What Do Data on Millions of U.S. Workers Reveal about Life-Cycle Earnings Risk?*

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

December 11, 2014

Abstract

We study the evolution of individual labor earnings over the life cycle using a very large panel data set from the U.S. Social Security Administration. Using fully nonparametric methods, our analysis reaches two broad conclusions. First, earnings shocks display substantial deviations from lognormality—the standard assumption in the incomplete markets literature. In particular, earnings shocks display strong negative skewness 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 vary significantly both over the life cycle and with the earnings level of individuals. We also estimate impulse response functions of earnings shocks and find important asymmetries: positive shocks to high-earnings individuals are quite transitory, whereas negative shocks are very persistent; the opposite is true for low-earnings individuals. Finally, we use these rich sets of moments to estimate econometric processes with increasing generality to capture these salient features of earnings dynamics.

JEL Codes: E24, J24, J31.
Keywords: Earnings dynamics, lifecycle earnings risk, nonparametric estimation, kurtosis, skewness, non-Gaussian shocks, normal mixture.

*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; www.fguvenen.com
‡Federal Reserve Bank of New York; fatih.karahan@ny.frb.org
§University of Toronto; serdar.ozkan@utoronto.ca
¶Social Security Administration; jae.song@ssa.gov

1 Introduction

This year about 2 million young American men will enter the labor market for the first time. Over the next 40 years, each of these men will go through his unique adventure in the labor market, involving a series of surprises—finding an attractive career, being offered a dream job, getting promotions and salary raises, and so on—as well as disappointments—experiencing unemployment, failing in one career and moving on to another one, suffering health shocks, and so on. These events will vary not only in their initial significance (upon impact), but also in how durable their effects turn out to be in the long run.[1]

An enduring question for economists is whether these wide-ranging labor market histories experienced by a diverse set of individuals display sufficiently simple regularities that would allow researchers to characterize some general properties of earnings dynamics over the life cycle. Despite a vast body of research since the 1970s, it is fair to say that many aspects of this question still remain open. For example, What does the probability distribution of earnings shocks look like? Is it more or less symmetric, or does it display important signs of skewness? More generally, how well is it approximated by a lognormal distribution, an assumption often made out of convenience? And, perhaps more important, how do these properties differ across low- and high-income workers or change over the life cycle? A host of questions also pertain to the dynamics of earnings. For example, how sensible is it to think of a single persistence parameter to characterize the durability of earnings shocks? Do positive shocks exhibit persistence that is different from negative shocks? Clearly, we can add many more questions to this list, but we have to stop at some point. If so, which of these many properties of earnings shocks are the most critical in terms of their economic importance and therefore should be included in this short list, and which are of second-order importance?

One major reason why many of these questions remain open has been the heretofore unavailability of sufficiently rich panel data on individual earnings histories.[2] Against

[1] In this paper, we focus on the earnings dynamics for men so as to abstract away from the complexities of the female nonparticipation decision. We intend to undertake a similar study that focuses on earnings dynamics for women.

[2] With few exceptions, most of the empirical work in this area has been conducted using the Panel Study of Income Dynamics (including the previous work of the authors of this paper), which contains between 500 to 2000 households per year depending on the selection criteria, and suffers from shortcomings that are typical of survey data, such as survey response error, attrition, and so on.

this backdrop, the goal of this paper is to characterize the most salient aspects of life-cycle earnings risk using a large and confidential panel data set from the U.S. Social Security Administration. The substantial sample size—of more than 200 million individual-year observations from 1978 to 2010—allows us to employ a fully nonparametric approach and take "high-resolution pictures" of individual earnings histories.

In deciding what aspects of the earnings data to focus on, this paper is motivated by a growing body of theoretical work (reviewed in the next section), which attributes a central role to skewness and kurtosis of economic variables for questions ranging from the effects of monetary policy to optimal taxation, and from the determinants of wealth inequality to asset prices. Therefore, we focus on the first four moments of earnings changes over the life cycle. This analysis reaches two broad conclusions. First, the distribution of individual earnings shocks displays important deviations from lognormality. Second, the magnitude of these deviations (as well as a host of other statistical properties of earnings shocks) varies greatly both over the life cycle and with the earnings level of individuals. Under this broad umbrella of "non-normality and life-cycle variation," we establish four sets of empirical results.

First, starting with the first moment, we find that average earnings growth over the life cycle varies strongly with the *level* of lifetime earnings: The median individual by lifetime earnings experiences an earnings growth of 38% from ages 25 to 55, whereas for individuals in the 95th percentile, this figure is 230%; and for those in the 99th percentile, this figure is almost 1500%. We show that the workhorse model in the literature—a stochastic process featuring a persistent plus a transitory component with normal innovations—fails to generate this heterogeneity for plausible parameter values.[3] This heterogeneity can be generated more easily by either using a mixture of autoregressive processes or allowing heterogeneous earnings growth rates.

Second, turning to the third moment, we see that earnings shocks are negatively skewed, and this skewness becomes more severe as individuals get older or their earnings increase (or both). Furthermore, this increasing negativity is due entirely to upside earnings moves becoming smaller before age 45 and to increasing "disaster" risk (the risk of a sharp fall in earnings) after age 45. Although these implications are quite plausible,

[3] A positive relationship between lifetime earnings and life-cycle earnings growth is to be expected (since, all else equal, those who experience fast earnings growth will have higher lifetime earnings). What is surprising is the magnitudes involved, which turns out to be hard to match empirically.

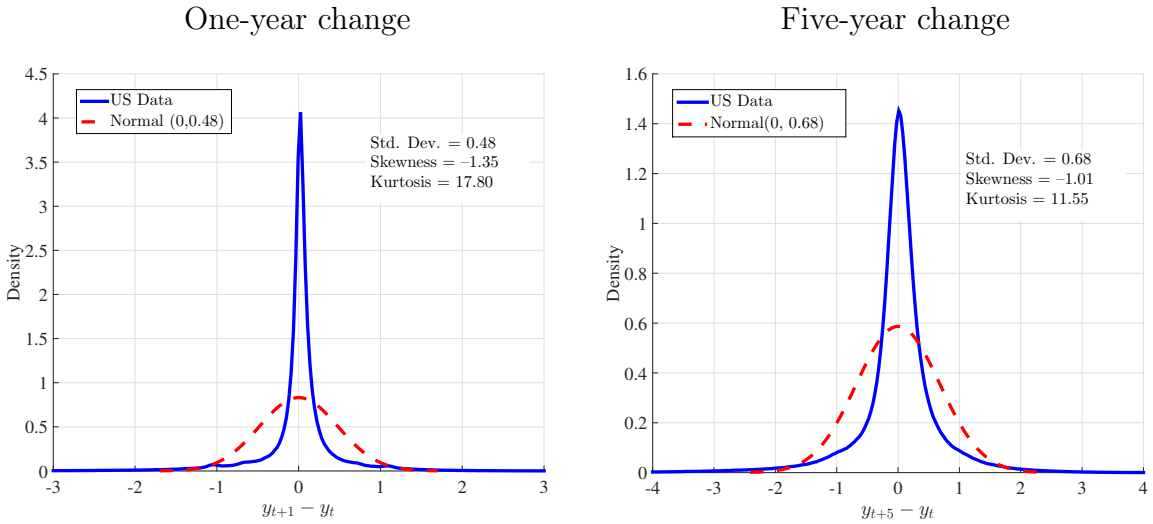


FIGURE 1 – Histogram of Log Earnings Changes. Note: The first year t is 1995 and the data are for all workers in the base sample defined in Section 2.

they are not captured by a lognormal specification, which implies zero skewness.

Third, studying the fourth (standardized) moment, we find that earnings changes display very high kurtosis. What kurtosis measures is most easily understood by looking at the histogram of log earnings changes, shown in Figure 1 (left panel: annual change; right panel: five-year change). Notice the sharpness in the peak of the empirical density, 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. Thus, there are far more people with very small earnings changes in the data compared with what would be predicted by a normal density. Furthermore, this *average* kurtosis masks significant heterogeneity across individuals by age and earnings: 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.

Fourth, we characterize the dynamics of earnings shocks by estimating non-parametric impulse response functions conditional on the recent earnings of individuals and on the size of the shock that hits them. We find two types of asymmetries. One, fixing the shock size, positive shocks to high-earnings individuals are quite transitory, whereas negative shocks are very persistent; the opposite is true for low-earnings individuals. Two, fixing

the earnings 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. To our knowledge, both of these findings are new in the literature. These kinds of asymmetries are hard to detect via the standard approach in the literature, which relies on the autocovariance matrix of earnings—the second cross-moments of the panel data.[4] 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 in aggregate data (e.g., Christiano *et al.* (2005), Borovicka *et al.* (2014)).

While this nonparametric approach allows us to establish key features of earnings dynamics in a robust fashion, a tractable parametric process is indispensable for conducting quantitative economic analyses. The standard approach in the earnings dynamics literature is to estimate the parameters of linear time-series models by matching the variance-covariance matrix of log earnings residuals. There are two difficulties with this approach. First, the strong deviations from lognormality documented in this paper calls into question the wisdom of focusing exclusively on covariances at the expense of the rich variation in higher-order moments, which will miss key features of earnings risk faced by workers. Second, the covariance matrix approach makes it difficult to select among alternative econometric processes, because it is difficult to judge the relative importance—from an economic standpoint—of the covariances that a given model matches well and those that it does not. This is an important shortcoming given that virtually every econometric process used to calibrate economic models is statistically rejected by the data.

With these considerations in mind, in Section 5, we follow a different route and target the four sets of empirical moments—broadly corresponding to the first four moments of earnings changes—described above, employing a method of simulated moments (MSM) estimator. We believe this is a more transparent approach: economists can more easily judge whether or not each of these moments is relevant for the economic questions 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 tolerable shortcoming. They can similarly judge the success of a given stochastic process in matching some moments and not others. We estimate a set of stochastic processes with increasing generality

[4] This is because a covariance lumps together all sorts of earning changes—large, small, positive, and negative—to produce a single statistic. This approach, while economical in its use of scarce data, masks lots of interesting heterogeneity revealed by our analysis.

to provide a reliable "user's guide" for applied economists. We find that a stochastic process that features a heterogenous growth component and, more importantly, a mixture of AR(1) processes where the mixture probabilities depend on age and recent earnings can match most of the documented features fairly well.

In Section 2, we describe the data and the empirical approach. Section 3 presents the findings on the cross-sectional moments of earnings growth and Section 4 presents the impulse response analysis. Section 5 describes the parametric estimation and discusses its results. Section 6 concludes.

Related Literature

Since its inception in the late 1970s,[5] the earnings dynamics literature has worked with the implicit or explicit assumption of a Gaussian framework, thereby making no use of higher-order moments beyond the variance-covariance matrix. One of the few exceptions is Geweke and Keane (2000) who emphasized the non-Gaussian nature of earnings shocks and fit a normal mixture model to earnings innovations. More recently, Bonhomme and Robin (2009) analyze French earnings data over short panels and model the transitory component as a mixture of normals and the dependence patterns over time using a copula model. They find the distribution of this transitory component to be left-skewed and leptokurtic. In this paper, we go beyond the overall distribution and establish the substantial variation in the degree of non-normality with age and earnings levels. Furthermore, the impulse response analysis (Section 4) shows the need for a different persistence parameter for large and small shocks, which we find to be better captured as a mixing of AR(1) processes, which is a step beyond the normal mixture model.[6]

Incorporating higher-order moments of earnings dynamics into economic models is still in its infancy.[7] In an early attempt, Castañeda *et al.* (2003) build a richly-specified incomplete markets model and choose the parameters of the earnings process so as to match the wealth inequality in the U.S. data. The resulting process turns out to display excess kurtosis for earnings changes, consistent with our direct measurements from

[5] Lillard and Willis (1978); Lillard and Weiss (1979); Hause (1980); and MaCurdy (1982).

[6] Geweke and Keane (2007) study how regression models can be smoothly mixed, and our modeling approach shares some similarities with their framework.

[7] Higher-order moments are gaining a more prominent place in recent work in monetary economics (e.g., Midrigan (2011) and Berger and Vavra (2011); see Nakamura and Steinsson (2013) for a survey) as well as in the firm dynamics literature (e.g., Bloom *et al.* (2011) and Bachmann and Bayer (2014)).

the U.S. data. Guvenen *et al.* (2014) document that the skewness of individual income shocks are strongly countercyclical, whereas the variance is acyclical. Building on this observation, Constantinides and Ghosh (2014) show that an incomplete markets asset pricing model with countercyclical skewness shocks generates plausible asset pricing implications, and McKay (2014) studies aggregate consumption dynamics in a business cycle model that is calibrated to match these skewness shocks. Turning to fiscal policy, Golosov *et al.* (2014) show that using an earnings process with negative skewness and excess kurtosis (targeting the empirical moments reported in this paper) implies a marginal tax rate on labor earnings for top earners that is substantially higher than under a traditional calibration with Gaussian shocks with the same variance.

Methodologically, our work is most closely related to two important recent contributions. Altonji *et al.* (2013) estimate a joint process for earnings, wages, hours, and job changes, targeting a rich set of moments via indirect inference. Browning *et al.* (2010) also employ indirect inference to estimate an earnings process featuring "lots of heterogeneity" (as they call it). However, neither paper explicitly focuses on higher-order moments or their life-cycle evolution. The latter paper does model heterogeneity across individuals in innovation variances, as do we, and finds a lot of heterogeneity along that dimension in the data. In ongoing research, Arellano *et al.* (2014) also explore differences in the mean-reversion patterns of earnings shocks across households that differ in their earnings histories. Using data from the Panel Study of Income Dynamics, they find asymmetries in mean reversion consistent with those we document in Section 4.

Relatively little work has been done on the life-cycle evolution of earnings dynamics, which is the main focus of this paper. A few papers (including Baker and Solon (2003), Meghir and Pistaferri (2004), Karahan and Ozkan (2013), and Blundell *et al.* (2014)) allow age-dependent innovation variances, but do not explore variation in higher-order moments. Our conclusion on the variance is consistent with this earlier work, indicating a decline in variance from ages 25 to 50, with a subsequent rise. However, quantitatively, the decline is modest relative to the variation in other moments.[8]

[8] Meghir and Pistaferri (2004) allow for a GARCH component in earnings, but they do not discuss the implications of the estimated process for higher-order moments.

2 Empirical Analysis

2.1 The SSA Data

The data for this paper come from the Master Earnings File (MEF) of the U.S. Social Security Administration records. The MEF is the main source of earnings data for the SSA and contains information for 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 number for the last digit and select all individuals in 1978 whose SSN ends with that number.[9] This process yields a 10% random sample of all SSNs issued in the United States in or before 1978. Using SSA death records, we drop individuals who are deceased in or before 1978 and further restrict the sample to those between ages 25 and 60. In 1979, we continue with this process of selecting the same last digit of the SSN. Individuals who survived from 1978 and who did not turn 61 continue to be present in the sample, whereas 10% of new individuals who just turn 25 are automatically added (because they will have the last digit we preselected), and those who died in or before 1979 are again dropped. Continuing with this process yields a 10% representative sample of U.S. males in every year from 1978 to 2010. Finally, the MEF has a small number of extremely high earnings observations. In each year, we cap (winsorize) observations above the 99.999th percentile in order to avoid potential problems with these outliers.

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

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 for which the statistic is constructed.

For a given year, the base sample is constructed as follows. First, we restrict attention to individuals between the ages of 25 and 60 to focus on working-age population. Second, we select workers whose annual wage/salary earnings exceeds a time-varying minimum threshold, denoted by $Y_{\min,t}$, defined as one-fourth of a full-year full-time (13 weeks at 40 hours per week) salary at half of the minimum wage, which amounts to 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 earnings 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 exceed a threshold level, defined as the maximum of $Y_{\min,t}$ 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

In the nonparametric analysis conducted in Sections 3 to 5, our main focus will be on individual-level log earnings changes (or growth) at one-year and five-year horizons. These earnings changes provide a simple and useful construct for discussing the dynamics of earnings without making strong parametric assumptions. In Section 5, we will link these "changes" to underlying "shocks" or "innovations" to an earnings process by means of a parametric estimation.

To examine how the properties of earnings growth vary over the life cycle and in the cross section, we proceed as follows. Let $\tilde{y}^i_{t,h}$ denote the log earnings of individual i who is h years old in year t. For each one- and five-year horizon starting in period t, we group individuals based on their age and *recent earnings* (hereafter, RE—to be defined precisely in a moment) as of time $t-1$. If these groupings are done at a sufficiently fine level, we can think of all individuals within a given age/recent-earnings group to be

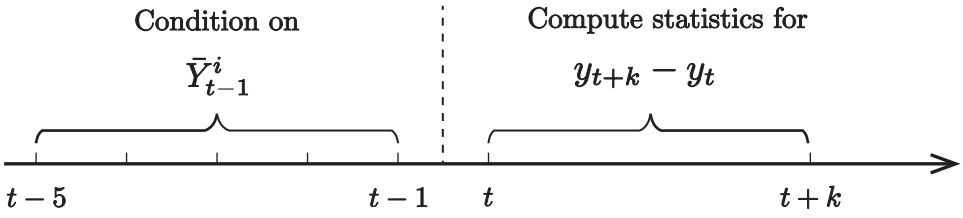


FIGURE 2 – Timeline For Rolling Panel Construction

ex ante identical (or at least very similar). Then, for each such group, we can compute the cross-sectional moments of earnings changes between t and $t+k$ ($k = 1, 2...$), which can then be viewed as corresponding to the properties of shocks that individuals within each bin can expect to face (see Figure 2 for this rolling sample construction). This approach has the advantage that we can compute higher-order moments precisely, as each bin contains several hundred thousands of individuals.[10] (Table I reports sample size statistics.)

We implement this approach by first grouping workers into five-year age bins based on their age in year $t-1$: 25–29, 30–34,..., 50–54, and 55–60. Then taking all individuals within an age group in a given year t, we compute each worker's average past earnings between years $t-1$ and $t-5$, denoted with $\tilde{Y}^i_{t-1} \equiv \sum_{s=1}^{5} \exp(\tilde{y}^i_{t-s,h-s})$. We set earnings observations below $Y_{\min,t}$ to the threshold for this computation. We also further control for age effects, because even within these narrowly defined age groups, age differences of a few years can systematically skew rankings in favor of older workers. To avoid this, we first estimate age dummies, denoted d_h, corresponding to log average earnings at each age,[11] and construct five-year average earnings from ages $h-5$ to $h-1$: $\sum_{s=1}^{5} \exp(d_{h-s})$. We then normalize $\tilde{Y}^i_{t-1}$ with this measure to clean age effects. Thus, our measure of

[10] A second possibility is that the properties of shocks depend more intimately on the characteristics individuals (such as the health, relationship skills, stamina, etc.), and different *types* of workers—for example, identified by their lifetime earnings—might face shocks with different properties. This suggests that perhaps 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 similar substantive conclusions, so we omit these results from the main text. These results are available upon request.

[11] These are estimated from a pooled regression of log earnings on age and cohort dummies. Further details are given later in Section 3.2.

TABLE I – Sample Size Statistics for Cross-Sectional Moments

	# Observations in Each RE Percentile Group			
Age group	Mean	Median	Min	Max
25-29	408,713	381,606	337,603	674,986
30-34	638,350	640,596	566,235	712,307
35-39	618,909	642,226	441,425	721,966
40-44	571,512	606,344	356,700	707,502
45-49	506,989	524,177	320,935	686,753
50-54	425,539	383,887	291,117	619,987
55-59	276,340	240,273	215,842	407,997

Note: Each entry reports the statistics of the number of observations in each of the 100 RE 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.

recent earnings (hereafter, RE) is

$$\bar{Y}^i_{t-1} \equiv \frac{\tilde{Y}^i_{t-1}}{\sum_{s=1}^{5} \exp(d_{h-s})}.$$

In much of the subsequent analysis, we shall group workers based on this variable—typically into 100 percentile bins—and will refer to these as "recent earnings groups."

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 (one-year) and long (five-year) horizons. For computing moments of earnings growth, we work with the time difference of y^i_t, which is log earnings *net* of the age effect. Thus:

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

We compute the cross-sectional moments of $\Delta_k y^i_t$ for each year, $t = 1980, 1981, ..., 2009$ and then average these across all years.[12]

[12] We use $t = 1980$ as the first year of our analysis and therefore group individuals in 1979 based on

Final Sample for Cross-Sectional Moments. To construct these moments, our *final sample* includes all observations that are in the base sample in $t-1$ and in at least two more years between $t-5$ and $t-2$, and furthermore satisfies the age (25–60) and no-self-employment condition in years t and $t+k$. We also drop the oldest age group (55- to 60-year-olds) and assign individuals within each group into 100 percentile bins based on their recent earnings, $\bar{Y}^i_{t-1}$. Table I reports the summary statistics of the number of observations in each age/earnings cell (summed over all years). As seen here, the sample size is very large—the smallest cell size exceeds 200,000 observations and the average is close to 500,000—which allows us to compute all statistics very precisely.

3.1 First Look at Data

Before delving into a disaggregated analysis of the data, we take a first look at some broad statistics. First, recall that the left panel of Figure 1 plotted the histogram of annual earnings changes in the U.S. data (solid blue 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 (annual) standard deviation of 0.48, a negative skewness of –1.35, and most strikingly, an extremely high kurtosis of 17.80.[13] The right panel plots the histogram for five-year earnings changes. Although the distribution is a little less concentrated around zero, it is still far from a normal distribution, with a kurtosis of 11.55 and a negative skewness of –1.01; it also has a large standard deviation of 0.68.

However, if we stop here we will be missing lots of heterogeneity that lies underneath this overall picture. To show what we mean, Figure 3 plots 12 histograms, each for a different age and RE percentile group. The top row plots annual earnings changes for young workers (ages 25 to 34), and the next row plots the same for prime-age males (ages 45 to 50). From left to right, the three panels plot the histograms for individuals who are in the 10th, 50th, and 90th percentiles of the past average earnings distribution (indicated with P10, P50, and P90 in the figure). Although we will elaborate on these features in the

their recent earnings computed over 1978 and 1979. Similarly, 2009 is the last feasible year for t, which allows us to construct the moments of one-year earnings changes between 2009 and 2010.

[13] A distribution with a kurtosis of 5 or 6 is considered to be highly leptokurtic.

coming sections, 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 important variation over the life cycle and across earnings levels.

3.2 First Moment: Mean of Log Earnings Growth

We begin our analysis with the first moment—average earnings growth—and examine how it varies with age (i.e., over the life cycle) and, for reasons that will become clear in a moment, across groups of individuals that differ in their *lifetime earnings* (and not recent earnings). But first, to provide a benchmark, we follow the standard procedure in the literature, (e.g., Deaton and Paxson (1994)) to estimate the average life-cycle profile of log earnings. Although the procedure is well understood, its details matter for some of the discussions below, so we go over it in some detail.

The average life-cycle profile is obtained from panel data or repeated cross sections by first grouping earnings observations of all individuals with the same age and in the same birth cohort into age-cohort cells. If there is a maximum of H ages and C cohorts, there are a maximum of $H \times C$ non-empty such cells. We then compute mean earnings within each cell and then regress the logarithm of these means ($H \times C$ observations) on H age dummies and C cohort dummies. The estimated age dummies are plotted as circles in Figure 4 and represent the average life-cycle profile of log earnings. It has the usual hump-shaped pattern that peaks around age 50. (On a side note, these age dummies turn out to be indistinguishable from a fourth order polynomial of age,[14] a point also observed by Murphy and Welch (1990) in Current Population Survey data.)

