What Do Data on Millions of U.S. Workers Say About Life Cycle Income Risk?*

Fatih Guvenen† Fatih Karahan‡ Serdar Ozkan§ Jae Song¶

Preliminary and Incomplete. Please Do Not Cite.

Abstract

We study the evolution of individual labor income over the life cycle using a very large panel dataset from the U.S. Social Security Administration. Using fully non-parametric methods, our analysis reaches two broad conclusions. First, the distribution of individual earnings shocks display substantial deviations from log-normality—the standard assumption in the incomplete markets literature. In particular, earnings shocks display strong negative skewness (individual “disaster” shocks) and extremely high kurtosis—as high as 35 compared with 3 for a Gaussian distribution. The high kurtosis implies that in a given year most individuals experience very small earnings shocks, but very few experience extremely large shocks. Second, these statistical properties of earnings shocks change substantially both over the life cycle and with the income level of individuals. We also estimate impulse response functions of income shocks and find significant asymmetries: positive shocks to high-income individuals are quite transitory, whereas negative shocks are very persistent; the opposite is true for low-income individuals. Finally, we use these rich sets of moments to estimate econometric processes with increasing generality to capture these salient features of earnings dynamics.

*The views expressed herein are those of the authors and do not represent those of the Social Security Administration, the Federal Reserve Bank of New York, or the Board of Governors of the Federal Reserve System.
†University of Minnesota and NBER; guvenen@umn.edu
‡Federal Reserve Bank of New York; fatih.karahan@ny.frb.org
§Federal Reserve Board; serdar.ozkan@frb.gov
¶Social Security Administration; jae.song@ssa.gov
1 Introduction

The goal of this paper is to shed new light on the nature of life cycle income risk using a large and confidential dataset covering 1978 to 2010 from the U.S. Social Security Administration. The substantial sample size (more than 200 million individual-year observations) allows us to take “high-resolution pictures” of individual earnings data and document five empirical facts regarding earnings risk over the life cycle.

First, earnings changes display extremely high kurtosis. What kurtosis measures is most easily understood by looking at the histogram of earnings changes, shown in Figure 1 (left panel: annual change; right panel: 5-year change). Notice how sharp the peak of the empirical density is, how little mass there is on the “shoulders” (i.e., the region around $\pm \sigma$), and how long the tails are compared with a normal density chosen to have the same standard deviation as in the data (of 0.48). Thus, there are far more people with very small income changes in the data compared with what would be predicted by a normal density. For example, under the normal distribution, about 8 percent of individuals would experience an annual income change (of either sign) smaller than 5%. The true fraction of such individuals in the US data is 35 percent, illustrating the concentration of individuals near very small income changes. Furthermore, this average kurtosis masks significant heterogeneity across individuals by age and income: Prime-age males with recent earnings of $100,000 (in 2005 dollars) face earnings shocks with a kurtosis as high as 35, whereas young workers with recent earnings of $10,000 face a kurtosis of only 5. This life cycle variation in the nature of earnings shocks is one of the key focuses of the present paper.

Second, we document that earnings shocks are negatively skewed, which becomes more severe as individuals get older and/or their earnings increase. This finding implies that individuals face limited upside surprises in their income as they get older and richer, and face an increasing “disaster” risk—the risk of a sharp fall in their earnings. Although this implication is quite plausible and may not seem surprising, it is not captured by a lognormal specification, which implies zero skewness and excess kurtosis. Overall, these results show that the higher order moments of earning changes—which is ignored in the workhorse lognormal framework in quantitative economics—display interesting and rich properties, which is likely to matter for a broad range of questions in economics.
Third, the rise in earnings over the life cycle (e.g., from ages 25 to 55) varies strongly with the level of lifetime earnings. For example, the median individual in the population (by lifetime earnings) experiences an earnings growth of 38% from ages 25 to 55, whereas for individuals in the 95th percentile, this figure is 250%; and for those in the 99th percentile, this figure is 1500%! We show that a stochastic process for earnings featuring a persistent plus a transitory component, which is the workhorse model in the literature, completely misses the kind of heterogeneity we find in the data. This is true even when its parameters are chosen to match these moments as well as possible.\(^1\)

Fourth, we characterize the dynamics of income shocks by estimating non-parametric impulse response functions separately by age, and conditional on the past average income of the individuals and on the size of the shock that hits them. We find two types of important asymmetries. One, fixing the size of the shock, positive shocks to high-income individuals are quite transitory, whereas negative shocks are very persistent; the opposite is true for low-income individuals. Two, fixing the income level of individuals, the strength of mean reversion differs by the size of the shock: large shocks tend to be much more transitory than small shocks. These kinds of asymmetries are hard to

\(^1\)Clearly, some of this relationship can be generated through luck: those with high earnings growth (due to say positive earnings shocks) will end up with high levels of lifetime income. Although this is true, we show that this effect cannot explain the vast differences we find in the data.
detect via the usual approach in income dynamics literature, which relies on covariances of income at different points in time. In this regard, our approach is in the spirit of the recent macroeconomics literature that views impulse responses as key to understanding time-series dynamics (e.g., Christiano et al. (2005), Borovicka et al. (2014)).

These features of individual earnings changes over the life cycle turn out to be difficult to capture with standard econometric specifications used in the existing literature. In Section 6, we estimate a set of stochastic processes with increasing generality to capture these salient features of earnings dynamics to provide a reliable “user’s guide” for applied economists. [This section is still incomplete and will be updated with new results soon.]

Literature Review

[To be Written.]

2 Empirical Analysis

2.1 The SSA Data

The data for this paper come from the Master Earnings File (MEF) of the US Social Security Administration records. The MEF is the main source of earnings data for the SSA and contains every individual in the United States who was ever issued a Social Security number. Basic demographic variables, such as date of birth, place of birth, sex, and race are available in the MEF along with several other variables. The earnings data in the MEF are derived from the employee’s W-2 forms, which U.S. employers have been legally required to send to the SSA since 1978. The measure of labor earnings is annual and includes all wages and salaries, bonuses, and exercised stock options as reported on the W-2 form (Box 1). Furthermore, the data are uncapped (no top coding) since 1978. We convert nominal earnings records into real values using the Personal Consumption Expenditure (PCE) deflator, taking 2005 as the base year. For background information and detailed documentation of the MEF, see Panis et al. (2000) and Olsen and Hudson (2009).
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 (n arbitrary) number for the last digit and select all individuals in 1978 whose SSN ends with that number. This 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 turn 25 are automatically added (because they will have the last digit we pre-selected), and those who died in or before 1979 are again dropped. Continuing with this process yields a 10% representative sample of US males in every year from 1978 to 2010. Finally, there is a small number of extremely high earnings observations in the MEF. In each year, we cap (winsorize) observations above the 99.999th percentile in order to avoid potential problems with these outliers.

Base Sample. Sample selection works in two steps. First, for each year we define a “base sample,” which includes all observations that satisfy three criteria, to be described in a moment. Second, to select the “final sample” for a given statistic that we analyze below, we select all observations that belong in the base sample in a collection of years, the details of which vary by the statistic and the year the statistic is constructed for.

