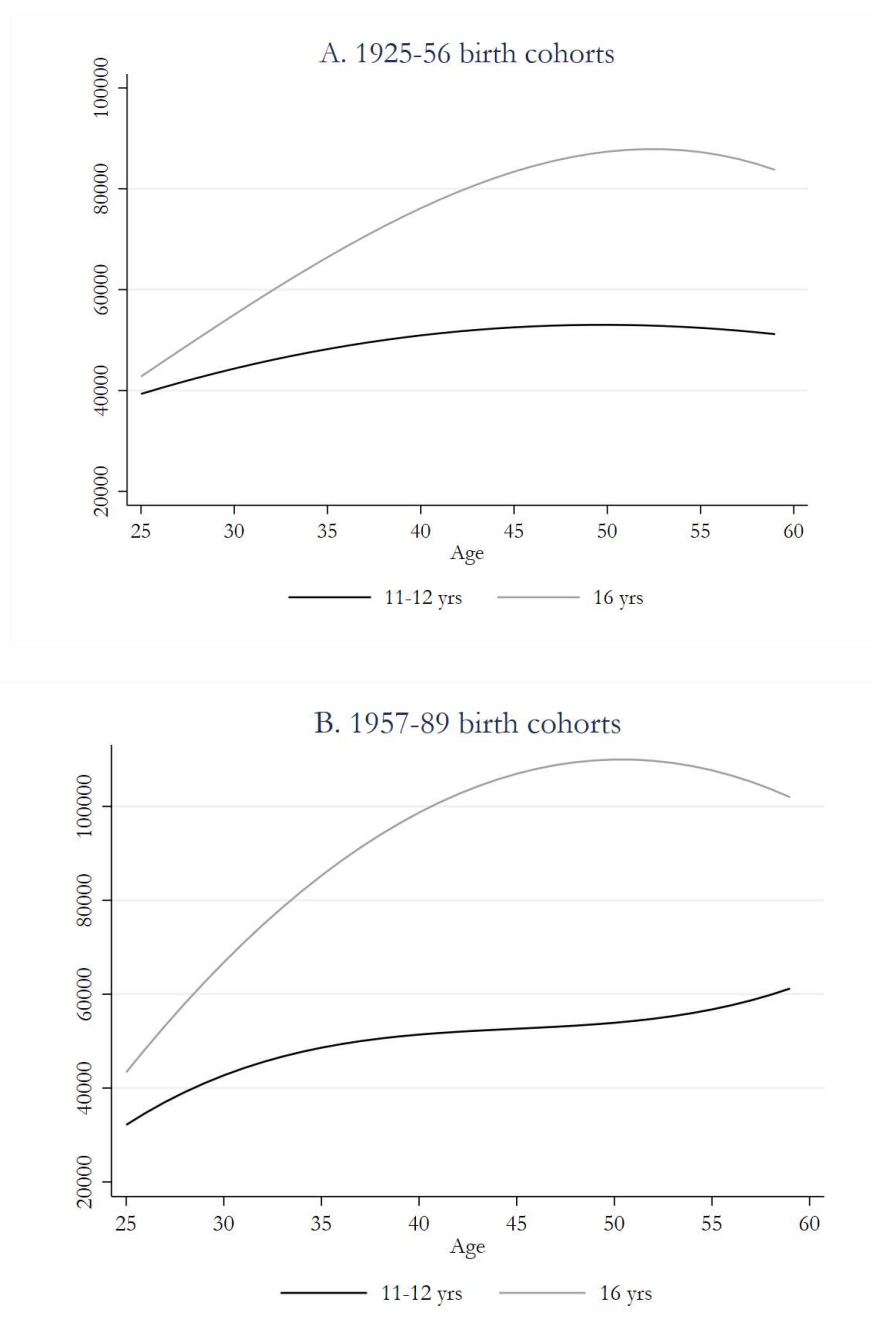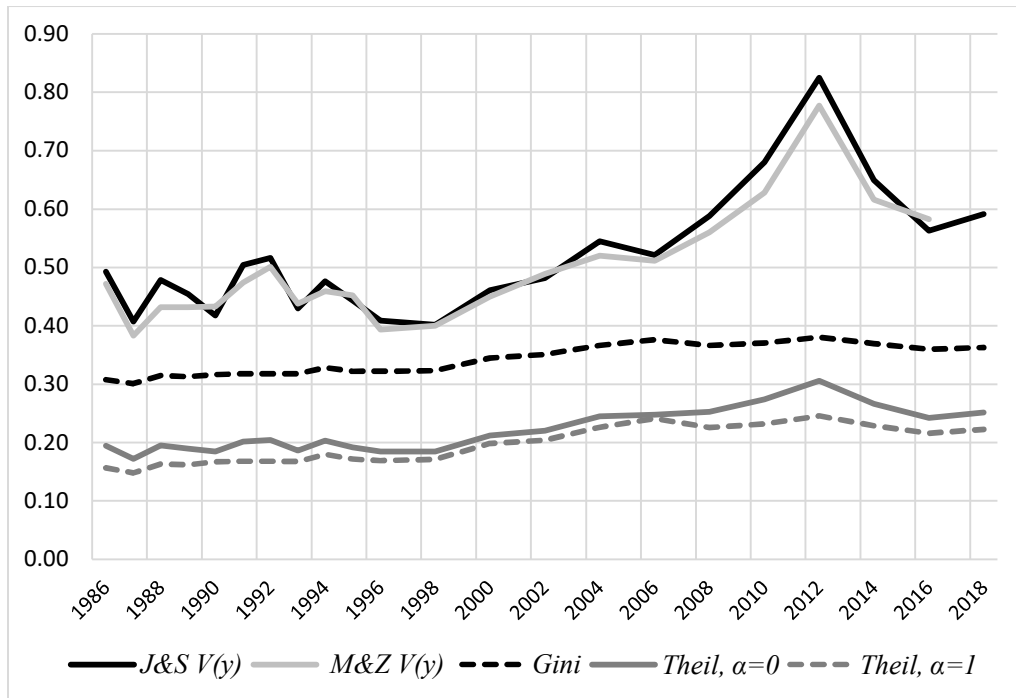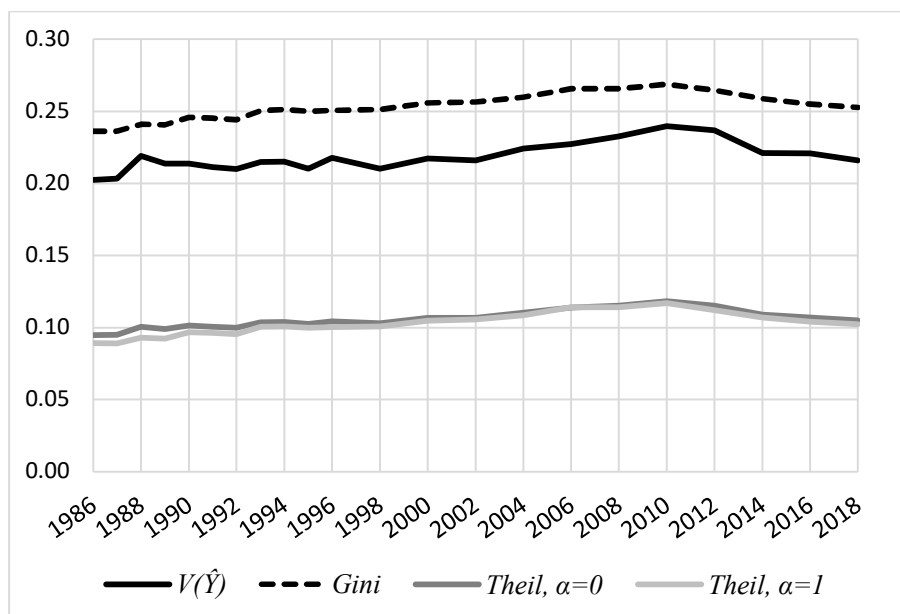
**Figure A2.** *Inequality in annual earnings, 1986-2018, annually to 1996, biannually thereafter*