One of the most important aspects of a life-cycle profile is the implied growth in average earnings over the life cycle (e.g., from ages 25 to 55). It is well understood that the magnitude of this rise matters greatly for many economic questions, because it is a strong determinant of borrowing and saving motives.[15] In our data, this rise is about 80 log points, which is about 127%.[16] Notice that feeding this life-cycle profile into a

[14] Regressing the age dummies on a fourth order polynomial of age yields an R^2 of 0.999.

[15] See, for example, Deaton (1991); Attanasio *et al.* (1999); and Gourinchas and Parker (2002), among others.

[16] This figure lies on the high end of previous estimates from data sets such as the PSID, but not unseen before (cf. Attanasio *et al.* (1999)).

FIGURE 3 – Histogram of Log Earnings Growth, by RE and Age Groups

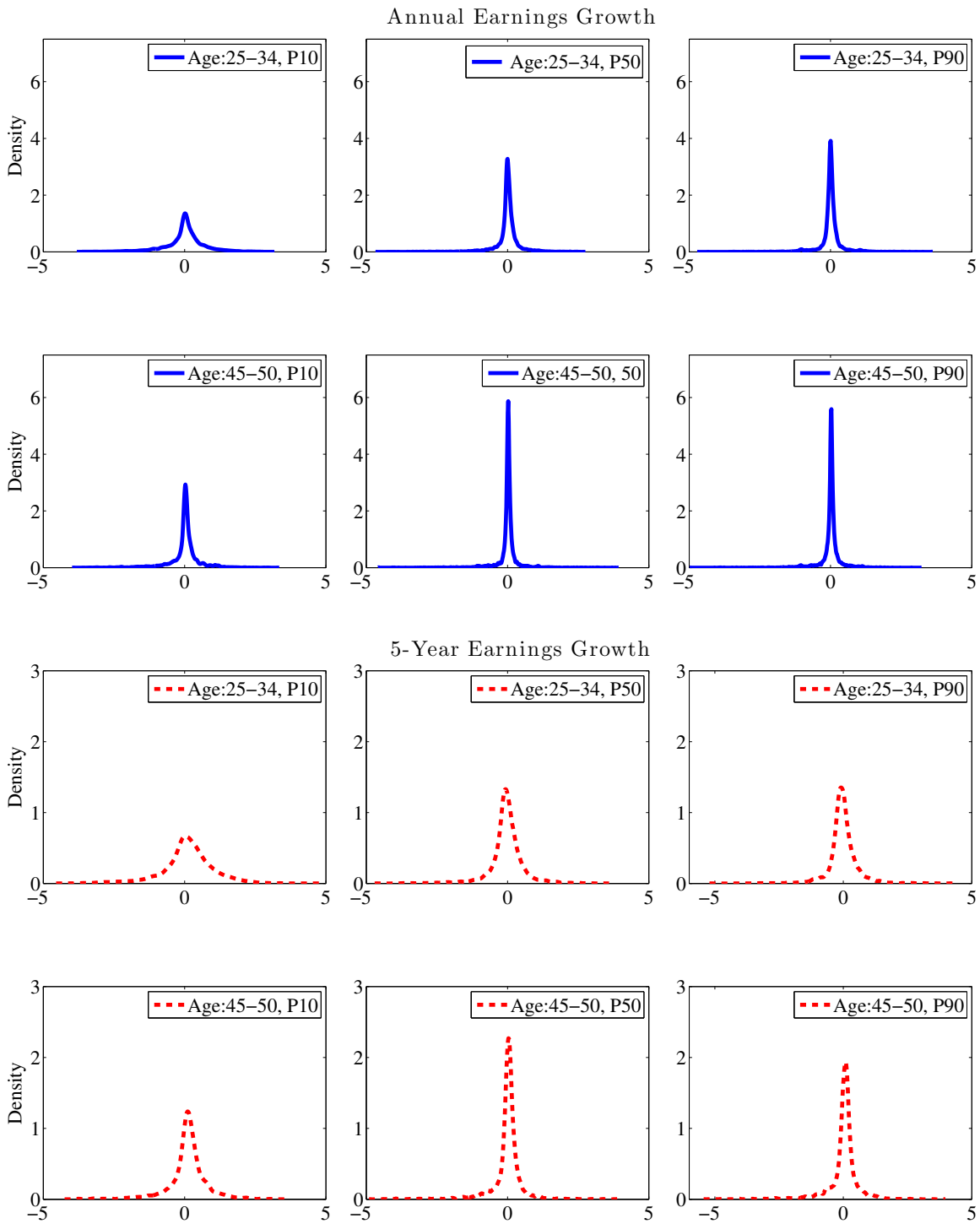
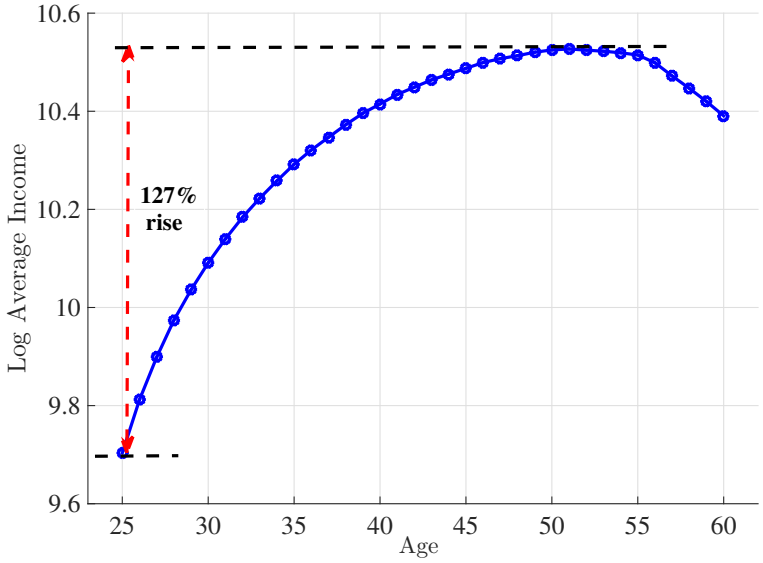


FIGURE 4 – Life-Cycle Profile of Average Log Earnings



calibrated life-cycle model will imply that the median individual in the simulated sample experiences (on average) a rise of this magnitude from ages 25 to 55. One question we now address is whether this implication is consistent with what we see in the data. In other words, if we rank male workers in the U.S. data by their lifetime earnings, does the median worker experience an earnings growth of approximately 127%?

This question can be answered directly with our data. First, we need to compute lifetime earnings for each individual. For this purpose, we select a subsample of individuals that have at least 32 years of data between the ages of 25 and 60. We further restrict our sample to individuals who (i) have earnings above $Y_{\min,t}$ for at least 15 years and (ii) are not self-employed for more than 8 years. We rank individuals based on their lifetime earnings, computed by summing their earnings from ages 25 through 60. Earnings observations lower than $Y_{\min,t}$ are set to this threshold. For individuals in a given lifetime earnings (hereafter, LE) percentile group, denoted LEj, $j = 1, 2, ..., 99, 100$, we compute growth in average earnings between any two ages h_1 and h_2 as $\log(\overline{Y}_{h_2,j}) - \log(\overline{Y}_{h_2,j})$, where $\overline{Y}_{h,j} \equiv \mathbb{E}(Y_h^i | i \in \text{LE}j)$ and Y_h^i for a given individual may be zero.

Figure 5 plots the results for $h_1 = 25$ and $h_2 = 55$. Here, there are several takeaways. First, individuals in the median lifetime earnings group experience a growth rate of 38%, about one-third of what was predicted by the profile in Figure 4. Moreover, we have to look all the way above LE90 to find an average growth rate of 127%. However,

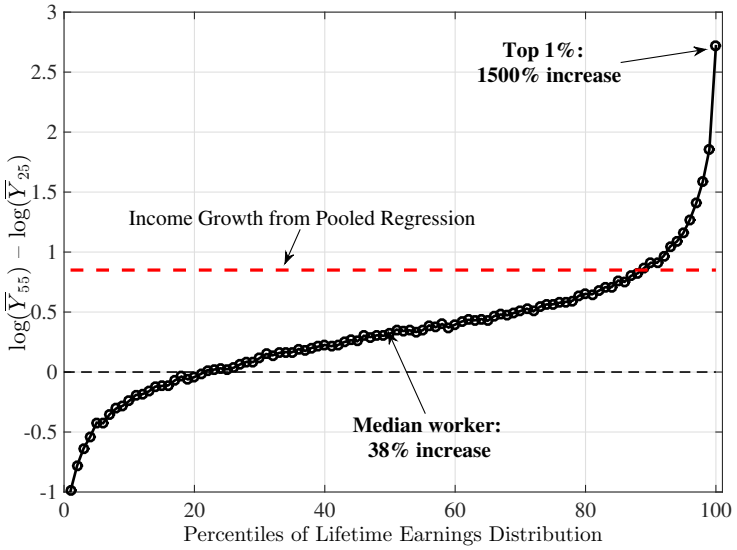


FIGURE 5 – Life-Cycle Earnings Growth Rates, by Lifetime Earnings Group

earnings growth is very high for high-income individuals, with those in the 95th percentile experiencing a growth rate of 230% and those in the 99th percentile experiencing a growth rate of 1450%. Although some of this variation could be expected due to the fact that individuals with high earnings growth are more likely to have high lifetime earnings, these magnitudes are too large to be accounted for by that channel, as we show below.

Earnings Growth by Decades. How is earnings growth over the life-cycle distributed over different decades of the life cycle? Figure 6 answers this question by plotting, separately, earnings growth from ages 25 to 35, 35 to 45, and 45 to 55. Across the board, the bulk of earnings growth happens during the first decade. In fact, for the median LE group, average earnings growth from ages 35 to 55 is zero (notice that the solid blue line and grey line with circles overlap at LE50). Second, with the exception of those in the top 10% of the LE distribution, all groups experience negative growth from ages 45 to 55. So, the peak year of earnings is strongly related to the lifetime earnings percentile. After age 45, the only groups that are experiencing growth on average are those who are in the top 2% of the LE distribution.

How do the results change if we consider a slightly later starting age? Figure 7 plots earnings growth starting at age 30 (solid blue line) and 35 (dashed red line). As can be anticipated from the previous discussion, from ages 35 to 55, average growth is zero for the median LE group and is very low for all workers below L70. Top earners still do

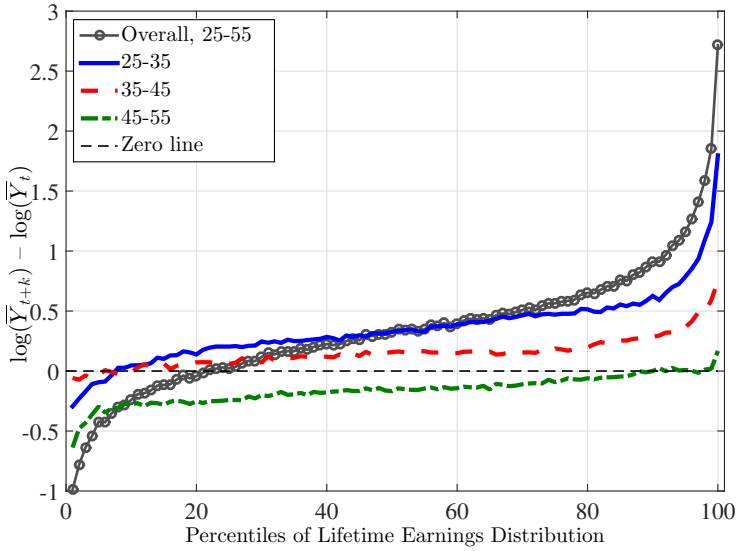


FIGURE 6 – Earnings Growth Rates in Each Decade of the Life Cycle

very well though, experiencing a rise of 200 log points (or 640%) from ages 30 to 55 and a rise of 90 log points (or 146%) from ages 35 to 55. Those at the bottom of the LE distribution display the opposite pattern: average earnings drops by 70 log points (or 50%) from ages 35 to 55.

3.3 Second Moment: Variance

How does the dispersion of earnings shocks vary over the life cycle and by earnings groups? To answer this question, Figure 8 plots the standard deviation of one-year and five-year earnings growth by age and recent earnings (hereafter, RE) groups (as defined above, Section 2.2). The following patterns hold true for both short- and long-run growth rates. First, for every age group, there is a pronounced U-shaped pattern by RE levels, implying that earnings changes are less dispersed for individuals with higher RE up to about the 90th percentile (along the x-axis). This pattern reverts itself inside the top 10% as dispersion increases rapidly with recent earnings. Second, over the life cycle, the dispersion of shocks declines monotonically up to about age 50 (with the exception of very top earners) and then rises slightly for middle- to high-earning individuals from ages 50 to 55.

The life-cycle pattern is quite different for top earners who experience a monotonic *increase* in dispersion of shocks over the life cycle. In particular, for one-year changes,

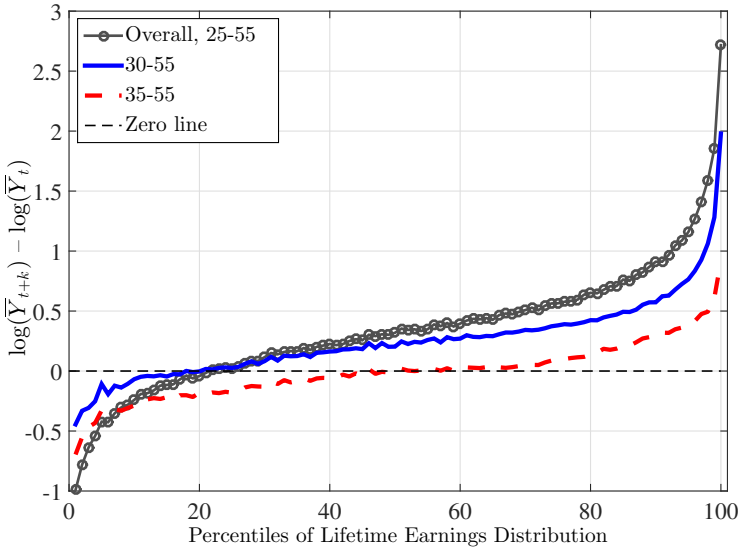
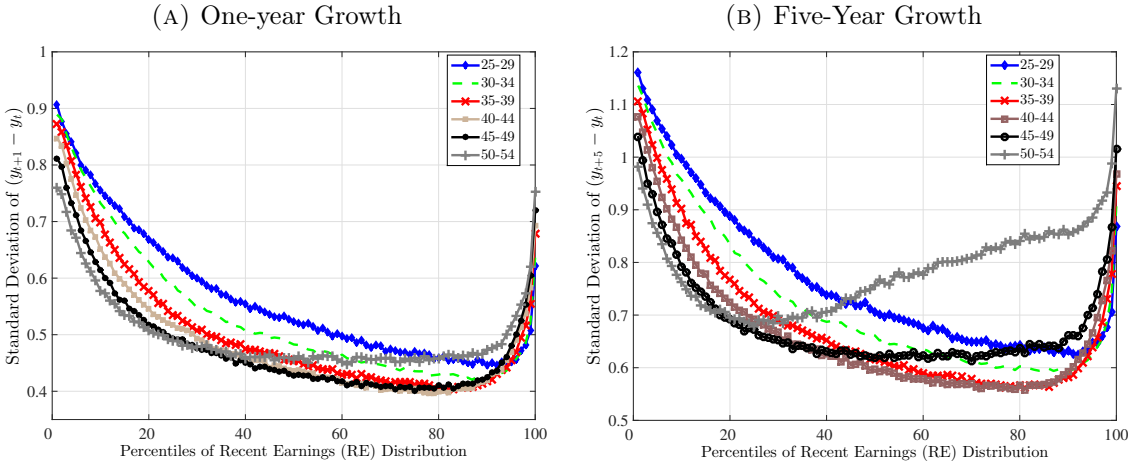


FIGURE 7 – Life-Cycle Earnings Growth Rates, Different Starting Ages

FIGURE 8 – Standard Deviation of Earnings Growth



individuals at the 95th percentile of the RE distribution experience a slight increase from 0.45 in the youngest age group up to 0.51 in the oldest group (50–54). Those in the top 1% experience a larger increase from 0.62 in the first age group up to 0.75 in the oldest. Therefore, we conclude that the lower 95 percentiles and the top 5 percentiles display patterns with age and recent earnings that are the opposite of each other. The same theme will emerge again in our analysis of higher-order moments.

Standard Deviation of (Log) Earnings Levels. Although the main focus of this section is on earnings *growth*, the life-cycle evolution of the dispersion of earnings *levels* has been at the center of the incomplete markets literature since the seminal paper of Deaton and Paxson (1994). For completeness, and comparability with earlier work, we have estimated the within-cohort variance of log earnings over the life cycle and report it in Figure A.2 in Appendix A.1.

3.4 Third Moment: Skewness (or Asymmetry)

The lognormality assumption implies that the skewness of earnings shocks is zero. Figure 9 plots the skewness, measured here as the third central moment, of one-year (top) and five-year (bottom) earnings growth. The first point to observe is that every graph in both panels of Figure 9 lie below the zero line, indicating that earnings changes are negatively skewed at every stage of the life cycle and for all earnings groups. The second point, however, is that 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 down, and the less room he has left to move up. And this is true for both short-run and long-run earnings changes. Curiously, and as was the case with the standard deviation, the life-cycle pattern in skewness becomes much weaker at the very top of the earnings distribution.

Another measure of asymmetry is provided by Kelly's measure of skewness, which is defined as

$$\mathcal{S}_{\mathcal{K}} = \frac{(P90 - P50) - (P50 - P10)}{P90 - P10}, \tag{1}$$

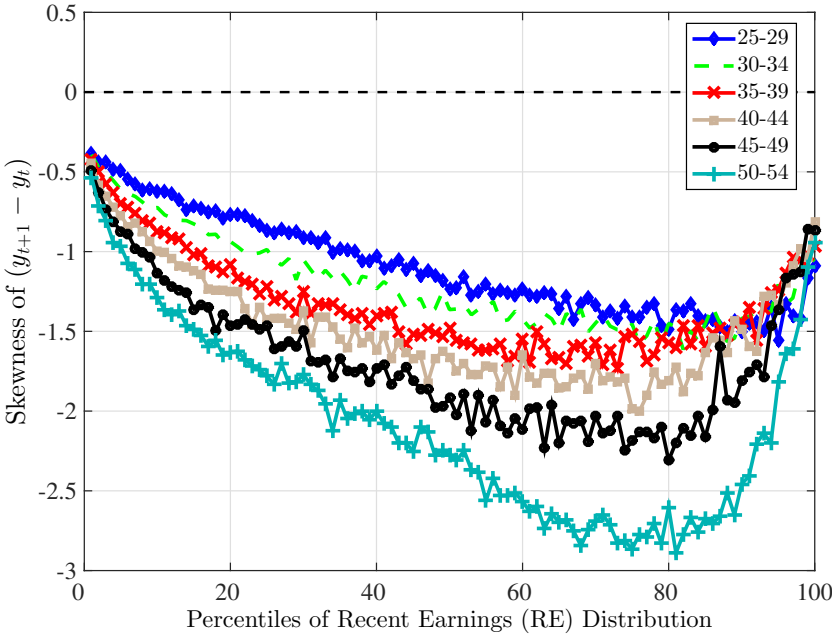
where Pxy refers to percentile xy of the distribution under study. So, basically $\mathcal{S}_{\mathcal{K}}$ measures the relative fractions of the overall dispersion (P90–P10) accounted for by the upper and lower tails. An appealing feature of Kelly's skewness relative to the third central moment is that a particular value is easy to interpret. To see this, rearrange (1) to get

$$\frac{P90 - P50}{P90 - P10} = \frac{\mathcal{S}_{\mathcal{K}} + 1}{2}.$$

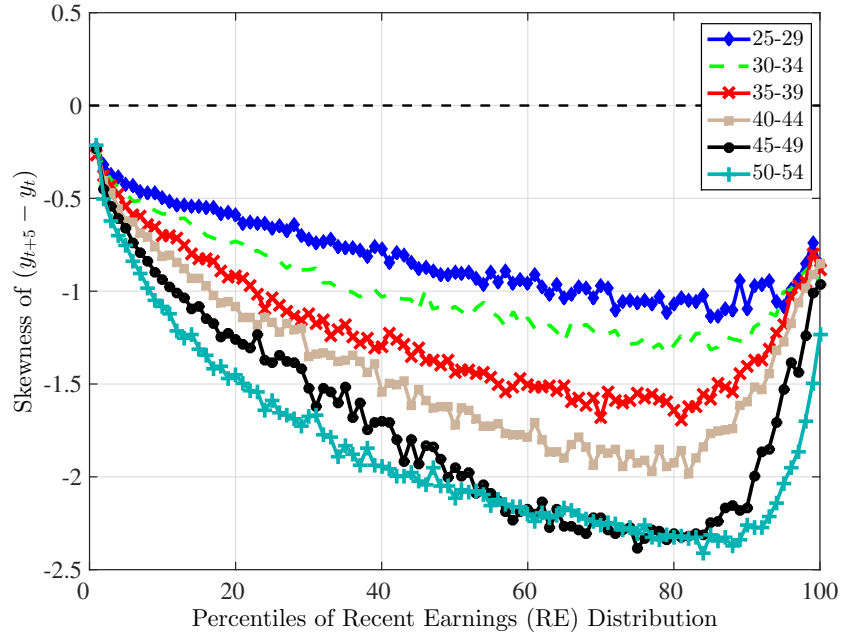
Thus, a negative value of $\mathcal{S}_{\mathcal{K}}$ implies that the lower tail (P50-P10) is longer than the upper tail (P90-P50), indicating negative skewness. Another property of Kelly's measure

FIGURE 9 – Skewness (Third Moment) of Earnings Growth

(A) One-Year Change



(B) Five-Year Change



is that it is less sensitive to extremes (above the 90th or below the 10th percentile of the shock distribution). Instead, it captures the weight distribution in the middling section of the shock distribution, whereas the third moment also puts a large weight on the relative lengths of each tail. (We shall examine the tails in more detail in the next subsection.)

In the top panel of Figure 10, we plot Kelly's skewness, which is also negative throughout and becomes more negative with age, especially below RE60. However, it does not always get more negative with higher RE. This difference from the third central moment (Fig. 9a) indicates that as RE increases, it is mostly the extreme shocks (measured in the third moment) that become more negatively skewed, rather than the more middling shocks.

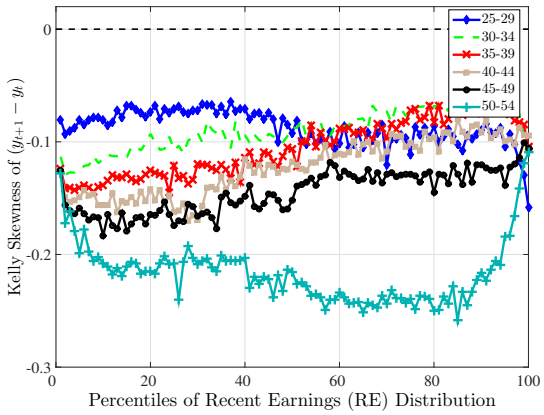
Figure 10b plots Kelly's skewness for five-year changes, which reveals essentially the same pattern as with the third moment in Figure 9b: each measure shows a strong increase in left-skewness with both age and earnings (except for the very-high earners). Furthermore, the magnitude of skewness is substantial. For example, the Kelly's skewness for five-year earnings change of –0.35 for individuals aged 45–49 and in the 80th percentile of the RE distribution implies that the P90-P50 accounts for 32% of P90-P10, whereas P50-P10 accounts for the remaining 68%. This is clearly different from a lognormal distribution, which is symmetric—both tails contribute 50% of the total.