First, we are restricting our base sample to individuals between the ages of 25 and 60 to focus on working-age population. Second, our base sample includes only workers whose annual wage/salary earnings exceeds a time-varying minimum threshold, denoted by $Y_{\min,t}$. The threshold $Y_{\min,t}$ is defined as one-fourth of a full year-full time (13 weeks at 40 hours per week) salary at the half of the minimum wage, which amounts to an annual earnings of approximately $1,300 in 2005. This condition helps us avoid issues with taking the logarithm of small numbers and makes our analysis more comparable to the empirical income dynamics literature, where this condition is fairly standard (see, among others, Abowd and Card (1989), Meghir and Pistaferri (2004), and Storesletten et al. (2004)). Third, the base sample excludes individuals whose self-employment earnings

\footnote{In reality, each individual is assigned a transformation of their SSN number for privacy reasons, but the same method applies.}
exceed a threshold level, defined as the maximum of \( Y_{\text{min},t} \) and 10% of the individual’s wage/salary earnings in that year. These steps complete the selection of the base sample. The selection of the final sample for a given statistic is described further below.

### 2.2 Empirical Approach

We now begin to characterize the properties of earnings shocks and document how these properties vary across workers that differ in observable labor market characteristics. One such characteristic is age and therefore a major focus throughout the paper is to examine the variation over the life cycle. Another observable by which earnings shocks could differ is income. But what measure of income is relevant? It seems plausible that workers may face income shocks with different characteristics depending on where they currently rank in the income distribution. This would suggest that as individuals move across the income distribution, the nature of shocks they face also changes with their income. This would be the case, for example, if these characteristics are linked to their occupations or their seniority in their jobs, which changes over the life cycle. This view suggests that we classify workers based on their recent earnings and study how the properties of income shocks of groups of workers that are similar to each other by this measure.\(^3\)

#### 2.2.1 Cross-sectional Moments

Suppose that income shocks vary with two characteristics: age and an individual’s recent earnings. In this case, if we can group individuals into age/income bins (as of time \( t - 1 \)) that are sufficiently fine, we can think of all individuals within a given bin to be ex ante identical. We can then compute the cross-sectional moments of earnings changes (say \( y_{t+k} - y_t^i \), for \( k = 1, 2, ... \)) for each group, which can then be viewed as corresponding to the properties of shocks individuals within each bin can expect to face. This approach has the advantage that we can compute higher order moments (which typically require large

\(^3\)A second possibility however is that shocks are more intimately related to individuals and different types of workers—for example, identified by their lifetime earnings—might face shocks with different properties. This suggests that we should group workers based on their lifetime earnings and study the properties of shocks for each group. We have an analogous set of results obtained by adopting this alternative perspective. It turns out that both approaches yield very similar substantive conclusions, so we omit these results from the main text.
samples to be estimated precisely) with this approach as each bin contains sometimes
tens of thousands of individuals.

When studying the distribution of income growth between \( t \) and \( t + k \), we first group
workers based on their observable characteristics up to year \( t - 1 \). We first group workers
into seven age bins with respect to their ages in year \( t - 1 \): 25–29, 30–34, ..., 50–54,
55–60.\(^4\) A second variable we use to classify individuals at a point in time is a proxy for
lifetime income constructed using earnings in the previous five years. To this end, we
first compute each worker’s average past income between years \( t - 1 \) and \( t - 5 \). We set
income observations below the minimum threshold to the threshold during this period.
Furthermore we control for age effects because even within narrowly defined age groups,
as characterized above, age variation can skew the rankings in favor of older workers. For
this purpose, we normalize the average past income of each worker using age dummies
from the pooled regression of log earnings on age and cohort dummies.\(^5\) In particular,
we divide the average past income of a worker by the average of the age dummies over
the period in which average income is computed. Namely, let worker \( i \) be at age \( h \) in
year \( t \) and let’s denote his log labor income in year \( t \) with \( \tilde{y}^i_{t,h} \) and the age dummy at age
\( h \) with \( d_h \), then:

\[
\bar{Y}^i_{t-1} = \frac{\sum_{s=1}^{5} e^s \tilde{y}^i_{t,h-s}}{\sum_{s=1}^{5} e^s d_{h-s}}.
\]

3 Cross-sectional Moments of Earnings Growth

We begin our analysis by documenting empirical facts about the first four moments of
earnings growth at short (1-year) and long (5- and 30-year) horizons. Income growth
for a worker is defined as the time difference of \( y^i_t \), which is log earnings net of the age
effect. Thus:

\[
\Delta_k y_t \equiv (y^i_{t+k} - y^i_t) = (\tilde{y}^i_{t+k} - d_{h+k}) - (\tilde{y}^i_t - d_h).
\]

We compute the cross-sectional moments of \( \Delta_k y_t \) (i.e., standard deviation, skewness,
kurtosis) for each year, \( t = 1980, 1981, ..., 2009 \) and then aggregate these up to obtain

\(^4\)The first six groups are 5-year bins and the last one covers six years.
\(^5\)For details of the pooled regression see Guvenen et al. (2013)
Table I: Sample Size Statistics for Cross-Sectional Moments

<table>
<thead>
<tr>
<th>Age group</th>
<th># Observations in Each Percentile Group of $\bar{Y}_{i-1}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
</tr>
<tr>
<td>25-29</td>
<td>408,713</td>
</tr>
<tr>
<td>30-34</td>
<td>638,350</td>
</tr>
<tr>
<td>35-39</td>
<td>618,909</td>
</tr>
<tr>
<td>40-44</td>
<td>571,512</td>
</tr>
<tr>
<td>45-49</td>
<td>506,989</td>
</tr>
<tr>
<td>50-54</td>
<td>425,539</td>
</tr>
<tr>
<td>55-59</td>
<td>276,340</td>
</tr>
</tbody>
</table>

Note: Each entry reports the statistics of the number of observations in each of 100 income percentile groups for each age. Cross-sectional moments are computed for each year and then averaged over all years, so sample sizes refer to the sum across all years of a given age by percentile group.

3.1 First Look at Data

Before delving into a disaggregated analysis of the data, we begin by looking at some broad statistics of 1-year and 5-year earnings growth. The left panel of Figure 1 plots the histogram of annual earnings changes in the US data (blue solid line), superimposed with the histogram of a normal distribution with the same mean and standard deviation. The contrast is rather striking: the data histogram looks nothing like normal, with a
substantial peak, very long tails and very low shoulders (mass near $\pm \sigma$). The figure also reports the second to fourth moments of the data, revealing a very high standard deviation of 0.48, a negative skewness of -1.35, and most strikingly, an extremely high kurtosis of 17.80.$^6$ The right panel plots the histogram for 5-year earnings changes, again superimposed with a normal density. Although, the distribution is a little bit less concentrated around zero, it is still very far from a normal, with a kurtosis of 11.55; it has a large standard deviation of 0.68 and a negative skewness of -1.01.

