# The Nature of Countercyclical Income Risk

## Fatih Guvenen

*University of Minnesota and National Bureau of Economic Research*

## Serdar Ozkan

*Federal Reserve Board*

## Jae Song

*Social Security Administration*

We study business cycle variation in individual earnings risk using a confidential and very large data set from the US Social Security Administration. Contrary to past research, we find that the variance of idiosyncratic shocks is not countercyclical. Instead, it is the left-skewness of shocks that is strongly countercyclical: during recessions, large upward earnings movements become less likely, whereas large drops in earnings become more likely. Second, we find that the fortunes during recessions are predictable by observable characteristics before the recession. Finally, the cyclicality of earnings risk is dramatically different for the top 1 percent compared with the rest of the population.

## I.  Introduction

How does labor earnings risk vary over the business cycle? The answer to this question turns out to be central for understanding important phe-

nomena in macroeconomics and finance. For example, some research-
ers have argued that a number of puzzling observations about asset prices
can be easily understood in a standard incomplete markets model, as long
as income shocks have countercyclical variances (e.g., Constantinides and
Duffie 1996; Krusell and Smith 1997; Storesletten, Telmer, and Yaron 2007).
Other researchers have argued that it is the countercyclical left-skewness
of shocks that is critical instead (Mankiw 1986; Brav, Constantinides, and
Geczy 2002; Kocherlakota and Pistaferri 2009). Recent papers that study
the Great Recession of 2007–9 now routinely use countercyclical risk as
one of the key drivers in business cycle models (see, e.g., Krebs 2007; Ed-
mond and Veldkamp 2009; Chen, Michaux, and Roussanov 2011; Braun
and Nakajima 2012).[1]

What is common to all of these theoretical and quantitative investi-
gations is that they need to rely on empirical studies to first establish the
basic facts regarding the cyclical nature of income risk. Unfortunately,
apart from a few important exceptions discussed below, there is little em-
pirical work on this question, largely because of data limitations. Against
this backdrop, the main contribution of this paper is to exploit a unique,
confidential, and large data set in order to shed new light on the precise
nature of business cycle variation in labor income risk. Our main panel
data set is a representative 10 percent sample of all US working-age males
from 1978 to 2011. This data set has three important advantages. First,
earnings records are uncapped (no top coding), allowing us to study in-
dividuals with very high earnings. Second, the substantial sample size al-
lows us to employ flexible nonparametric methods and still obtain ex-
tremely precise estimates. Third, thanks to their records-based nature,
the data contain little measurement error, which is a common problem
with survey-based microdata sets. One drawback is possible underreport-
ing (e.g., cash earnings), which can be a concern at the lower end of the
earnings distribution.

More specifically, this paper asks two questions. First, how does the dis-
tribution of purely idiosyncratic earnings shocks (i.e., conditional on key
observable characteristics) change over the business cycle? Second, are

    [1] Following a seminal paper by Bloom (2009), a related strand of the macro literature
models "uncertainty shocks" typically in the form of countercyclical productivity risk; see,
among many others, Bloom et al. (2011), Fernandez-Villaverde et al. (2011), and Arellano,
Bai, and Kehoe (2012).

there any observable (potentially time-varying) characteristics of a worker that can help us predict his fortunes during a business cycle episode? To answer these two questions, we decompose earnings growth over the business cycle into a component that can be predicted on the basis of observable characteristics prior to the episode (i.e., a factor structure) and a separate "residual" component that represents purely idiosyncratic shocks that hit individuals who are ex ante identical. The first one represents the "between-group" or "systematic" component of business cycle risk, whereas the second can be thought of as the "within-group" or "idiosyncratic" component.

Our main findings are as follows. First, contrary to past research, we find that idiosyncratic shock variances are not countercyclical. However, uncertainty does have a significant countercyclical component, but it comes from the left-skewness increasing during recessions. The two scenarios—countercyclical variance versus left-skewness—are shown in figure 1. Thus, during recessions, the upper end of the earnings growth distribution collapses—large upward earnings movements become less likely—whereas the bottom end expands—large downward movements become more likely. At the same time, the center of the shock distribution (i.e., the median) remains stable and moves little compared with either tail, causing the countercyclicality of left-skewness. Therefore, relative to the earlier literature that argued for increasing variance—which results in some individuals receiving larger positive shocks during recessions—our results are more pessimistic: uncertainty increases in recessions without an increasing chance of upward movements (fig. 10 below).

Next, we turn to the systematic component (factor structure) of business cycle risk. We find large and robust differences between groups of individuals who enter a recession with different levels of average earnings. For example, when we rank prime-age male workers on the basis of their 2002–6 average earnings, those in the 10th percentile of this distribution experienced a fall in their earnings during the Great Recession (2007–10) that was about 18 percent worse than that experienced by those who ranked in the 90th percentile.[2] In fact, average earnings change during this recession was almost a linear (upward-sloping) function of prerecession average earnings all the way up to the 95th percen-

---

[2] The fact that the initial ranking of individuals is based on past earnings raises the issue of mean reversion going forward. But note that, as we discuss in greater detail in Sec. VI, the effect of mean reversion would work in the opposite direction of the described finding: it would imply higher future growth for those with lower past earnings. Thus, accounting for mean reversion would reveal an even stronger factor structure than these numbers suggest.
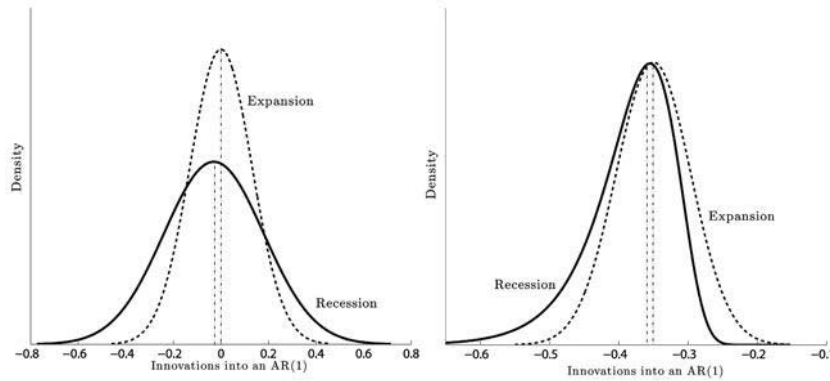
FIG. 1.—Countercyclical variance or countercyclical left-skewness? *Left*, countercyclical variance. *Right*, countercyclical left-skewness.

tile (fig. 13 below). Interestingly, this good fortune of high-income workers did not extend to the very top: those in the top 1 percent, on the basis of their 2002–6 average earnings, experienced an average loss that was 21 percent worse than that of workers in the 90th percentile. Although these magnitudes are largest for the Great Recession, the same general patterns emerged in the other recessions too. For example, the 1980–83 double-dip recession is quite similar to the Great Recession for all but the top five percentiles. But the large earnings loss for the top 1 percent was not observed during that recession at all. In fact, this phenomenon appears to be more recent: the worst episode for the top 1 percent was the otherwise mild 2000–2002 recession, when their average earnings loss exceeded that of those in the 90th percentile by almost 30 percent.

Our results on the business cycle behavior of top incomes complement and extend the findings in Parker and Vissing-Jørgensen (2010). In particular, that paper used repeated cross sections to construct synthetic groups of individuals on the basis of their earnings level. They then documented the strong cyclicality of high earnings groups over the business cycle. With panel data, we are able to track the same individuals over time, which allows us to control for compositional change and measure how persistent the effects of such fluctuations are. Our results confirm their finding that the top earners have extremely cyclical incomes and further reveal the high persistence of these fluctuations. For example, individuals who were in the top 0.1 percent as of 1999 experienced a 5-year average earnings loss between 2000 and 2005 that exceeded 50 log points! Similarly large persistent losses are found for the top income earners during the 5-year period covering the Great Recession (2004–9) as well as the 1989–94 period.

The analysis in this paper is deliberately nonparametric, made possible by the large sample size. The substantial nonlinearities revealed by this analysis justify this approach because a more parametric approach could easily miss or obscure these empirical patterns. An added benefit is that our approach allows us to present our main findings in the form of figures and easy-to-interpret statistics, which makes the results transparent. Nevertheless, a parametric specification is indispensable for calibrating economic models. To provide useful input into those studies, in Section VII, we estimate a simple parametric model of earnings dynamics that allows for mean-reverting shocks and cyclical variation in both the variance and the skewness.

*Related literature.*—The cyclical patterns of idiosyncratic labor earnings risk have received attention from both macro and financial economists. In an infinite-horizon model with permanent shocks, Constantinides and Duffie (1996) showed that one can generate a high equity premium if idiosyncratic shocks have countercyclical variance. Storesletten, Telmer, and Yaron (2004) used a clever empirical identification scheme to estimate the cyclicality of shock variances.[3] Using the Panel Study of Income Dynamics (PSID), they estimated the variance of AR(1) innovations to be three times higher during recessions. Probably because of the small sample size, they did not, however, investigate the cyclicality of the skewness of shocks, nor did they allow for a factor structure as we do here. Moreover, note that the question of interest is "the cyclical *changes* in the *dispersion* of earnings *growth* rates," which involves triple differencing. Answering such a question without a very large and clean data set is extremely challenging. Our findings are more consistent with those of Mankiw (1986), who showed that one can resolve the equity premium puzzle if idiosyncratic shocks have countercyclical left-skewness—as found in the current paper. In a related context, Brav et al. (2002) found that accounting for the countercyclical skewness of individual consumption growth helps generate a high equity premium with a low risk aversion parameter. Finally, the paper by Schulhofer-Wohl (2011) is an important precursor to our paper that uses Social Security data (with capped earnings) and analyzes the cyclicality of labor income. However, he does not examine the business cycle variation in idiosyncratic earnings risk.

This paper is also related to some recent work that emphasizes the effects of job displacement risk on the costs of business cycles.[4] In par-

---

[3] If shocks are persistent and countercyclical, cohorts that have lived through more recessions should have a larger cross-sectional dispersion of earnings at the same age than those that have not.

[4] Since the influential work of Jacobson, LaLonde, and Sullivan (1993), a large empirical literature has documented large and persistent costs of cyclical job displacement; see Farber (2005) for a review of available evidence and Von Wachter, Song, and Manchester (2009) for recent estimates using administrative data.

ticular, Krebs (2007) has argued persuasively that higher job displacement risk in recessions gives rise to countercyclical left-skewness of earnings shocks, generating costs of business cycles that far exceed earlier calculations by Lucas (1987, 2003) and others. Our findings complement this work in two ways. First, we directly measure the overall cyclicality of earnings changes and document that left-skewness is indeed strongly countercyclical. Second, our results show that this outcome is due not only to increased downside risk during recessions but also equally to the compression of the upper half of the earnings growth distribution. Therefore, the effects of recessions are not confined to a relatively small subset of the population that faces job displacement risk but are pervasive across the population.

## II.    The Data

We employ a unique, confidential, and large panel data set on earnings histories from the US Social Security Administration records. For our baseline analysis, we draw a 10 percent random sample of US males—covering 1978–2011—directly from the Master Earnings File (MEF) of Social Security records.[5]

### A.    The Master Earnings File

The MEF is the main source of earnings data for the Social Security Administration and grows every year with the addition of new earnings information received directly from employers (Form W-2 for wage and salary workers).[6] The MEF includes data for every individual in the United States who has a Social Security number. The data set contains basic demographic characteristics, such as date of birth, sex, race, type of work (farm or nonfarm, employment or self-employment), self-employment taxable earnings, and several other variables. Earnings data are uncapped (no top coding) and include wages and salaries, bonuses, and exercised stock options as reported on the W-2 form (box 1). For more information, see Panis et al. (2000) and Olsen and Hudson (2009). Finally, all nominal

---

[5] Our focus on males is motivated by the fact that this group had a relatively stable employment rate and labor supply during this period. In contrast, female labor participation increased substantially during this period. Because our data set contains only labor earnings but no hours information, including women in the analysis would have introduced an important confounding factor, which we wished to avoid.

[6] Although the MEF also contains earnings information for self-employed individuals, these data are top coded at the taxable limit until 1994. Because of this, we do not use these data in this paper. In an earlier version, we conducted all the analysis using total labor earnings (and included self-employed individuals) and found no difference in our substantive conclusions.

variables were converted into real ones using the personal consumption expenditure deflator with 2005 taken as the base year.

## B. Creating the 10 Percent Sample

To construct a nationally representative panel of males, we proceed as follows. For 1978, a sample of 10 percent of US males are selected on the basis of a fixed subset of digits of (a transformation of) the Social Security number (SSN). Because these digits of the SSN are randomly assigned, this procedure easily allows randomization. For each subsequent year, new individuals are added to account for the newly issued SSNs; those individuals who are deceased are removed from that year forward. This process yields a representative 10 percent sample of US males every year.