Note: "J&S V(y)" is the log variance of annual earnings, as in Figure 2. "M&Z V(y)" is Moffitt and Zhang's (2020) PSID-based measure of the variance of log earnings (their Appendix Table 1). "Theil, a=0 and Theil, a=1" are the Theil generalized entropy measures defined in footnote 8.

Figure A3. *Inequality in estimated lifetime earnings, 1986-2018, annually to 1996, biannually thereafter*

Note: V(Ŷ) is the variance of log lifetime earnings. Theil, α=0 and Theil, α=1 are the Theil generalized entropy measures defined in footnote 8, for lifetime earnings.

41

Figure A4. *Inequality in annual and estimated lifetime earnings and residual volatility, males aged 31-59 and 25-59, with untrimmed earnings, 1986-2012, 1% trim, annually to 1996, biannually thereafter*
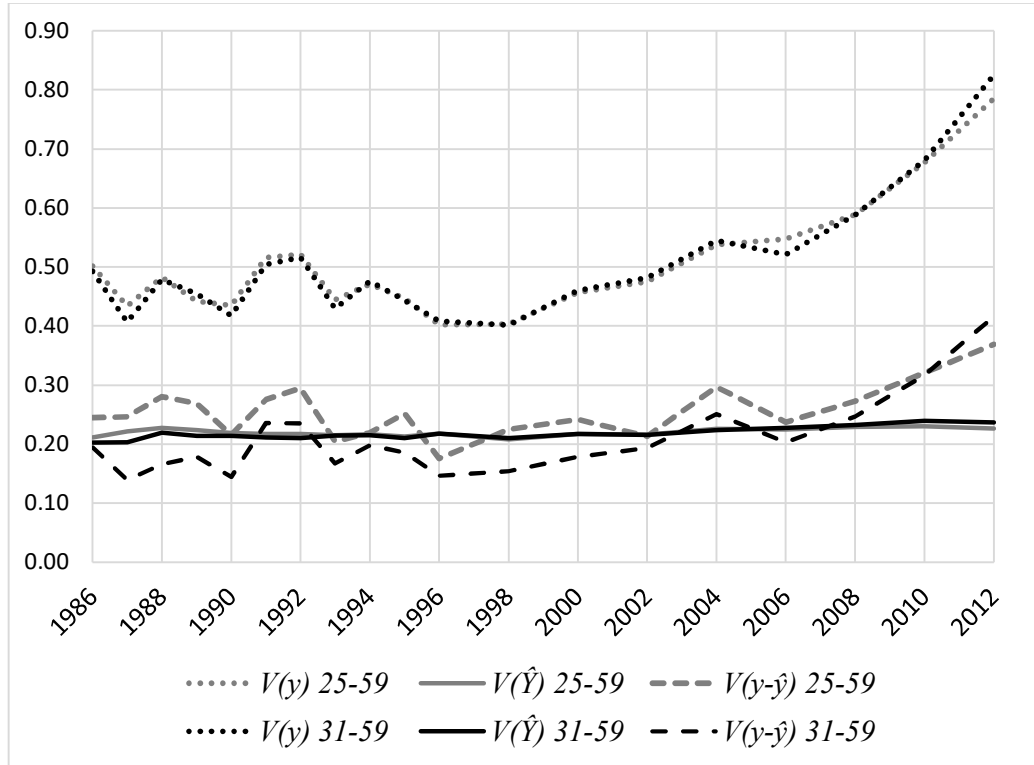
Note: V(y) is the log variance of annual earnings; V(Ŷ) is the log variance of estimated lifetime earnings; and V(y-ŷ) is the variance of the log difference between actual and predicted earnings, our measure of residual volatility.

Figure A5. *The two-year arc difference and log difference of earnings, two short-panel measures of volatility, for males aged 31-59 and 25-59, 1986-2012, biannually*
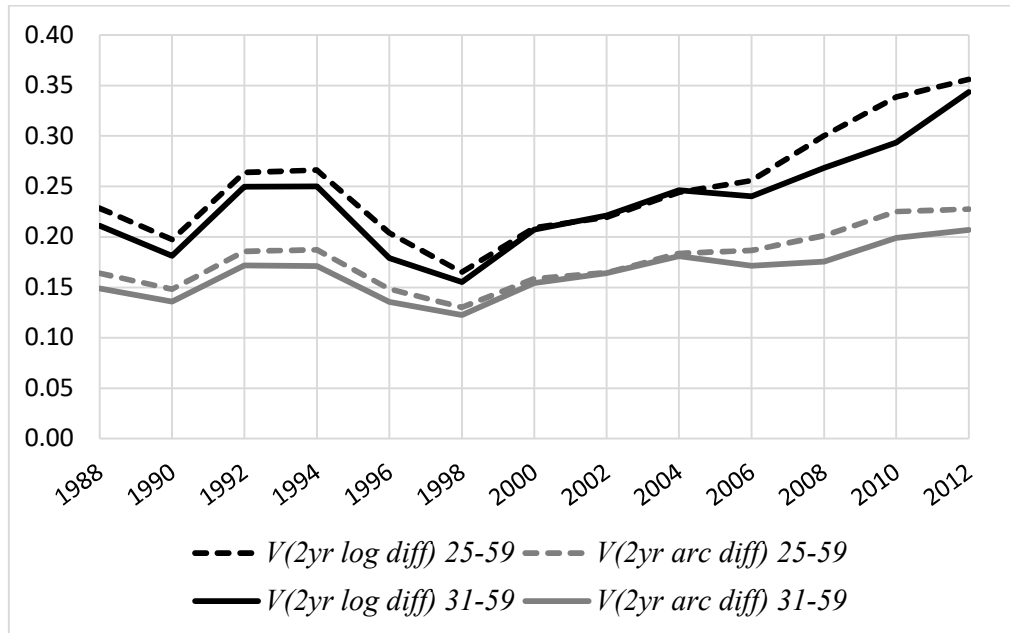