The first main takeaway from this figure is that earnings changes, both at short- and long-horizon are very different from log normal, displaying substantial concentration near zero and very long tails, along with negative skewness. However, if we stop here, we will be missing lots of heterogeneity that lies underneath this overall picture. To see what we mean, Figure 2 plots 12 histograms, for earnings changes for different age and past income groups. The top row plots annual income changes for young workers and the next row plots the same for prime-age males. From left to right, the three panels plots the histograms for individuals who are in the 10th, 50th, and 90th percentile of the past average earnings distribution (indicated with P10, P50, and P90 in the figure). Although we will elaborate on these features later, at this point it suffices to observe how much the histograms change with these two characteristics. Therefore, simply saying that earnings changes have negative skewness and excess kurtosis misses lots of variation over the life cycle and across income levels. Once the analysis in this section and the next is completed, it will become more apparent that this figure summarizes many of the conclusions we reach about the properties of earnings shocks.

### 3.2 Fourth Moment: Kurtosis

It is useful to spend some time discussing what kurtosis measures. A useful interpretation of kurtosis has been suggested by Moors (1986, 1988), who described it as measuring how dispersed a probability distribution is away from $\mu \pm \sigma$. This can be easily seen by introducing a standardized variable $Z = (x - \mu)/\sigma$ and noting that kurtosis is

$$
\kappa = \mathbb{E}(Z^4) = \text{var}(Z^2) + \mathbb{E}(Z^2)^2 = \text{var}(Z^2) + 1.
$$

$^6$A distribution with a kurtosis of 5 or 6 is considered to be highly leptokurtic.
Figure 2: Histogram of Earnings Growth, By Past Average Earnings and Age Groups
Table II: Fraction of Individuals with Selected Ranges of Log Income Change

<table>
<thead>
<tr>
<th>$x$</th>
<th>All workers</th>
<th>Ages 45-50, P90</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Data</td>
<td>$N(0, 0.48^2)$</td>
</tr>
<tr>
<td>0.05</td>
<td>0.35</td>
<td>0.08</td>
</tr>
<tr>
<td>0.10</td>
<td>0.54</td>
<td>0.16</td>
</tr>
<tr>
<td>0.20</td>
<td>0.71</td>
<td>0.32</td>
</tr>
<tr>
<td>0.50</td>
<td>0.86</td>
<td>0.70</td>
</tr>
<tr>
<td>1.00</td>
<td>0.94</td>
<td>0.96</td>
</tr>
</tbody>
</table>

Note: The empirical distribution used in this calculation is for 1995-96, the same as the one in the plotted histograms.

So $\kappa$ can be thought of as the dispersion of $Z^2$ around its expectation, which is 1. Or, the dispersion of $Z$ around +1 and −1. This is consistent with how a distribution with excess kurtosis often looks like: a sharp/pointy center, long tails, and little mass near $\pm \sigma$. A corollary to this description is that for a distribution with high kurtosis the usual way we think about standard deviation is not very useful: very few realizations of the random variable is of a magnitude close to the standard deviation. Instead, most realizations are either close to zero (or the mean) or are in the tails.

Armed with this definition, let us now examine the earnings growth data. First, and most importantly, annual income growth displays extremely high kurtosis—ranging from 14 to 18—compared with a normal distribution, whose kurtosis is 3. An even more interesting picture emerges when we group individuals by age and (past average) income. Figure 3 plots the kurtosis of one-year income change ($y_{t+1} - y_t$) for individuals grouped by age and by their past 5-year average income (on the x-axis). Notice first that kurtosis increases monotonically with past income up to the 80th to 90th percentiles for all age groups. That is, high-income individuals experience even smaller income changes of either sign, with few experiencing very large changes. Second, kurtosis increases with age, for every level of past income, except perhaps the top 5% of $\bar{Y}$. Furthermore, and most significantly, the peak levels of kurtosis reached ranges from a low of 20 for the youngest group, all the way up to 30 for the middle age group (40–54).

These figures represent important deviations from the log-normality assumption and
raise serious concerns about the current focus in the extant literature on the covariances (second moments) alone. In particular, targeting the covariances only can vastly overestimate the typical income shock received by the average worker and miss out the substantial but infrequent jumps experienced by few.

There are well-known statistical and economic frameworks that can generate very high kurtosis. One example of a statistical model is one where income shocks follow a Poisson arrival process. Thus, income does not change regularly—most of the time there is no change—and once in a while there is a big up or down move (promotion, job loss, etc.). Alternatively, consider an economic model of job search that can be modeled as a mixture of normals: every period each worker draws a random variable which tells him whether he is going to be changing jobs or not. If he does, he draws a new income realization from a normal distribution with a large variance—and vice versa when he does not change his job. The overall income change distribution can easily be made to have very high kurtosis.
Figure 4: Kurtosis of 5-Year Income Change, By Age and Past Income

Figure 5: Skewness of Annual Earnings Growth
3.3 Third Moment: Skewness

The log-normality assumption also implies that the skewness of income shocks is zero. Figure 5 (left) plots the skewness of annual income changes, conditional on past income as done above. First, notice that income shocks (on average) always have negative skewness, regardless of the stage of the life cycle or where the individual stands in the past income distribution. However, and further, skewness becomes even more negative as we move to the right (higher income levels) and as individuals get older. Thus, it seems that the higher an individual’s past average income, the more room he has to fall down, and the less room he has left to move up. The right panel plots Kelley’s measure of skewness, which is defined as:

\[ S_K = \frac{(P_{90} - P_{50}) - (P_{50} - P_{10})}{P_{90} - P_{10}}, \]

and basically measures the relative fractions of the overall dispersion (P90-P10) accounted for by the upper and lower tails. A negative number implies that the lower tail is longer than the upper tail, indicating negative skewness. Kelley’s measure is less sensitive to extremes (above the 90th or below the 10th percentile of the shock distribution) and captures the weight shift in the center of the shock distribution. In the right panel, Kelley’s skewness is negative throughout, becomes more negative with age (in general). However, it does not always get more negative with higher income. This difference between the two panels indicates that as income increases it is mostly the extreme shocks that become more negatively skewed, rather than the more middling shocks.

Turning to more persistent earnings changes, Figure 6 plots the two measures of skewness for 5-year changes. Here the two panels are more consistent with each other—each showing a strong increase in left-skewness with both age and income levels (except for the very high earners). Furthermore, the magnitude of skewness is substantial. For example, the Kelley’s skewness of −0.35 for individuals aged 45-49 and in the P80 of the past income distribution implies that the log 90-50 differential of \( y_{t+5} - y_t \) accounts for 32% of the log 90-10 differential, whereas the log 50-10 differential (the left tail) accounts for the remaining 68%. This is very different from a log normal distribution which is symmetric (and therefore both tails contribute 50% of the total).
Third central moment

Kelley’s measure

Figure 6: Skewness of 5-Year Earnings Growth

P90-P50; 5-year change

P50-P10; 5-year change

Figure 7: Kelly’s Skewness Decomposed: P90-P50 and P50-P10; 5-year income growth
3.4 Second Moment: Variance

Figure 8a plots the variance of income shocks as a function of past income. First, at all ages, there is a pronounced U-shaped pattern, implying that income shocks are less dispersed for higher income individuals up to about the 90th percentile on the x-axis. This pattern reverts itself inside the top 10% and dispersion of shocks increases rapidly as past income increases. Second, over the life cycle, the dispersion of shocks declines monotonically, and almost for all income groups up to about age 50, and then rises slightly for middle to high income individuals from age 50 to 55. Notice that this life cycle pattern in dispersion is much weaker for top income individuals who experience a smaller change with age.