For a statistic computed using data for (not necessarily consecutive) years $t_1, t_2, \ldots, t_n$, an individual observation is included if the following three conditions are satisfied for all these years: the individual (i) is between the ages of 25 and 60, (ii) has annual wage/salary earnings that exceed a time-varying minimum threshold, and (iii) is not self-employed (i.e., has self-employment earnings less than the same minimum threshold). This minimum, denoted $Y_{\min,t}$, is equal to one-half of the legal minimum wage times 520 hours (13 weeks at 40 hours per week), which amounts to annual earnings of approximately \$1,300 in 2005. This condition allows us to focus on workers with a reasonably strong labor market attachment (and avoids issues with taking the logarithm of small numbers). It also makes our results more comparable to the income dynamics literature, where this condition is standard (see, among others, Abowd and Card [1989], Meghir and Pistaferri [2004], and Storesletten et al. [2004], as well as Juhn, Murphy, and Pierce [1993] and Autor, Katz, and Kearney [2008] on wage inequality). Finally, the MEF contains a small number of extremely high earnings observations each year. To avoid potential problems with outliers, we cap (Winsorize) observations above the 99.999th percentile.

Figure 2 displays the number of individuals who satisfy these selection criteria, as well as the total number of individuals in each year. The sample starts with about 3.7 million individuals in 1978 and grows to about 5.4 million individuals by the mid-2000s. Notice that the number of individuals in the sample does not follow population growth one for one (grey line marked with diamonds) because inclusion in the base sample also requires participating in the labor market in a given year (hence the slowdown in sample growth in the 2000s and the fall during the Great Recession).

Online Appendix A reports a broad set of summary statistics for our sample. The lowest earnings that qualify a male worker in the top 10 per-
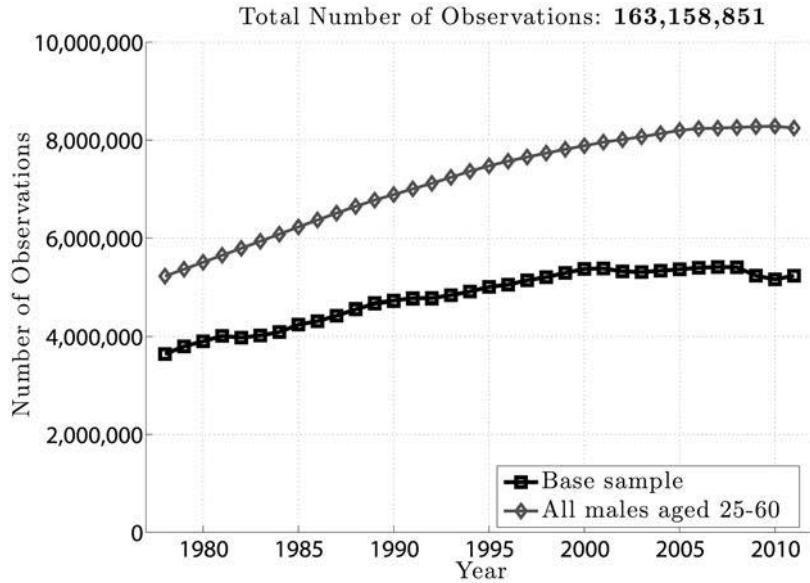
FIG. 2.—Number of observations by year, 1978–2011

cent (e.g., above the 90th percentile) have been steady at approximately $98,000 (in 2005 dollars) since year 2000. In 2011, a worker must be making more than $302,500 to be in the top 1 percent. This threshold was highest in 2007, when it reached $318,000.[7]

### C.   Recessionary versus Expansionary Episodes

The start date of a recession is determined as follows. If the National Bureau of Economic Research (NBER) peak of the previous expansion takes place in the first half of a given year, that year is classified as the first year of the new recession. If the peak is in the second half, the recession starts in the subsequent year.[8] The ending date of a recession is a bit more open to interpretation for our purposes because the NBER troughs are often not followed by a rapid fall in the unemployment rate and a rise in individual wages. This can be seen in figure 3. For example, whereas the NBER announced the start date of the expansion as March 1991, the unemployment rate peaked in the summer of 1992. Similarly, while the NBER trough was November 2001, the unemployment rate re-

---

[7] Further, Sec. B in App. A contains a comparison of inequality trends revealed by the base sample to those found in the Current Population Survey (CPS) data.

[8] Two recessions start in the first quarter (1980 and 2001) and one starts in the fourth quarter (2007), so the classification of these is clear. Only one recession starts in the third quarter of 1990, and we shift the starting date to 1991 as per the rule described.
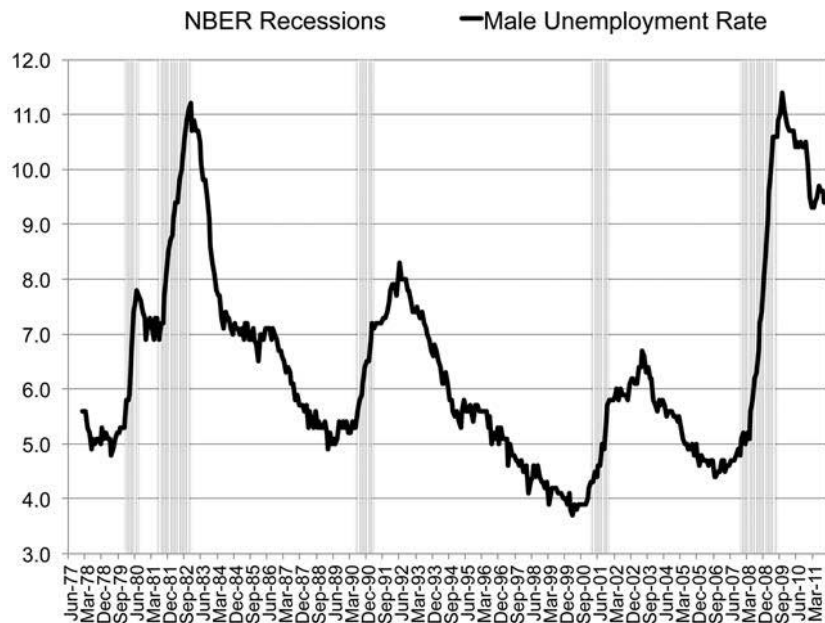
FIG. 3.—US male unemployment rate, 1978–2011

mained high until mid-2003. With these considerations in mind, we settled on the following dates for the last three recessions: 1991–92, 2001–2, and 2008–10. We treat the 1980–83 period as a single recession, given the extremely short duration of the intervening expansion, the anemic growth it brought, and the lack of a significant fall in the unemployment rate. On the basis of this classification, there are three expansions and four recessions during our sample period.[9]

## III.   Earnings Risk over the Business Cycle: First Look

Before delving into the full-blown panel data analysis in the next section, we begin by providing a bird's-eye view of the business cycle patterns in earnings risk. Specifically, we exploit the panel dimension of the MEF data set to document how the dispersion and skewness of the earnings growth distribution vary over the business cycle.[10]

---

[9] As a complementary approach, in Sec. G of online App. B, we study business cycle variation by analyzing the comovement of the earnings growth distribution with cyclical variables, such as the male unemployment rate, GDP per capita, and Standard & Poors 500 returns.

[10] In the text, we alternatively refer to earnings growth as "earnings change" and, with a mild abuse of language, as "earnings shocks" to prevent monotonicity.

It will be useful to distinguish between earnings growth over short and long horizons. To this end, we examine 1- and 5-year log earnings growth rates (denoted by $\tilde{y}_t - \tilde{y}_{t-k}$ for $k = 1, 5$) and think of these as roughly corresponding to "transitory" and "persistent" earnings shocks. A more rigorous justification for this interpretation will be provided in the next section.

The left panel of figure 4 plots the evolution of the log differential between the 90th and 50th percentiles of the $\tilde{y}_t - \tilde{y}_{t-1}$ distribution (hereafter L90–50) as well as the log differential between the 50th and 10th percentiles (L50–10). The first important observation is that the top and bottom ends of the shock distributions clearly move in opposite directions over the business cycle. In particular, L50–10 rises strongly during recessions, implying that there is an increased chance of larger downward movements during recessions. In contrast, the top end (L90–50) dips consistently in every recession, implying a smaller chance of upward movements during recessions. In other words, relative to the median growth rate, the top end compresses, whereas the bottom end expands during recessions. Similarly, the right panel of figure 4 plots the corresponding graph for persistent (5-year) shocks. The comovement of L90–50 and L50–10 is clearly seen here, even more strongly than in the transitory shocks (the correlation of the two series is −.67).

A couple of remarks are in order. First, the fact that L90–50 and L50–10 move in opposite directions implies that L90–10 (which is a measure of overall dispersion of shocks) changes little over the business cycle because the fall in L90–50 partially cancels out the rise in L50–10. An alternative measure of shock dispersion—the standard deviation—is plotted in figure 5 for both persistent and transitory shocks, which shows that dispersion does not increase much during recessions (notice the small cyclical variation on the $y$-axis). Perhaps the only exception is the 2001–2 recession, during which time the transitory shock variance increases. In
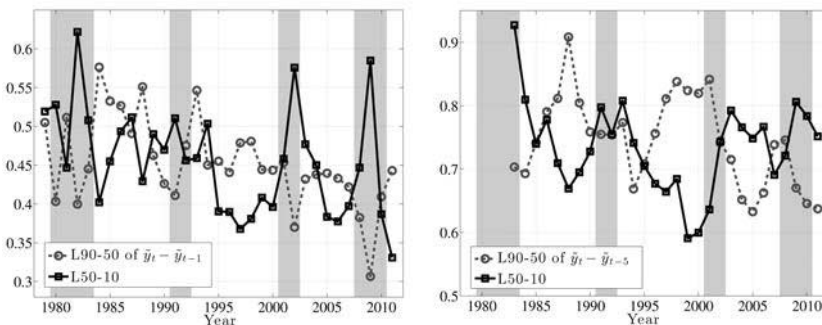


FIG. 4.—Top and bottom ends of the earnings growth distribution. *Left*, transitory. *Right*, persistent.
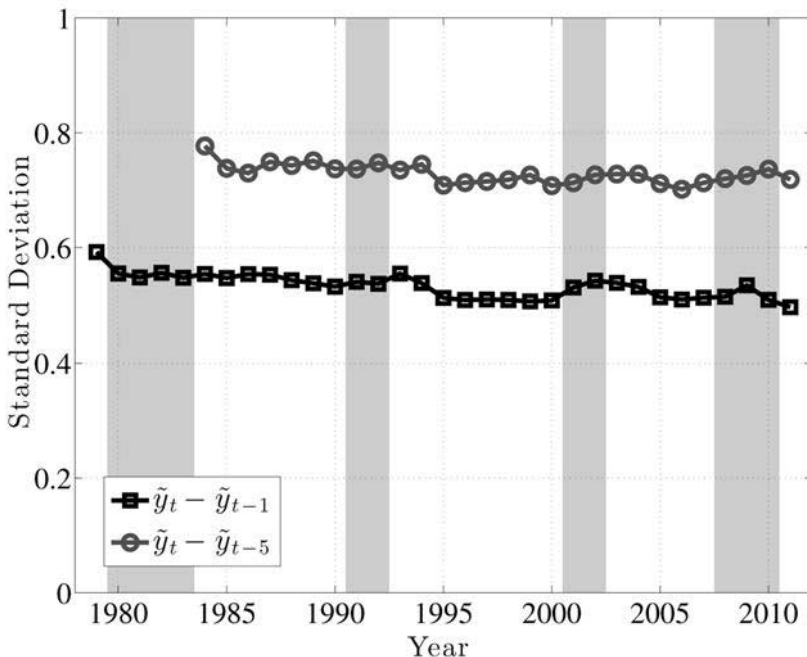
Fig. 5.—Standard deviation of transitory and persistent earnings growth

the coming sections, this point will be examined further and will be made more rigorously. This observation will provide one of the key conclusions of this paper, given how clearly it contradicts the commonly held belief that idiosyncratic earnings shock variances are strongly countercyclical.

