Econometrica Supplementary Material

SUPPLEMENT TO "WHAT DO DATA ON MILLIONS OF U.S. WORKERS REVEAL ABOUT LIFECYCLE EARNINGS DYNAMICS?"

(*Econometrica*, Vol. 89, No. 5, September 2021, 2303–2339)

FATIH GUVENEN
Department of Economics, University of Minnesota, FRB of Minneapolis, and NBER

FATIH KARAHAN
Microeconomic Studies Function, FRB of New York

SERDAR OZKAN
Department of Economics, University of Toronto

JAE SONG
Social Security Administration

APPENDIX A: DATA APPENDIX

CONSTRUCTING A NATIONALLY REPRESENTATIVE PANEL OF MALES from the MEF is relatively straightforward. The last four digits of the SSN are randomly assigned, which allows us to pick a number for the last digit and select all individuals in 1978 whose SSN ends with that number.[1] This process yields a 10% random sample of all SSNs issued in the United States in or before 1978. Using SSA death records, we drop individuals who are deceased in or before 1978 and further restrict the sample to those between ages 25 and 60. In 1979, we continue with this process of selecting the same last digit of the SSN. Individuals who survived from 1978 and who did not turn 61 continue to be present in the sample, whereas 10% of new individuals who just turned 25 are automatically added (because they will have the last digit we preselected), and those who died in or before 1979 are again dropped. Continuing with this process yields a 10% representative sample of U.S. males in every year from 1978 to 2013.

The measure of wage earnings in the MEF includes all wages and salaries, tips, restricted stock grants, exercised stock options, severance payments, and many other types of income considered remuneration for labor services by the IRS as reported on the W-2 form (Box 1). This measure does not include any pre-tax payments to IRAs, retirement annuities, independent child care expense accounts, or other deferred compensation. We apportion 2/3 of the self-employment income as labor income. Given the lack of direct data on this, the 2/3 allocation has been the convention adopted by the literature as well as the PSID. In a previous version, we ignored self-employment income altogether and found similar results, leading us to believe that the exact allocation matters very little.

Finally, the MEF has a small number of extremely high earnings observations. For privacy and confidentiality reasons, we cap (winsorize) observations above the 99.999th percentile of the year-specific income distribution. For background information and detailed

Fatih Guvenen: guvenen@umn.edu
Fatih Karahan: fatih.karahan@ny.frb.org
Serdar Ozkan: serdar.ozkan@utoronto.ca
Jae Song: jae.song@ssa.gov

[1]In reality, each individual is assigned a transformation of their SSN number for privacy reasons, but the same method applies.

 https://doi.org/10.3982/ECTA14603

TABLE A.I

SAMPLE SIZE STATISTICS FOR CROSS-SECTIONAL MOMENTS OF FIVE-YEAR EARNINGS GROWTH[a]

	# Observations in Each RE Percentile Group			
Age group	Median	Min	Max	Total ('000s)
28–34	141,914	75,417	147,867	13,593
35–44	202,203	103,688	210,169	19,193
45–54	171,043	91,058	180,318	16,312

[a]Note: Each row reports several statistics for the number of observations used to compute the cross-sectional moments of five-year earnings changes for a given age group. Since cross-sectional moments are computed for each age-year-recent earnings percentile cell and then averaged over all years, sample sizes refer to the sum across all years of a given age by percentile group. The last column ("Total") reports the sum of observations across all 100 RE percentile groups for the age group indicated.

documentation of the MEF, see Panis, Euller, Grant, Bradley, Peterson, Hirscher, and Stinberg (2000) and Olsen and Hudson (2009).

Table A.I shows some sample size statistics regarding the sample used in the cross-sectional moments. Recall that we compute these statistics for each age-year-recent earnings percentile and aggregate them across years. Therefore, sample sizes refer to the sum across all years of a given age by percentile group. Each row reports the median, minimum, maximum, and total number of observations used to compute the cross-sectional moments for a given age group. Note that even the smallest cell has a sample size of more than 75,000 on which the computation of higher-order moments is based.

A.1. *Imputation of Self-Employment Income Above SSA Taxable Limit*

We restrict our main sample for cross-sectional and impulse response moments to years between 1994 and 2013 during which neither self-employment income nor wage/salary income is capped. However, this sample period–covering only 20 years–is too short to construct reliable measures of lifetime incomes of individuals. For this purpose, lifetime income moments in Section 5 are computed using the whole sample that covers 36 years between 1978 and 2013. But self-employment income is capped by the SSA maximum taxable earnings limit before 1994. In this section, we introduce a methodology to impute self-employment income above the top code for years before 1994, and show that imputing self-employment income has a negligible effect on our results.

Let $y_t^{\max}$ be the official SSA maximum taxable earnings limit in year t. Our goal is to impute the uncapped (unobservable) self employment income measure, $\tilde{y}_{i,t}^{\mathrm{SE}}$ for individuals who have self-employment income around the maximum taxable earnings limit reported in the MEF data specified by threshold $\chi y_t^{\max}$ (i.e., $y_{it}^{\mathrm{SE}} \geq \chi y_t^{\max}$), where $\chi < 1$.[2] For this purpose, we take the uncapped self-employment income measure in 1996, $y_{i,1996}^{\mathrm{SE}}$, and regress it on observables that can also be constructed for the period before 1994.[3] In particular, we first group workers into three bins based on their age in year 1996: 28–

[2]We assume $\chi = 0.95 < 1$, because the MEF data have several observations above the SSA taxable limit implying measurement error around the limit.

[3]The first year with uncapped self-employment income is 1994, but we use 1996 self-employment income in the regression due to measurement issues in the 1992 data for self-employment income.

29, 30–34, and 35–40.[4] Next, within each age group h, we estimate quantile regressions of uncapped self-employment income in 1996 for 75 equally spaced quantiles τ. Thus, in total we estimate the following specification 3×75 times—one for each age group and quantile:

$$\begin{aligned}
&\log y^{\mathrm{SE}}_{i,1996} \\
&\quad = \sum_{k=0}^{3} \alpha^{h,\tau}_{1,k} \mathbb{I}\{y^{W}_{i,1996-k} < Y_{\min,1996-k}\} + \sum_{k=0}^{3} \alpha^{h,\tau}_{2,k} \mathbb{I}\{y^{W}_{i,1996-k} \geq Y_{\min,1996-k}\} \log y^{W}_{i,1996-k} \\
&\qquad + \sum_{k=1}^{3} \alpha^{h,\tau}_{3,k} \mathbb{I}\{y^{\mathrm{SE}}_{i,1996-k} > \chi y^{\max}_{1996-k}\} + \sum_{k=1}^{3} \alpha^{h,\tau}_{4,k} \mathbb{I}\{y^{\mathrm{SE}}_{i,1996-k} < Y_{\min,1996-k}\} \\
&\qquad + \sum_{k=1}^{3} \alpha^{h,\tau}_{5,k} \mathbb{I}\{y^{\mathrm{SE}}_{i,1996-k} \geq Y_{\min,1996-k}\} \min\bigl(\log y^{\mathrm{SE}}_{i,1996-k}, \log \chi y^{\max}_{1996-k}\bigr) + \varepsilon_{it},
\end{aligned} \tag{A.1}$$

where $y^{W}_{i,t}$ is the wage and salary income of individual i in year t, $\mathbb{I}$ is the indicator function, and ε_{it} is the residual term. The right-hand-side variables are as follows: (i) a dummy variable of whether the worker's wage earnings $y^{W}_{i,t}$ is less than the minimum income threshold $Y_{\min,t}$; (ii) if it is higher than $Y_{\min,t}$, the log of wage earnings $\log y^{W}_{i,t}$; (iii) a dummy variable of whether the self-employment income $y^{\mathrm{SE}}_{i,t}$ is above the maximum cap $\chi y^{\max}_{t}$; (iv) a dummy variable of whether $y^{\mathrm{SE}}_{i,t}$ is less than the minimum threshold $Y_{\min,t}$; (v) if it is higher than $Y_{\min,t}$, the log self-employment income capped at the maximum threshold $\log(\min(y^{\mathrm{SE}}_{i,t}, \chi y^{\max}_{t}))$. We also include three lags of these as independent variables. Then, $\alpha^{h,\tau}_{i,k}$ denotes the regression coefficient of variable i with lag k for age group h, quantile τ.

We then use these regression coefficients to impute the uncapped self-employment income before 1994 for individuals who have SE income above the limit $\chi y^{\max}_{t}$ reported in the MEF data. For this purpose, we randomly assign individuals to quantiles $\tau = 1, \dots, 75$ in our lifetime income sample. Then, the imputed self-employment income for an individual in age group h with quantile τ who has recorded self-employment income above the limit $\chi y^{\max}_{t}$ in year $t = 1981, 1981, \dots, 1993$ is given by the following equation:[5]

$$\begin{aligned}
\log \tilde{y}^{\mathrm{SE}}_{i,t} &= \sum_{k=0}^{3} \alpha^{h,\tau}_{1,k} \mathbb{I}\{y^{W}_{i,t-k} < Y_{\min,t-k}\} + \sum_{k=0}^{3} \alpha^{h,\tau}_{2,k} \mathbb{I}\{y^{W}_{i,t-k} \geq Y_{\min,t-k}\} \log y^{W}_{i,t-k} \\
&\quad + \sum_{k=1}^{3} \alpha^{h,\tau}_{3,k} \mathbb{I}\{y^{\mathrm{SE}}_{i,t-k} > \chi y^{\max}_{t-k}\} + \sum_{k=1}^{3} \alpha^{h,\tau}_{4,k} \mathbb{I}\{y^{\mathrm{SE}}_{i,t-k} < Y_{\min,t-k}\} \\
&\quad + \sum_{k=1}^{3} \alpha^{h,\tau}_{5,k} \mathbb{I}\{y^{\mathrm{SE}}_{i,t-k} \geq Y_{\min,t-k}\} \min\bigl(\log y^{\mathrm{SE}}_{i,t-k}, \log \chi y^{\max}_{t-k}\bigr).
\end{aligned} \tag{A.2}$$

[4]Our imputed lifetime income sample employs a balanced panel that selects all individuals who are between ages 25 and 28 in 1981 (who were born between 1954 and 1957). This condition ensures that we have 33 years of earnings-histories between ages 25 and 60 for each individual (which might include years with zero earnings). The same condition also implies that, in this sample, only workers younger than 40 are affected by the top coding until 1993 and we impute their capped self-employment income.

[5]The imputed lifetime income sample starts with year 1981 because, to impute self-employment income, we need to observe wage and self-employment income in the previous three years between 1978 and 1980.

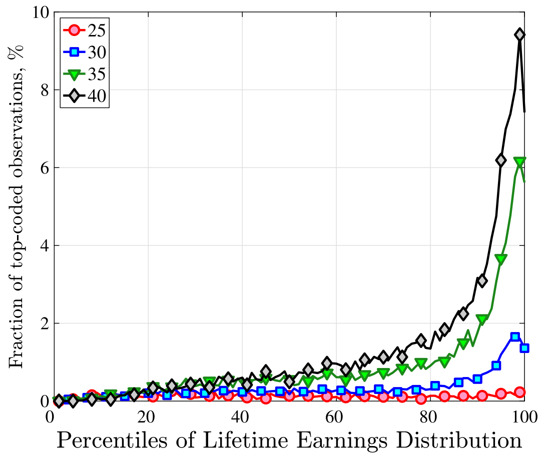


FIGURE A.1.—Fraction of top-coded self-employment income observations.

Figure A.1 plots the fraction of top-coded self-employment income observations against the percentiles of lifetime earnings distribution at ages 25, 30, 35, and 40 in our imputed lifetime income sample.[6] Almost no observations are top coded for individuals below the 20th percentile of the lifetime earnings distribution, in particular at young ages. As expected, the fraction of top-coded observations increases with age and with lifetime earnings and is highest for workers in the 99th percentile when they are 40 years old.

Furthermore, Figure A.2 plots the lifetime income growth between 25 and 55 against lifetime earnings percentiles using imputed and nonimputed data, which is already shown in Figure 11(a). The two series are almost indistinguishable, indicating that top coding has very little effect on lifetime income growth. This is because only a very small number of workers are affected by the top coding; those who had very high self-employment income before 1994 or when the cohort was younger than age 41.

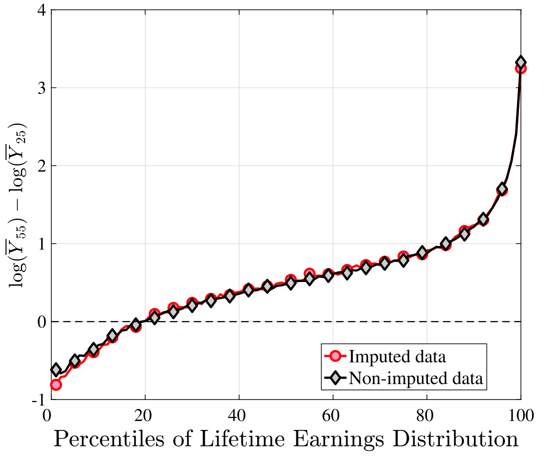


FIGURE A.2.—Income growth for imputed and nonimputed data.

[6]Recall that in this sample, only workers younger than 40 are affected by the top coding.

APPENDIX B: DERIVATION OF HIGHER-ORDER MOMENTS OF LOG CHANGE

Let us suppose that the earnings dynamics are given by the commonly used random-walk permanent/transitory model in which i.i.d. permanent (η_t^i) and transitory (ε_t^i) innovations are drawn from some general distributions F_η and F_ε, respectively. Then, the k-year log growth of earnings is given by

$$\Delta_{\log}^k y_t^i = y_{t+k}^i - y_t^i = \sum_{j=t+1}^{t+k} \eta_j^i + \varepsilon_{t+k}^i - \varepsilon_t^i.$$

Let us denote the variance, skewness, and excess kurtosis of distribution F_x, $x \in \{\eta, \varepsilon\}$ by σ_x^2, $\mathcal{S}_x$, and $\mathcal{K}_x$, respectively. Then the variance is given by

$$\sigma^2(\Delta_{\log}^k y_t^i) = k\sigma_\eta^2 + 2\sigma_\varepsilon^2.$$

In order to derive the skewness of $\Delta_{\log}^k y_t^i$, we use the following properties:

$$\begin{aligned}
\mathcal{S}(kx) &= \mathcal{S}_x, \quad \text{for any } k > 0,\\
\mathcal{S}(x+y) &= \left(\frac{\sigma_x}{\sigma_{x+y}}\right)^3 \times \mathcal{S}_x + \left(\frac{\sigma_y}{\sigma_{x+y}}\right)^3 \times \mathcal{S}_y,\\
\mathcal{S}(x-y) &= \left(\frac{\sigma_x}{\sigma_{x+y}}\right)^3 \times \mathcal{S}_x - \left(\frac{\sigma_x}{\sigma_{x+y}}\right)^3 \times \mathcal{S}_y.
\end{aligned}$$

Then:

$$\begin{aligned}
\mathcal{S}(\Delta_{\log}^k y_t^i) &= \sum_{j=t+1}^{t+k} \left(\frac{\sigma_\eta}{\sigma^2(\Delta_{\log}^k y_t^i)}\right)^3 \times \mathcal{S}_\eta \\
&\quad + \left(\frac{\sigma_\varepsilon}{\sigma^2(\Delta_{\log}^k y_t^i)}\right)^3 \times \mathcal{S}_\varepsilon - \left(\frac{\sigma_\varepsilon}{\sigma^2(\Delta_{\log}^k y_t^i)}\right)^3 \times \mathcal{S}_\varepsilon \\
&= \frac{k\sigma_\eta^3 S_\eta}{\sigma^3(\Delta_{\log}^k y_t^i)}.
\end{aligned}$$

In order to derive the kurtosis of $\Delta_{\log}^k y_t^i$, we use the following properties:

$$\begin{aligned}
\mathcal{K}(kx) &= \mathcal{K}_x, \quad \text{for any } k > 0,\\
\mathcal{K}\left(\sum_{j=1}^k x_j\right) &= \sum_{j=1}^k \left[\left(\frac{\sigma_{x_j}}{\sigma\left(\sum_j x_j\right)}\right)^4 \cdot \mathcal{K}_{x_j}\right].
\end{aligned}$$

We obtain

$$\mathcal{K}(\Delta_{\log}^k y_t^i) = \frac{k \times \sigma_\eta^4}{\sigma^4(\Delta_{\log}^k y_t^i)} \mathcal{K}_\eta + \frac{2 \times \sigma_\varepsilon^4}{\sigma^4(\Delta_{\log}^k y_t^i)} \mathcal{K}_\varepsilon.$$

APPENDIX C: ROBUSTNESS AND ADDITIONAL FIGURES

This section reports additional results from the data. Section C.1 reports the cross-sectional moments of one-year earnings growth. Section C.2 shows the cross-sectional moments of one-year and five-year arc-percent changes of earnings. Section C.3 presents several features of the data on log earnings changes that are mentioned in the paper but are relegated to the Appendix. In Section C.4, we show several moments of persistent earnings changes based on the measure introduced in Section 3.4. Section C.5 provides further analysis regarding the higher-order moments of job-stayers and -switchers. Section C.6 documents cross-sectional moments using a much broader sample and shows that the changes in higher-order moments are not driven by the particular sample selection criteria used in the main text. In Section C.6, we present properties of earnings changes in survey data. Section C.8 investigates the role of Social Security Disability Income in our findings. Finally, Section C.9 shows several results about the lifecycle profile of earnings and employment that were left out of the main analysis.

C.1. *Cross-Sectional Moments of One-Year Log Earnings Growth*

Throughout the main text, we showed the cross-sectional moments of five-year (log) earnings growth. Figures C.1–C.4 show analogous features of the data for one-year earnings growth.

C.2. *Arc-Percent Moments*

In the main text, we documented moments of log earnings changes. In doing so, we are forced to drop observations close to zero to obtain sensible statistics. However, as we discuss in Section 2, such observations contain potentially valuable information, as they inform us about very large changes in earnings caused by events such as long-term nonemployment. To complement our analysis, this section reports the cross-sectional moments of arc-percent changes defined in Section 2, which we reproduce here for convenience:

$$\text{arc-percent change:} \quad \Delta_{\text{arc}} Y_{t,k}^i = \frac{Y_{t+k}^i - Y_t^i}{(Y_{t+k}^i + Y_t^i)/2}.$$

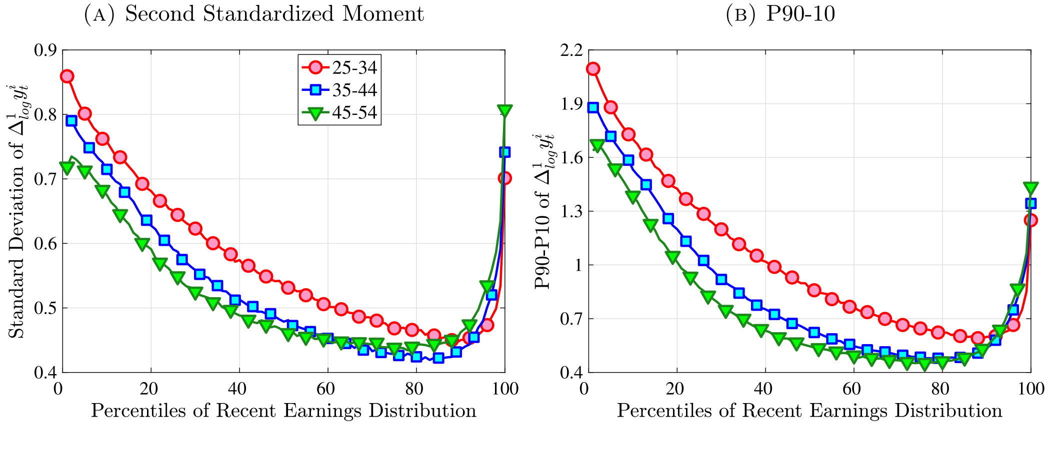


FIGURE C.1.—Dispersion of one-year log earnings growth.

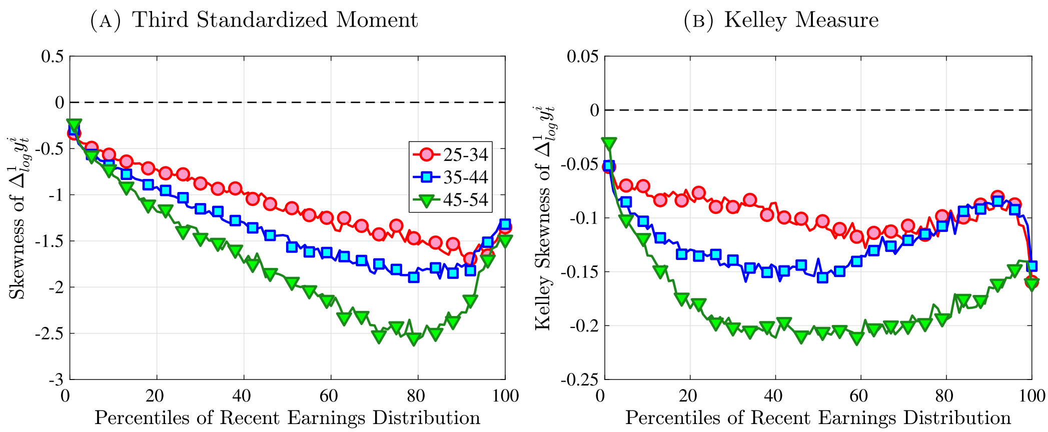


FIGURE C.2.—Skewness of one-year log earnings growth.

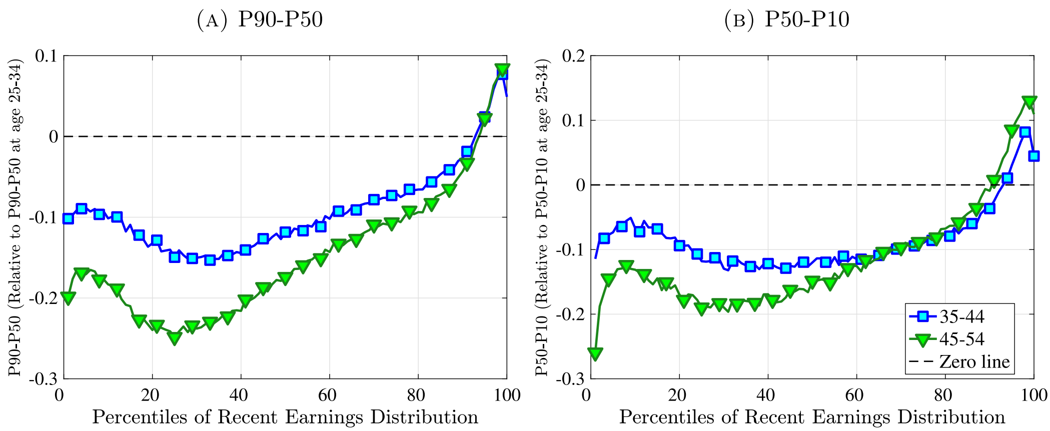


FIGURE C.3.—Kelley's skewness decomposed: change in P90–P50 and P50–P10 relative to age 25–29 (one-year log growth).