Figure A6. *Short-panel measures of volatility, and our residual measure, $V(y - \hat{y})$, males aged 31-59 top 1% trimmed with earnings in the year above a dollar threshold equal to 13 weeks of full-time work at half the minimum wage, 1986-2018, biannually*
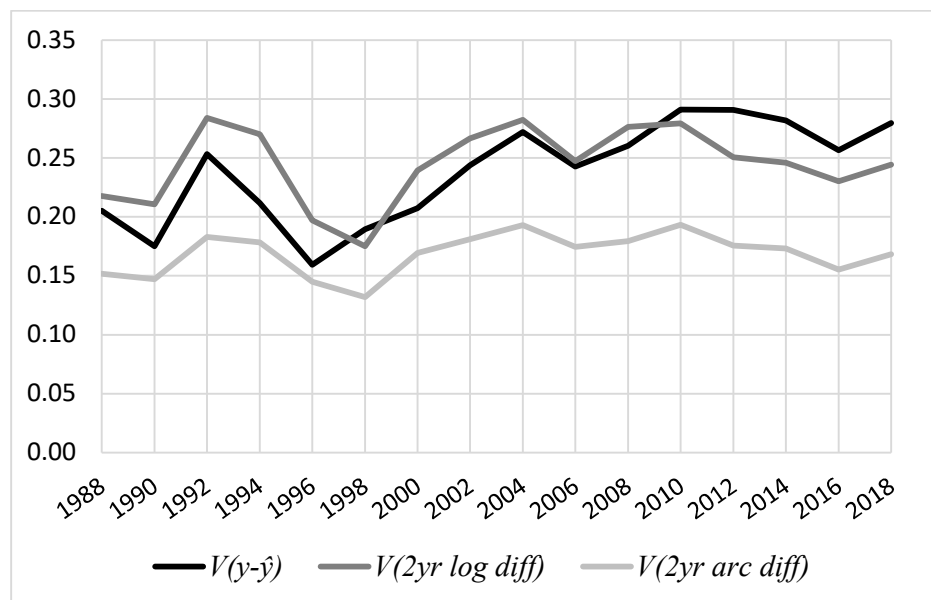
Figure A7. *PSID and NLSY age-earnings profiles, by years of schooling, white males born in 1957-64; derived from a regression of earnings on a