What Do Data on Millions of U.S. Workers Reveal about Lifecycle Earnings Dynamics?

by Fatih Guvenen, Fatih Karahan, Serdar Ozkan, Jae Song
Econometrica, Volume 89, Issue 5, September 2021,
Pages 2303–2339

James J. Heckman

THE UNIVERSITY OF CHICAGO

Econ 350, Spring 2022

1. Introduction

GOAL OF THIS PAPER

- To characterize the most salient properties of individual earnings dynamics over the life cycle, focusing on nonnormalities and nonlinearities.
- First, by studying its higher-order moments (specifically, skewness and kurtosis), we investigate the distribution of earnings changes and whether it can be well approximated by a normal distribution.
- Second, we explore mean reversion patterns of earnings changes that may differ between positive and negative changes as well as by size.
- Finally, we study how these properties vary over the life cycle and across the earnings distribution.

- Our descriptive analysis covers (i) the properties of the distributions of earnings changes, (ii) the extent of mean reversion during the 10 years following earnings changes, and (iii) workers' long-term outcomes covering their entire working lives, such as cumulative earnings growth and the incidence of nonemployment.

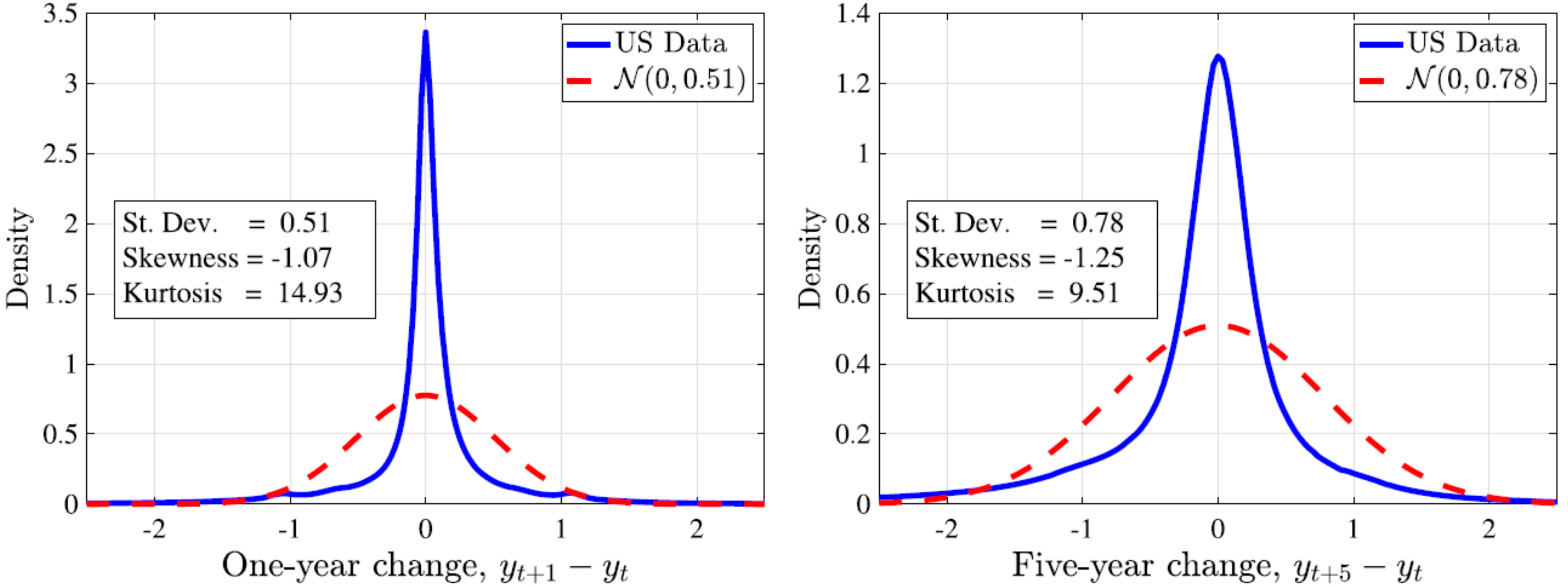
- Starting with the distribution of earnings changes, we find that it is left- (negatively) skewed, and this left-skewness becomes more severe as individuals get older or their earnings increase (or both).

- In contrast, young low-income workers face an almost symmetric distribution.

- The rise in left-skewness over the life cycle is entirely due to a reduction in opportunities for large gains from ages 25 to 45 and to the increasing likelihood of a sharp fall in earnings after age 45.

- In addition, earnings growth displays a very high kurtosis relative to a Gaussian density (Figure 1).

Figure 1. Histograms of one- and five-year log earnings changes.



Notes: This figure plots the empirical densities of one- and five-year earnings changes superimposed on Gaussian densities with the same standard deviation. The data are for all workers in the base sample defined in Section 2 and $t = 1997$.

- There are far more people in the data with very small or with extreme earnings changes and fewer people with middling ones.
- Also, a typical worker sees a change larger than three standard deviations with a 2.4% chance, which is about one-ninth as likely under a normal distribution.
- Importantly, the average kurtosis masks significant heterogeneity.

- To shed some light on the sources of these nonnormalities, we analyze data from the Panel Study of Income Dynamics (PSID) on work hours, hourly wages, and a rich set of additional covariates that are not available in the SSA data.

- We find that hourly wage changes exhibit little left-skewness but an excess kurtosis with a magnitude and lifecycle variation similar to earnings changes.

- Furthermore, wage changes are at least as important as changes in hours, even in the tails of the distribution.

- Moreover, workers experiencing extreme changes are likely to have gone through nonemployment or job or occupation changes, or to have experienced health shocks, suggesting that the tails are not a statistical artifact or measurement error in the survey data.

- Next, we characterize the mean reversion patterns of earnings changes by estimating nonparametric impulse response functions conditional on recent earnings and on the size and sign of the change.

- We find two types of asymmetry:

  1. Fixing the size of the change, positive changes to high-earnings individuals are quite transitory, while negative ones are persistent; in contrast, the opposite is true for low-earnings individuals.

  2. With a fixed level of earnings, the strength of the mean reversion differs by the size of the change: Large changes tend to be much more transitory than small ones.

- These asymmetries are difficult to detect in a covariance matrix, in which all sorts of earnings changes—large, small, positive, and negative—are masked by a single statistic.

- Finally, we document two facts regarding long-term outcomes covering individuals' entire working lives.

  1. The cumulative earnings growth over the life cycle varies systematically and substantially across groups of workers with different lifetime earnings.

  2. There is substantial variation in individuals' lifetime nonemployment rate—which we define as the fraction of a lifetime (ages 25 to 60) spent as (full-year) nonemployed.

- These numbers imply an extremely high persistence in the long-term nonemployment state.

- While the nonparametric approach allows us to establish key features of earnings dynamics in a transparent way, a tractable parametric process is indispensable because (i) it allows us to connect earnings changes to underlying innovations or shocks to earnings, and (ii) it can be used as an input to calibrate quantitative models with idiosyncratic risk.

- Therefore, in Section 6, we target the empirical moments described above to estimate a range of income processes.

- We start with the familiar linear-Gaussian framework (i.e., the persistent plus transitory model with Gaussian shocks) and build on it incrementally until we arrive at a rich, yet tractable, benchmark specification that can capture the key features of the data.

- Our benchmark process incorporates two key features to the linear-Gaussian framework: normal mixture innovations to the persistent and transitory components and, more importantly, a long-term nonemployment shock with a realization probability that depends on age and earnings.
- This *state-dependent* employment risk generates recurring nonemployment with scarring effects concentrated among young and low-income individuals and helps capture the lifecycle and income variation of the moments.
- Our empirical facts require non-Gaussian features in *persistent* innovations; these can be achieved by such income-dependent nonemployment shocks or non-Gaussian shocks to the persistent component, but not by a uniform nonemployment risk that is transitory in nature.

2. Data and Variable Construction

2.1. The SSA Data Set

- We draw a representative 10% panel sample of the U.S. population from the Master Earnings File (MEF) of the SSA, which combines various data sets that go back as far as 1978.
- Wage income is not top coded throughout our sample, whereas self-employment income was capped at the SSA taxable limit until 1994.
- Although this top coding affects only a small number of individuals, we restrict our sample to the 1994–2013 period to ensure that our analysis is not affected by this issue.
- The data set also has some important drawbacks, such as limited demographic information, the absence of capital income, and the lack of hours (and thus hourly wage) data.
- To overcome some of these limitations, we supplement our analysis with survey data whenever possible.
- Another important limitation is the lack of household-level data.

2.2. Sample Selection

- Our *base sample* is a revolving panel consisting of males with some labor market attachment that is designed to maximize the sample size (important for precise computation of higher-order moments in finely defined groups) and keep the age structure stable over time.

- First, in order for an individual-year income observation to be *admissible* to the *base sample*, the individual (i) must be between 25 and 60 years old (the working lifespan) and (ii) have earnings above the minimum income threshold $Y_{\min,t}$, equivalent to earnings from one quarter of full-time work (13 weeks at 40 hours per week) at half of the legal minimum wage in year t (e.g., approximately $1885 in 2010).

- The revolving panel for year t then selects individuals that are admissible in $t-1$ and in at least two more years between $t-5$ and $t-2$.

- This ensures that the individual was participating in the labor market and we can compute a reasonable measure of average recent earnings.

Recent Earnings

- The average income of a worker i between years $t-1$ and $t-5$ is given by $\widehat{\boldsymbol{Y}}_{t-1}^{i} = \frac{1}{5}\sum_{j=1}^{5} \max\{\tilde{Y}_{t-j}^{i}, Y_{\min,t}\}$, where $\widetilde{\boldsymbol{Y}}_{t}^{i}$ denotes his earnings in year t.

- We then control for age and year effects by regressing $\widehat{\boldsymbol{Y}}_{t-1}^{i}$ on age dummies separately for each year, and define the residuals as *recent earnings* (hereafter RE), $\overline{\boldsymbol{Y}}_{t-1}^{i}$.

- In Sections 3 and 4, we will group individuals by age and by $\overline{\boldsymbol{Y}}_{t-1}^{i}$ to investigate how the properties of income dynamics vary over the life cycle and by income levels.

3. Cross-Sectional Moments of Earnings Growth

3.1. Empirical Methodology: A Graphical Construct

- Our main focus is on how the moments of earnings growth vary with recent earnings and age.

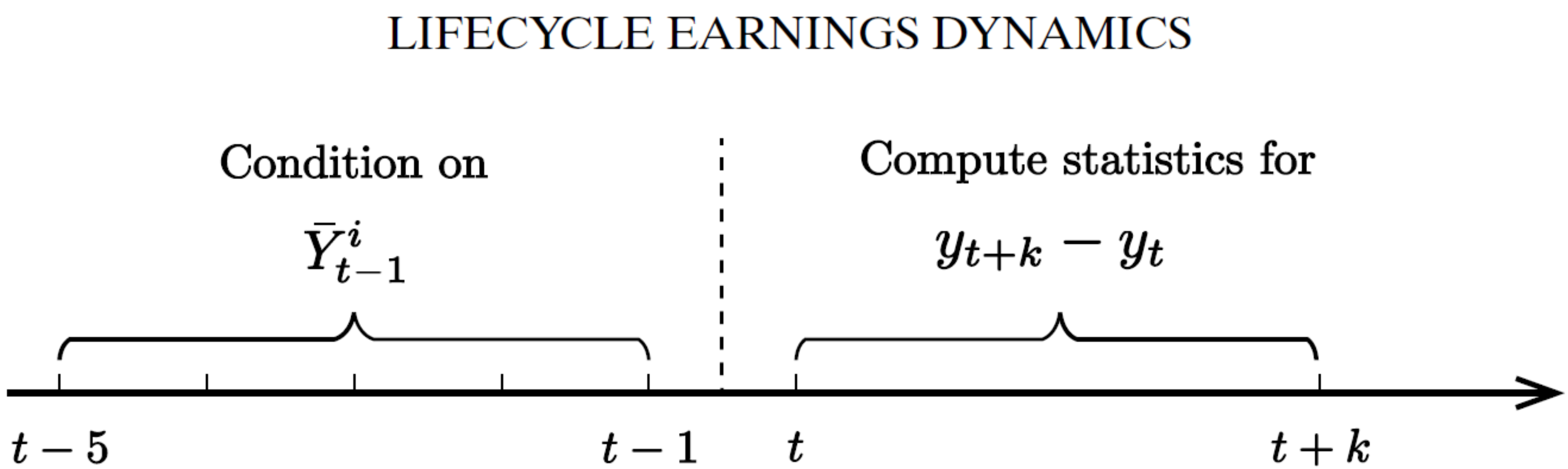
- To this end, for each year t, we divide individuals into six groups based on their age in $t-1 (25\text{–}29, \ldots, 45\text{–}54)$, and then within each age group, sort individuals into 100 percentile groups by their recent earnings $\bar{Y}_{t-1}^{i}$.

- If these groupings are done at a sufficiently fine level, we can think of all individuals within a given age/RE group to be ex ante identical (or at least very similar).

- Then, for each such group, the cross-sectional moments of earnings growth between t and $t+k$ can be viewed as the properties of earnings uncertainty that workers within that group expect to face looking ahead (see Figure 2).