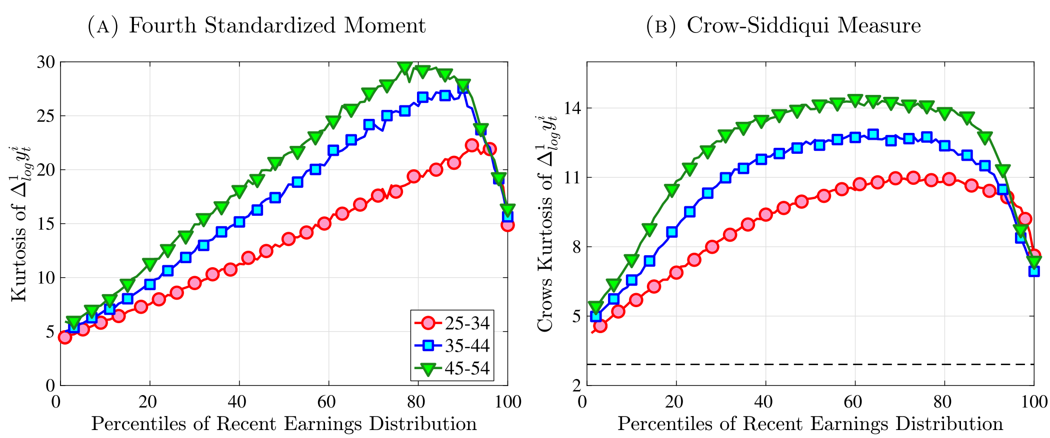


FIGURE C.4.—Kurtosis of one-year log earnings growth.

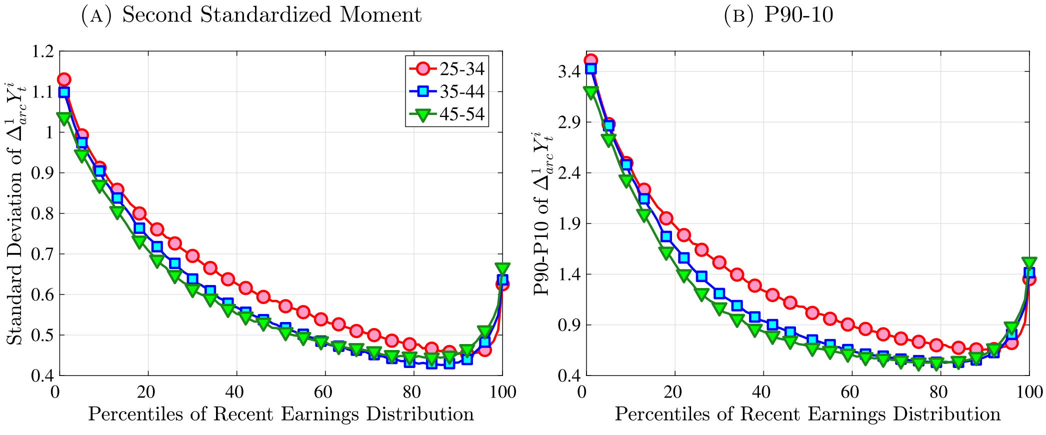


FIGURE C.5.—Dispersion of annual arc-percent earnings change.

This measure allows computation of earnings growth even when the individual has zero income in one of the two years t and $t+k$. Section C.2.1 shows the moments of one-year arc-percent change, and Section C.2.2 shows the moments of five-year change.

C.2.1. *Moments of Annual Arc-Percent Changes*

Figures C.5(a)–C.8(b) show the standardized moments of one-year arc-percent changes.

C.2.2. *Moments of Five-Year Arc-Percent Changes*

Figures C.9(a)–C.12(b) show the standardized moments of -year arc-percent changes.

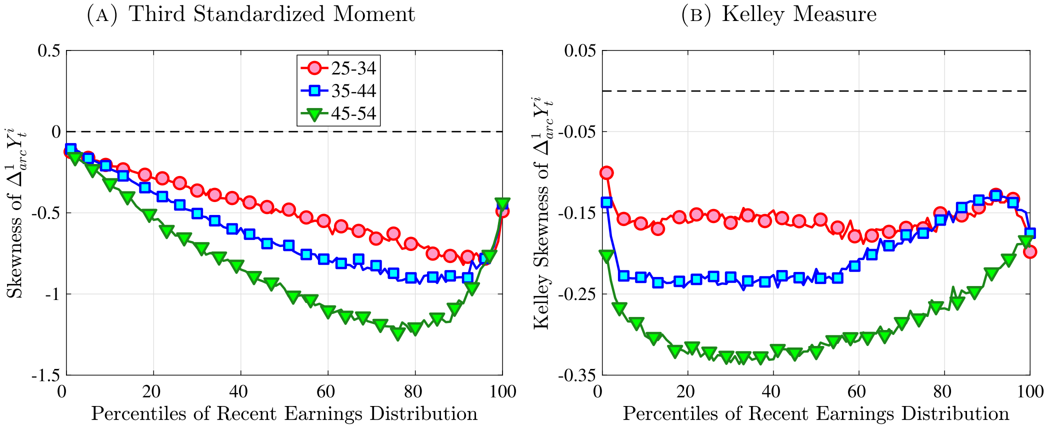


FIGURE C.6.—Skewness of annual arc-percent earnings change.

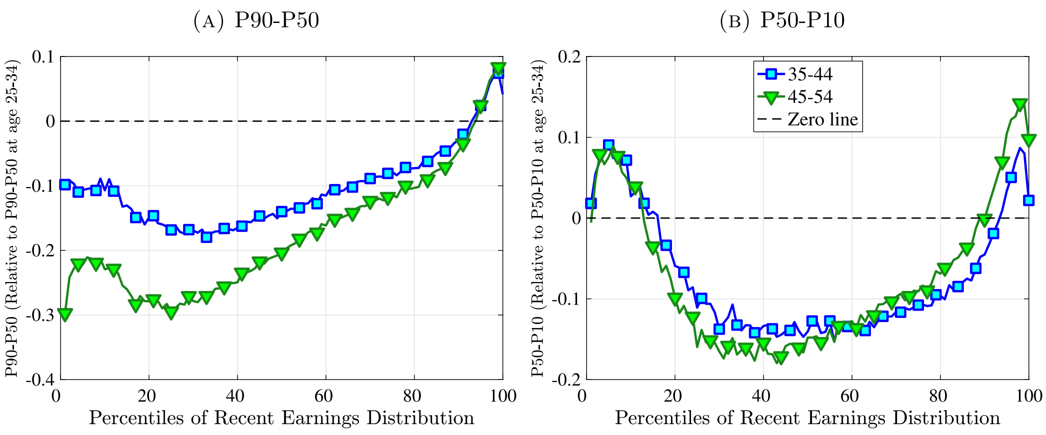


FIGURE C.7.—Kelley's skewness decomposed (annual arc-percent growth): change in P90–P50 and P50–P10 relative to age 25–34.

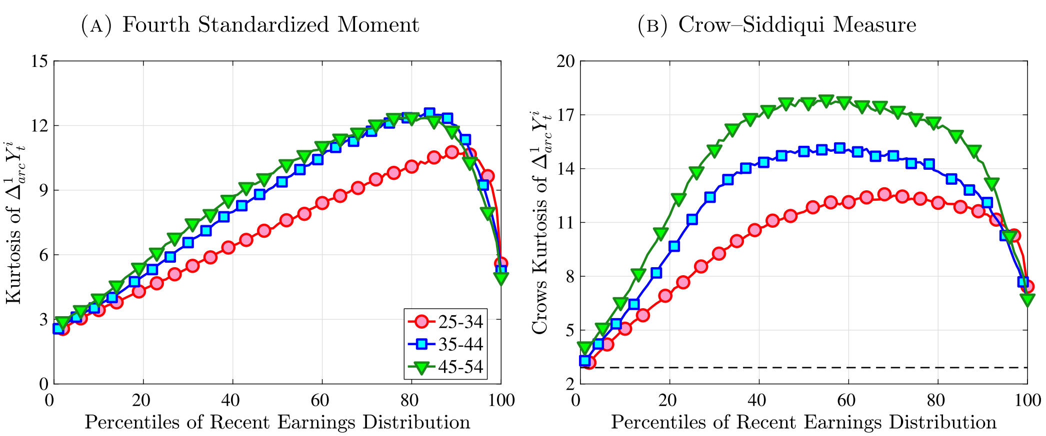


FIGURE C.8.—Kurtosis of annual arc-percent earnings change.

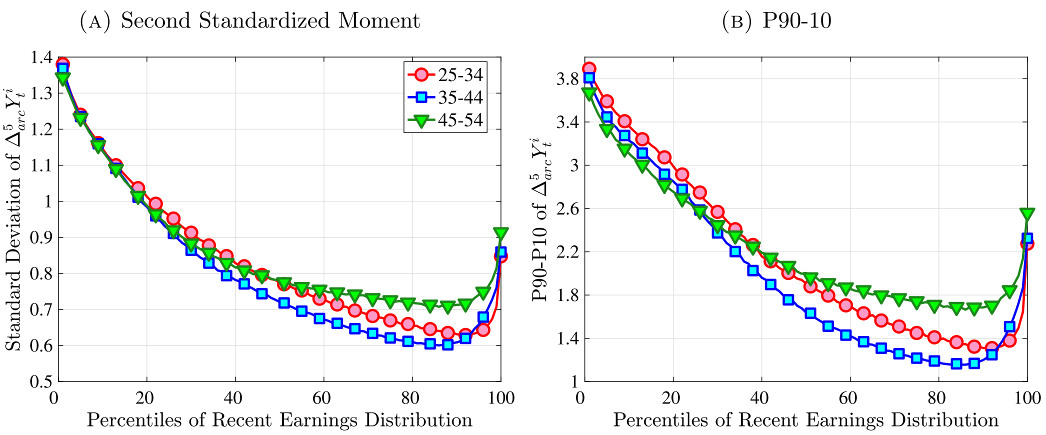


FIGURE C.9.—Dispersion of five-year arc-percent earnings change.

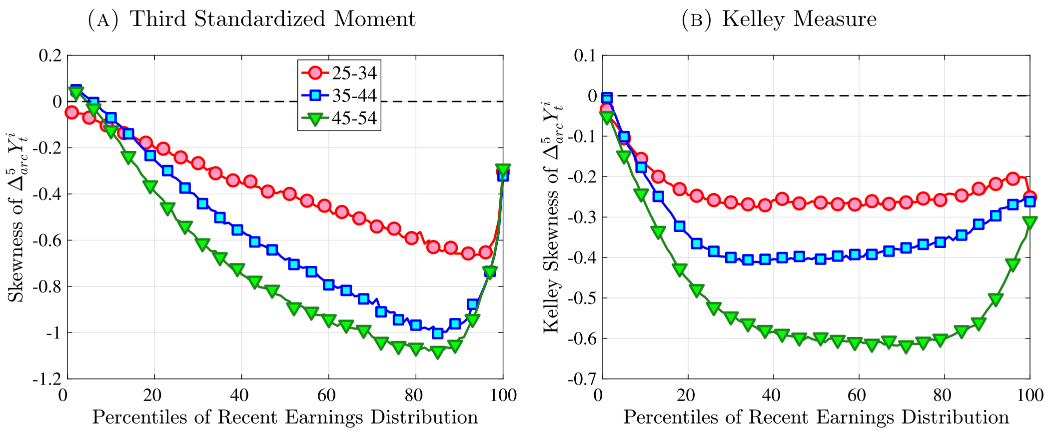


FIGURE C.10.—Skewness of five-year arc-percent earnings change.

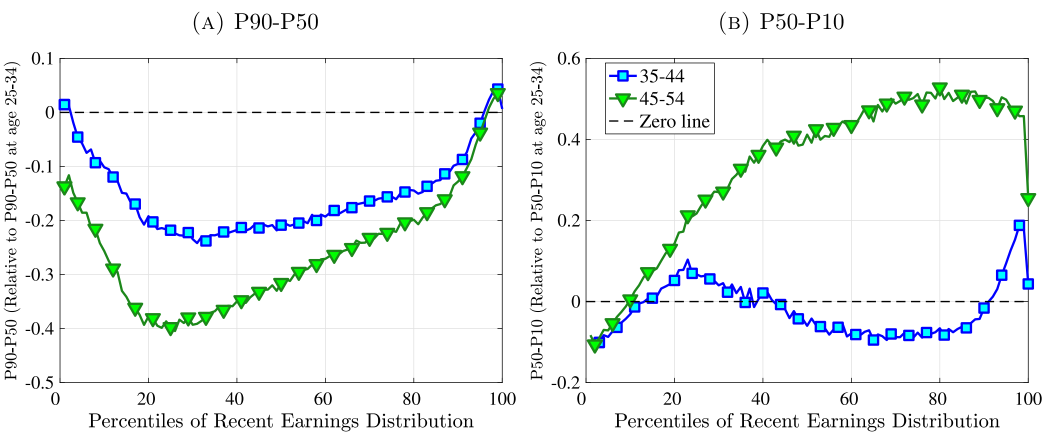


FIGURE C.11.—Kelley's skewness decomposed (5-year arc-percent growth): change in P90–P50 and P50–P10 relative to age 25–34.

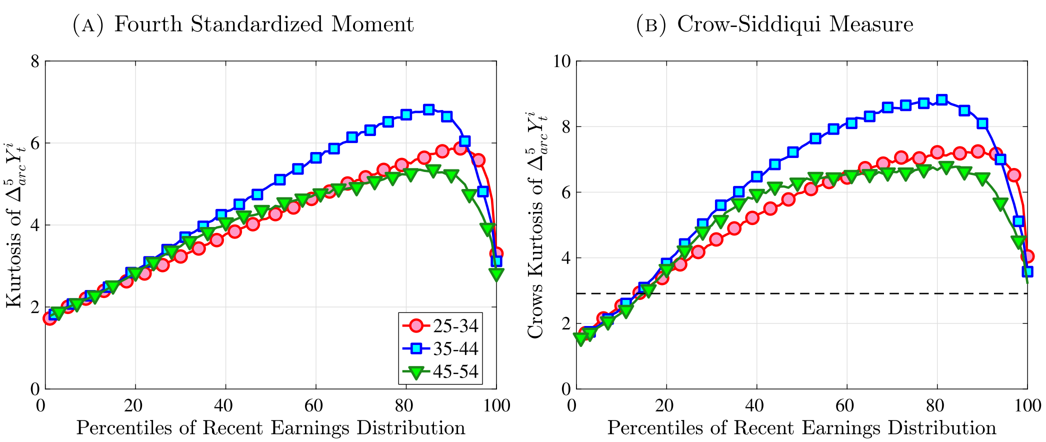
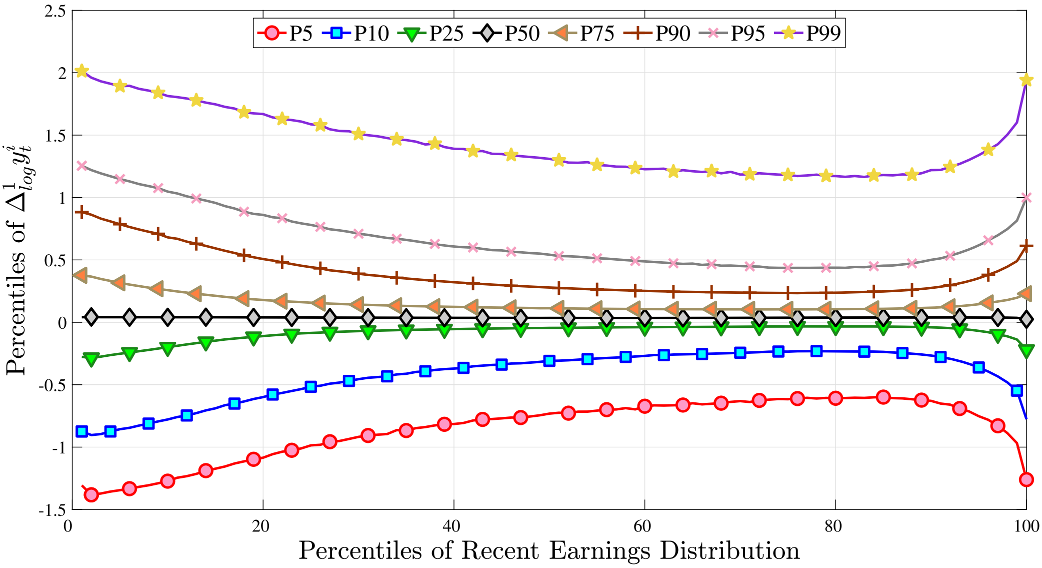


FIGURE C.12.—Kurtosis of five-year arc-percent earnings change.

(A) One-Year Changes



(B) Five-Year Changes

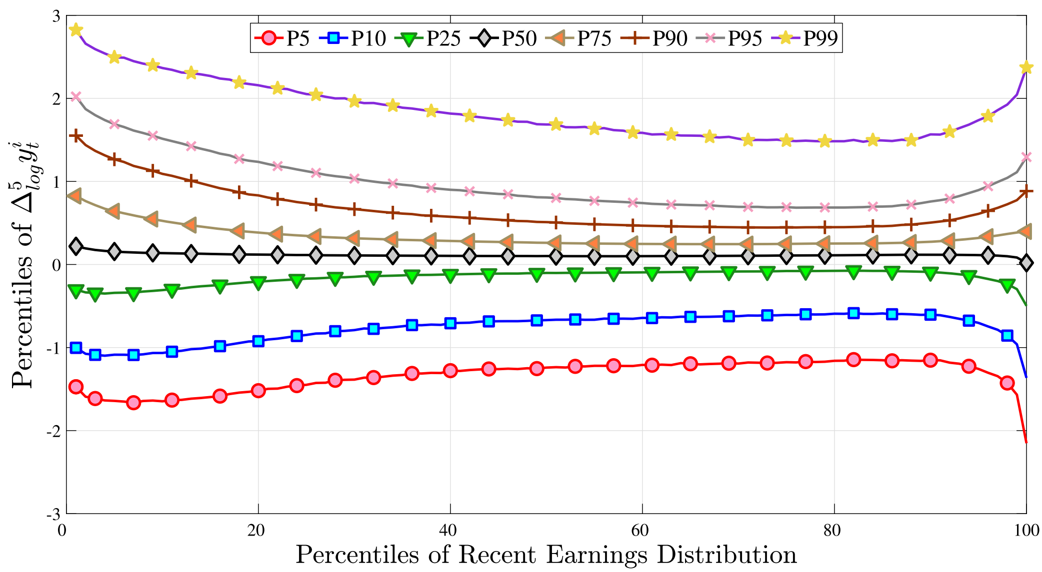


FIGURE C.13.—Selected percentiles of log earnings changes.

C.3. *Further Moments of Log Earnings Change*

In this section, we report some additional figures of interest that are omitted from the main text due to space constraints. First, Figure C.13 plots selected percentiles of the annual and five-year log earnings change distribution for every RE percentile.

Second, Figure C.14 shows an additional measure of kurtosis proposed by Moors (1988) for one- and five-year earnings changes. Similar to the measure proposed by Crow and

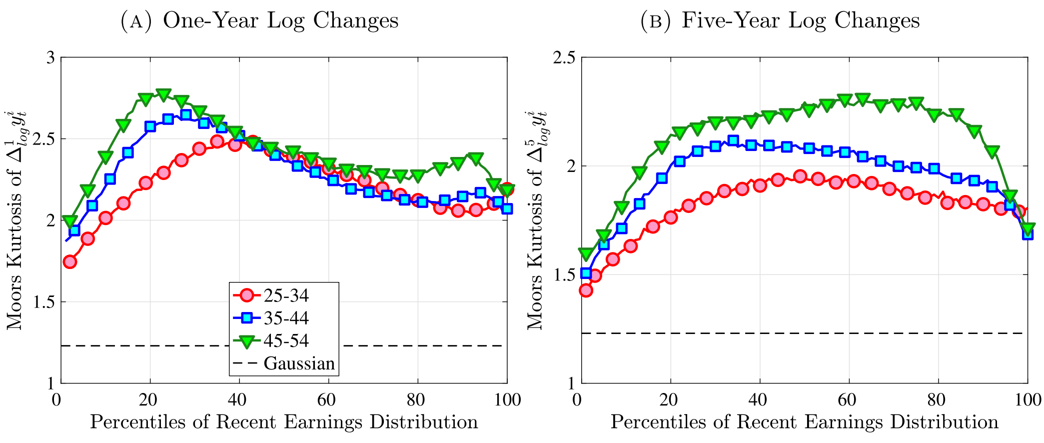


FIGURE C.14.—Moors's kurtosis of log earnings changes.

Siddiqui (1967), this measure is robust to outliers in the tails. Moors's kurtosis, κ_M, is defined as

$$\kappa_M = \frac{(\text{P87.5} - \text{P62.5}) + (\text{P37.5} - \text{P12.5})}{\text{P75} - \text{P25}}.$$

For a Gaussian distribution, Moors's kurtosis is 1.23 (shown on dashed lines).

C.4. *An Alternative Measure of Persistent Earnings Changes*

In Section 3, we studied the distribution of five-year earnings changes, and explained that the five-year changes reflect more of the distribution of the persistent innovations rather than transitory innovations. We also considered an alternative measure ($\overline{\Delta}^5_{\log}(\bar{y}_t{}^i) \equiv \bar{y}^i_{t+4} - \bar{y}^i_{t-1}$, where $\bar{y}^i_{t+4} \equiv \log(\bar{Y}^i_{t+4})$ and $\bar{y}^i_{t-1} \equiv \log(\bar{Y}^i_{t-1})$) to deal with the caveat that our baseline measure is contaminated by transitory changes in years t and $t+k$. The main text showed the standardized moments of this alternative measure; Figure C.15 shows the quantile-based moments.

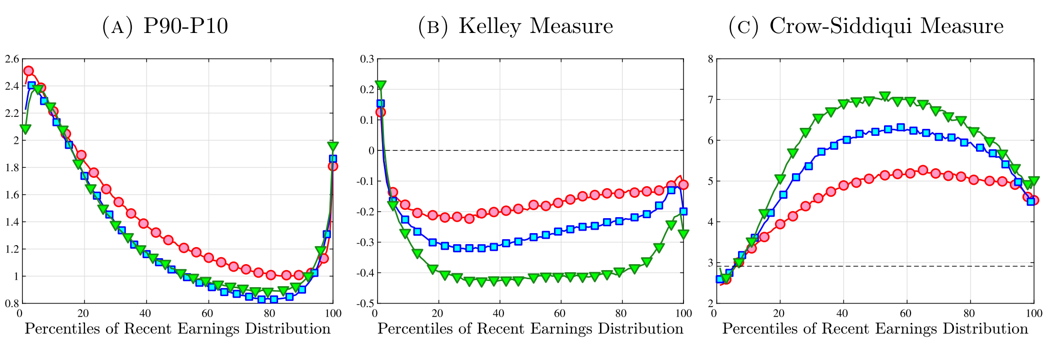


FIGURE C.15.—Alternative measure of persistent earnings changes, $\overline{\Delta}^5_{\log}(\bar{y}_t{}^i)$: quantile-based moments.

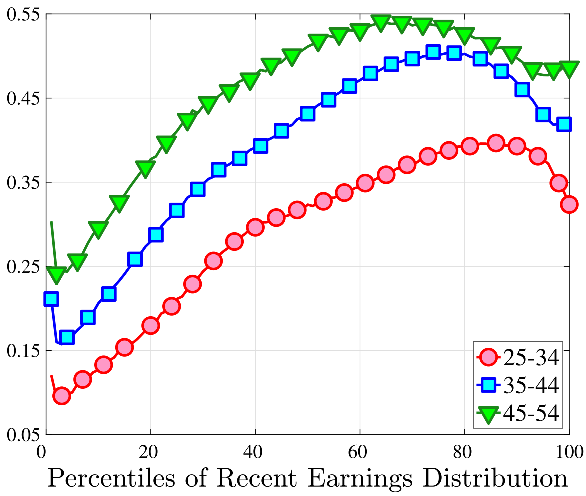


FIGURE C.16.—Fraction staying at jobs between t and $t+1$ (baseline definition).

