Changes in the Distribution of Income Volatility

Shane T. Jensen and Stephen H. Shore*†‡

October 30, 2007

Abstract

Research documenting an increase in income risk in recent decades has treated its proxy – income volatility, the expected magnitude of income changes – as if it were the same for everyone. In reality, some people face more risk than others. We develop a Markovian hierarchical Dirichlet process (MHDP) prior to model the evolution of the distribution of income volatilities, not just the mean. This augments the recently-developed hierarchical Dirichlet process (HDP) prior to accommodate the serial dependence of panel data. We find that increases over time in mean volatility can be attributed solely to increases in volatility at the right tail; the risky are getting riskier while the rest of the distribution remains unchanged. This has strong welfare implications, as those facing increasing volatility may also be the most risk tolerant, as evidenced by their willingness to take on income risk.

*Jensen: Wharton School, Department of Statistics, stjensen@wharton.upenn.edu. Shore: Johns Hopkins University, Department of Economics, shore@jhu.edu. Please contact Shore at: 358 Mergenthaler Hall, 3400 N. Charles Street, Baltimore, MD, 21218; 410-516-5564.

†JEL Classification: D31 - Personal Income, Wealth, and Their Distributions.

‡We thank Dylan Small for helpful comments, as well as seminar participants at the Wharton School of the University of Pennsylvania. Shore thanks Ning Ma for excellent research assistance.

Contents

List of Figures

List of Tables

1 Introduction

Idiosyncratic income risk is arguably the largest source of risk people face. To the degree that people are risk-averse, income risk may carry substantial welfare costs. Income volatility – the variance of squared income changes – can be used to measure income risk. There has been a great deal of recent interest by economists in the evolution of income risk (Dahl, DeLeire, and Schwabish, 2007; Dynan, Elmendorf, and Sichel, 2007; Gottschalk and Moffitt, 1994, 2002; Hacker, 2006). Many but not all of these papers argue that income changes have gotten larger in recent decades or that large drops in income have become more likely. Some work (Hacker, 2006) argues that since people are risk-averse, this increase in risk is a bad thing.

To date, research on the evolution of income volatility has treated this volatility as if it were a single number in any given year. In reality, income risk will not be the same for everyone, even among those with similar demographics. Volatility will vary not just over time but also across individuals. Population volatility in a given year is a *distribution* of parameters and not a single parameter. Effectively, previous research has been estimating the evolution of the population *mean*. These changes in mean may not capture all important changes in the full distribution. Volatility will be increasing for some people and at some times and decreasing for others (Meghir and Pistaferri, 2004). Since these changes need not cancel out, the 5th, 50th, and 95th percentiles of the volatility distribution may evolve very differently.

In this paper, we provide non-parametric estimates of the evolution of the distribution of income volatility. We find that increases in the mean of income volatility can be attributed solely to the right tail of volatility moving farther right. While the riskiest end of the distribution has gotten riskier, volatility has not changed for the median person (and not changed very much for the 10th, 25th, or 75th percentile person). This has important welfare implications. To the degree that those who are most risk-tolerant will take on the most income risk (Schulhofer-Wohl, 2006), an increase in volatility for those people may not be a bad thing.

We use panel data on labor income for men from the PSID to estimate the parameters of a flexible income process. This income process incorporates both permanent and transitory shocks, and allows them to enter and damp out in a general way. In the absence of heterogeneity, our iterative solution method would be relatively straightforward. We alternate between a) using a Kalman filter to decompose income changes into permanent and transitory shocks (given model parameters) and then b) using a regression to estimate model parameters (given these shocks).

The chief innovation in this paper is the use of hierarchical Bayes, so that volatility parameters for a given individual at an instant in time are drawn from a distribution of values. We begin with a Dirichlet process (DP) prior model that allows the distribution of volatility to take a very general shape. In a DP model, the parameters – in this case, volatility – can take any one of N values, where these values and the number of them are determined by the data. To allow for the possibility that the many observations for a given individual are likely to be more similar to each other than to the observations from other individuals, we use the recently-developed hierarchical Dirichlet process (HDP, see Teh, Jordan, Beal, and Blei, 2007) prior

model. This is the first application of an HDP model in economics.

The DP and HDP models are not set up for panel data, where repeated observations for an individual have a natural order: time. We extend the existing DP and HDP models to allow for Markovian dependencies between consecutive years within each individual. Unlike these models, our Markovian hierarchical Dirichlet Process (MHDP) model allows the distribution of volatility to evolve over time (or not) based on the data.

Our method gives a posterior estimate of both any individual's volatility parameters and of the distribution of parameters in the population at any point in time. It is then straightforward to look at evolution over time of any individual's volatility parameters and also the the distribution in the population. Our chief result is that mean volatility has increased since 1968, but that this can be attributed exclusively to increases in volatility among at the volatile tail of the distribution.

2 Data

Data is drawn from the Panel Study of Income Dynamics (PSID). The PSID was designed as a nationally representative panel of U.S. households. It tracked families annually from 1968 to 1997 and in odd-numbered years after that; this paper uses data through 2005. The PSID includes data on education, income, hours worked, employment status, and age for members of households. In this paper, we limit the analysis to men age 22 to 60 and use annual labor income as a measure of income.[1] Table 1 presents summary statistics from these data.

We want to ensure that changes in income are not driven by changes in the top-code (the maximum value for income entered that can be entered in the PSID). The lowest top code for income in real terms was \$99,999 in 1982 (equivalent to \$202,281 in 2005 dollars, after which the top-code rises to \$9,999,999). So that top-codes will be standardized in real terms, this minimum top-code is imposed on all years in real terms, so the top-code is \$99,999 in 1982 and \$202,281 in 2005. We also want to ensure that results for the log of income are not dominated by small changes in the level of income near zero (which will imply huge or infinite changes in the log of income). To address this concern, we replace income values that are very small or zero with a non-trivial lower bound. We choose as this lower-bound the income that would be earned from a half-time job (1,000 hours per year) at the real equivalent of the 2005 federal minimum wage (\$5.15 per hour). This imposes a bottom-code of \$5,150 in 2005 and \$2,546 in 1982. Note that the difference in log income between the top- and bottom-code is constant over time. The vast majority of the values below this bound were exactly zero. This bound allows us to exploit transitions into and out of the labor force. At the same time, the bound prevents economically

[1]Labor income in 1968 is labeled v74 for husbands and has a constant definition through 1993. From 1994, we use the sum of labor income (HDEARN94 in 1994) and the labor part of business income (HDBUSY94), with a constant definition through 2005. Note that data is collected on household "heads" and "wives" (where husbands are always so the "head" of any couple), so that men who are not household heads (as would be the case if they lived with their parents) are excluded.

Table 1: Summary Statistics

	mean	st. dev.	min	max
year	1986.4	9.6	1968	2005
age (years)	37.5	10	22	60
education (years)	12.2	3.4	0	17
# of observations/person	13.9	8.4	1	34
married (1 if yes, 0 if no)	0.83	.	.	.
black (1 if yes, 0 if no)	0.28	.	.	.
annual income (2005 $s)	$41,322	$44,314	0	$3,000,000
annual income ($s)	$24,178	$34,775	0	$3,000,000
family size	3.4	1.7	1	19

This table summarizes data from 124,792 observations on 8,949 male household heads.

unimportant changes that are small in levels but huge in logs (because the level of income is close to zero) from dominating the results. Results are robust to other values for this lower bound, such as the income from full-time work (2,000 hours per year) at the 2005 minimum wage (in real terms).

In this paper, we model the evolution of "excess" log income. This is taken as the residual from a regression of the natural log of labor income (top- and bottom-coded as described) on host of regressors. We use as regressors: a cubic in age for each level of educational attainment (none, elementary, junior high, some high school, high school, some college, college, graduate school), the presence and number of infants, young children, and older children in the household, the total number of family members in the household, and dummy variables for each calendar year. Including calendar year dummy variables eliminates the need to convert nominal income to real income explicitly. While this step is standard in the literature that estimates income process parameters, it is not necessary to obtain our results.