- In our figures, we plot the average of these moments for each age/RE group over the years between 1997 and 2013-k.

Figure 2. Timeline for rolling panel construction.



Growth Rate Measures

- The first measure of income change that we use is log growth rate of income between t and $t+k$, $\Delta^k_{\log} y^i_t \equiv y^i_{t+k} - y^i_t$, where $y^i_t = \tilde{y}^i_t - d_{t,h(i,t)}$ denote the log income ($\tilde{y}^i_t$) of individual i in year t at age $h(i,t)$ net of age and year effects $d_{t,h(i,t)}$. $\{d_{t,h}\}^{60}_{h=25}$ are obtained by regressing $\tilde{y}^i_t$ on a full set of age dummies separately in each year.

- While its familiarity makes the log change a good choice for the descriptive analysis, it has a well-known drawback that observations close to zero need to be dropped or winsorized at an arbitrary value.

- When we use $\Delta^k_{\log} y^i_t$, we drop individuals from the sample with earnings less than $Y_{\min}$ in t or $t+k$, and lose information in the extensive margin.

Growth Rate Measures, Cont'd

- Our second measure of income growth—arc-percent change—is not prone to this caveat and is commonly used in the firm-dynamics literature, where firm entry and exit are key margins.

- We define $\Delta_{\text{arc}}^{k} Y_t^i = \frac{Y_{t+k}^i - Y_t^i}{(Y_{t+k}^i + Y_t^i)/2}$, where earnings level $Y_t^i = \frac{\bar{\boldsymbol{Y}}_{\boldsymbol{t}}^{\boldsymbol{i}}}{\widetilde{\boldsymbol{d}}_{\boldsymbol{t},\boldsymbol{h}(\boldsymbol{i},\boldsymbol{t})}}$ is net of average earnings in age h and year t, $\widetilde{\boldsymbol{d}}_{t,h(i,t)}$.

Transitory versus Persistent Income Changes

- As is well understood, longer-term earnings changes (i.e., $\Delta_{\log}^{k} y_t^i$ with larger k) reflect more persistent innovations.

- To see this intuition, consider the commonly used random-walk permanent/transitory model in which permanent (η_t^i) and transitory (ε_t^i) innovations are drawn from distributions F_η and F_ε, respectively.

- We 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.

Transitory versus Persistent Income Changes, Cont'd

- Then the second to fourth moments of k-year log income growth $\Delta_{\log}^k y_t^i$ are given by:

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

$$\mathcal{S}(\Delta_{\log}^k y_t^i) = \underbrace{\frac{k \times \sigma_\eta^3}{(k\sigma_\eta^2 + 2\sigma_\varepsilon^2)^{3/2}}}_{<1} \mathcal{S}_\eta, \tag{1}$$

$$\mathcal{K}(\Delta_{\log}^k y_t^i) = \underbrace{\frac{k \times \sigma_\eta^4}{(k\sigma_\eta^2 + 2\sigma_\varepsilon^2)^2}}_{<1} \mathcal{K}_\eta + \underbrace{\frac{2 \times \sigma_\varepsilon^4}{(k\sigma_\eta^2 + 2\sigma_\varepsilon^2)^2}}_{<1} \mathcal{K}_\varepsilon.$$

Transitory versus Persistent Income Changes, Cont'd

- Equation (1) shows that as k increases, the variance and kurtosis of k-year log change $\Delta_{\log}^{k} y_t^i$ reflect more of the distribution of η_t^i than that of ε_t^i.
- Also, skewness is solely driven by permanent changes.
- Finally, the distribution of $\Delta_{\log}^{k} y_t^i$ is closer to normal than the underlying distributions of F_η and F_ε, because as innovations η_t^i and ε_t^i accumulate, the distribution of $\Delta_{\log}^{k} y_t^i$ converges toward Gaussian, per the central limit theorem.

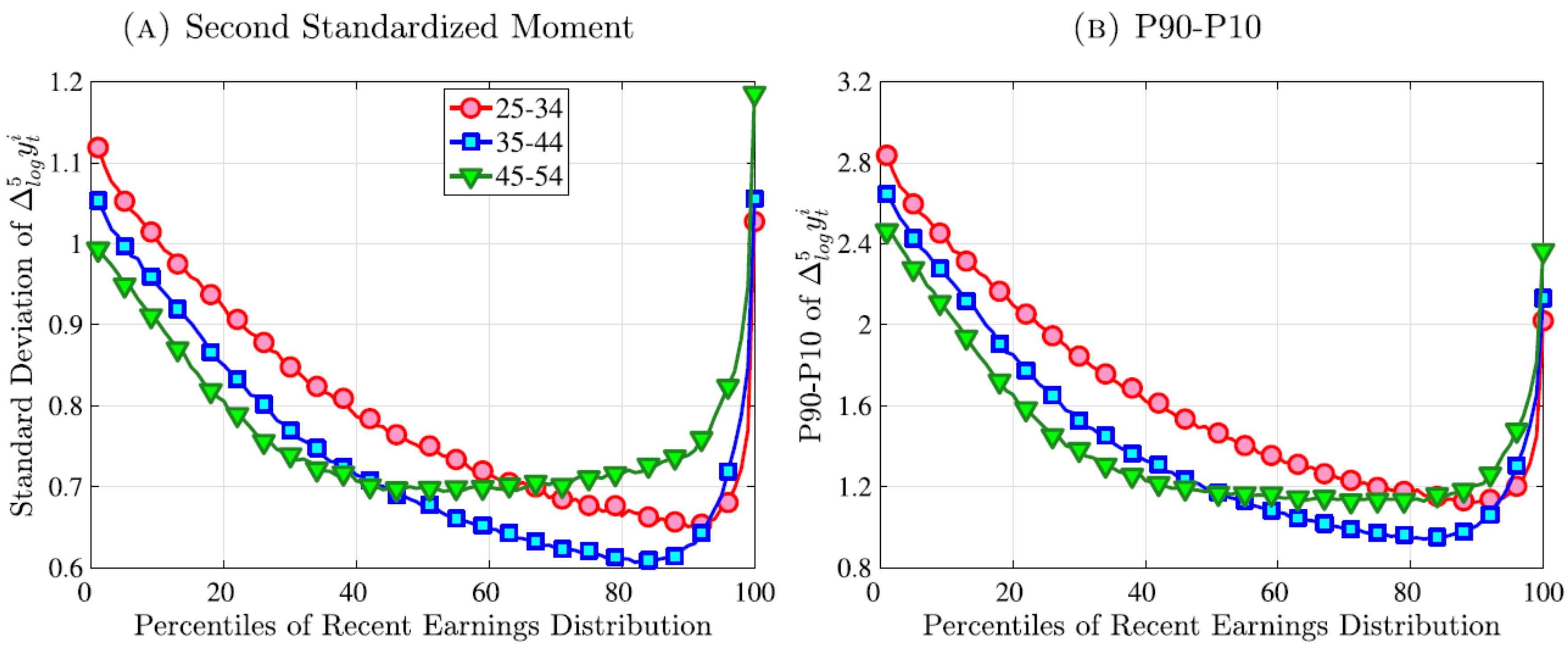
3.2. Second Moment: Standard Deviation

- Figure 3(a) plots the standard deviation of five-year residual earnings growth by age and recent earnings (for clarity, we use one marker for every fourth RE percentile group).

- In the right panel, we also report the difference between the 90th and 10th percentiles of log earnings changes, denoted by P90–P10, which is robust to outliers.

- Both measures show a pronounced U-shaped pattern by RE for every age group.

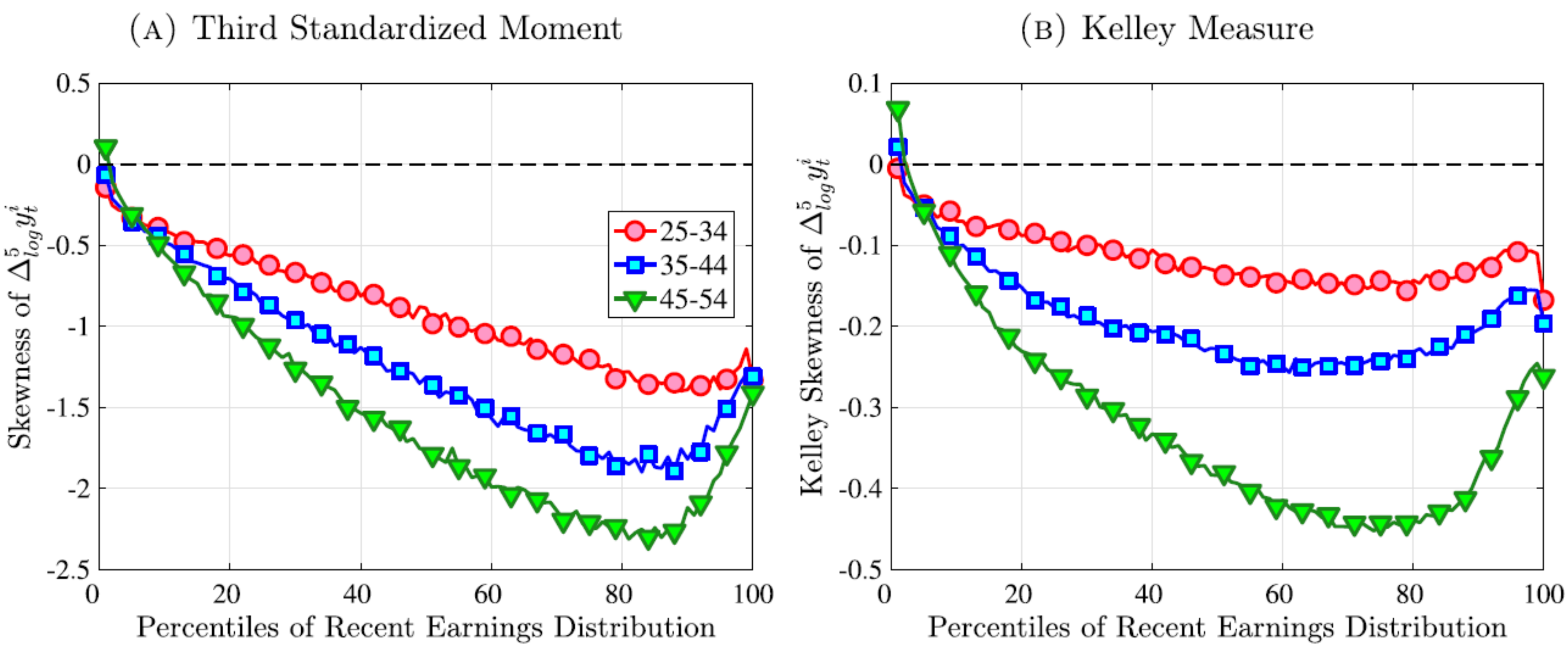
Figure 3. Dispersion of five-year log earnings growth.



3.3. Third Moment: Skewness (Asymmetry)

- Figure 4(a) plots the skewness of five-year earnings growth, measured as the third standardized moment.
- First, notice that earnings changes are negatively (left) skewed at every stage of the life cycle and for (almost) all earnings groups.
- Second, skewness is increasingly more negative for individuals with higher earnings and as individuals get older.
- Thus, it seems that the higher an individual's current earnings, the more room he has to fall and the less room he has left to move up.
- Note that the variation in skewness with age is more muted for individuals at the bottom or top of the (recent) earnings distribution (similar to the dispersion patterns above).

Figure 4. Skewness of five-year log earnings growth.



- Is negative skewness as measured by the third central moment driven by extreme observations?

- While the information on tails is important, we also look at Kelley (1947) skewness, $\mathcal{S}_{\mathcal{K}} = \frac{(\text{P90–P50})-(\text{P50–P10})}{\text{P90–P10}}$, which is robust to observations above the 90th or below the 10th percentile of the distribution.

- Basically, $\mathcal{S}_{\mathcal{K}}$ measures the relative fractions of the overall dispersion (P90–P10) accounted for by the upper and lower tails. Specifically, $\mathcal{S}_{\mathcal{K}}$ <0 implies that the lower tail (P50–P10) is longer than the upper tail (P90–P50).

- Kelley's skewness exhibits essentially the same pattern (Figure 4(b)).