3.5 Decomposing the Moments

Going back to Topel and Ward (1992), a number of studies have shown important differences between the income changes that occur during an employment relationship and those that occur across jobs (see more recently, Low et al. (2010)). We now examine how the empirical patterns we documented so far relate to within- and between-job income changes. An important advantage of our dataset is that it contains employer identifiers for each job that a worker holds in a given year, which allows us to conduct such an analysis.

One challenge we face is that many workers hold multiple jobs in a given year, which requires us to be careful about how to think of job changes. We have explored several definitions and found broadly very similar results. Here we describe one reasonable classification: A worker is said to be a “job stayer” between years $t$ and $t+1$ if a given EIN (employer identification number) provides the largest part of his annual income (out of all his EINs in that year) in years $t-1$ through $t+2$ and further, if the same EIN provides at least 90% of his total income in years $t$ and $t+1$. A worker is defined as a “job switcher” if he is not a job stayer.\footnote{Clearly, this classification is quite stringent for classifying somebody as a job stayer, meaning that some individuals will be classified as job switchers even though they did not change a job. An alternative definition we have explored defines a job switcher directly as somebody who has an EIN that provides more than 50% of his annual income in year $t$ and provides less than 10% of his annual income in year $t+1$.}
Figure 8: Standard Deviation of Earnings Growth

(a) 1-Year Growth

(b) 5-Year Growth
Figure 9: Standard Deviation of Earnings Growth: Stayers vs Switchers

(a) 1-Year Growth

(b) 5-Year Growth
Table III: Moments with Trimmed Data

<table>
<thead>
<tr>
<th></th>
<th>Original</th>
<th>Trim top/bot. 1%</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Std. dev.</td>
<td>Skew.</td>
</tr>
<tr>
<td><strong>1-Year</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All</td>
<td>0.49</td>
<td>−1.49</td>
</tr>
<tr>
<td>Stayers</td>
<td>0.31</td>
<td>−1.51</td>
</tr>
<tr>
<td>Switchers</td>
<td>0.67</td>
<td>−1.09</td>
</tr>
<tr>
<td><strong>5-Year</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All</td>
<td>0.66</td>
<td>−1.06</td>
</tr>
<tr>
<td>Stayers</td>
<td>0.41</td>
<td>−1.04</td>
</tr>
<tr>
<td>Switchers</td>
<td>0.78</td>
<td>−0.92</td>
</tr>
</tbody>
</table>

As can be seen in Figures 9, 10, and 11, the results are quite consistent with what we might expect. Job-stayers face a dispersion of income shocks that is much smaller than job-changers; job-changers face shocks that are very negatively skewed as opposed to job stayers who face either zero or slightly positively skewed shocks; and job-stayers experience shocks with much higher kurtosis than job-changers. These implications are broadly consistent with standard search-theoretic models of life cycle job changes.

Before concluding this section, we examine how the tails of the income change distribution affect the computed statistics and also examine how this effect varies by stayers and switchers. Unlike with survey-based data, here we are not too concerned that these tails might be dominated by measurement error, as most of these changes are likely to be genuine. Instead, we are simply interested in understanding what parts of the distributions are critical for the levels and the differences in different moments. Table III reports the 2nd to 4th moments for the original sample used so far (left panel) as well as for a sample where we drop tail observations, defined as those in the top or bottom 1% of the earnings change distribution.

---

*t + 1; and also has an EIN that provides less than 10% of his income in *t* and more than 50% in *t + 1*. The results were very similar to those reported here.
Figure 10: Skewness of Earnings Growth: Stayers vs. Switchers

(a) 1-Year Growth

(b) 5-Year Growth
Figure 11: Kurtosis of Earnings Growth: Stayers vs. Switchers

(a) 1-Year Growth

(b) 5-Year Growth
3.6 First Moment: Mean of Log Income Growth

The following analysis requires us to first compute lifetime earnings for each individual. To this end, we select a subsample of individuals who has at least 32 years of data between the ages of 25 and 60. We further restrict our sample to individuals that (i) have earnings above the threshold for at least 15 years and (ii) are not self-employed for more than 8 years. We rank individuals based on their lifetime income, defined as the sum of their earnings from ages 25 through 60. Income observations lower than the minimum threshold are set to this threshold. We compute income growth between any two ages \( h_1 \) and \( h_2 > h_1 \) as \( \log E_{y_{h_2}} - \log E_{y_{h_1}} \).

Figure 12 plots the life cycle profile of log average income, which is obtained by extracting age effects from panel data using the usual procedure, controlling for cohort effects (Deaton and Paxson (1994)). One side note here is that the circles shown in this figure are full age dummies, although they turn out the be indistinguishable from a 4th order polynomial fitted to the same data, a point first observed by Murphy and Welch (1990) in CPS data.

More substantively, we are interested in the rise in earnings over the life cycle, here shown from age 25 to 55. It is well understood that the magnitude of this rise matters...
greatly for many questions of economic interest, as it is a strong determinant of borrowing and saving motives over the life cycle. In our data, this rise is about 80 log percent, which is about 127%. This figure lies on the high end of some estimates from datasets such as the PSID, but not unseen before (c.f. Attanasio et al. (1999)). Notice that a life cycle model calibrated with this profile will imply that the median individual in the simulated sample experiences (on average) a rise of this magnitude from age 25 to 55. One question we would like to address is, if this prediction is a good approximation to what we see in the data. That is, we would like to see if an individual with median lifetime earnings indeed experiences income growth over the life cycle of approximately 127%.

Figure 13 plots the results. There are several takeaways. First, the median individual experiences a growth rate of 38%, about 1/3 of what the profile we extracted above predicts. Moreover, we have to look all the way above the 90th percentile of lifetime earnings to see individuals who on average experiences income growth of 127%. Second, and however, another striking fact is that income growth is very high for high income individuals, with those in the 95th percentile experiencing an income growth of 200% and those in the 99th percentile experiencing a growth of 1500% over the life cycle!

Although some of this variation could be expected simply due to endogeneity (i.e., that those individuals with high income growth are more likely to have high lifetime earnings).
income), the magnitude observed here is too large to be accounted for by that channel. For example, a standard persistent-transitory model estimated in the literature (such as in Hubbard et al. (1995); Storesletten et al. (2004)) would predict that individuals in the top 1% of the lifetime income distribution should have earnings growth over the life cycle that exceeds the median individual by only 5 percentage point. The actual gap in Figure 13 is 275 log points, which corresponds to 1560 percentage points!