While the preceding decomposition is useful, it does not answer a key question: is the increasingly more negative skewness over the life cycle primarily due to a compression of the upper tail (fewer opportunities to move up) or due to an expansion in the lower tail (increasing risk of falling a lot)? For the answer, we need to look at the *levels* of the P90-50 and P50-P10 separately over the life cycle. The left panel of Figure 11 plots P90-50 for different age groups minus the P90-50 for 25- to 29-year-olds, which serves as a normalization. The right panel plots the same for P50-P10. One way to understand the link between these two graphs and skewness is that keeping P50-P10 fixed over the life cycle, if P90-P50 (left panel) declines with age, this causes Kelly's skewness to become more negative. Similarly, keeping P90-P50 fixed, a rise in P50-P10 (right panel) has the same effect.

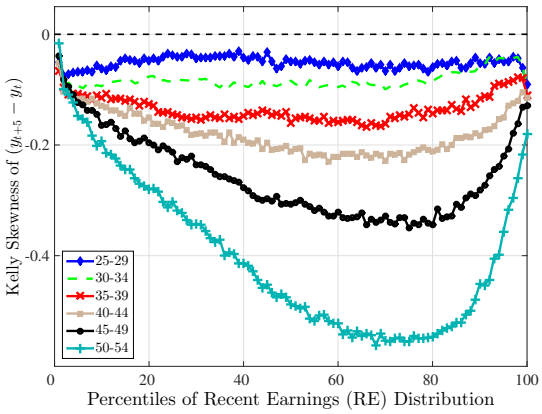
Turning to the data, up until age 45, both P90-P50 and P50-P10 decline with age (across most of the RE distribution). This leads to the declining dispersion that we have

FIGURE 10 – Kelly's Skewness of Earnings Growth

(A) One-year Change



(B) Five-Year Change



seen above. The shrinking P50-P10 would also lead to a *rising* skewness if it were not for the faster compression of P90-P50 during the same time. Therefore, from ages 25 to 45, the increasing negativity of skewness is *entirely* due to the fact that the upper end of the shock distribution compresses more rapidly than the compression of the lower end. After age 45, P50-P10 starts expanding rapidly (larger earnings drops becoming more likely), whereas P90-P50 stops compressing any further (stabilized upside). Thus, during this phase of the life cycle, the increasing negativity in Kelly's skewness is due to increasing downward risks and not the disappearance of upward moves. The only exception to this pattern is, again, the top earners (RE95 and above) for whom P90-50 actually never compresses over the life cycle, whereas the P50-P10 gradually rises as they get older. Therefore, as they climb the wage ladder, these individuals do not face a tightening ceiling, but do suffer from an increasing risk of falling a lot.

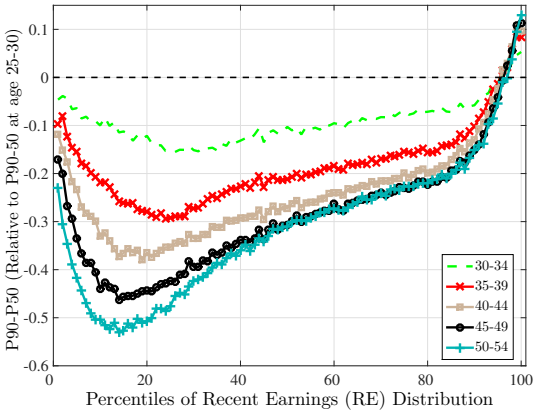
3.5 Fourth Moment: Kurtosis (Peakedness and Tailedness)

It is useful to begin by discussing what kurtosis measures. A useful interpretation has been suggested by Moors (1986, 1988), who described kurtosis as measuring how dispersed a probability distribution is away from $\mu \pm \sigma$.[17] This is consistent with how a

[17]This can easily be seen by introducing a standardized variable $Z = (x - \mu)/\sigma$ and noting that kurtosis is $\kappa = \mathbb{E}(Z^4) = \text{var}(Z^2) + \mathbb{E}(Z^2)^2 = \text{var}(Z^2) + 1$. So κ 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.

FIGURE 11 – Kelly's Skewness Decomposed: Change in P90-P50 and P50-P10 Relative to Age 25–30; Five-Year Earnings Growth

(A) P90-P50



(B) P50-P10

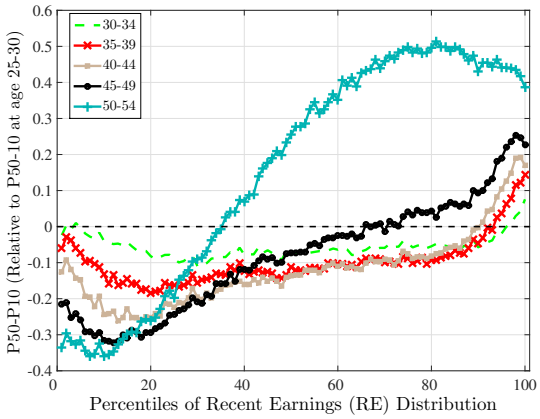


TABLE II – Fraction of Individuals with Selected Ranges of Log Earnings Change

	$\text{Prob}(\lvert y^i_{t+1} - y^i_t \rvert < x)$				$\text{Prob}((y^i_{t+1} - y^i_t) \in \textit{Range})$		
x :	Data*	$\mathcal{N}(0, 0.48^2)$	Ratio	*Range:*†	Data	$\mathcal{N}(0, 0.48^2)$	Ratio
0.05	0.35	0.08	4.38	Center	0.653	0.250	2.61
0.10	0.54	0.16	3.38	Shoulders	0.311	0.737	0.42
0.20	0.71	0.32	2.23	Tails	0.036	0.013	2.77
0.50	0.86	0.70	1.22				
1.00	0.94	0.96	0.98				

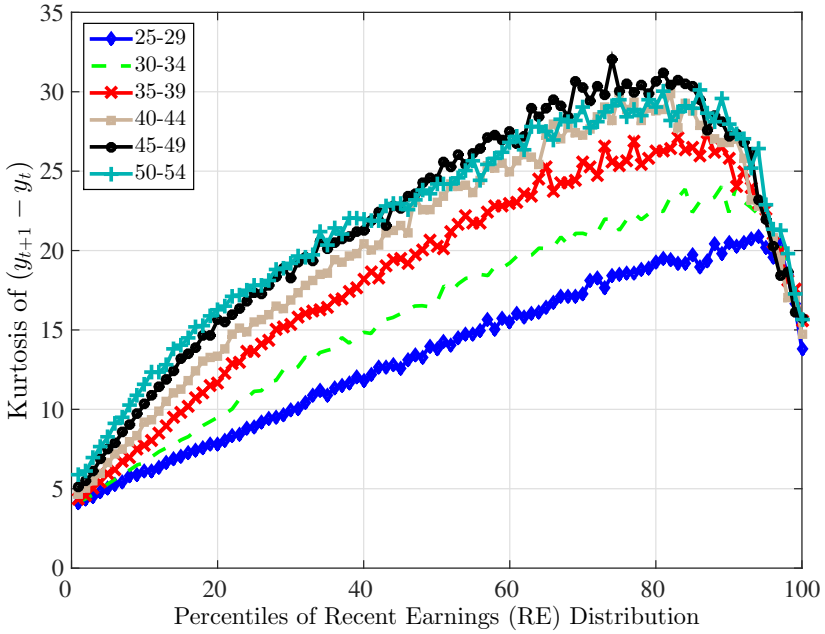
Notes: *The empirical distribution used in this calculation is for 1995-96, the same as in Figure 1. †The intervals are defined as follows: "Center" refers to the area inside the first intersection between the two densities in Figure 1 [-0.122 to 0.187]. "Tails" refer to the areas outside the intersection point at the tails: $(-\infty, -1.226] \cup [1.237, \infty)$. "Shoulders" refer to the remaining areas of the densities.

distribution with excess kurtosis often looks like: a sharp/pointy center, long tails, and little mass near $\mu \pm \sigma$. A corollary to this description is that for a distribution with high kurtosis, the usual way we think about standard deviation—as representing the size of the typical shock—is not very useful. This is because very few realizations will be of a magnitude close to the standard deviation; instead, most will be either close to the median or in the tails.

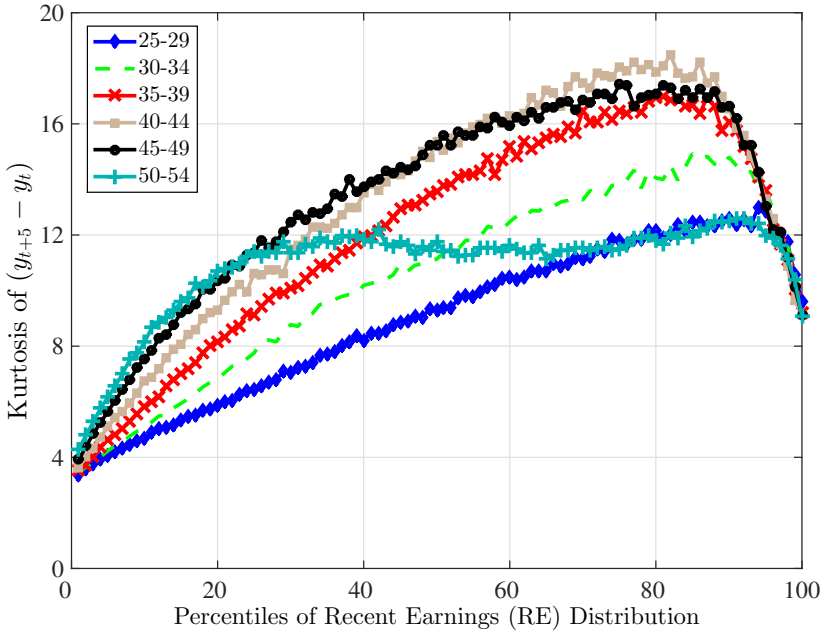
With this definition in hand, let us now examine the earnings growth data. Figure

FIGURE 12 – Kurtosis of Earnings Changes

(A) Annual Change



(B) Five-Year Change

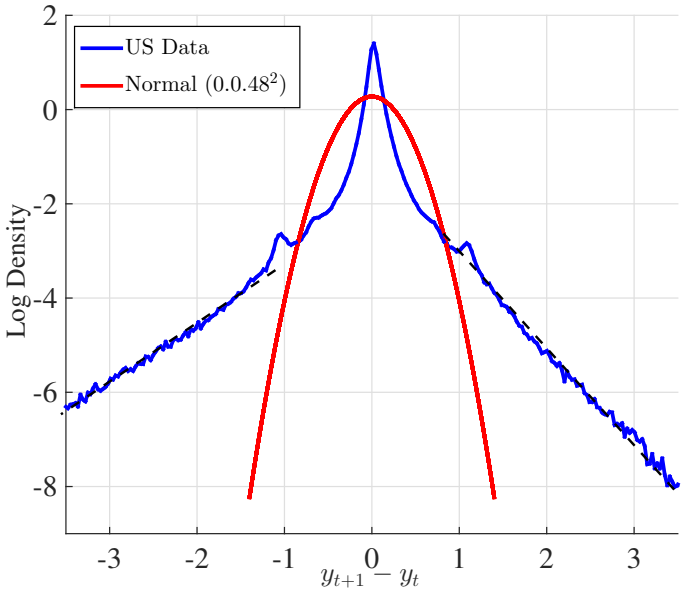


12a plots the kurtosis of annual earnings changes. First, notice that kurtosis increases monotonically with recent earnings up to the 80th to 90th percentiles for all age groups. That is, high-earnings individuals experience *even smaller* earnings changes of either sign, with few experiencing very large changes. Second, kurtosis increases over the life cycle, for all RE levels, except perhaps the top 5 percent. Furthermore, the peak levels of kurtosis range from a low of 20 for the youngest group, all the way up to 30 for the middle-age group (40–54).

To provide a more familiar interpretation of these kurtosis values, it is useful to calculate measures of concentration. The first three columns of Table II report the fraction of individuals experiencing a log earnings change (of either sign) of less than a threshold $x = 0.05, 0.10, 0.20, 0.50$, and 1.00, under alternative assumptions about the data-generating process. For the entire sample, the standard deviation of $y^i_{t+1} - y^i_t$ is 0.48. Assuming that the data-generating process is a Gaussian density with this standard deviation, only 8 percent of individuals would experience an annual earnings change of less than 5%. The true fraction in the data is 35 percent. Similarly, the Gaussian density predicts a fraction of 16 percent when the threshold is 0.10, whereas the true fraction is 54 percent. As an alternative calculation, we calculate the areas under the densities in three different ranges determined by the intersections of the two densities in the left panel of Figure 1. The center is the area inside the first set of intersections, and the Gaussian density has 25% of its mass in this area compared with 65% in the data. The shoulders are the second set of areas, marked again by the intersections, and the Gaussian density has almost three-quarters of its mass in this area, compared with only 31% in the data. Finally, the Gaussian density has only 1.3% of its mass in the tails compared with almost three times that amount in the data.

We now take a closer look at the tails of the earnings growth distribution compared with a normal density. Figure 13 plots the log density of the one-year change in the data versus the Gaussian density. This is essentially the same as the left panel of Figure 1 but with the y-axis now in logs. The lognormal density is an exact quadratic, whereas the data display a more complex pattern. Two points are worth noting. One, the data distribution has much thicker *and* longer tails compared with a normal, and the tails decline almost linearly, implying a Pareto distribution at both ends, with significant

FIGURE 13 – Tails of the Distributions: U.S. Data vs. Normal Density



weight at extremes.[18] Two, the tails are asymmetric, with the left tail declining much more slowly than the right, contributing the negative third central moment documented above. In fact, fitting linear regression lines to each tail yields a tail index of 2 for the right tail and 1.2 for the left tail—the latter showing especially high thickness.[19]

Overall, these findings show that earnings changes in the U.S. data exhibit important deviations from lognormality and raise serious concerns about the focus in the current literature on the covariances (second moments) alone. In particular, targeting the covariances alone can vastly overestimate the typical earnings shock received by the average worker and miss the substantial but infrequent jumps experienced by few.

Economic Models behind Skewness and Kurtosis. While the lognormal framework is often adopted for technical and empirical tractability, negative skewness and excess kurtosis are naturally generated by standard structural models of job search over the life cycle. For example, job ladder models in which workers do on-the-job search and move from job to job as they receive better offers and fall off the job ladder due to

[18] Double Pareto distribution is when both tails are Pareto with different tail indices.

[19] Notice that while the Pareto tail in the earnings distribution is well known, here the Pareto tails emerges in earnings *changes,* on which much less empirical evidence exists.

unemployment not only will generate negative skewness but also will imply that skewness becomes more negative with age. This is because as the worker climbs the job ladder, the probability of receiving a wage offer much higher than the current wage will be declining. At the same time, as the worker is higher up in the wage ladder, falling down to a flat unemployment surface (or disability) implies that there is more room to fall. Furthermore, as the attractiveness of job offers declines with the current wage, more job offers will be rejected, and therefore job to job transitions will also decline with age, implying that most wage changes will be small (within-job) changes. This increases the concentration of earnings changes near zero, which in turn raises the kurtosis of changes. Having said that, the magnitudes of variation over the life cycle and by earnings levels that we documented in these moments are so large that it is an open question whether existing models of job search can be consistent with these magnitudes, and, if not, what kinds of modifications should be undertaken to make them consistent.[20]

3.6 Robustness and Extensions

The results documented in the previous sections show important deviations from log-normality as well as clear patterns with age and past earnings. This raises the question whether some of these findings are due to simple statistical artifacts—say, due perhaps to extreme shocks experienced by few individuals—and whether the age and earnings 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.

Decomposing Moments: Job-Stayers vs. Job-Switchers. Going back to Topel and Ward (1992), economists have found important differences between the earnings changes that occur during an employment relationship and those that occur across jobs (see, more recently, Low *et al.* (2010), Altonji *et al.* (2013), and Bagger *et al.* (2014)). Therefore, it is of interest to ask how the empirical patterns we documented so far relate to within- and between-job earnings changes. Our data set contains a unique employer

[20] A high kurtosis in earnings changes could partly be due to heterogeneity across individuals in shock variances, as documented by Chamberlain and Hirano (1999) and Browning *et al.* (2010). We have explored this possibility by estimating earnings shock variances for each individual. While we do find significant heterogeneity in variances, consistent with these papers, the remaining kurtosis is still substantial. We do model and estimate individual-specific variances in Section 5.

identification number (EIN) for each job that a worker holds in a given year, which allows us to conduct such an analysis. In the interest of space, we report those statistics in Appendix A.2.

Disentangling The Effects of Age and Recent Earnings. Recall that in the analysis so far, we have grouped workers first by age, and then within each age group, we have ranked and divided them into recent earnings percentiles. The implication is that RE percentiles in these figures are age-group-dependent. So when we fix an RE percentile and examine how statistics vary with age, we are simultaneously looking at changes with earnings, since earnings vary with age. The advantage of this approach was that it ensured each RE group contained (more or less) the same number of observations, whereas grouping workers based on the RE distribution in the overall sample will result in too many younger workers appearing in lower RE percentiles and vice versa for middle-age workers. As a consequence, the latter approach might result in different sampling variability in different parts of each graph, which we wished to avoid. Despite this concern, it is useful to check what the results would look like under this alternative approach.

In Figures 14 and 15, we plot 3-D graphs of skewness and kurtosis, where we first group workers based on the RE distribution in the overall sample, then within each RE group, we classify workers by age. Inspecting these graphs carefully and comparing with their counterparts above shows that the main substantive conclusions described above are robust.

Averaging Earnings Over Neighboring Years. Recall that the statistics above were constructed by taking differences between two years, t and $t+k$. One concern is that some of the interesting findings could be due to the timing of earnings. For example, suppose that an individual's income has been shifted from the last few months of year $t+k$ into the beginning of $t+k+1$. If true, this would represent an earnings fluctuation that is easy to smooth, but could appear as a big negative shock between t and $t+k$. A similar comment applies to period t.

To address this issue, we have constructed the same set of statistics for the second to fourth moments by using two-year average earnings. For the short-run and long-run

FIGURE 14 – Skewness of One-year Earnings Change, Jointly Conditioning on Age and Recent Earnings

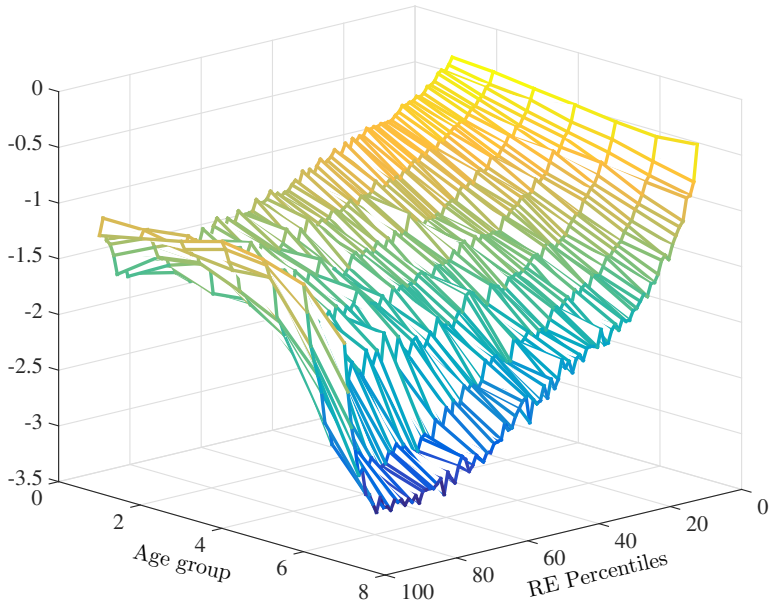
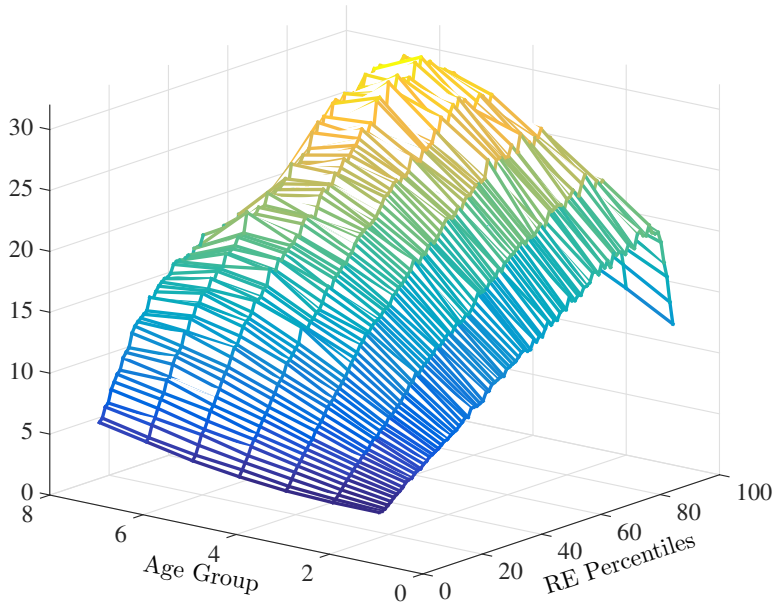


FIGURE 15 – Kurtosis of One-year Earnings Change, Joint Conditioning on Age and Recent Earnings



variations, we use, respectively

$$\tilde{\Delta} y_t^i = \log(Y_{t+3}^i + Y_{t+2}^i) - \log(Y_t^i + Y_{t+1}^i)$$

and

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

Notice that the first measure is not a one-year difference but is more like a two-year difference, whereas the second one is closer to a five-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* Earnings. Even though we condition on recent earnings over the past five 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 earnings 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 the short-run and long-run variations:

$$\Delta_{\text{short}} y^i = \log(Y_{t+1}^i) - \log(\overline{Y}^i_{t-1}),$$

and

$$\Delta_{\text{long}} y^i = \log(Y_{t+5}^i) - \log(\overline{Y}^i_{t-1}).$$

These are longer differences than before, 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, 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 third and fourth 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 RE percentiles, and then calculate the third and fourth moments.

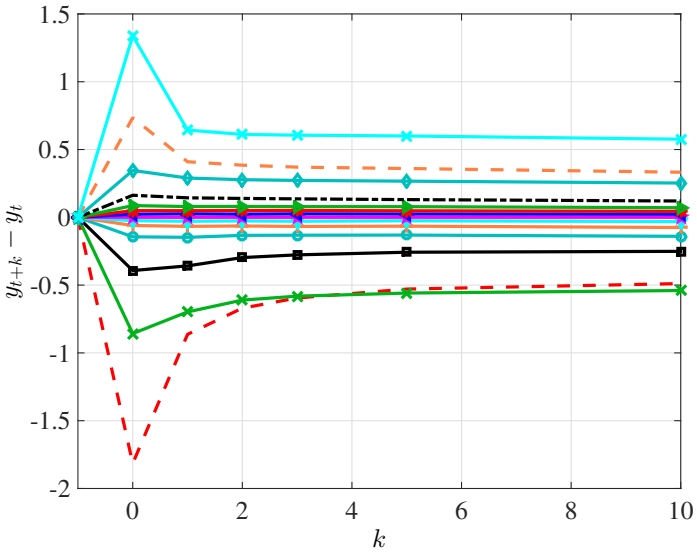
Second, recall that we consider an individual to be non-employed and drop from the computation of a statistic if his earnings is below $Y_{\min,t}$ in periods t or $t+k$. Because this threshold does not vary with earnings, high-income individuals can experience a larger fall in earnings *and* still remain in the sample, whereas low-income individuals would exit the sample with the same fall. This asymmetry might give the appearance of a more negative skewness for higher-income individuals. To explore this issue, we drop all observations in which an individual's earnings in period t or $t+k$ is below 5% of *his own recent earnings*. Because the threshold is now indexed to the level of his earnings, the mechanical relationship is no longer a concern.