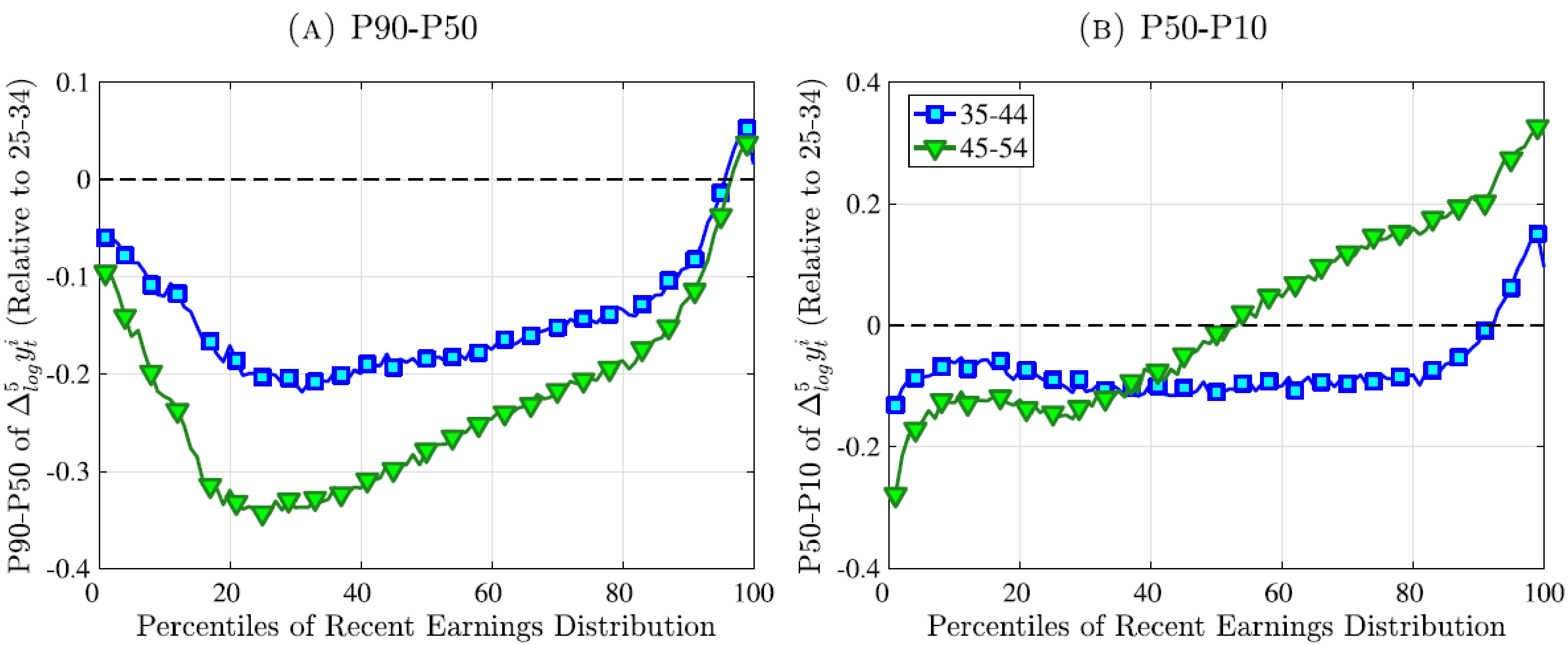
- Thus, the asymmetry is prevalent across the entire distribution rather than being driven just by the tails.

- Furthermore, the magnitudes are substantial.

- Another question is whether skewness becomes more negative over the life cycle because of a compression of the upper tail (fewer opportunities for large gains) or because of an expansion in the lower tail (higher risk of large declines).

- To answer this question, we investigate how the P90–P50 and P50–P10 change over the life cycle from their levels at ages 25–34 (Figure 5).

Figure 5. Skewness decomposed: P90–P50 and P50–P10 relative to age 25–34.



Notes: The y-axes show the change in P90–P50 and P50–P10 from the youngest age group to the two older age groups.

3.4. Fourth Moment: Kurtosis (Peakedness and Tailedness)

- We can think of kurtosis as a measure of the tendency of a density to stay away from $\mu \pm \sigma$.

- Thus, a leptokurtic distribution typically has a sharp/pointy center, long tails, and little mass near $\mu \pm \sigma$ (relative to a Gaussian distribution).

- A corollary to this description is that with excess kurtosis, the usual way we interpret standard deviation—as representing the size of the typical observation—is not very useful because most realizations will be either close to the center or out in the tails.

- To illustrate this point, we calculate concentration measures for earnings growth.

- Table I reports the fraction of individuals experiencing an absolute log earnings change less than a threshold, $|\Delta^1_{\log} y_t| \leq 0.05, 0.10$, and so on.

Table I. Fraction of Individuals within Selected Ranges of Annual Earnings Growth[a]

	Prob($\lvert \Delta^1_{\log} y_t \rvert \in S$)				
S:	≤ 0.05	≤ 0.10	≤ 0.20	$\geq 2\sigma (\approx 1.0)$	$\geq 3\sigma (\approx 1.5)$
Data	30.6	48.8	66.5	6.64	2.37
$\mathcal{N}(0, 0.51)$	7.7	15.4	30.2	4.55	1.46
Ratio	3.88	3.27	2.23	1.46	8.77

[a]Notes: The empirical distribution used in this calculation is for 1997–1998, the same as in Figure 1.

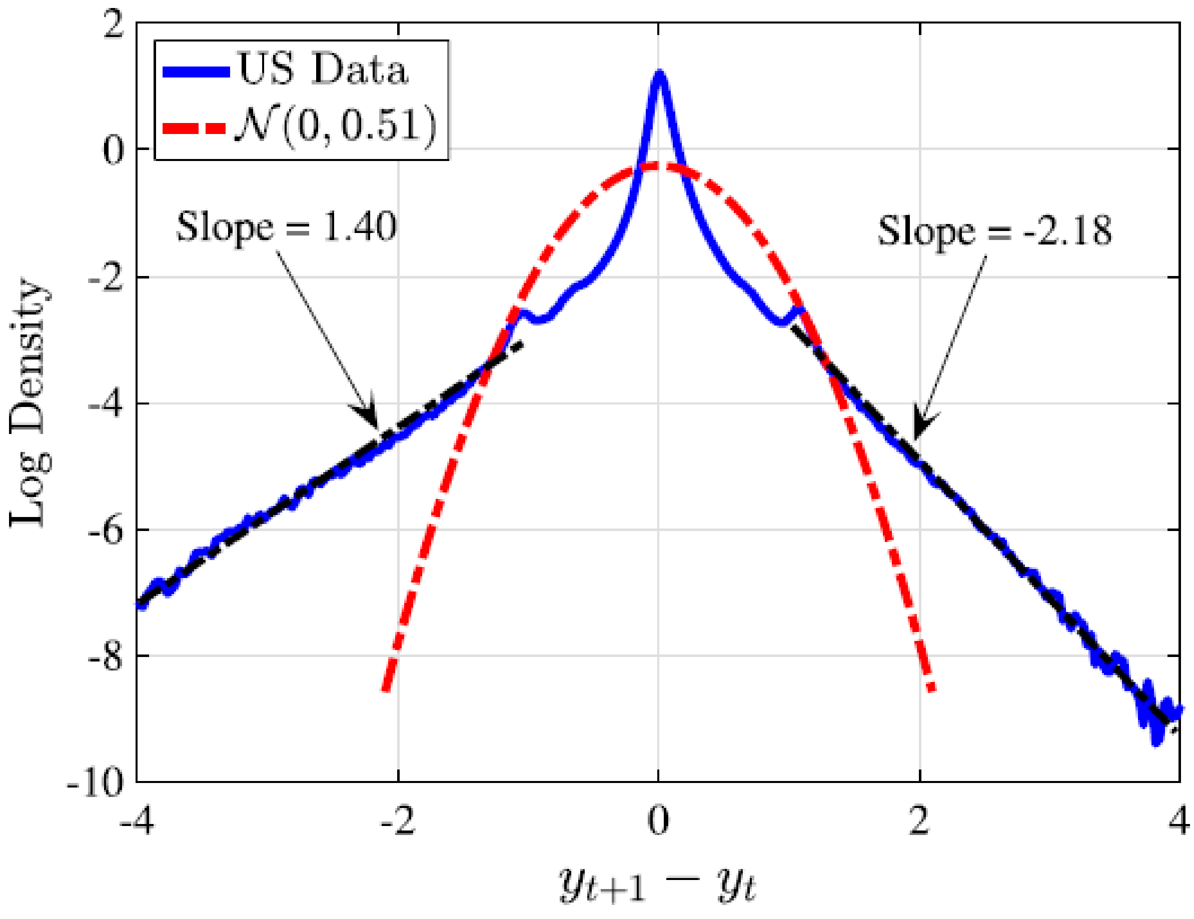
- The high likelihood of extreme events in the data motivates us to take a closer look at the tails of the earnings growth distribution by examining its empirical log density versus the Gaussian log density (which is an exact quadratic).

- First, in line with our previous discussion, the data have much thicker and longer tails compared with a normal distribution (Figure 6).

- Second, the tails decline almost linearly, implying a Pareto distribution at both ends.

- Third, they are asymmetric, with the left tail declining much more slowly than the right, which contributes to the left-skewness documented above.

Figure 6. Double-Pareto tails of the U.S. annual earnings growth distribution.

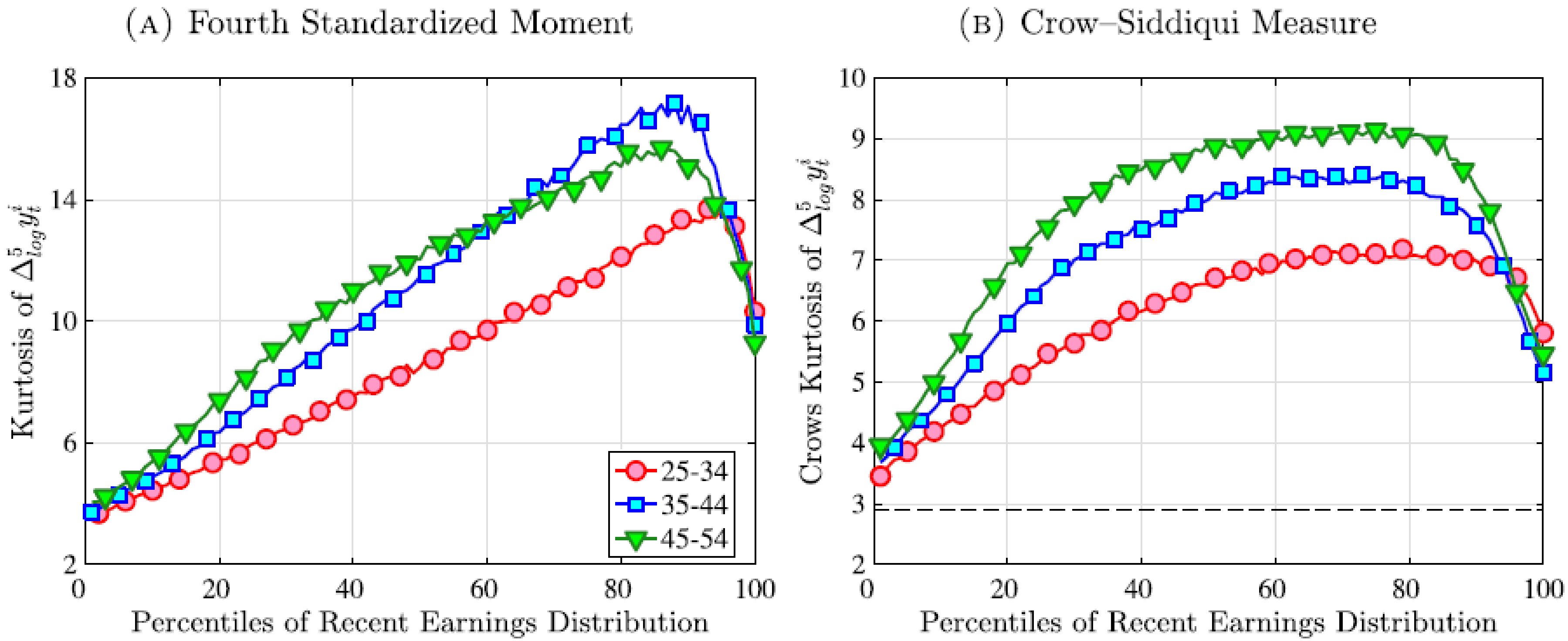


Notes: The empirical distribution in this figure is for 1997–1998, the same as in Figure 1 but with the y-axis now in logs.

- Next, to see how kurtosis varies by age and income, we report two statistics in Figure 7 that are analogous to the ones we used for skewness: the fourth standardized moment and the quantile-based Crow and Siddiqui (1967) measure, which is defined as $\kappa_{C-S} = \frac{\text{P97.5}-\text{P2.5}}{\text{P75}-\text{P25}}$ and is equal to 2.91 for a Gaussian distribution.

- As with dispersion and skewness, kurtosis varies substantially with age and recent earnings.

Figure 7. Kurtosis of five-year log earnings growth.