Second, the finding described above—that the top end of the shock distribution compresses during recessions while at the same time the bottom end expands—suggests that one important cyclical change could be found in the skewness of shocks. Indeed, as seen in figure 6, both the 1- and 5-year earnings growth distributions become more left-skewed (negative skewness increases) during recessions and the magnitude of change is large.[11]

*Trends in volatility: a brief digression.*—Looking at the left panel of figure 4, notice that L90–50 displays a clear downward trend during this time period. A fitted linear trend reveals a drop of 11 log points from 1979 to 2011. The interpretation is that the likelihood of large upward movements has become smaller during this period. We see a similar, if a bit less pronounced, trend in L50–10, which indicates that the likelihood

---

[11] To see whether these results are robust to mean reversion in $\tilde{y}_t$, in Sec. A of App. B, we plot the analogues of figs. 4 and 5 using quasi-differencing: $\tilde{y}_t - \rho\tilde{y}_{t-1}$ for $\rho = 0.80, 0.90$, and 0.95, which show that the results reported here are robust to such mean reversion.
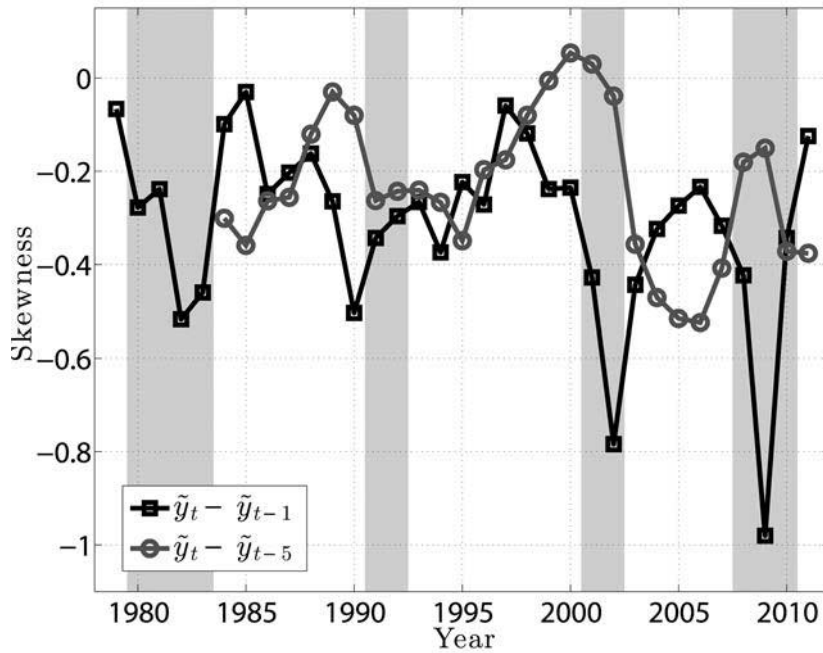
FIG. 6.—Skewness of transitory and persistent earnings growth

of large falls has also become somewhat smaller. Overall, both the L90–10 and the standard deviation of earnings growth (fig. 5) display a clear downward trend. This finding is in contrast to the conventional wisdom in the literature that earnings shock variances have generally risen since the 1980s (Moffitt and Gottschalk 1995). However, it is consistent with a number of recent papers that use administrative data (e.g., Sabelhaus and Song 2010). In this paper, we will not dwell much on this trend in order to keep the analysis focused on the cyclical changes in earnings risk.

## IV.   Panel Analysis

The analysis so far provided a general look at how earnings shocks vary over the business cycle. However, one can imagine that the properties of earnings shocks vary systematically with individual characteristics and heterogeneity: for example, young and old workers can face earnings shock distributions different from those of prime-age workers with more stable jobs. Similarly, workers at different parts of the earnings distribution could experience different types of earnings risks. The large sample size allows us to account for such variation without making strong parametric assumptions.

## A. A Framework for Empirical Analysis

Let $\tilde{y}_t^i$ denote individual $i$'s log labor earnings in year $t$ and let $\mathbf{V}_{t-1}^i$ denote a vector of (possibly time-varying) individual characteristics that will be used to group individuals as of period $t-1$. For each business cycle episode, we will examine how earnings growth varies between these groups defined by $\mathbf{V}_{t-1}^i$. We shall refer to this first type of variation alternatively as a factor structure or systematic risk. Of course, even individuals within these finely defined groups will likely experience different earnings growth rates during recessions and expansions, reflecting within-group or idiosyncratic earnings shocks. We will also quantify the cyclical nature of such shocks.

## B. Grouping Individuals into $\mathbf{V}_{t-1}^i$

Let $t$ denote the generic time period that marks the beginning of a business cycle episode. We now describe how we group individuals on the basis of their characteristics at time $t-1$. Each individual is identified by three characteristics that can be used to form groups. Not every characteristic will be used in the formation of groups in every experiment.[12]

1. *Age.*—Individuals are divided into seven age groups. The first six groups are 5 years wide (25–29, 30–34, . . . , 50–54) and the last one covers 6 years: 55–60.

2. *Pre-episode average earnings.*—A second dimension individuals differ along is their average earnings. For a given year $t$, we consider all individuals who were in the base sample (i) in year $t-1$ and (ii) in at least two more years between $t-5$ and $t-2$. For example, an individual who is 23 years old in $t-5$ (and hence is not in the base sample that year) will be included in the final sample for year $t$ if he has earnings exceeding $Y_{\min}$ in every year between $t-3$ and $t-1$.

We are interested in average earnings to determine how a worker ranks relative to his peers. But even within the narrow age groups defined above, age variation can skew the rankings in favor of older workers. For example, between ages 25 and 29, average earnings grows by 42 percent in our sample, and between 30 and 34, it grows by 20 percent. So, unless this life cycle component is accounted for, a 29-year-old worker in the first age group would appear in a higher earnings percentile than the same worker when he was 25. This variation would confound age and earnings differences.

To correct for this, we proceed as follows. First, using all earnings observations from our base sample from 1978 to 2011, we run a pooled

---

[12] Education is a potentially relevant worker characteristic that is not recorded in the MEF.

regression of $\tilde{y}_{t,h}^i$ on age $(h)$ and cohort dummies without a constant to characterize the age profile of log earnings. We then scale the age dummies (denoted by $d_h$) so as to match the average log earnings of 25-year-old individuals used in the regression. Using these age dummies, we compute the average earnings between years $t-5$ and $t-1$ for the average individual of age $h$ in year $t$. Then for a given individual $i$ of age $h$ in year $t$, we first average his earnings from $t-5$ to $t-1$ (and set earnings below $Y_{\min,t}$ equal to the threshold) and then normalize it by the population average computed using the age dummies:

$$\overline{Y}_{t-1}^i \equiv \frac{\sum_{s=1}^{5} e^{\tilde{y}_{t-s}^i}}{\sum_{s=1}^{5} e^{d_{h-s}}}.$$

This 5-year average (normalized) earnings is denoted by $\overline{Y}_{t-1}^i$. We also define $y_t^i \equiv \tilde{y}_{t,h}^i - d_h$ as the log earnings in year $t$ net of life cycle effects, which will be used in the analysis below.

3. *Pre-episode earnings growth.*—A third characteristic is an individual's (recent) earnings growth. This could be an indicator of individuals whose careers are on the rise, as opposed to being stagnant, even after controlling for average earnings as done above. We compute $\Delta_5(y_{t-1}^i) \equiv (y_{t-1}^i - y_{t-s}^i)/(s-1)$, where $s$ is the earliest year after $t-6$ in which the individual has earnings above the threshold. In the main text, we focus on the first two characteristics and, to save space, report the results with pre-episode earnings growth in Section E of online Appendix B.

## V.   Within-Group (Idiosyncratic) Shocks

One focus of this analysis will be on simple measures of earnings shock volatility, conditional on individual characteristics. For the sake of this discussion, suppose that log earnings (net of life cycle effects) is composed of a random walk component with innovation $\eta_t^i$ plus a purely transitory term $\varepsilon_t^i$. Then, computing the within-group variance, we get

$$\text{Var}\,(y_{t+k}^i - y_t^i | \mathbf{V}_{t-1}^i) = \left[ \sum_{s=1}^{k} \text{Var}(\eta_{t+s}^i | \mathbf{V}_{t-1}^i) \right]$$
$$+ [\text{Var}\,(\varepsilon_t^i | \mathbf{V}_{t-1}^i) + \text{Var}(\varepsilon_{t+k}^i | \mathbf{V}_{t-1}^i)].$$

Two points can be observed from this formula. First, as we consider longer time differences, the variance reflects more of the permanent shocks, as seen by the addition of the $k$ innovation variances and given that there are always two variances from the transitory component. For example, computing this variance over a 5-year period that spans a recession (say 1979–84 or 1989–94) would allow us to measure how the

variance of permanent shocks changes during recessions. It will also contain transitory variances, but for two years that are not part of a recession (1979 and 1984, e.g.). Second, looking at short-term variances, say $k = 1$, yields a formula that contains only one permanent shock variance and two transitory shock variances. So, as we increase the length of the period over which the variance is computed, the statistic shifts from being informative about transitory shock variances toward more persistent variation.

In the analysis below, we consider $k = 1$ and $k = 5$. The choice of $k = 5$ is motivated by the fact that recessions last 2–3 years, so that by year $t + 5$ the unemployment rate will have declined from its peak and will, in most cases, be close to the prerecession level (in year $t$). This feature will facilitate the interpretation of our findings, as we discuss later.

## A.    A Graphical Construct

Most of the empirical analysis in this paper will be conducted using the following graphical construct: we plot the quantiles of $\overline{Y}^i_{t-1}$ for a given age group on the $x$-axis against the distribution of future earnings growth rates for that quantile on the $y$-axis: $\mathbb{F}(y^i_{t+k} - y^i_t | \overline{Y}^i_{t-1})$. The properties of these conditional distributions for each $\overline{Y}^i_{t-1}$ will be informative about the nature of within-group variation. To study these properties more closely, we will use the same construct to plot various statistics from these conditional distributions, such as various percentiles, the mean, the variance, the skewness, and so on.

Figure 7 is the first use of this graphical construct and contains a lot of information that will be referred to in the rest of this section. The top panel displays P90 (the 90th percentile), P50 (the median), and P10 of the distribution of long-run changes, $y^i_{t+5} - y^i_t$, on the $y$-axis against each percentile of $\overline{Y}^i_{t-1}$ on the $x$-axis. To compare recessions and expansions, we averaged each percentile graph separately over the four recessions (solid black lines) and three expansions (dashed grey lines) during our sample period.[13] Similarly, because these figures look quite similar across age groups, to save space here, we also averaged across age groups. (We report the complete set of figures by age group in App. B.)

[13] For 5-year changes, recession years can be defined in a number of ways, since many 5-year periods cover a given recession. We have experimented with different choices and found them to make little difference to the substantive conclusions drawn here. The reported results are for a simple definition that includes one 5-year change for each recession that starts 1 year before the recession begins. Specifically, the recession graph averages over four 5-year periods starting in $t = 1979$, 1989, 1999, and 2006 (since this is the latest possible 5-year change covering the Great Recession). Expansions average over all 5-year changes that do not coincide with a recession year, i.e., periods starting in $t = 1983$, 1984, 1993, 1994, and 2002.
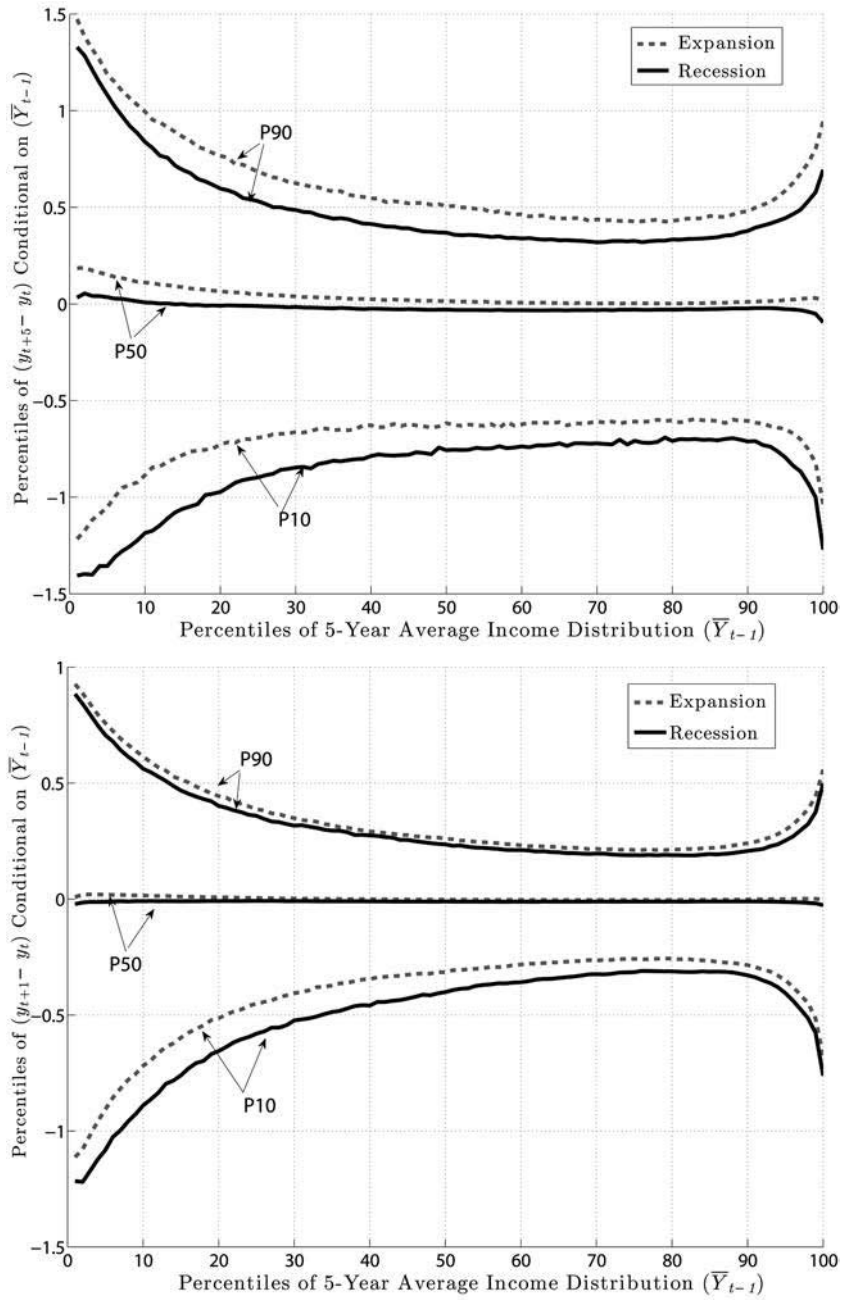
Fig. 7.—Percentiles of the earnings growth distribution: recession versus expansion. *Top*, persistent change. *Bottom*, transitory change.