C.5. *Cross-Sectional Moments for Job-Stayers versus Job-Switchers*

In the main text, we analyzed the properties of earnings growth separately for job-stayers and job-switchers by showing quantile-based moments of five-year earnings changes. Here, we first complement our analysis by showing several features of the data that were omitted in the main text to save space. Second, we consider an alternative definition of job-stayers and investigate the cross-sectional moments of earnings according to that definition.

Figure C.16 shows the fraction of job-stayers according to our baseline definition as a function of recent earnings and age. The probability of staying with the same employer increases with recent earnings and age. For the youngest age group (25–34), the probability of staying in the same job is around 20% at the bottom of the recent earnings distribution. This fraction increases with recent earnings and reaches a peak around 60% at the 95th percentile of the RE distribution. This pattern reverses itself slightly at the top of the RE distribution. As workers age, the probability of staying with the same employer increases across the RE distribution.

Second, Figures C.17 and C.18 show the age profile of higher-order moments shown in Section 3.5. The age patterns are broadly similar across switchers and stayers: P90–10 declines slightly, skewness becomes more negative, and kurtosis increases for both job-stayers and job-switchers over the life cycle.

Next, we complement our analysis of job-stayers and job-switchers by investigating the standardized moments of one- and five-year earnings changes (as opposed to quantile-based moments analyzed in the main text). We plot these moments in Figure C.19.

The results are consistent with what one might expect. Job-stayers face earnings changes that (i) have half the dispersion of job-switchers, (ii) are less negatively skewed as opposed to job-switchers, who face very negatively skewed changes, and (iii) have a much higher kurtosis than job-switchers. In fact, kurtosis is as high as 40 for annual changes and 25 for five-year changes for job-stayers, but is less than 10 for job-switchers at both horizons.

One caveat worth emphasizing again is that constructing a clean measure of job-stayers and job-switchers is not possible in our data set, primarily because of the annual nature

(A) P90-P10 (B) Kelley's Skewness (C) Crow–Siddiqui Kurtosis

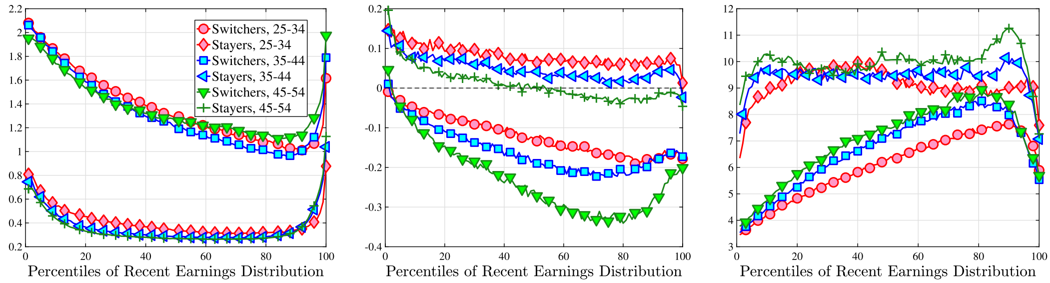


FIGURE C.17.—Quantile-based moments of one-year log earnings growth: stayers versus switchers.

of the data: We observe when employment spells begin and end only at annual frequency, and therefore cannot infer if a worker worked multiple jobs at the same time or if he switched employers at some point during the year, and if so, whether the change was a direct job-to-job switch or involved a nonemployment spell in between. In our baseline measure, we have opted to be very conservative when defining job-stayers. This measure probably understates job-stayers and overstates true switchers.

We now consider an alternative definition, which is a much more conservative definition of job-switchers. According to this definition, we call an individual a job-switcher in year t if (i) the largest paying employer is different in years t and $t+1$, (ii) the largest paying employers in years t and $t+1$ contribute to at least 75% of the worker's total salary, (iii) the worker either has no income in year $t+1$ from the main employer of year t, or if he does, that income in $t+1$ is less than 25% of what he made in t (from the same employer). Figure C.20 compares the share of job-stayers according to this alternative definition (left panel) to our baseline (right panel) and Figure C.21 compares the cross-sectional moments of one-year earnings changes. By construction, the fraction of individuals identified as job-stayers and job-switchers is quite different across the two definitions. Nevertheless, all of the substantive conclusions go through regarding the differences in the cross-sectional moments of job-stayers and job-switchers.

C.6. *Cross-Sectional Moments by Age Without Sample Selection*

When choosing the sample for cross-sectional moments, we required an individual to have an earnings level above the minimum threshold in $t-1$ and in at least two more

(A) P90-P10 (B) Kelley's Skewness (C) Crow–Siddiqui Kurtosis

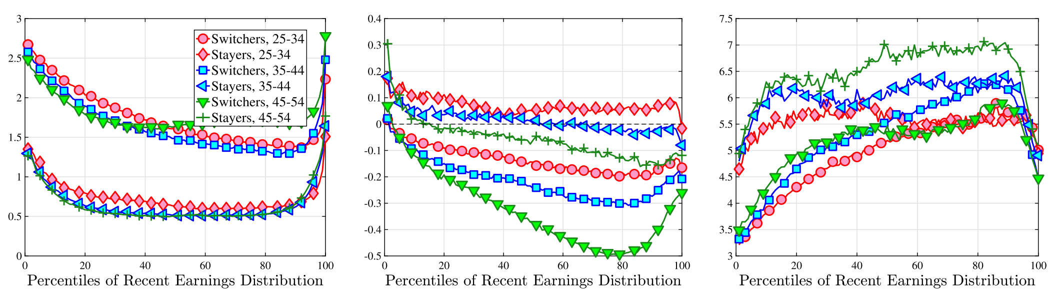


FIGURE C.18.—Quantile-based moments of five-year log earnings growth: stayers versus switchers.

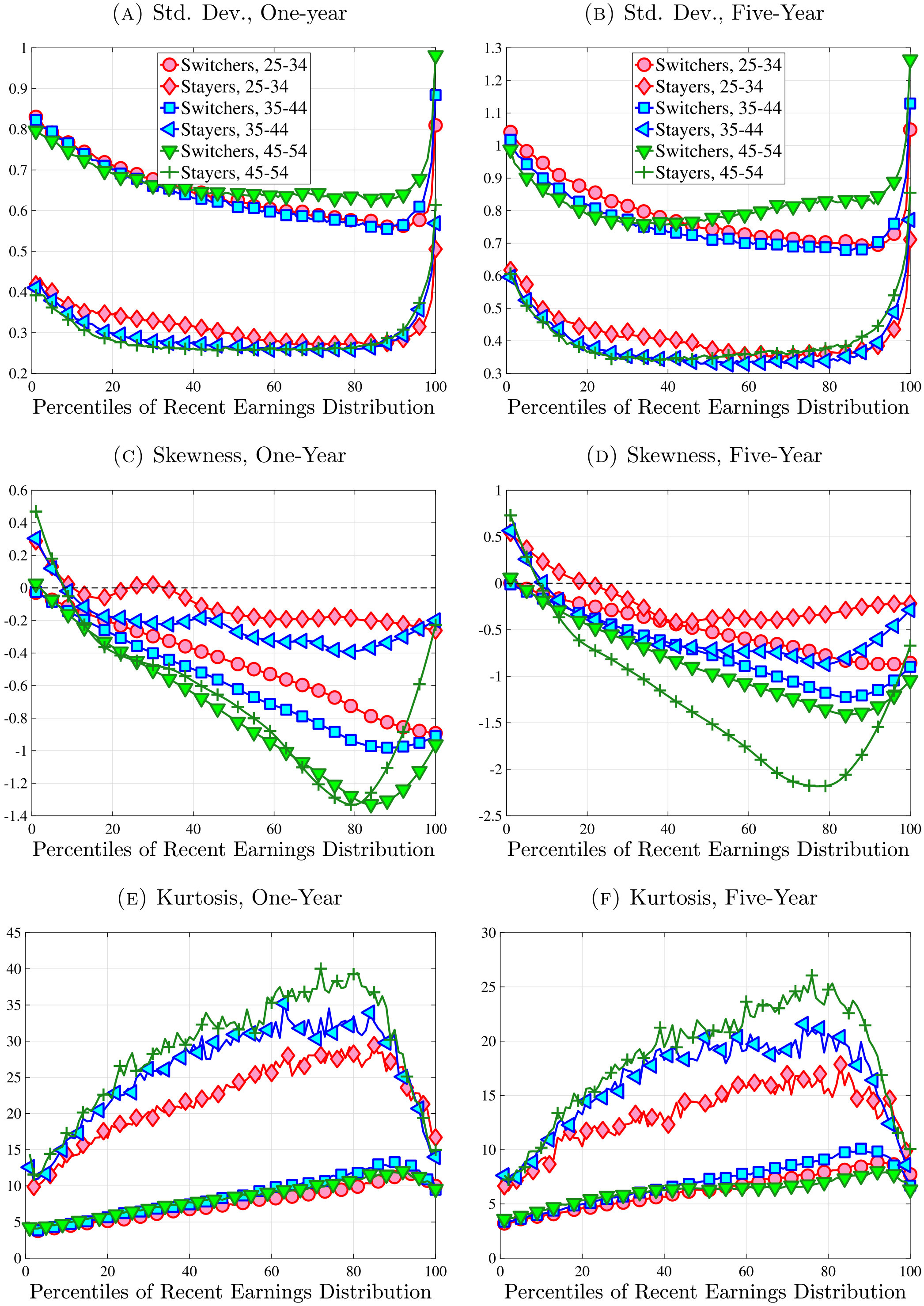


FIGURE C.19.—Standardized moments of log earnings growth: stayers versus switchers.

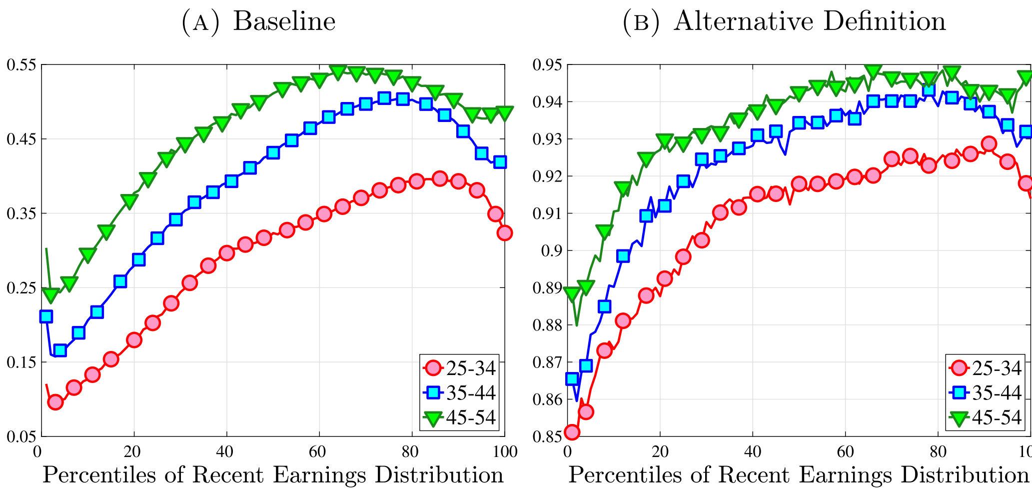


FIGURE C.20.—Fraction of job stayers.

years between $t-5$ and $t-2$. Figure C.22 shows that these conditions result in a substantial share of the initial sample being dropped from the analysis. This large selectivity opens up the possibility that some of our results might be specific to our final sample. Here, we relax the selection criteria and include any person for whom earnings change can be computed. Figures C.23–C.28 show the standardized and quantile-based moments of one-year and five-year earnings changes. We find that our substantive conclusions are unchanged: Dispersion of earnings changes declines with age for most of the life cycle, and earnings changes become more negatively skewed and more leptokurtic.

C.7. *Survey Data and Higher-Order Moments*

Panel Study of Income Dynamics (PSID)

The Panel Study of Income Dynamics has a smaller sample size compared to the CPS, but it has the advantage of following the same household over a much longer period of time. The PSID started collecting data annually in 1968 on a sample of around 5000 households, of which 3000 households were representative and the remaining were low-income families (the Census Bureau's Survey of Economic Opportunities sample, SEO). We restrict our study to households in the core sample and do not use households in the SEO or the Latino subsamples. The questions on income are retrospective, meaning that respondents in a given year are asked about the previous calendar year. We analyze data for the period 1999–2013. During this period, the survey was biennial.

Our measure of labor income (variable ER16463 in 1999) is the sum of wages and salaries, bonuses, pay for overtime, tips, commissions, professional practice or trade, market gardening, miscellaneous labor income, and extra job income. To remain consistent with the rest of the paper, we focus on male heads of household between ages 25 and 55. We deflate annual earnings by the price level with the base year 2010.[7] We drop observations that report earnings less than \$1500. We residualize earnings, wages, and hours by

[7]We use the consumer price index for all urban consumers (CPI-U) published by the Bureau of Labor Statistics.

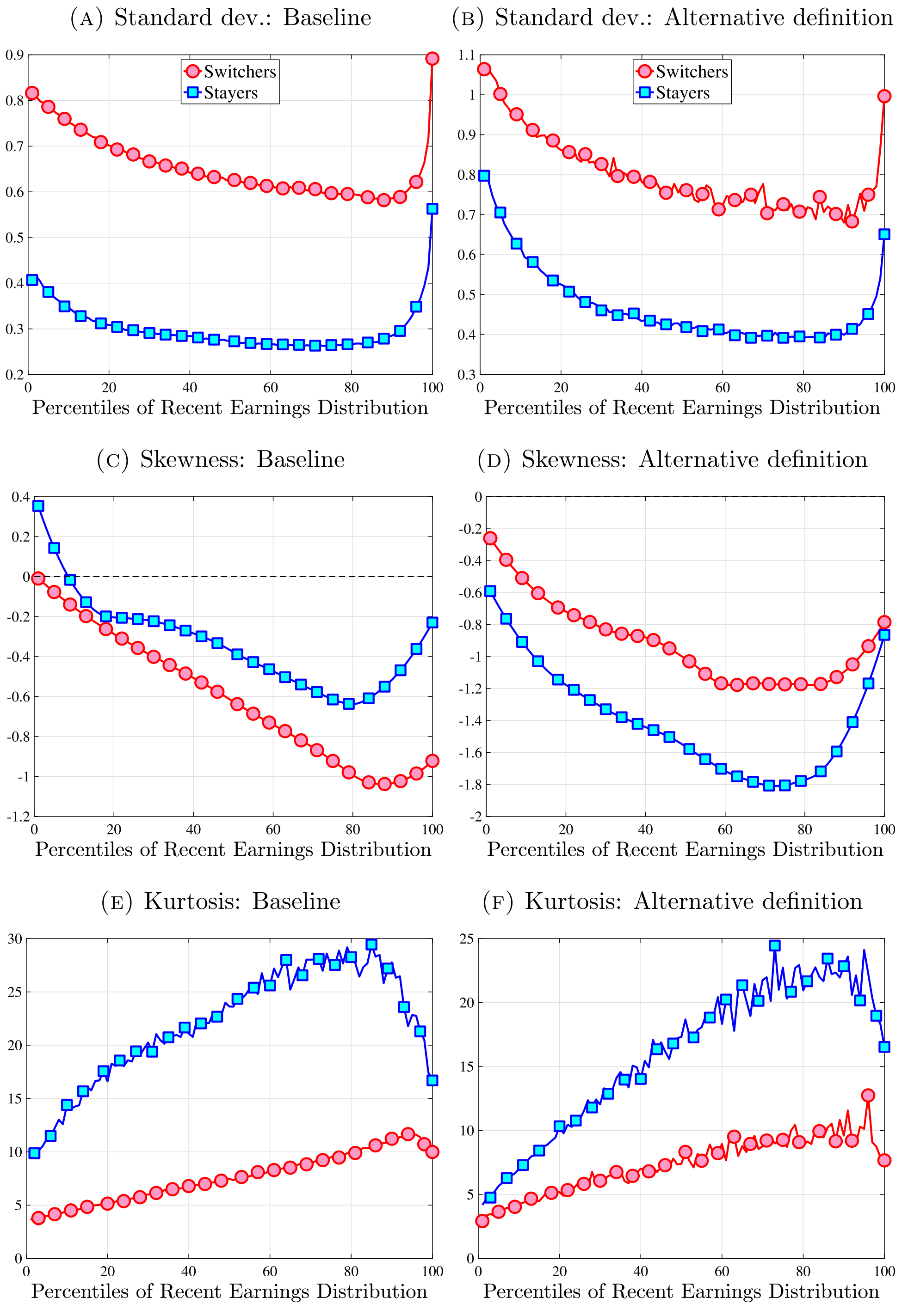


FIGURE C.21.—Alternative job-stayer definition: cross-sectional moments of one-year earnings growth.

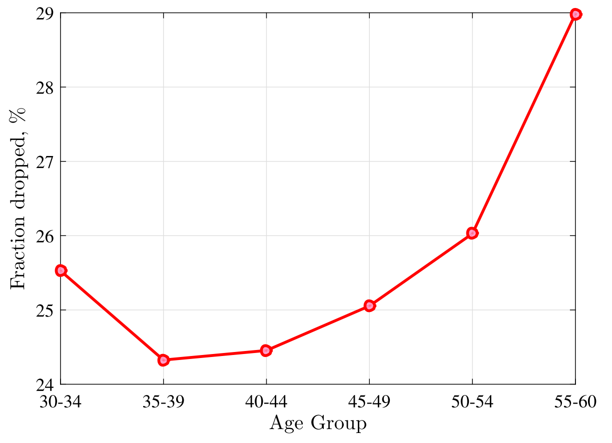


FIGURE C.22.—Fraction of observations dropped by age.

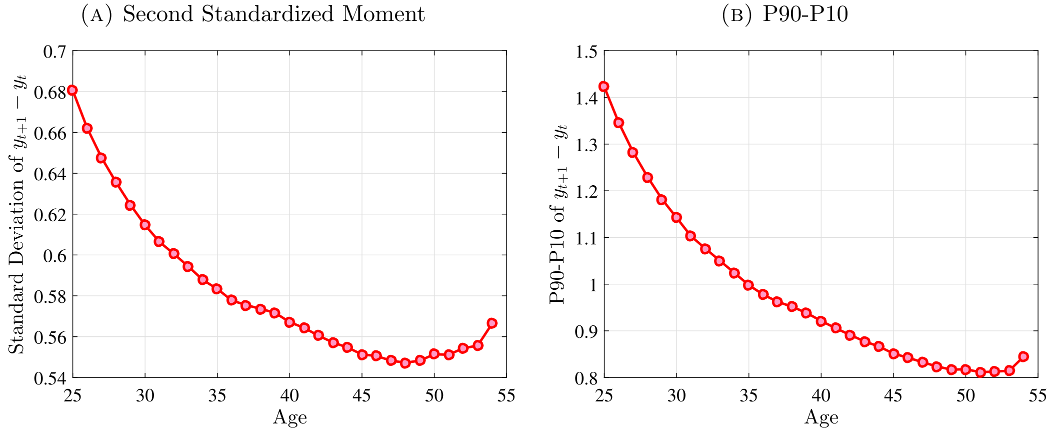


FIGURE C.23.—Dispersion of one-year log earnings growth by age.

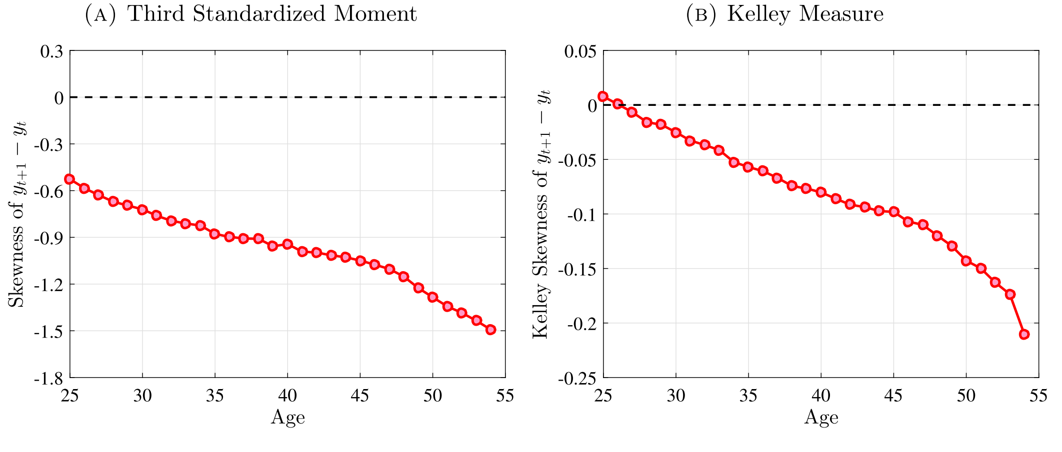


FIGURE C.24.—Skewness of one-year log earnings growth by age.

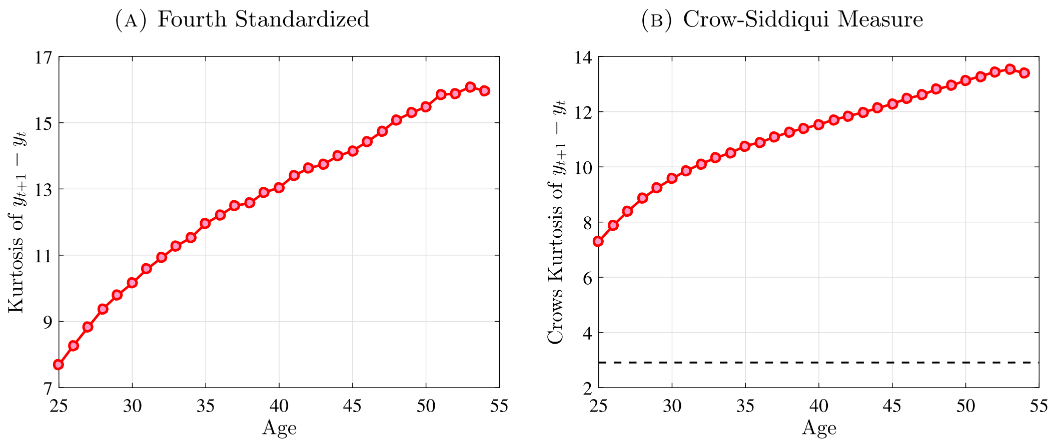


FIGURE C.25.—Kurtosis of one-year log earnings growth by age.

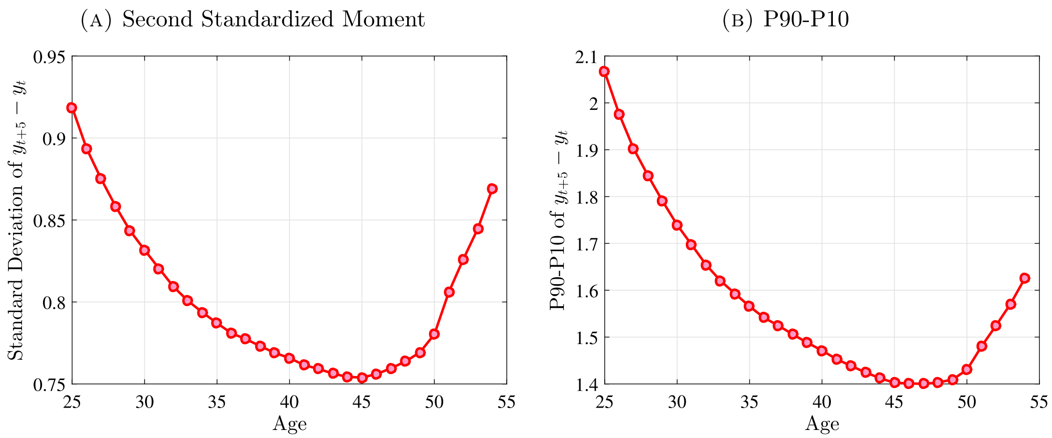


FIGURE C.26.—Dispersion of five-year log earnings growth by age.