cubic polynomial in age and its interaction with education and race (equation 1) and individual random effects. Heights of the graphs are averages of the random effects for white males, by schooling, in each data set.*
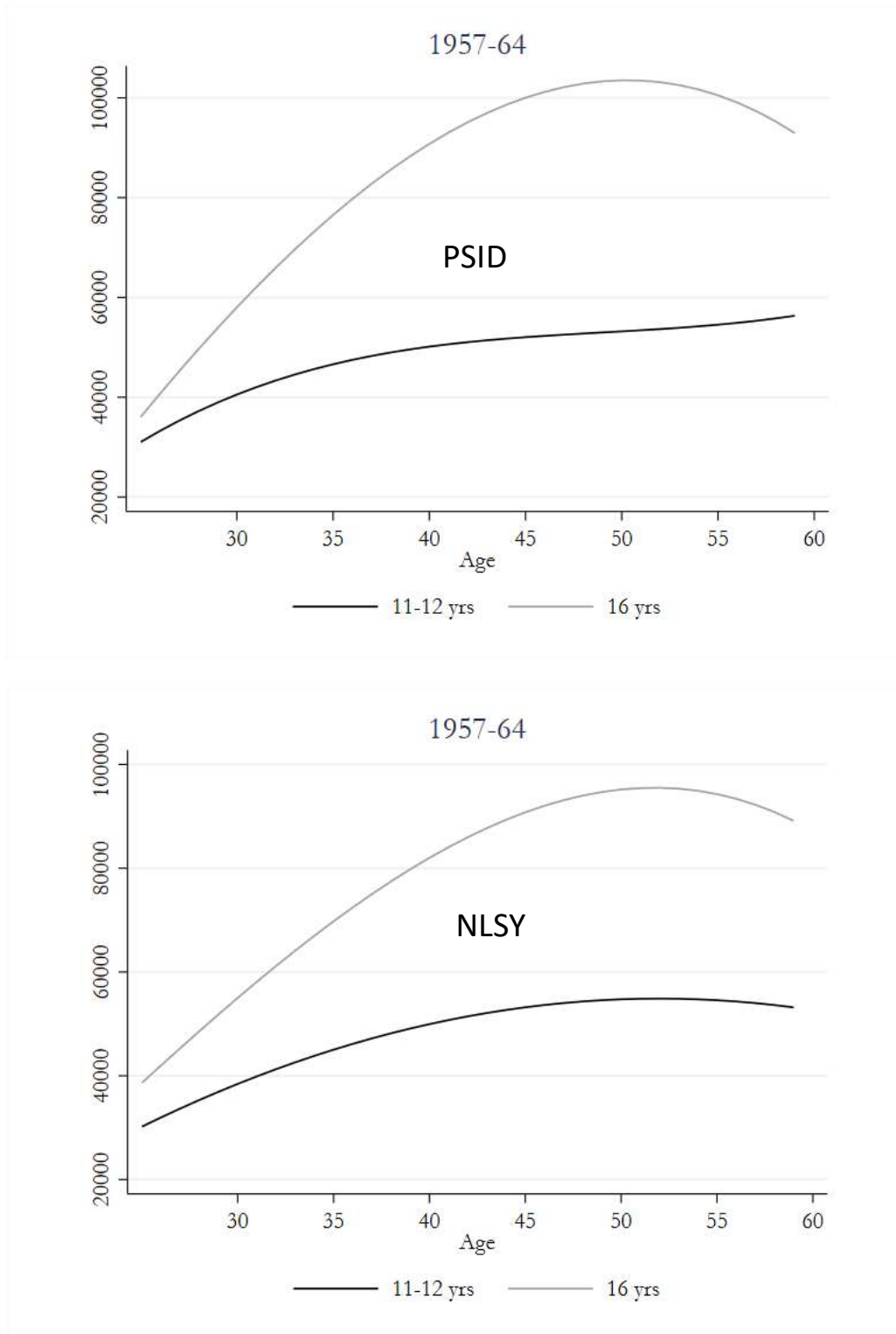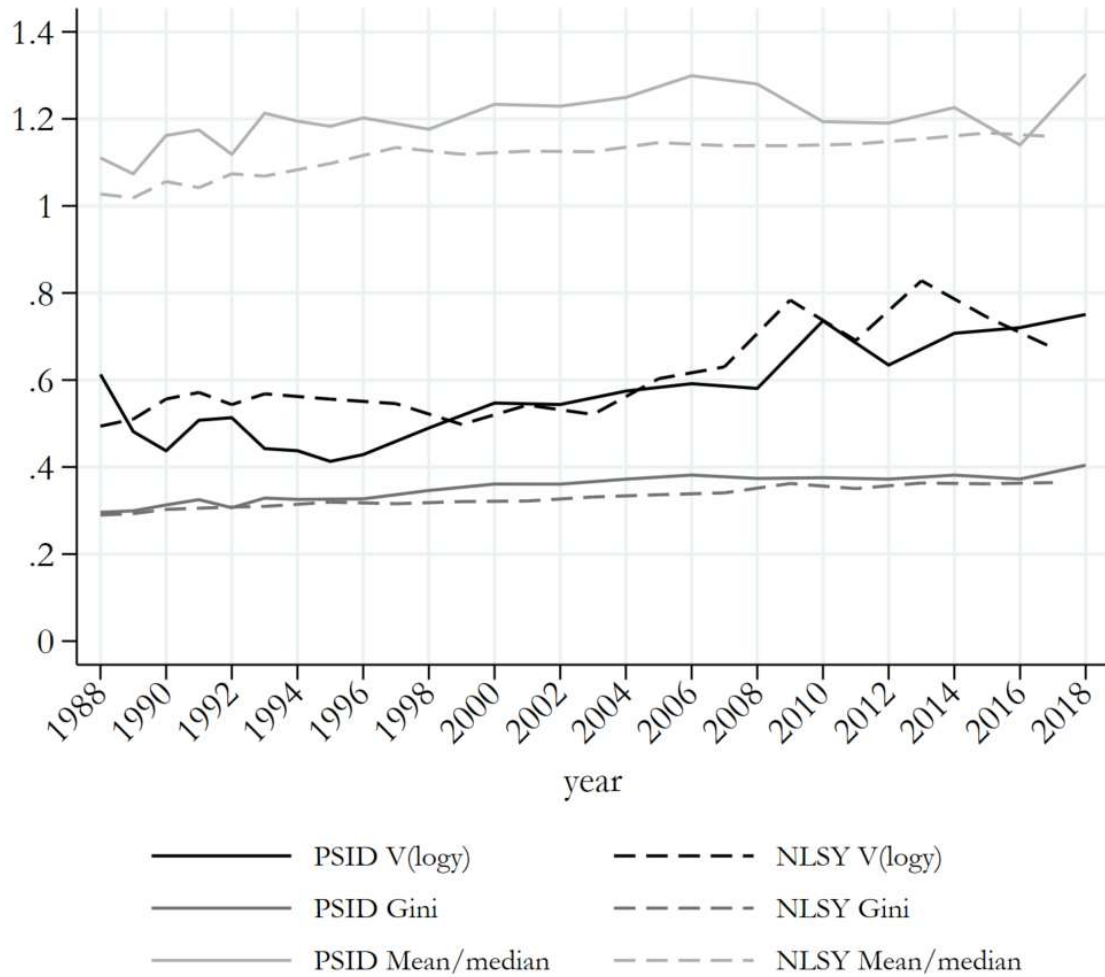
Figure A8. *PSID and NLSY inequality in annual earnings, males born in 1957-64, aged 31-59 and with untrimmed earnings (2018 prices) in each year*