Table 2 presents data on the distribution of real annual income for men in column 1 (imposing top- and bottom-code restrictions in parentheses). While the mean real income is nearly identical with and without top- and bottom-code restrictions ($41,558 versus $40,957), these restrictions on extreme values reduce the standard deviation of real income from $44,596 to $31,624). Column 2 shows the distribution of "excess" log income. Since excess log income is the residual from a regression, its mean is zero. The inter-quartile range of excess log income is -0.36 to 0.51, which with a log-linearized approximation corresponds to income increasing 87 percent when moving from the 25^{th} percentile to the 75^{th} percentile. Column 3 presents the distribution of one-year changes in excess log income. Naturally, the mean of one-year changes is close to zero. The inter-quartile range of one-year changes is -0.12 to 0.16; excess income does not change more than 12 to 16 percent from year to year for most individuals. However, there are extreme changes in income, so the standard deviation of changes to log income is 0.50. This implies either that changes to income

Table 2: Distribution of Income, Excess Log Income, and Income Changes for Men

	Real Income	Excess Log Income	One-Year Change	Five-Year Change
Mean	\$41,322 (\$41,123)	-0.088	-0.008	-0.047
St. Dev.	\$44,314 (\$33,123)	0.654	0.418	0.593
Observations	124,792	124,792	100,314	75,635
Minimum	\$0 (\$5,124)	-1.297	-2.205	-2.205
5^{st} Percentile	\$0 (\$5,124)	-1.297	-0.709	-1.220
25^{th} Percentile	\$19,508	-0.468	-0.115	0
50^{th} Percentile	\$35,213	0.037	0	0
75^{th} Percentile	\$53,458	0.394	0.123	0.229
95^{th} Percentile	\$97,086	0.885	0.636	0.862
Maximum	\$3,000,000 (\$202,241)	0.908	2.205	2.205

Table 2 describes the distribution of labor income for men in the PSID over the period from 1968 to 2005. See Section 2 for a detailed description of the income variable and the top- and bottom-coding procedure. Column 1 shows the distribution of real annual income for men. The numbers in parentheses are the values with top- and bottom-coding restrictions. Column 2 shows the distribution of "excess" log income, the residual from the regression of log labor income on the covariates enumerated in Section 2. Column 3 presents the distribution of one-year changes in excess log income. This implies either that changes to income have fat tails, or that there is heterogeneity in volatility. Column 4 repeats the results for column 3, but presents five-year changes instead of one-year changes.

---

have fat tails, or alternatively that there is heterogeneity in volatility. Unless a model is identified from parametric assumptions, these are observationally equivalent when there is only a cross-section of volatility estimates. However, heterogeneity and fat tails have different implications for the time-series of volatility, and we exploit these in this paper. Column 4 repeats the results from column 3, but presents five-year changes instead of one-year changes. These long-term changes have only slightly higher standard deviations than the one-year change, 0.71 vs. 0.50, suggesting some mean-reversion in income.

This mean-reversion can be seen explicitly in Table 3, which shows the autocorrelation of $y_{i,t} - y_{i,t-1}$, the autocorrelation in one-year changes in men's excess log incomes. One-year auto-correlation in income is strongly negative (-0.30); income tends to increase in the year following a decrease. This negative autocorrelation between consecutive one-year changes in income is a central feature earnings dynamics data; any process for labor income must accommodate negative auto-correlation. Auto-correlation at longer lags is negative but small and decreasing. The non-zero values for higher-order lags are driven primarily by the inclusion of individuals without income. Such individuals leave the labor force (income falls) and then subsequently re-enter the labor force (income rises). Excluding these observations eliminates higher-order autocorrelation almost entirely. Section 3 presents a standard income

Table 3: Autocorrelation of Changes in Excess Log Income

correlation	$y_{i,t} - y_{i,t-1}$
$y_{i,t-1} - y_{i,t-2}$	-0.300
$y_{H,t-2} - y_{H,t-3}$	-0.053
$y_{H,t-3} - y_{H,t-4}$	-0.029
$y_{H,t-4} - y_{H,t-5}$	-0.023

Autocorrelation of one-year changes in excess log incomes at one- through four-year lags. The distribution of these one-year changes are shown in Table 2.

process that accomodates these features of the data.

The paragraph on missing data is not complete but will go here. Households occasionally have missing data. More problematically, the PSID does not collect any data in even-numbered years following 1997. Our estimation procedure has no explicit way to account for this. Instead, we will fill in these missing observations with "guesses" of income. These guesses could be obtained as: a) the subsequent value (which assumes the entire change in income occurs immediately and will tend to over-estimate income volatility around missing data; b) the mid-point of previous and subsequent values, which will under-estimate income volatility around missing data; or c) as a random draw of income drawn from the set of income changes elsewhere in the data with a similar overall change during a period of this interval. Of these, the a) and b) are simplest while c) is most likely to ensure that any changes over time in our results cannot be driven by changes in the frequency of missing values. For the initial estimation for this draft (Tables 4 onward), data are drawn from the set of married men, data from missing years are ignored so that data from 1999 and 2001 are placed next to each other, and Winsorizing is done at the 5% and 95% level instead of at the real top- and bottom-codes discussed here. These shortcomings will be remedied in the subsequent draft.

3 Income Process

Here, we present a standard general process for income (following Carroll and Samwick, 1997; Meghir and Pistaferri, 2004, and many others). Recall that excess log labor income for individual i at time t is $y_{i,t}$. We assume that excess log income is composed

of permanent (p) and transitory (ξ) components:

$$\begin{aligned} y_{i,t} &= p_{i,t} + \xi_{i,t} \\ p_{i,t} &= p_{i,0} + \sum_{k=1}^{t-\Omega} \omega_{i,k} + \sum_{k=t-\Omega+1}^{t} \theta_{t-k}\omega_{i,k}. \\ \xi_{i,t} &= \sum_{k=t-\epsilon+1}^{t} \phi_{t-k}\varepsilon_{i,k} \end{aligned} \tag{1}$$

Excess log income is the sum of permanent income, $p_{i,t}$, and transitory income, $\xi_{i,t}$. Permanent income is an initial income, $p_{i,0}$, plus the weighted sum of all shocks to permanent income, $\omega_{i,t}$. Transitory income is the weighted sum of recent shocks to transitory income, $\varepsilon_{i,t}$, but with the additional constraint that the weights sum to one ($\sum_k \phi_k = 1$). Permanent and transitory shocks are assumed to be independent of one another, over time, and across individuals. Subsequently, "permanent variance" refers to the variance of permanent shocks, $\sigma^2_{i,t} \equiv E\left[\omega^2_{i,t}\right]$; "transitory variance" refers to the variance of transitory shocks, $\tau^2_{i,t} \equiv E\left[\varepsilon^2_{i,t}\right]$; "variance parameters" or "volatility parameters" refers to $(\sigma^2_{i,t}, \tau^2_{i,t})$ jointly. Here, permanent shocks come into effect over Ω periods, and transitory shocks fade completely after ϵ periods. As an example of our notation, θ_2 denotes the weight placed on a permanent shock from two periods ago in current log income; ϕ_2 denotes the weight placed on a permanent shock from two periods ago in current log income. Note that while we use the word "shock" for parsimony, these innovations to income may be predictable to the individual, even if they look like shocks in the data.

While this model provides a useful stylized process for income, income does not evolve exactly according to this process. Most obviously, labor income will be zero – and log income undefined – when an individual is not in the labor force; transitions into and out of employment the labor force are not modeled. While bottom-codes imposed on income allow zeros to be accommodated, the income process does not account for them explicitly.

4 Heterogeneity

Note that $(\sigma^2_{i,t}, \tau^2_{i,t})$ describe the magnitude of income changes for an individual at a point in time. These volatility parameters can and will differ across individuals and over time. To date, the literature on the evolution of income volatility has assumed homogeneity at any point in time, $\sigma^2_{i,t} = \sigma^2_t$ and $\tau^2_{i,t} = \tau^2_t$. While a few papers have considered heterogeneity in income processes across individuals or over time, these were not designed to estimate an evolving distribution of parameters (Alvarez, Browning, and Ejrnaes, 2001; Meghir and Pistaferri, 2004).