In Appendix A, we report the figures analogous to those above under these three robustness checks. Although the figures are quantitatively different, the difference is almost always small, and therefore the substantive conclusions of this analysis remain intact.

4 Dynamics of Earnings

A key dimension of life-cycle earnings risk is the persistence of earnings 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 by the rate of decline of autocovariances with the lag order. 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 approach might be appropriate in survey data with small sample sizes, it imposes restrictions on the data that might be too strong, 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 characterize persistence without making parametric assumptions.

FIGURE 16 – Impulse Responses, Prime-Age Workers with Median RE

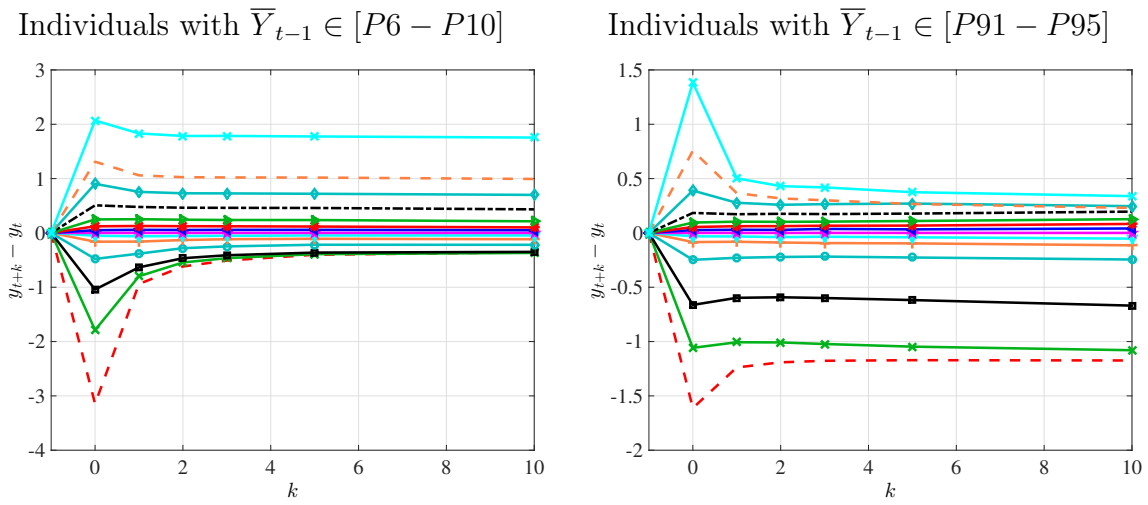


Final Sample for Impulse Response Analysis

The final sample for this analysis is slightly different from the one used in the previous section. In particular, our *final sample* includes all observations that are in the base sample in $t-1$ and in at least two more years between $t-5$ and $t-2$, and furthermore satisfies the age (25–60) and no-self-employment condition in years t through $t+3$, and in $t+5$ and $t+10$. To reduce the number of graphs to a manageable level, we aggregate individuals across demographic groups. First, we combine the first two age groups (ages 25 to 34) into "young workers," and the last four groups (ages 35 to 55) into "prime-age males."

To this end, we rank and group individuals based on their average earnings from $t-5$ to $t-1$, then within each such group, we rank and group again by the size of the earnings change between $t-1$ and t. Hence, all individuals within a given group obtained by crossing the two conditions have the same average earnings up to time $t-1$ and experience the same earnings "shock" from $t-1$ to t. For each such group of individuals, we then compute their average earnings change from t to $t+k$, for all values of $k = 1, 2, 3, 5, 10$. Specifically, we construct 21 groups based on their RE percentiles: 1–5, 6–10, 11–15, . . . , 86–90, 91–95, 96–99, 100. Then, we construct 20 equally-divided groups based on the percentiles of the shock, $y_t^i - y_{t-1}^i$: percentiles 1–5, 6–10, 11–15, . . . , 91–95,

FIGURE 17 – Impulse Responses, Prime-Age Workers with Low or High RE



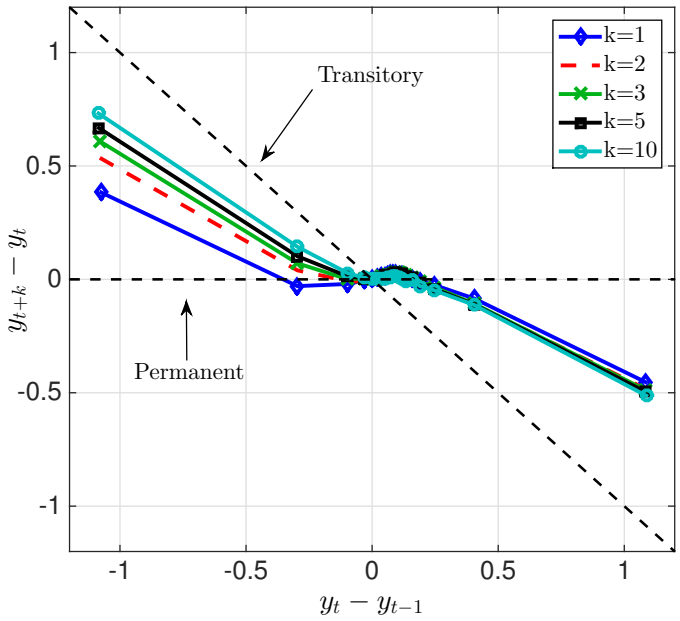
96–100. Therefore, for every year t, we have 2 age groups, 21 RE groups and 20 groups for shock size (for a total of 840 groups). As before, we construct these groups separately for each t and assign workers based on these averages. Then, for workers in each group, we compute the average of log k-year earnings growth, $\mathbb{E}\left[y^i_{t+k} - y^i_t | \overline{Y}^i_{t-1}, y^i_t - y^i_{t-1}\right]$, for $k = 1, 2, 3, 5, 10$.

4.1 Impulse Response Functions

For prime-age males with median RE level (as of $t-1$), Figure 16 plots 14 impulse response functions (one for each "shock" size, $y^i_t - y^i_{t-1}$). As seen here, the most positive 5% of shocks at time t are about 140 log points and the most negative are about –180 log points. Notice that the mean reversion pattern varies 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 the 10-year horizon, a nonnegligible 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 functions but now for workers that are in the 10th percentile (top panel) and 90th percentile (bottom) of RE distribution. Notice that, for low-income individuals, negative shocks mean-revert much more quickly, whereas positive shocks are more persistent than before. The opposite is true for high-income individuals.

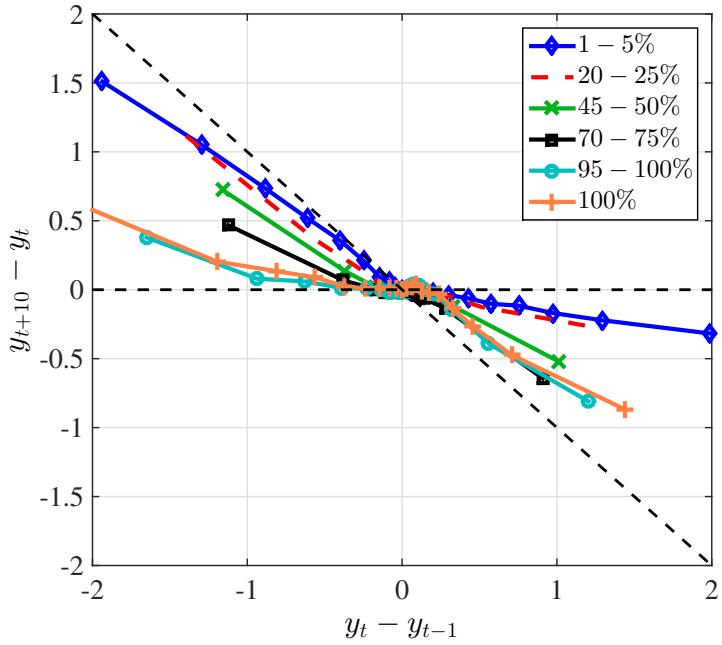
FIGURE 18 – Impulse Responses (Rotated View), Prime-Age Workers with Median RE



To illustrate these patterns more clearly, Figure 18 plots the shock, $y_t^i - y_{t-1}^i$, on the x-axis and the fraction of each shock that has mean-reverted, $y_{t+k}^i - y_t^i$, on the y-axis for the median RE group. Thus, this figure contains the same information as in Figure 16 but is reported differently. Several remarks are in order. First, negative earnings changes tend to be less persistent than positive earnings changes. For example, a worker whose earnings rise by 100 log points between $t-1$ and t loses about 50% of this gain in the following 10 years. It is also interesting to note that almost all of this mean reversion happens after one year, implying that whatever mean reversion there is happens very quickly. Turning to earnings losses: a worker whose earnings fall 100 log points recovers one-third of that loss by $t+1$ and recovers more than two-thirds of the total within 10 years. Moreover, unlike with positive shocks, the recovery (hence mean reversion) is more gradual in response to negative shocks.

Second, the degree of mean reversion varies with the magnitude of earnings shocks. This is evident in Figure 18, where small shocks (i.e., those less than 10 log points in absolute value) look very persistent, whereas there is substantial mean reversion following larger earnings changes. A univariate autoregressive process with a single persistence parameter will fail to capture this behavior. In the next section, we will allow for multiple

FIGURE 19 – Asymmetric Mean Reversion: Butterfly Pattern



AR(1) processes to accommodate the variation in persistence by shock size.

Next, we investigate how these mean reversion patterns vary across workers that differ in their recent earnings, which is a proxy for where these individuals rank overall relative to their peers. To make the comparison clear, we focus on a fixed horizon, 10 years, and plot the total mean reversion between t and $t + 10$ for the 8 RE groups in Figure 19. Starting from the lowest RE group (those individuals in the bottom 5% of the recent earnings distribution), notice that negative shocks are transitory, with an almost 80% mean reversion rate at the 10-year horizon. But positive shocks are quite persistent, with only about 20% mean reversion at the same horizon. As we move up the RE distribution, the positive and negative branches of each graph start rotating in *opposite* directions, so that for the highest RE group, we have the opposite pattern: only 20 to 25% of negative shocks mean-revert at the 10-year horizon, whereas almost 75% of positive shocks mean-revert at the same horizon. We refer to this shape as the "butterfly pattern."

5 Estimation

With the few exceptions noted above, the bulk of the literature relies on the (often implicit) assumption that earnings shocks can be approximated reasonably well with a lognormal distribution. This assumption, combined with linear time series models (such as an ARMA(p,q) process) 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 earnings either in levels or in first difference form, with very few exceptions.[21]

There are two problems with this approach. First, the broad range of evidence presented in the previous sections implies that this approach is likely to miss important aspects of the data and produce a picture of earnings 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 it is difficult to judge the relative importance—from an economic standpoint—of the covariances that a given model matches well and those that it does not match well. This is an especially important shortcoming given that virtually every econometric process used to calibrate economic models is statistically rejected by the data.

With these considerations in mind, we propose and implement a different approach that relies on matching the kinds of moments presented above. We believe that economists can much more easily judge whether or not each one of these sets of moments is relevant for the economic questions 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 an acceptable shortcoming. They can similarly judge the success of a given stochastic process in matching some moments and not others.

More concretely, in the first stage, we use each set of moments presented above as diagnostic tools to determine the basic components that should be included in the stochastic process that we will then fit to the data. Clearly, this stage requires extensive pre-testing and exploratory work. For example, to generate the lifecycle earnings growth

[21] Exceptions include Browning *et al.* (2010), Altonji *et al.* (2013), and Guvenen and Smith (2014). Clearly, GMM or minimum distance estimation that is used to match such moments does not require the assumption of lognormality for consistency. But abstracting away from moments higher than covariances is a reflection of the belief that higher-order moments do not contain independent information, which does require the lognormality assumption.

patterns documented in Section 3.2, we considered three basic ingredients: (i) an AR(1) + i.i.d process, (ii) a quadratic HIP process with no shocks, and (iii) a mixture of two AR(1) process where each component receives a non-zero innovation with a certain probability. We picked the first two ingredients because of the widespread attention they garnered in the previous literature, and the third one based on our conjecture that it might perform well. We found that the first ingredient, on its own, could not generate the rich patterns of earnings growth revealed by the data, whereas the HIP process performed fairly well, and the normal mixture process performed the best. Therefore, we concluded that a stochastic process for earnings should include either (ii) or (iii) as one of its components. Although earnings growth data on its own could not determine which one of these pieces is more important, together with the other moments, we will be able to obtain sharper identification of the parameters of these two components.

We conducted similar diagnostic analyses on the other cross-sectional moments (standard deviation, skewness, and kurtosis) as well as on impulse response moments. The variation in the second to fourth moments over the life cycle and with earnings levels seemed impossible to match without introducing explicit dependence of shocks on these two characteristics. Allowing the mixing probabilities to depend on age and earnings delivered much better results and we make this specification part of our benchmark. We then estimate the parameters of such a process to target moments whose economic significance is more immediate, including the distribution of lifetime earnings, the kurtosis and skewness of earnings changes, as well as how these moments vary with age and rising earnings. These moments are used as targets using a method of simulated moments (or more generally, an indirect inference) estimator.

5.1 A Flexible Stochastic Process

The extensive preliminary analyses we conducted led us to our benchmark specification with the following features: (i) a heterogeneous earnings profiles component of (up to) quadratic form; (ii) a mixture of two AR(1) processes, denoted by z and x, where each component receives a new innovation in a given year with probability $p_j \in [0, 1]$ for $j = z, x$; and (iii) a random walk component that is realized with probability p_v; and

(iv) an i.i.d. transitory shock. Here is the full specification:

$$\begin{aligned}
\widetilde{y}_t^i &= \left(\alpha^i + \beta^i t + \gamma^i t^2\right) + z_t^i + x_t^i + v_t^i + \varepsilon_t^i \\
z_t^i &= \rho_z z_{t-1}^i + \eta_{zt}^i \\
x_t^i &= \rho_x x_{t-1}^i + \eta_{xt}^i \\
v_t^i &= v_{t-1}^i + \eta_{vt}^i,
\end{aligned}$$

where for $j = z, x, v$:

$$\eta_{jt}^{*i} \sim \mathcal{N}(\mu_j, \sigma_j^i) \quad \text{and} \quad \eta_{jt}^i = \eta_{jt}^{*i} \times \mathbb{I}\{s_{i,t} \in I_{p_j}\} \tag{2}$$

$$\log\sigma_j^i \sim \mathcal{N}(\overline{\sigma}_j - \frac{\sigma_{jj}^2}{2}, \sigma_{jj}^2), j = z, x, \qquad \sigma_v^i \equiv \sigma_v. \tag{3}$$

Here, (α, β, γ) follows a multivariate normal distribution with zero mean and covariance matrix Σ. The realizations of the three innovations $(\eta_{jt}, j = z, x, v)$ are mutually exclusive—only one of the three shocks is received per period. This is implemented by first drawing a uniform random variable, $s_{i,t}$, for a given individual at age t and dividing the unit interval into three pieces: $I_{p_z} = [0, p_z]$, $I_{p_x} = (p_z, p_z + p_x]$, and $I_{p_v} = (p_z + p_x, 1]$, where $p_z + p_x \leq 1$ and $p_v = 1 - p_z - p_x$. Depending on which interval $s_{i,t}$ falls in, that innovation is set to its random draw, η_{jt}^{*i}, and the others are set equal to zero.

The specification in (3) implies that the innovation standard deviation for each AR(1) process has an individual-specific component that is lognormal, with mean $\overline{\sigma}_j$ and standard deviation proportional to σ_{jj}. To economize on parameters, we assume that the permanent innovation is identically distributed across individuals. Regarding the initial conditions of the persistent processes, z_0^i and x_0^i, we assume that they are drawn from a normal distribution with zero mean and standard deviation $\sigma_{j,0}$, $j = z, x$.[22] Finally, to avoid indeterminacy in the estimation, without loss of generality we impose $\rho_x > \rho_z$. We will consider a variety of specifications similar to this benchmark.

The infrequent nature of earnings shocks to either of the two AR(1) components as well as the heterogeneity in innovation variances help us generate excess kurtosis. In addition, they capture the non-linear pattern in life-cycle earnings growth documented

[22] The initial variance of the permanent component cannot be identified from the fixed effect variance and hence is normalized to zero.

in Section 3.2. Allowing the means of innovations to differ helps with generating non-zero (in our case negative) skewness. Furthermore, allowing infrequent shocks with a 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 it.

To capture the variation in these moments over the life cycle and across recent earnings levels, we allow the mixing probabilities to depend on age and on the idiosyncratic component of earnings:

$$p_j(y^i_{t-1}) = a_j + b_j \times \left(y^i_{t-1}\right) + c_j \times t + d_j \times \left(y^i_{t-1}\right) \times t. \tag{4}$$

Recall that in Section 3, we defined y^i_t as log earnings net of age effects. Here, y^i_t can alternatively be defined as $z^i_t + x^i_t + v^i_t + \varepsilon^i_t$. Therefore, in the estimation analyses below, when a HIP component is allowed, y^i_t is net of the entire HIP component, whereas in specifications without HIP it is simply net of the common age component. Finally, since p_j's are probabilities, they must remain between [0,1]. Thus, equation (4) is truncated outside of this range.[23]

5.2 Estimation Method

We estimate the parameters of the described stochastic process using a method of simulated moments (MSM) estimator. The empirical moments are chosen to capture the three sets of empirical facts documented in the previous two sections. Specifically, we target: (i) moments from the life-cycle profile of average earnings, summarized in Figures 5 to 7, (ii) the standard deviation, skewness, and kurtosis of one-year and five-year

[23] We have also considered an alternative specification where the innovation variances are functions of earnings and age:

$$\sigma_{jt}(\widetilde{y}^i_t) = a_j + b_j \times (\widetilde{y}^i_t) + c_j \times t + d_j \times (\widetilde{y}^i_t) \times t.$$

In this specification, as earnings changes with age, this could capture both an age structure and a cross-sectional structure. We have experimented with this specification extensively but found it not to fit the moments nearly as well as the main specification presented in the text.

earnings growth presented in Sections 3.3 to 3.5, and (iii) moments describing the impulse response functions presented in Section 4. In addition, and as noted above, the within-cohort variance of log earnings levels is a key dimension of the data extensively studied in previous research. For both completeness and consistency with earlier work, we include these variances (summarized in Figure A.2) as a fourth set of moments.

Accounting for Zeros. Recall that in order to construct the cross-sectional moments (set (ii)), we have dropped individuals who had very low earnings—below $\overline{Y}_{\min}$—in year t or $t+k$ so as to allow taking logarithms in a sensible manner.[24] Although this made sense for documenting empirical facts that are easy to understand, for the estimation exercise, we would like to also capture the patterns of these "zeros" (or very low earnings observations), given that they clearly contain valuable information. This is feasible thanks to the flexibility of the MSM approach, and in Appendix B we describe how we modify the cross-sectional moments to allow these zeros in the estimation.

Aggregating Moments. If we were to match all data points in all these moments—i.e, for every RE percentile and every age group—it would yield more than 10,000 moments. Although this is doable, not much is likely to be gained from such a level of detail, but it would make the diagnostics, that is, judging the performance of the estimation, quite difficult. To avoid this, we aggregate 100 RE percentiles into 10 to 15 groups and the 6 age groups into two (ages 25–34 and 35–55). Full details of how this aggregation is performed and which moments are targeted are included in Appendix B. After the aggregation procedure, we are left with 120 moments capturing average earnings over the life cycle (targeted ages are 25, 30, 35, 40, 45, 50, 55, and 60); 156 moments for cross-sectional statistics (standard deviation, skewness and kurtosis of one-year and five-year earnings growth); and 1120 moments coming from the impulse response functions. Adding the 36 moments on the variance of log earnings, in sum, we target a total of $156 + 120 + 1120 + 36 = 1432$ moments.

Let m_n for $n = 1, ..., N = 1432$ denote a generic empirical moment, let $d_n(\theta, X)$ be the corresponding model moment that is simulated for a given set of earnings process parameters, θ and a given vector of random variables, X. We simulate the entire earnings

[24] We were able to include those below the threshold in sets (i) and (iii) because for those moments it made sense to first take averages, including zeros, and then take the logarithms of those averages.

histories of 200,000 individuals who enter labor market at age 25 and work until age 60. When computing the model moments, we apply precisely the same sample selection criteria and employ the same methodology 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, define

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

where $\gamma_n > 0$ 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 not unusual (e.g., the median of log earnings changes). To account for this, we choose γ_n to be equal to the 10th percentile of the distribution of the absolute value of the moments in a given set. The MSM estimator is

$$\hat{\theta} = \arg\min_{\theta} \boldsymbol{F}(\theta)' W \boldsymbol{F}(\theta), \tag{5}$$

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

$$\boldsymbol{F}(\theta) = \left[F_1(\theta), ..., F_N(\theta)\right].$$

The weighting matrix, W, is chosen such that the life-cycle average earnings growth moments and impulse response moments are assigned a relative weight of 0.25 each, the cross-sectional moments of earnings growth receive a relative weight of 0.35, and the variance of log earnings is given a relative weight of 0.15. The objective function is highly jagged in certain directions and highly nonlinear in general, owing to the fact that we target higher-order moments and percentiles of the distribution (the latter creating discontinuities). Therefore, we employ a global optimization routine, described in further detail in Guvenen (2013), to perform the minimization in (5). Further details can be found in Appendix B.

5.3 Results

Table III reports the parameter estimates from the benchmark model (column 1) as well as from a model with a linear HIP component (column 2) and one with no HIP component (column 3). First, the systematic component of earnings growth displays significant heterogeneity: In column 1, the dispersion in the fixed (level) effect is $\sigma_\alpha = 0.35$, the dispersion in growth rates is $\sigma_\beta = 0.0068$ (or approximately 0.68 percent), and the dispersion in the quadratic component is also sizable. The fixed effect is negatively correlated with the growth rate. Estimates of the simpler linear HIP specification reveal a cross-sectional standard deviation of growth rates of about 1 percent per year. Overall, the estimates of growth rate heterogeneity is quantitatively large. Compared to earlier estimates, such as Haider (2001) and Guvenen (2009), who used survey data and matched the covariance matrix of earnings residuals (with no allowance for mixtures), the estimates of growth rate heterogeneity here are somewhat smaller. For example, Guvenen (2009) estimated a standard deviation of 1.8% per year; and Haider (2001)'s estimates were comparable. This suggests that the more general earnings process considered here attributes some of the properties captured

Turning to the stochastic component driving the dynamics, for a young worker that just entered the labor market with zero idiosyncratic earnings (i.e., setting $h = 0$ and $\tilde{y} = 0$), the first AR(1) process receives a shock with 4 percent annual probability, whereas the second AR(1) and the random walk components receive innovations with 34 and 62 percent annual probabilities (column 1). Notably, the mixing probabilities display substantial systematic heterogeneity. In particular, the first AR(1) process receives innovations with a probability that tends to increase with age and decrease with idiosyncratic income with a positive interaction between the two. The probability of receiving an innovation to the second AR(1) process, however, is decreasing in age but increasing in idiosyncratic income.