An Alternative Measure of Persistent Changes

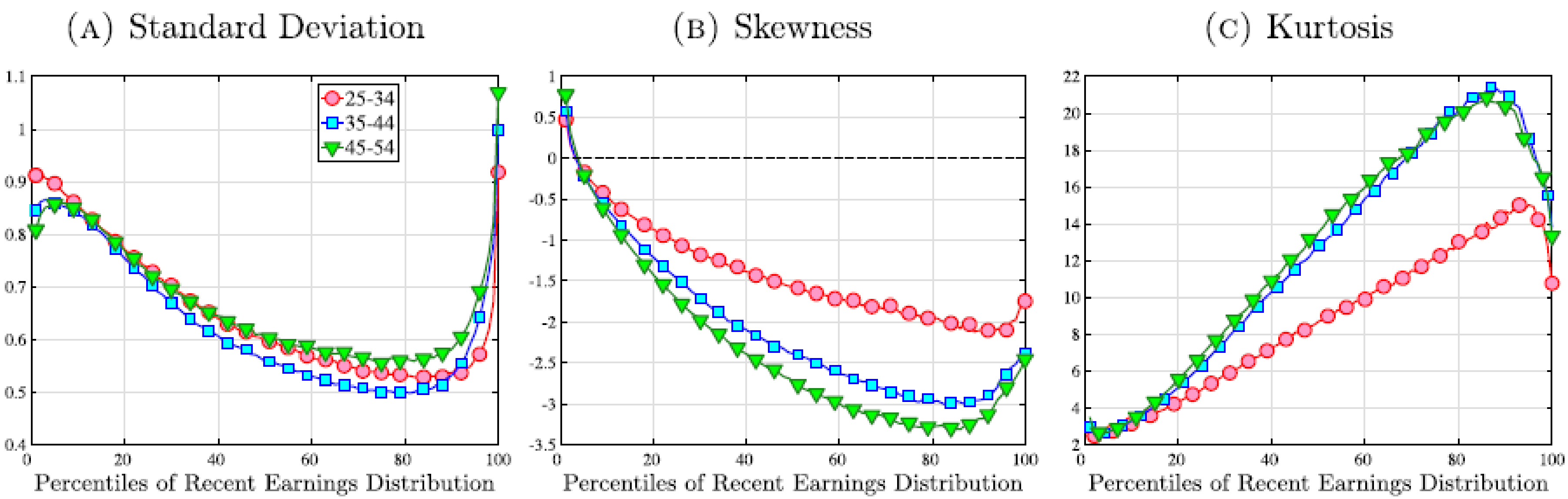
- As we noted earlier, while the five-year income growth measure reveals a good deal about persistent changes in earnings, it still contains possible transitory innovations in years t and $t+5$, which can potentially confound the inferences we draw about persistent changes.

- To check the robustness of our results, we consider an alternative measure that is based on the change between two consecutive five-year averages of earnings: $\bar{\Delta}^5_{\log}\left(\bar{y}^i_t\right) \equiv \log \bar{Y}^i_{t+4} - \log(\bar{Y}^i_{t-1})$, where $\bar{Y}^i_{t+4}$ is calculated the same as $\bar{Y}^i_{t-1}$ but over the period t to $t+4$.

- Averaging earnings before differencing purges transitory changes and better isolates the persistent ones.

- Figure 8 plots the standardized moments of this alternative measure, which show essentially identical patterns to their counterparts using our baseline five-year growth measure.
- In fact, if anything, this measure shows a slightly larger negative skewness and a higher excess kurtosis.
- These results confirm our conclusion that the nonnormalities are stronger in persistent earnings changes.

Figure 8. Alternative measure of persistent changes $\overline{\Delta}^5_{\log}(\bar{y}^i_t)$: standardized moments.



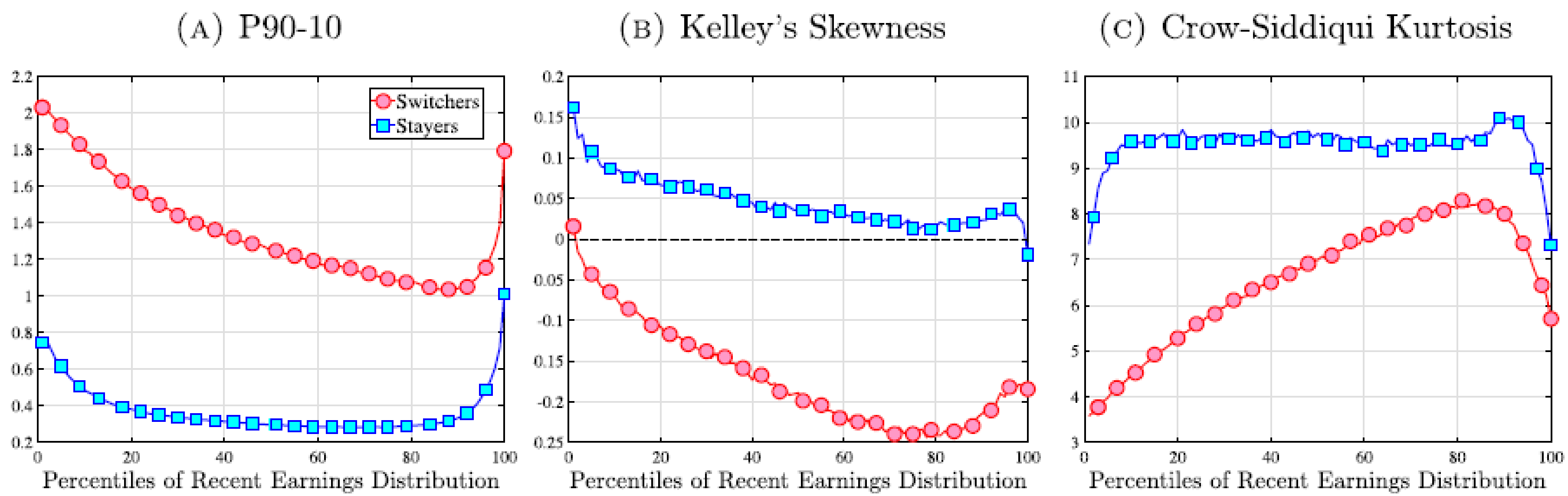
3.5. Job-Stayers and Job-Switchers

- We show in Figure 9 how the quantile-based second to fourth moments of annual earnings growth for stayers and switchers vary with recent earnings.

- Relative to job-switchers, job-stayers experience earnings changes that have a smaller dispersion (about one-third for median-income workers), and are more leptokurtic, especially for low-RE workers.

- Changes are symmetric or slightly right skewed for stayers and left skewed for switchers.

Figure 9. Higher-order moments of earnings growth, $\Delta^1_{\log} y^i_t$: stayers vs. switchers.



3.6. What Are the Sources of Nonnormalities in Earnings Growth?

- For many economic questions, it is important to know the extent to which nonnormalities in earnings dynamics are driven by wages versus hours.

- We start by investigating the non-Gaussian features of two-year earnings changes in the PSID (Table II).

- The standardized third moment and the Kelley measure point to a weakly left-skewed distribution, possibly due to added noise in the PSID to the extent that measurement error is symmetric.

- Excess kurtosis is a more striking feature: Both measures of kurtosis from the PSID are quite close to their SSA counterparts.

Table II. Higher-Order Moments of Two-Year Changes in the PSID[a]

		All		25–39		40–55	
	Normal	Earnings	Wages	Earnings	Wages	Earnings	Wages
Skewness	0.0	−0.26	−0.14	−0.17	−0.20	−0.34	−0.09
Kelley Skew.	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 Kurt.	2.91	6.83	5.59	6.33	5.02	7.33	6.11

[a]Note: Wages are obtained by dividing annual earnings of male heads of households by their annual hours in the PSID using data over the period 1999–2013, during which data are biennial.

- Motivated by the importance of extreme earnings changes for excess kurtosis, we investigate the roles of hours and wages in the tails of the earnings growth distribution.

- For this purpose, we distribute workers into six groups based on their two-year residual earnings change.

- As in the SSA data, most workers experience only small earnings changes (col. 1 of Table III).

- For each group, we compute the average change in residual earnings, hours, and wages (Table III, cols. 2–4).

- Our results show that wage changes are at least as important as hours changes.

- Moreover, wage changes seem to be even more important for smaller earnings changes (e.g., more than 70% of $|\Delta y| < 0.25$ can be attributed to wages).

Table III. Important Lifecycle Events and Earnings Changes[a]

Group $\Delta y \in$	Share % (1)	Mean Δy (2)	Mean Δw (3)	Mean Δh (4)	Δ wks not empl. (5)	Occup. switch % (6)	Employer switch % (7)	Disab. Flow in % (8)
$(-\infty, -1)$	3.8%	−1.65	−1.01	−0.64	10.01	26.1	45.6	9.2
$[-1, -0.25)$	14.4%	−0.48	−0.34	−0.14	1.62	14.9	29.1	4.4
$[-0.25, 0)$	31.2%	−0.11	−0.08	−0.03	0.17	6.9	13.3	3.5
$[0, 0.25)$	31.1%	0.11	0.08	0.03	−0.03	5.3	9.7	2.8
$[0.25, 1)$	16.5%	0.47	0.34	0.13	−1.30	8.6	16.9	2.9
$(1, \infty)$	3.0%	1.64	1.06	0.58	−7.51	18.0	30.7	3.8

[a]Notes: This table shows hours and wage growth (Δh and Δw, respectively) and the various lifecycle events for people in different biennial earnings change (Δy) groups. In column 5, "weeks not employed" is the sum of weeks unemployed and out of the labor force. Columns 6 and 7 show the fraction of workers that switch occupation and employer within each earnings change group, respectively. Column 8 shows the fraction of workers who become disabled in that period.

- We link large earnings changes to various lifecycle events.
- We start with a natural suspect: nonemployment spells.
- The group with the largest earnings decline also reports the largest increase in the incidence of nonemployment—10 weeks (Table III, col. 5).
- Similarly, the group with the largest earnings increase reports the largest decline in nonemployment.
- These results underline the importance of the extensive margin for the tails of the earnings change distribution.

- Next, we study occupation and job mobility, both of which are known to be associated with large changes in earnings.
- The likelihood of occupation and employer switches follows a distinct U-shaped pattern with earnings changes (Table III, cols. 6 and 7, respectively).
- Compared to the workers with small changes ($|\Delta y| < 0.25$), the top and bottom earnings-change groups are three to four times more likely to make these switches.
- The sources of mobility are possibly very different at the top and the bottom earnings-change groups.

- Finally, we investigate health shocks, which are known to have large effects on earnings.
- We focus on disabilities that affect individuals' work performance.
- We find higher transition rates into disability for workers with earnings declines, with the highest transition (9.2%) in the bottom earnings-change group (Table III, col. 8).
- These results suggest that the extreme earnings changes are not purely a statistical artifact or measurement error.

4. Dynamics of Earnings

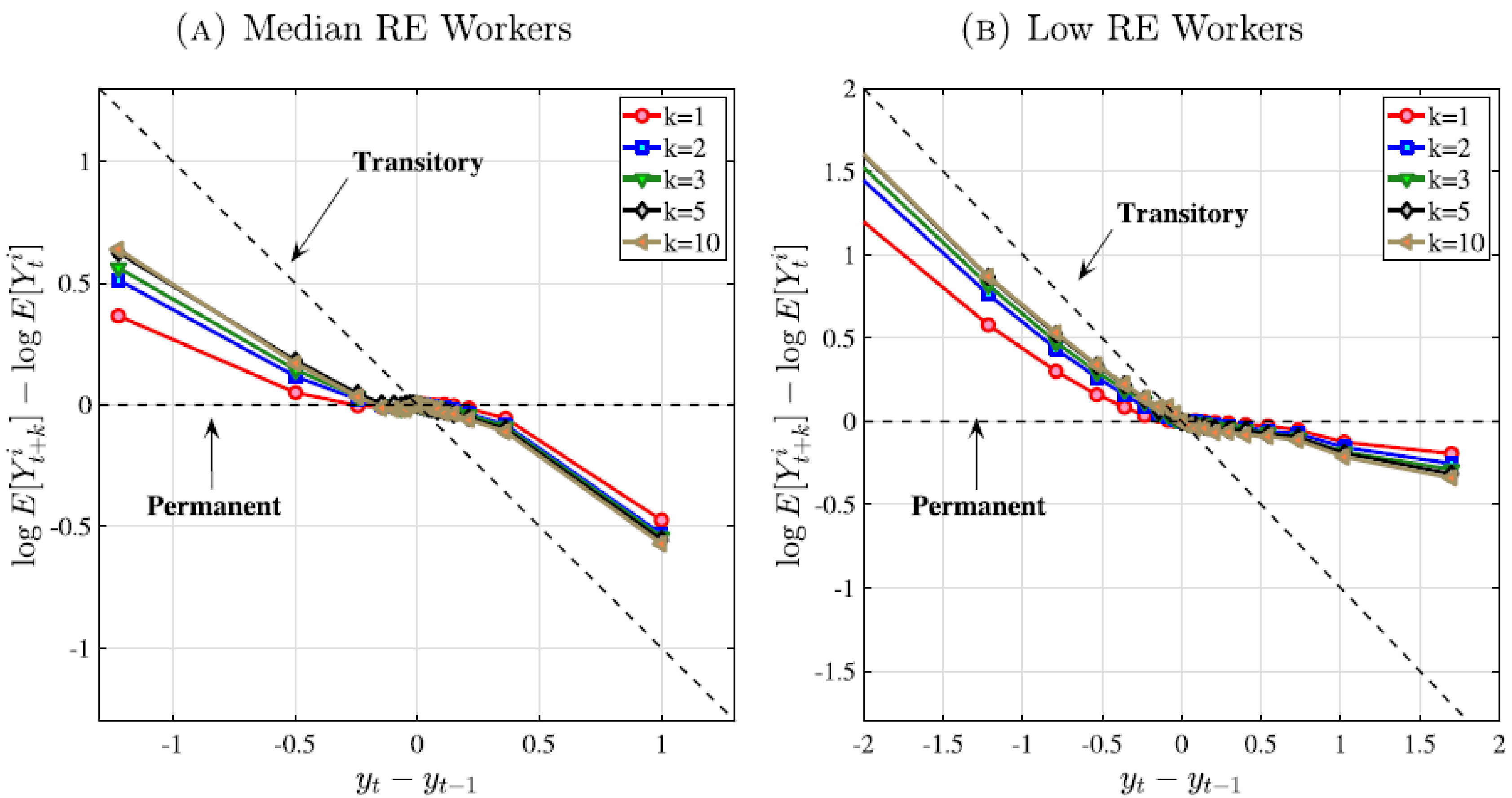
4.1. Impulse Response Functions Conditional on Recent Earnings

- In Figure 10, we show the mean reversion of different sizes of earnings changes $y_t^i - y_{t-1}^i$ for prime-age workers over a 10-year period.