The next figure (14) decomposes the overall difference in lifecycle income growth into three decades: growth from ages 25 to 35, from ages 35 to 45 and, from ages 45 to 55. Several points are easily noted. First, most of the growth over the life cycle happens from ages 25 to 35, for all lifetime income groups. In fact for the median income group, average growth from ages 35 to 55 is zero, so growth during this early period is equal to growth over the entire 25–55 period (notice that the blue solid line and grey line with circles overlap at P50). Second, with the exception of workers in the top 10%, all income groups experience negative growth from age 45 to 55. So the peak year of earnings is strongly related to the lifetime earnings percentile. For those in the top 1-2%, growth happens throughout the life cycle, although it slows down considerably with each decade passing.

It is useful to also see how the lifecycle growth rates look like when we shift the first
year from 25 to 30 and then to 35. As can be anticipated from the previous graph, the growth rate falls substantially for the 35-55 age range, the median rate being zero. Top earners still do very well though, experiencing a rise of 200 log points (or 740%) from ages 30 to 55 and a rise of 90 log points (or 250%) from ages 35 to 55. The lowest earners display a symmetric pattern with income loss experienced throughout the life cycle: average income drops by 70 log points (or 50%) from ages 35 to 55 for those in the lowest percentile of lifetime income.

### 3.7 Robustness

The results documented in the previous sections show important deviations from log-normality as well as clear patterns with age and past income. This raises the question whether some of these findings are due to simple statistical artifacts, say maybe due to extreme shocks experienced by few individuals, and whether the age and income patterns might be due to sample selection or other assumptions made in the construction of these statistics. In this section, we explore several of these possibilities.
**Averaging Neighboring Years.** First, we constructed the previous statistics by taking income differences between two years, \( t \) and \( t + k \). One concern could be that what we are measuring a big shock at time \( t \) or \( t + k \) is a timing issue: that for some reason the individual’s income has been shifted from the end of \( t + k \) into the beginning of \( t + k + 1 \). If true this would be a fluctuation that is easy to smooth, but could appear as a big negative shock from \( t \) to \( t + k \). A similar comment applies to period \( t \). To address this, we construct the same set of statistics for the second to fourth moments by using 2-year average earnings. For the short-run and long-run variation we use, respectively:

\[
\tilde{\Delta} y_t = \log(Y_{t+3} + Y_{t+2}) - \log(Y_t + Y_{t+1})
\]

and

\[
\tilde{\Delta}_5 y_t = \log(Y_{t+5} + Y_{t+6}) - \log(Y_t + Y_{t+1}).
\]

Notice that the first measure is not a 1-year difference but is more like a 2-year difference, whereas the second one is closer to a 5-year difference as before. However, we are mostly interested in whether statistics are broadly robust and the qualitative patterns remain unchanged, so these are reasonable choices.

**Difference from Usual Income.** Even though we condition average income over the past 5 years and require all individuals in the sample to be employed in year \( t - 1 \), it is conceivable that some individuals receive large positive shocks in period \( t \), and the subsequent drop in income from \( t \) to \( t + k \) is simply mean reversion and not a new shock. The same argument applies for a large negative shock in \( t \). To see if this might be important we also construct the same statistics using an alternative difference measure, again for short-run and long-run:

\[
\Delta y = \log(Y_{t+1}) - \log(\overline{Y}_{t-1}),
\]

and

\[
\Delta y = \log(Y_{t+5}) - \log(\overline{Y}_{t-1}).
\]

These are longer differences than 1- and 5-year since the base year is now centered
around $t - 3$.

**Trimming the Tails.** As noted before, earnings growth displays very long tails, and even though measurement error is unlikely to explain it given the quality of the data, it is still of interest to know, for example, how much of the very large kurtosis and negative skewness is due to the extreme observations and how much is due to the bulk of the distribution. Since the 3rd and 4th moments are sensitive to tails this is worth exploring (although Kelly’s skewness is already reassuring for the skewness). Therefore, we construct the statistics in two different ways. First, we drop all observations in the top and bottom 1% of the earnings growth distribution by age and past income percentile, and then calculate the 3rd and 4th moments.

Second, recall that we consider an individual to be non-employed and drop from the computation of a statistic if his income is below a fixed income threshold in period $t$ or $t + k$. Because this threshold is fixed, individuals with higher incomes have a longer distance to fall and still remain in the sample relative to low income individuals who would exit the sample with a smaller downward fall. This might give the appearance of a more negative skewness for higher income individuals. To explore this issue, we drop all observations where an individual’s earnings in period $t$ or $t + k$ is below 5% of his past average income. Because the threshold is now indexed to the level of his income, the mechanical relationship is no longer a concern.

In Appendix A, we report the figures analogous to those above under these three robustness checks. While the figures are quantitatively different, qualitatively, the broad message of this analysis remains intact.

## 4 Dynamics of Earnings

In this section, we study the dynamic properties of earnings shocks. Our main focus is on the nature of persistence in income changes and how the dynamics differ across the income distribution as well as how it varies depending on the sign and magnitude of income shocks.
Sample Selection for Impulse Responses. The sample selection criteria here follows our methodology for cross-sectional facts. In addition to the same sample selection criteria, in each year \( t \), we exclude all workers that are self-employed in any of the years \( t-5 \) through \( t+3 \), \( t+5 \), and \( t+10 \). We focus on two age groups: young workers (25-34) and prime-age workers (35-55). In each year, individuals are assigned to groups based on where they stand in the average past income distribution among individuals of the same age. The following list defines these groups in terms of past income percentiles: 1-5, 6-10, 11-30, 31-50, 51-70, 71-90, 91-95, 96-100. We then compute the income growth between \( t-1 \) and \( t \) as \( \tilde{y}_t^i - \tilde{y}_{t-1}^i \). We pool all such samples for various years, and assign people to groups based on the percentiles of the income change distribution. The following list defines our groups: 1-2, 3-5, 6-10, 11-20, 21-30, 31-40, 41-50, 51-60, 61-70, 71-80, 81-90, 91-95, 96-98, 99-100. We compute and report the dynamics in two distinct ways. First, for workers in each of the group, we compute the average of log income growth in years \( k \) years for \( k = 1, 2, 3, 5, 10 \); i.e. \( \mathbb{E}[y_{t+k} - y_t] \). To avoid issues that could arise due to the presence of very large income changes, we also compute and report the average income growth for a “representative agent” in a group. More specifically, this measure is defined as the log change in mean earnings in \( t+k \) and \( t \): \( \ln \mathbb{E}(Y_{t+k}) - \ln \mathbb{E}(Y_t) \).