Figure 20 below shows the dependence of mixing probabilities on age and idiosyncratic income for the quadratic HIP specification. Shocks to the AR(1) components tend to be fairly transitory with first order autocorrelations of $\rho_1 = 0.23$ and $\rho_2 = 0.38$. However, these innovations (when they are realized) tend to be very large, with average standard deviations of 0.76 and 0.36 for the first and second AR(1) components, respectively. Consistent with earlier work, there is sizable heterogeneity in the standard deviation of

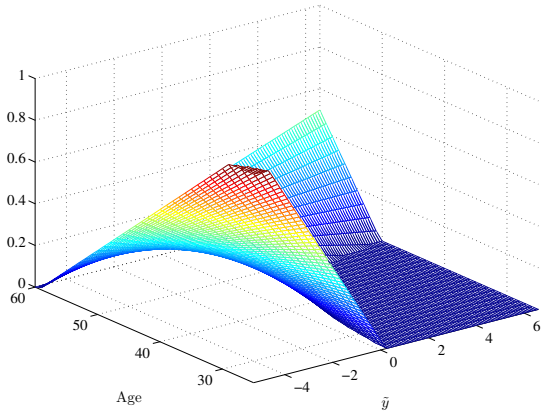
TABLE III – Estimation Results

	Benchmark: Quadratic HIP	Specification 2: Linear HIP	Specification 3: No HIP
σ_α	0.3504	0.3619	0.3274
$\sigma_\beta \times 10$	0.103	0.1034	
$\sigma_\gamma \times 100$	0.0494		
$\sigma_{\alpha\beta}$	–0.3093	–0.6111	
$\sigma_{\alpha\gamma}$	–0.3490		
$\sigma_{\beta\gamma}$	–0.1199		
σ_{z0}	0.0955	0.1355	0.1578
σ_{x0}	0.1063	0.0793	0.1165
p_z	$0.041 + 0.047h - 0.211y + 0.071hy$	$0.052 + 0.051h - 0.206y + 0.076hy$	$0.123 + 0.097h - 0.355y + 0.083hy$
p_x	$0.345 - 0.083h + 0.096y - 0.548hy$	$0.285 - 0.038h + 0.045y - 0.500hy$	$0.331 - 0.205h + 0.087y - 0.003hy$
ρ_z	0.2316	0.2830	0.2901
ρ_x	0.3799	0.3342	0.5671
μ_z	–0.5584	–0.4116	–0.2889
μ_x	–0.0175	–0.0426	0.0604
μ_v	0.0562	0.0605	0.0921
$\bar{\sigma}_z$	0.7606	0.7744	0.7302
σ_{zz}	0.1841	0.1825	0.1864
$\bar{\sigma}_x$	0.3641	0.3757	0.2192
σ_{xx}	0.2190	0.1573	0.1381
σ_v	0.0985	0.1011	0.0877
σ_ε	0.0414	0.0386	0.0413
Obj.	0.4152	0.4203	0.4225

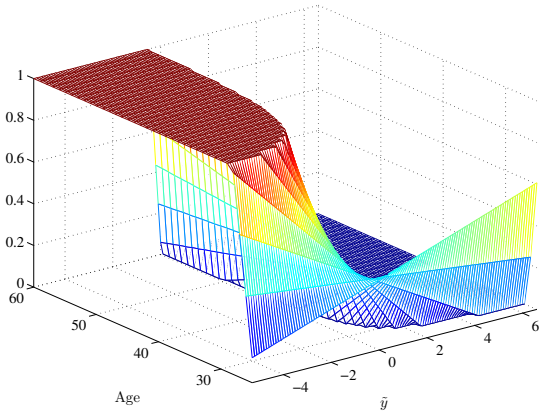
Note:

FIGURE 20 – 3-D Plot of Mixing Probabilities

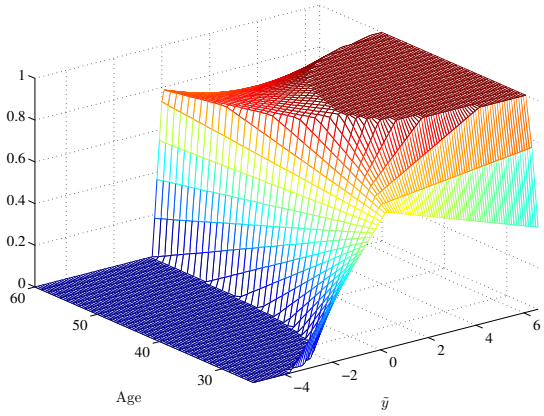
(A) Mixing Probability for the First AR(1) Component (z_t)



(B) Mixing Probability for the Second AR(1) Component (x_t)



(C) Mixing Probability for the Random Walk Component (ν_t)



the AR(1) innovations. The histograms for each specification are shown in Figure **??**. A robust finding across all specifications is that the standard deviation of innovations to the first AR(1) component is much more dispersed than those for the second component. The size of the purely i.i.d. shock is estimated very small, with a standard deviation of $\sigma_\epsilon = 0.04$ in the quadratic HIP specification.

Next, we turn to the fit of the model to the targeted moments. To that end, Figure 21 displays the fit of the three specifications on cross-sectional moments of prime-age workers. Note that all three specifications generate levels and patterns of skewness and kurtosis consistent with the data. However, the model underestimates the standard

deviation of one-year earnings changes in the data. Finally, we examine the fit of the estimated models on income growth moments.

Figure 22 shows the implications of the three specifications for income growth moments documented in section 3.2. We find that all three specifications do a good job in generating the nonlinear profile of life cycle income growth. Lastly, Figure 23 shows the fit of the model on the variance of log earnings over the life cycle. While all specifications imply a continuous rise in earnings over the entire life cycle, they overestimate its magnitude. The linear HIP specification is the closest to the empirical log earnings profile.

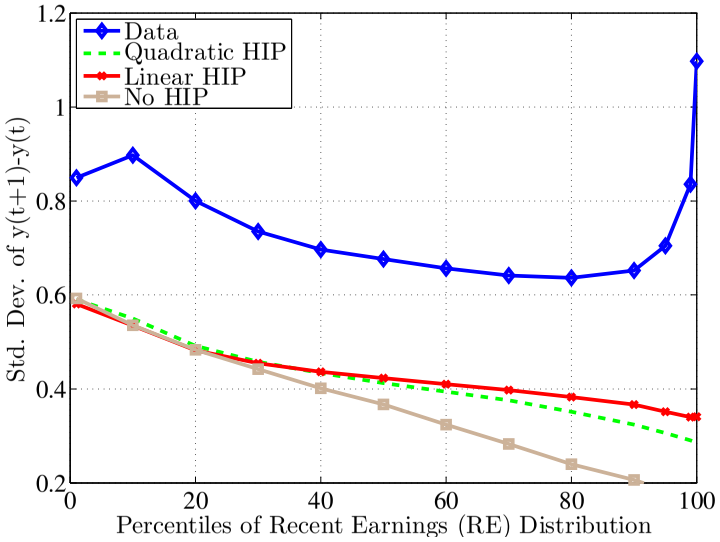
6 Conclusions

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 lognormal. 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 earnings level of individuals. We also estimate impulse response functions of earnings shocks and find significant asymmetries: positive shocks to high-earnings individuals are quite transitory, whereas negative shocks are very persistent; the opposite is true for low-earnings 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 earnings process broadly consistent with these salient features of the data.

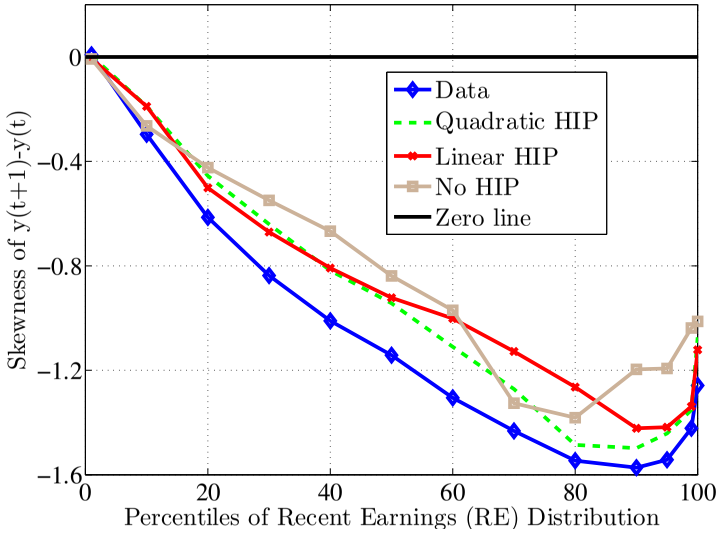
A broader message of this paper is a call to reconsider the way researchers approach the study of earnings dynamics. The covariance matrix approach that dominates current work is too opaque and a bit mysterious: it is difficult to judge what it means to match or miss certain covariances from an economics perspective. With the current trend toward

FIGURE 21 – Fit of Estimated Model to Key Cross-Sectional Data Moments

(A) Standard Deviation of One-Year Earnings Growth



(B) Skewness of One-Year Earnings Growth



(C) Kurtosis of One-Year Earnings Growth

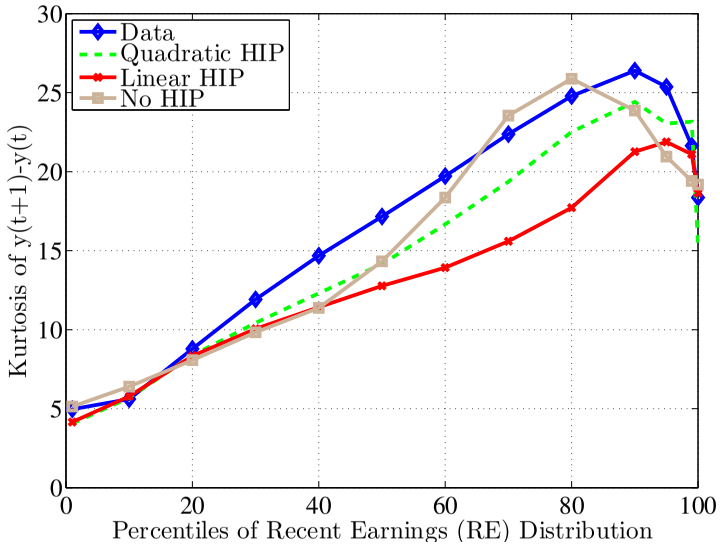
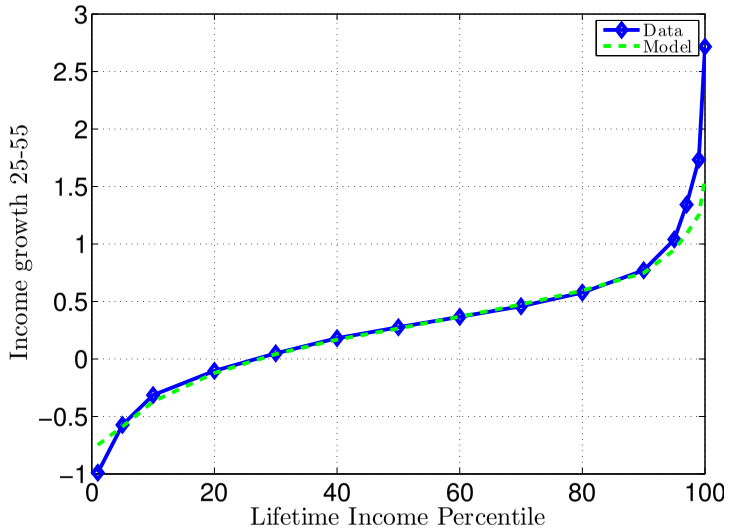
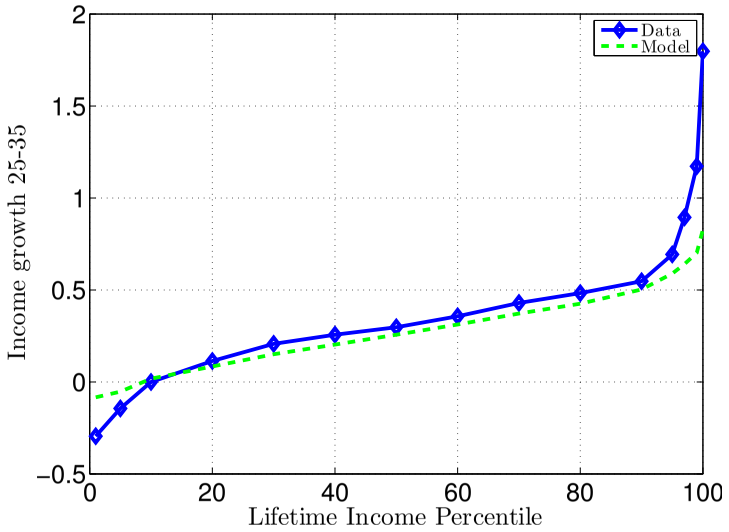


FIGURE 22 – Fit of Estimated Model to Key Cross-Sectional Data Moments

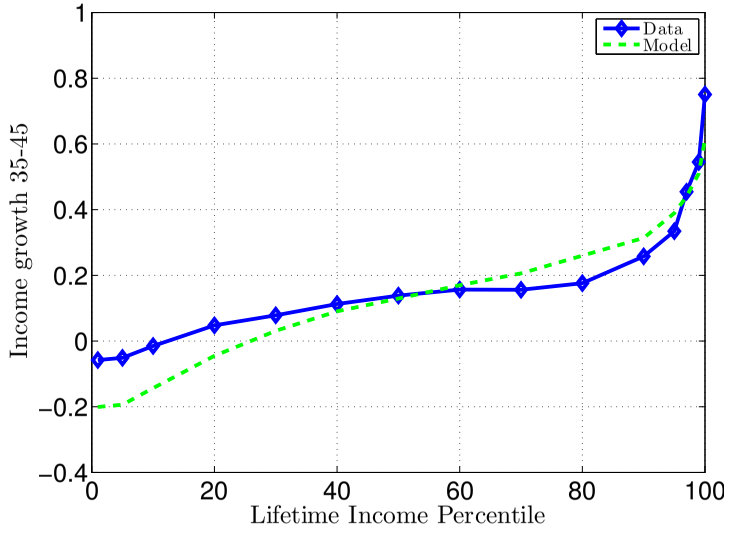
(A) Life Cycle Income Growth: 25–55



(B) Life Cycle Income Growth: 25–35



(C) Life Cycle Income Growth: 35–45



(D) Life Cycle Income Growth: 45–55

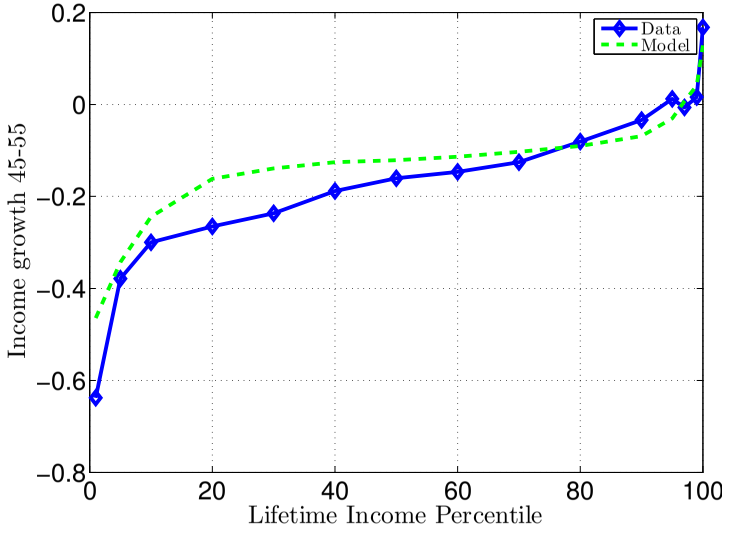
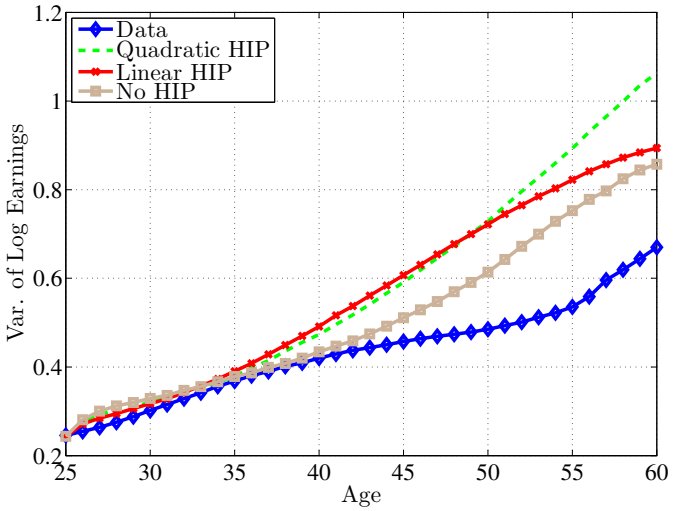


FIGURE 23 – Fit of Estimated Model to Variance of Log Earnings



the increasing availability of very large panel data set (which should only get better going forward), our priority in choosing methods should shift from efficiency concerns to transparency. The approach adopted here is an example of the latter, and we believe that economists can better judge what each moment implies for economic questions. We hope the empirical findings about the nature of earnings dynamics documented in this paper feed back into economic research and policy analyses.

References

ABOWD, J. M. and CARD, D. (1989). On the covariance structure of earnings and hours changes. *Econometrica*, **57** (2), 411–45.

ALTONJI, J., SMITH, A. A. and VIDANGOS, I. (2013). Modeling earnings dynamics. *Econometrica*, **81** (4), 1395–1454.

ARELLANO, M., BLUNDELL, R. and BONHOMME, S. (2014). *Household Earnings and Consumption: A Nonlinear Framework*. mimeo, University College London.

ATTANASIO, O. P., BANKS, J., MEGHIR, C. and WEBER, G. (1999). Humps and bumps in lifetime consumption. *Journal of Business & Economic Statistics*, **17** (1), 22–35.

Bachmann, R. and Bayer, C. (2014). Investment dispersion and the business cycle. *American Economic Review*, **104** (4), 1392–1416.

Bagger, J., Fontaine, F., Postel-Vinay, F. and Robin, J.-M. (2014). Tenure, experience, human capital and wages: A tractable equilibrium search model of wage dynamics. *American Economic Review*, **104** (6), 1551–1596.

Baker, M. and Solon, G. (2003). Earnings dynamics and inequality among Canadian men, 1976–1992: Evidence from longitudinal income tax records. *Journal of Labor Economics*, **21** (3), 289–321.

Berger, D. and Vavra, J. (2011). *Dynamics of the U.S. Price Distribution*. Working paper, Yale University.

Bloom, N., Floetotto, M., Jaimovich, N., Saporta-Eksten, I. and Terry, S. J. (2011). *Really Uncertain Business Cycles*. Working paper, Stanford University.

Blundell, R., Graber, M. and Mogstad, M. (2014). Labor income dynamics and the insurance from taxes, transfers, and the family. *Journal of Public Economics*.

Bonhomme, S. and Robin, J.-M. (2009). Assessing the equalizing force of mobility using short panels: France, 1990 a 2000. *Review of Economic Studies*, **76** (1), 63–92.

Borovicka, J., Hansen, L. P. and Sheinkman, J. A. (2014). *Shock Elasticities and Impulse Responses*. Working paper, University of Chicago.

Browning, M., Ejrnaes, M. and Alvarez, J. (2010). Modelling income processes with lots of heterogeneity. *Review of Economic Studies*, **77**, 1353–1381.

Castañeda, A., Díaz-Giménez, J. and Ríos-Rull, J.-V. (2003). Accounting for the U.S. earnings and wealth inequality. *The Journal of Political Economy*, **111** (4), 818–857.

Chamberlain, G. and Hirano, K. (1999). Predictive distributions based on longitudinal earnings data. *Annales d'Économie et de Statistique*, **55-56**, 211–242.

Christiano, L. J., Eichenbaum, M. and Evans, C. L. (2005). Nominal rigidities and the dynamic effects of a shock to monetary policy. *Journal of Political Economy*, **113** (1), 1–45.

Constantinides, G. M. and Ghosh, A. (2014). *Asset Pricing with Countercyclical Household Consumption Risk*. Working paper, University of Chicago.

Deaton, A. (1991). Saving and liquidity constraints. *Econometrica*, **59** (5), 1221–48.

— and Paxson, C. (1994). Intertemporal choice and inequality. *Journal of Political Economy*, **102** (3), 437–67.

Geweke, J. and Keane, M. (2000). An empirical analysis of earnings dynamics among men in the PSID: 1968-1989. *Journal of Econometrics*, **96**, 293–356.

— and — (2007). Smoothly mixing regressions. *Journal of Econometrics*, **138**, 252–290.

Golosov, M., Troshkin, M. and Tsyvinski, A. (2014). *Redistribution and Social Insurance*. Working paper, Princeton University.

Gourinchas, P.-O. and Parker, J. A. (2002). Consumption over the life cycle. *Econometrica*, **70** (1), 47–89.

Guvenen, F. (2009). An empirical investigation of labor income processes. *Review of Economic Dynamics*, **12** (1), 58–79.

— (2013). Quantitative economics with heterogeneity. Book Manuscript, University of Minnesota.

—, Ozkan, S. and Song, J. (2014). The nature of countercyclical income risk. *Journal of Political Economy*, **122** (3), 621–660.

— and Smith, A. A. (2014). Inferring labor income risk and partial insurance from economic choices. *Econometrica*.

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

Hause, J. C. (1980). The fine structure of earnings and the on-the-job training hypothesis. *Econometrica*, **48** (4), 1013–1029.

Karahan, F. and Ozkan, S. (2013). On the persistence of income shocks over the life cycle: Evidence, theory, and implications. *Review of Economic Dynamics*, **16** (3), 452–476.

Lillard, L. A. and Weiss, Y. (1979). Components of variation in panel earnings data: American scientists 1960-70. *Econometrica*, **47** (2), 437–454.

— and Willis, R. J. (1978). Dynamic aspetcs of earnings mobility. *Econometrica*, **46** (5), 985–1012.

Low, H., Meghir, C. and Pistaferri, L. (2010). Wage risk and employment risk over the life cycle. *American Economic Review*, **100** (4), 1432–1467.

MaCurdy, T. E. (1982). The use of time series processes to model the error structure of earnings in a longitudinal data analysis. *Journal of Econometrics*, **18** (1), 83–114.

McKay, A. (2014). *Time-Varying Idiosyncratic Risk and Aggregate Consumption Dynamics*. Working paper, Boston University.

Meghir, C. and Pistaferri, L. (2004). Income variance dynamics and heterogeneity. *Econometrica*, **72** (1), 1–32.

Midrigan, V. (2011). Menu costs, multiproduct firms, and aggregate fluctuations. *Econometrica*, **79** (4), 1139–1180.

Moors, J. J. A. (1986). The meaning of kurtosis: Darlington reexamined. *The American Statistician*, **40**, 283–284.

— (1988). A quantile alternative for kurtosis. *Journal of the American Statistical Association. Series D (The Statistician)*, **37** (1), 25–32.

Murphy, K. M. and Welch, F. (1990). Empirical age-earnings profiles. *Journal of Labor Economics*, **8** (2), 202–29.

Nakamura, E. and Steinsson, J. (2013). Price rigidity: Microeconomic evidence and macroeconomic implications. *Annual Review of Economics*, **5** (133-163).

Olsen, A. and Hudson, R. (2009). Social security administration's master earnings file: Background information. *Social Security Bulletin*, **69** (3), 29–46.

Panis, C., Euller, R., Grant, C., Bradley, M., Peterson, C. E., Hirscher, R. and Stinberg, P. (2000). *SSA Program Data User's Manual*. Baltimore, MD: Social Security Administration.

Storesletten, K., Telmer, C. I. and Yaron, A. (2004). Cyclical dynamics in idiosyncratic labor market risk. *Journal of Political Economy*, **112** (3), 695–717.

Topel, R. H. and Ward, M. P. (1992). Job mobility and the careers of young men. *The Quarterly Journal of Economics*, **107** (2), 439–79.

Zhang, H., Conn, A. R. and Scheinberg, K. (2010). A derivative-free algorithm for least-squares minimization. *SIAM Journal on Optimization*, **20** (6), 3555–3576.

Supplemental Online Appendix

NOT FOR PUBLICATION

A Appendix: Additional Figures

A.1 Further Figures

In this section, we report some additional figures of interest that is omitted from the main text due to space constraints. First, Figure A.1 plots the selected percentiles of the future earnings change distribution for every RE percentile. Second, Figure A.1 plots the cross-sectional variance of log earnings over the life cycle, constructed along the lines described in Deaton and Paxson (1994).