We could relax the assumption of homogeneity while imposing strong parametric assumptions, for example that $(\sigma^2_{i,t}, \tau^2_{i,t})$ are jointly drawn from an inverse Wishart distribution. The problem with such an approach is that it imposes a particular parametric shape on the distribution of volatilities that cannot change over time.

This would preclude us from looking for, much less finding, changes in the shape of the volatility distribution.

Instead, we develop a new methodology based on recent advances in non-parametric Bayesian modeling. Specifically, we model $(\sigma^2_{i,t}, \tau^2_{i,t})$ as being jointly drawn from a completely unknown (non-parametric) distribution. Sharing of information across individuals and across years is accomplished by using a Dirichlet process (DP) prior distribution. In a DP prior model, observations are placed in clusters, with all observations in the same cluster having the same parameters. The number of clusters and the parameters that describe each cluster are not known *a priori* but rather identified from the data. This allows volatility parameters for a given individual in a given year to be identical to values for other person-years, or to take on unique values.

While the DP prior model can be used to approximate any other distribution, it does have an implicit structure. For a given observation, the probability of taking on particular parameter values is proportional to the number of other observations with those particular volatility parameter values. A popular metaphor for this model is called the Chinese restaurant process. When the first patron arrives at an empty Chinese restaurant, she chooses a seat at her preferred table. When the next patron arrives, he must choose whether to sit with the first patron, or to sit at an unoccupied table. The third patron arrives and decides to sit with one of the first two, or sit at an unoccupied table, and so on. People entering this restaurant prefer popular tables, and so are more likely to join a table that already contains many patrons. In this case, each patron refers to an individual in a year and tables correspond to possible values of the volatility parameters for that person-year. See Muller and Quintana (2004) for a general introduction to Bayesian non-parametric modeling with the Dirichlet process.

The standard DP model is appealing because it allows the distribution of volatility to take a general shape. However, this one-level model is inadequate for our purposes because it does not model systematic differences between people in income volatility. The standard DP model assumes that all pairs of observations are *a priori* equally likely to have the same parameters, regardless of whether or not they come from the same person or from different people. This shortcoming is overcome in the two-level hierarchical Dirichlet process (HDP) model, recently introduced in Teh, Jordan, Beal, and Blei (2007). This generalization of the standard DP model allows groups of observations (in our case, multiple observations from one individual in many years) to be *a priori* more likely to share parameters. Unfortunately, the HDP model cannot account for panel data, where time gives a natural order to observations. In an HDP model, all pairs of observations from the same person are *a priori* equally likely to share parameters, regardless of whether or not they are close together in time. In other words, that model cannot account for autocorrelation.

We allow for autocorrelation by extending the DP model to a three-level Markovian hierarchical DP (MHDP) model, illustrated in Figure 1. In our model, each individual begins with a given volatility (or more precisely a pair of volatility parameters $(\sigma^2_{i,t}, \tau^2_{i,t})$), the distribution of which may come from the standard DP process. In subsequent years, the volatility parameter will evolve according to this three-level hierarchy (also shown in Figure 1, with the number of the equation giving the prob-

Figure 1: Model Hierarchy

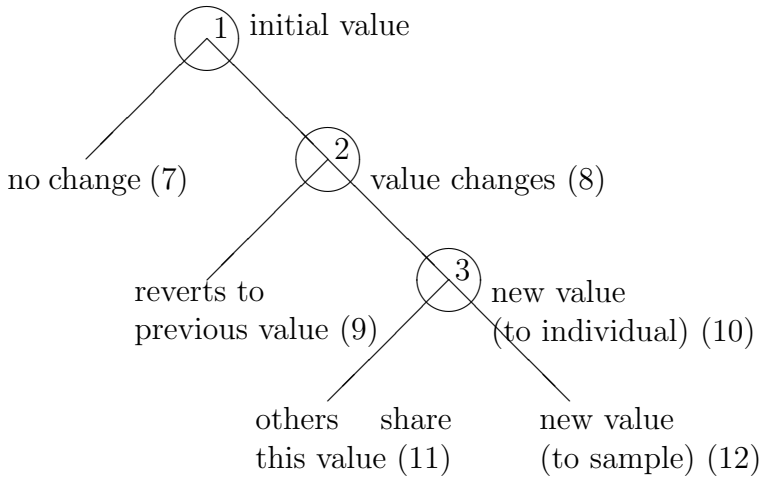


Diagram describes evolution of volatility parameters. The numbers 1, 2, and 3 in circles at each decision node correspond to the levels of the hierarchy described on page 10. The numbers (7) through (12) identify the equation number giving the probability of reaching that branch.

---

ability of each possibility in parentheses):

Level 1 A volatility parameter can remain unchanged from the *previous period* (7) or can change (8). If the volatility parameter changes, it can change to;

Level 2 A parameter from the set of *other values for that individual* (9) or can take on new value (a value the individual has not had before) (10). If the volatility parameter takes on a new value, this value can be;

Level 3 A volatility parameter held by *other individuals* (11) or can be new value not shared with other individuals (12).

Level 3 of our MHDP model would also be found in a DP or HDP model. Here, an observation may or may not take on the same parameter value as other observations in the sample. Level 2 of our MHDP model would also be found in an HDP model (but not a standard DP model). Here, an observation may take on the same parameter value as other observations for that individual. Level 1 is unique to our model. Here, an observation may remain unchanged, namely take on the same parameter as the immediately previous observation. Allowing a Markovian dependence on the immediately previous state is a novel methodological development that allows us to consider time-series or panel data. This structure also allows us to differentiate heterogeneity in volatility from fat-tailed shocks. While these two possibilities are observationally equivalent for a cross-section of estimates, the latter implies that estimates of volatility are not autocorrelated.

Table 4: Summary of Notation

0. Data used to estimate parameters and shocks:

   $\boldsymbol{y}$ (N by $T+1$ matrix) : Excess log incomes

1. Parameters, constant over time and across individuals:

   $\boldsymbol{\theta}$ ($\Omega - 1$ by 1 matrix): Rate at which permanent shocks ($\omega_{i,t}$) enter
   $\boldsymbol{\phi}$ ($\epsilon - 1$ by 1 matrix): Rate at which transitory shocks ($\varepsilon_{i,t}$) damp out
   γ^2 : Variance of homogeneous transitory noise

2. Shocks:

   $\boldsymbol{\omega}$ (N by T matrix) : Realized permanent shocks ($\omega_{i,t}$)
   $\boldsymbol{\varepsilon}$ (N by T matrix) : Realized transitory shocks ($\varepsilon_{i,t}$)

3. Parameters, vary over time and across individuals:

   $\boldsymbol{\sigma}^2$ (N by T matrix) : Variances of permanent shocks ($\omega_{i,t}$)
   $\boldsymbol{\tau}^2$ (N by T matrix) : Variances of transitory shocks ($\varepsilon_{i,t}$)

See Section 5 for details on these variables. Vectors and matrices are written in bold.

---

5 Estimation

The model uses annual data on excess log income for N individuals over T years. Our results include estimates of the parameters that describe the income process, the permanent and transitory shocks that drive changes in income, and the permanent and transitory volatility parameters for each individual in each year. These are summarized in Table 4 and described in detail below. Vectors and matrices are written in bold:

0. Data ($\boldsymbol{y}$): We use $\boldsymbol{y}$ to denote the N by T matrix of excess log incomes where the element in the ith row of the tth column is the excess log income for individual i in year t, $y_{i,t}$. This matrix will be ragged because not all individuals have data in all years.

1. Income Process Parameters ($\boldsymbol{\theta}, \boldsymbol{\phi}, \gamma$): We use $\boldsymbol{\theta}$ to denote the $\Omega - 1$ by 1 vector whose kth element is θ_k, the rate at which permanent shocks enter into income; we use $\boldsymbol{\phi}$ to denote the $\epsilon - 1$ by 1 vector whose kth element is ϕ_k, the rate at which transitory shocks damp out of income. These are constant over time and across individuals. γ^2 is the variance of homogeneous transitory noise. While the homogeneous γ^2 is separated from the heterogeneous transitory variance parameters, $\boldsymbol{\tau}^2$, as part of the estimation strategy, both capture transitory income changes.