Note: We use the 1978-2017 waves the of the NLSY79, collected annually to 1993 and biannually thereafter.

Appendix B: Two examples of how changes in the age distribution of lifetime earnings can affect inequality in annual earnings

The following two numerical examples illustrate how changes in the age distribution of lifetime income can affect inequality in annual earnings through its effect on the covariance of lifetime earnings with life-cycle variation in earnings. The first describes a gradually implemented structural change; the second, cyclical variation in the age distribution of occupations.

In both examples, we consider an overlapping generations economy with a fixed population of eight individuals, each living two periods, which we follow over time. Half the population in the first (and each successive) period is younger (Y) and half is older (O); and all are employed in one of two occupations, one that is lower-skilled, lower earnings (L) and the other higher-skilled, higher-earning (H). Each occupation has a fixed earnings trajectory, and we assume there are no transitory shocks, so actual earnings equal predicted earnings: $Y_{LY} = 20$, $Y_{HO} = 60$, $Y_{HY} = 40$, and $Y_{HO} = 120$; and average lifetime earnings are $Y_L = 40$ and $Y_H = 80$. Taking logarithms, we write: $y_{it} = \hat{Y}_i + (y_{it} - \hat{Y}_i)$, where t is a period index, $t = 1,2,3$, and use this to decompose the log variance of current earnings as the sum of three terms: $V(y_{it}) = V(\hat{Y}_i) + V(y_{it} - \hat{Y}_i) + 2Cov(y_{it} - \hat{Y}_i, \hat{Y}_i)$. In this numerical example, the variance of life-cycle change $V(y_{it} - \hat{Y}_i)$ is unaffected, so changes in inequality in annual earnings are driven by changes in lifetime inequality and in the covariance of lifetime earnings with life-cycle variation in earnings among the active population in each period.

First example. Initially, in the first period, half of each generation is low-skilled, and the other is high-skilled. Table B1 presents annual and average lifetime earnings by generation in each period. Changes in the economy then increase demand for high-skilled workers, which is achieved gradually. In the second period, the first-period older generation 1 exits the economy, the younger generation becomes older, and a new younger generation enters the economy with

a higher skill ratio, one is low-skilled and three are high-skilled; it is more skilled than the older generation. In the third period, the new younger generation again comprises one low-skilled and three high-skilled workers, so the two active generations again have the same occupational structure, completing the transition to a new, higher-skilled steady state.