First, notice the variation in these percentiles as we move to the right along the $x$-axis. Interestingly, the following pattern holds in both recessions and expansions: At any point in time, individuals with the lowest levels of past average earnings face the largest dispersion of earnings shocks $(y_{t+k}^i - y_t^i)$ looking forward. That is, L90–10 is widest for these individuals and falls in a smooth fashion moving to the right. Indeed, workers who are between the 70th and 90th percentiles of the $\overline{Y}_{t-1}^i$ distribution face the smallest dispersion of shocks looking ahead. As we continue moving to the right (into the top 10 percent), the shock distribution widens again. Notice that the P10 and P90 of the $y_{t+5}^i - y_t^i$ distribution look like the mirror image of each other relative to the median, so the variation in L90–10 as we move to the right is driven by similar variations in P90 and P10 individually.

In the bottom panel, the same graph is plotted now for $y_{t+1}^i - y_t^i$ (transitory shocks).[14] Precisely the same qualitative features are seen here, with low- and high-income individuals facing a wider dispersion of persistent shocks than those in the "safer" zones—between the 70th and 90th percentiles. Of course, the scales of both graphs are different: the overall dispersion of persistent shocks is much larger than that of transitory shocks, which is to be expected. To summarize, both graphs reveal strong and systematic variation in the dispersion of persistent and transitory earnings shocks across individuals with different past earnings levels.[15]

Now we turn to two key questions of interest. First, what happens to idiosyncratic shocks in recessions? For example, are shock variances countercyclical? And second, how does any potential change in the distribution of idiosyncratic shocks vary across earnings levels (i.e., the cross-partial derivative)? In other words, do we see the shock distribution of individuals in different earnings levels being affected differently by recessions?

## B.  Are Shock Variances Countercyclical? No

The existing literature has largely focused on the cyclicality of persistent shocks, so this is where we also start (top panel of fig. 7). First, note that both P90 and P10 shift downward by similar amounts from expansion to recession. Consequently, the L90–10 gap varies little over the business cycle, as we shall soon see. Furthermore, following the same steps as the one used to construct these graphs, one can also compute the standard deviation of $y_{t+5}^i - y_t^i$ conditional on $\overline{Y}_{t-1}^i$ during recessions and expan-

[14] For 1-year changes, recession years are those with $t$ = 1980, 1981, 1982, 1990, 1991, 2000, 2001, 2007, 2008, and 2009. The remaining years are considered as expansion years.
[15] This finding clearly contradicts one of the standard assumptions in the income dynamics literature: that the variance of earnings shocks does not depend on the current or past level of earnings. We study these features in Guvenen et al. (2013).

sions, which is plotted in the left panel of figure 8. The two graphs (for expansions and recessions) virtually overlap over the entire range of pre-episode earnings levels. For transitory shocks (right panel), there is more of a gap, but the two lines are still quite close to each other.

To make the measurement of countercyclicality more precise, the left panel of figure 9 plots the ratios (recession/expansion) of (i) the standard deviations and (ii) the L90–10s of 5-year earnings changes. Both measures are only about 2 percent higher in recessions than in expansions. For comparison, Storesletten et al. (2004) used indirect methods to estimate a standard deviation of 0.12 for innovations into a persistent AR(1) process during expansions and 0.21 for recessions. The ratio is 1.75 compared with the 1.02 we find in this paper. In fact, above the 30th percentile of the past earnings distribution, the average ratio in figure 9 is precisely 1.00.[16]

A second question that was raised above was whether recessions affect the distribution of shocks differently in different parts of the earnings distribution. It is probably evident by now that the answer is "no": as seen in figure 9, the ratios of L90–10s and standard deviations are flat. A similar question is whether countercyclicality might be present for some age groups, which may not be apparent once age groups are combined as in the left panel of figure 9. To check for this, we plot the ratio of standard deviations separately for each age group in the right panel. As seen here, the graph is nearly flat for all age groups and remains trapped within a narrow corridor between 0.95 and 1.05.

To summarize, we conclude that when it comes to the variance of persistent shocks, the main finding is one of homogeneity: the variance remains virtually flat over the business cycle for every age and earnings group.

## C.   Countercyclical Left-Skewness: A Tale of Two Tails

So, do recessions have any effect on earnings shocks? The answer is yes, which could already be anticipated from figure 7 by noting that while P90 and P10 move down together during recessions, P50 (the median of the shock distribution) remains stable and moves down by only a little. This has important implications: L90–50 gets compressed during recessions, whereas L50–10 expands. In other words, for every earnings level $\overline{Y}^i_{t-1}$, when individuals look ahead during a recession, they see a

---

[16] The same ratios can also be computed for transitory shocks. Here we see a bit more movement: the standard deviation is higher by about 4 percent (averaged across the *x*-axis) and L90–10 is higher by about 6 percent. So, to the extent that recessions involve a larger dispersion of shocks, these are to be found in short-term shocks without much long-term effect. That said, these numbers are still negligible compared with the values typically used in the literature.
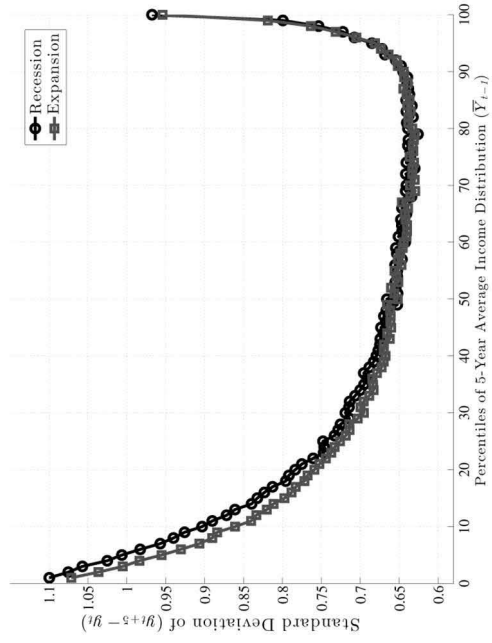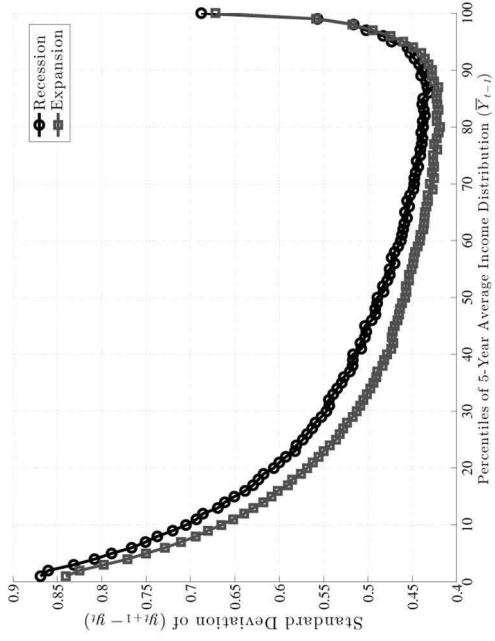
FIG. 8.—Dispersion of the earnings growth distribution: recession versus expansion. *Left*, persistent. *Right*, transitory
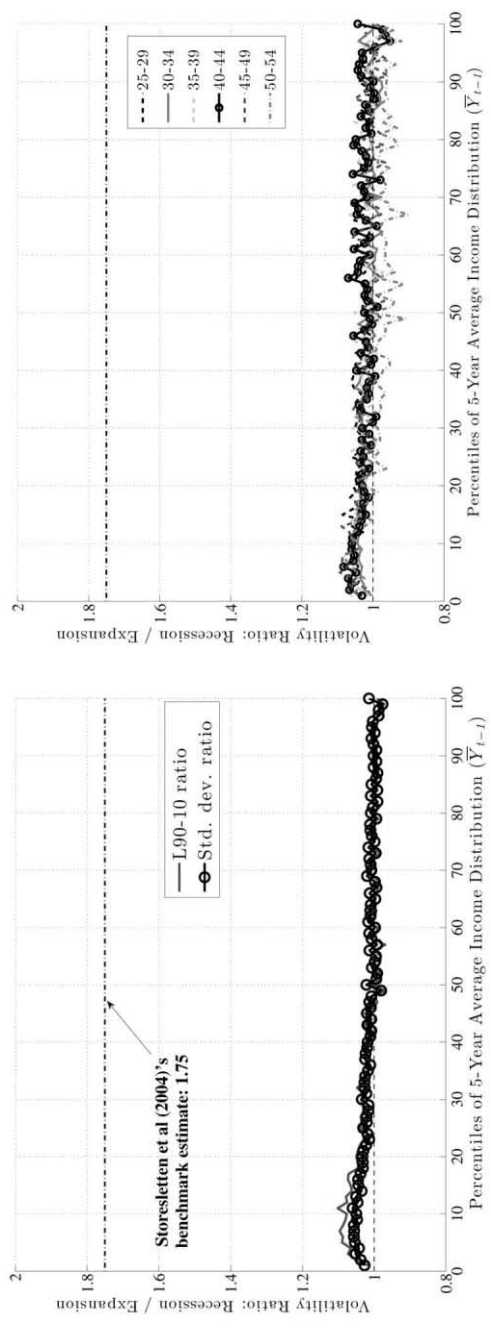
FIG. 9.—Ratio of (5-year) volatility measures: recession/expansion. *Left*, all ages combined. *Right*, by age groups

smaller chance of upward movements (relative to an expansion) and a higher chance of large downward movements.

   This result is not specific to using P90 or P10 but is pervasive across the distribution of future earnings growth rates. This can be seen in figure 10, which plots the change in selected percentiles above (and including) the median from an expansion to a recession (top panel). The bottom panel shows selected percentiles below the median. Starting from the top and focusing on the middle part of the *x*-axis, we see that P99 falls by about 30 log points from an expansion to a recession, whereas P95 falls by 20, P90 falls by 15, P75 falls by 6, and P50 falls by 5 log points, respectively. As a result, the entire upper half of the shock distribution gets squeezed toward the median. In other words, the half of the population who experience earnings change above the median now experience ever smaller upward moves during recessions. Turning to the bottom panel, we see the opposite pattern: P50 falls by 5 log points, whereas P25 falls by 7 and P10 falls by 15 log points, respectively. Consequently, the bottom half of the shock distribution now expands, with "bad luck" meaning even "worse luck" during recessions.
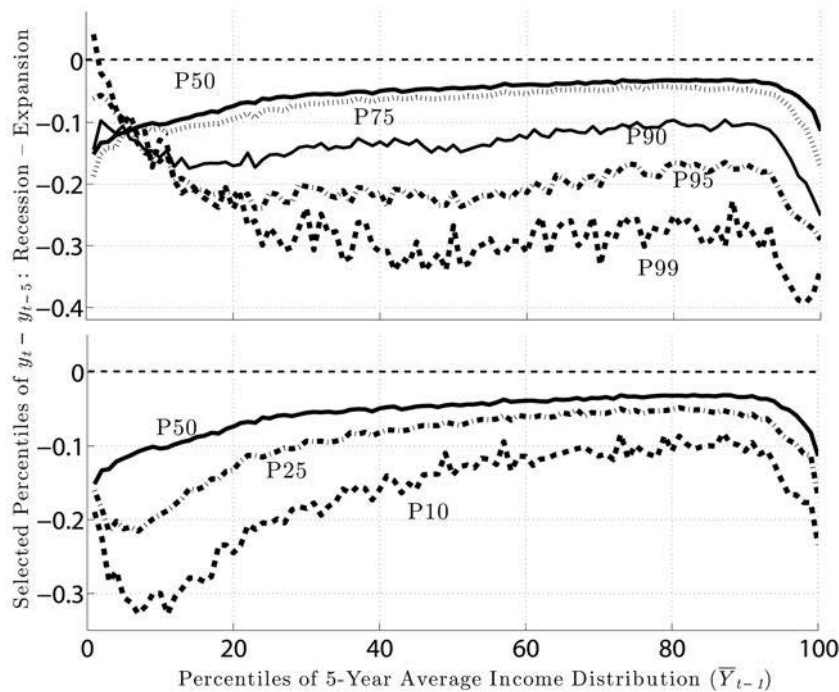
FIG. 10.—Cyclical change in the percentiles of 5-year earnings growth distribution

From this analysis, a couple of conclusions can be drawn. First, idiosyncratic risk is countercyclical. However, this does not happen by a widening of the entire distribution (e.g., variance rising), but rather a shift toward a more left-skewed shock distribution. Another useful way to document this latter point is to compute some summary statistics for skewness. With higher-order moments, one has to be careful about extreme observations. These are not likely to be outliers as with survey data, but even if they are genuine observations, we may want to be careful that a few observations do not affect the overall skewness measure. For this purpose, our preferred statistic is "Kelley's measure" of skewness, which relies on the quantiles of the distribution and is robust to outliers (left panel of fig. 11). It is computed as the relative difference between the upper and lower tail inequalities: (L90–50 − L50–10)/L90–10. A negative number indicates that the lower tail is wider than the upper tail, and vice versa for a positive number. For completeness, we also plot the third central moment in the right panel. The substantive conclusions we draw from both statistics are essentially the same.