2. Shocks ($\boldsymbol{\omega}, \boldsymbol{\varepsilon}$): We use $\boldsymbol{\omega}$ to denote the N by T matrix of realized (but not directly observed) permanent shocks where the element in the ith row of the tth column is the permanent shock for individual i between years $t-1$ and t, $\omega_{i,t}$. We use $\boldsymbol{\varepsilon}$ to denote the N by T matrix of realized (but not directly observed) transitory shocks

where the element in the ith row of the tth column is the transitory shock for individual i between years $t-1$ and t, $\varepsilon_{i,t}$. Like $\boldsymbol{y}$ these matrices will be ragged; without data there is no shock to decompose into permanent and transitory components.

3. Volatility Parameters $\boldsymbol{\sigma}^2$ to denote the N by T matrix of permanent variances where the element in the ith row of the tth column is $\sigma^2_{i,t}$, the permanent variance of individual i between years $t-1$ and t; we use $\boldsymbol{\tau}^2$ to denote the N by T matrix of transitory variances where the element in the ith row of the tth column is $\tau^2_{i,t}$, the transitory variance of individual i between years $t-1$ and t. The rows of $\boldsymbol{\sigma}^2$ and $\boldsymbol{\tau}^2$, $\boldsymbol{\sigma}^2_i$ and $\boldsymbol{\tau}^2_i$ respectively, represent the distribution of volatility parameters $\sigma^2_{i,t}$ and $\tau^2_{i,t}$ over time t for an individual i; the columns of $\boldsymbol{\sigma}^2$ and $\boldsymbol{\tau}^2$, $\boldsymbol{\sigma}^2_t$ and $\boldsymbol{\tau}^2_t$ respectively, represent the distribution of volatility parameters $(\sigma^2_{i,t}, \tau^2_{i,t})$ at time t across all individuals i. This paper aims to identify the evolution of the permanent and transitory variance $(\boldsymbol{\sigma}^2, \boldsymbol{\tau}^2)$ over time. Like $\boldsymbol{y}$, these matrices will be ragged.

We take a fully Bayesian approach to the estimation of these parameters: we estimate the joint posterior distribution of all unknown parameters conditional on our observed data as:

$$p(\boldsymbol{\theta}, \boldsymbol{\phi}, \gamma, \boldsymbol{\omega}, \boldsymbol{\varepsilon}, \boldsymbol{\sigma}^2, \boldsymbol{\tau}^2 | \boldsymbol{y}) \propto p(\boldsymbol{y} | \boldsymbol{\theta}, \boldsymbol{\phi}, \gamma, \boldsymbol{\omega}, \boldsymbol{\varepsilon}) \cdot p(\boldsymbol{\omega}, \boldsymbol{\varepsilon} | \boldsymbol{\sigma}^2, \boldsymbol{\tau}^2) \cdot p(\boldsymbol{\sigma}^2, \boldsymbol{\tau}^2) \tag{2}$$

Following Bayes rule, the distribution of parameters given the data – $p(\boldsymbol{\theta}, \boldsymbol{\phi}, \gamma, \boldsymbol{\omega}, \boldsymbol{\varepsilon}, \boldsymbol{\sigma}^2, \boldsymbol{\tau}^2 | \boldsymbol{y})$ – is proportional to the product of the distribution of the data given those parameters – $p(\boldsymbol{y} | \boldsymbol{\theta}, \boldsymbol{\phi}, \gamma, \boldsymbol{\omega}, \boldsymbol{\varepsilon})$ – and the probability of those parameters – $p(\boldsymbol{\omega}, \boldsymbol{\varepsilon} | \boldsymbol{\sigma}^2, \boldsymbol{\tau}^2) \cdot p(\boldsymbol{\sigma}^2, \boldsymbol{\tau}^2)$. We will estimate the posterior distribution of our unknown parameters by Markov Chain Monte Carlo (MCMC) simulation, specifically the Gibbs sampler (Geman and Geman, 1984). The Gibbs sampler estimates the full posterior distribution in equation (2) by iteratively sampling a value for each unknown parameter conditional on the current values of the other unknown parameters. In other words, we iterate over the following steps. The number of each step corresponds to subsection number in this section, so that Subsection 5.1 describes Step 1 in detail and so on. The notation for variables estimated in a given step are summarized in Table 4, so that entry 1 in Table 4 summarizes the variables estimated in Step 1 and so on.

Step 1: Sample new values of $(\boldsymbol{\theta}, \boldsymbol{\phi}, \gamma)$ from $p(\boldsymbol{\theta}, \boldsymbol{\phi}, \gamma | \boldsymbol{y}, \boldsymbol{\omega}, \boldsymbol{\varepsilon}, \boldsymbol{\sigma}^2, \boldsymbol{\tau}^2)$

Step 2: For each person i and year t, sample new values of $(\omega_{i,t}, \varepsilon_{i,t})$ from $p(\omega_{i,t}, \varepsilon_{i,t} | \boldsymbol{y}, \boldsymbol{\theta}, \boldsymbol{\phi}, \gamma, \boldsymbol{\sigma}^2, \boldsymbol{\tau}^2)$

Step 3: For each person i and year t, sample new values of $(\sigma^2_{i,t}, \tau^2_{i,t})$ from $p(\sigma^2_{i,t}, \tau^2_{i,t} | \boldsymbol{y}, \boldsymbol{\theta}, \boldsymbol{\phi}, \gamma, \boldsymbol{\omega}, \boldsymbol{\varepsilon})$ using three-level MHDP model.

These sampling steps form a Markov chain that is iterated until the set of all parameters has converged to their joint posterior distribution. We present our evaluation of convergence in the supplemental materials. In the following subsections, we outline in detail the sampling process for the three steps given above.

5.1 Step 1: Sampling income process parameters $(\boldsymbol{\theta}, \boldsymbol{\phi}, \gamma)$

Our model for the income process is based on a decomposition of the excess log income into permanent and transitory components, as outlined in Section 3. Reorganizing equation (1) and setting limits of $\Omega = 3$ periods for permanent shocks to come into effect and $\epsilon = 3$ periods for transitory shocks to fade completely (a conservative choice according to Abowd and Card, 1989), we get the following dynamic linear model,

$$y_{i,t} = \sum_{k=0}^{t-3} \omega_{i,k} + \sum_{k=t-2}^{t} \theta_{t-k}\omega_{i,k} + \sum_{k=t-2}^{t} \phi_{t-k}\varepsilon_{i,k} \tag{3}$$

For each individual i, the dynamic linear model for their excess log income ($\boldsymbol{y}_i$) is a combination of the homogeneous parameters ($\boldsymbol{\theta}, \boldsymbol{\phi}$) and realized shocks ($\boldsymbol{\omega}_i, \boldsymbol{\varepsilon}_i$). In our Gibbs sampling model implementation, we take advantage of the fact that sampling new values of the homogeneous parameters conditional on fixed values of the realized shocks is relatively simple, and vice versa.