- Specifically, we plot $\log \mathbb{E}\left[Y_{t+k}^i\right] - \log \mathbb{E}[Y_t^i]$ of each $y_t^i - y_{t-1}^i$ quantile on the y-axis against its average on the x-axis.

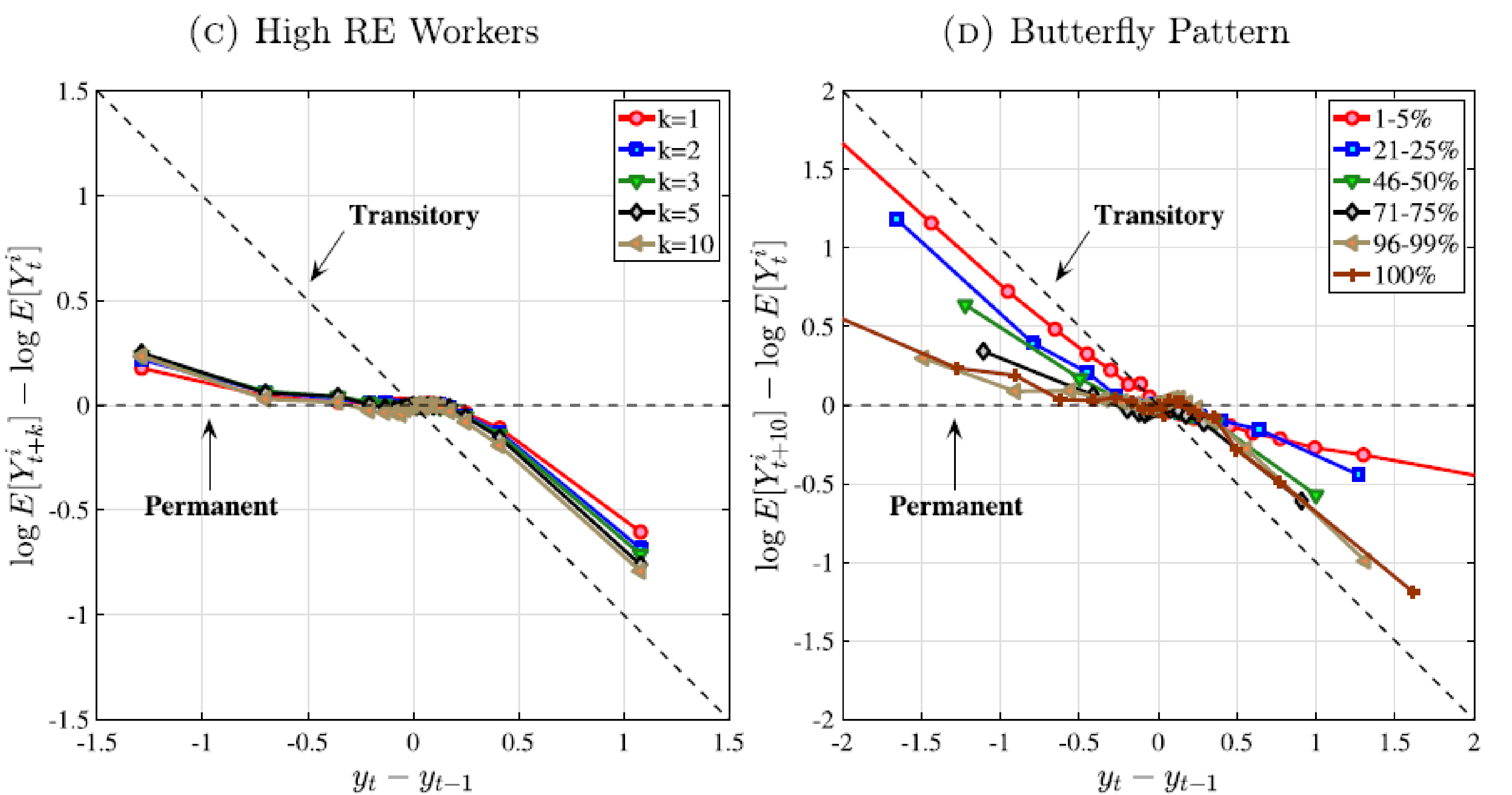
- This graphical construct contains the same information as a standard impulse response function but allows us to see the heterogeneous mean reversion patterns more clearly.

Figure 10. Impulse responses, prime-age workers.



Notes: Median-, low-, and high-RE in panels A, B, and C refer to workers with $\bar{Y}_{t-1}$ in (P46–P55), (P6–P10), and (P91–P95), respectively. Prime age refers to age 35 to 50.

Figure 10. Impulse responses, prime-age workers, cont'd.



Notes: Median-, low-, and high-RE in panels A, B, and C refer to workers with $\bar{Y}_{t-1}$ in (P46–P55), (P6–P10), and (P91–P95), respectively. Prime age refers to age 35 to 50.

- This butterfly pattern broadly resonates with the earnings dynamics in job ladder models.

- For high-RE workers—who are at the higher rungs of the ladder—a job loss leads to a more persistent earnings decline relative to low-RE workers because of search frictions.

- Similarly, for low-RE workers, large increases are likely due to unemployment-to-employment or job-to-job transitions, which have long-lasting effects on earnings.

5. Earnings Growth and Employment: The Long View

- In this section, we turn to two questions that complete the picture of earnings dynamics over the life cycle.

- The first one is about average earnings growth: How much cumulative earnings growth do individuals experience over their working life, and how does that vary across individuals with different lifetime incomes?

- The second question investigates the lifetime nonemployment rate—defined as the fraction of an individual's working life spent as full-year nonemployed.

- Although the incidence of long-term nonemployment is of great interest for many questions in economics, documenting it requires long panel data with no sample attrition, a phenomenon most common among long-term nonemployed.

- The administrative nature of the MEF data set and its long panel dimension provide an ideal opportunity to study this question.

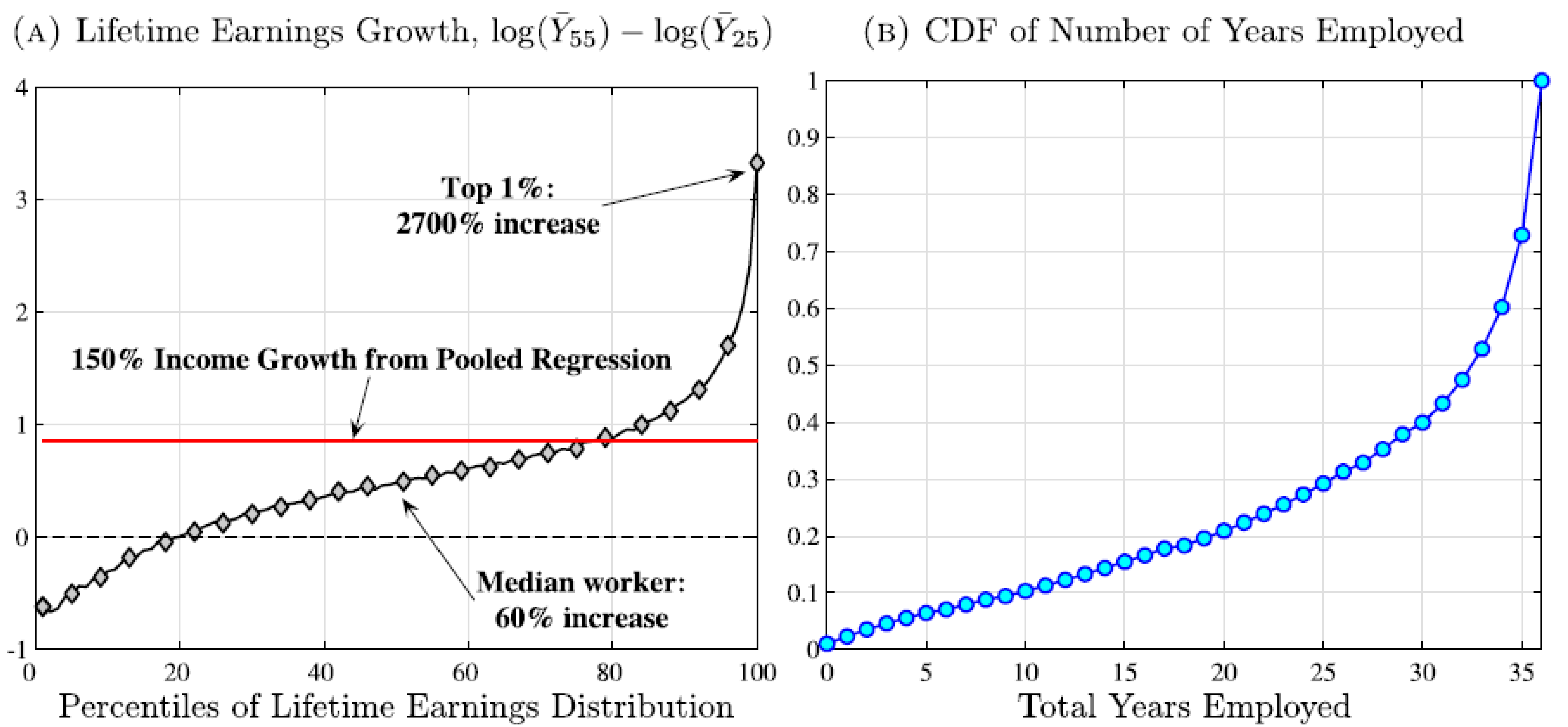
5.1. Lifecycle Earnings Growth and Its Distribution

- The results in Figure 11(a) show that between ages 25 and 55, the median worker (by LE) experiences a smaller earnings growth—about 60%—than a 150% mean growth estimated from a Deaton–Paxson pooled regression.

- More importantly, higher-LE workers experience a much higher earnings growth over the life cycle compared with the rest of the distribution.

- While an upward slope per se is not surprising (as it is partly mechanical—faster growth will deliver higher LE, everything else held constant), the variation at the top end is so large and the curvature is so steep, that it turns out to be difficult to capture using simple earnings processes, as we discuss in the next section.

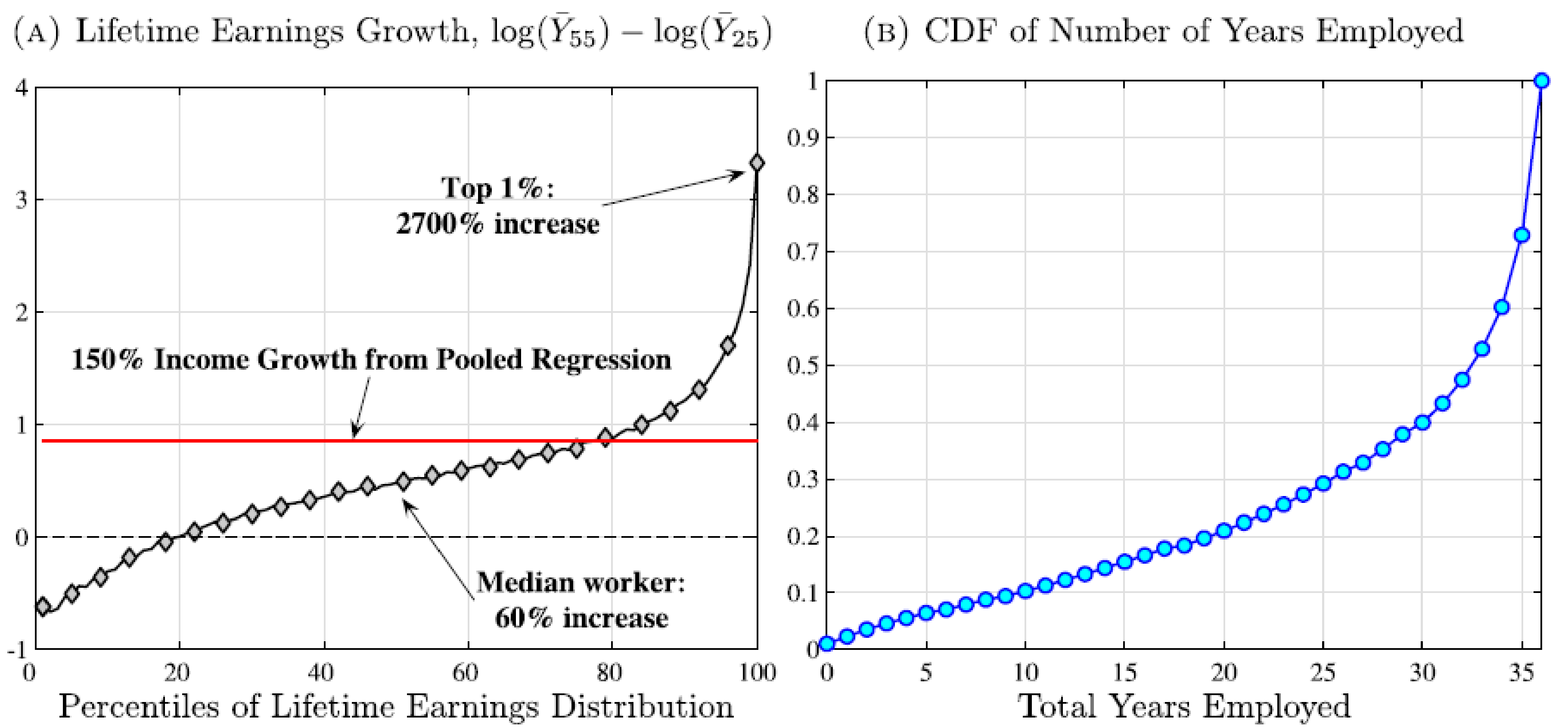
Figure 11. Earnings growth and employment: the long view.