Inspecting the graphs in the left panel of figure 11, first, notice that individuals in higher earnings percentiles face persistent shocks that are more negatively skewed than those faced by individuals ranked lower, consistent with the idea that the higher an individual's earnings are, the more room he has to fall. Second, and more important, this negative skewness increases during recessions for both transitory and persistent shocks. Another advantage of Kelley's skewness measure is that its value has a straightforward interpretation. For example, for individuals at the median of the $\overline{Y}_{t-1}$ distribution, Kelley's measure for persistent shocks averages −0.125 during expansions. This means that the dispersion of shocks above P50 accounts for 44 percent of overall L90–10 dispersion. Similarly, dispersion below P50 accounts for the remaining 56 percent (hence [44 percent − 56 percent]/100 percent = −0.12) of L90–10. In
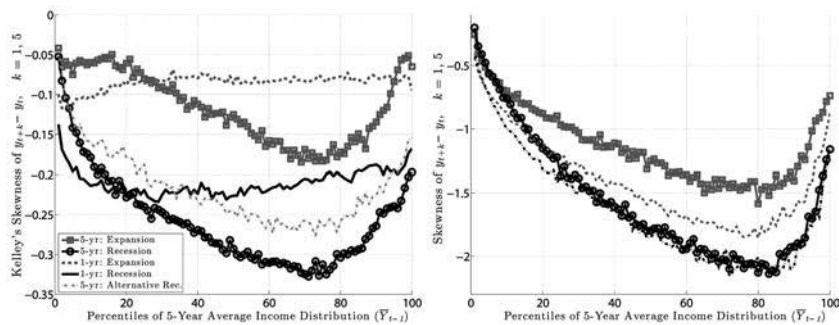


Fig. 11.—Skewness of the earnings growth distribution: recession versus expansion. *Left*, Kelley's skewness. *Right*, third central moment.
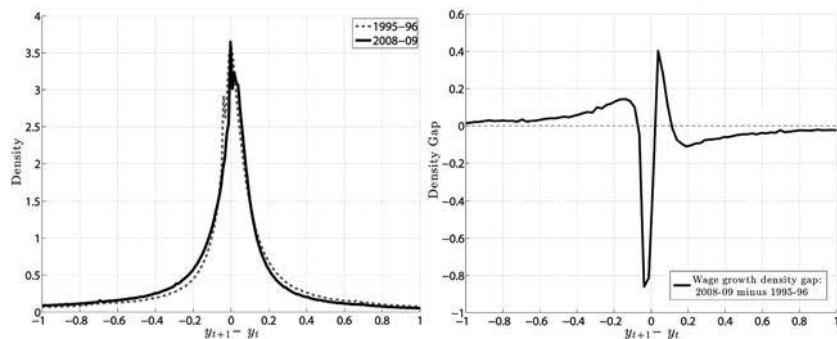
FIG. 12.—Histogram of $y_{t+1} - y_t$: US data, 1995–96 versus 2008–9. *Left*, densities: recession versus expansion. *Right*, gap: recession minus expansion.

recessions, however, this figure falls to $-0.30$, implying that L90–50 accounts for 35 percent of L90–10 and the remaining 65 percent is due to L50–10. This is a substantial shift in the shape of the persistent shock distribution over the business cycle. The change in the skewness of transitory shocks is similar, if somewhat less pronounced. It goes from $-0.08$ down to $-0.226$ at the median (the share of L90–50 going down from 46 percent of L90–10 down to 39 percent). As seen in figure 11, the rise in left-skewness takes place with similar magnitudes across the earnings distribution (with the exception of very low-income individuals).

To understand how different this conclusion is from a simple countercyclical variance formulation, recall figure 1, which plots the densities of two normal random variables: one with zero mean and a standard deviation of 0.13 (expansion) and a second one with a mean of $-0.03$ and a standard deviation of 0.21 (recession; both numbers from Storesletten et al. [2004]). As seen here, the substantial increase in variance and small fall in the mean imply that many individuals will receive larger positive shocks in recessions than in expansions under this formulation. For comparison, the left panel of figure 12 plots the empirical densities of earnings growth from the US data, comparing the 1995–96 period to the worst of the Great Recession (2008–9). To highlight how the density changes, the right panel plots the difference between the two densities. As seen here, the probability mass on the right side shifts from large positive shocks to more modest ones; on the left side, it shifts from small negative shocks to even larger negative ones.[17]

---

[17] A natural question is whether there are cyclical changes in moments beyond the third (skewness), e.g., in the fourth moment—the kurtosis. Although the answer is yes—the kurtosis is lower in recessions compared with expansions—the differences are quite modest. We omit those results for brevity.

*What role does unemployment play?*—Is the countercyclicality of left-skewness all due to unemployment, which increases during recessions? Although unemployment risk is unlikely to explain the change in skewness coming from the compression of the upper half of the shock distribution, this is still a valid question for the expansion of the bottom half. In Section F of Appendix B, we investigate this question using data from both the MEF and the CPS. We conclude that, while the cyclical changes in unemployment are clearly nonnegligible, they are not large enough to generate the bulk of the expansion of the bottom half of the shock distribution.

## VI. Between-Group (Systematic) Business Cycle Risk

We now turn to the between-group, or systematic, component of earnings risk. The goal here is to understand the extent to which earnings growth during a business cycle episode can be predicted by observable characteristics prior to the episode. A natural way to measure the between-group variation is to define the mean log earnings change conditional on characteristics as of $t-1$:

$$f_1(\mathbf{V}_{t-1}^i) \equiv \mathbb{E}(y_{t+k}^i - y_t^i | \mathbf{V}_{t-1}^i). \tag{1}$$

The shape of $f_1$ tells us how individuals who differ in characteristics $\mathbf{V}_{t-1}^i$ before a business cycle episode fare during the episode.[18] However, one drawback of this measure is that it can be computed using only individuals with positive earnings in years $t$ and $t + k$. While, on average, the number of individuals who are excluded is small, the number varies both over the business cycle and across groups $\mathbf{V}_{t-1}^i$, which could be problematic. This concern leads us to our second measure of systematic risk:

$$f_2(\mathbf{V}_{t-1}^i) \equiv \log \mathbb{E}(Y_{t+k}^i | \mathbf{V}_{t-1}^i) - \log \mathbb{E}(Y_t^i | \mathbf{V}_{t-1}^i), \tag{2}$$

where $Y_t^i \equiv \exp(y_t^i)$. This measure now includes both the intensive margin and the extensive margin of earnings changes between two periods. For the empirical analysis in this section, $f_2$ will be our preferred measure. In Section D of Appendix B, we present the analogous results obtained with $f_1$ and discuss the (small) differences between the two measures.

### A. Caution: Mean Reversion Ahead

The interpretations of $f_1$ and $f_2$ require some care when $y_t^i$ has a mean-reverting component, which seems plausible. The reason is that when $y_t^i$

---

[18] To be precise, $f_1$ should have a time subscript since we will allow it to vary over time. However, to keep the notation clean, we will suppress the subscript in this paper.

is a mean-reverting process and we condition on past earnings (such as $\overline{Y}_{t-1}^{i}$), these measures will be a decreasing function of $\overline{Y}_{t-1}^{i}$ in the absence of any factor structure. For the sake of this discussion, let us assume that $y_{t+1}^{i} = \rho y_{t}^{i} + \eta_{t+1}^{i}$. The $k$th difference of $y_{t}^{i}$ can be written as

$$y_{t+k}^{i} - y_{t}^{i} = (\eta_{t+k}^{i} + \rho\eta_{t+k-1}^{i} + \cdots + \rho^{k-1}\eta_{t+1}^{i}) + (\rho^{k} - 1)y_{t}^{i}.$$

When we take the expectation of both sides with respect to $\overline{Y}_{t-1}^{i}$, the terms in parentheses vanish, since future innovations ($\eta_{t+k}^{i}$'s) have zero mean and are independent of past earnings. Now consider $f_1$ (which is analytically more tractable than $f_2$, but the same point applies to $f_2$ as well):

$$f_1(\overline{Y}_{t-1}^{i}) = \mathbb{E}(y_{t+k}^{i} - y_{t}^{i}|\overline{Y}_{t-1}^{i}) = \mathbb{E}((\rho^{k} - 1)y_{t}^{i}|\overline{Y}_{t-1}^{i})$$

$$\Rightarrow \frac{\partial f_1(\overline{Y}_{t-1}^{i})}{\partial \overline{Y}_{t-1}^{i}} = (\rho^{k} - 1) \times \frac{\partial}{\partial \overline{Y}_{t-1}^{i}} \mathbb{E}(y_{t}^{i}|\overline{Y}_{t-1}^{i}) < 0,$$

where the last inequality follows straightforwardly from the fact that, on average, $y_{t}^{i}$ is higher when $\overline{Y}_{t-1}^{i}$ is higher and $\rho^{k} < 1$. Therefore, when $y_{t}^{i}$ contains a mean-reverting component, between-group differences will be downward sloping as a function of past average earnings. Hence, if we estimate $f_1$ or $f_2$ to be upward sloping (overcoming this potential downward bias), this would be a strong indication of a factor structure.

### B.    Variation between $\overline{Y}_{t-1}^{i}$ Groups

We estimate $f_2(\overline{Y}_{t-1}^{i})$ for each recession and expansion and separately for each of the six age groups defined above. As we show in Section C of Appendix B, the four age groups between ages 35 and 54 behave similarly to each other over the business cycle. Motivated by this finding, from this point on we combine these individuals into one group and refer to them as "prime-age males." We also combine the first two age groups and refer to them as "young workers" (ages 25–34). For brevity, we focus on prime-age males in the main text and present the results for young workers in Appendix B.

### 1.    Recessions

Figure 13 plots $f_2$ for the four recessions during our sample period. For the Great Recession (black line with squares), $f_2$ is upward sloping in an almost linear fashion and rises by about 17 log points between the 10th and 90th percentiles of $\overline{Y}_{t-1}^{i}$. So workers with prerecession average earn-
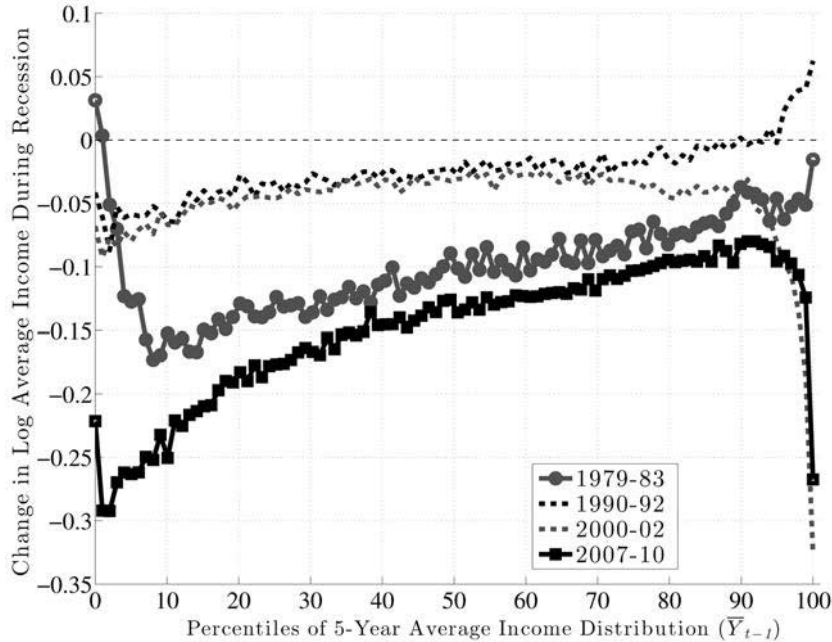
FIG. 13.—Change in log average earnings during recessions, prime-age males

ings in the 10th percentile saw their earnings decline by about 25 log points during the recession, compared with a decline of only 8 log points for workers in the 90th percentile.[19] Clearly, this factor structure leads to a significant widening of earnings inequality over much of the distribution. However, this good fortune of high-income individuals does not extend to the very top: $f_2$ first flattens beyond the 90th percentile, and then for the top 1 percent, it actually falls steeply. Specifically, these individuals experienced an average loss of 27 log points compared with 12.5 log points for those in the second-highest percentile. One conclusion we draw is that individuals near the 90th percentile of the average earnings distribution (about \$100,000 per year) as of 2006 have suffered the smallest loss of any earnings group.