Table B2 shows numerically, and Figure B1 describes graphically, how changes in age distribution of skills, which correspond to lifetime earnings, principally drive large changes in the log variance of annual earnings and its decomposition over the three periods. In the first period the two generations have the same skill structure, and the covariance of lifetime earnings with life-cycle variation in earnings is zero. The disproportionate entry of the higher-skilled young generation in period 2 temporarily reduces the covariance of lifetime earnings with life-cycle variation in earnings and compresses the earnings distribution, which is the main factor driving lower inequality in current earnings, along with a much smaller decline in lifetime inequality. In period 3, the recurring influx of higher-skilled younger generation restores the generational balance in skill distribution, and the covariance of lifetime earnings with life-cycle variation in earnings rises back to its first-period level of zero, and though this large rise is partly offset by a further decline in lifetime inequality, it causes inequality in current earnings to *rise* to a level slightly lower than its level in the first period.

Table B1: Annual earnings and average lifetime earnings over three periods

		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

Table B2: Decomposition of changes in the inequality of annual earnings over three periods

Decomposition	$V(y_{it})$ Inequality of annual earnings	$V(\hat{Y}_i)$ Inequality of lifetime earnings	$V(y_{it} - \hat{Y}_i)$ Variance of life-cycle change	$2Cov(y_{it} - \hat{Y}_i, \hat{Y}_i)$ Covariance of life-cycle change and lifetime earnings
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
Share of change		7.3%	0	92.7%
Change, period 2 to 3	0.073	-0.023	0	0.095
Share of change		-31.0%	0	131.0%
Change, period 1 to 3	-0.030	-0.030	0	0
Share of change		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 a rise in demand for high-skilled workers is met gradually*
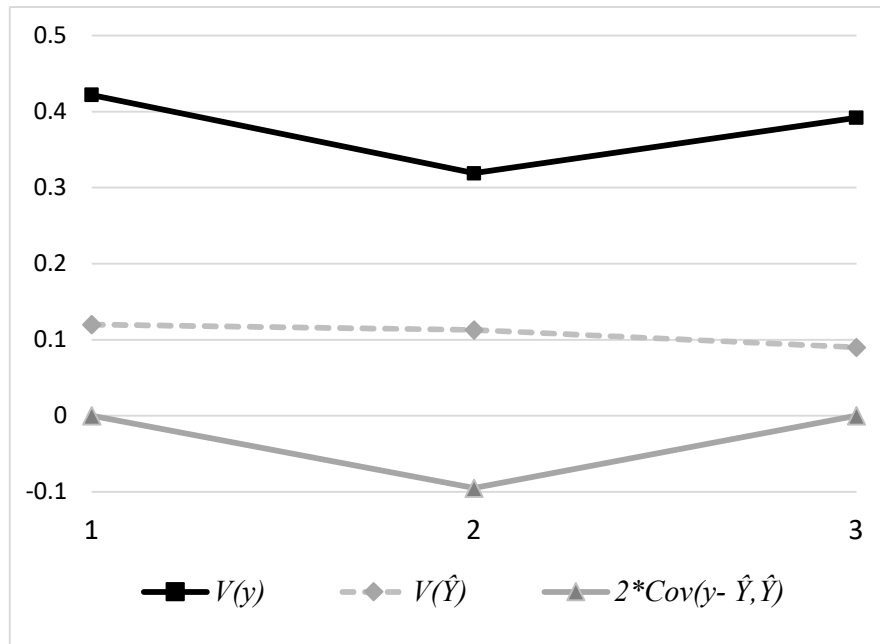
Second example. In this example, the economy is always evenly divided between lower-skilled and higher-skilled workers, but their age distribution changes. In the first period, three of the younger generation are lower-skilled and one is higher-skilled, while older-generation worker is lower-skilled and three are higher-skilled. In the second period, the first-period older generation exits the economy, the younger generation becomes the older generation, and a new younger generation enters the economy with the same skill composition as the exiting older generation—three higher-skilled and one lower-skilled—so that employment remains evenly divided between the lower-skilled and the higher-skilled. This two-year cycle then repeats itself. Table B3 presents annual and average lifetime earnings in each period. Lifetime inequality, $V(\hat{Y}_i)$, and the variance of life-cycle change, $V(y_{it} - \hat{Y}_i)$, are unaffected by these changes. The variation in current inequality stems entirely from changes in the covariance of lifetime earnings with life-cycle variation in earnings: it is higher in years when the higher-skilled are mostly older, and lower when they are mostly younger. Table B4 and Figure B2 show, numerically and graphically, the log variance of annual earnings and its decomposition.