A.2 Decomposing Moments: Job-Stayers vs. Job-Switchers

In this section, we present the the cross-sectional statistics analyzed in Section 3 by first splitting the sample in each year depending on whether a worker switched employers or whether he stayed at the same job.

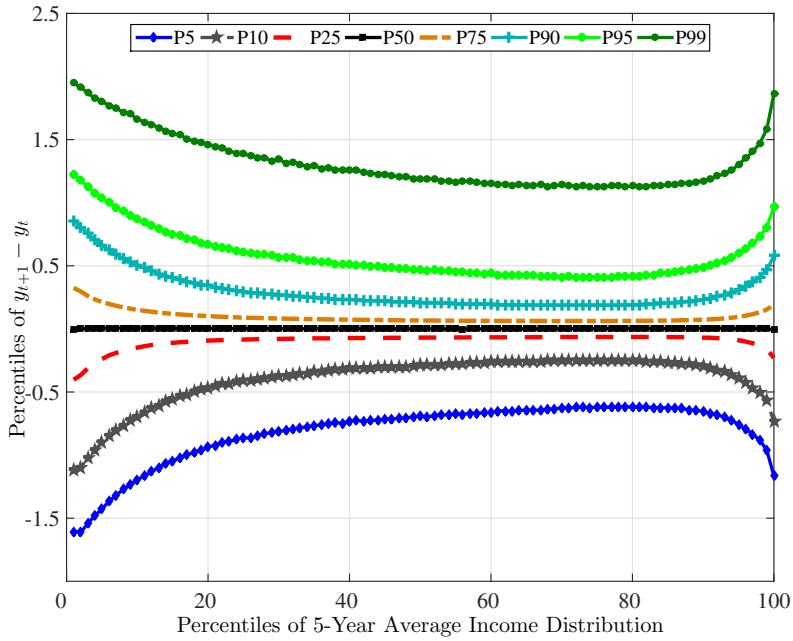
One challenge we face is that many workers hold multiple jobs in a given year, which requires us to be careful about how to think of job changes. We have explored several plausible definitions and found broadly very similar results. Here we describe one reasonable classification. A worker is said to be a "job-stayer" between years t and $t+1$ if a given EIN (employer identification number) provides the largest share of his annual earnings (out of all his EINs in that year) in years $t-1$ through $t+2$, and if the same EIN provides at least 90% of his total annual earnings in years t and $t+1$. A worker is defined as a "job-switcher" if he is not a job-stayer.[25]

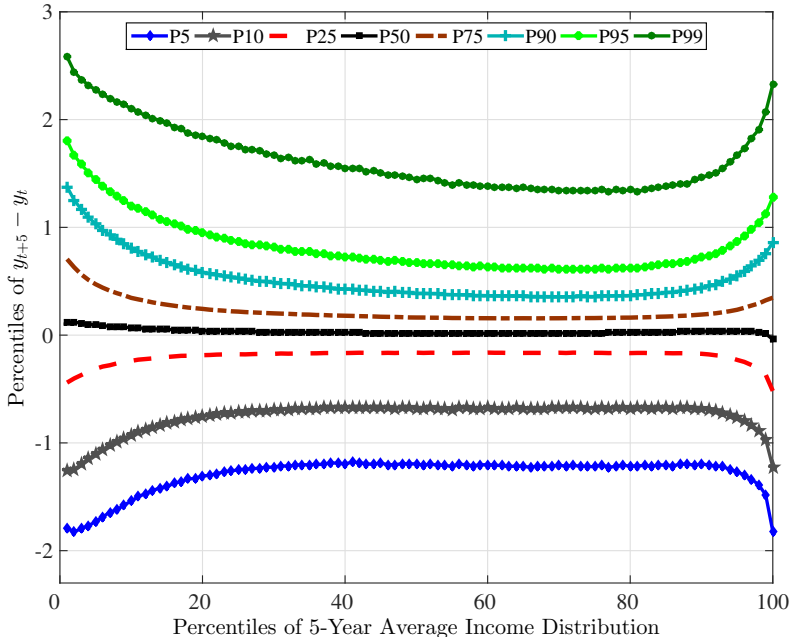
As seen in Figure A.4, the results are consistent with what we might expect. Job-stayers (i) face a dispersion of earnings changes that is less than half that of job-changers, (ii) face shocks that have zero or slightly positive skewness as opposed to job-switchers who face shocks that are very negatively skewed, and (iii) experience shocks with much higher kurtosis than job-switchers. In fact, kurtosis is as high as 43 for annual changes

[25] 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 earnings in year t and provides less than 10% of his annual earnings in year $t+1$; and also has an EIN that provides less than 10% of his earnings in t and more than 50% in $t+1$. The results were very similar to those reported here.

FIGURE A.1 – Quantiles of Earnings Changes

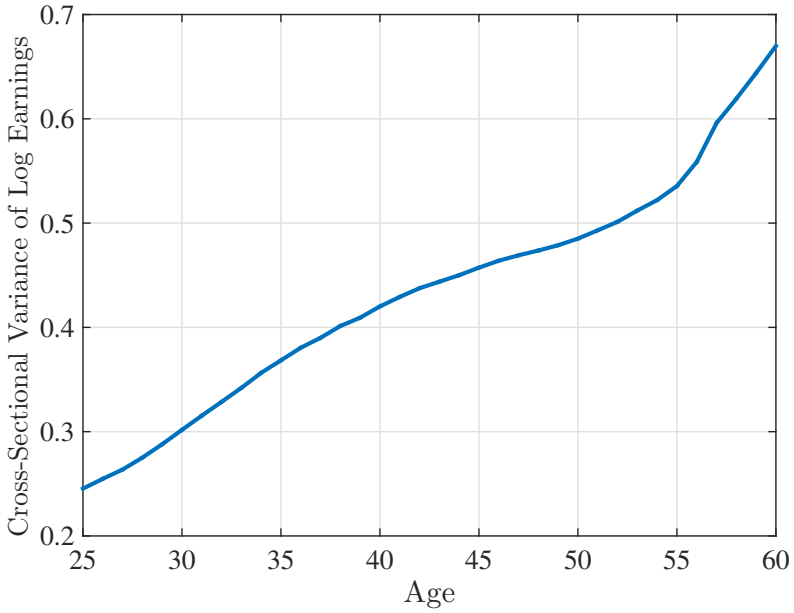
(A) Quantiles of One-Year Earnings Changes





(B) Quantiles of Five-Year Earnings Changes

FIGURE A.2 – Within-Cohort Variance of Log Earnings



and 28.5 for five-year changes for job-stayers, but is less than 10 for job-switchers at both horizons.

Before concluding this section, we examine how the tails of the earnings 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 distribution are critical for the different moments that we estimated so far. Table A.1 reports the 2nd to 4th moments for the original sample used so far (left panel) as well as for a sample where we drop extreme observations, defined as those in the top or bottom 1% of the earnings change distribution. As expected, discarding the tails reduces all statistics, but does not change the ranking between stayers and switchers.

FIGURE A.3 – Fraction Switching Jobs Between t and $t+1$

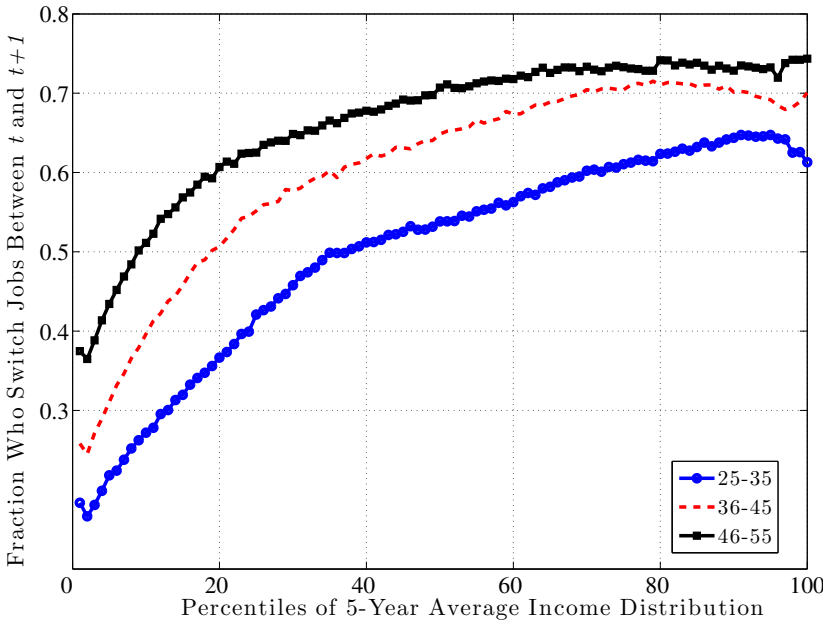
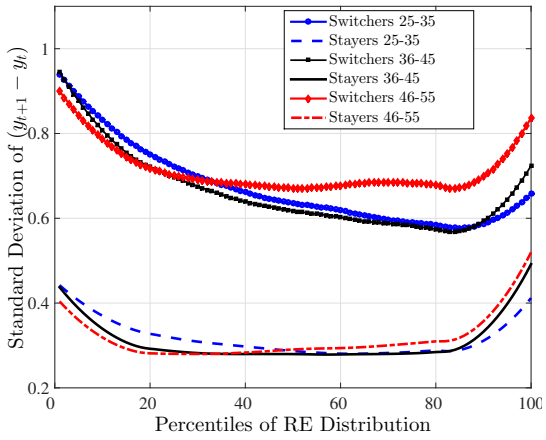


TABLE A.1 – Moments with Trimmed Data

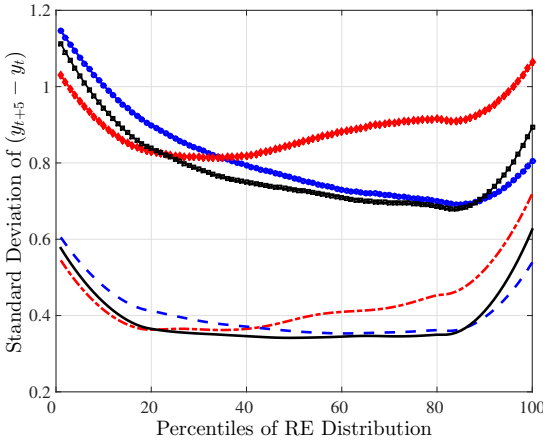
	Original			Trim top/bot. 1%		
	Std. dev.	Skew.	Kurtosis	Std. dev.	Skew.	Kurtosis
	One-Year					
All	0.49	–1.49	18.35	0.36	–0.89	8.91
Stayers	0.31	–1.51	42.81	0.20	–0.51	13.50
Switchers	0.67	–1.09	9.69	0.56	–0.85	6.17
	Five-Year					
All	0.66	–1.06	12.16	0.53	–0.60	6.58
Stayers	0.41	–1.04	28.54	0.29	0.10	7.16
Switchers	0.78	–0.92	8.86	0.65	–0.67	5.61

FIGURE A.4 – Second to Fourth Moments of Earnings Growth: Stayers vs Switchers

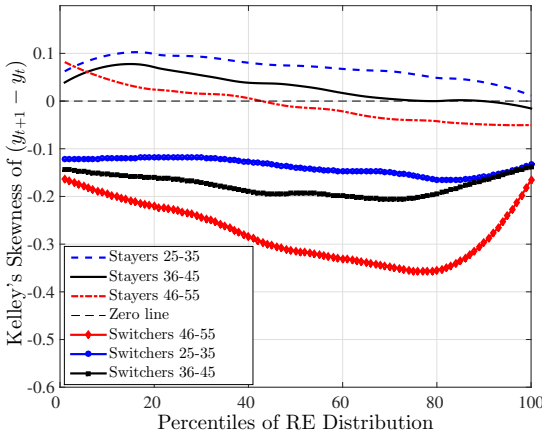
(A) Std. Dev., One-year



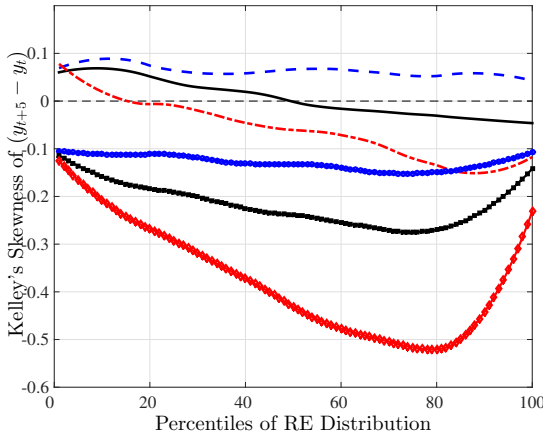
(B) Std. Dev., Five-Year



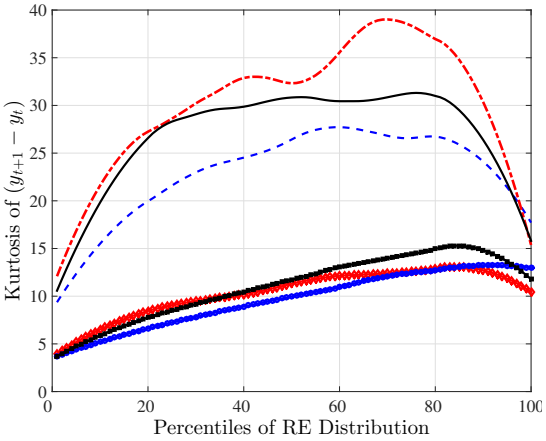
(C) Kelly's Skewness, One-Year



(D) Kelly's Skewness, Five-Year



(E) Kurtosis, One-Year



(F) Kurtosis, Five-Year

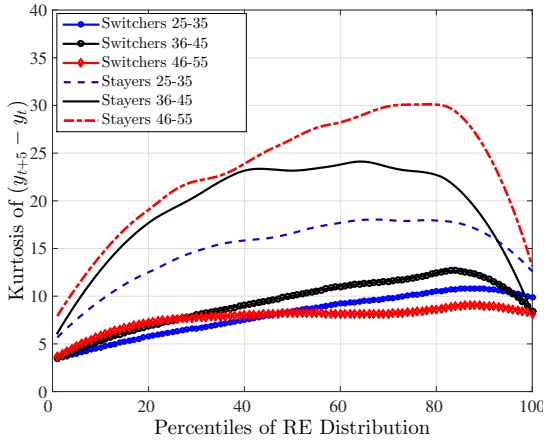


TABLE A.2 – One-Year Transition Probability Matrix: Prob($y_{t+1} \in G_{t+1}$(column)$|y_t \in G_t$(row))

		Wage Income Percentile Group G_{t+1}															
	Pctile	0-1	1-5	5-10	10-20	20-30	30-40	40-50	50-60	60-70	70-80	80-90	90-95	95-97	97-98	98-99	99-100
Wage Income Percentile Group G_t	0-1	**7.74**	15.37	10.81	9.81	3.85	1.87	1.07	0.69	0.48	0.33	0.25	0.10	0.04	0.02	0.01	0.01
	1-5	4.06	**20.32**	14.84	13.68	5.30	2.51	1.41	0.88	0.59	0.43	0.32	0.13	0.04	0.02	0.02	0.02
	5-10	2.25	11.75	**24.48**	21.86	8.03	3.75	2.02	1.21	0.76	0.52	0.37	0.15	0.05	0.03	0.02	0.02
	10-20	1.06	5.38	10.07	**41.46**	16.40	5.83	3.03	1.77	1.12	0.74	0.48	0.17	0.05	0.02	0.02	0.02
	20-30	0.47	2.38	4.00	14.53	**43.96**	16.25	5.14	2.60	1.55	1.04	0.80	0.25	0.08	0.03	0.03	0.02
	30-40	0.26	1.32	2.14	5.78	14.56	**44.81**	16.50	4.66	2.21	1.37	1.07	0.42	0.11	0.04	0.04	0.02
	40-50	0.17	0.82	1.30	3.30	4.86	15.50	**45.39**	16.59	4.26	1.88	1.24	0.75	0.18	0.07	0.05	0.03
	50-60	0.11	0.55	0.84	2.07	2.69	4.34	16.24	**46.85**	16.44	3.70	1.68	0.95	0.35	0.11	0.08	0.04
	60-70	0.08	0.38	0.57	1.38	1.66	2.19	3.95	16.55	**49.59**	15.82	3.06	1.07	0.65	0.22	0.16	0.07
	70-80	0.06	0.29	0.40	0.93	1.14	1.38	1.85	3.45	16.18	**54.30**	14.34	1.66	0.76	0.44	0.31	0.14
	80-90	0.05	0.22	0.29	0.61	0.86	1.09	1.26	1.65	2.85	14.66	**61.37**	9.55	1.35	0.70	0.76	0.41
	90-95	0.04	0.19	0.24	0.47	0.56	0.88	1.47	1.90	2.15	3.09	19.41	**52.99**	9.05	1.98	1.61	1.30
	95-97	0.04	0.18	0.22	0.39	0.43	0.61	0.98	1.68	3.12	3.69	6.26	23.22	**37.92**	10.45	4.60	3.08
	97-98	0.04	0.16	0.20	0.34	0.36	0.48	0.75	1.28	2.14	4.35	6.62	9.22	21.67	**28.59**	14.65	5.76
	98-99	0.03	0.14	0.19	0.31	0.32	0.38	0.55	0.91	1.70	3.13	7.42	7.53	8.54	14.94	**35.25**	15.01
	99-100	0.03	0.13	0.19	0.26	0.24	0.27	0.32	0.48	0.80	1.69	4.41	6.61	5.89	5.28	14.33	**54.83**
S.E. Income Percentile Group G_t	0-5	1.00	4.21	5.14	8.05	4.79	2.74	1.67	1.12	0.73	0.54	0.37	0.14	0.05	0.02	0.03	0.02
	5-10	0.77	3.25	4.04	6.81	5.80	4.13	2.47	1.49	1.00	0.67	0.46	0.17	0.06	0.02	0.02	0.02
	10-20	0.56	2.48	3.11	5.00	4.05	3.67	3.30	2.64	1.78	1.02	0.62	0.25	0.08	0.04	0.03	0.02
	20-30	0.43	1.85	2.42	3.72	2.81	2.39	2.22	2.19	2.24	1.96	1.11	0.35	0.12	0.05	0.05	0.03
	30-40	0.35	1.43	1.92	2.99	2.16	1.78	1.60	1.55	1.57	1.71	1.71	0.57	0.17	0.08	0.07	0.04
	40-50	0.27	1.12	1.42	2.39	1.78	1.40	1.28	1.18	1.22	1.31	1.57	0.89	0.26	0.10	0.09	0.06
	50-60	0.22	0.83	1.08	1.81	1.44	1.18	1.04	0.99	0.99	1.08	1.37	0.95	0.43	0.16	0.13	0.09
	60-70	0.17	0.64	0.78	1.32	1.15	0.98	0.88	0.84	0.88	0.96	1.25	0.91	0.50	0.27	0.22	0.14
	70-80	0.12	0.48	0.56	0.96	0.83	0.78	0.75	0.73	0.79	0.90	1.17	0.89	0.51	0.31	0.35	0.24
	80-90	0.09	0.34	0.39	0.66	0.58	0.55	0.55	0.58	0.67	0.82	1.24	1.01	0.58	0.36	0.44	0.46
	90-95	0.07	0.23	0.28	0.45	0.40	0.38	0.40	0.46	0.51	0.65	1.15	1.15	0.74	0.46	0.54	0.69
	95-98	0.05	0.17	0.19	0.29	0.23	0.24	0.24	0.28	0.34	0.42	0.77	0.91	0.79	0.58	0.70	0.96
	98-99	0.05	0.13	0.15	0.20	0.16	0.16	0.15	0.17	0.23	0.28	0.53	0.61	0.61	0.56	0.83	1.32
	99-100	0.04	0.12	0.14	0.21	0.15	0.10	0.12	0.16	0.16	0.22	0.43	0.51	0.56	0.45	0.80	2.51
	Death	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	$<Y_{\min}$	1.18	3.52	2.73	2.62	1.17	0.66	0.43	0.31	0.25	0.21	0.20	0.10	0.05	0.02	0.03	0.03

Table VI continues.

		Self Employment Income Percentile Group G_{t+1}															
	Pctile	0-5	5-10	10-20	20-30	30-40	40-50	50-60	60-70	70-80	80-90	90-95	95-98	98-99	99-100	Death	$<Y_{\min}$
Wage Income Percentile Group G_t	0-1	0.62	0.47	0.72	0.56	0.46	0.36	0.27	0.19	0.13	0.09	0.03	0.01	0.00	0.00	0.65	42.96
	1-5	0.63	0.50	0.77	0.59	0.47	0.36	0.26	0.19	0.12	0.08	0.03	0.01	0.00	0.00	0.56	30.87
	5-10	0.58	0.47	0.73	0.59	0.48	0.35	0.25	0.18	0.12	0.07	0.02	0.01	0.00	0.00	0.45	18.40
	10-20	0.43	0.38	0.55	0.43	0.35	0.28	0.21	0.15	0.10	0.06	0.02	0.01	0.00	0.00	0.34	9.05
	20-30	0.25	0.31	0.44	0.31	0.25	0.20	0.16	0.13	0.09	0.06	0.02	0.01	0.00	0.00	0.25	4.39
	30-40	0.15	0.22	0.39	0.26	0.20	0.16	0.14	0.11	0.09	0.06	0.02	0.01	0.00	0.00	0.21	2.66
	40-50	0.10	0.14	0.36	0.25	0.18	0.15	0.13	0.10	0.09	0.06	0.02	0.01	0.00	0.00	0.18	1.84
	50-60	0.07	0.09	0.29	0.25	0.18	0.14	0.12	0.10	0.09	0.07	0.02	0.01	0.00	0.00	0.16	1.36
	60-70	0.05	0.06	0.20	0.25	0.19	0.14	0.12	0.11	0.10	0.08	0.03	0.01	0.00	0.00	0.14	1.10
	70-80	0.04	0.04	0.12	0.21	0.20	0.15	0.13	0.12	0.11	0.10	0.04	0.01	0.00	0.00	0.13	0.94
	80-90	0.03	0.03	0.08	0.13	0.19	0.19	0.17	0.15	0.15	0.15	0.07	0.03	0.01	0.00	0.11	0.85
	90-95	0.03	0.03	0.07	0.09	0.13	0.20	0.23	0.23	0.23	0.26	0.14	0.08	0.02	0.01	0.10	0.86
	95-97	0.02	0.03	0.06	0.08	0.10	0.15	0.24	0.30	0.32	0.37	0.24	0.16	0.04	0.03	0.09	0.91
	97-98	0.03	0.02	0.05	0.08	0.09	0.12	0.19	0.31	0.38	0.45	0.30	0.22	0.08	0.05	0.08	0.93
	98-99	0.02	0.02	0.05	0.06	0.08	0.11	0.15	0.26	0.42	0.56	0.36	0.27	0.13	0.10	0.08	0.97
	99-100	0.03	0.02	0.05	0.05	0.07	0.08	0.13	0.19	0.31	0.60	0.49	0.40	0.19	0.34	0.08	1.21
S.E. Income Percentile Group G_t	0-5	**11.40**	6.73	8.36	5.45	3.90	2.96	2.22	1.57	1.05	0.61	0.17	0.06	0.01	0.01	0.27	24.60
	5-10	6.76	**9.76**	11.71	6.88	4.68	3.38	2.50	1.79	1.15	0.65	0.16	0.05	0.01	0.01	0.24	19.11
	10-20	4.12	5.74	**16.73**	10.51	6.41	4.48	3.21	2.21	1.43	0.80	0.20	0.06	0.01	0.01	0.22	15.22
	20-30	2.68	3.38	10.25	**17.47**	11.26	6.92	4.57	3.05	1.88	1.03	0.26	0.08	0.01	0.01	0.22	12.98
	30-40	1.86	2.30	6.27	11.10	**18.78**	12.27	7.21	4.37	2.56	1.36	0.33	0.11	0.02	0.01	0.20	11.54
	40-50	1.41	1.62	4.29	6.64	12.00	**20.44**	13.26	7.14	3.92	1.93	0.46	0.14	0.02	0.02	0.18	10.20
	50-60	1.02	1.18	3.01	4.32	6.89	12.95	**22.65**	14.24	6.76	3.08	0.69	0.21	0.04	0.02	0.17	8.99
	60-70	0.73	0.81	2.04	2.74	4.08	6.80	13.90	**26.25**	15.27	5.83	1.14	0.33	0.06	0.04	0.15	7.96
	70-80	0.49	0.52	1.28	1.66	2.30	3.53	6.32	14.81	**32.73**	15.81	2.42	0.62	0.11	0.06	0.15	6.86
	80-90	0.29	0.29	0.71	0.91	1.18	1.71	2.75	5.38	15.18	**43.43**	10.47	1.86	0.29	0.16	0.13	5.94
	90-95	0.16	0.15	0.36	0.43	0.58	0.80	1.20	2.02	4.33	20.38	**42.50**	11.32	1.13	0.56	0.12	5.42
	95-98	0.09	0.09	0.20	0.22	0.28	0.38	0.57	0.91	1.77	5.55	18.87	**48.71**	8.05	2.37	0.10	4.71
	98-99	0.08	0.06	0.10	0.12	0.15	0.22	0.32	0.49	0.92	2.46	4.96	24.61	**40.91**	14.07	0.08	4.31
	99-100	0.05	0.04	0.08	0.08	0.11	0.12	0.20	0.33	0.57	1.44	2.34	6.05	13.83	**62.83**	0.10	5.16
	Death	0.00	100.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	**100**	0.00
	$<Y_{\min}$	0.43	0.34	0.55	0.46	0.41	0.35	0.29	0.25	0.20	0.16	0.07	0.04	0.01	0.01	0.79	**82.12**

Notes: This table is constructed similar to the previous table except that the last two rows indicate out-of-sample, respectively, because of (i) death or aging (turning 61), (ii) having wage/salary and S.E. earnings below $Y_{\min}$.