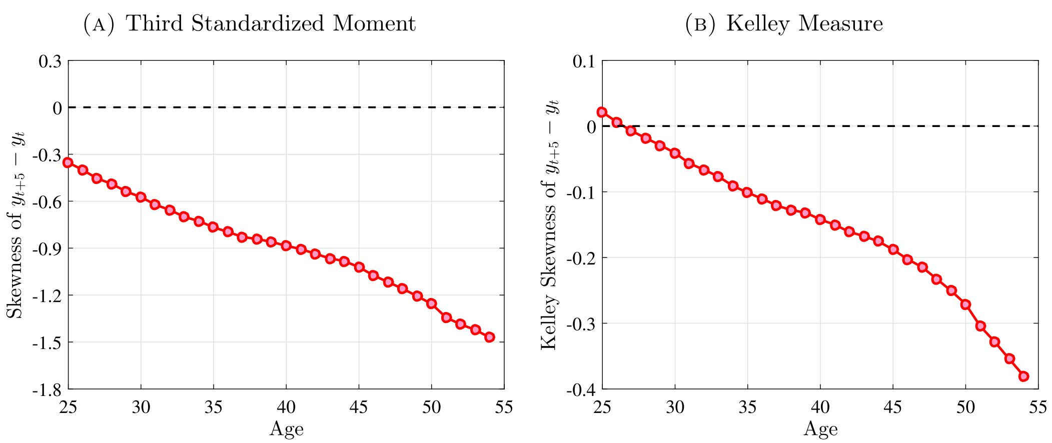


FIGURE C.27.—Skewness of five-year log earnings growth by age.

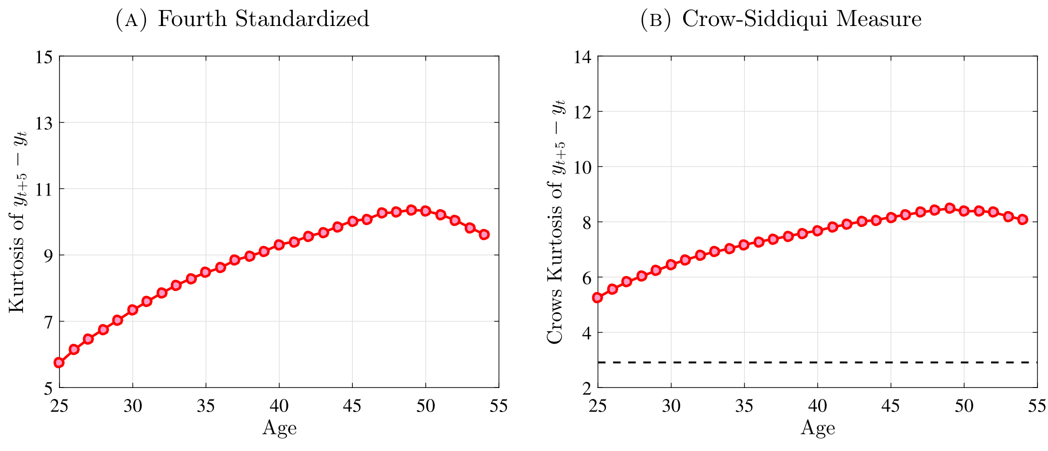


FIGURE C.28.—Kurtosis of five-year log earnings growth by age.

regressing them on a full set of age dummies, controlling for 3 race, 3 education, and 8 region dummies. The three education levels are: less than 11 years (less than high school), 12 years (high school), and more (college dropout, BA degree, or more). We take the maximum grade achieved as the relevant education level of an individual throughout the sample. Race dummies correspond to white, black, and the remaining race and ethnicities. We clean the age variable so that it increases by 2 for each individual across two surveys. We run these regressions separately by year, obtain the residuals, and analyze the change in the residuals between consecutive interviews. We group observations into seven bins, depending on the magnitude of this change.

We utilize different variables in the PSID to identify individuals who experience a change in health, start experiencing disability, or experience time out of work or a job or occupation change. We now describe specifically which variables are used to construct the various measures in Table III.

Bad Health. The head is asked the following question (ER15447 in 1999): “Would you (HEAD) say your health in general is excellent, very good, good, fair, or poor?” We classify someone in bad health if he reports a poor health condition (==5). Transitions into poor health are identified as someone who reported being in excellent, very good, good, or fair health in the previous survey and reports being in poor health in the current survey. This variable is available throughout our sample period.

Disability. The head is asked the following question on disability (ER15451 in 1999): “For work you can do, how much does it limit the amount of work you can do—a lot, somewhat, just a little, or not at all?” We classify someone as having some disability if he reports having an issue that affects his work a lot (==1), somewhat (==3), or just a little (==5). A new disability is coded as someone who did not have such an issue the last time and reports an issue in the current survey. This variable is available throughout our sample period.

Weeks Unemployed. Some individuals report time spent unemployed in units of months (ER21322 in 2003), whereas some report it in units of weeks (ER21320 in 2003). We combine these two variables by taking the maximum reported unemployment duration in weeks. These variables are available starting in 2003.

TABLE C.I
INVOLUNTARY MOVES AND (THE TAILS OF) THE EARNINGS CHANGE DISTRIBUTION

Group $\Delta y \in$	$(-\infty, -1)$	$[-1, -0.25)$	$[-0.25, 0)$	$[0, 0.25)$	$[0.25, 1)$	$(1, \infty)$
Share %	3.8%	14.4%	31.2%	31.1%	16.5%	3.0%
Invol. move %	6.9%	4.5%	3.2%	2.6%	3.6%	4.3%

Weeks out of Labor Force. The PSID asks about head's total weeks out of the labor force in the previous calendar year (ER24087 in 2003). This variable is available throughout our sample period.

Move in Response to Outside Events (Involuntary Reasons). The PSID asks about geographic moves (whether the head changed residence) and the reasons for the move (variable ER13080 in 1999). We classify someone as having moved for involuntary reasons if they report having moved for being evicted, armed services, health reasons, divorce, and health-related retirement. Other observations are classified as nonmovers. This variable is available throughout our sample period. See Table C.I.

Occupation Change. The PSID asks about the head's occupation in the main employer, labeled as job 1 (ER21145 in 2003). This variable is available starting in 2013 and uses the 3-digit occupation code from the 2000 Census of Population and Housing. This variable is available every year since 2003. We code someone as having changed occupations if (i) his occupation in the current year t is different than in the previous survey $t-2$, (ii) he reports having changed jobs (explained below), and (iii) his occupation in the next survey $t+2$ is different from his occupation in year $t-2$. The last condition is used to deal with potential coding errors of occupations prevalent in most survey data.

Job Change. We use the start year of the current main job (job 1) to identify job changes (ER21130 in 2003). If the head reports having started the job in the same year as the survey or the year before, we code him as a job-switcher. This variable is available every year since 2003.

Current Population Survey (CPS)

The CPS is a rotating panel based on addresses. Each address in the survey is interviewed for four consecutive months, then leaves the sample for eight months, and then returns for another four consecutive months. Because of this rotating nature, it is possible to match at most three quarters of respondents across months. Since the survey is based on addresses and does not follow households, if households move, they leave the sample and may be replaced by others that move in to the same address. To have a reliable panel, we match individual records using rotation groups, household identifiers, individual line numbers, race, sex, and age.

The Annual Social and Economic Supplement (ASEC) of the CPS, a supplement to the CPS in March, asks respondents about their earnings and hours and weeks worked during the past calendar year (variables incwage, wkslyr and hrslyr, respectively.) We use data for the period 1968–2013 and focus on males between ages 25 and 55. Similarly to the SSA sample, we impose a minimum earnings threshold that corresponds to working for 13 weeks for 40 hours a week at half the minimum wage. We focus on three measures: annual

TABLE C.II

HIGHER-ORDER MOMENTS OF INCOME CHANGES IN PSID AND CPS[a]

		PSID					
		All		25–39		40–55	
	Gaussian	Earnings	Wages	Earnings	Wages	Earnings	Wages
Skewness	0.0	−0.26	−0.14	−0.17	−0.20	−0.34	−0.09
Kelley Skewness	0.0	−0.02	−0.02	0.03	0.016	−0.06	−0.04
Kurtosis	3.0	12.26	13.65	10.44	9.00	14.01	17.10
Crow Kurtosis	2.91	6.83	5.59	6.33	5.02	7.33	6.11

		CPS					
		All		25–39		40–55	
		Earnings	Wages	Earnings	Wages	Earnings	Wages
Skewness	0.0	−0.15	−0.09	−0.09	−0.023	−0.21	0.004
Kelley Skewness	0.0	−0.02	−0.01	0.002	−0.008	−0.03	−0.017
Kurtosis	3.0	9.29	11.2	9.12	10.60	9.53	12.1
Crow Kurtosis	2.91	7.15	5.93	6.97	5.72	7.29	6.05

[a]Note: Wages are obtained by dividing annual earnings by annual hours in the PSID, and by the weekly wage variable in the CPS.

earnings, average weekly wages, and average hourly wages. We regress each measure on a full set of age dummies and a race dummy (white and nonwhite). We run these regressions separately for each year and education group (college and noncollege), thereby allowing the coefficients on age and race dummies to depend on age and education. We then obtain the residuals from this regression and analyze the changes in the residuals. We use the CPS weights throughout this analysis.

Higher-Order Moments in CPS. In the main text, we reported higher-order moments of two-year changes in earnings and wages by age groups. Table C.II provides similar results from the CPS. The growth measure in the CPS is annual and is therefore not easily comparable to the figures from the PSID. However, the findings are qualitatively similar.

C.8. *The Role of Disability Income*

In this section, we investigate the robustness of our findings to the inclusion of income from Social Security disability benefits (SSDI). This is particularly relevant for thinking about the tails of the distribution of earnings changes. To this end, we link to our data set information about disability benefits from the SSA records. We then define a measure of "total income" as the sum of labor income and annual disability income. Section C.8.1 compares the cross-sectional moments of earnings changes to our baseline and Section C.8.2 does the same for lifetime income growth.

C.8.1. *Cross-Sectional Moments*

Figures C.29–C.34 show several moments of one-year and five-year earnings changes. In each figure and for each age group, we show these moments for labor income and total income separately, where total income is labor income plus disability benefits. For each

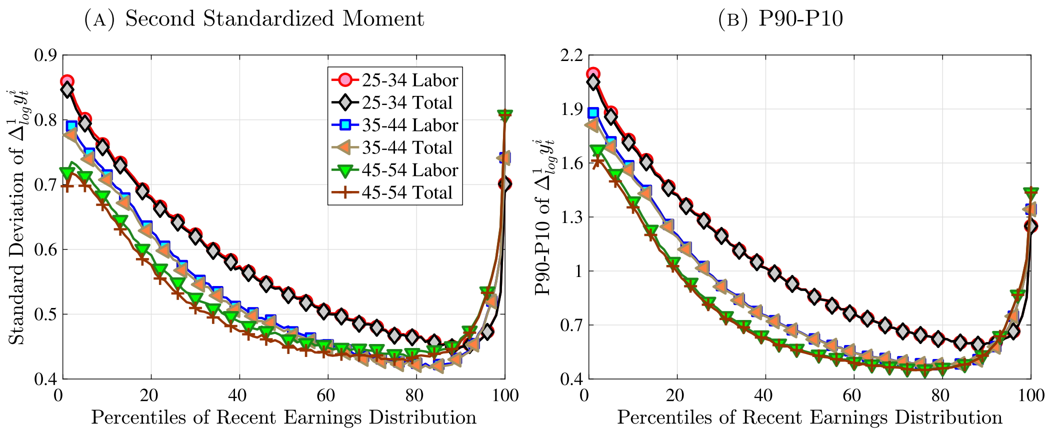


FIGURE C.29.—Dispersion of one-year log earnings growth.

measure of income, recent earnings are recalculated using that measure. Otherwise, these graphs are calculated in a fashion analogous to that in Section 3. The main finding here is that the inclusion of disability income has little effect on cross-sectional moments, even at low levels of earnings.

C.8.2. *Lifetime Income Growth Moments*

Figure C.35 shows growth in average earnings over the life cycle. The left panel shows (log) growth in average earnings between ages 25 and 55, whereas the right panel does the same for 30 and 55. We consider two measures: labor income and total income (including disability benefits). Figure C.35 plots income growth for the two measures against lifetime income. To allow comparability across the two measures, we use labor income to construct each individual's lifetime income and impose the sample selection criteria based on this measure. This allows us to keep the same overall sample as well as the same people in each lifetime income group. The differences in the two series are therefore only due to disability payments. We find that SSDI makes a difference for the income growth of the

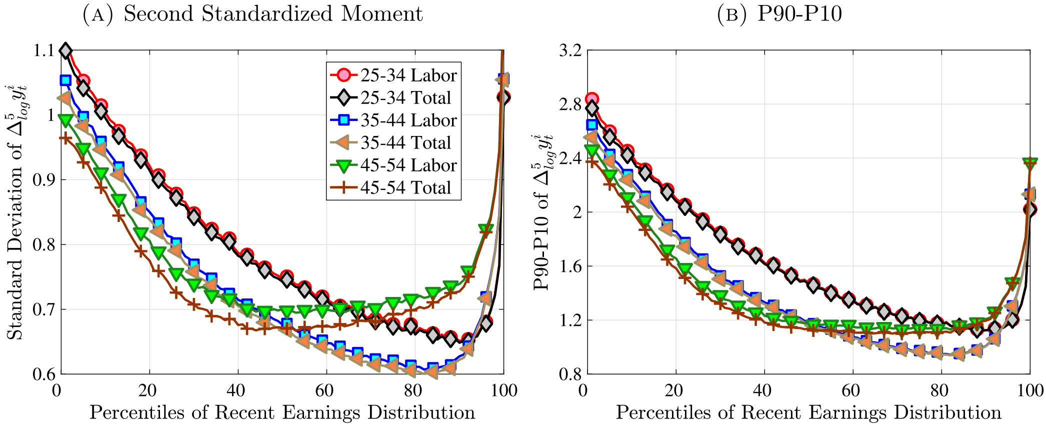


FIGURE C.30.—Dispersion of five-year log earnings growth.

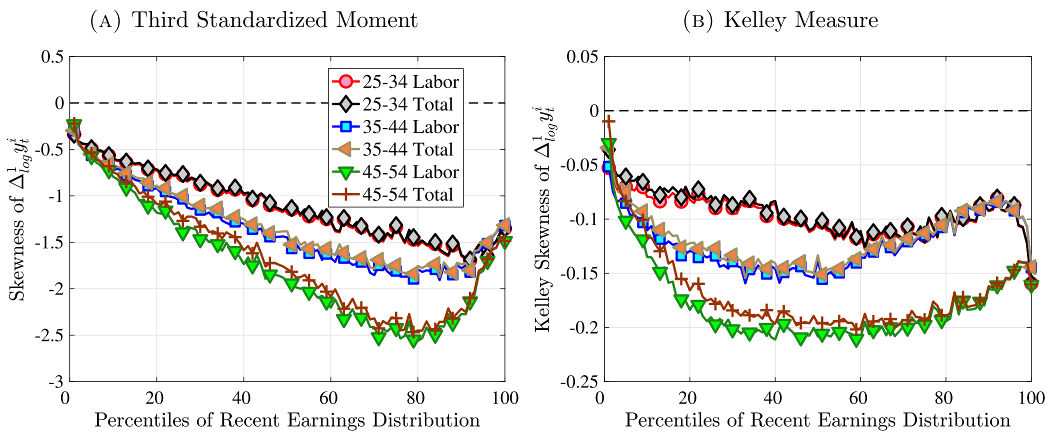


FIGURE C.31.—Skewness of one-year log earnings growth.

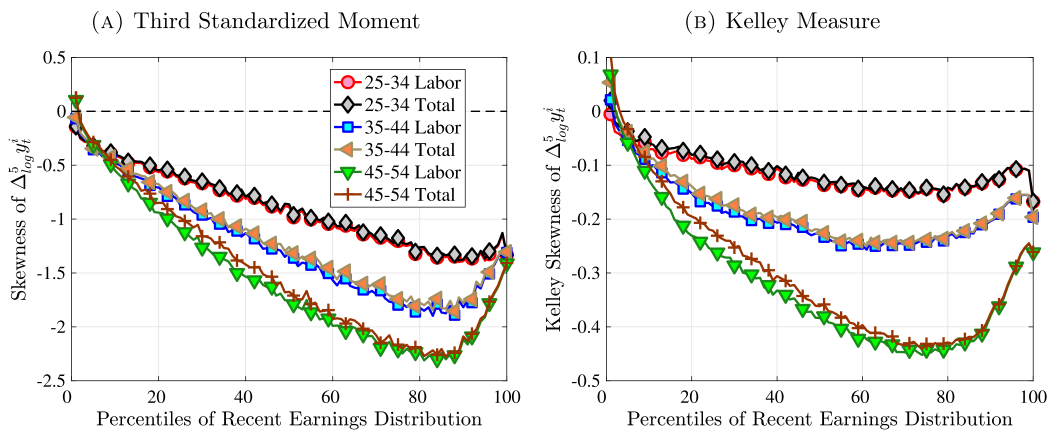


FIGURE C.32.—Skewness of five-year log earnings growth.

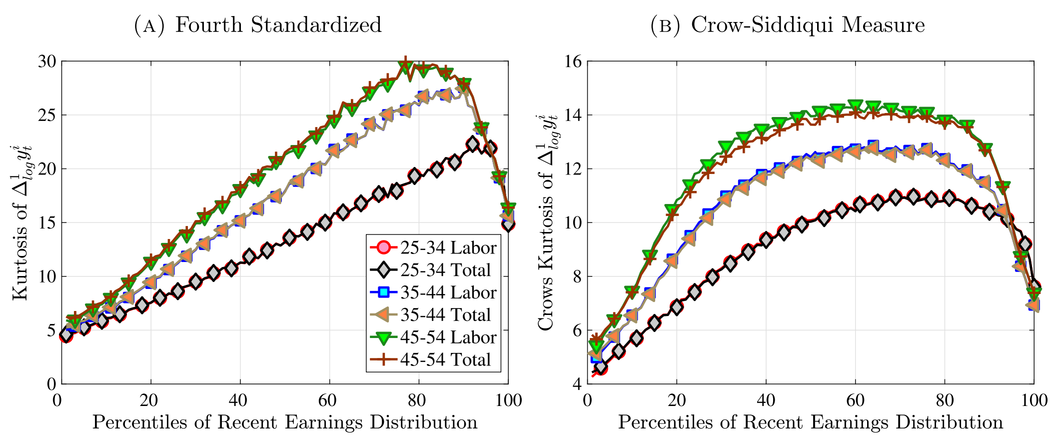


FIGURE C.33.—Kurtosis of one-year log earnings growth.

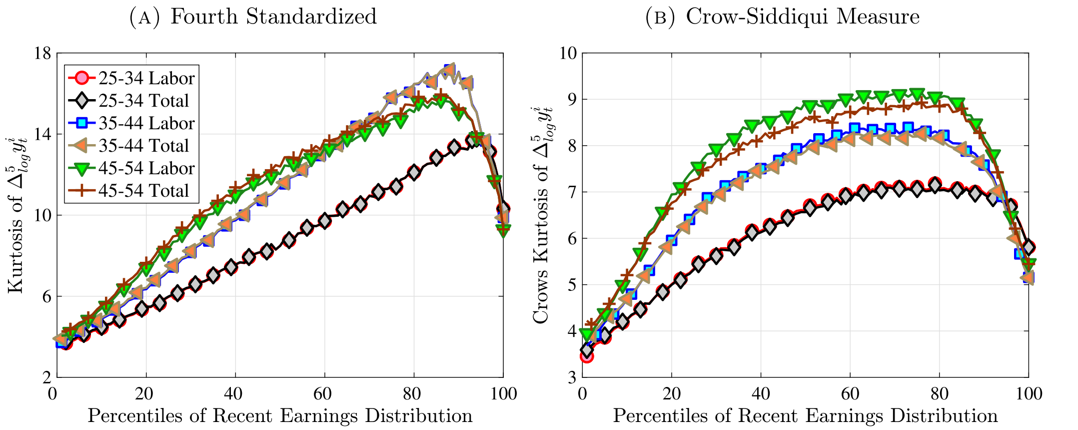


FIGURE C.34.—Kurtosis of five-year log earnings growth.

bottom LE individuals. For the bottom 1%, we find that SSDI can undo around 50% of earnings declines over the lifetime. This contribution declines gradually and vanishes by around the 20th percentile of the LE distribution. This result is not very surprising since bottom LE workers are more likely to be claiming disability benefits.

C.9. *Additional Figures on Lifecycle Patterns of Earnings*

To provide a benchmark for the analysis in Section 5, we estimate the average lifecycle profile of log earnings using a standard pooled regression of log individual earnings on a full set of age and (year-of-birth) cohort dummies using the admissible observations (as defined in Section 2) between 1994 and 2013.[8] The estimated age dummies are plotted as circles in Figure C.36 and represent the average lifecycle profile of log earnings. It has the

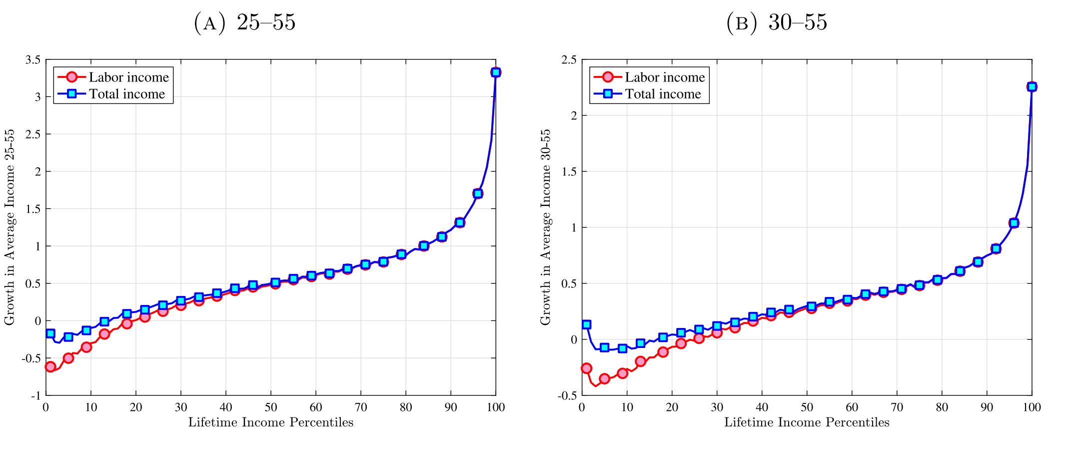


FIGURE C.35.—Lifecycle earnings growth rates, by lifetime earnings group.

[8]This procedure is standard in the literature; see, for example, Deaton and Paxson (1994) and Storesletten, Telmer, and Yaron (2004).