Table B3: *Annual earnings and average lifetime earnings during a cyclical change in the age distribution of earnings*

		Young				Old			
Period 1	Annual earnings	20	20	20	40	120	120	120	60
	Avg lifetime earnings	40	40	40	80	80	80	80	40
Period 2	Annual earnings	20	40	40	40	60	60	60	120
	Avg lifetime earnings	40	80	80	80	40	40	40	80
Period 3	Annual earnings	20	20	20	40	120	120	120	60
	Avg lifetime earnings	40	40	40	80	80	80	80	40
Period 4	Annual earnings	20	40	40	40	60	60	60	120
	Avg lifetime earnings	40	80	80	80	40	40	40	80

Table B4: Decomposition of changes in the inequality of annual earnings over four periods

Decomposition	$V(y_{it})$ Inequality of annual earnings	$V(\hat{Y}_i)$ Inequality of lifetime earnings	$V(y_{it} - \hat{Y}_i)$ Variance of life-cycle change	$2Cov(y_{it} - \hat{Y}_i, \hat{Y}_i)$ Covariance of life-cycle change and lifetime earnings
Period 1 value	0.612	0.120	0.302	0.190
Period 2 value	0.231	0.120	0.302	-0.190
Period 3 value	0.612	0.120	0.302	0.190
Period 4 value	0.231	0.120	0.302	-0.190
Change, period 1 to 2	-0.380	0	0	-0.380
Change, period 2 to 3	0.380	0	0	0.380
Change, period 3 to 4	-0.380	0	0	-0.380

Table B2: Changes in the inequality of annual earnings, $V(y_{it})$ and lifetime earnings, $V(\hat{Y}_i)$, and in twice the covariance of life-cycle variation in earnings and lifetime earnings, $2Cov(y_{it} - \hat{Y}_i, \hat{Y}_i)$, as the economy cycles between age-earnings distributions
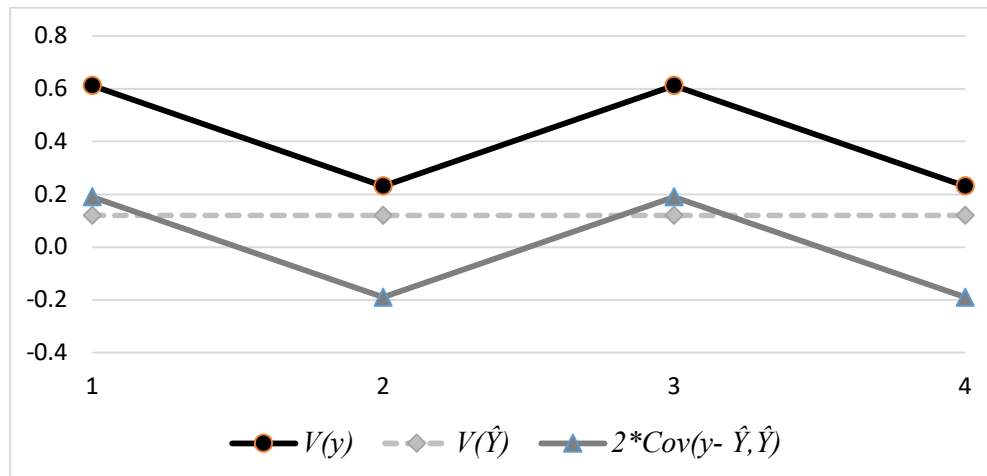
In both these examples, though there are no transitory shocks, no change in the dispersion of earnings over the life-cycle, $V(y_{it} - \hat{Y}_i)$, and little or no change in the inequality of lifetime earnings, $V(\hat{Y}_i)$, there is substantial change in the inequality of current earnings, $V(y_{it})$, driven by changes in the life-cycle variation in earnings and lifetime earnings, $2Cov(y_{it} - \hat{Y}_i, \hat{Y}_i)$.