5.2. Lifetime Employment Rate and Its Distribution

- Next, we investigate the lifetime nonemployment rates across individuals.
- Using the same criteria as before—working life defined as the period between ages 25 and 60, and full-year nonemployed defined as annual earnings below $Y_{\min}$—we examine the cumulative distribution of total lifetime years employed in Figure 11(b).
- The results show that, first, a large fraction of individuals are very strongly attached to the labor market: 28% of individuals were never nonemployed during their working life, and almost half (48%) were nonemployed for less than three years.
- But second, the distribution has a long left tail, showing a surprisingly large fraction of men who spend half of their working life or more without employment: 18.3% of men spend 18 years—or half of their working life—as full-year nonemployed, and 12.3% spend at least 24 years as nonemployed.

Figure 11. Earnings growth and employment: the long view.



6. Econometric Models for Earnings Dynamics

6.1. A Flexible Stochastic Process

- The models we estimate are special cases of the following general framework, which includes:

  (i) an AR(1) process (z_t^i) with innovations drawn from a mixture of normals;

  (ii) a nonemployment shock whose incidence probability ($p_\nu^i(t, z_t)$) can vary with age or z_t or both, and whose duration (ν_t^i) is exponentially distributed;

  (iii) a heterogeneous income profiles component (HIP); and

  (iv) an i.i.d. normal mixture transitory shock (ε_t^i).

$$\text{Level of earnings:} \quad \tilde{Y}_t^i = (1 - \nu_t^i) e^{(g(t) + \alpha^i + \beta^i t + z_t^i + \varepsilon_t^i)}, \tag{2}$$

$$\text{Persistent component:} \quad z_t^i = \rho z_{t-1}^i + \eta_t^i, \tag{3}$$

$$\text{Innovations to AR(1):} \quad \eta_t^i \sim \begin{cases} \mathcal{N}(\mu_{\eta,1}, \sigma_{\eta,1}) & \text{with prob. } p_z, \\ \mathcal{N}(\mu_{\eta,2}, \sigma_{\eta,2}) & \text{with prob. } 1 - p_z, \end{cases} \tag{4}$$

$$\text{Initial condition of } z_t^i\text{:} \quad z_0^i \sim \mathcal{N}(0, \sigma_{z_0}), \tag{5}$$

$$\text{Transitory shock:} \quad \varepsilon_t^i \sim \begin{cases} \mathcal{N}(\mu_{\varepsilon,1}, \sigma_{\varepsilon,1}) & \text{with prob. } p_\varepsilon, \\ \mathcal{N}(\mu_{\varepsilon,2}, \sigma_{\varepsilon,2}) & \text{with prob. } 1 - p_\varepsilon, \end{cases} \tag{6}$$

$$\text{Nonemployment duration:} \quad \nu_t^i \sim \begin{cases} 0 & \text{with prob. } 1 - p_\nu(t, z_t^i), \\ \min\{1, \exp(\lambda)\} & \text{with prob. } p_\nu(t, z_t^i), \end{cases} \tag{7}$$

$$\text{Prob of Nonemp. shock:} \quad p_\nu^i(t, z_t) = \frac{e^{\xi_t^i}}{1 + e^{\xi_t^i}}, \quad \text{where } \xi_t^i \equiv a + bt + cz_t^i + dz_t^i t. \tag{8}$$

- In equation (2), $g(t)$ is a quadratic polynomial, where $t = (\text{age} - 24)/10$ is normalized age, that captures the lifecycle profile of earnings common to all individuals.

- The random vector (α^i, β^i) determines ex ante heterogeneity in the level and in the growth rate of earnings and is drawn from a multivariate normal distribution with zero mean and a covariance matrix to be estimated.

- The innovations, η_t^i, to the AR(1) component are drawn from a mixture of two normals.

- An individual draws a shock from $\mathcal{N}(\mu_{\eta,1}, \sigma_{\eta,1})$ with probability p_z and otherwise from $\mathcal{N}(\mu_{\eta,1}, \sigma_{\eta,2})$.

- Without loss of generality, we normalize η to have zero mean (i.e., $\mu_{\eta,1} p_z + \mu_{\eta,2}(1 - p_z) = 0$) and assume $\mu_{\eta,1} < 0$ for identification.

- Heterogeneity in the initial conditions of z_t is captured by $z_0^i \sim \mathcal{N}(0, \sigma_{z_0})$.

- Transitory shocks, ε_t^i, are also drawn from a mixture of two normals (eq. (6)), with analogous identifying assumptions (zero mean and $\mu_{\varepsilon,1} < 0$).

- The last component of the earnings process—and as it turns out, a critical one—is a nonemployment shock (eq. (7)) that is intended to primarily capture movements in the extensive margin.

- Specifically, a worker is hit with a nonemployment shock with probability p_v whose duration $\nu_t > 0$ follows an exponential distribution with mean $1/\lambda$ and is truncated at 1 (corresponding to full-year nonemployment with zero annual income).

- This shock differs from z_t and ε_t by scaling the level of annual income—not its logarithm—which allows the process to capture the sizable fraction of workers who transition into and out of full-year nonemployment every year.

- None of the components introduced so far depend explicitly on age or recent earnings, whereas variation along these dimensions is a key characteristic of the empirical patterns we saw.

- One promising way we found for introducing such variation was by making the nonemployment incidence p_ν depend on age t and z_t through the logistic function shown in equation (8).

- The dependence of p_ν on z_t—which we refer to as "state dependence"—turns out to be especially important as it induces persistence in nonemployment from one year to the next (despite ν_t itself being independent over time).

6.2. Results: Estimates of Stochastic Processes

- We now present the estimation results for six different specifications (Table IV).
- We start from the canonical linear-Gaussian model and add new features step by step until we reach our preferred benchmark process.
- Figure 12 plots the fit of each model to the six sets of moments targeted in the estimation.
- We also show the fit to selected impulse response functions separately in Figure 13.

Table IV. Estimates of Stochastic Process Parameters[a]

Model:	(1)	(2)	(3)	(4)	(5)	(6)	
						Benchmark process	
	Gaussian process					Parameters	Std. Err.
AR(1) *Component*	G	G	mix	mix	mix	mix	mix
↪*Prob. age/income*	—	—	no/no	yes/yes	no/no	no/no	no/no
Nonemployment shocks	no	yes	no	no	yes	yes	yes
↪*Prob. 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
Parameters							
ρ	1.005	0.967	1.010	0.992	0.991	0.959	0.0001
p_z			5.0%	—†	17.6%	40.7%	0.0005
$\mu_{\eta,1}$			−1.0*	−1.0*	−0.524	−0.085	0.0006
$\sigma_{\eta,1}$	0.134	0.197	1.421	1.070	0.113	0.364	0.0004
$\sigma_{\eta,2}$			0.010	0.032	0.046	0.069	0.0002
σ_{z_0}	0.343	0.563	0.213	0.446	0.450	0.714	0.0005
λ		0.030			0.016	0.0001	0.0003
p_ε			11.8%	8.8%	4.4%	13.0%	0.0004
$\mu_{\varepsilon,1}$			−0.826	0.311	0.134	0.271	0.0009
$\sigma_{\varepsilon,1}$	0.696	0.163	1.549	0.795	0.762	0.285	0.0006
$\sigma_{\varepsilon,2}$			0.020	0.020	0.055	0.037	0.0003
σ_α	1.182	0.655	0.273	0.473	0.472	0.300	0.0009
σ_β						0.196	0.0002
$\text{corr}_{\alpha\beta}$						0.768	0.0015

(continues)

Table IV. Estimates of Stochastic Process Parameters,[a] Cont'd

Objective value	74.87	35.69	56.66	40.97	27.22	22.64
Decomposition:						
(i) Standard deviation	9.48	6.28	7.66	6.11	5.65	5.77
(ii) Skewness	43.03	14.12	23.75	15.02	14.12	9.96
(iii) Kurtosis	26.53	5.90	12.95	9.22	5.83	5.81
(iv) Impulse resp. short	18.35	13.51	20.04	16.70	9.85	8.65
(v) Impulse resp. long	28.34	22.87	32.86	27.00	16.27	12.12
(vi) Lifetime inc. growth	37.12	10.33	24.21	11.82	7.96	6.89
(vii) Age-ineq. profile	20.70	8.24	16.42	12.78	1.52	3.32
(viii) Nonempl. CDF	3.63	10.97	9.05	4.42	6.95	8.13
Model Selection p-val.						
Test 1	0.000	0.000	0.000	0.000	0.000	0.000
Test 2	0.000	0.000	0.000	0.000	0.000	—

[a]Notes: The top panel provides a summary of the features of each specification, the middle panel shows the estimated values of key parameters (the rest are reported in Table D.III), and the bottom panel reports the weighted percentage deviation between the data and simulated moments for each set of moments (the total objective value is the square root of the sum of the squares of objective values of each component) as well as the p-values for model selection. The *'s indicate that in columns 3 and 4, the value of $\mu_{\eta,1}$ is constrained by the lower bound we impose in the estimation. †: p_2 is not a number but a function in this specification and reported in Table D.III. The standard errors of parameter estimates (using a parametric bootstrap with 100 repetitions) are extremely small, thanks to the very large sample size. Hence, we do not report them except for the benchmark process.

Figure 12. Estimated model versus data: key moments.

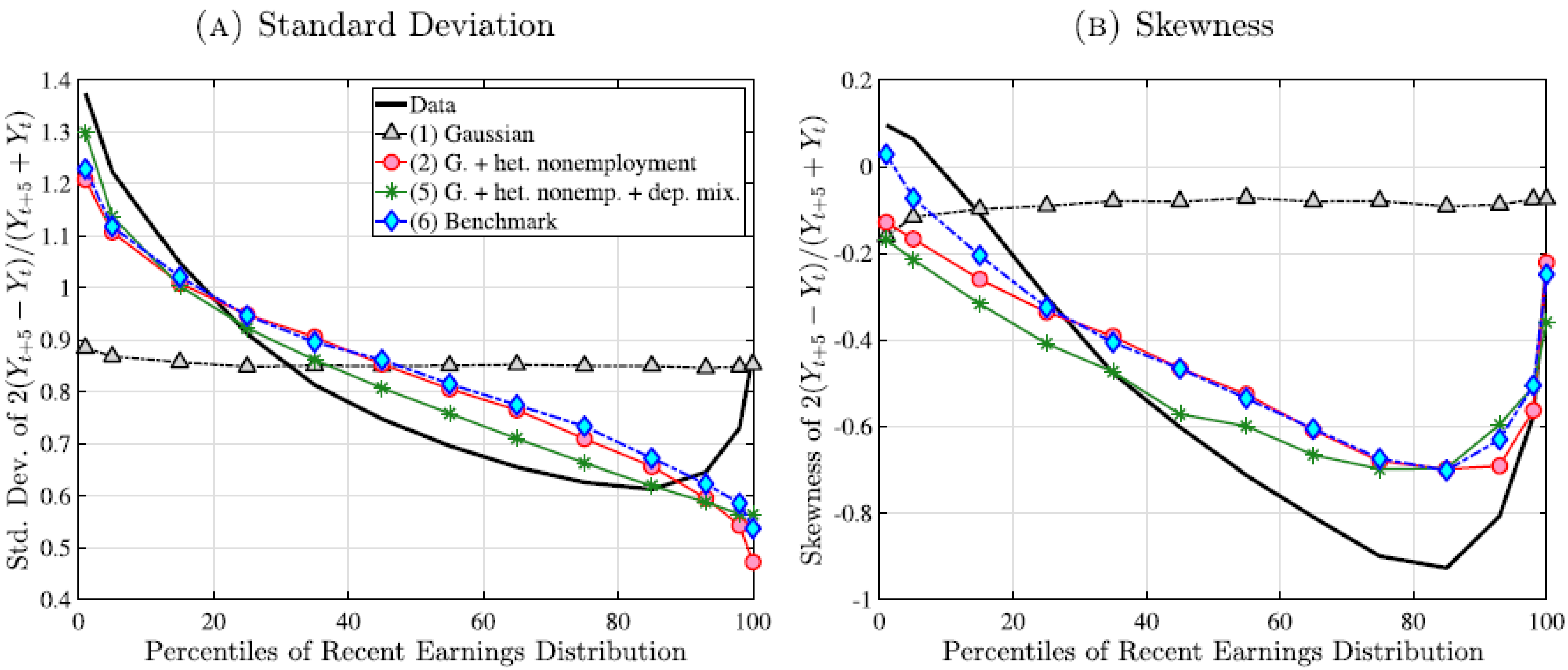


Figure 12. Estimated model versus data: key moments, cont'd.