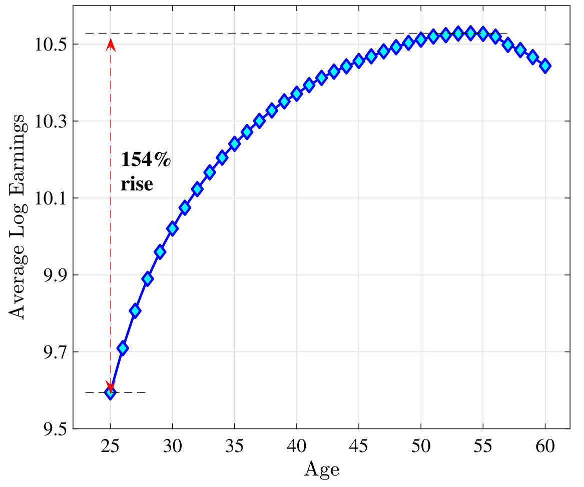


FIGURE C.36.—Lifecycle profile of average log earnings.

usual hump-shaped pattern that peaks around age 50. These age dummies turn out to be indistinguishable from a fourth-order polynomial in age:

$$y_h = -0.0240 + 0.2013 \times h - 0.6799 \times h^2 + 1.2222 \times h^3 + 9.4895 \times h^4,$$

where $h = (\text{age} - 24)/10$. Figure C.37 contains two panels on the distribution of lifecycle growth rates that complement the analysis in Section 5.

C.10. *Concentration of Nonemployment by Lifetime Earnings Group*

In this section, we investigate how concentrated (full-year) nonemployment is. We rank individuals by their lifetime earnings and group them into percentiles. For each lifetime earnings group, we compute what fraction of full-year nonemployment at a given age is

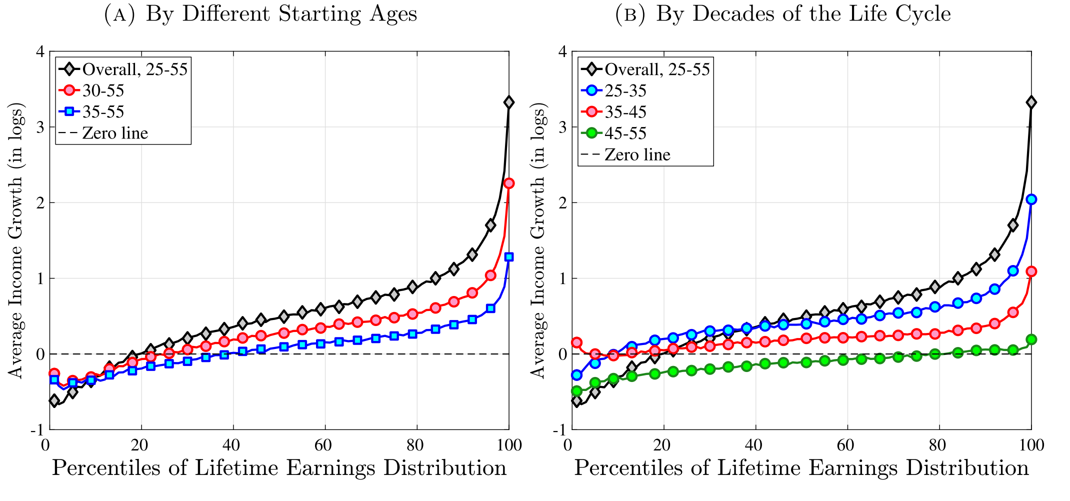


FIGURE C.37.—Log earnings growth over subperiods of life cycle.

(A) Fraction Nonemployed wrt Labor Income (B) Fraction Nonemployed wrt Total Income

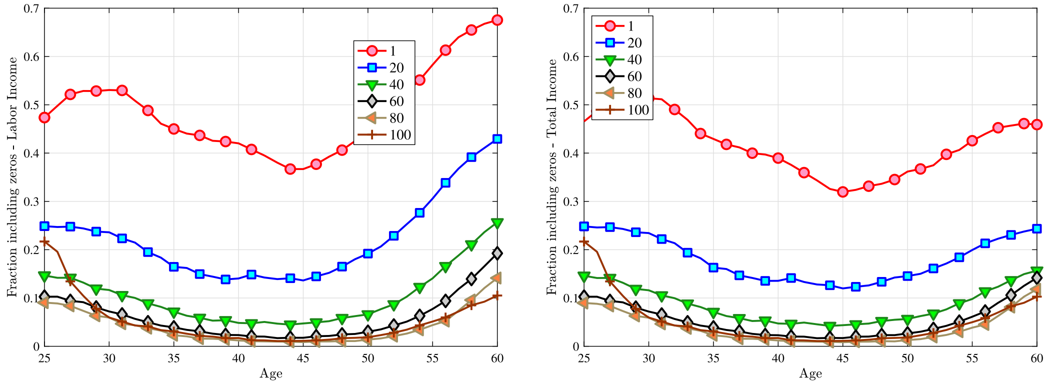


FIGURE C.38.—Nonemployment concentration over the life cycle, by lifetime earnings group.

(A) Average Employment by Decades wrt Lifetime Employment (B) Fraction Nonemployed over the Life Cycle wrt Number of Years Nonemployment

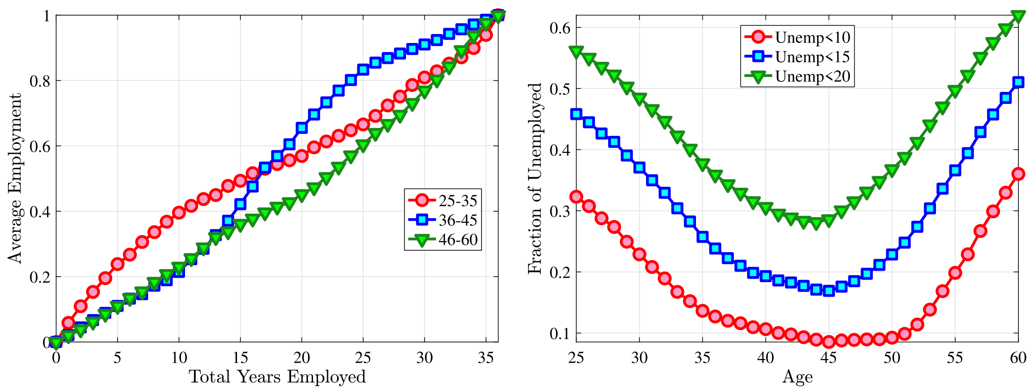


FIGURE C.39.—Employment CDF by age groups.

accounted for by that group. These are shown in Figure C.38. For example, the bottom 1% of the lifetime earnings distribution accounts for around 50% of total nonemployment at ages 25–30.

APPENDIX D: ESTIMATION

In this section, we describe the steps of our estimation procedure of method of simulated moments (MSM) in more detail and provide additional estimation results.

D.1. *Moment Selection and Aggregation*

Accounting for Zeros. Recall that in order to construct the cross-sectional moments of log growth, we have dropped individuals who had very low earnings—below $\overline{Y}_{\min}$—in year t or $t+k$ so as to allow taking logarithms in a sensible manner. Although this approach made sense for documenting empirical facts that are easy to interpret, for the estimation exercise, we would like to also capture the patterns of these "zeros" (or very low earnings observations), given that they clearly contain valuable information. Targeting log growth moments also creates technical issues with the optimization due to (little) jumps in the objective function as workers cycle in and out of employment. To this end, instead of targeting moments of log earnings change, we target moments of arc-percent change, as defined in Section 3. According to this measure, an income change from any positive level to 0 corresponds to an arc-percent change of -2, whereas an income change from 0 to any positive level indicates an arc-percent change of 2.

Aggregating Moments. If we were to match all data points for every RE percentile and every age group, it would yield more than 10,000 moments. Although such an estimation is not infeasible, not much is likely to be gained from such a level of detail, and it would make the diagnostics—that is, judging the performance of the estimation—quite difficult. To avoid this, we aggregate the 100 RE percentiles and the 6 age groups into fewer homogeneous groups. We now describe the details of this aggregation and the resulting list of moments targeted in our estimation.

1. Cross-Sectional Moments of Earnings Changes. To capture the variation in the cross-sectional moments of earnings changes along the age and recent earnings dimensions, we condition the distribution of earnings changes on these variables. For this purpose, we first group workers into 6 age bins (five-year age bins between 25 and 54) and within each age bin into 13 selected groups of RE percentiles in age $t-1$. The RE percentiles are grouped as follows: 1, 2–10, 11–0, 21–30, ..., 81–90, 91–95, 96–99, 100. Thus, we compute the three moments of the distribution of one- and five-year earnings changes for $6 \times 13 = 78$ different groups of workers. We then aggregate the 6 age bins into 3 age groups, A^i_{t-1}, by taking an average of moments within each age group. The first age group is defined as young workers between ages 25 and 34, the second is between ages 35 and 44, and the third age group is defined as workers between the ages of 35 and 54. Consequently, we target three standardized moments (i.e., standard deviation, skewness, and kurtosis) of one- and five-year arc-percent change for three age and 13 recent earnings groups, giving us $3 \times 2 \times 3 \times 13 = 234$ cross-sectional moments. These moments are shown in Figure D.1.

2. Lifecycle Earnings Profile. The second set of moments captures the heterogeneity in log earnings growth over the working life across workers who are in different percentiles of the LE distribution. We target the average dollar earnings at 8 points over the life cycle: ages 25, 30, ..., and 60 for different LE groups. We combine LE percentiles into larger groups to keep the number of moments at a manageable number, yielding 15 groups consisting of percentiles of the LE distribution: 1, 2–5, 6–10, 11–20, 21–30, ..., 81–90, 91–95, 96–97, 98–99, and 100. The total number of moments we target in this set is $8 \times 15 = 120$.

3. Impulse Response Functions. We target average arc-percent changes in earnings over the next k years for $k = 1, 2, 3, 5, 10$ conditional on groups formed by crossing age, recent earnings $\overline{Y}_{t-1}$, and earnings change between $t-1$ and t $\Delta^1_{\text{arc}} Y^i_{t-1}$: $\mathbb{E}[\Delta^{k+1}_{\text{arc}} Y^i_{t-1} \mid \text{age}, \overline{Y}_{t-1}, \Delta^1_{\text{arc}} Y^i_{t-1}]$.[9] In each year, we first group workers into two age bins, denoted by

[9]Notice that, different from the moments we have shown in Section 4, we target earnings growth between $t+k$ and $t-1$. This is because all workers have $\tilde{Y}^i_{t-1} \geq Y_{\min,t-1}$ in $t-1$ by construction of the RE sample. Thus, we can compute the arc-percent growth between $t+k$ and $t-1$ for all workers, which keeps the composition of workers constant in each k.

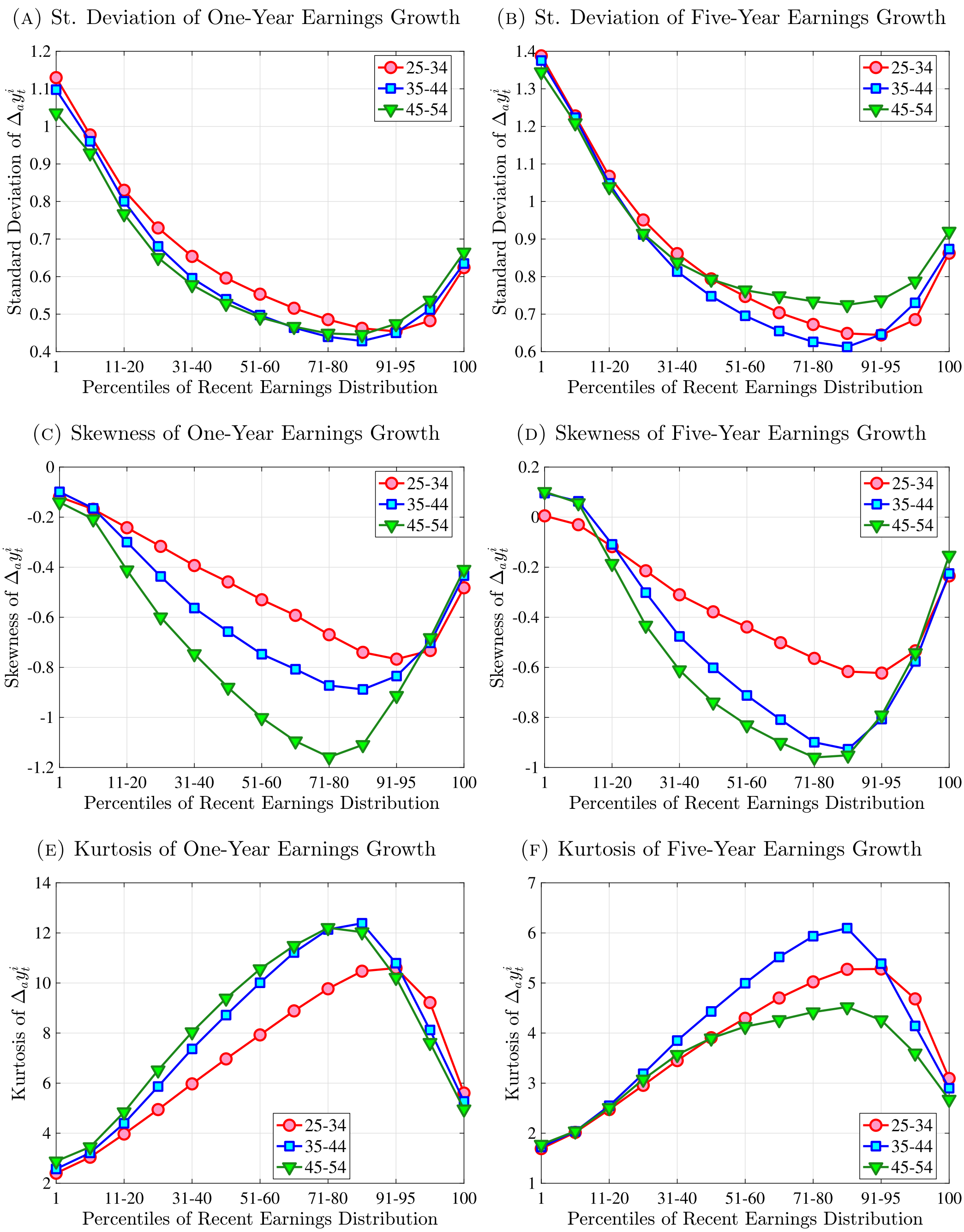


FIGURE D.1.—Standardized cross-sectional moments of arc-percent growth targeted in the estimation.

h: young workers (25–34) and prime-age workers (35–55). Then, within each age group, individuals are ranked into the following RE percentiles, denoted by j: 1–5, 6–10, 11–30, 31–50, 51–70, 71–90, 91–95, 96–100.

Within each age h and RE group j, we then estimate our targets for the persistence of earnings growth. For each k-year expected future earnings growth, we create a piecewise linear function of arc-percent growth between t and $t-1$, that is, $\mathbb{E}_{h,j}[\Delta^{k+1}_{\text{arc}} Y^i_{t-1} \mid \Delta^1_{\text{arc}} Y^i_{t-1}] = f^k_{h,j}(\Delta^1_{\text{arc}} Y^i_{t-1})$. For this purpose, we condition workers in the data into 23 groups with respect to $\Delta^1_{\text{arc}} Y^i_{t-1}$. We first group all workers who are full-year nonemployed in t in the first bin. Then we rank the rest of the workers into the following percentiles: 1–2, 3–5, 6–10, 11–15, 16–20, 21–25, …, 91–95, 96–98, 99–100. Then the piecewise linear function $f^k_{h,j}(\Delta^1_{\text{arc}} Y^i_{t-1})$ for year k is determined by the linear interpolation of 23 data points of average earnings growth between $t-1$ and t, $\mathbb{E}[\Delta^1_{\text{arc}} Y^i_{t-1}]$, and their corresponding k-year future expected growth, $\mathbb{E}[\Delta^k_{\text{arc}} Y^i_t]$.

For the model-simulated data, we group workers into 2×8 age and RE groups—similar to the data moments. But within each age h and RE group j, we now rank workers into 10 $\Delta^1_{\text{arc}} Y^i_{t-1}$ groups defined by the following percentiles: 1–2, 3–5, 6–10, 11–30, 31–50, 51–70, 71–90, 91–95, 96–98, 99–100. In the estimation of income processes, we minimize the distance between 10 different values of $\mathbb{E}^{\text{model}}_{h,j}[\Delta^{k+1}_{\text{arc}} Y^i_{t-1} \mid \Delta^1_{\text{arc}} Y^i_{t-1}]$ (for each age and RE group and k-year expectation) from the model and its corresponding data moment from the piecewise linear function, $f^k_{h,j}(\Delta^1_{\text{arc}} Y^i_{t-1})$. As a result, we have a total of $2 \times 8 \times 5 \times 10 = 800$ moments based on impulse responses.

The impulse response functions targeted in the estimation are plotted in Figures D.2(a)–D.2(d) (to keep the figures similar to their counterparts in Section 4, we plot $\mathbb{E}[\Delta^{k+1}_{\text{arc}} Y^i_{t-1} \mid \Delta^1_{\text{arc}} Y^i_{t-1}] - \mathbb{E}[\Delta^1_{\text{arc}} Y^i_{t-1}]$). More specifically, Figure D.2(a) plots, for prime-age workers with median recent earnings, the mean reversion patterns at various horizons. Figures D.2(b) and D.2(c) do the same for workers at the 90th and 10th percentiles of the recent earnings distribution, respectively. Last, Figure D.2(d) shows the variation of these impulse response functions with recent earnings.

4. Age Profile of Within-Cohort Variance of log Earnings. Although the main focus of this section is on earnings *growth*, the lifecycle evolution of the dispersion of earnings *levels* has been at the center of the incomplete markets literature since the seminal paper of Deaton and Paxson (1994). For completeness and comparability with earlier work, we have estimated the within-cohort variance of log earnings over the life cycle by controlling for cohort dummies in a sample of cross-sectional moments in the data (Figure D.3). In our estimation, we compute this set of moments for a sample with income observations above the minimum income threshold. We have a total of 36 moments based on the variance of log earnings, one for each age.

5. CDF of Employment Over the Life Cycle. We target the distribution of total number of years employed ($\tilde{Y}^i_{t,h} \geq \overline{Y}_{\min,t}$) over the life cycle. In particular, we target the cumulative distribution of total lifetime years employed as shown in Figure 11(b). Thus, in total, there are 35 such moments targeted in our estimation.

In sum, we target a total of $J = 234 + 120 + 800 + 36 + 35 = 1227$ moments in our estimation.[10]

D.2. *2-State Process*

We also estimate a more flexible income process which has *two* AR(1) components, denoted by z_1 and z_2, each subject to innovations from a mixture of two normals with

[10]The full set of moments targeted in the estimation are reported (in Excel format) as part of an online appendix available from the authors' websites.

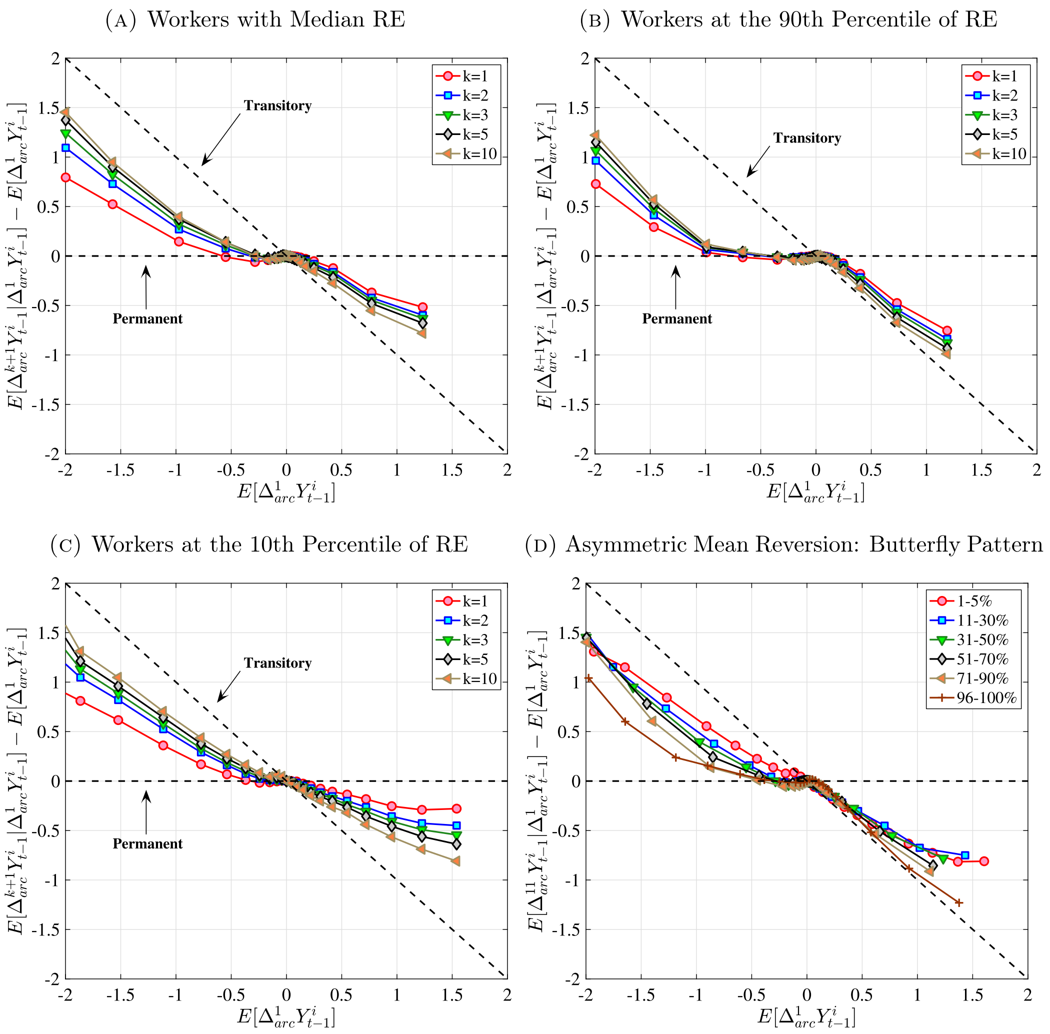


FIGURE D.2.—Impulse response moments targeted in the estimation, prime-age workers.

age- and income-dependent shock probabilities. Here is the full specification where $t = (\text{age} - 24)/10$ denotes normalized age and for $j = 1, 2$:

$$Y_t^i = (1 - \nu_t^i)\exp\big(g(t) + \alpha^i + \beta^i t + z_{1,t}^i + z_{2,t}^i + \varepsilon_t^i\big), \tag{D.1}$$

$$z_{1,t}^i = \rho_1 z_{1,t-1}^i + \eta_{1t}^i, \tag{D.2}$$

$$z_{2,t}^i = \rho_2 z_{2,t-1}^i + \eta_{2t}^i, \tag{D.3}$$

$$\text{Innovations to AR(1):} \quad \eta_{j,t}^i \sim \begin{cases} \mathcal{N}(\mu_{z_j,1}, \sigma_{z,1}) & \text{with pr. } p_{z_j,t}, \\ \mathcal{N}(\mu_{z_j,2}, \sigma_{z,2}) & \text{with pr. } 1 - p_{z_j,t}, \end{cases} \tag{D.4}$$

$$\text{Initial value of AR(1) process:} \quad z_{j0}^i \sim \mathcal{N}(0, \sigma_{j0}), \quad j = 1, 2, \tag{D.5}$$

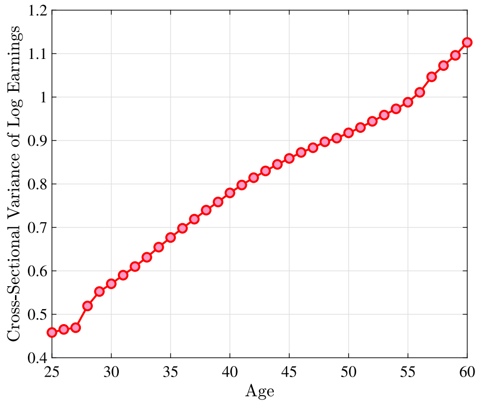


FIGURE D.3.—Within-cohort variance of log earnings.

$$\text{Nonemployment shocks:}\quad \text{equation (7)}, \tag{D.6}$$

$$\text{Transitory shock:}\quad \text{equation (6)}. \tag{D.7}$$

Each AR(1) component, z_1 and z_2, receives a shock drawn from a mixture of two Gaussian distributions as in our benchmark specification. We again normalize the mean of innovations to the persistent components to zero; that is, $\mu_{z_j,1}p_{z_j} + \mu_{z_j,2}(1 - p_{z_j}) = 0$.[11] We also allow for heterogeneity in the initial conditions of the persistent processes, $z^i_{1,0}$ and $z^i_{2,0}$, given in equation (D.6). Since the specifications of z_1 and z_2 are the same so far, we need an identifying assumption to distinguish between the two, so, without loss of generality, we impose $\rho_1 < \rho_2$.