4.1 Impulse Response Functions

A key dimension of life cycle income risk is the persistence of income changes. Typically, this persistence is modeled as an AR(1) process or a low-order ARMA process (typically, ARMA(1,1)), and the persistence parameter is pinned down from the rate of decline of autocovariances with time. The AR(1) structure, for example, predicts a geometric decline and the rate of decline is directly given by the mean reversion parameter. While this might be appropriate in survey data, given the data limitations, it imposes some restrictions on the data that one might be skeptical about, such as the uniformity of mean reversion for positive and negative shocks, for large and small shocks, and so on. Here, the substantial sample size allows us to get a much higher resolution picture of the data, and in particular, characterize persistence without making parametric assumptions. To this end, we rank and group individuals based on their average income from \( t-5 \) to \( t-1 \), then within each such group, we rank and group again by the size of the income change.
between \( t - 1 \) and \( t \). Hence, all individuals within a given group obtained by crossing the two conditions, have the same average income up to time \( t - 1 \) and experienced the same income “shock” from \( t - 1 \) to \( t \). For each such group of individuals, we then compute their average income change from \( t \) to \( t + k \), for all values of \( k = 1, 2, 3, 5, 10 \).

Figure 16 takes all prime age workers who have median average earnings as of \( t - 1 \), and plots the impulse responses to shocks between \( t - 1 \) and \( t \) that are divided into 20 size categories. As seen here the most positive 5% of shocks at time \( t \) are about 170 log points and the most negative are of similar magnitude (but opposite sign). The mean reversion pattern also seems to vary with the size of the shock with much stronger mean reversion in \( t + 1 \) for large shocks and smaller reversion for smaller shocks. Furthermore, even at 10 year horizon, a non-negligible fraction of the shocks effect is still present, indicating a permanent component to these shocks.

In the next figure (17), we plot the same kind of impulse response function but now for workers that are in the 10th percentile (top panel) and 90th percentile (bottom) of past average earnings distribution. Comparing these figures to the previous one shows that for low income individuals, negative shocks seem to mean revert much more quickly, whereas positive shocks are more persistent. And the opposite is true for high income individuals.
Figure 17: Impulse Response Functions, Prime-Age Workers With Low and High Recent Incomes

Individuals with $\bar{Y}_{t-1} = P10$

Individuals with $\bar{Y}_{t-1} = P90$
To illustrate these patterns more clearly, we now plot the “shock” \((y_t - y_{t-1})\) on the x-axis and the fraction of that shock that has mean reverted on the y-axis \((y_{t+k} - y_t)\).

Figure 18 plots the same information as in Figure 16 (that is for individuals with median past average earnings) but using this alternative construct. Several remarks are in order. First, negative income changes tend to be less persistent than positive income changes. In fact, a worker that gets a 100 log point rise in earnings in time \(t\) loses about 50 percent of this gain in 10 years. It is interesting to note that almost all of this mean reversion happens after one year. As for an income decline of a similar magnitude, we see that the worker gains back about 70 percent of the initial income fall. Moreover, the recovery is more gradual. Only about 30 log points are recovered in the first year. Another important observation is with regard to the magnitude of income shocks. It is apparent in figure 18 that small shocks are more persistent than large shocks. The differences are striking. Income shocks that are smaller than 10 log points in absolute value look very persistent, whereas there is a substantial amount of mean reversion following larger income changes.
We now turn to the heterogeneity across the income distribution. The middle and bottom panels of figure 20 show the same plot for those at the 10th percentile (6-10) and 90th percentile (90-94) of the past income distribution. There are important differences. First, for workers that are poor with respect to average past income in the last 5 years, a negative income shock almost completely reverts to the mean in the next 10 years. However, there is very little mean reversion following a positive income shock. The opposite is true for the income rich. For them, it is the positive shocks that are almost completely transitory and the negative shocks that are almost completely permanent.

Figure 17 presents these facts with a canonical impulse function plot. It is easy to note that large income changes at time 0 are associated with more mean reversion compared to smaller income changes. The middle and bottom panels show the heterogeneity across the income distribution.

Figure 21 plots income growth in 10 years following an income shock in time $t$ for various income groups. It is evident that the pattern we observed in Figure 20 generalizes to more income groups.
Figure 20: Impulse Response to Income Shocks for Workers With Low and High Past Income, Prime-Age Workers

10th Percentile

90th Percentile
Figure 21: Asymmetric Mean Reversion: Butterfly Pattern, Baseline

Figure 22: Asymmetric Mean Reversion: Butterfly Pattern, Representative Agent
5 Estimation [To be Completed]

With the few exceptions noted above, the bulk of the literature relies on the (often implicit) assumption that income shocks can be approximated reasonably well with a log normal distribution. This assumption, combined with an AR(1) or random walk specification to capture the accumulation of such shocks, has made higher order moments irrelevant and allowed researchers to focus their estimation to match the covariance matrix of log income either in levels or in first difference form, with very few exceptions.\(^8\)

There are two issues with this approach. First, the broad range of evidence presented in the previous sections strongly suggests that this approach is likely to miss important aspects of the data—revealed in the rich structure of higher order moments—and produce a picture of income risk that does not capture salient features of the risks faced by workers. Second, the covariance matrix estimation method makes it difficult to select among various models of earnings risk, because the covariances that a given model matches well and those that it does not match well are difficult to compare from an economic standpoint.

Therefore, we follow a different route and estimate a stochastic process to match the moments of the data documented above. We believe that economists can much more easily judge whether each one of these sets of moments is relevant or irrelevant for the economic question they have in hand. Therefore, they can decide whether the inability of a particular stochastic process to match a given moment is a catastrophic failure or a simple nuisance. They can similarly judge the success of a given stochastic process in matching some moments and not others.

In this section, we use each one of the statistical properties documented above to infer what components are needed in a parametric model of earnings dynamics. We then estimate such a process to target moments whose economic significance is more immediate, including the distribution of lifetime income, the kurtosis and skewness of income changes, as well as how these moments vary with age and rising incomes. These moments are used as targets using a method of simulated moments (or more generally, an indirect inference) estimator.

\(^8\)Exceptions include Guvenen and Smith (2009), Browning et al. (2010), and Altonji et al. (2013).
5.1 A Flexible Stochastic Process

The baseline process we consider has the following features: (i) a heterogeneous income profiles component of quadratic form, (ii) a mixture of two AR(1) processes, denoted by \( z \) and \( x \), that vary in their persistence and innovation variance, where each AR(1) process receives a new innovation in a given year with probability \( p_j \in [0,1] \) for \( j = z, x \); and (iii) an i.i.d transitory shock. We also estimate simpler versions that are nested by this specification as well as more general versions of it.

Here is the full specification

\[
\begin{align*}
\bar{y}_t^i &= (\alpha^i + \beta^i t + \gamma^i t^2) + z_t^i + x_t^i + \varepsilon_t^i \\
z_t^i &= \rho_z z_{t-1}^i + \eta_{zt}^i \\
x_t^i &= \rho_x x_{t-1}^i + \eta_{xt}^i 
\end{align*}
\]

where for \( j = z, x \):

\[
\eta_{jt}^i \begin{cases} 
= 0 & \text{w.p } 1 - p_j \\
\sim \mathcal{N}(\mu_j, \sigma_j) & \text{w.p } p_j
\end{cases}
\]

To avoid indeterminacy, we impose \( \rho_x > \rho_z \), both without loss of generality.