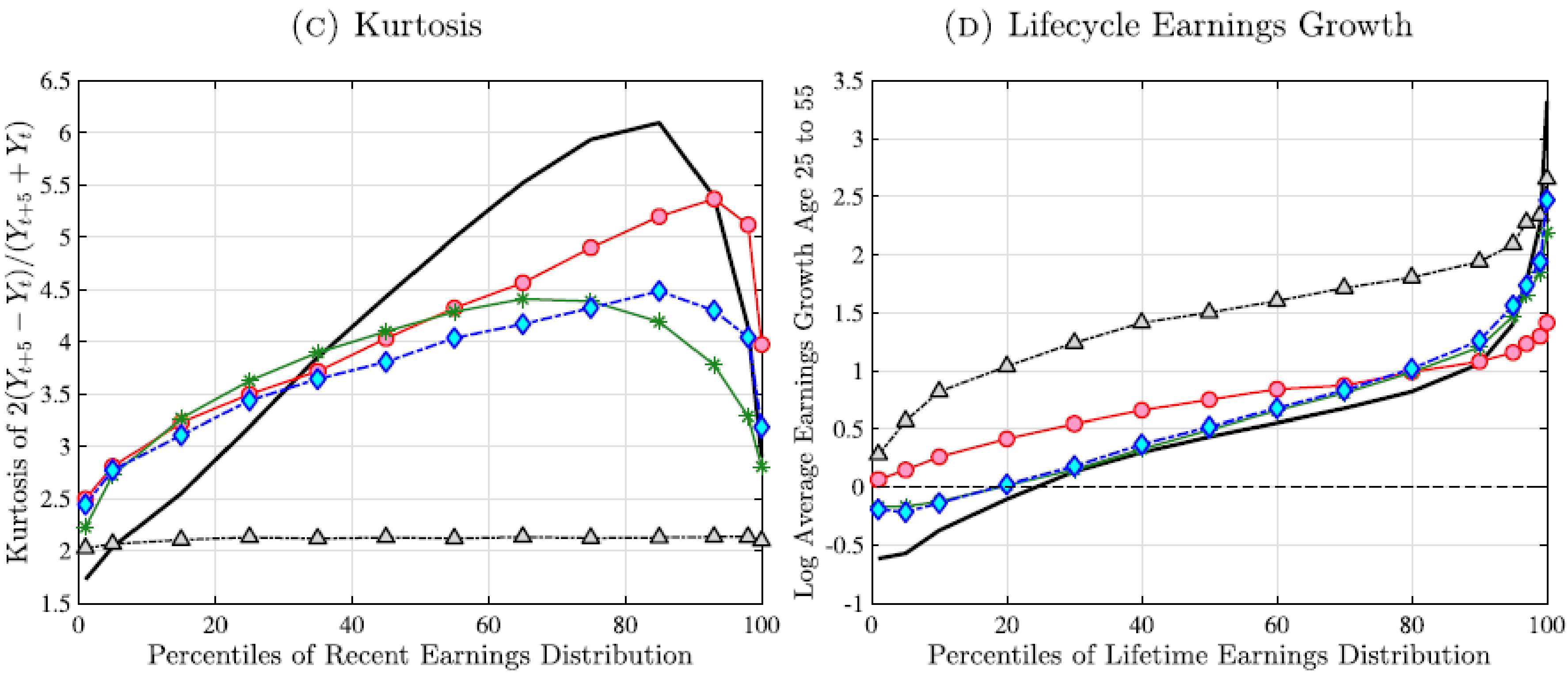


Figure 12. Estimated model versus data: key moments, cont'd.

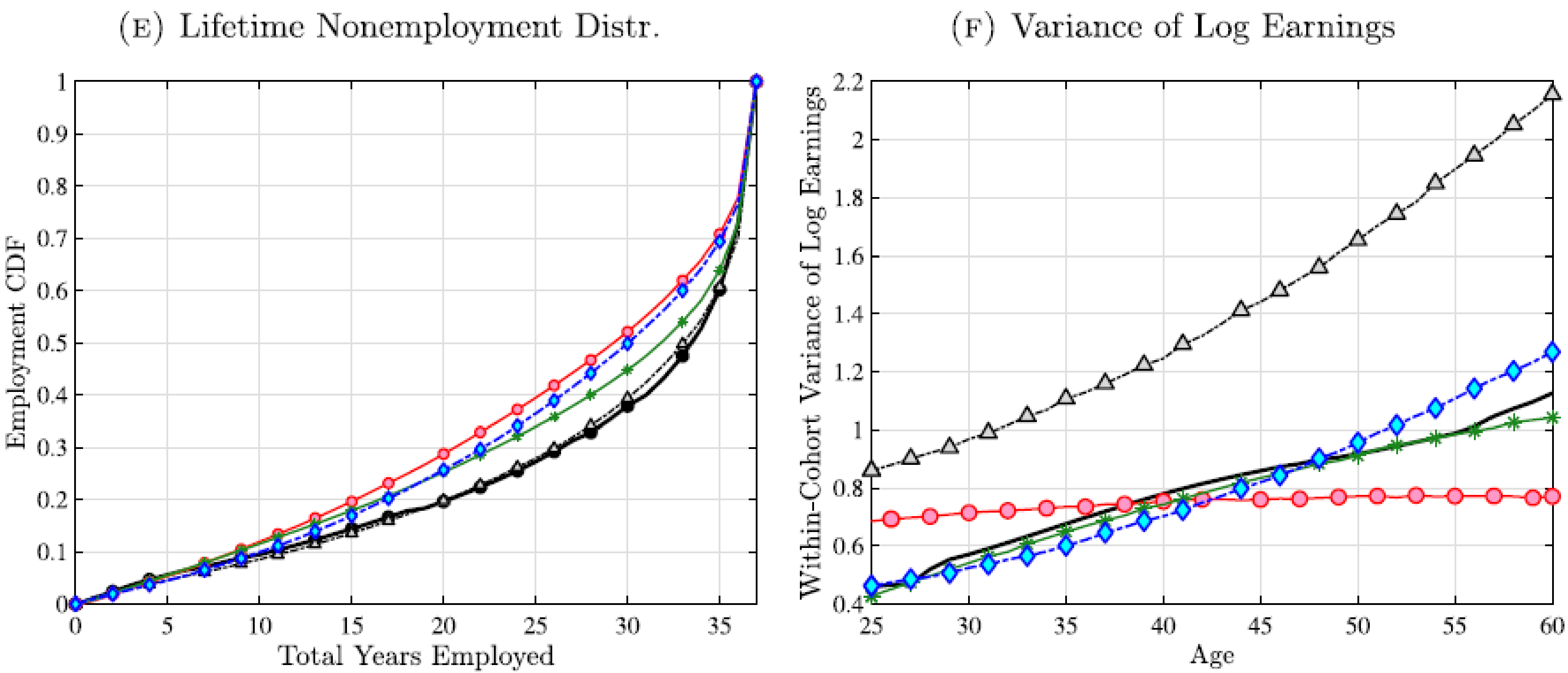


Figure 13. Estimated model versus data: selected impulse response moments.

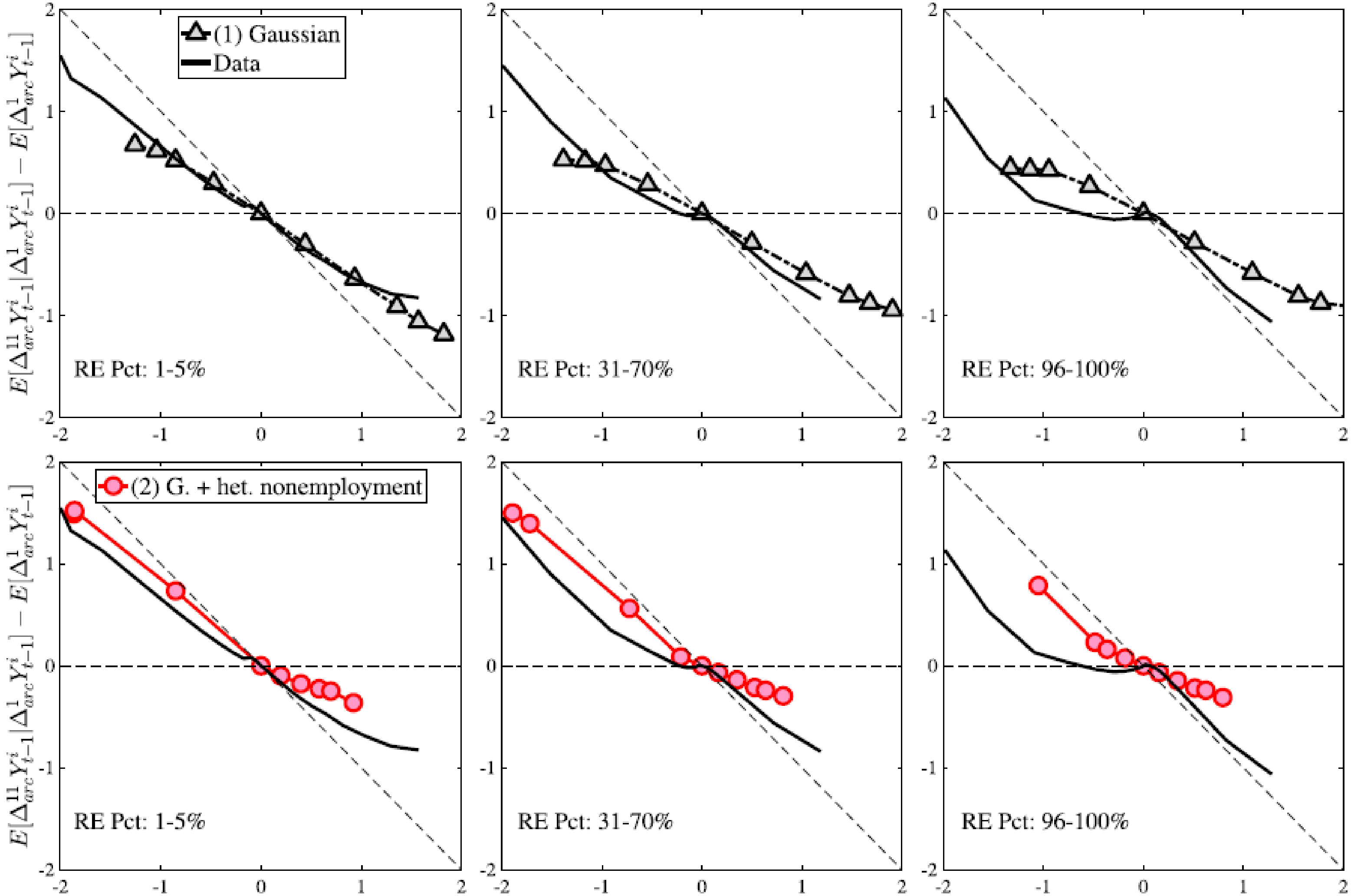
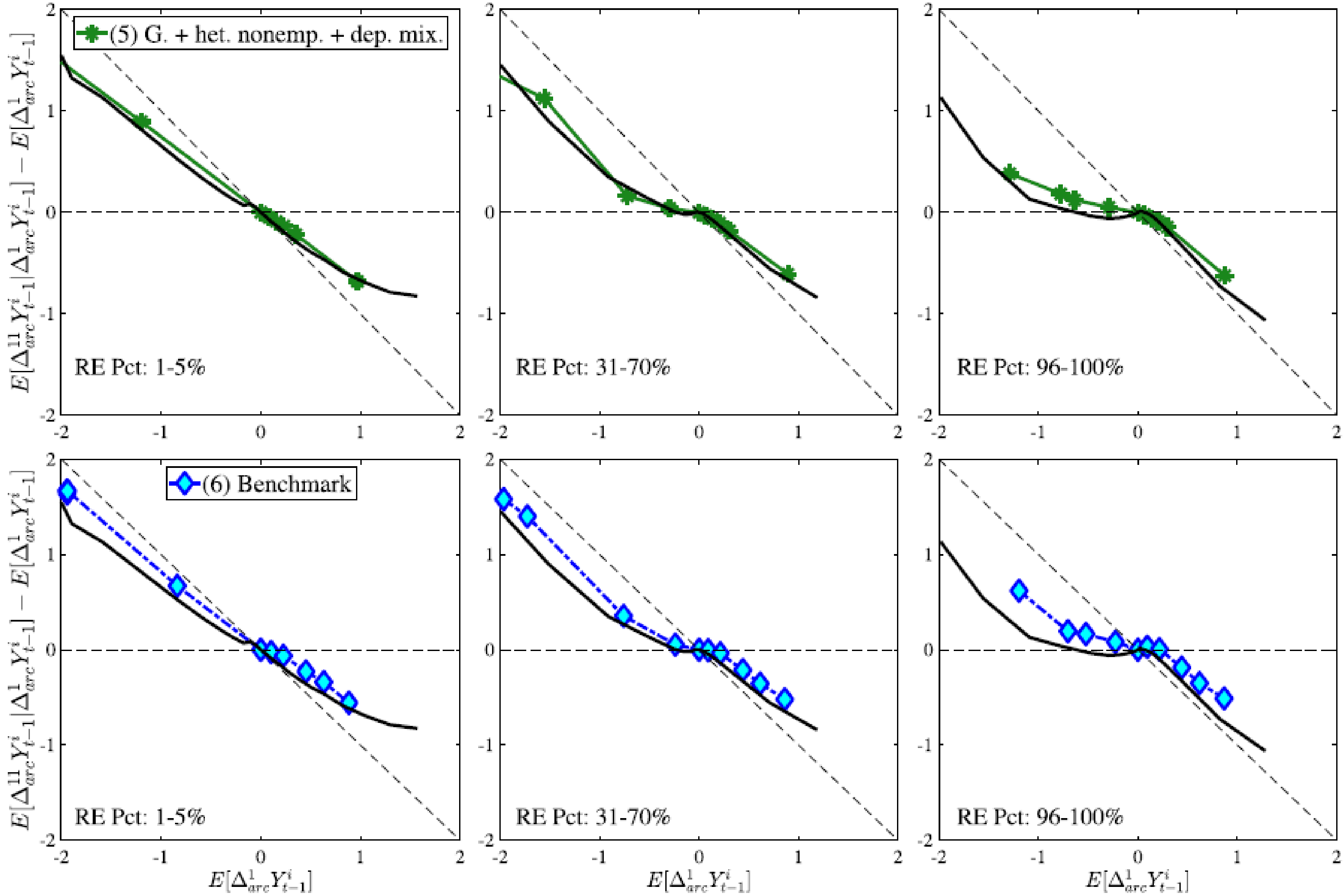