The age and income dependence of moments is captured by allowing the mixture probabilities to depend on age and the sum of persistent components $(z_1 + z_2)$:[12]

$$p^i_{z_j,t} = \frac{\exp(\xi^i_{j,t-1})}{1+\exp(\xi^i_{j,t-1})}, \tag{D.8}$$
$$\xi^i_{jt} = a_{z_j} + b_{z_j} \times t + c_{z_j} \times (z^i_{1,t} + z^i_{2,t}) + d_{z_j} \times (z^i_{1,t} + z^i_{2,t}) \times t,$$

for $j = 1, 2$. The equation for $p_{\nu t}$ is the same as (D.8) but $\xi^i_{j,t-1}$ is replaced with ξ^i_{jt}. This completes the description of the 2-state process.

This 2-state process provides a significantly better fit to the targeted moments and matches top income inequality as well as the income variation in nonemployment risk (Figures D.4(b) and D.4(a)). We find that the two AR(1) components are quite different from each other, especially in terms of their persistence with $\rho_2 = 0.98$ versus $\rho_1 = 0.79$ (Table D.I). We report the probability of drawing a nonemployment shock for various age and RE percentile groups for workers who satisfy the conditions of the base sample in Table D.II. The composition of large negative shocks changes from (hard-to-insure) more

[11]We do not have to make the identification assumption of $\mu_{z_j,1} < 0$ as we did for the benchmark process, because the first Gaussian, $\mathcal{N}(\mu_{z_j,1}, \sigma_{z,1})$, is already different than the second one, $\mathcal{N}(\mu_{z_j,2}, \sigma_{z,2})$, by having a mean $\mu_{z_j,1}$ constant over age and income, whereas $\mu_{z_j,2}$ varies by income and age. The latter is because p_{z_j} is a function of persistent components and age.

[12]We have also considered an alternative specification where the innovation variances are functions of earnings and age. After extensive experimentation, we have found it to perform poorly.

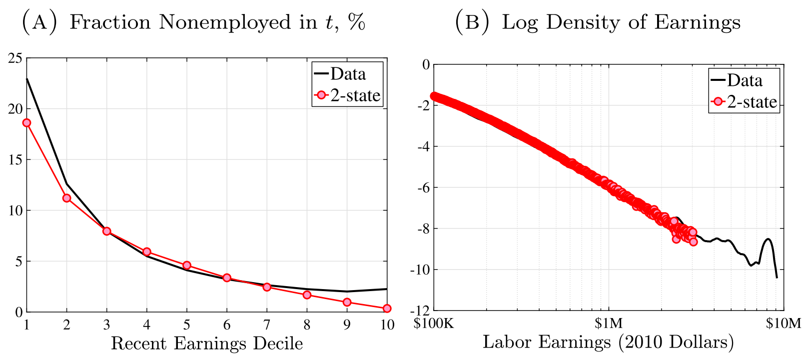


FIGURE D.4.—Model fit: nontargeted statistics.

persistent innovations to the less persistent ones over the life cycle. The probability of receiving at least one large shock to one of the two AR(1) components or a nonemployment shock in a given year is declining in recent earnings, ranging from 29% at the low end to 9% for individuals above the 90th percentile. Finally, the age and income variation of nonemployment risk in the 2-state specification is qualitatively similar to that in the benchmark process.

TABLE D.I

PARAMETER ESTIMATES FOR 2-STATE BENCHMARK PROCESS[a]

Specification		*Parameters*	est.	se.	*Parameters*	est.	se.
AR(1) *Component*	2 mixtures	*Persistent Components*			*Normal Mixture Probability*		
↪*Probability age/inc.*	yes/yes	ρ_1	0.791	0.0006	$a_{z_1} \times 1$	-1.927	0.0052
↪*Probability age/inc.*	yes/yes	ρ_2	0.976	0.0006	$b_{z_1} \times t$	0.778	0.0033
Nonemployment shocks	yes	$\mu_{\eta_1,1}$	-0.393	0.0013	$c_{z_1} \times z_{t-1}$	-1.432	0.0040
↪*Probability age/inc.*	yes/yes	$\mu_{\eta_2,1}$	-0.215	0.0007	$d_{z_1} \times t \times z_{t-1}$	-1.445	0.0044
Transitory Shocks	mix	$\sigma_{\eta_1,1}$	0.561	0.0012	$a_{z_2} \times 1$	-0.099	0.0049
HIP	yes	$\sigma_{\eta_1,2}$	0.078	0.0005	$b_{z_2} \times t$	-0.915	0.0028
		$\sigma_{\eta_2,1}$	0.591	0.0007	$c_{z_2} \times z_{t-1}$	-1.122	0.0027
		$\sigma_{\eta_2,2}$	0.002	0.0006	$d_{z_2} \times t \times z_{t-1}$	0.632	0.0018
Objective value	19.59	$\sigma_{z_1,0}$	0.200	0.0008			
Decomposition:		$\sigma_{z_2,0}$	0.693	0.0007	*Nonemployment Shocks*		
(i) Standard deviation	5.05	*Transitory Shocks*			λ	0.001	0.0004
(ii) Skewness	6.93	prob_ε	7.8%	0.0002	$a_\nu \times 1$	-2.992	0.0043
(iii) Kurtosis	4.73	$\mu_{\varepsilon,1}$	0.467	0.0011	$b_\nu \times t$	-1.036	0.0033
(iv) Impulse response-short	6.22	$\sigma_{\varepsilon,1}$	0.420	0.0010	$c_\nu \times z_t$	-3.391	0.0040
(v) Impulse response-long	12.70	$\sigma_{\varepsilon,2}$	0.020	0.0004	$d_\nu \times t \times z_t$	-2.120	0.0043
(vi) Lifetime inc. growth	6.98	*Individual Fixed Effect*			*Deterministic Lifecycle Profile*		
(vii) Within cohort ineq.	2.28	σ_α	0.274	0.0008	$a_0 \times 1$	2.492	0.0013
(viii) Nonemployment CDF	5.83	σ_β	0.160	0.0002	$a_1 \times t$	0.600	0.0011
		$\text{corr}_{\alpha\beta}$	0.826	0.0010	$a_2 \times t^2$	-0.135	0.0003

[a]Note: We define the deterministic lifecycle profile as a quadratic function of t, $g(t) = a_0 + a_1 t + a_2 t^2$, where $t = (\text{age} - 24)/10$. $y_t = z_t$ for the 1-state income process with one AR(1) component and $y_t = z_1 + z_2$ for 2-state income process with two AR(1) components.

TABLE D.II

MIXTURE PROBABILITIES FOR 2-STATE PROCESS[a]

	Age groups			RE (Percentile) groups				
	25–34	35–49	45–60	1–10	21–30	41–60	71–80	91–100
p_{z_1} ($\rho_{z_1} = 0.79$)	0.122	0.140	0.167	0.399	0.204	0.114	0.053	0.013
p_{z_2} ($\rho_{z_2} = 0.98$)	0.251	0.162	0.111	0.205	0.184	0.174	0.163	0.158
p_ν (nonemp.)	0.067	0.054	0.052	0.188	0.080	0.040	0.017	0.004
Pr (any large shock)	0.337	0.279	0.266	0.531	0.353	0.270	0.208	0.169

[a]Notes: This table reports how the probabilities of innovations with large standard deviations vary by age and past income. In particular, the first row reports the probability of drawing innovations to the z_1 persistent component from the first normal distribution, $z_{1,1} \sim \mathcal{N}(-0.393, 0.561)$. And similarly, the second row presents the probability of drawing innovations to the z_2 persistent component from the first normal distribution, $z_{2,1} \sim \mathcal{N}(-0.215, 0.591)$. The last row reports the probability of any one of the events happening in the first three rows.

D.3. *Numerical Method for Estimation*

Objective Function. Let d_j for $j = 1, \dots, J = 1227$ denote a generic empirical moment, and let $\tilde{d}_j(\theta)$ be the corresponding model moment that is simulated for a given vector of earnings process parameters, θ. We simulate the entire earnings histories of 100,000 individuals who enter the labor market at age 25 and work until age 60. When computing the model moments, we apply precisely the same sample selection criteria and employ the same methodology with the simulated data as we did with the actual data. To deal with potential issues that could arise from the large variation in the scales of the moments, we minimize the *scaled* arc-percent deviation between each data target and the corresponding simulated model moment. For each moment j, define

$$m_j(\theta) = \frac{\tilde{d}_j(\theta) - d_j}{0.5\big(\big|\tilde{d}_j(\theta)\big| + |d_j|\big) + \psi_j}, \tag{D.9}$$

where $\psi_j > 0$ is an adjustment factor. When $\psi_j = 0$ and d_j is positive, m_j is simply the (arc-) percentage deviation between data and model moments. This measure becomes problematic when the data moment is very close to zero, which is not unusual (e.g., impulse response of arc-percent earnings changes close to zero). To account for this, we choose ψ_j to be equal to the 10th percentile of the distribution of the absolute value of the moments in a given set. The MSM estimator is

$$\hat{\theta} = \arg\min_{\theta} \boldsymbol{m}(\theta)' W \boldsymbol{m}(\theta), \tag{D.10}$$

where $\boldsymbol{m}(\theta)$ is a column vector in which all moment conditions are stacked, that is,

$$\boldsymbol{m}(\theta) = \big[m_1(\theta), \dots, m_J(\theta)\big]'.$$

We choose a weighting matrix that corresponds to essentially first averaging the moments within each of the seven sets, and then assigning equal weight (1/7) to each set of moments. For example, each of the 117 cross-sectional moments (standard deviation, skewness, kurtosis) of one-year earnings growth receives a weight of $1/(7 \times 117)$, each of the 480 short-term impulse response moments receives a weight of $1/(7 \times 480)$, and so on. Recall again that the seven sets of moments are as follows: (i) the standard deviation, skewness, and kurtosis of one-year and (ii) five-year earnings growth; (iii) impulse response moments over short (at one-, two-, and three-year) horizons and (iv) long (at fiveand ten-year) horizons; (v) average earnings of each LE group over the life cycle; (vi) the cumulative distribution of nonemployment; and (vii) the age profile of the within-cohort variance of log earnings.

Numerical Method. The global stage is a multi-start algorithm where candidate parameter vectors are uniform Sobol (quasi-random) points. We typically take about 250,000 initial Sobol points for pre-testing and select the best 1000 (i.e., ranked by objective value) for the multiple restart procedure depending on the number of parameters to be estimated. For processes with a large number of parameters to be estimated (e.g., the benchmark process or the 2-state process), we also tried using 300,000 initial Sobol points and used the best 2000 of them. We found this wider search for parameter values to be inconsequential for our estimates. The local minimization stage is performed with a mixture of Nelder-Mead's downhill simplex algorithm (which is slow but performs well on difficult objectives) and the DFNLS algorithm of Zhang, Conn, and Scheinberg (2010), which is much faster but has a higher tendency to be stuck at local minima. We have found that the combination balances speed with reliability and provides good results.

D.4. *Model Selection*

Clearly, income processes with more parameters deliver a better fit to the data. To what extent should we prefer the richer parameterized processes in Table IV? To answer this question, we now implement a procedure for model selection to the specifications in Table IV. Specifically, we carry out two tests.

Test 1. The first procedure tests the null hypothesis that a given specification is the true data-generating process. It does so, as in existing specification tests in the literature, by using the asymptotic distribution of the objective value—the J-statistic in the GMM context. Our objective value is *different* than the traditional J-statistic, since we do not use the efficient weighting matrix, and therefore it does not follow the chi-squared distribution. Therefore, we first derive analytically the asymptotic distribution of our objective value, which we label as the pseudo-J statistic.

Let W denote the weighting matrix and $\boldsymbol{m}(\theta)$ the moment conditions defined by $\tilde{d}(\theta)$ and d (equation (D.9)) for sample size N. Let the matrix L be the Cholesky decomposition of the variance-covariance matrix of moment conditions, $S = \mathbb{E}\boldsymbol{m}\boldsymbol{m}'$ so that $LL' = S$:[13]

$$\begin{aligned} \text{pseudo-}J_n &= N\boldsymbol{m}'W\boldsymbol{m} && \text{(D.11)} \\ &= \left(\sqrt{N}L^{-1}\boldsymbol{m}\right)'L'WL\left(\sqrt{N}L^{-1}\boldsymbol{m}\right) \\ &\sim z'L'WLz, \quad z \sim \mathcal{N}(0, I). && \text{(D.12)} \end{aligned}$$

The last line holds because $\sqrt{N}L^{-1}\boldsymbol{m} \to_d \mathcal{N}(0, I)$ (per the central limit theorem). (If W is the efficient weighting matrix, that is, $W = [\mathbb{E}\boldsymbol{m}\boldsymbol{m}']^{-1}$, equation (D.12) boils down to the commonly used chi-squared J-statistic in Hansen (1982).)

[13] If S is positive semi-definite instead of positive definite, such a decomposition can be obtained from the LDL decomposition, which exists for semi-definite matrices. To see this, let $\tilde{L}$ and D be such that $\tilde{L}D\tilde{L}' = S$ and define $L = \tilde{L}\sqrt{D}$, where $\sqrt{D}$ is the diagonal matrix containing the square root of the diagonal elements of D.

(A) Test 1: Distribution of Pseudo J-statistic, F_1 (B) Test 2: The Distribution of Objective Values, F_2

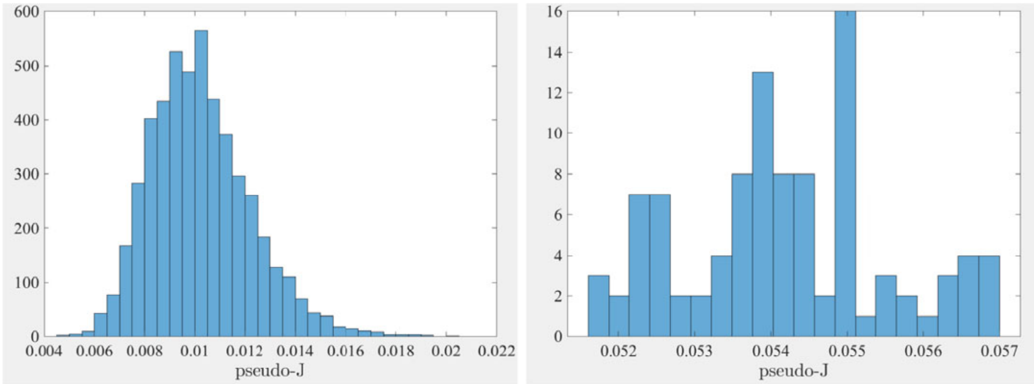


FIGURE D.5.—Distribution of test statistics.

To calculate the distribution of this statistic, we take 5000 draws of the moment conditions ($\boldsymbol{m}$) from the benchmark process by repeatedly simulating data with different seeds of random variables. We use these draws to compute $S = \mathbb{E}\boldsymbol{m}\boldsymbol{m}'$, which is then used to construct L. Then, we simulate random draws from a standard normal distribution, compute pseudo-J for each draw using (D.12), and use these to obtain the distribution of pseudo-J (see Figure D.5(a)). Let F_1 denote the CDF of this distribution.

Our test computes the probability that the pseudo-J obtained from a given specification, denoted as ζ, comes from this distribution; that is, $1 - F_1(\zeta)$. We apply this test to the eight specifications in the main part of the draft and, as explained in Section 6.2, reject all of them (see the bottom panel of Table IV).

Test 2. We develop a second procedure that tests a specification against a benchmark (in this case, column (8) in Table IV). First, we obtain the distribution of the objective values that can be attained from the 1-state distribution via Monte Carlo methods. Specifically, we draw 100 seeds of random variables and estimate our benchmark process by running a local minimization around the current estimates. We then use these objective values to construct the nonparametric distribution denoted by F_2 (see Figure D.5(b)). Our test compares the objective value of the specification at hand (ζ) against this distribution and reports $1 - F_2(\zeta)$ (see the bottom panel of Table IV).

To sum up, our investigation reveals that the benchmark process offers the best fit to the data in a statistical sense. The data reject the hypothesis that the simpler versions analyzed in this paper can provide a similar fit.

Bootstrap Standard Errors. The last column in Table IV reports the standard errors of our benchmark process using a parametric bootstrap. In calculating the bootstrap standard errors, we first simulate data using the parameter estimates reported in Table IV and create moments from simulated data. We then run the estimation for 100 different seeds of random variables by targeting these moments obtained in the previous step. For each seed of random variables, we run the estimation once by employing a simplex algorithm around the original parameter estimates. We compute the standard errors using the resulting 100 parameter vectors.

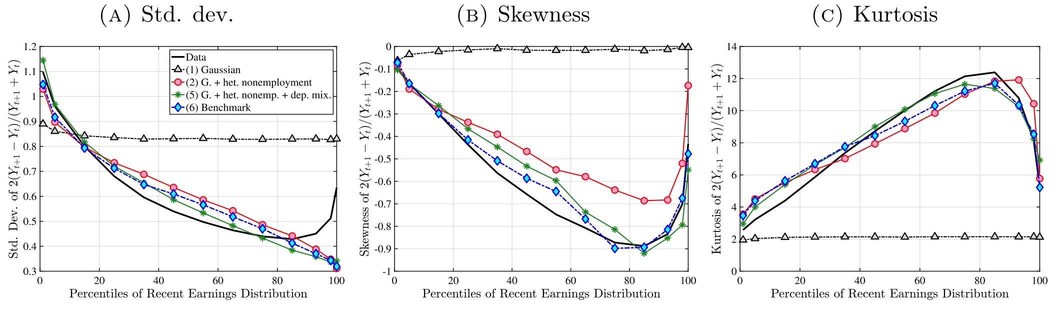


FIGURE D.6.—Model fit to cross-sectional moments of one-year earnings changes.

D.5. *Additional Estimation Results*

This section contains estimation results not reported in the main text. We first report the estimates of additional specifications. Then, for all of the estimated income processes, we report some of the parameters that were not reported in Table IV. A comprehensive set of parameters for all income processes are available online for download as an Excel file on authors' websites.

Model Fit: Additional Figures. Figure D.6 plots the fit on cross-sectional moments of one-year earnings changes against recent earnings, averaging over the life cycle. Figures D.7 and D.8 show how the estimated models fit the lifecycle variation in the cross-sectional moments of one- and five-year earnings changes (averaging over recent earnings). Figures D.9 and D.10 show the fit on the variation by both recent earnings *and* age.

Deviations of Estimated Models From Targeted Values. In the main text, we compared the model counterparts of targeted moments to the data. In this section, we show how the fit looks through the lens of our objective function in (D.9). Figure D.11 shows these for several key moments. More specifically, we plot equation (D.9) for each set of moments, with the exception of income growth moments. Recall that in our estimation, we target the *levels* of income at various ages of different LE percentiles and not the lifecycle growth rates.

Models (3) and (4)'s Fit to the Data. Figures D.12 and D.13 show how models (3) and (4)—that were omitted in the main text—fit selected moments of the data.

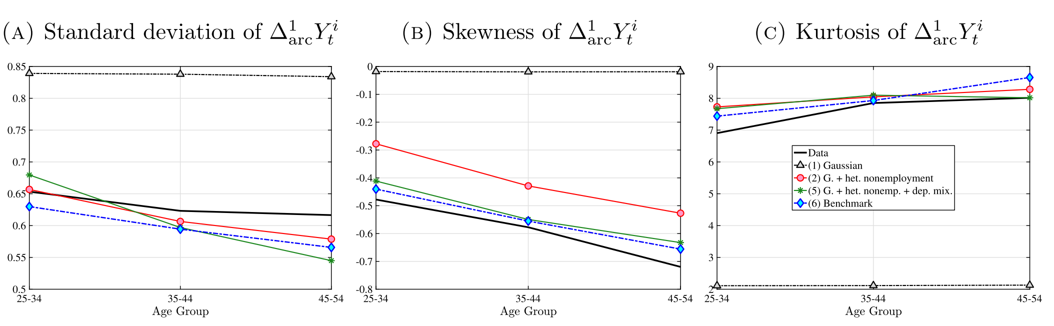


FIGURE D.7.—Fit on cross-sectional moments of one-year earnings changes by age.

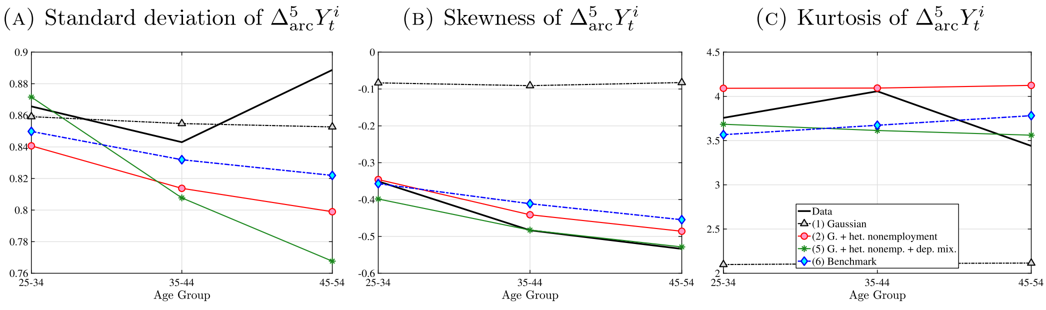


FIGURE D.8.—Fit on cross-sectional moments of five-year earnings changes by age.

Additional Parameter Estimates. Table D.III contains several parameters that were not reported in the main text in Table IV due to space constraints. These parameters include deterministic lifecycle profiles and the coefficients on age and income in the probability functions. Since it is difficult to interpret the magnitudes of the coefficients on shock probabilities, in Figure D.14 we show the 3D figure for the estimated relationship between the nonemployment shock probability and age and the persistent component for the benchmark specification.

Results for Uniform Nonemployment Risk and Non-Gaussian Persistent Shocks. Table D.IV shows estimates of two specifications that appeared in the previous version of the paper. Model (1b) is an intermediate case between Models (1) and (2): It adds *uniform* nonemployment risk to the Gaussian process in (1). The estimates from this model implies that 2.1% of workers are hit with a nonemployment shock each year, and about 42% of those experience full-year nonemployment. The nonemployment shocks soaks up some of the transitory variation in earnings, especially in the tails, in turn reducing the estimated standard deviation of ε relative to Model (1). The improvement in the objective value is quite limited (73.39 versus 74.87), mainly because this model manages to generate some excess kurtosis but very little negative skewness, and it largely misses the age and income variation in the moments. Furthermore, the estimated persistence is even higher ($\rho = 1.015$) than in Model (1), moving the model further away from stationarity. It also implies an unusually large initial heterogeneity ($\sigma_\alpha = 1.26$). As a result, the fit deteriorates slightly for the impulse response and lifecycle income growth moments.