The infrequent nature of income shocks to either of the two AR(1) components help us generate the excess kurtosis in the data. Allowing the means of innovations to differ helps with negative skewness that is quite salient in the data. Furthermore, allowing infrequent shocks with non-zero mean to components with different degrees of mean reversion helps us fit the heterogeneity in mean reversion. For example, it could be that one AR(1) component has very large innovations but a lower persistence. In this case, shocks to this component are likely to show up more in the tails. This composition could explain why large shocks are more mean reverting than small shocks. We do not impose any such restriction on the process, but instead let the data speak for itself.

Note that the distribution of shocks varies greatly both over the life cycle and over the income distribution. To allow the model to fit these features of the data, we model the standard deviation of innovations as a function of age and the persistent component.
of income:

\[ \sigma_{jt}(z + x) = a_j + b_j \times (z + x) + c_j \times t + d_j \times (z + x) \times t. \]

Here, we allow the shock variance to change with the persistent component of earnings, \( z + x \), as well as with age \( t \) and the interaction of the two. Again, because this is a specification for variance, we impose a positivity condition on it.\(^9\)

### 5.2 Estimation Method

We estimate the parameters of the income process just described using the method of simulated moments (MSM). There are three sets of target moments from the data we aim to match in our estimation:

1. **Cross-sectional moments of earnings changes.** The first set of moments we target in our estimation is the selected percentiles of the distribution of 1- and 5-year earnings changes \( \Delta y_{it+k} = y_{it+k} - y_{it}, k = 1, 5 \). \( \log y_{it} \) is defined as the log of residual earnings net of age effects. There are 30 percentiles from the distribution of earnings changes that we target in our estimation: 1st through 10th, 20th, 25th, 30th, 40th, 50th, 60th, 75th, 80th, 90th, 91st, 92nd,...99th, 99.5th. Furthermore, in order to capture the variation in the cross sectional moments of earnings changes along the age and average past income dimensions, we condition the distribution of earnings changes on these variables. For this purpose, we first group workers into 6 age bins (5-year age bins between 25 and 54) and within each age bin into 100 average past income \( \bar{Y}_{i_{t-1}} \) percentiles in age \( t - 1 \). Thus, we compute the selected percentiles from the distribution of earnings changes for 600 different groups of workers. To reduce the computational time in our estimation, we aggregate these

\[^9\]We have also considered an alternative specification where the mixing probabilities are functions of income and age:

\[ p_j(v + z + x) = a_j + b_j \times (v + z + x) + c_j \times t + d_j \times (v + z + x) \times t. \]

In this specification, the probability of each shock being realized will be changing with income. As income changes with age, this could capture both an age structure and a cross-sectional structure. Since this is a probability however, \( a_j, b_j, c_j \) must be chosen so as to ensure that \( p_j \) stays bounded between 0 and 1.
600 groups of workers into 2 age groups, $A_{i-1}$, and 13 average past income groups, $\tilde{Y}_{i-1}$ (i.e., $\Delta y_{it+k} \sim F(y_{it+k} - y_{it-1} | A_{i-1}, \tilde{Y}_{i-1})$). The first age group is defined as young workers between ages 25 through 34, whereas the second age group is defined as prime-age workers between the ages of 35 and 54. The average past income percentiles are grouped as follows: 1, 2-10, 11-0, 21-30, ..., 81-90, 91-95, 96-99, 100. Consequently, we end up with $30 \times 2 \times 13 = 780$ cross-sectional moments.

2. **Mean of Log Income Growth.** The second set of moments captures the heterogeneity in log income growth over the working life across workers that are in different percentiles of the lifetime income distribution. To that end, we target the average dollar earnings (in terms of $1000$) at age 25, 30, ..., and 60 for different lifetime income groups. First, we select our sample according to the criteria discussed in section 4.5 and construct each individual’s lifetime income as discussed in the same section. Then, we rank individuals according to their lifetime income. The lifetime income groups used in our estimation are: 1st, 2nd-5th, 6th-10th, 11th-20th, 21st-30th, 31st-40th, ..., 81st-90th, 91st-95th, 96th-97th, 98th-99th, 100th percentiles of the lifetime income distribution. Thus, the total number of moments we target in this set is $8 \times 15 = 120$.

3. **Impulse Response Functions.** Our final set of moments is informative about the dynamics of earnings change. These moments help us identify the persistence of labor income shocks. Here, we target average log income growth over the next $k$ years for $k = 1, 2, 3, 5, 10$; i.e., $E[y_{t+k} - y_t]$, conditional on age and average past income in $t-1$ as well as the income growth between $t$ and $t-1$ ($y_t - y_{t-1}$; see section 6 for details). We focus on two age groups: young workers (25-34) and prime-age workers (35-55). In each year, individuals are assigned to groups based on where they stand in the average past income distribution among individuals of the same age. The following list defines these groups in terms of past income percentiles: 1-5, 6-10, 11-30, 31-50, 51-70, 71-90, 91-95, 96-100. We then group workers based on the percentiles of the income change between $t$ and $t-1$. The following list defines such groups: 1-2, 3-5, 6-10, 11-20, 21-30, 31-40, 41-50, 51-60, 61-70, 71-80, 81-90, 91-95, 96-98, 99-100. As a result we have a total of $2 \times 8 \times 14 \times 5 = 1120$ moments regarding the dynamics of earnings changes.
In sum, we target a total of $780 + 120 + 1120 = 2020$ moments. Let $m_n$ be one of the moments from the data, where $n = 1, \ldots, N = 2020$ and let $d_n(\theta, X)$ be the corresponding model moment which is simulated for a given set of income process parameters, $\theta$ and a given vector of random variables $X$. We simulate the entire earnings histories of 100,000 individuals who enter labor market at age 25 and retire at age 60. When computing the model moments, we apply the same sample selection criteria and employ the same methodology to the simulated data as we did with the actual data.

To deal with potential issues that could arise due to the large variation in the scales of the moments, we minimize the “scaled” deviation between each data target and the corresponding simulated model moment. For each moment $n$, let’s define:

$$F_n(\theta) = \left(\frac{d_n(\theta, X) - m_n}{|m_n| + \gamma_n}\right)$$

where $\gamma_n$ is an adjustment factor. When $\gamma_n = 0$ and $m_n$ is positive, $F_n$ is simply the percentage deviation between data and model moments. This measure becomes problematic when the data moment is very close to zero, which is the case for many moments (e.g., median of log income changes). To account for this, we choose $\gamma_n$ to be the 10th percentile of the absolute values of all moments in a given set.\(^{10}\)

The MSM estimator is

$$\hat{\theta} = \arg \min_{\theta} F(\theta)'W F(\theta)$$

(1)

where $F(\theta)$ is a vector in which all moment conditions are stacked, i.e.,

$$F(\theta) = [F_1(\theta), \ldots, F_N(\theta)].$$

The weighting matrix, $W$ is chosen such that each set of moments is given the same weight in the objective function. Namely, each cross-sectional moment is weighted by $1/780$, each mean income growth moment is weighted by $1/120$, and finally, each im-

\(^{10}\)We divided the first set of moments into four subsets: 1-year change of young workers, 5-year change of young workers, 1-year change of prime-age workers, 5-year change of prime-age workers (the corresponding $\gamma_n$ values are 0.94, 1.44, 0.90, 1.53, respectively). Furthermore, $\gamma_n = 38$ and $\gamma_n = 0.35$ for mean of log income growth moments and impulse response function moments, respectively.
pulse response function moment is weighted by $1/1120$. The objective function is highly jagged in certain directions and highly non-linear in general, owing to the fact that we target higher-order moments and percentiles of the distribution (the latter creating kinks). Therefore, we employ a global optimization routine, described in further detail in Guvenen (2013), to perform the minimization in (1).