If we are given values of the realized shocks ($\boldsymbol{\omega}_i, \boldsymbol{\varepsilon}_i$), we can calculate the scalar $y^{\star}_{i,t}$ and the 1×7 vector $X_{i,t}$,

$$y^{\star}_{i,t} = y_{i,t} - \sum_{k=0}^{t-3} \omega_{i,k} \qquad\qquad X_{i,t} = (\omega_{i,t-2}, \omega_{i,t-1}, \omega_{i,t}, \varepsilon_{i,t-2}, \varepsilon_{i,t-1}, \varepsilon_{i,t})$$

Let $\boldsymbol{y}^{\star}$ be the $N(T-3) \times 1$ vector of all $y^{\star}_{i,t}$s across individuals i and time t, and let $\boldsymbol{X}$ be the $N(T-3) \times 7$ matrix whose rows are all $X_{i,t}$s across individuals and time points within individual. We can then write equation (3) as a simple linear regression model,

$$\boldsymbol{y}^{\star} = \boldsymbol{X} \cdot \boldsymbol{\beta} + \boldsymbol{e} \qquad \text{where} \quad \boldsymbol{e} \sim \text{Normal}(\boldsymbol{0}, \gamma^2 \cdot \boldsymbol{I})$$

where $\boldsymbol{\beta} = (\theta_2, \theta_1, \theta_0, \phi_2, \phi_1, \phi_0)$ are the homogeneous parameters of interest. We have also introduced an additional parameter γ^2 which is the variance of the residual error in this linear model. Under our Bayesian approach, we use non-informative prior distributions for both γ^2 and $\boldsymbol{\beta}$, which leads to the following simple posterior distributions:

$$\begin{aligned} \gamma^2 &\sim \text{Inv} - \text{Gamma}\left(\frac{TN}{2}\,,\, \frac{(\boldsymbol{y}^{\star} - \boldsymbol{X}\hat{\beta})'(\boldsymbol{y}^{\star} - \boldsymbol{X}\hat{\boldsymbol{\beta}})}{2}\right) \\ \boldsymbol{\beta} &\sim \text{Normal}\left(\hat{\boldsymbol{\beta}}\,,\, \gamma^2 \cdot (\boldsymbol{X}'\boldsymbol{X})^{-1}\right) \end{aligned} \tag{4}$$

where $\hat{\boldsymbol{\beta}} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y}^{\star}$ is the usual least-squares estimate of $\boldsymbol{\beta}$. We sample new values of γ^2 and $(\boldsymbol{\theta}, \boldsymbol{\phi})$ from the distributions in (4), but with the additional constraint that $\sum_k \phi_k = 1$. For a detailed review of the Bayesian approach to regression models, see Gelman, Carlin, Stern, and Rubin (1995).

5.2 Step 2: Sampling realized shocks $(\boldsymbol{\omega}, \boldsymbol{\varepsilon})$

If we are now given values of the homogeneous parameters $(\boldsymbol{\theta}, \boldsymbol{\phi})$, then the only unmeasured variables in our dynamic linear model (3) are the realized shocks $(\boldsymbol{\omega}_i, \boldsymbol{\varepsilon}_i)$. The Kalman filter (Kalman, 1960) is a popular approach for calculating maximum likelihood estimates of the realized shocks in our dynamic linear model. The maximum likelihood estimates from the Kalman filter can then be used to sample new values of the realized shocks $(\boldsymbol{\omega}_i, \boldsymbol{\varepsilon}_i)$, as outlined in (Carter and Kohn, 1994). Given the homogeneous parameters $(\boldsymbol{\theta}, \boldsymbol{\phi}, \gamma)$ and the collection of volatility parameters $(\boldsymbol{\sigma}^2, \boldsymbol{\tau}^2)$, each individual's income process is independent, which means that we can run the Kalman filter and sampling procedure for the realized shocks $(\boldsymbol{\omega}_i, \boldsymbol{\varepsilon}_i)$ for each individual i separately.

5.3 Step 3: Sampling volatility parameters $(\boldsymbol{\sigma}^2, \boldsymbol{\tau}^2)$

Our primary parameters of interest are the permanent and transitory variance $(\sigma^2_{i,t}, \tau^2_{i,t})$ for each person i in each year t. The information about $(\sigma^2_{i,t}, \tau^2_{i,t})$ comes from our sampled permanent and transitory shocks $(\omega_{i,t}, \varepsilon_{i,t})$ as well as the permanent and transitory variance parameters from other years and other people, which we denote $(\boldsymbol{\sigma}^2_{-(i,t)}, \boldsymbol{\tau}^2_{-(i,t)})$. They are linked through the posterior distribution of the volatility parameters,

$$p(\sigma^2_{i,t}, \tau^2_{i,t} | \omega_{i,t}, \varepsilon_{i,t}, \boldsymbol{\sigma}^2_{-(i,t)}, \boldsymbol{\tau}^2_{-(i,t)}) \propto p(\omega_{i,t}, \varepsilon_{i,t} | \sigma^2_{i,t}, \tau^2_{i,t}) \cdot p(\sigma^2_{i,t}, \tau^2_{i,t} | \boldsymbol{\sigma}^2_{-(i,t)}, \boldsymbol{\tau}^2_{-(i,t)}) \quad (5)$$

The first term of equation (5) comes from the likelihood of our realized shocks $(\omega_{i,t}, \varepsilon_{i,t})$ from our dynamic linear model,

$$p(\omega_{i,t}, \varepsilon_{i,t} | \sigma^2_{i,t}, \tau^2_{i,t}) \propto \left(\sigma^2_{i,t}\tau^2_{i,t}\right)^{-\frac{1}{2}} \exp\left(\frac{1}{2\sigma^2_{i,t}}\omega^2_{i,t} + \frac{1}{2\tau^2_{i,t}}\varepsilon^2_{i,t}\right) \quad (6)$$

The second term of equation (5) is our Markovian hierarchical Dirichlet process (MHDP) prior $p(\sigma^2_{i,t}, \tau^2_{i,t} | \boldsymbol{\sigma}^2_{-(i,t)}, \boldsymbol{\tau}^2_{-(i,t)})$ which consists of a weighted mixture of the other volatility parameters in the population. Sampling new values $(\sigma^2_{i,t}, \tau^2_{i,t})$ from the posterior distribution (5) is a multi-step process that acknowledges the structure of our population. First, we sample proposal values $\sigma^2_\star$ and $\tau^2_\star$ from a continuous distribution $f(\cdot)$. We will set $(\sigma^2_{i,t}, \tau^2_{i,t})$ equal to these proposal values only if we can not find a suitable set of values from among our population of other $(\boldsymbol{\sigma}^2_{-(i,t)}, \boldsymbol{\tau}^2_{-(i,t)})$ values across time within the same person and across other people.

5.3.1 Level 1: Is volatility unchanged from last year?

We first consider sampling $(\sigma^2_{i,t}, \tau^2_{i,t})$ to be the same as the previous year $(\sigma^2_{i,t-1}, \tau^2_{i,t-1})$,

$$p((\sigma^2_{i,t}, \tau^2_{i,t}) = (\sigma^2_{i,t-1}, \tau^2_{i,t-1})) \propto p(\omega_{i,t}, \varepsilon_{i,t} | \sigma^2_{i,t-1}, \tau^2_{i,t-1}) \quad (7)$$

$$p((\sigma^2_{i,t}, \tau^2_{i,t}) \neq (\sigma^2_{i,t-1}, \tau^2_{i,t-1})) \propto p(\omega_{i,t}, \varepsilon_{i,t} | \sigma^2_\star, \tau^2_\star) \quad (8)$$

This choice is made stochastically by flipping a weighted coin with weights equal to the probabilities in (7) and (8). Note that these probabilities are based on the likelihood of our realized shocks – equation (6) – given the set of candidate volatility parameters. If choice (7) is made, we are done and we set $(\sigma^2_{i,t}, \tau^2_{i,t}) = (\sigma^2_{i,t-1}, \tau^2_{i,t-1})$.

5.3.2 Level 2: Is volatility the same as in another year?

If choice (8) is made, we next consider all other (σ^2, τ^2) values across the other time points for only person i. Assume there are K unique values $(\sigma^2_{i,k}, \tau^2_{i,k})$ within the time series for person i. We next consider sampling $(\sigma^2_{i,t}, \tau^2_{i,t})$ equal to one of these values,

$$p((\sigma^2_{i,t}, \tau^2_{i,t}) = (\sigma^2_{i,k}, \tau^2_{i,k})) \propto n_k \cdot p(\omega_{i,t}, \varepsilon_{i,t} | \sigma^2_{i,k}, \tau^2_{i,k}) \quad k = 1, \ldots, K \tag{9}$$

$$p(\sigma^2_{i,t}, \tau^2_{i,t}) \neq \text{any } (\sigma^2_{i,k}, \tau^2_{i,k})) \propto p(\sigma^2_{i,t}, \tau^2_{i,t} | \sigma^2_{\star}, \tau^2_{\star}) \tag{10}$$

where n_k is the number of time-points within person i that currently have the values $(\sigma^2_{i,k}, \tau^2_{i,k})$ as their volatility parameters. These probabilities are still based on the likelihood of our realized shocks (6) given the set of candidate volatility parameters but with an additional weighting by n_k that allows popular or common values of volatility for person i to have a greater chance of being selected. The choice between (9) and (10) is again made stochastically by flipping a weighted coin with weights equal to the probabilities given above. If choice (9) is made, we are done and we set $(\sigma^2_{i,t}, \tau^2_{i,t}) = (\sigma^2_{i,k}, \tau^2_{i,k})$.