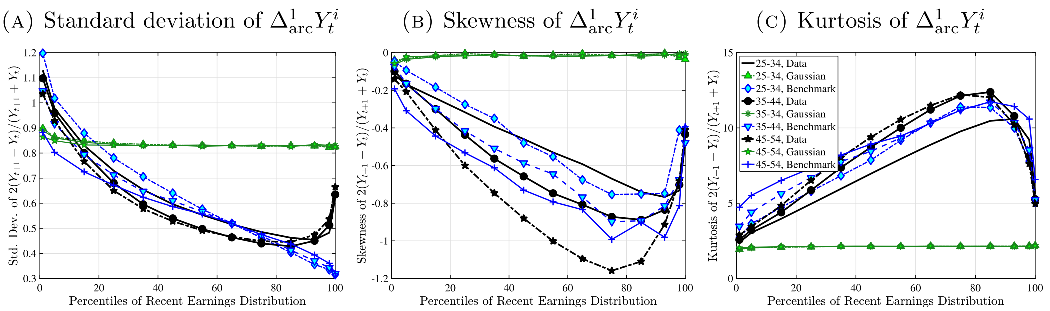


FIGURE D.9.—Cross-sectional moments of one-year earnings changes by age and recent earnings: benchmark, gaussian, and data.

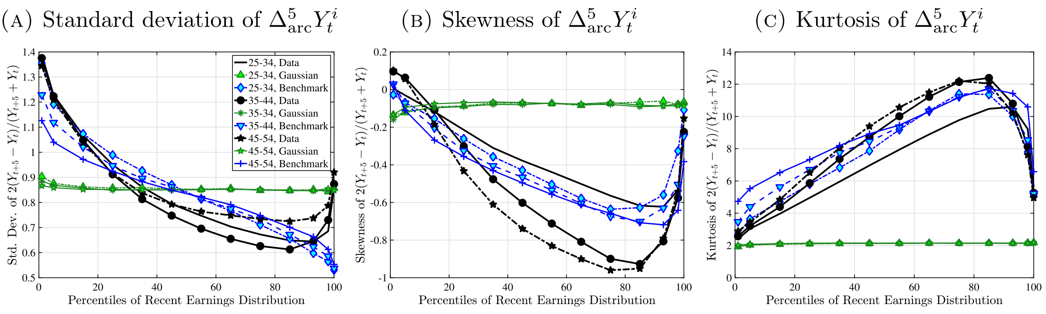


FIGURE D.10.—Cross-sectional moments of five-year earnings changes by age and recent earnings: benchmark, gaussian, and data.

Model (3b) investigates the relative importance of having a mixture component in only persistent innovations. In that sense, it serves as a bridge between models (1) and (3), which features normal mixtures for both transitory and persistent shocks. The objective value falls from 74.9 to 62.1, about two thirds of the improvement of Model (3) relative to (1), indicating that the data demand non-Gaussian features in persistent innovations more than in transitory ones.

Results for Additional Specifications. Table D.V presents estimates from an earlier version of the paper that used a slightly different weighting matrix in the estimation. In particular, the weighting matrix used in this table first assigns 15% relative weight to the employment CDF moments. The rest of the moments share the remaining 85% weight according to the following scheme: the cross-sectional moments (standard deviation, skewness, kurtosis) collectively receive a relative weight of 35%, the lifecycle earnings growth moments and impulse response moments each receive a weight of 25%, and the variance of log earnings by age receives a weight of 15%.

The columns 7 to 15—report the estimates of eight different specifications that are not reported in the main text. Columns (7), (8), and (9) are different versions of Column (5) of Table IV. Namely, we model nonemployment probability as a logistic function of a number of combinations of individual fixed effect α and persistent component z (similar to equation (8)). In particular, in column (7), nonemployment probability is a quadratic function of α. In column (8), nonemployment probability p_ν is assumed to be a linear function of $\alpha + z$ and age and their interaction. In column (9), nonemployment probability depends linearly on α and z and their interaction.

Columns (10) and (11) are similar to our benchmark specification. Again they only differ in how the nonemployment shock probability is modeled. In column (10), p_ν is a quadratic function of z. In column (11), p_ν depends on z, z^2, and age, and the interaction of z and age.

In column (12), we introduce variance heterogeneity to the 1-state benchmark process. In particular, we allow the variance of each innovation from $\mathcal{N}(\mu_{z,1}, \sigma^i_{z,1})$ to the persistent component be individual-specific, with a lognormal distribution with mean $\overline{\sigma}_{z,1}$ and a standard deviation proportional to $\widetilde{\sigma}_{z,1}$, that is, $\log(\sigma^i_{z,1}) \sim \mathcal{N}(\log \overline{\sigma}_{z,1} - \frac{\widetilde{\sigma}^2_{z,1}}{2}, \widetilde{\sigma}_{z,1})$.

In the next income process (column (13)), in the 1-state benchmark process we incorporate age and income dependence into the mixture probability in innovations to the persistent component similar to equation (D.8).

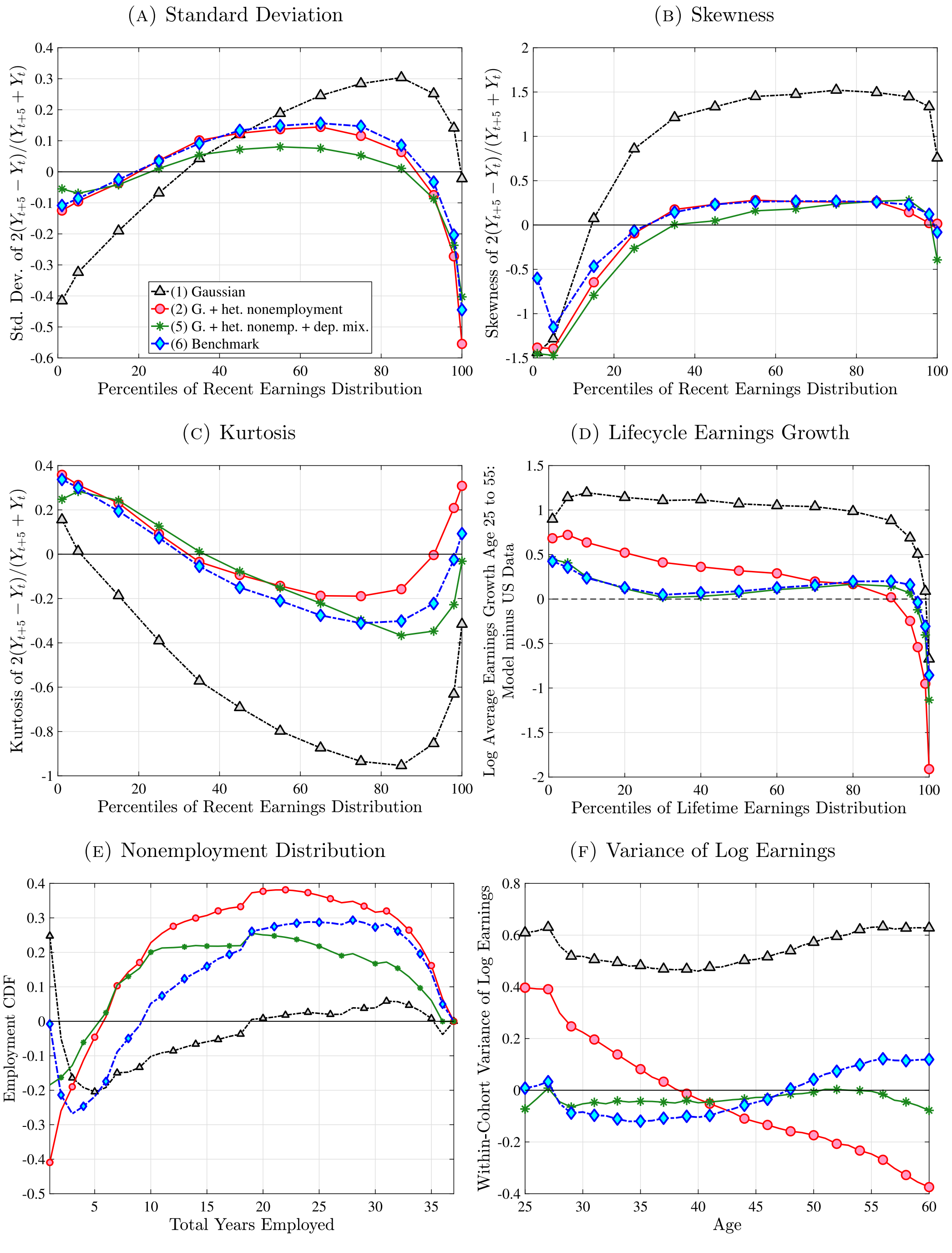


FIGURE D.11.—Deviations of model moments from data moments.

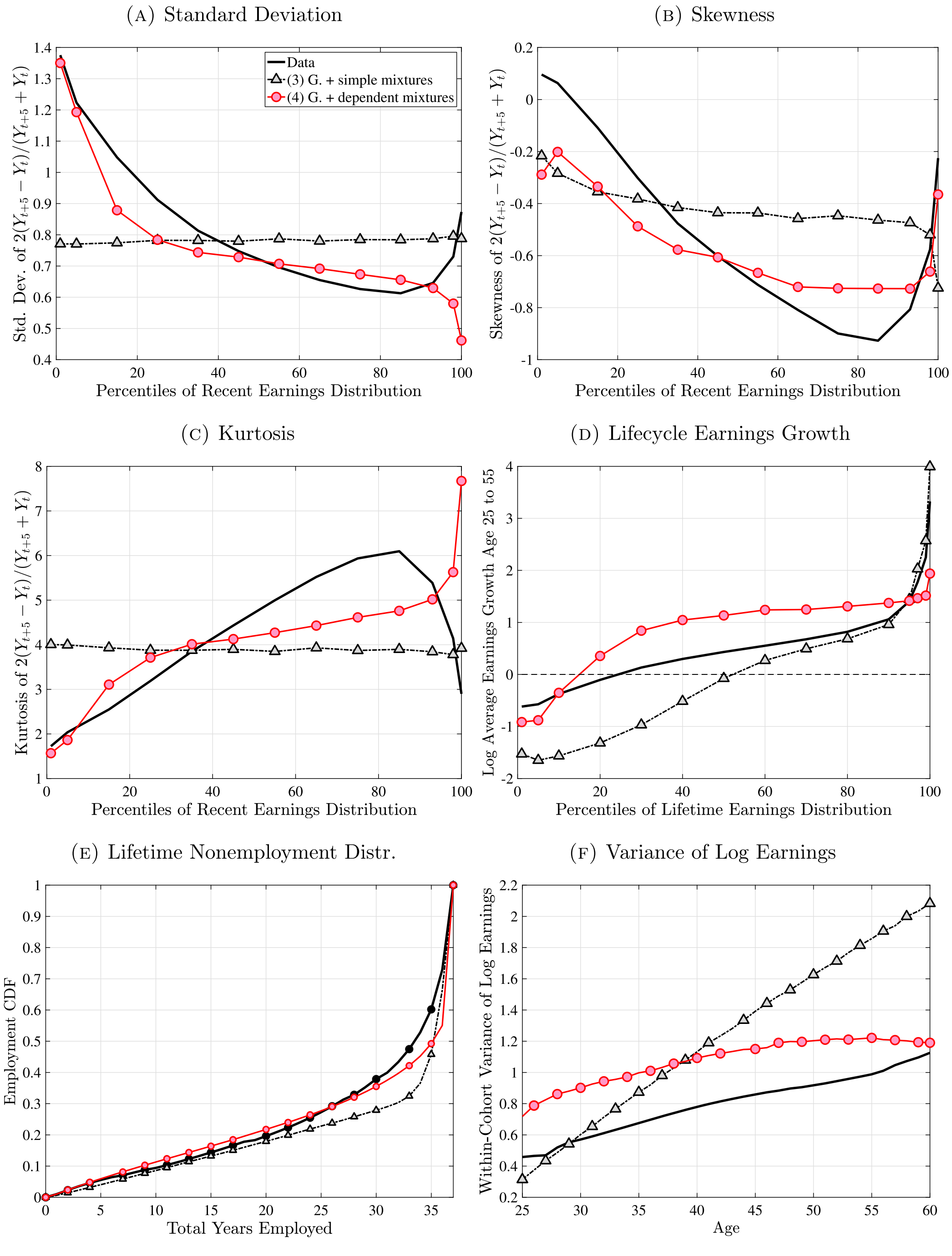


FIGURE D.12.—Estimated model versus data: key moments (Models 3 and 4).

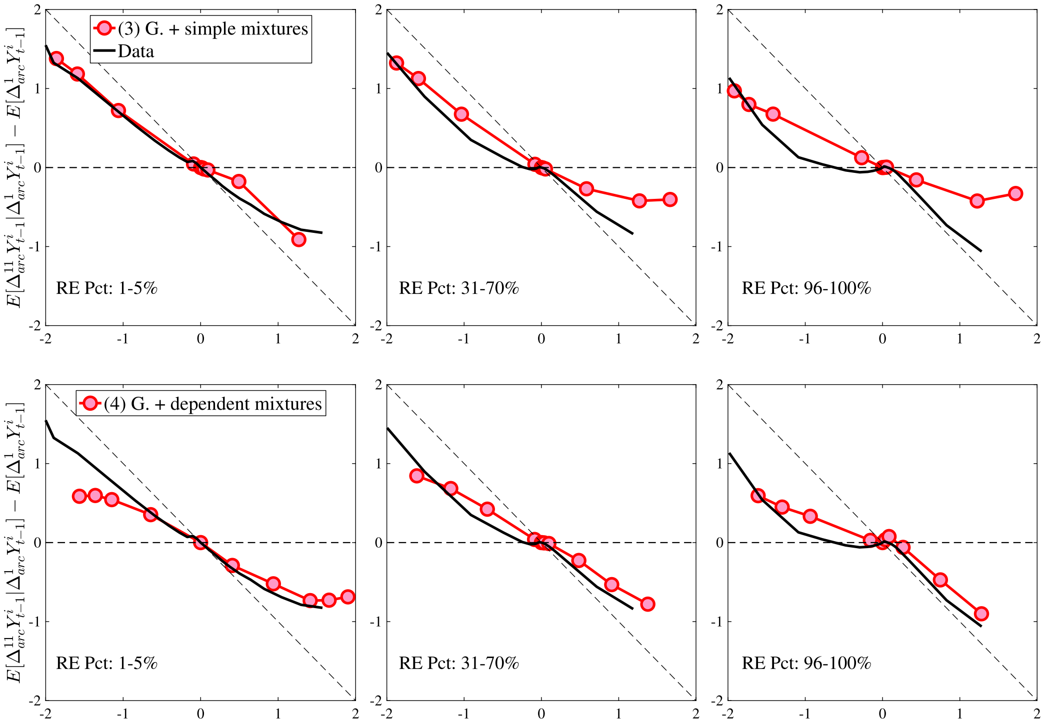


FIGURE D.13.—Fit of Models 3 and 4 on selected impulse response moments.

In column (14), we introduce ex ante variance heterogeneity in the 2-state benchmark process. Thus, the variance of each innovation from $\mathcal{N}(\mu_{z_j,1}, \sigma^i_{z_j,1})$ to the persistent component j is individual-specific, with a lognormal distribution with mean $\overline{\sigma}_{z_j,1}$ and a standard deviation proportional to $\widetilde{\sigma}_{z_j,1}$, that is, $\log(\sigma^i_{z_j,1}) \sim \mathcal{N}(\log \overline{\sigma}_{z_j,1} - \frac{\widetilde{\sigma}^2_{z,1}}{2}, \widetilde{\sigma}_{z_j,1})$.

The last column (column (15)) shows the parameter estimates for the specification presented in column (3) but without imposing a lower bound for the mean of persistent shocks.

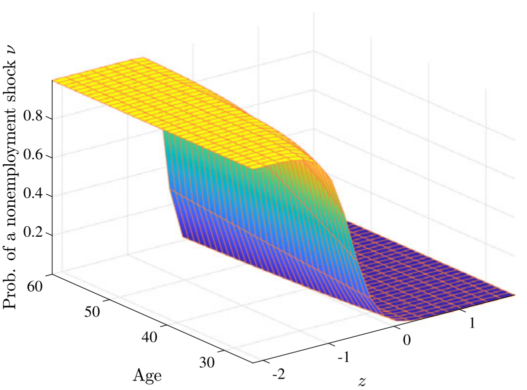


FIGURE D.14.—Benchmark process: 3-D Plot of nonemployment shock probability p_ν.

TABLE D.III

ADDITIONAL PARAMETER ESTIMATES FOR PROCESSES IN TABLE IV[a]

Model:		(1)	(2)	(3)	(4)	(5)	(6)	
		Gaussian process					*Benchmark Process*	
Parameters							Parameters	Std. Err.
AR(1) *Component*		G	G	mix	mix	mix	mix	mix
↪*Probability age/inc.*		—	—	no/no	yes/yes	no/no	no/no	no/no
Nonemployment shocks		no	yes	no	no	yes	yes	yes
↪*Probability age/inc.*		—	yes/yes	—	—	yes/yes	yes/yes	yes/yes
Transitory Shocks		G	G	mix	mix	mix	mix	mix
HIP		no	no	no	no	no	yes	yes
		Deterministic Lifecycle Profile Parameters						
a_0	$\times 1$	0.740	2.569	2.547	2.176	2.746	2.581	0.0018
a_1	$\times t$	0.337	0.766	−0.144	0.169	0.624	0.812	0.0011
a_2	$\times t^2$	0.070	−0.152	−0.059	−0.100	−0.167	−0.185	0.0003
		Nonemployment Shock Probability Function Parameters						
a_ν	$\times 1$		−3.036			−2.495	−3.353	0.0039
b_ν	$\times t$		−0.917			−1.037	−0.859	0.0031
c_ν	$\times z_t$		−5.397			−5.051	−5.034	0.0064
d_ν	$\times t \times z_t$		−4.442			−1.087	−2.895	0.0036
		Normal Mixture Probability Function Parameters						
a_{z_1}	$\times 1$			0.05	−0.474	0.176	0.407	0.0005
b_{z_1}	$\times t$				1.961			
c_{z_1}	$\times z_{t-1}$				−3.183			
d_{z_1}	$\times t \times z_{t-1}$				−0.187			

[a]Notes: We define the deterministic lifecycle profile as a quadratic function of t, $g(t) = a_0 + a_1 t + a_2 t^2$, where $t = (\text{age} - 24)/10$. $y_t = z_t$ for the 1-state income process with 1 AR(1) component, and $y_t = z_1 + z_2$ for the 2-state income process with two AR(1) components.

TABLE D.IV

UNIFORM NONEMPLOYMENT RISK AND NON-GAUSSIAN PERSISTENT SHOCKS

Model:	(1b) Uniform Nonemp.	(3b) Mix. Trans.	*Model*:	(1b) Uniform Nonemp.	(3b) Mix. Trans.
AR(1) Component	G	mix	*Objective value*	73.39	62.11
↪Prob. age/income	—	no/no	Decomposition:		
Nonemployment shocks	yes	no	(i) Standard deviation	8.53	7.56
↪Prob. age/inc.	no/no	—	(ii) Skewness	39.78	21.14
Transitory shocks	G	G	(iii) Kurtosis	19.15	17.99
HIP	no	no	(iv) Impulse resp. short	20.72	22.74
Parameters			(v) Impulse resp. long	32.24	37.73
ρ	1.015	0.998	(vi) Lifetime inc. growth	39.63	26.60
p_z		5.9%	(vii) Age-ineq. profile	17.63	16.64
$\mu_{\eta,1}$		−1.0	(viii) Nonempl. CDF	3.96	10.29
$\sigma_{\eta,1}$	0.085	1.580	*Model Selection p-val.*		
$\sigma_{\eta,2}$		0.0291	Test 1	0.000	0.000
σ_{z_0}	0.183	0.340	Test 2	0.000	0.000
λ	0.547				
$\sigma_{\varepsilon,1}$	0.488	0.371			

TABLE D.V

ADDITIONAL SPECIFICATIONS[a]

Specification:	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
	Nonemp. depends on α	Nonemp. depends on $\alpha + z$ & age	Nonemp. depends on α and z	Nonemp. depends on z and z^2	Nonemp. depends on z, z^2, & age	1-state Benchmark w/ var. het.	1-state Benchmark w/ η het.	2-state Benchmark w/ var. het.	Column (3) No Bound for $\mu_{z,1}$
AR(1) *Component*	mix	mix	mix	mix	mix	mix	mix	2 mixtures	mix
↪*Probability age/inc.*	no/no	no/no	no/no	no/no	no/no	no/no	yes/yes	yes/yes	no/no
Nonemployment shocks	yes	yes	yes	yes	yes	yes	yes	yes	no
↪*Probability age/inc.*	yes/α	yes/$\alpha + z$	no/α, z	no/z, z^2	yes/z, z^2	yes/yes	yes/yes	yes/yes	—
Transitory Shocks	mix	mix	mix	mix	mix	mix	mix	mix	mix
HIP	no	no	no	yes	yes	yes	yes	yes	no
Variance Heterogeneity	no	no	no	no	no	yes	no	yes	no
Parameters									
ρ_1	0.847	0.978	0.976	0.968	0.964	0.960	0.965	0.824	1.005
ρ_2								0.979	
p_{z_1}	0.044	0.150	0.091	0.219	0.427	0.274			0.0124
$\mu_{\eta_1,1}$	−0.961	−0.327	−0.576	−0.335	−0.120	−0.107	−0.436	−0.393	−5.63
$\mu_{\eta_2,1}$								−0.270	
$\sigma_{\eta_1,1}$	1.396	0.689	0.575	0.304	0.345	0.433	0.788	0.620	0.471
$\sigma_{\eta_1,2}$	0.066	0.083	0.168	0.174	0.061	0.0827	0.195	0.116	0.148
$\sigma_{\eta_2,1}$								0.564	
$\sigma_{\eta_2,2}$								0.001	
$\sigma^i_{\eta_1,1}$						0.162		0.002	
$\sigma^i_{\eta_2,1}$								0.013	
$\sigma_{z_1,0}$	0.339	0.154	0.193	0.719	0.689	1.5042	0.530	0.189	0.227
$\sigma_{z_2,0}$								0.603	
λ	0.196	0.104	0.022	0.001	0.002	0.005	0.014	0.001	
$a_\nu \times$ 1	−3.875	−2.399	−3.191	−4.131	−3.115	−2.958	−3.773	−3.045	
$b_\nu \times$ t, α	−5.366	−1.241	−1.932		−1.108	−1.341	0.468	−1.092	
$c_\nu \times$ $[z_t, \alpha, (\alpha + z_t)]$		−3.550	−5.528	−5.477	−3.914	−4.222	−3.635	−3.219	
$d_\nu \times$ α^2, z^2	0.293			1.611	0.832				
$e_\nu \times$ $t[z_t, (\alpha + z_t)]$		−1.837	0.700		−2.403	−3.450	−3.560	−2.240	
prob_ε	0.227	0.104	0.239	0.290	0.105	0.130	0.115	0.095	0.210
$\mu_{\varepsilon,1}$	0.115	0.249	0.146	0.159	0.296	0.223	0.340	0.340	−0.09
$\sigma_{\varepsilon,1}$	0.449	0.562	0.239	0.142	0.247	0.384	0.283	0.438	1.024
$\sigma_{\varepsilon,2}$	0.061	0.042	0.072	0.022	0.064	0.048	0.081	0.025	0.024
σ_α	0.808	0.520	0.640	0.289	0.274	0.313	0.267	0.273	0.4367
σ_β				0.194	0.205	0.217	0.204	0.182	
$\text{corr}_{\alpha\beta}$				0.630	0.676	0.489	0.974	0.435	
Objective value	40.33	28.28	26.19	23.32	22.11	21.81	21.74	18.43	47.9
Decomposition:									
(i) Standard deviation	4.86	4.81	5.31	5.86	5.93	5.09	5.95	4.95	7.92
(ii) Skewness	21.15	15.76	14.67	12.760	10.65	10.09	9.52	7.06	22.99
(iii) Kurtosis	6.67	6.04	4.80	5.63	5.88	6.32	6.70	4.70	13.99
(iv) Impulse resp.	27.98	18.19	15.82	13.54	13.50	13.25	13.70	11.58	36.40
(v) Inc. growth	15.10	10.81	10.39	9.21	8.39	8.24	7.84	8.37	8.94
(vi) Inequality	7.70	3.38	1.66	3.36	3.47	3.68	3.16	2.23	4.99
(vii) Nonemployment CDF	6.46	5.71	7.66	5.94	6.43	7.14	6.52	5.86	8.91

[a]Note: In this table, we present estimates from an earlier version of the paper for which we use a slightly different weighting matrix in the estimation. In particular, the weighting matrix used in this table first assigns 15% relative weight to the employment CDF moments. The rest of the moments share the remaining 85% weight according to the following scheme: the cross-sectional moments (standard deviation, skewness, kurtosis) collectively receive a relative weight of 35%, the lifecycle earnings growth moments and impulse response moments each receive a weight of 25%, and the variance of log earnings by age receives a weight of 15%.