Figure 13. Estimated model versus data: selected impulse response moments, cont'd.



6.3. Parameter Estimates of the Benchmark Process

- Starting with the AR(1) process, the persistent shock is drawn about every 2.5 years ($p_z = 40.7\%$) from an “unfavorable” distribution—with a negative mean and large standard deviation ($\mu_{\eta,1} = -0.085$ and $\sigma_{\eta,1} = 0.364$)—and in the other years from a “favorable” one—with a positive mean and small standard deviation ($\mu_{\eta,2} = 0.058$, $\sigma_{\eta,2} = 0.069$).

- This mixture of normals implies that innovations to the persistent component are both strongly left skewed (skewness of -0.87) and leptokurtic (kurtosis around 6.3).

- In contrast, transitory shocks (ε_t) are typically smaller: In most years (with 87% probability), they are drawn from a tight distribution, $\mathcal{N}(-0.041, 0.037^2)$, and every eight years or so ($p_\varepsilon = 13.0\%$) from a distribution with large positive mean and dispersion, $\mathcal{N}(0.271, 0.285^2)$.

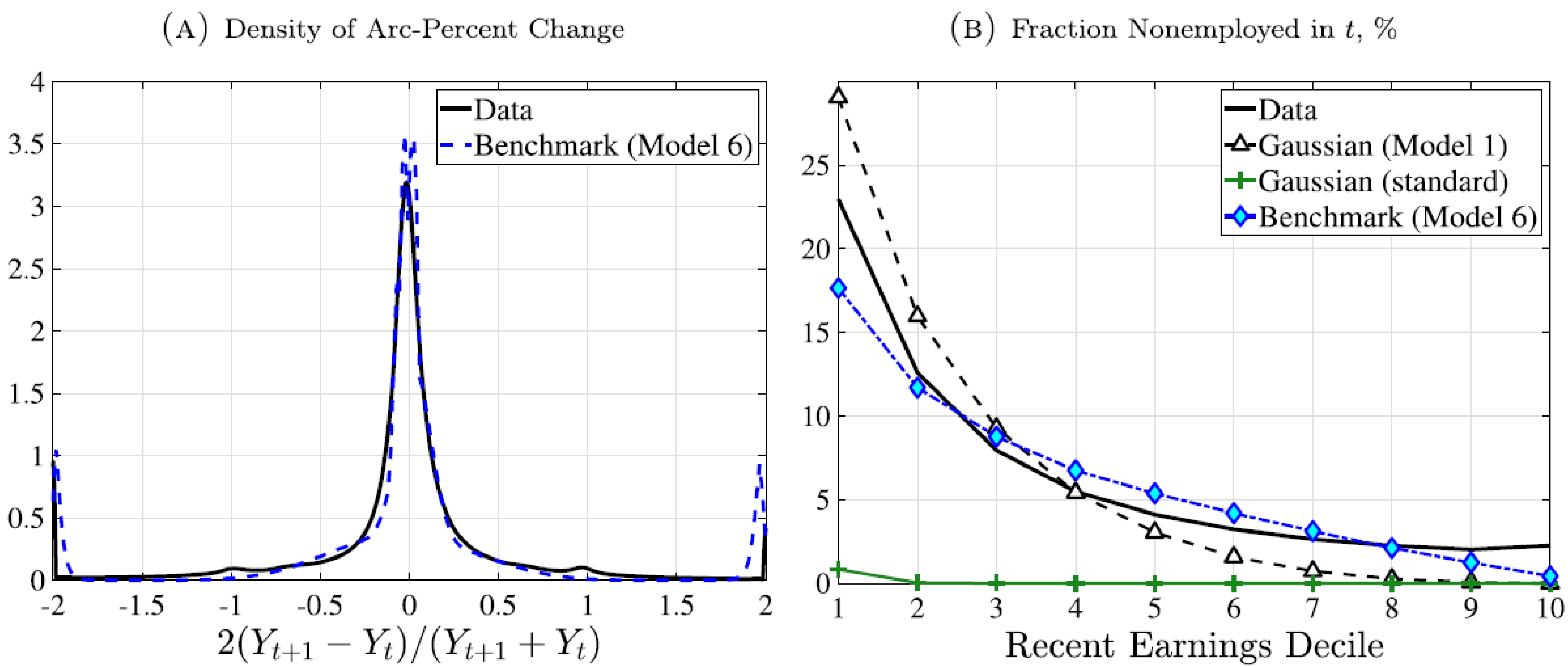
- Consequently, transitory shocks feature a skewness of 3.2 and a kurtosis of 15.4.

- However, η_t and ε_t are not the only sources of higher-order moments; workers also face nonemployment risk, which allows the model to generate a leptokurtic density for arc-percent changes with spikes at both ends (Figure 14(a)).

- The state dependence in nonemployment risk also leads to age- and income-varying skewness and kurtosis in persistent earnings changes.

- We conclude that persistent innovations are key drivers of non-Gaussian features in the data.

Figure 14. Model fit: nontargeted statistics.



Notes: The parameters of the Gaussian process come from Model 1 in Table IV, whereas Gaussian (standard) is the same process estimated without targeting the employment CDF. The data series on Panel (B) is conditional on past 2 years' income (in $t-1$ and $t-2$).

- Almost all workers hit by nonemployment shocks experience full-year nonemployment ($\lambda = 0.0001$).

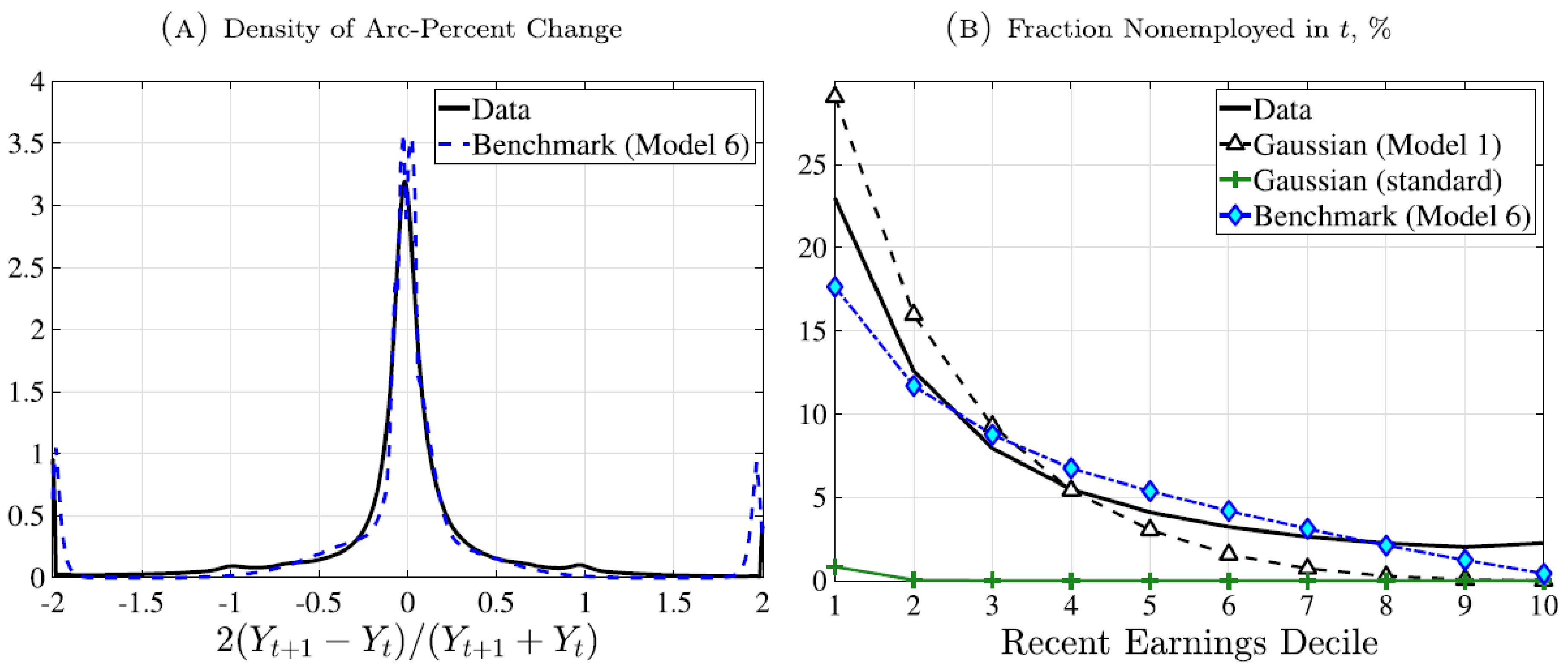
- How often are these shocks realized? We investigate this probability for various age and RE quantiles for workers who satisfy the conditions of the RE sample: Nonemployment risk declines modestly over a working life, from 6.9% over 25–34 to 6.1% in the next 10 years, and to 5.5% over 45–54.

- Differences in nonemployment risk between income groups are much more pronounced (Figure 14(b)).

- As noted earlier, however, this specification's ability to capture the nonemployment CDF comes at the cost of an implausibly steep inequality profile (Figure 12(f)).

- In fact, under a more plausible parameterization, only less than 4% of individuals ever experience a full year of nonemployment over their working life.

Figure 14. Model fit: nontargeted statistics.



Notes: The parameters of the Gaussian process come from Model 1 in Table IV, whereas Gaussian (standard) is the same process estimated without targeting the employment CDF. The data series on Panel (B) is conditional on past 2 years' income (in $t-1$ and $t-2$).

- How does the model capture the nonlinear earnings dynamics with a single AR(1) component?

- The autocorrelation of persistent shocks is not precisely captured by ρ, because modeling p_ν as a function of z_t implies that nonemployment is autocorrelated even though v_t is drawn in an i.i.d. manner.

- Moreover, since this function is highly nonlinear, how income responds on impact to a given shock η_t and how persistent this response is depend very much on the persistent component z_t and the sign and magnitude of the shock.

- This feature, along with the normal mixture shocks, generates the asymmetric mean reversion in impulse responses in Figure 13.

- To illustrate its persistence, we examine the future nonemployment risk of workers who are nonemployed in t.
- As usual, we further condition workers on their RE in $t-1$.
- Nonemployment is fairly persistent overall and more so for low-income workers: Between ages 25 and 35, 48% of the workers in the bottom RE decile experience another nonemployment spell five years after the initial nonemployment (Table V).
- This number declines monotonically over the RE distribution to 35% for the top decile.
- Furthermore, nonemployment risk becomes more persistent over the working life, particularly for high earners.

Table V. Persistence of Nonemployment Risk, Benchmark Process (Model (6))

	25–35		*36–45*		*46–55*	
Nonemp. at $t \rightarrow$	$t+1$	$t+5$	$t+1$	$t+5$	$t+1$	$t+5$
RE Groups						
1–10	0.543	0.481	0.583	0.500	0.623	0.519
41–60	0.455	0.426	0.544	0.475	0.604	0.503
91–100	0.334	0.352	0.491	0.440	0.582	0.498

7. Concluding Thoughts