5.3.3 Level 3: Is volatility the same as another person's?

If choice (10) is made, we next consider all other (σ^2, τ^2) values from the population outside of person i. Assume there are L unique values (σ^2_l, τ^2_l) across all time points within all people in the population outside of person i. We next consider sampling $(\sigma^2_{i,t}, \tau^2_{i,t})$ equal to one of these values,

$$p((\sigma^2_{i,t}, \tau^2_{i,t}) = (\sigma^2_l, \tau^2_l)) \propto n_l \cdot p(\omega_{i,t}, \varepsilon_{i,t} | \sigma^2_l, \tau^2_l) \qquad l = 1, \ldots, L \tag{11}$$

$$p((\sigma^2_{i,t}, \tau^2_{i,t}) \neq \text{any } (\sigma^2_l, \tau^2_l)) \propto p(\omega_{i,t}, \varepsilon_{i,t} | \sigma^2_{\star}, \tau^2_{\star}) \tag{12}$$

where n_l is the number of time-points in the population of all people outside of person i that currently have the values (σ^2_l, τ^2_l) for their volatility parameters. These probabilities are still based on the likelihood of our realized shocks (6) given the set of candidate volatility parameters but with an additional weighting by n_l that allows popular or common values of volatility in the population to have a greater chance of being selected. The choice between (11) and (12) is again made stochastically by flipping a weighted coin with weights equal to the probabilities given above. If choice (11) is made, we are done and we set $(\sigma^2_{i,t}, \tau^2_{i,t}) = (\sigma^2_l, \tau^2_l)$. If choice (12) is made, we are still done but we set $(\sigma^2_{i,t}, \tau^2_{i,t}) = (\sigma^2_{\star}, \tau^2_{\star})$, which are values that have not yet been seen in the population.

The steps we have outlined above produce new sampled values for the permanent and transitory variance $(\sigma^2_{i,t}, \tau^2_{i,t})$ for year t of person i, and these steps must be

repeated for all other years and all other people in the population. The entire sampling procedure outlined in this section is designed to allow popular values of $(\sigma^2_{i,t}, \tau^2_{i,t})$ to be shared over time within an individual as well as across the population of individuals, while still allowing for new values to be introduced into the population if needed to effectively model the realized shocks sampled in Section 5.2.

This algorithm is programmed in Python and run on a cluster of computers. One run of this model (with 2,000 iterations) takes approximately one week, though multiple runs can be done simultaneously.

Table 5: Basic Model Results

Distribution of Variances

	Permanent Variance	Transitory Variance
Mean	0.020	0.098
St. Dev.	0.064	0.343
Observations	58,844	58,844
Minimum	0.013	0.037
1^{st} Percentile	0.014	0.038
5^{th} Percentile	0.014	0.038
10^{th} Percentile	0.014	0.039
25^{th} Percentile	0.014	0.039
50^{th} Percentile	0.015	0.040
75^{th} Percentile	0.015	0.042
90^{th} Percentile	0.016	0.058
95^{th} Percentile	0.024	0.200
99^{th} Percentile	0.089	1.757
Maximum	6.218	7.864

Distribution of posterior means

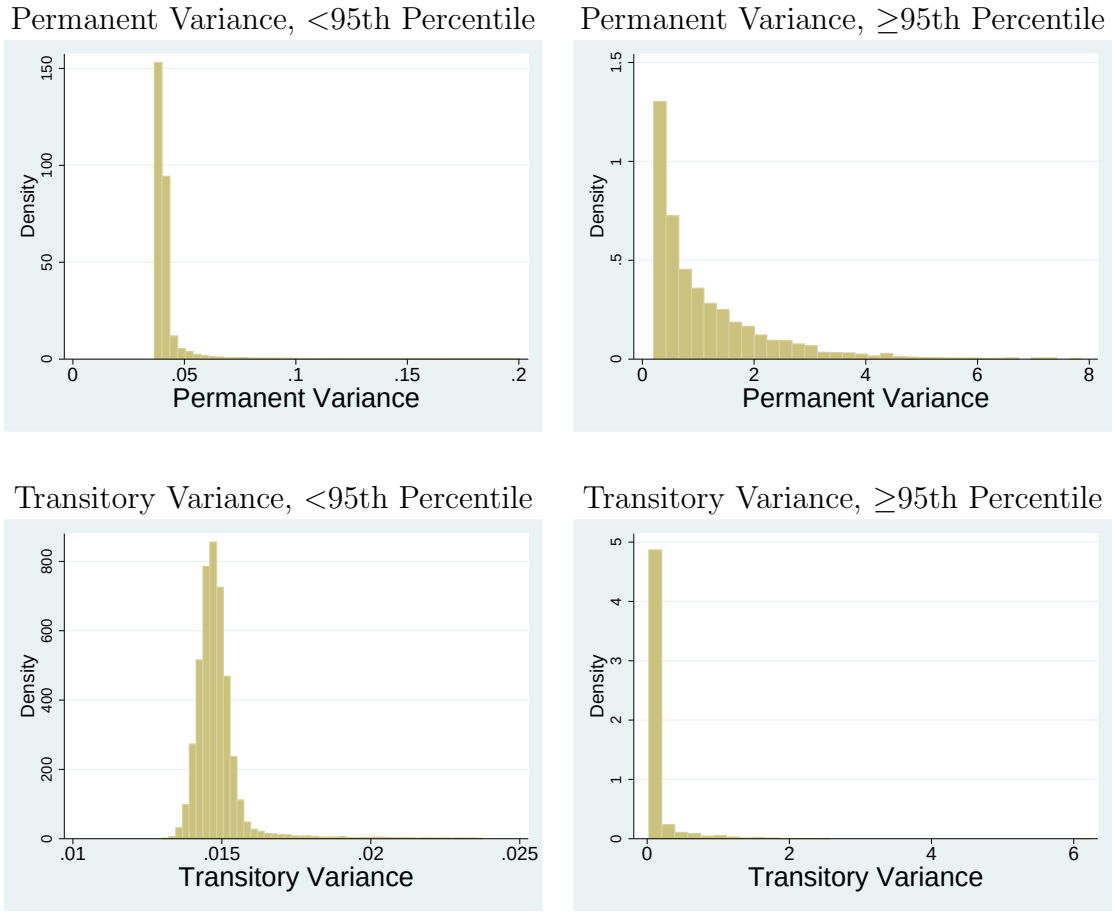
Shocks' Rate of Entry/Exit

lag, k	θ_k	ϕ_k
$k = 0$	0.9907	0.7915
(s.e.)	(0.0022)	(0.0005)
$k = 1$	1.3963	0.1521
(s.e.)	(0.0016)	(0.0004)
$k = 2$	1.3921	0.0564
(s.e.)	(0.0020)	(0.0004)

θ_k: impact of permanent shock from k periods ago
ϕ_k: impact of transitory shock from k periods ago

The left panel presents the estimates ofthe permanent and transitory variance distribution $(\boldsymbol{\sigma}^2, \boldsymbol{\tau}^2)$. These are the distribution of posterior means estimated from the data. These posteriors of the permanent variance and transitory variance are calculated for each individual in each year, as described in Section 5. The distributions presented here consider all years and all individuals together. The right panel of this table presents $\boldsymbol{\theta}$ and $\boldsymbol{\phi}$. $\boldsymbol{\theta}$ refers to the rate at which permanent shocks enter into income data, so that θ_k shows the impact of a permanent shock from k periods ago in current excess log income. These are assumed to be 1 for $k > \Omega$. $\boldsymbol{\phi}$ refers to the rate at which transitory shocks damp out of income data, so that ϕ_k shows the impact of a transitory shock from k periods ago in current excess log income. These are assumed to be 0 for $k > \epsilon$. The impact of these coefficients is detailed in equation 1.