With regard to the other major recession in our sample—the 1979–83 episode—$f_2$ looks similar to that in the Great Recession period between the 10th percentile and about the 95th percentile, with the same linear shape and a slightly smaller slope. However, for individuals with the lowest

---

[19] Recall that the earnings measure used in these computations, $y_t$, is net of earnings growth due to life cycle effects as explained in Sec. IV. This adjustment shifts the intercept of the $f_2$ function downward, which should be considered when interpreting the reported earnings growth figures.

average earnings (below the 10th percentile), the graph is downward sloping, indicating some mean reversion during the recession.[20] Also, and perhaps surprisingly, there is no steep fall in earnings for the top 1 percent during this recession. In fact, these individuals fared better than any other income group during this period. Overall, however, for the majority of workers, the 1979–83 recession was quite similar to—slightly milder than—the Great Recession, in terms of both its between-group implications and its average effect.

As for the remaining two recessions during this period, both of them feature modest falls in average earnings: about 3 log points for the median individual in these graphs. The 1990–92 recession also features mild but clear between-group differences, with $f_2$ rising linearly by about 7 log points between the 10th and 90th percentiles.[21] The 2000–2002 recession overlaps remarkably well with the former up to about the 70th percentile and then starts to diverge downward. In particular, there is a sharp drop after the 90th percentile. In fact, for the top 1 percent, this recession turns out to have the worst outcomes of all recessions: an average drop of 33 log points in 2 years!

Inspecting the behavior of $f_2$ above the 90th percentile reveals an interesting pattern. For the earlier two recessions, very high-income individuals fared better than anybody, whereas for the latest two recessions, there has been a reversal of fortunes, and this group has suffered the most.

To summarize, there is a clear systematic pattern to individual earnings growth during recessions. For all but the highest earnings groups, earnings loss during a recession decreases almost linearly with the pre-recession earnings level. The slope of this relationship also varies with the severity of the recession: the severe recessions of 1979–83 and 2007–10 saw a gap between the 90th and 10th percentiles in the range of 15 log points, whereas the milder recessions of 1990–92 and 2000–2002 saw a gap of 4–7 log points. Second, the fortunes of very high-income individuals require a different classification, one that varies over time: more recent recessions have seen substantial earnings losses for high-income individuals, unlike anything seen in previous ones. Below we will further explore the behavior of the top 1 percent over the business cycle.

[20] This probably has more to do with the fact that for the 1979–83 recession, we were limited to using only earnings in 1978 to form groups (rather than taking 5-year averages as we did for other periods), which led to a higher degree of mean reversion than would otherwise have been the case.

[21] These two recessions last half as long as the other two longer recessions, so the slope of these graphs should be interpreted in this context. However, normalizing total earnings growth (the vertical axis in these graphs) by the duration of each recession is not necessarily a satisfactory solution because even during longer recessions, the largest earnings falls have been concentrated within 1- or 2-year periods (2008–9, e.g.).

## 2. Expansions

Unlike recessions, $f_2$ is U-shaped during expansions (fig. 14). In particular, for workers who enter an expansion with average earnings above the 70th percentile, $f_2$ is an upward-sloping function, indicating further spreading out of the earnings distribution at the top. For workers with earnings below the median before the expansion, the pattern of earnings growth varied across expansions. The 1990s expansion was the most favorable, with a strong mean reversion raising the incomes of workers at the lower end relative to the median. The other two expansions showed little factor structure in favor of low-income workers: the function is quite flat, indicating that earnings changes were relatively unrelated to past earnings.

The pronounced U shape in the 1990s can be viewed as a stronger version of what Autor, Katz, and Kearney (2006) called "wage polarization" during this period. Basically, these authors compared the percentiles of the wage distribution at different points in time and concluded that the lower and higher percentiles grew more during the 1990s than the middle percentiles. Figure 14 goes one step further by following the same individuals over time and showing that it is precisely those individuals whose pre-1990s earnings were lowest and highest that experienced the fastest growth during the 1990s.
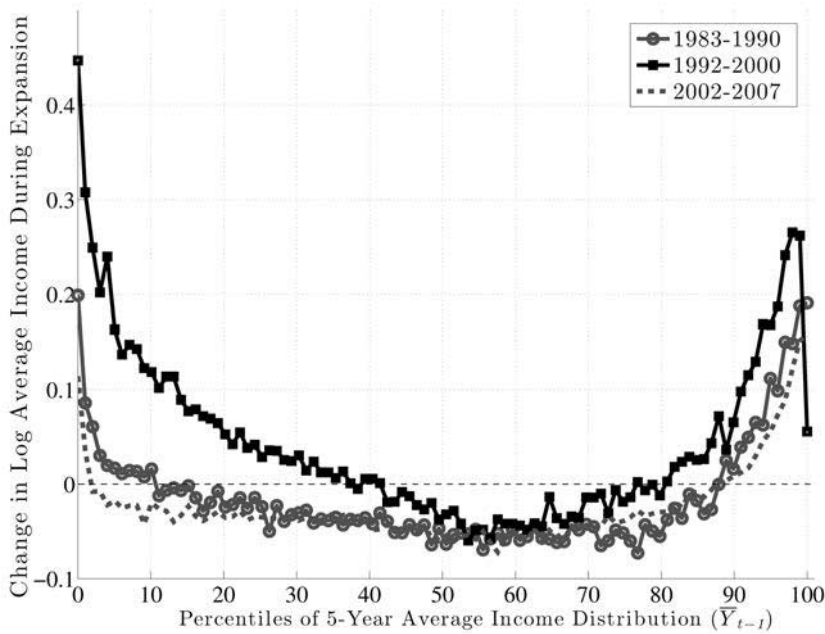


FIG. 14.—Change in log average earnings during expansions, prime-age males

### 3.   Putting Recessions and Expansions Together

To summarize these patterns, figure 15 aggregates $f_2$ across all six age groups (ages 25–54) and combines separate recessions and expansions. Overall, $f_2$ is U-shaped during expansions, indicating a compression of the earnings distribution at the bottom and expansion at the top. In contrast, recessions reveal an upward-sloping figure, implying a widening of the entire distribution except at the very top (above the 95th percentile). Thus, the main systematic component of business cycle risk is felt below the median and at the very top, two groups for which incomes rise fast in expansions and fall hard during recessions. Put together, these factor structures seen in figure 15 explain how the earnings distribution expands in recessions and contracts in expansions (resulting in countercyclical earnings inequality) without within-group (idiosyncratic) shocks having countercyclical variances.

### C.   The Top 1 Percent

Before concluding this section, we now take a closer look at top earners. To understand the differences and similarities within the top 1 percent,
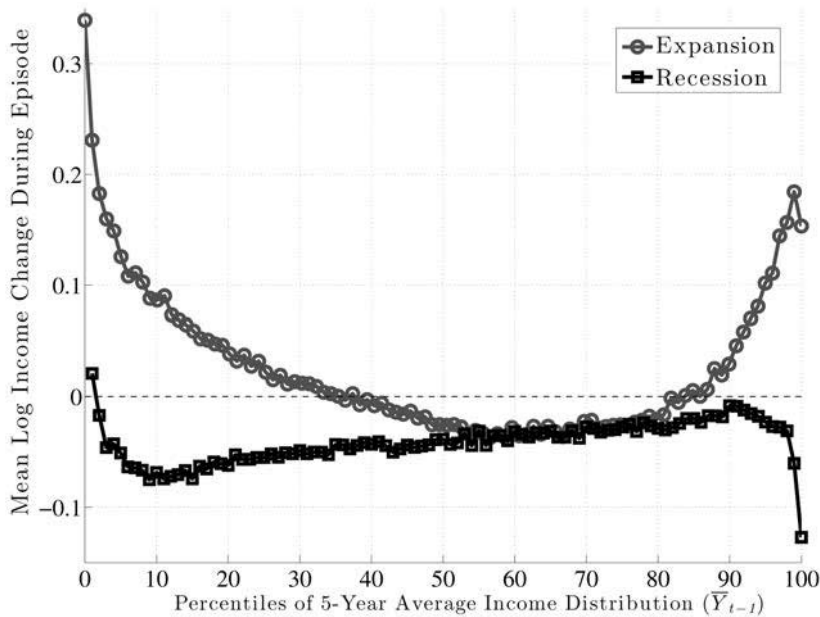


Fig. 15.—Change in log average earnings: expansions versus recessions, all workers. The recession graph has been scaled upward by 6.5 log points to be tangent to the expansion graph.

we divide this group into 10 quantiles and focus on the first, fourth, seventh, and tenth quantiles, denoted by P99.1, P99.4, P99.7, and the top 0.1 percent. Figure 16 (top panel) plots the annual change using the $f_2$ measure for each of these quantiles. First, notice that the four groups move quite closely to each other until the late 1980s, after which point a clear ranking emerges: higher quantiles become more cyclical than lower ones. In particular, individuals in higher quantiles have seen their earnings plummet in recessions relative to lower quantiles but did not see a larger bounce-back in the subsequent expansion, which would have allowed them to catch up. In fact, during expansions, the average earnings in each group grew by similar amounts.

The implication is that these differential losses during recessions across earnings quantiles are also persistent (fig. 16, bottom): individuals who were in the top 0.1 percent as of 1999 saw their earnings fall by an average of 50 log points between 2000 and 2005! Similarly large losses were experienced by the same group from 1989 to 1994 and from 2004 to 2009. By comparison, the 5-year loss for those in P99.1 ranges from 10 to 20 log points during these recessions. Thus, cyclicality increases strongly with the level of earnings.[22]

## VII.   Countercyclical Risk: Parametric Estimates

The nonparametric nature of the preceding analysis was essential for establishing our main empirical results by imposing as few assumptions as possible. At the same time, an important use of these empirical results is for calibrating economic models, for which parametric estimates are indispensable. With this in mind, this section provides parametric estimates of business cycle risk that can serve as inputs into quantitative models.

An important challenge we face in this task could be partly anticipated from the results established so far: earnings growth rates display important deviations from normality, which makes higher-order moments matter for earnings risk. Guvenen et al. (2013) show that an econometric process that aims to fully capture these features would have to be very complex. Estimating such a process—while also allowing for business cycle risk (which Guvenen et al. abstract from)—is beyond the scope of this paper. Furthermore, such a complex process would not be suitable for calibrating most economic models, where parsimony is of paramount importance. With these considerations in mind, we augment the "persistent-plus-transitory" specification (commonly used in the earnings dynamics

---

[22] Recall that $f_2$ averages each group's earnings before taking logs, which could be affected by a few extremely large earnings levels. By contrast, $f_1$ is the mean of logs, which is less sensitive to this problem. We include the graphs with $f_1$ in Sec. H of App. B, which shows the same qualitative patterns.
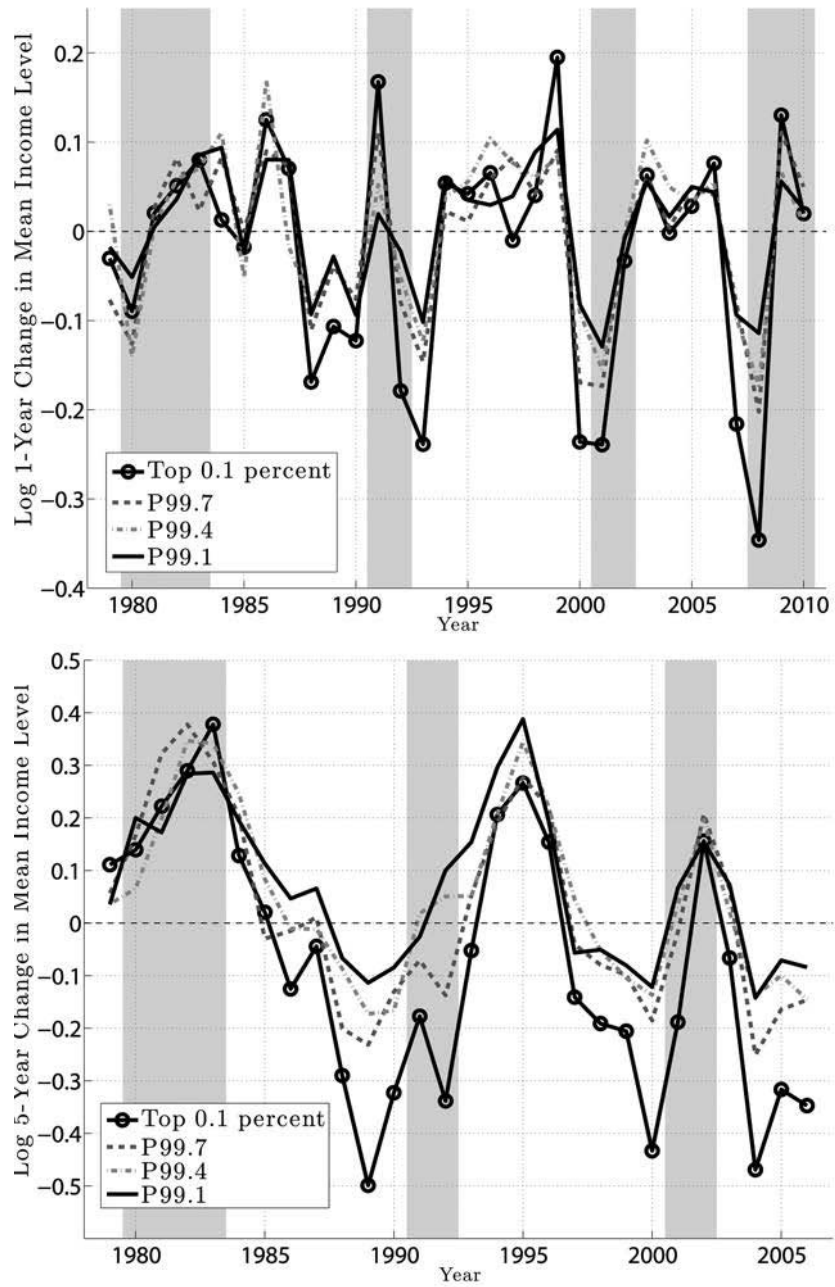
Fig. 16.—Earnings growth, 1- and 5-year, top 1 percent of prime-age males. *Top*, log 1-year growth in mean earnings ($f_2$). *Bottom*, log 5-year growth in mean earnings ($f_2$).