The global stage is a multi-start algorithm where candidate parameter vectors are uniform Sobol’ (quasi-random) points. We typically take about 10,000 initial Sobol’ points for pre-testing and select the best 2000 points (i.e., ranked by objective value) for the multiple restart procedure, although using about one-fifth of these numbers often delivers the same global minimum. 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 et al. (2010), which is much faster but has a higher tendency to be stuck at local minima. The combination balances speed with reliability and provides the best results.

5.3 Results

[Coming soon..]

6 Conclusions

Our analysis of life cycle earnings histories of millions of U.S. workers has reached two broad conclusions. One, the higher order moments of individual earnings shocks are important, because the distribution of shocks is very different from log-normal. In particular, earnings shocks display strong negative skewness (individual “disaster” shocks) and extremely high kurtosis—as high as 35 compared with 3 for a Gaussian distribution. The high kurtosis implies that in a given year, most individuals experience very small earnings shocks, but very few experience extremely large shocks. Second, these statistical properties of earnings shocks change substantially both over the life cycle and with the income level of individuals. We also estimate impulse response functions of income shocks and find significant asymmetries: positive shocks to high-income individuals
are quite transitory, whereas negative shocks are very persistent; the opposite is true for low-income individuals. While these statistical properties are typically ignored in quantitative analyses of life cycle models, they are fully consistent with search-theoretic models of careers over the life cycle. After establishing these empirical facts non-parametrically, we estimated what we think is the simplest income process broadly consistent with these salient features of the data.

References


A Appendix: Additional Figures

A.1 Cross-Sectional Facts

A.1.1 Standard Deviation

A.1.2 Skewness

A.1.3 Kurtosis
Figure 23: Standard Deviation of Annual Earnings Growth
Figure 24: Standard Deviation of Biann 5-Year Earnings Growth

Figure 25: Standard Deviation of 5-Year Earnings Growth, Usual
Figure 26: Skewness of (Bi-Ann) Annual Earnings Growth

Figure 27: Skewness of Annual Earnings Growth, Usual
Figure 28: Skewness of (Bi-Ann) 5-Year Earnings Growth

Figure 29: Skewness of 5-Year Earnings Growth, Usual
Figure 30: Skewness of 1-Year Earnings Growth, Excluding bottom and top 1% of income change

Figure 31: Skewness of 5-Year Earnings Growth, Excluding bottom and top 1% of income change
Figure 32: Skewness of 1-Year Earnings Growth, Excluding those with income lower than 5% of average past income

Figure 33: Skewness of 5-Year Earnings Growth, Excluding those with income lower than 5% of average past income
Figure 34: Kurtosis of Annual Income Change, By Age and Past Income, (2 Year Average Income)

Figure 35: Kurtosis of Annual Income Change, By Age and Past Income, Usual
Figure 36: Kurtosis of 5-Year Income Change, By Age and Past Income, (2 Year Average Income)

Figure 37: Kurtosis of 5-Year Income Change, By Age and Past Income, Usual
Figure 38: Kurtosis of 1-Year Income Change, By Age and Past Income, Excluding bottom and top 1% of income change

![Figure 38](image1.png)

Figure 39: Kurtosis of 5-Year Income Change, By Age and Past Income, Excluding bottom and top 1% of income change

![Figure 39](image2.png)
(a) Kurtosis of 1-Year Income Change, Excluding those with income lower than 5% of average past income

(b) Kurtosis of 5-Year Income Change, Excluding those with income lower than 5% of average past income

Figure 40: Kurtosis, robustness
B Time-Series Moments by Individual

Sample Selection for Individual Regressions. We rank individuals based on their lifetime income, defined as the sum of their earnings in ages 25 through 60. Income observations lower than the minimum threshold are set to this threshold. We require that an individual has at least 32 years of data. We further restrict our sample to individuals that i) have earnings above the threshold for at least two thirds of the time they are in our sample and ii) are not self-employed for more than 8 years.\textsuperscript{11} This criterion leaves us with 4 cohorts that were born between 1951 and 1954.

For each worker, we compute annual income growth rates as differences in logs as well as the residuals from a regression of log earnings on a quadratic polynomial of age. For each of these measures, we then compute the following individual-specific statistics: standard deviation, skewness, kurtosis, a robust measure of dispersion, and the median.\textsuperscript{12} We then compute and report several sample statistics regarding the distribution of of individual-specific statistics, both for the overall sample and also by the percentiles of the lifetime income distribution.

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{figure41.png}
\caption{Mean Years Non-Employed, by Lifetime Income Percentile}
\end{figure}

\textsuperscript{11}Self employment is defined as before; i.e. having a self employment income larger than 10\% of total wage and salary income and the minimum income threshold.

\textsuperscript{12}The robust dispersion measure is defined as the difference between the third largest and the third smallest observation.
Figure 42: Within-Group Moments of Time Series Kurtosis of Individual Income Growth

B.1 Kurtosis

B.2 Skewness

Figure 43: Within-Group Mean of Time-Series Skewness of Individual Income Growth
B.3 Standard Deviation

Figure 44: Within-Group Std. Dev. of Time-Series Median of Individual Income Growth

Figure 45: Within-Group Mean of Time-Series Std. Dev. of Individual Income Growth

The bottom panel of the same figure plots the standard deviation for 5-year changes. The general patterns are quite similar, with the exception of a stronger rise in dispersion after age 50 for middle to high income individuals. However, notice that this group will
be close to age 60 in year $t + 5$ so some of this higher variance is likely to be associated with reduction in hours. The skewness results we discuss in a moment supports this hypothesis.

Overall, it is worth noting that the strong dependence of the standard deviation on past levels of income documented here is not something that existing stochastic processes for earnings allow for. We will incorporate this feature into the econometric process we estimate in the coming sections.

**B.4 Average Income Growth**

Figure ?? plots the percentiles of average annual income growth over the lifecycle for individuals ranked by lifetime earnings. As seen here, income growth rate is robustly higher for individuals with higher lifetime income with the exception of the bottom 10% of the distribution.
Figure 47: Within-Group Skewness of Time-Series Std. Dev. of Individual Income Growth

Figure 48: Within-Group Moments of Time-Series Mean of Individual Income Growth
Figure 49: Correlation Between Mean and Std. Dev. of Growth Rates

Figure 50: Correlation Between Skewness and 1st and 2nd Moments
Figure 51: Correlation Between Kurtosis and Dispersion
Figure 52: Correlation Between Years Non-Employed and Life Cycle Statistics