Figure 2: Distribution of Permanent and Transitory Variance



This figure presents the distribution of $\boldsymbol{\tau}^2$ and $\boldsymbol{\sigma}^2$. These are the distribution of posterior means estimated from the data, as presented numerically in Table 5. These posteriors of the permanent variance and transitory variance are calculated for each individual in each year, as described in Section 5. The distributions presented here consider all years and all individuals together. The top panels present the distribution of the permanent variance, $\boldsymbol{\tau}^2$. The lower panels present the distribution of the transitory variance, $\boldsymbol{\sigma}^2$. The left panels present the distributions for all values less than the 95th percentile values. The right panels present the distributions for all values greater than or equal to the 95th percentile values. Data are too skewed for the distribution of either $\boldsymbol{\tau}^2$ or $\boldsymbol{\sigma}^2$ to be shown in a single picture.

6 Results

Here, we present the model parameters estimated using the methods from Section 5. While the chief object of interest is the evolution of the distribution of permanent and transitory volatility over time, $(\boldsymbol{\sigma}^2, \boldsymbol{\tau}^2)$, we begin with more basic results. In subsection 6.1, we present estimates of the homogeneous parameters that determine the rate at which shocks enter into and exit from income, $\boldsymbol{\omega}$ and $\boldsymbol{\varepsilon}$, as well as the overall distribution of volatility, $\boldsymbol{\tau}^2$ and $\boldsymbol{\sigma}^2$. In subsection 6.2, we show how the distribution of volatility in the population has evolved over time. Our main finding is that the increase in mean volatility can be attributed solely to an increase in volatility at the volatile tail; those who in other years would be expected to have large income changes are now likely to have even larger ones. In subsection 6.3, we present evidence on the time-series behavior of individual volatility parameters.

6.1 Basic results

Table 5 presents the basic parameter estimates obtained from fitting our model to the PSID income data described in Section 5. The left panel shows the distribution of risk in the population, $\boldsymbol{\tau}^2$ and $\boldsymbol{\sigma}^2$. Formally, we present the distribution of posterior means of permanent and transitory variance parameters. The right panel show estimates of the rate at which permanent shocks enter into income, $\boldsymbol{\theta}$, and the rate at which transitory ones exit, $\boldsymbol{\phi}$. While we estimate a distribution of volatility parameters, allowing them to vary over time and across individuals, we impose the constraint that $\boldsymbol{\theta}$ and $\boldsymbol{\phi}$ are constant.

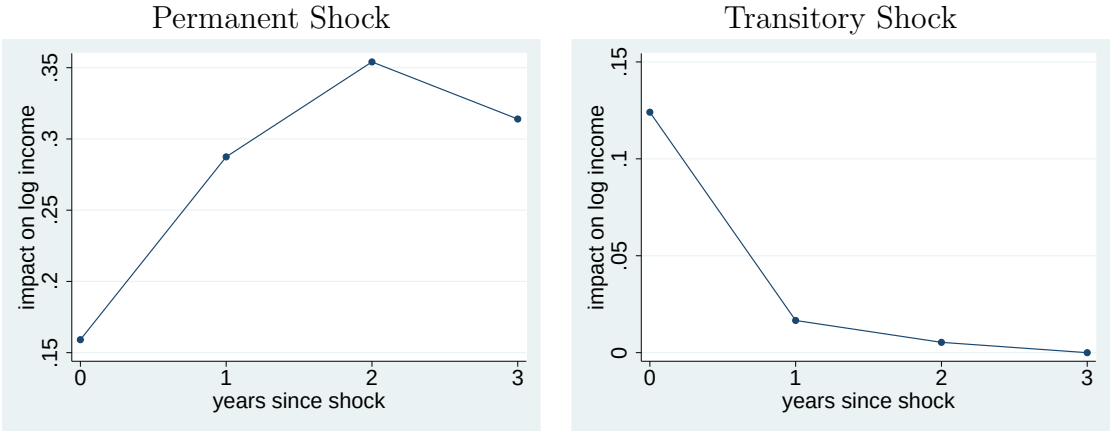
The main thing to note about the distribution of $\boldsymbol{\tau}^2$ and $\boldsymbol{\sigma}^2$ in the left panel of Table 5 is the extreme skew and kurtosis. While medians are modest, means exceed the 90^{th} percentile values. At the median, transitory shocks have a standard deviation of approximately 12% annually; permanent shocks have a standard deviation of approximately 20% annually. However, the highest volatility observations imply shocks with standard deviations well above 100% annually. Figure 2 plots these skewed and fat-tailed distributions by showing the right tail separately (right panels) from the rest of the distribution (left panels).

The main thing to note about estimates of $\boldsymbol{\theta}$ and $\boldsymbol{\phi}$, shown in the right panel of Table 5, is that permanent shocks enter in quickly (θ_k are close to one) while transitory shocks damp out quickly (ϕ_k fall to zero). The impact of a shock on the evolution of income is presented in Figure 3. These present impulse response functions for a permanent (left panel) and transitory (right panel) shock. Shocks were calibrated as a one standard-deviation shock for an individual with volatility parameters at the estimated means (pulled from Table 5).

6.2 Evolution of the volatility distribution

Here, we show how the distribution of posterior means of variance parameters has evolved over time. The evolution of the mean and median of the posterior means is shown in Figure 4. While the population median is constant over time (changes are

Figure 3: Impulse Response Function for Permanent and Transitory Shocks



This figure presents an estimated impulse response function for a permanent (left panel) and transitory (right panel) shock. Shocks were calibrated as a one standard-deviation shock for an individual with volatility parameters at the estimated means. This is a permanent shock of 0.314 (31.4 % of income, given log-linear approximation) and a transitory shock of 0.146 (14.6 % of income, given log-linear approximation).

small in magnitude and statistically indistinguishable from zero), the mean increases substantially. What explains this divergence?

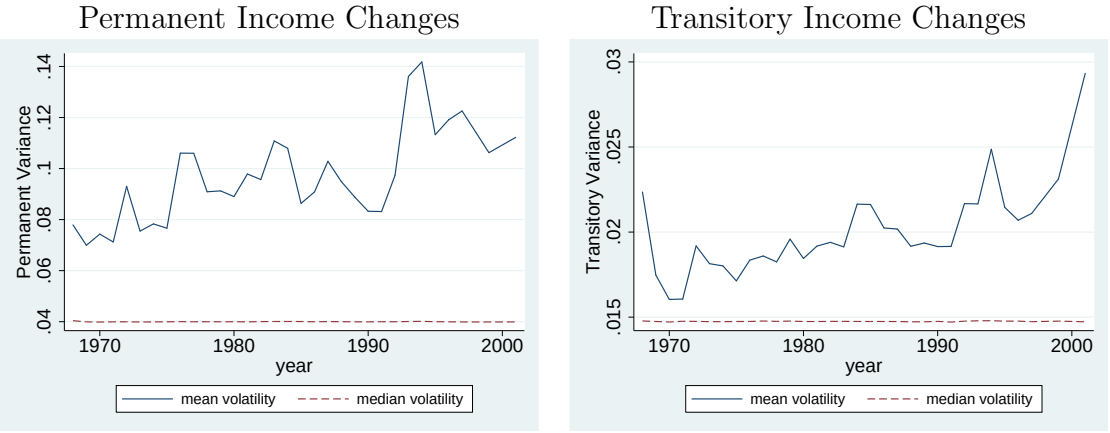
In the bottom panels of Figure 5, we plot the 1st, 5th, 10th, 25th, 50th, and 75th percentile values of the posterior mean of the permanent (left) and transitory (right) variance parameters by year. These are very stable and show no clear upward trend. While the 1st, 5th, 10th and 25th percentiles for the posterior means of the permanent variance have increased in a statistically significant way. The size of this increase is extremely small economically. It can account for almost none of the increase in the mean. In the case of the distribution of the posterior means of the transitory variance, there is no upward trend. Looking at all but the "risky" tail of the distributions, the distributions look very stable.