literature) with an error structure featuring a mixture of normals. Specifically,

$$y_t^i = z_t^i + \varepsilon_t^i, \tag{3}$$

$$z_t^i = \rho z_{t-1}^i + \eta_t^i, \tag{4}$$

where $\varepsilon_t^i \sim \mathcal{N}(0, \sigma_\varepsilon)$ and

$$\eta_t^i = \begin{cases} \eta_{1,t}^i \sim \mathcal{N}(\mu_{1s(t)}, \sigma_1) & \text{with prob. } p_1 \\ \eta_{2,t}^i \sim \mathcal{N}(\mu_{2s(t)}, \sigma_2) & \text{with prob. } 1 - p_1. \end{cases} \tag{5}$$

The subscript $s(t) = E, R$ indicates whether $t$ is an expansion or a recession year ($R = $ 1980–83, 1991–92, 2001–2, 2008–10). This specification allows for deviations from normality—for example, negative skewness and excess kurtosis—in earnings growth rates. Business cycle variation enters through changes in the means of the normal distributions ($\mu_{1s}$, $\mu_{2s}$).[23]

For our baseline case (model 1), we estimate the vector of parameters

$$(\rho, p_1, \mu_{1E}, \mu_{2E}, \mu_{1R}, \mu_{2R}, \sigma_1, \sigma_2, \sigma_\varepsilon)$$

by targeting (i) the mean, (ii) P90, (iii) P50, and (iv) P10 of 1-, 3-, and 5-year earnings changes, for a total of 372 moments.[24] We use a method of simulated moments estimator in which each moment is the percentage deviation between a data target and the corresponding simulated statistic. Online Appendix C contains further details of the estimation method (and reports all the data series used in the estimation to allow replication).

The parameter estimates are reported in the first column of table 1. A few points are worth noting. First, the two innovations are mixed with almost equal probability ($p_1 = .49$). Second, one of the two normal innovations is fairly large ($\sigma_1 = 0.325$), whereas the second one is drawn from a nearly degenerate distribution: $\sigma_2 = 0.001$. Although these figures may seem a bit strange at first blush, they are in fact consistent with a plausible economic environment in which $\eta_1^i$ and $\eta_2^i$ represent between-job and within-job earnings changes, respectively. In a given year, some workers do not change jobs, and their earnings growth is determined

---

[23] We have also experimented with a specification in which $p_1$ can vary over the business cycle. That specification did not fit the moments we examine below nearly as well.

[24] Data statistics are detrended by fitting a linear trend and rescaling each residual by the sample average to preserve the level of each statistic.

TABLE 1
ESTIMATES FROM THE PARAMETRIC MODEL

| Parameters | Model 1 Baseline | Model 2 | Model 3 | Model 4 |
|---|---|---|---|---|
| $\rho$ | .979 | .999 | .967 | .953 |
| $p_1$ | .490 | .473 | .896 | 1.00* |
| $\mu_{1E}$ | .119 | .088 | .067 | .065 |
| $\mu_{2E}$ | −.026 | −.011 | −.026 | . . . |
| $\mu_{1R}$ | −.102 | −.186 | .039 | .010 |
| $\mu_{2R}$ | .094 | .065 | −.317 | . . . |
| $\sigma_1$ | .325 | .327 | .185 | .242(E)/.247(R) |
| $\sigma_2$ | .001 | .017 | .493 | . . . |
| $\sigma_\varepsilon$ | .186 | .193 | .187 | .173 |

NOTE.—Model 1 (baseline) targets the mean, P90, P50, and P10 of 1-, 3-, and 5-year earnings changes. Model 2 defines a recession as a year with negative earnings growth from col. 5 of table A1. Model 3 is the same as model 1 but adds the third central moment as a fifth moment to target.

\* Model 4 reestimates the process in Storesletten et al. (2004), with the moments used for model 1. The two values for $\sigma_1$ in the last column are the estimates for expansions and recessions, respectively. See the text for details.

mainly by aggregate factors (such as GDP growth and inflation) and lead to earnings moving together by an amount roughly equal to $\mu_{2s}$. The rest of the population changes jobs and draws a new wage/earnings, $\eta_1^i$, and such changes have large dispersion. This structure has found empirical support in previous work (e.g., Topel and Ward 1992; Low, Meghir, and Pistaferri 2010), and our results provide further support for it.

An alternative way to write this process is to consider the limiting case of $\sigma_2 = 0$. Define $\lambda_t \equiv \eta_{2,t}^i = \mu_{2s(t)}$ to be an *aggregate* shock experienced by all workers, and modify (4) to read $z_t^i = \rho z_{t-1}^i + \lambda_t + \eta_{1,t}^i$. In addition, each individual faces a Poisson arrival process for the *idiosyncratic* shock: $\eta_{1,t}^i \sim \mathcal{N}(\mu_{1s(t)} - \mu_{2s(t)}, \sigma_1)$ with probability $p_1$ and is zero otherwise. This structure generates the same probability distribution for $y_t^i$ as (3)–(5).

Now we turn to business cycle variation in earnings risk. Figure 17 plots the fit of the baseline model (thick solid grey line) to the four sets of moments of $y_t^i - y_{t-1}^i$: mean, standard deviation, L90–50, and L50–10 (top to bottom). Because the estimated process allows for time variation only across expansions and recessions, it makes sense to also plot the US data by averaging each statistic over each business cycle episode (plotted as the thick black solid line). The baseline model captures (i) the consistent dip in the mean earnings growth rate in every recession as well as (ii) the dip in L90–50 and (iii) the rise in L50–10 in recessions. Consequently, the model matches the countercyclical left-skewness seen in the data fairly well, as seen in the top panel of figure 18, which plots Kelley's skewness measure. As for the standard deviation, the baseline model fails to match its high level; however, it does nicely capture the
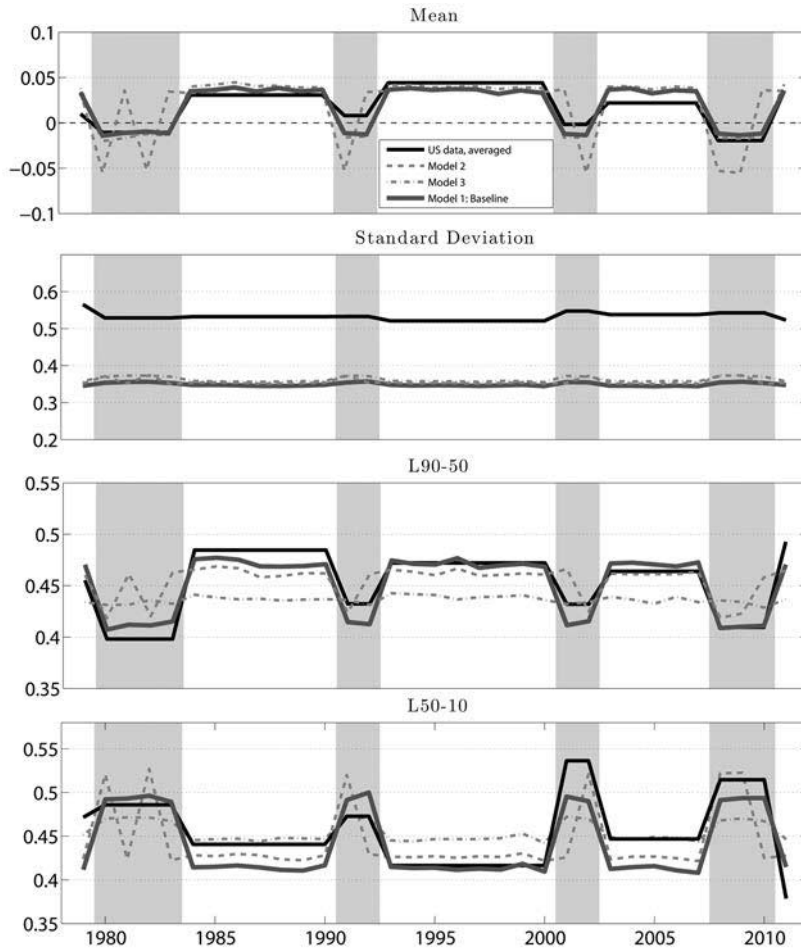
Fɪɢ. 17.—Moments of 1-year earnings change: data versus estimated parametric model

lack of cyclicality observed in the data (second panel of fig. 17). Finally, figure 19 plots the histograms of $y_t^i - y_{t-1}^i$ generated by the baseline model, which shows the same kind of shift in the distribution of shocks as figure 1 anticipated.

One assumption we made in model 1 was to identify recessions with NBER business cycle dates. Although this assumption is not controversial, as discussed earlier, it is not perfect either. Model 2 considers a plausible alternative, which essentially classifies year $t$ as a recession if average earnings growth in our sample was negative in that year (as reported in App. table A1). This assumption yields a smaller set of recession years: $R = 1980, 1982, 1991, 2002, 2008,$ and $2009$. Figure 17 plots the
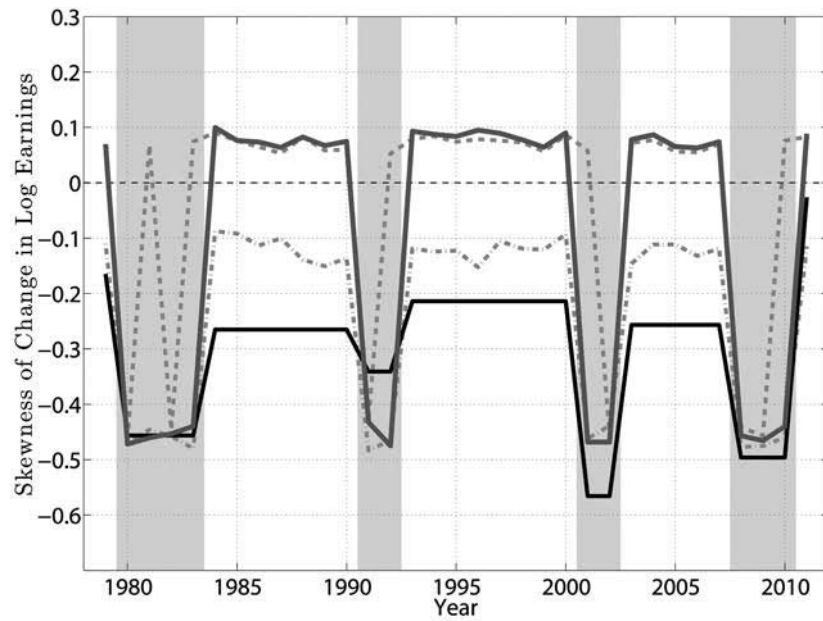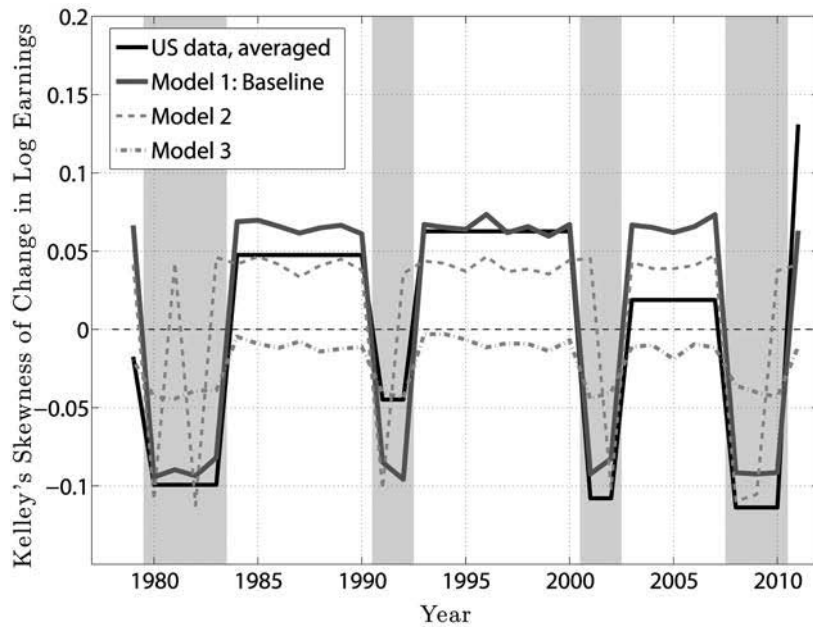
Fɪɢ. 18.—Skewness of 1-year earnings changes: data versus estimated parametric model
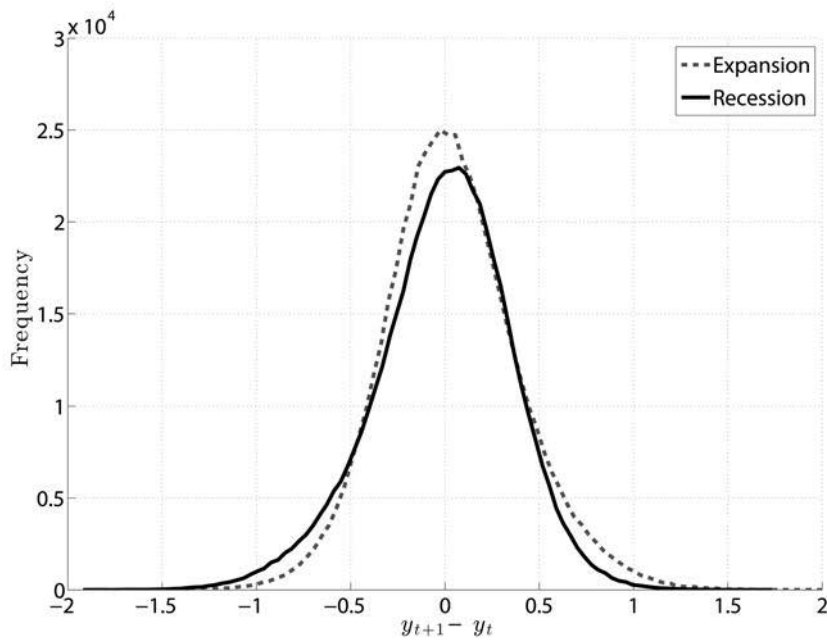
FIG. 19.—Histogram of annual earnings change, baseline parametric model