TABLE A.3 – Five-Year Transition Probability Matrix: Prob($y_{t+5} \in G_{t+5}$(column)$|y_t \in G_t$(row))

		Wage Income Percentile Group G_{t+5}															
	Pctile	0-1	1-5	5-10	10-20	20-30	30-40	40-50	50-60	60-70	70-80	80-90	90-95	95-97	97-98	98-99	99-100
Wage Income Percentile Group G_t	0-1	**3.49**	9.12	7.63	9.62	5.46	3.49	2.37	1.76	1.41	1.09	0.87	0.38	0.14	0.06	0.05	0.04
	1-5	2.56	**10.62**	9.53	11.87	6.74	4.16	2.84	2.09	1.62	1.28	1.01	0.42	0.15	0.07	0.07	0.05
	5-10	1.81	8.23	**12.48**	16.87	8.91	5.37	3.60	2.56	1.89	1.45	1.12	0.47	0.17	0.08	0.07	0.06
	10-20	1.15	5.21	8.23	**23.55**	15.63	7.86	4.79	3.22	2.26	1.67	1.24	0.50	0.17	0.08	0.07	0.05
	20-30	0.71	3.19	4.75	14.12	**23.76**	16.41	8.07	4.59	2.95	2.03	1.47	0.58	0.20	0.09	0.07	0.05
	30-40	0.49	2.22	3.20	8.01	14.54	**23.41**	16.75	8.12	4.35	2.64	1.81	0.72	0.23	0.10	0.09	0.06
	40-50	0.38	1.67	2.34	5.39	7.84	15.23	**23.37**	16.84	7.91	3.94	2.25	0.98	0.30	0.12	0.10	0.06
	50-60	0.30	1.32	1.79	3.84	4.83	7.93	15.87	**24.33**	17.06	7.32	3.27	1.29	0.44	0.17	0.14	0.09
	60-70	0.25	1.06	1.40	2.82	3.21	4.43	7.88	16.52	**26.50**	17.48	6.25	1.73	0.71	0.27	0.23	0.16
	70-80	0.21	0.86	1.09	2.09	2.23	2.75	4.02	7.47	17.18	**30.80**	17.64	3.15	1.05	0.49	0.38	0.25
	80-90	0.16	0.65	0.80	1.49	1.60	1.89	2.35	3.44	6.58	17.36	**38.68**	12.09	2.38	0.99	0.87	0.52
	90-95	0.13	0.53	0.65	1.15	1.20	1.52	2.03	2.78	3.66	6.67	22.82	**30.78**	9.69	3.20	2.45	1.68
	95-97	0.12	0.46	0.58	1.03	1.03	1.26	1.67	2.59	3.94	5.45	11.93	21.55	**18.41**	8.79	6.43	4.31
	97-98	0.11	0.44	0.54	0.93	0.93	1.09	1.46	2.07	3.27	5.35	10.28	14.21	15.58	**12.54**	11.81	7.88
	98-99	0.10	0.41	0.50	0.85	0.83	0.96	1.22	1.75	2.69	4.46	9.56	11.24	10.88	10.54	**16.91**	14.87
	99-100	0.11	0.41	0.48	0.76	0.67	0.70	0.81	1.13	1.70	2.91	6.60	8.66	7.54	6.62	12.87	**34.11**
S.E. Income Percentile Group G_t	0-5	0.94	3.74	4.70	8.10	6.25	4.60	3.49	2.61	2.03	1.56	1.17	0.49	0.18	0.08	0.09	0.07
	5-10	0.80	3.32	4.20	7.62	6.96	5.78	4.19	3.15	2.35	1.77	1.33	0.54	0.20	0.09	0.08	0.06
	10-20	0.65	2.82	3.64	6.40	5.77	5.34	4.78	4.08	3.19	2.25	1.55	0.64	0.23	0.10	0.09	0.07
	20-30	0.58	2.38	3.18	5.48	4.74	4.21	3.86	3.60	3.43	3.02	2.20	0.81	0.27	0.13	0.12	0.10
	30-40	0.51	2.11	2.82	4.97	4.22	3.70	3.29	2.96	2.83	2.75	2.63	1.08	0.36	0.16	0.15	0.12
	40-50	0.46	1.86	2.42	4.38	3.82	3.37	2.96	2.65	2.50	2.42	2.51	1.32	0.47	0.21	0.20	0.16
	50-60	0.39	1.62	2.13	3.83	3.45	3.09	2.80	2.54	2.34	2.27	2.41	1.43	0.64	0.29	0.25	0.20
	60-70	0.34	1.38	1.78	3.23	3.05	2.84	2.64	2.47	2.30	2.25	2.42	1.49	0.74	0.39	0.36	0.28
	70-80	0.28	1.13	1.49	2.73	2.63	2.52	2.48	2.36	2.33	2.35	2.55	1.63	0.83	0.48	0.50	0.42
	80-90	0.22	0.91	1.18	2.17	2.08	2.03	2.14	2.17	2.31	2.57	3.24	2.18	1.14	0.66	0.73	0.72
	90-95	0.17	0.69	0.88	1.52	1.50	1.55	1.60	1.69	1.93	2.40	3.76	3.27	1.89	1.00	1.13	1.15
	95-98	0.13	0.48	0.59	0.97	0.92	0.93	0.99	1.17	1.35	1.73	3.00	3.29	2.63	1.76	1.96	1.90
	98-99	0.10	0.34	0.45	0.67	0.58	0.62	0.60	0.73	0.88	1.11	1.96	2.48	2.52	2.13	3.01	3.50
	99-100	0.10	0.35	0.44	0.58	0.46	0.45	0.45	0.54	0.60	0.78	1.40	1.86	2.08	1.78	2.73	6.50
	Death	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	$<Y_{\min}$	1.08	3.68	3.65	5.24	3.30	2.19	1.56	1.21	1.02	0.92	0.95	0.54	0.25	0.14	0.14	0.13

Table VII continues.

		Self Employment Income Percentile Group G_{t+5}															
	Pctile	0-5	5-10	10-20	20-30	30-40	40-50	50-60	60-70	70-80	80-90	90-95	95-98	98-99	99-100	Death	$<Y_{\min}$
Wage Income Percentile Group G_t	0-1	0.52	0.49	0.83	0.74	0.68	0.60	0.45	0.37	0.27	0.19	0.08	0.04	0.01	0.01	0.74	46.99
	1-5	0.54	0.50	0.87	0.80	0.71	0.60	0.47	0.38	0.28	0.21	0.07	0.03	0.01	0.01	0.65	38.79
	5-10	0.54	0.50	0.87	0.80	0.73	0.64	0.51	0.39	0.29	0.20	0.07	0.03	0.01	0.01	0.54	28.74
	10-20	0.46	0.44	0.76	0.68	0.62	0.56	0.46	0.37	0.28	0.19	0.07	0.03	0.01	0.00	0.44	18.95
	20-30	0.35	0.38	0.67	0.57	0.52	0.48	0.43	0.37	0.29	0.20	0.06	0.03	0.01	0.00	0.35	12.24
	30-40	0.28	0.32	0.61	0.51	0.46	0.42	0.40	0.37	0.30	0.21	0.06	0.03	0.01	0.01	0.29	8.98
	40-50	0.23	0.26	0.55	0.48	0.42	0.39	0.38	0.37	0.33	0.25	0.07	0.03	0.01	0.00	0.25	7.25
	50-60	0.20	0.21	0.48	0.47	0.40	0.37	0.36	0.36	0.35	0.29	0.10	0.03	0.01	0.01	0.22	6.14
	60-70	0.17	0.17	0.39	0.43	0.38	0.36	0.35	0.36	0.37	0.36	0.13	0.05	0.01	0.01	0.20	5.35
	70-80	0.14	0.14	0.31	0.36	0.37	0.35	0.34	0.35	0.38	0.41	0.19	0.07	0.02	0.01	0.17	4.73
	80-90	0.12	0.12	0.24	0.28	0.33	0.35	0.36	0.39	0.43	0.54	0.32	0.15	0.03	0.02	0.15	4.30
	90-95	0.10	0.10	0.21	0.23	0.29	0.35	0.42	0.47	0.55	0.76	0.56	0.38	0.09	0.06	0.13	4.35
	95-97	0.11	0.10	0.19	0.22	0.26	0.32	0.41	0.54	0.67	0.94	0.75	0.67	0.24	0.16	0.12	4.73
	97-98	0.11	0.09	0.19	0.20	0.24	0.29	0.40	0.55	0.75	1.09	0.85	0.84	0.41	0.31	0.11	5.08
	98-99	0.09	0.08	0.18	0.19	0.23	0.27	0.36	0.52	0.75	1.17	0.96	0.95	0.56	0.54	0.10	5.26
	99-100	0.11	0.10	0.18	0.17	0.21	0.25	0.33	0.45	0.67	1.24	1.13	1.10	0.60	1.19	0.10	6.10
S.E. Income Percentile Group G_t	0-5	**4.21**	3.20	5.40	4.19	3.48	3.05	2.51	2.05	1.56	1.03	0.34	0.12	0.02	0.02	0.38	28.35
	5-10	3.23	**3.39**	5.95	4.71	3.93	3.22	2.75	2.21	1.60	1.07	0.31	0.11	0.02	0.01	0.36	24.68
	10-20	2.60	2.90	**6.60**	5.86	4.69	3.92	3.28	2.57	1.88	1.23	0.36	0.13	0.02	0.02	0.34	21.98
	20-30	2.16	2.39	5.69	**6.98**	6.31	5.25	4.21	3.24	2.34	1.49	0.45	0.15	0.03	0.03	0.31	20.86
	30-40	1.82	2.00	4.84	6.31	**7.58**	7.16	5.70	4.25	2.95	1.84	0.56	0.19	0.04	0.03	0.30	19.73
	40-50	1.59	1.71	4.04	5.23	7.02	**8.70**	7.99	6.00	4.05	2.41	0.71	0.24	0.05	0.04	0.28	18.23
	50-60	1.30	1.40	3.28	4.14	5.58	7.76	**10.04**	8.92	6.06	3.55	0.97	0.32	0.07	0.05	0.26	16.59
	60-70	1.02	1.07	2.52	3.08	4.10	5.77	8.72	**12.08**	10.24	5.85	1.48	0.48	0.10	0.07	0.24	15.22
	70-80	0.76	0.78	1.79	2.13	2.71	3.82	5.90	10.01	**16.32**	11.76	2.65	0.79	0.16	0.11	0.21	13.37
	80-90	0.52	0.52	1.12	1.33	1.66	2.18	3.33	5.66	11.55	**23.01**	7.75	2.02	0.37	0.24	0.18	12.11
	90-95	0.34	0.30	0.68	0.77	0.88	1.17	1.70	2.69	5.18	16.39	**21.78**	8.80	1.28	0.76	0.15	11.00
	95-98	0.21	0.19	0.39	0.44	0.52	0.68	0.91	1.43	2.55	7.12	16.12	**26.65**	6.52	2.79	0.14	9.54
	98-99	0.15	0.13	0.29	0.29	0.34	0.45	0.60	0.84	1.58	3.86	7.03	21.55	**19.94**	12.46	0.13	8.68
	99-100	0.16	0.13	0.24	0.21	0.29	0.34	0.43	0.61	1.10	2.42	3.87	8.86	12.25	**36.98**	0.11	10.92
	Death	0.00	100.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	**100**	0.00
	$<Y_{\min}$	0.45	0.42	0.77	0.72	0.67	0.59	0.51	0.42	0.34	0.26	0.10	0.05	0.01	0.01	0.72	**67.93**

Notes: This table is constructed similar to the previous table except that the last two rows indicate out-of-sample, respectively, because of (i) death or aging (turning 61), (ii) having wage/salary and S.E. earnings below $Y_{\min}$.

A.3 Cross-Sectional Facts

A.4 State Transition Matrices for One- and Five-Year Earnings

B Details of Estimation Method

B.1 Moment Selection and Aggregation

1. 1. Cross-sectional moments of earnings changes. In Section 3, we have dropped individuals with low earnings ($\exp(y_{it}) < \overline{Y}_{\min}$) in t or $t+k$ (or both) when constructing the moments of the cross-sectional distribution. We now modify this procedure slightly so as to allow the incorporation of these previously dropped observations. The basic idea is to transform these observations by mapping them into earnings values that are slightly above $\overline{Y}_{\min}$. Specifically, the transformed earnings values are obtained as

$$\tilde{y}_{i,t} = \ln\left(\overline{Y}_{\min} + \frac{\overline{Y}_{\min}}{10\left(\overline{Y}_{\min}+10\right)}\left(\exp y_{i,t} + 10x\right)\right),$$

where $x \sim U(0,1)$ follows a uniform distribution.

In order to have a smooth earnings change distribution, we replace earnings observations lower than our minimum earnings threshold by mapping them to levels above. This imputation has the following three goals: First, there is a mass of individuals with an earnings observation of zero, which would otherwise be dropped when constructing log earnings changes. We would like to use these observations as they contain information about the nature of earnings risk. Second, we would like to avoid having any mass in the distribution to avoid possible jaggedness in the objective function. Third, we want to preserve the ranking among individuals with less than threshold earnings.

In order to capture the variation in the cross-sectional moments of earnings changes along the age and recent earnings dimensions, we condition the distribution of earnings changes on these variables. For this purpose, we first group workers into 6 age bins (five-year age bins between 25 and 54) and within each age bin into 13 selected groups of RE percentiles in age $t-1$. The RE percentiles are grouped as follows: 1, 2–10, 11–0, 21–30, ..., 81–90, 91–95, 96–99, 100. Thus, we compute the three moments of the distribution of one- and five-year earnings changes for $6 \times 13 = 78$ different groups of

workers. We aggregate these 6 age bins into 2 age groups, A^i_{t-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. Consequently, we end up with $2 \times 2 \times 13 \times 3 = 156$ cross-sectional moments.

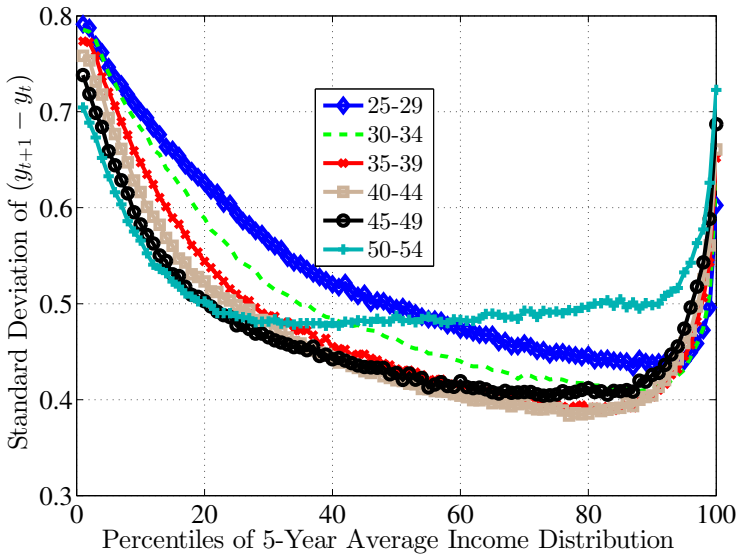
2. Mean of log earnings growth. The second set of moments captures the heterogeneity in log earnings growth over the working life across workers that are in different percentiles of the LE distribution. We target the average dollar earnings at 8 points over the life cycle: ages 25, 30, ..., and 60 for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

3. Impulse response functions. We target average log earnings growth over the next k years for $k = 1, 2, 3, 5, 10$, that is, $\mathbb{E}[y_{t+k} - y_t]$, conditional on groups formed by crossing age, $\overline{Y}_{t-1}$, and $y^i_t - y^i_{t-1}$. We impute income observations lower than our minimum income threshold following the same procedure as in the smoothing of the cross-sectional moments. We aggregate age groups into two: young workers (25-34) and prime-age workers (35-55). In each year, individuals are assigned to groups based on their ranking in the age-specific RE distribution. The following list defines these groups in terms of the RE distribution: 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 age- and RE-specific income change distribution, $y^i_t - y^i_{t-1}$: 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 based on impulse response.

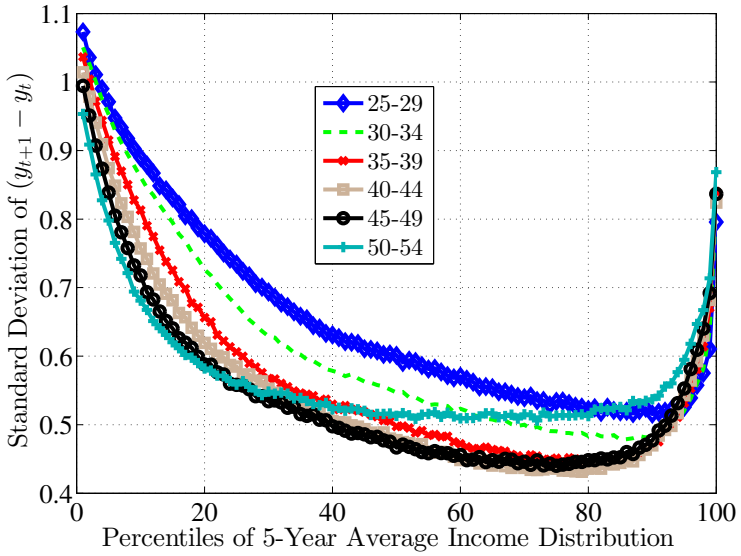
4. Variance of log earnings. We target the life cycle profile of the variance of log earnings for income observations larger than our minimum income threshold. As a result, we have a total of 36 moments based on the variance of log earnings.

B.2 Numerical Method for Estimation

The estimation objective is maximized as described now. 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. 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. We have found that the combination balances speed with reliability and provides good results.