In the middle and upper panels of Figure 5, we show the evolution of the "risky" tail of the distribution of posterior means. In this case, variance parameters increase strongly and significantly. This increase the right tail of the distribution explains the increase in the mean completely.

6.3 Evolution of individuals' volatility parameters

This section is not yet completed. It is important for two reasons. First, serial dependence provides a test of heterogeneity. While our model assumes normality, the results we observe are consistent with a homogeneous but fat-tailed distribution. In this case, people with very large realizations would mistakenly be classified as high-risk. If an increase in fat-tails of risks common to everyone explained our

Figure 4: Evolution of Mean and Median Volatility



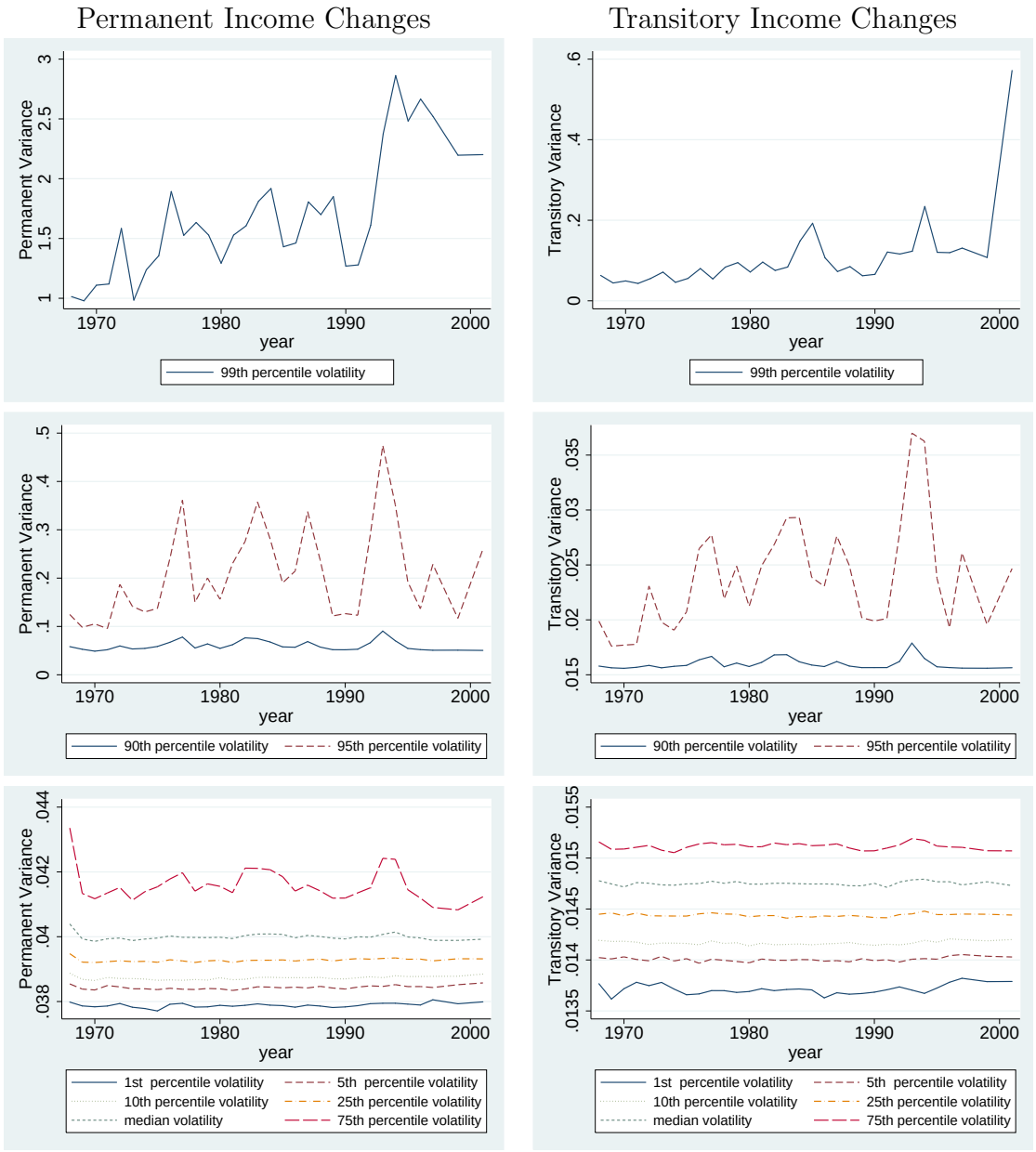
This figure presents an estimated impulse response function for a permanent (left panel) and transitory (right panel) shock. Shocks were calibrated as a one standard-deviation shock for an individual with volatility parameters at the estimated means.

findings, then serial dependence of posterior estimates of variance parameters should have fallen for those with high posterior values. Second, identifying the time-series properties of volatility estimates is a way of understanding and quantifying our model. If individuals change variance parameters often and at random, the notion of volatility as innate is undercut.

7 Conclusion

We have shown that increases in the size of income changes in the PSID can be attributed solely to the "right tail" of the volatility distribution. In other words, those who would have been high risk in past years are now even higher risk. Everyone else has had no substantial change in risk. This has important welfare implications. Those with the most income risk are presumably also the most risk tolerant; more risk averse individuals would want to select into occupations or life-paths with less risk. Since the increase in risk has been limited to those who are risk-tolerant or even risk-seeking, there will be much lower welfare costs of increased risk. There may even be welfare gains.

Figure 5: Evolution of Percentiles of Volatility Distribution



These figures show the evolution of various percentiles of the posterior mean of the permanent (left) and transitory (right) variance. This bottom panels present the evolution of the 1st, 5th, 10th, 25th, 50th, and 75th percentiles. The middle panels present the evolution of the 90th and 95th percentiles. The top panels present the evolution of the 99th percentiles.

References

Abowd, J., and D. Card (1989): "On the Covariance Structure of Earnings and Hours Changes," *Econometrica*, 57(2), 411–445.

Alvarez, J., M. Browning, and M. Ejrnaes (2001): "Modelling income processes with lots of heterogeneity," University of Copenhagen Working Paper.

Carroll, C. D., and A. A. Samwick (1997): "The Nature of Precautionary Wealth," *Journal of Monetary Economics*, 40(1), 41–71.

Carter, C., and R. Kohn (1994): "On Gibbs Sampling for State Space Models," *Biometrika*, 81, 541–553.

Dahl, M., T. DeLeire, and J. Schwabish (2007): "Trends in Earnings Variability Over the Past 20 Years," Working Paper: Congressional Budget Office:.

Dynan, K. E., D. W. Elmendorf, and D. E. Sichel (2007): "The Evolution of Earnings and Income Volatility," Working Paper, Federal Reserve Board.

Gelman, A., J. Carlin, H. Stern, and D. Rubin (1995): *Bayesian Data Analysis, Chapter 8*. Chapman and Hall/CRC, Boca Raton, FL.

Geman, S., and D. Geman (1984): "Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images," *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 6, 721–741.

Gottschalk, P., and R. Moffitt (1994): "The Growth of Earnings Instability in the U.S. Labor Market," *Brookings Papers on Economic Activity*, 0(2), 217–254.

——— (2002): "Trends in the Transitory Variance of Earnings in the United States," *Economic Journal*, 112, C68–73.

Hacker, J. (2006): *The Great Risk Shift: The Assault on American Jobs, Families, Health Care, and Retirement; And How You Can Fight Back*. Oxford University Press, Oxford.

Kalman, R. E. (1960): "A New Approach to Linear Filtering and Prediction Problems," *Transaction of the ASME—Journal of Basic Engineering*, 82, 35–45.

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

Muller, P., and F. A. Quintana (2004): "Nonparametric Bayesian Data Analysis," *Statistical Science*, 19, 95–110.

Schulhofer-Wohl, S. (2006): "Heterogeneity, Risk Sharing and the Welfare Costs of Risk," Working Paper.

Teh, Y., M. Jordan, M. Beal, and D. Blei (2007): "Hierarchical Dirichlet Processes," *Journal of the American Statistical Association*, 101(476), 1566–1581.