statistics from model 2 as the dashed grey line. Overall, the results look quite similar to the baseline, but this specification seems to match the ups and downs during some recessions better (compare it with the raw data graphs in figs. 4, 5, and 6).

Although both models 1 and 2 match the level and cyclicality of Kelley's skewness well, they both overestimate the level of the third central moment, which is plotted in the bottom panel of figure 18. The reason is that the latter is heavily influenced by the thick tails of the US earnings growth distribution, which the estimated processes fail to fully capture.[25] To see if this can be remedied, in model 3, we add the third central moment (again, of 1-, 3-, and 5-year changes) to the set of targets used for model 1. The estimated parameters are reported in table 1, and the grey dashed-dotted line in figure 17 plots the simulated moments. As seen here, while the mean and standard deviation are unaffected by this change, model 3 fails to generate sufficient cyclicality in L90–50 and L50–10 and, consequently, in Kelley's skewness. Model 3 does, however, match the lower

[25] In fact, this is intimately related to the failure of the model to match the level of the standard deviation mentioned above. In other words, the data exhibit too high a level of standard deviation relative to the L90–10 dispersion, indicative of excess kurtosis (i.e., long tails).

level of the third central moment better than models 1 and 2. Nevertheless, we do not view this as a sufficient reason to move away from our preferred baseline because the movements of L90–50 and L50–10 affect the nature of the shocks faced by the bulk of the population, and models 1 and 2 match them quite well.

Before we conclude this section, it seems useful, for completeness, to estimate the process considered by Storesletten et al. (2004). To this end, we set $p_1 = 1$ (which eliminates $\eta^i_{2,t}$, thereby reducing the process for $y^i_t$ to a normal distribution) and further allow the innovation variance $\sigma_1$ to change between recessions and expansions. Clearly, this process implies zero skewness by assumption, so the only substantive issue is whether the variance changes over the business cycle. We continue to target the same set of moments as in model 1. As seen in the last column, $\sigma_1$ barely moves, rising slightly from 0.242 to 0.247 from expansion to recession, consistent with our nonparametric finding from Section V that the variance of persistent shocks is acyclical.

## VIII. Discussion and Conclusions

This paper has studied between- and within-group variation in earnings growth rates over the business cycle. Using a large and confidential panel data set with little measurement error, it has documented three sets of empirical facts.

Our first set of findings concerns the cyclical nature of idiosyncratic shocks. During recessions, the upper end of the shock distribution collapses—that is, large upward earnings movements become less likely—whereas the bottom end expands—that is, large drops in earnings become more likely. Moreover, the center of the shock distribution (i.e., the median) is stable and moves little compared with either tail. What does change (more significantly) is the behavior of the tails, which swing back and forth in unison over the business cycle. These swings lead to cyclical changes in skewness, but not so much in overall dispersion. We conclude that recessions are best viewed as a small negative shock to the median and a large negative shock to the skewness of the idiosyncratic earnings shock distribution, with little change in the variance.

What accounts for the different conclusions reached by Storesletten et al. (2004) and this paper? A definitive answer would require an exact replication of that paper with our data set and a step-by-step elimination of each potential source of difference. While this step is beyond the scope of this paper, it is useful to point out some of the key differences that could potentially be responsible. First, that paper assumes an AR(1) specification for shocks, which restricts skewness to zero. Second, in the authors' estimation, the only parameter of the econometric process that is allowed to vary over the business cycle is the variance of shocks. Given

that earnings-level inequality is countercyclical (which is also true in our sample; see online fig. A3), the estimated variance would have to rise during recessions to account for the rising inequality. Furthermore, they assume that shock variances display no secular trend (from 1910 to 1993), despite much empirical evidence finding low-frequency movements in variances. Any one of these three identifying assumptions could potentially lead to a finding of countercyclical variance, since it is the only parameter that is allowed to vary in their estimation.[26]

Second, we examined the systematic component of business cycle risk. The pre-episode average earnings level turns out to predict a worker's earnings growth during subsequent business cycle episodes. During recessions, earnings growth is an increasing function of past earnings (except for very top earners), whereas during expansions it is a U-shaped function. Between-group differences are large and systematic. Put together, these factor structures are consistent with countercyclical earnings inequality without needing within-group (idiosyncratic) shocks that have countercyclical variances.

Third, the one deviation we find from these simple patterns is a remarkable nonlinearity for individuals who enter a recession with very high earnings—those in the top 1 percent. During the last two recessions, these individuals have experienced enormous and persistent earnings losses (about 30 log points), which dwarfs the losses of individuals even with slightly lower earnings. In fact, individuals who entered the last three recessions in the top 0.1 percent of the earnings distribution had earnings levels 5 years later that were at least 50 log points lower than their prerecession levels.

Overall, these empirical findings have important implications for how we think about earnings risk over the business cycle. The traditional approach to modeling recession risk consists of a (negative) aggregate shock and a positive shock to the variance of idiosyncratic shocks. Our results suggest that this simple view is inadequate. Instead, they turn our focus to the countercyclical variation in the third moment (skewness) of idiosyncratic shocks as central to understanding how the fortunes of ex ante similar individuals fare during recessions. Even the change in mean earnings (which we think of as an aggregate shock) is largely driven by the change in skewness. In addition, the factor structure results imply that business cycle risk is not entirely a surprise or a shock, but it has a component that can be predicted on the basis of information available to both individuals and economists at the beginning of business cycle episodes.

---

[26] The assumptions listed here merely reflect the limitations imposed by the small sample size of the PSID, which these researchers had to deal with.

## References

Abowd, J. M., and D. Card. 1989. "On the Covariance Structure of Earnings and Hours Changes." *Econometrica* 57 (March): 411–45.

Arellano, C., Y. Bai, and P. J. Kehoe. 2012. "Financial Frictions and Fluctuations in Volatility." Staff Report no. 466, Fed. Reserve Bank Minneapolis.

Autor, D. H., L. F. Katz, and M. S. Kearney. 2006. "The Polarization of the U.S. Labor Market." *A.E.R.* 96 (May): 189–94.

———. 2008. "Trends in U.S. Wage Inequality: Revising the Revisionists." *Rev. Econ. and Statis.* 90 (2): 300–323.

Bloom, N. 2009. "The Impact of Uncertainty Shocks." *Econometrica* 77 (3): 623–85.

Bloom, N., M. Floetotto, N. Jaimovich, I. Saporta-Eksten, and S. J. Terry. 2011. "Really Uncertain Business Cycles." Working paper, Stanford Univ.

Braun, A., and T. Nakajima. 2012. "Uninsured Countercyclical Risk: An Aggregation Result and Application to Optimal Monetary Policy." *J. European Econ. Assoc.* 10 (6): 1450–74.

Brav, A., G. Constantinides, and C. Geczy. 2002. "Asset Pricing with Heterogeneous Consumers and Limited Participation: Empirical Evidence." *J.P.E.* 110 (4): 793–824.

Chen, H., M. Michaux, and N. Roussanov. 2011. "Mortgage Refinancing and Macroeconomic Activity." Working paper, Sloan School, Massachusetts Inst. Tech.

Constantinides, G. M., and D. Duffie. 1996. "Asset Pricing with Heterogeneous Consumers." *J.P.E.* 104 (2): 219–40.

Edmond, C., and L. Veldkamp. 2009. "Income Dispersion and Counter-cyclical Markups." *J. Monetary Econ.* 56:791–804.

Farber, H. S. 2005. "What Do We Know about Job Loss in the United States? Evidence from the Displaced Workers Survey, 1984–2004." *Fed. Reserve Bank Chicago Econ. Perspectives* 29 (2): 13–28.

Fernandez-Villaverde, J., P. Guerron, J. F. Rubio-Ramirez, and M. Uribe. 2011. "Risk Matters: The Real Effects of Volatility Shocks." *A.E.R.* 101:2530–61.

Guvenen, F., F. Karahan, S. Ozkan, and J. Song. 2013. "What Do Data on Millions of U.S. Workers Say about Labor Income Risk?" Working paper, Univ. Minnesota.

Jacobson, L. S., R. J. LaLonde, and D. G. Sullivan. 1993. "Earnings Losses of Displaced Workers." *A.E.R.* 83 (4): 685–709.

Juhn, C., K. M. Murphy, and B. Pierce. 1993. "Wage Inequality and the Rise in Returns to Skill." *J.P.E.* 101 (June): 410–42.

Kocherlakota, N., and L. Pistaferri. 2009. "Asset Pricing Implications of Pareto Optimality with Private Information." *J.P.E.* 117 (3): 555–90.

Krebs, T. 2007. "Job Displacement Risk and the Cost of Business Cycles." *A.E.R.* 97 (3): 664–86.

Krusell, P., and A. A. Smith. 1997. "Income and Wealth Heterogeneity, Portfolio Choice, and Equilibrium Asset Returns." *Macroeconomic Dynamics* 1 (1): 387–422.

Low, H., C. Meghir, and L. Pistaferri. 2010. "Wage Risk and Employment Risk over the Life Cycle." *A.E.R.* 100 (4): 1432–67.

Lucas, R. E., Jr. 1987. *Models of Business Cycles.* New York: Blackwell.

———. 2003. "Macroeconomic Priorities." *A.E.R.* 93 (March): 1–14.

Mankiw, N. G. 1986. "The Equity Premium and the Concentration of Aggregate Shocks." *J. Financial Econ.* 17 (September): 211–19.

Meghir, C., and L. Pistaferri. 2004. "Income Variance Dynamics and Heterogeneity." *Econometrica* 72 (1): 1–32.

Moffitt, R. A., and P. Gottschalk. 1995. "Trends in the Variances of Permanent and Transitory Earnings in the U.S. and Their Relation to Earnings Mobility." Working Papers in Economics, no. 444 ( July), Boston Coll.

Olsen, A., and R. Hudson. 2009. "Social Security Administration's Master Earnings File: Background Information." *Soc. Security Bull.* 69 (3): 29–46.

Panis, C., et al. 2000. *SSA Program Data User's Manual.* Baltimore: Soc. Security Admin.

Parker, J. A., and A. Vissing-Jørgensen. 2010. "The Increase in Income Cyclicality of High-Income Households and Its Relation to the Rise in Top Income Shares." *Brookings Papers Econ. Activity* 41 (2): 1–70.

Sabelhaus, J., and J. Song. 2010. "The Great Moderation in Micro Labor Earnings." *J. Monetary Econ.* 57:391–403.

Schulhofer-Wohl, S. 2011. "Heterogeneity and Tests of Risk Sharing." *J.P.E.* 119 (5): 925–58.

Storesletten, K., C. I. Telmer, and A. Yaron. 2004. "Cyclical Dynamics in Idiosyncratic Labor Market Risk." *J.P.E.* 112 ( June): 695–717.

———. 2007. "Asset Pricing with Idiosyncratic Risk and Overlapping Generations." *Rev. Econ. Dynamics* 10 (October): 519–48.

Topel, R. H., and M. P. Ward. 1992. "Job Mobility and the Careers of Young Men." *Q.J.E.* 107 (May): 439–79.

Von Wachter, T., J. Song, and J. Manchester. 2009. "Long-Term Earnings Losses Due to Mass Layoffs during the 1982 Recession: An Analysis Using U.S. Administrative Data from 1974 to 2004." Technical report, Columbia Univ.