(A) Bi-Annual



(B) Relative to Usual

FIGURE A.5 – Standard Deviation of Annual Earnings Growth

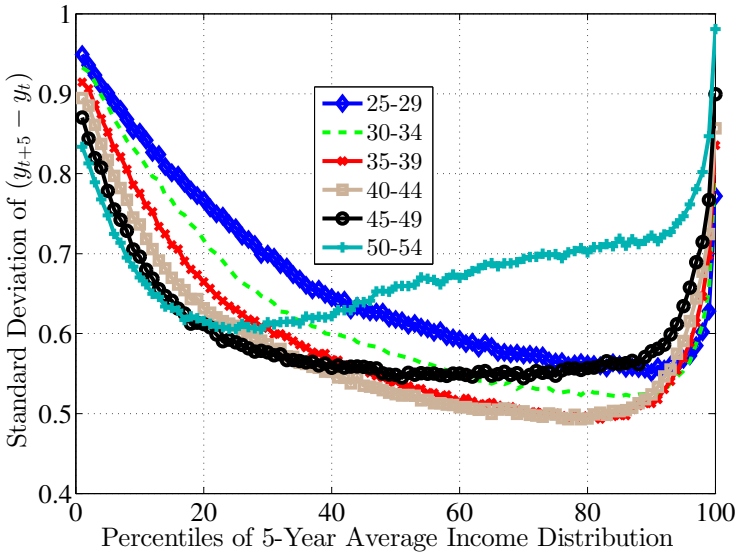


FIGURE A.6 – Standard Deviation of Biann Five-Year Earnings Growth

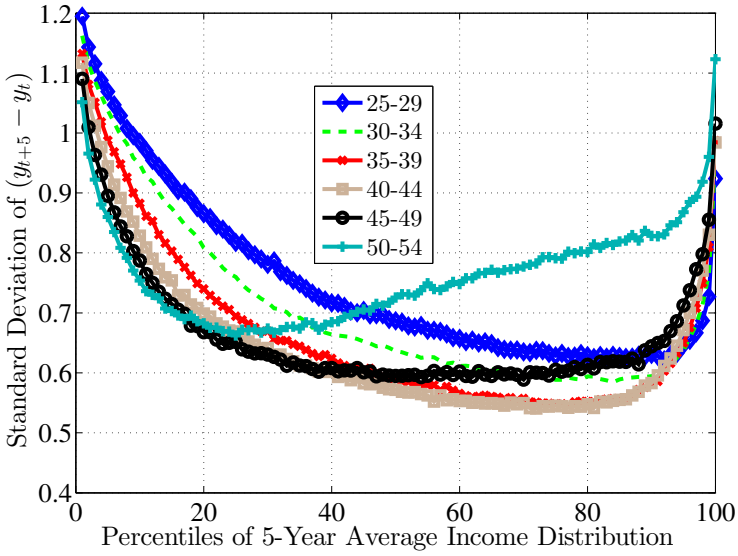


FIGURE A.7 – Standard Deviation of Five-Year Earnings Growth, Usual

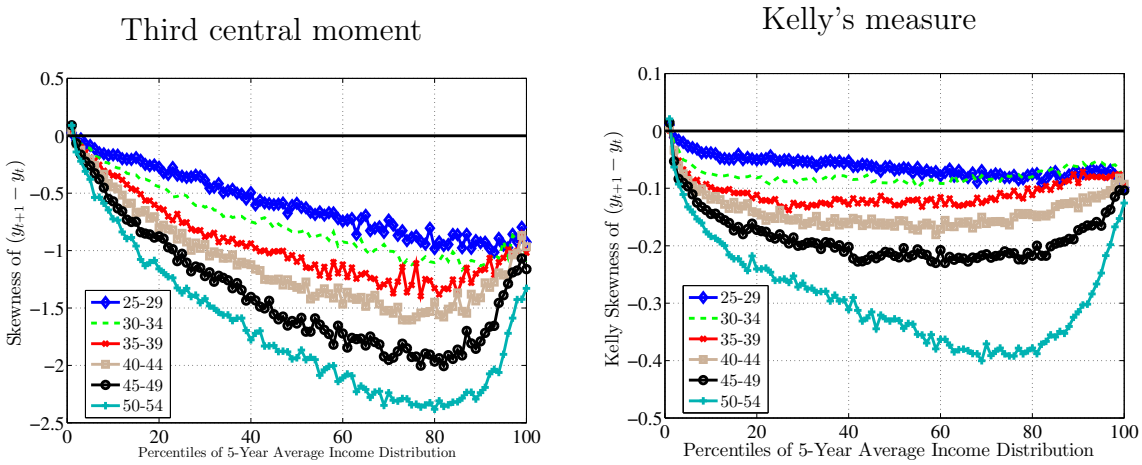


FIGURE A.8 – Skewness of (Bi-Ann) Annual Earnings Growth

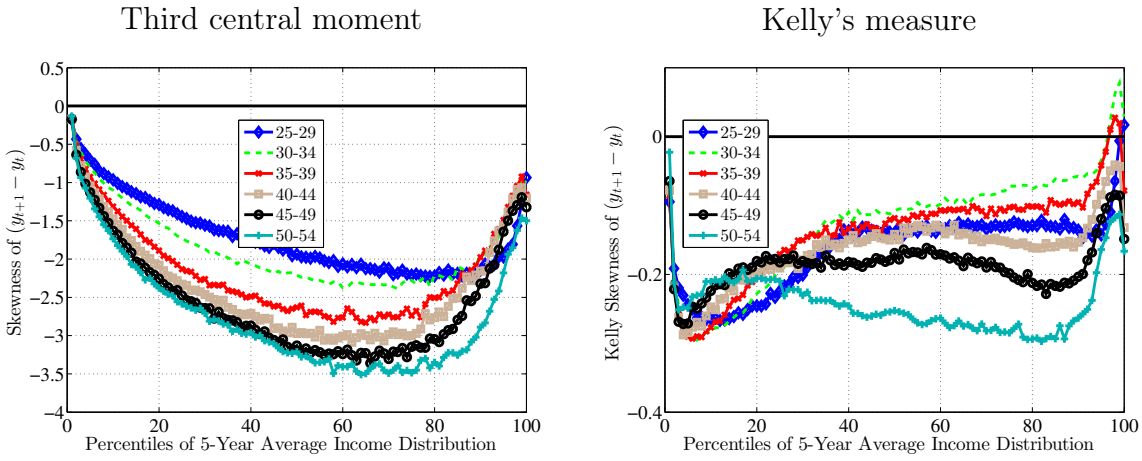


FIGURE A.9 – Skewness of Annual Earnings Growth, Usual

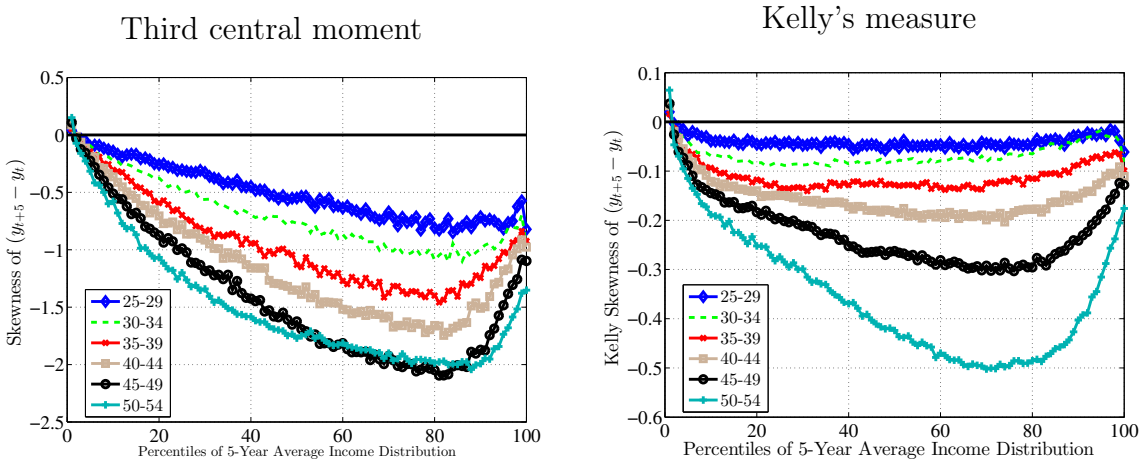


FIGURE A.10 – Skewness of (Bi-Ann) Five-Year Earnings Growth

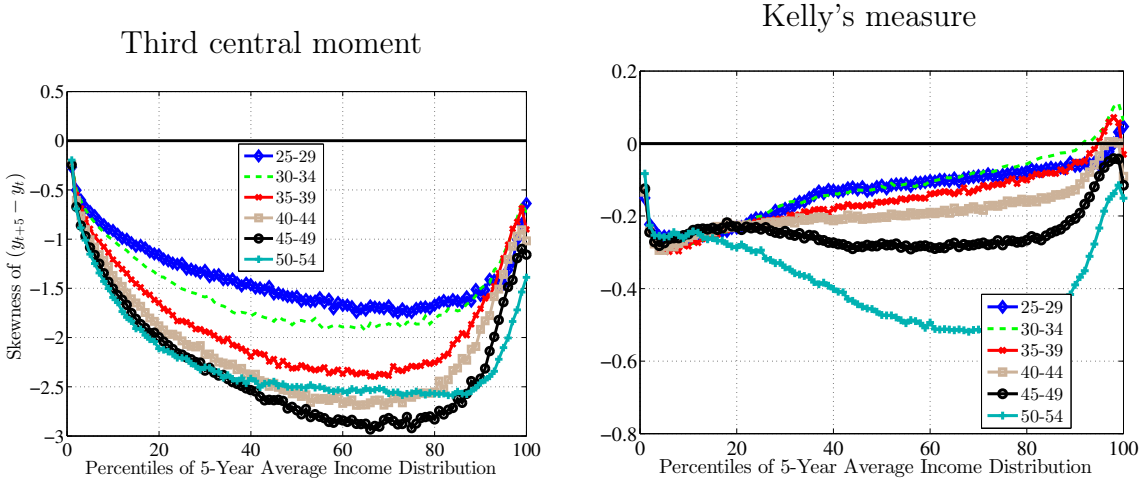


FIGURE A.11 – Skewness of Five-Year Earnings Growth, Usual

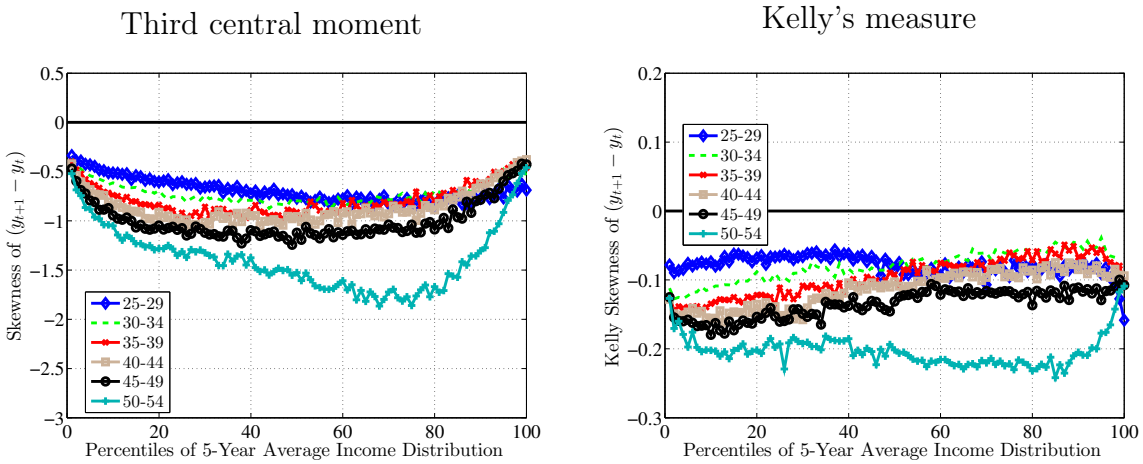


FIGURE A.12 – Skewness of One-Year Earnings Growth, Excluding bottom and top 1% of earnings change

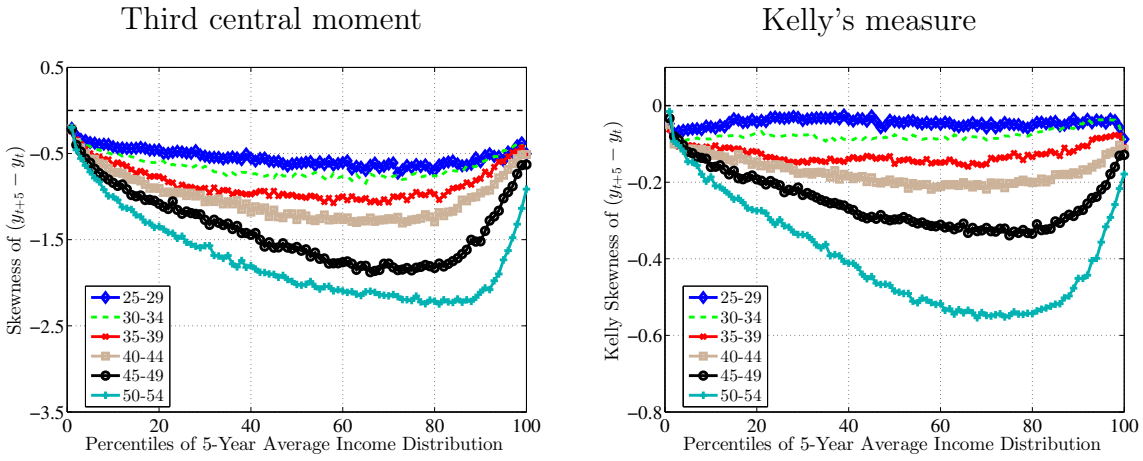


FIGURE A.13 – Skewness of Five-Year Earnings Growth, Excluding bottom and top 1% of earnings change

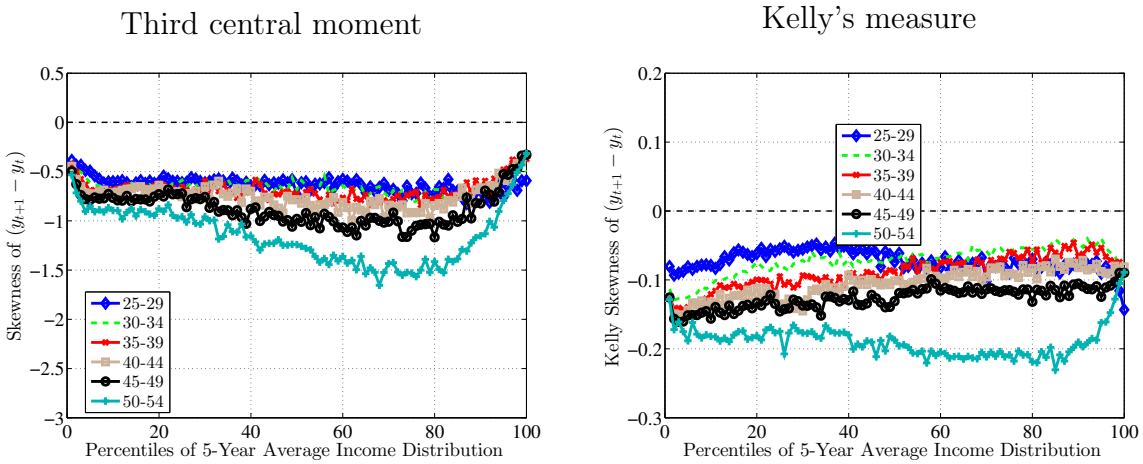


FIGURE A.14 – Skewness of One-Year Earnings Growth, Excluding those with earnings lower than 5% of average past earnings

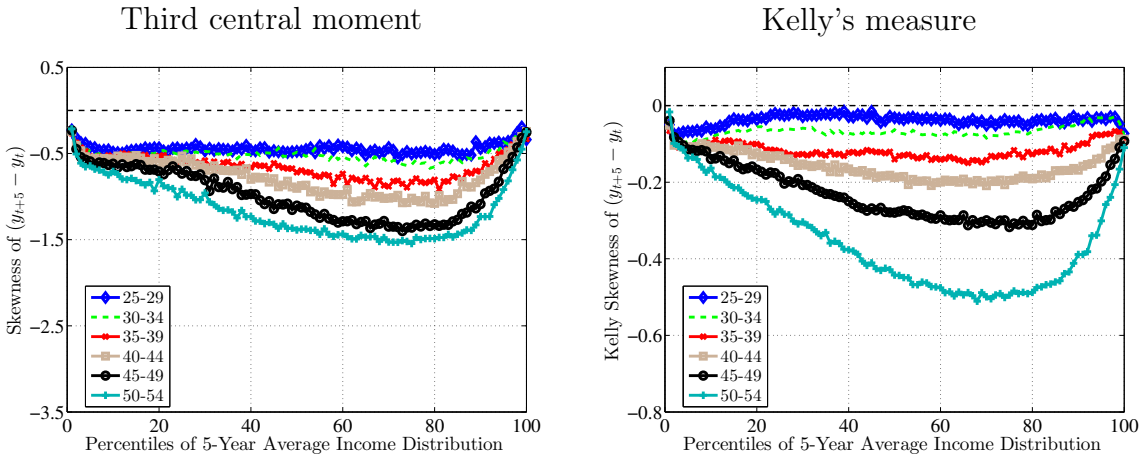


FIGURE A.15 – Skewness of Five-Year Earnings Growth, Excluding those with earnings lower than 5% of average past earnings

FIGURE A.16 – Kurtosis of Annual earnings Change, By Age and Past Earnings, (2 Year Average Earnings)

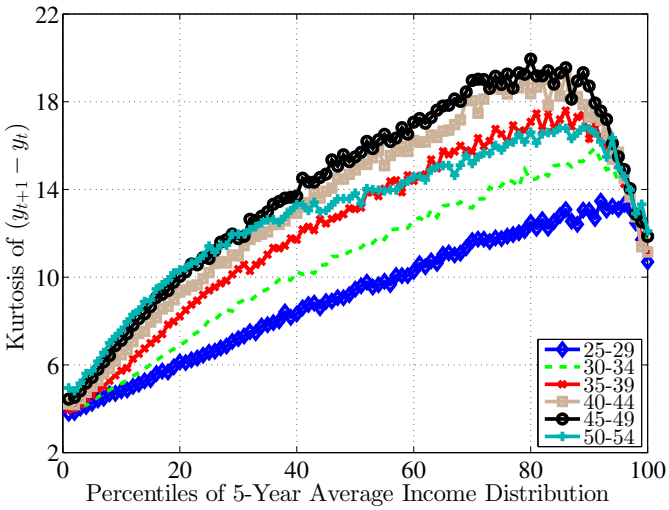


FIGURE A.17 – Kurtosis of Annual Earnings Change, By Age and Past Earnings, Usual

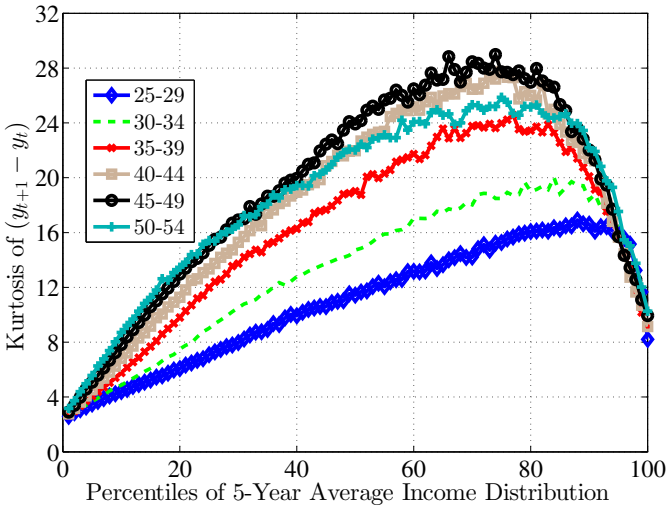


FIGURE A.18 – Kurtosis of Five-Year Earnings Change, By Age and Past Earnings, (2 Year Average Earnings)

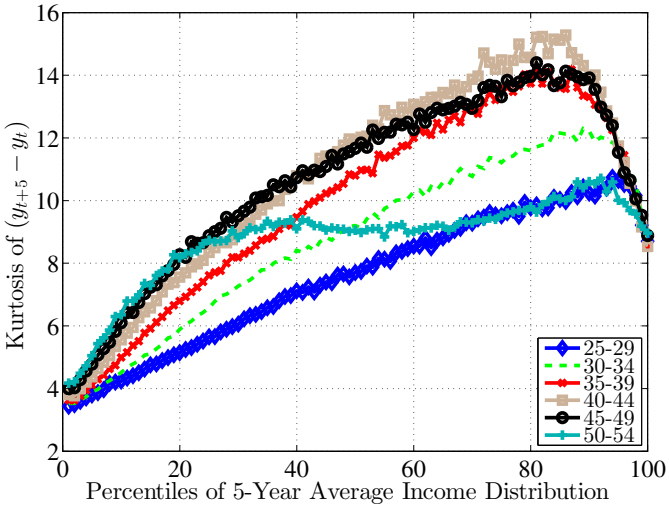


FIGURE A.19 – Kurtosis of Five-Year Earnings Change, Usual Earnings

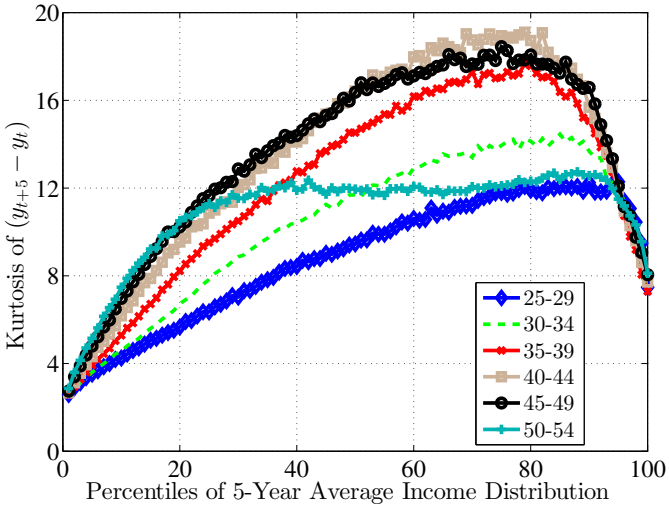


FIGURE A.20 – Kurtosis of One-Year Earnings Change, Excluding bottom and top 1% of Earnings change

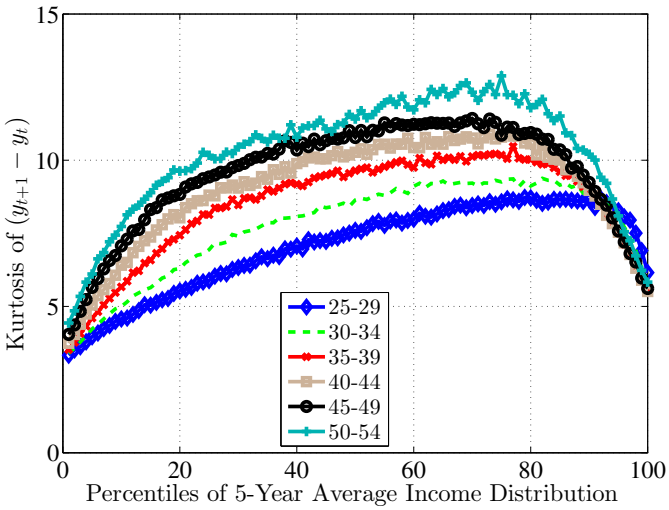
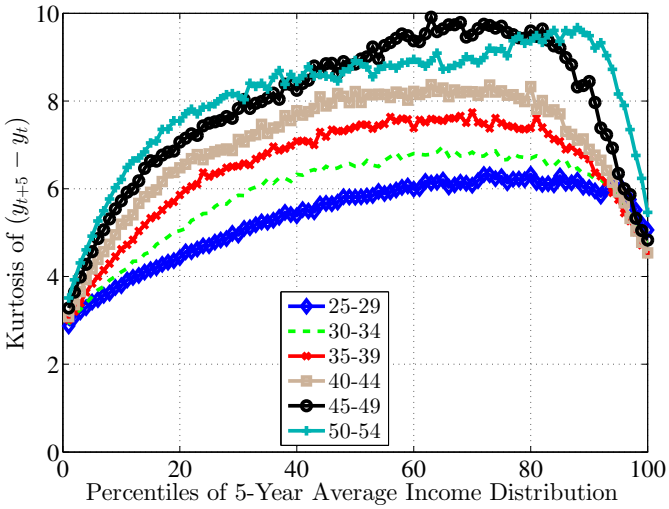
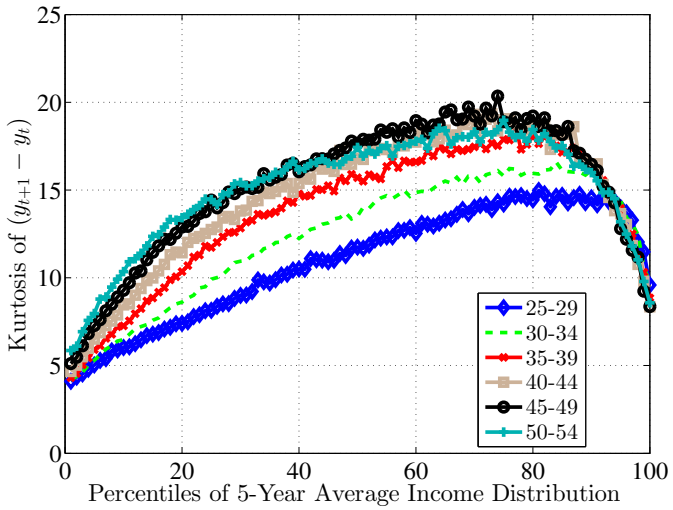
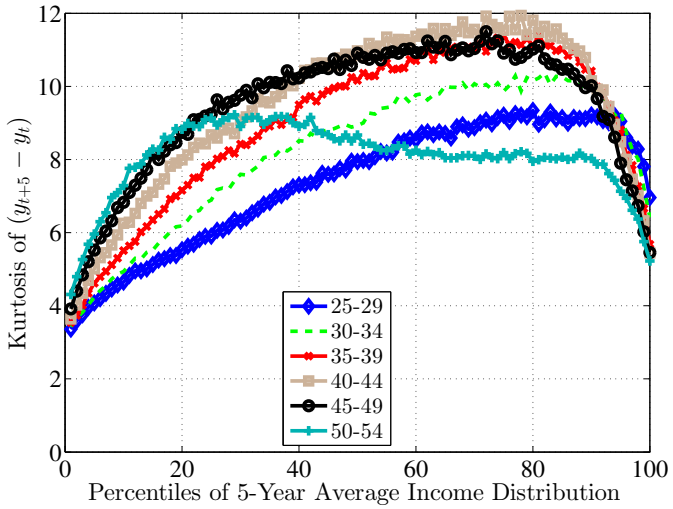


FIGURE A.21 – Kurtosis of Five-Year Earnings Change, Excluding bottom and top 1% of Earnings change





(A) Kurtosis of One-Year Earnings Change, Excluding those with Earnings lower than 5% of RE



(B) Kurtosis of Five-Year Earnings Change, Excluding those with Earnings lower than 5% of average past Earnings

FIGURE A.22 – Kurtosis, robustness