D.6. *Autocovariance Structure of Earnings*

Another set of moments that has been widely used for decades to estimate income processes is the autocovariance of earnings in level and changes (e.g., Abowd and Card (1989)). In this section, for completeness with the earlier literature, we document these moments from data along with their simulation counterparts from our benchmark specification. Tables D.VI and D.VII show autocovariance matrix of log earnings over the life cycle in levels and changes, respectively.

Following Abowd and Card (1989), MaCurdy (1982), in Figure D.15, we show the autocovariance of 1-year log earnings growth at several lags:

$$\operatorname{cov}\bigl[\log(y_{h+1}) - \log(y_h), \log(y_{h+k+1}) - \log(y_{h+k})\bigr] \quad \text{for } k > 1, \tag{D.13}$$

where $\log(y_h)$ is log earnings at age h. As usual, we only include observations that are above the minimum income threshold, $Y_{\min}$. The left panel of Figure D.15 shows that in our data, consistent with earlier work using survey data (Meghir and Pistaferri (2004)), the autocovariance of earnings growth is small and approaches to zero quickly after a couple of lags. Our model generates a similar pattern, although the autocovariance for $k = 1$ tends to be smaller and the lags $k > 1$ approach to zero (right panel of Figure D.15). Similar patterns are also clearly seen in autocorrelations of earnings from the data and our benchmark process (Figure D.16). Both in the data and in our benchmark process, the autocorrelations of earnings growth starts around -0.20 and approaches quickly to zero after a couple of lags (in the benchmark process for $k > 1$). The small discrepancy between our process and the data for shorter lags k can be addressed by modeling the transitory component as a moving average of order q (MA(q)) process (see Meghir and Pistaferri (2004)).

MaCurdy (1982) noted that if a HIP component is present ($\sigma_\beta^2 > 0$), the autocovariance of one-year log earnings growth should turn positive at longer lags. Figure D.15 shows that in the data, the autocovariance of log earnings growth does not increase above zero even after 35 years (left panel). Therefore, this test would not reject $\sigma_\beta^2 = 0$ in the data. Interestingly, MaCurdy's test reaches a similar conclusion when applied to data simulated from our benchmark specification, which features a sizable HIP component (right panel of Figure D.15).

Guvenen (2009) discussed why MaCurdy (1982) may reject $\sigma_\beta^2 > 0$ even if the true process features a HIP component. If the variance of the persistent component is large enough, the autocovariance of earnings growth may not be significantly greater than zero even after 20–30 years. This is indeed the case in our model. The theoretical autocovariance of earnings growth given by equation (D.13) for the estimated parameter values of our benchmark process becomes positive, and barely so, only with a 35 year lag. Furthermore, as Karahan, Ozkan, and Song (2019) and Guvenen (2009) showed, the HIP component may have a Pareto distribution, which would imply that β heterogeneity is negligible for most of the population, but the top of the distribution has a much larger β than the rest.

TABLE D.VI

AUTOCOVARIANCE MATRIX OF LOG EARNINGS

age	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60
25	0.68	0.41	0.34	0.30	0.28	0.26	0.25	0.24	0.23	0.23	0.22	0.21	0.21	0.20	0.20	0.20	0.20	0.19	0.19	0.19	0.18	0.18	0.18	0.17	0.17	0.16	0.16	0.15	0.14	0.14	0.13	0.12	0.11	0.10	0.10	0.08
26		0.69	0.43	0.37	0.33	0.31	0.29	0.28	0.27	0.27	0.26	0.25	0.25	0.25	0.24	0.24	0.23	0.23	0.23	0.22	0.22	0.22	0.22	0.21	0.20	0.20	0.19	0.18	0.18	0.17	0.16	0.15	0.14	0.13	0.12	0.11
27			0.70	0.45	0.39	0.36	0.34	0.32	0.31	0.30	0.29	0.29	0.28	0.28	0.27	0.27	0.26	0.26	0.26	0.25	0.25	0.25	0.24	0.24	0.23	0.23	0.22	0.21	0.21	0.20	0.19	0.18	0.17	0.16	0.15	0.14
28				0.72	0.46	0.41	0.38	0.36	0.35	0.34	0.33	0.32	0.31	0.31	0.30	0.30	0.29	0.29	0.28	0.28	0.28	0.27	0.27	0.27	0.26	0.25	0.25	0.24	0.23	0.22	0.22	0.20	0.19	0.19	0.18	0.16
29					0.73	0.48	0.43	0.40	0.38	0.37	0.36	0.35	0.34	0.34	0.33	0.33	0.32	0.32	0.31	0.31	0.31	0.30	0.30	0.29	0.29	0.28	0.27	0.27	0.26	0.25	0.24	0.23	0.22	0.21	0.20	0.19
30						0.74	0.50	0.45	0.42	0.41	0.39	0.38	0.37	0.37	0.36	0.36	0.35	0.34	0.34	0.33	0.33	0.32	0.32	0.31	0.31	0.30	0.29	0.29	0.28	0.27	0.26	0.24	0.23	0.22	0.21	0.20
31							0.76	0.52	0.47	0.44	0.43	0.41	0.40	0.40	0.39	0.38	0.37	0.37	0.36	0.36	0.35	0.35	0.34	0.33	0.33	0.32	0.32	0.31	0.30	0.29	0.27	0.26	0.26	0.24	0.23	0.22
32								0.77	0.53	0.49	0.46	0.45	0.43	0.42	0.41	0.41	0.40	0.39	0.39	0.38	0.37	0.37	0.36	0.35	0.35	0.34	0.33	0.33	0.32	0.31	0.29	0.28	0.27	0.26	0.25	0.24
33									0.79	0.55	0.51	0.48	0.46	0.45	0.44	0.43	0.42	0.42	0.41	0.40	0.40	0.39	0.39	0.38	0.37	0.36	0.35	0.35	0.34	0.32	0.31	0.30	0.29	0.28	0.27	0.26
34										0.81	0.57	0.53	0.50	0.48	0.47	0.46	0.45	0.44	0.43	0.43	0.42	0.41	0.41	0.40	0.39	0.38	0.37	0.36	0.36	0.35	0.33	0.32	0.31	0.29	0.28	0.27
35											0.82	0.59	0.54	0.52	0.50	0.49	0.47	0.47	0.45	0.45	0.44	0.43	0.43	0.41	0.41	0.40	0.39	0.38	0.37	0.36	0.35	0.34	0.32	0.31	0.30	0.28
36												0.83	0.60	0.56	0.53	0.52	0.50	0.49	0.48	0.47	0.46	0.45	0.45	0.44	0.43	0.42	0.41	0.40	0.39	0.38	0.37	0.36	0.34	0.33	0.31	0.30
37													0.84	0.62	0.57	0.55	0.53	0.51	0.50	0.49	0.48	0.47	0.47	0.45	0.45	0.44	0.42	0.42	0.41	0.39	0.38	0.37	0.35	0.34	0.33	0.31
38														0.86	0.63	0.59	0.56	0.54	0.53	0.52	0.51	0.50	0.49	0.48	0.47	0.46	0.45	0.43	0.42	0.41	0.40	0.39	0.37	0.36	0.34	0.33
39															0.88	0.65	0.60	0.57	0.56	0.54	0.53	0.52	0.51	0.49	0.49	0.47	0.46	0.45	0.44	0.43	0.41	0.40	0.39	0.37	0.36	0.34
40																0.89	0.66	0.61	0.59	0.57	0.55	0.54	0.53	0.51	0.50	0.49	0.48	0.47	0.45	0.44	0.43	0.42	0.40	0.39	0.37	0.36
41																	0.90	0.67	0.62	0.60	0.58	0.56	0.55	0.54	0.52	0.51	0.50	0.48	0.47	0.46	0.44	0.43	0.42	0.40	0.39	0.37
42																		0.91	0.68	0.63	0.61	0.59	0.57	0.56	0.54	0.53	0.51	0.50	0.49	0.48	0.46	0.44	0.43	0.41	0.40	0.38
43																			0.92	0.69	0.64	0.62	0.60	0.58	0.56	0.55	0.53	0.52	0.50	0.49	0.47	0.46	0.45	0.43	0.42	0.40
44																				0.93	0.70	0.66	0.63	0.61	0.59	0.57	0.55	0.54	0.52	0.51	0.49	0.48	0.46	0.45	0.43	0.42
45																					0.93	0.71	0.66	0.63	0.61	0.59	0.57	0.56	0.54	0.53	0.51	0.49	0.48	0.46	0.44	0.43
46																						0.95	0.72	0.67	0.64	0.62	0.60	0.58	0.56	0.55	0.53	0.51	0.50	0.48	0.46	0.44
47																							0.95	0.72	0.67	0.64	0.62	0.60	0.58	0.56	0.54	0.53	0.51	0.49	0.48	0.46
48																								0.95	0.73	0.68	0.65	0.62	0.60	0.58	0.56	0.54	0.52	0.50	0.49	0.47
49																									0.95	0.73	0.68	0.65	0.63	0.60	0.58	0.56	0.54	0.52	0.50	0.48
50																										0.96	0.73	0.69	0.65	0.63	0.60	0.58	0.56	0.54	0.52	0.50
51																											0.96	0.74	0.69	0.66	0.63	0.60	0.58	0.56	0.54	0.51
52																												0.97	0.74	0.69	0.66	0.63	0.60	0.58	0.56	0.53
53																													0.98	0.75	0.69	0.66	0.63	0.60	0.58	0.56
54																														0.99	0.75	0.70	0.66	0.63	0.60	0.58
55																															0.98	0.75	0.70	0.66	0.63	0.60
56																																1.00	0.76	0.71	0.67	0.64
57																																	1.03	0.78	0.73	0.68
58																																		1.04	0.80	0.74
59																																			1.06	0.81
60																																				1.08

TABLE D.VII

AUTOCOVARIANCE MATRIX OF LOG EARNINGS GROWTH

age	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59
25	0.413	−0.091	−0.028	−0.012	−0.006	−0.005	−0.002	−0.001	0.000	0.000	−0.001	−0.001	0.000	0.000	0.000	0.000	0.001	0.002	0.002	−0.001	0.001	0.001	−0.001	0.001	0.000	−0.001	0.001	0.001	0.001	−0.001	0.002	0.000	−0.003	−0.003	0.006
26		0.389	−0.088	−0.027	−0.010	−0.006	−0.004	−0.001	−0.001	0.000	0.000	−0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.000	0.000	−0.001	0.001	0.000	0.000	0.001	0.000	0.000	0.000	0.000	−0.001	0.001	0.000	0.001
27			0.372	−0.087	−0.026	−0.010	−0.006	−0.003	−0.003	0.000	0.000	0.000	0.000	0.000	0.000	−0.001	−0.001	−0.001	0.001	0.001	0.001	0.000	0.001	0.000	0.000	0.001	−0.001	0.001	0.001	0.002	−0.001	−0.001	−0.001	0.002	−0.001
28				0.359	−0.085	−0.024	−0.009	−0.005	−0.003	−0.002	−0.002	0.000	0.000	0.000	0.000	0.000	−0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	−0.001	0.000	0.001	0.001	0.001	0.000	−0.002
29					0.347	−0.084	−0.023	−0.008	−0.005	−0.003	−0.002	−0.002	−0.001	0.000	−0.001	0.000	0.001	−0.002	0.001	0.000	0.000	0.000	0.001	0.000	0.000	0.001	−0.001	−0.001	0.000	−0.001	0.001	0.001	0.000	−0.002	−0.001
30						0.339	−0.084	−0.022	−0.009	−0.005	−0.004	−0.002	0.000	−0.002	−0.001	0.000	−0.001	0.000	−0.001	0.000	0.000	−0.001	0.000	0.000	−0.001	0.001	−0.001	0.000	0.000	0.000	−0.001	0.002	0.001	0.000	−0.001
31							0.331	−0.083	−0.021	−0.009	−0.005	−0.003	−0.002	−0.001	−0.001	−0.001	0.000	0.000	−0.001	−0.001	0.000	−0.001	0.000	0.000	0.001	0.000	0.000	0.001	0.000	−0.002	0.001	−0.002	0.000	0.001	0.001
32								0.324	−0.083	−0.021	−0.008	−0.005	−0.003	−0.002	−0.001	−0.001	0.000	0.000	0.000	0.000	0.000	−0.001	0.000	0.000	0.000	0.000	0.001	0.000	−0.001	0.001	0.002	0.002	−0.001	0.001	0.000
33									0.320	−0.082	−0.021	−0.008	−0.005	−0.003	−0.003	−0.001	−0.001	−0.001	0.000	−0.001	−0.001	0.000	−0.001	0.001	−0.001	0.000	−0.001	0.001	0.000	−0.001	−0.001	−0.001	−0.001	0.001	0.001
34										0.312	−0.079	−0.020	−0.007	−0.005	−0.003	−0.001	−0.002	−0.002	−0.001	0.000	−0.001	0.000	−0.001	0.000	0.001	0.000	0.000	0.000	−0.001	0.001	0.001	−0.001	0.001	0.000	−0.002
35											0.305	−0.079	−0.019	−0.007	−0.004	−0.003	−0.003	−0.001	−0.001	0.000	−0.001	0.000	0.000	0.000	−0.001	0.000	0.000	0.000	0.001	0.000	0.000	0.000	−0.001	−0.001	0.000
36												0.301	−0.077	−0.019	−0.007	−0.005	−0.003	−0.002	−0.002	−0.001	0.000	−0.001	−0.001	0.000	−0.001	−0.002	0.001	0.000	0.000	0.000	−0.001	−0.001	−0.001	0.000	0.000
37													0.296	−0.078	−0.019	−0.007	−0.004	−0.004	−0.001	−0.002	−0.001	−0.001	0.000	−0.002	−0.001	0.000	−0.001	0.000	−0.001	0.001	0.000	0.000	0.001	−0.002	−0.001
38														0.293	−0.077	−0.018	−0.007	−0.004	−0.002	−0.002	−0.002	−0.001	−0.001	0.000	0.000	0.000	−0.001	−0.001	0.000	−0.001	−0.001	0.002	0.000	0.001	0.000
39															0.295	−0.078	−0.018	−0.007	−0.004	−0.003	−0.002	−0.001	−0.001	−0.001	0.000	0.000	0.000	−0.001	0.000	0.001	−0.001	0.000	0.000	−0.001	−0.001
40																0.290	−0.075	−0.017	−0.007	−0.004	−0.002	−0.003	−0.001	−0.002	−0.002	0.000	−0.001	−0.001	0.000	−0.002	0.000	−0.001	0.001	−0.001	0.000
41																	0.287	−0.076	−0.017	−0.006	−0.004	−0.003	−0.001	−0.001	−0.001	−0.002	−0.001	0.000	−0.001	−0.001	−0.001	0.000	−0.001	0.001	0.001
42																		0.284	−0.074	−0.016	−0.006	−0.004	−0.004	−0.001	−0.002	−0.001	0.000	0.000	−0.001	−0.001	−0.001	−0.001	0.000	−0.001	0.000
43																			0.280	−0.073	−0.016	−0.008	−0.003	−0.003	−0.002	−0.001	−0.001	0.000	0.000	0.000	−0.002	0.000	−0.001	−0.001	0.000
44																				0.278	−0.073	−0.015	−0.008	−0.004	−0.003	−0.003	0.000	−0.001	−0.001	0.000	0.000	−0.001	−0.001	0.000	−0.001
45																					0.274	−0.072	−0.015	−0.006	−0.003	−0.003	−0.002	−0.002	−0.001	−0.001	0.001	−0.002	−0.001	0.000	0.000
46																						0.273	−0.070	−0.016	−0.007	−0.003	−0.003	−0.001	−0.001	−0.002	−0.002	0.000	0.000	−0.001	−0.001
47																							0.271	−0.071	−0.015	−0.007	−0.004	−0.003	−0.002	−0.001	0.000	−0.001	−0.002	0.000	−0.001
48																								0.269	−0.069	−0.015	−0.006	−0.004	−0.003	−0.002	−0.002	−0.001	−0.001	−0.001	−0.001
49																									0.269	−0.069	−0.014	−0.007	−0.003	−0.003	−0.002	−0.002	0.000	−0.001	−0.001
50																										0.266	−0.068	−0.014	−0.006	−0.004	−0.002	−0.001	−0.002	−0.002	0.000
51																											0.271	−0.066	−0.015	−0.006	−0.004	−0.002	−0.002	−0.001	−0.001
52																												0.272	−0.068	−0.014	−0.006	−0.003	−0.003	−0.003	−0.001
53																													0.273	−0.065	−0.014	−0.008	−0.003	−0.003	−0.002
54																														0.274	−0.067	−0.013	−0.005	−0.003	−0.002
55																															0.285	−0.064	−0.014	−0.006	−0.004
56																																0.310	−0.065	−0.016	−0.008
57																																	0.308	−0.065	−0.013
58																																		0.310	−0.065
59																																			0.317

(A) Data (B) Benchmark Specification

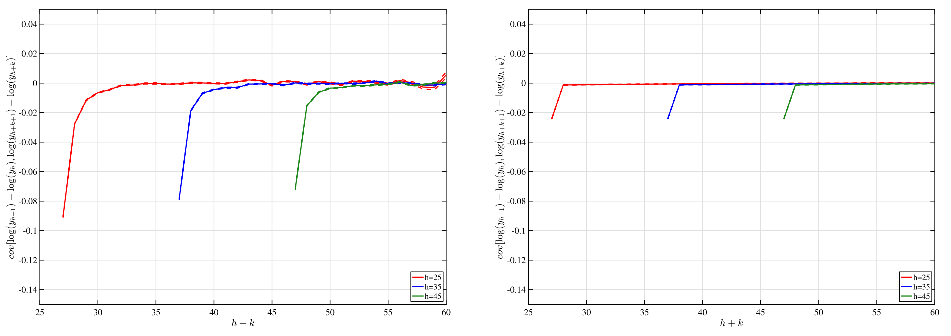


FIGURE D.15.—Autocovariance of earnings growth: data versus model.

(A) Data (B) Benchmark Specification

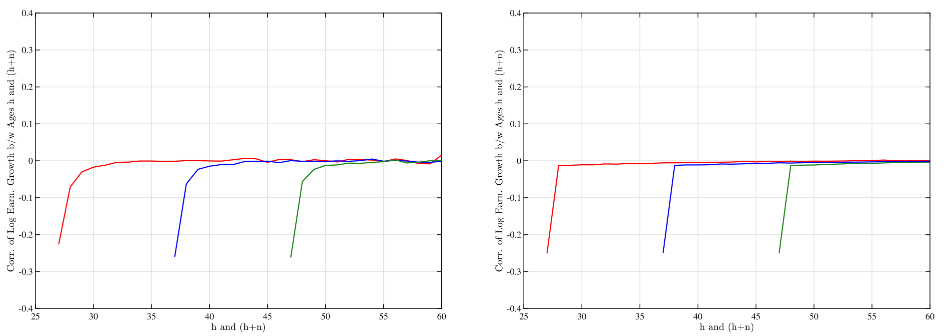


FIGURE D.16.—Autocorrelation of earnings growth: data versus model.

REFERENCES

ABOWD, J. M., AND D. CARD (1989): "On the Covariance Structure of Earnings and Hours Changes," *Econometrica*, 57 (2), 411–445. [45]

CROW, E. L., AND M. SIDDIQUI (1967): "Robust Estimation of Location," *Journal of the American Statistical Association*, 62 (318), 353–389. [11,12]

DEATON, A., AND C. PAXSON (1994): "Intertemporal Choice and Inequality," *Journal of Political Economy*, 102 (3), 437–467. [25,30]

GUVENEN, F. (2009): "An Empirical Investigation of Labor Income Processes," *Review of Economic Dynamics*, 12 (1), 58–79. [45]

HANSEN, L. P. (1982): "Large Sample Properties of Generalized Method of Moments Estimators," *Econometrica*, 50 (4), 1029–1054. [35]

KARAHAN, F., S. OZKAN, AND J. SONG (2019): "Anatomy of Lifetime Earnings Inequality: Heterogeneity in Job Ladder Risk vs. Human Capital," FRB of New York Staff Report (908). [45]

MACURDY, T. E. (1982): "The Use of Time Series Process to Model the Error Structure of Earnings in a Longitudinal Data Analysis," *Journal of Econometrics*, 18, 83–114. [45]

MEGHIR, C., AND L. PISTAFERRI (2004): "Income Variance Dynamics and Heterogeneity," *Econometrica*, 72 (1), 1–32. [45]

MOORS, J. J. A. (1988): "A Quantile Alternative for Kurtosis," *Journal of the American Statistical Association. Series D (The Statistician)*, 37 (1), 25–32. [11]

OLSEN, A., AND R. HUDSON (2009): "Social Security Administration's Master Earnings File: Background Information," *Social Security Bulletin*, 69 (3), 29–46. [2]

PANIS, C., R. EULLER, C. GRANT, M. BRADLEY, C. E. PETERSON, R. HIRSCHER, AND P. STINBERG (2000): *SSA Program Data User's Manual*. Baltimore, MD: Social Security Administration. [2]

STORESLETTEN, K., C. I. TELMER, AND A. YARON (2004): "Consumption and Risk Sharing Over the Life Cycle," *Journal of Monetary Economics*, 51 (3), 609–633. [25]

ZHANG, H., A. R. CONN, AND K. SCHEINBERG (2010): "A Derivative-Free Algorithm for Least-Squares Minimization," *SIAM Journal on Optimization*, 20 (6), 3555–3576. [35]

Co-editor Giovanni L. Violante handled this manuscript.

Manuscript received 4 August, 2016; final version accepted 26 February, 2021; available online 19 April, 2021.