# Cyclical Labor Income Risk*

Makoto Nakajima† Vladimir Smirnyagin‡

April 23, 2020

**Abstract**

We investigate cyclicality of variance and skewness of household labor income risk in a unified estimation framework using PSID data. We make five findings. First, we find that head's labor income exhibits countercyclical variance and procyclical skewness. Second, cyclicality of hourly wage is muted, suggesting that head's labor income risk is mainly coming from volatility of hours. Third, the second earner lowers the cyclicality of both volatility and skewness of labor income risk. Fourth, government taxes and transfers reduce fluctuations of income risk even further, making it nearly invariant with respect to business cycle. Finally, among heads with strong labor market attachment, cyclicality of labor income volatility becomes weaker, while cyclicality of skewness remains.

**Keywords:** Labor income risk, income inequality, business cycles

**JEL:** D31, E24, E32, H31, J31

*We thank the Opportunity and Inclusive Growth Institute of the Federal Reserve Bank of Minneapolis, where this project was initiated. We thank seminar participants at the MMM 2018 (Nashville), LACEA 2018 (Guayaquil) and SED 2019 (St. Louis). The views expressed here do not necessarily reflect those of the Federal Reserve Bank of Philadelphia, the Federal Reserve Bank of Minneapolis, or the Federal Reserve System.

†Federal Reserve Bank of Philadelphia. E-mail: makoto.nakajima@gmail.com.

‡University of Minnesota and the Federal Reserve Bank of of Minneapolis. E-mail: smirn016@umn.edu.

# 1 Introduction

How does individual labor income risk change between economic expansions and contractions? What drives the cyclical nature of labor income risk—wages or hours? How effective are private (through the second earner) and public (through government taxes and transfers) insurance channels in stabilizing labor income risk? We address these questions by estimating time-varying second and third moments of earnings shocks for several income types by drawing on the Panel Study of Income Dynamics (PSID).

At the conceptual level, we systematically analyze cyclicality of risk for several income definitions within an unified estimation framework. We extend the econometric technique proposed by Storesletten, Telmer and Yaron (2004), which is designed to estimate the parameters of income shocks with a time-varying second moment (variance), by extending it to the case with a time-varying third moment (skewness). The list of income definitions we consider includes individual (head's) hourly wages and labor income, joint (spousal) labor income, and post-government (taxes and transfers) joint labor income. We also consider head's labor income for a subsample of heads with strong labor market attachment (Abowd and Card, 1989; Meghir and Pistaferri, 2004; Guvenen, Ozkan and Song, 2014), identified by labor income exceeding a certain income threshold. For the sake of convenience, hereafter we refer to the head's labor income defined this way as a *narrowly defined* head's labor income.

We make several findings. First, head's labor income exhibits countercyclical income risk, in the sense that both variance (countercyclical variance) and right skewness (procyclical skewness) increase in contractions. Second, head's hourly wage is less cyclical than head's labor income in both variance and skewness. This implies that changes in hours, possibly due to unemployment, are behind the cyclicality of head's labor income risk.

Third, we investigate the role of private (through labor income of the second earner) and public (through government taxes and transfers) insurance in mitigating the level and cyclicality of labor income risk. We find that existence of the second earner lowers the overall *level* of skewness of income risk, as well as its *cyclicality*. Effects on variance of labor income risk are limited. On the other hand, government taxes and transfers are found to lower the cyclicality of both variance and skewness. These results are found to have important implications for the probability of disasters: both private and public insurance channels mitigate by a factor of 2 the fall in labor income following a large (3 standard deviations) negative income shock in recessions.

Finally, we reconcile seemingly conflicting evidence between Storesletten, Telmer and

Yaron (2004), who find countercyclical variance of labor income risk, and Guvenen, Ozkan and Song (2014) and Busch, Domeij, Guvenen and Madera (2018), who find acyclical variance and procyclical skewness of labor income risk. We find that using narrowly-defined individual labor income can partially explain the differences between the two. When we look at narrowly-defined individual labor income, cyclicality of variance is found to be sizably weaker than that of head's labor income. Besides, government taxes and transfers also ameliorate cyclicality of variance. On the other hand, narrowly-defined head's labor income exhibits cyclicality of skewness almost as strongly as head's labor income.

We contribute to the literature by studying cyclicality of both variance and skewness of labor income risk for various definitions of labor income. Storesletten, Telmer and Yaron (2004) serves as the classic benchmark in the literature. They propose a novel estimation methodology of the countercyclical income risk, and find that standard deviation of labor income shocks is 80% higher in recessions than in expansions. We extend their study in two ways. First, we allow for time-varying skewness in addition to time-varying volatility. Second, we investigate 5 definitions of labor income, which allows us to study roles of various factors affecting the cyclicality of labor income risk, such as second earners and government taxes and transfers.

Deepening our understanding of labor income risk is important as labor income risk, together with market incompleteness, is found to be crucial for many important questions in macroeconomics. The literature that studies the role of labor income risk in macroeconomics goes back at least to Deaton (1997). Storesletten, Telmer and Yaron (2001) argue that cost of business cycles changes significantly if cyclical movement of labor income risk is taken into account. Kaplan and Violante (2010) find that the degree of consumption smoothing depends on the nature of labor income risk—whether it is persistent or transitory. Heathcote, Storesletten and Violante (2010b) explore the welfare implications of the increasing wage volatility in the U.S., and find that it benefits recent generations of workers as higher educational premium improves college attainment and redistributes labor within the household. Guvenen, Karahan, Ozkan and Song (2019) document a related empirical regularity: they show that in the data, income growth rate is very small for most individuals, while there is a considerable mass of people with very large growth rates. Therefore, high-order moments (kurtosis in this case) are important features of income growth distribution. In macro-finance literature, Constantinides and Duffie (1996) and Storesletten, Telmer and Yaron (2007) find that considering labor income risk could partially solve the risk premium puzzle.

The rest of paper is organized as follows. Sections 2 and 3 describe the data and lay out the estimation methodology. In Section 4, we estimate and analyze individual labor income risk with time-varying variance for various definitions of labor income. We extend the methodology to allow for time-varying skewness in Section 5. We provide economic interpretation of our results in Section 6. Section 7 concludes.

## 2 Data

**Background** We draw on the Panel Study of Income Dynamics (PSID) data, which is the longest publicly-available panel data on the U.S. population. PSID started in 1968 with more than 2,000 U.S. families being interviewed on a broad set of topics. The "split-off" families (when family members move out and establish their own households) are also interviewed. PSID spans the time period 1968-2014.

**Different Labor Income Types** The advantage of PSID for our purposes is the possibility to simultaneously observe several types of labor income. In particular, we consider 5 different definitions of labor income:

1. head's hourly wage[1],
2. head's labor income,
3. head's labor income (narrow definition),
4. joint labor income (head and spouse combined),
5. post-government (taxes and transfers) joint labor income.[2]

As it was mentioned in Section 1, *narrowly defined* head's labor income refers to those observations for which the labor income exceeds some minimum threshold. This is intended to capture individuals with a strong labor-market attachment. In particular, the income threshold is defined as half of an hourly minimum wage multiplied by 520 hours (13 weeks at 40 hours per week). We use the minimum wage data from the Federal Reserve Economic Data.[3] The post-government joint labor income is equal to the joint labor income (head

[1] Throughout the paper, we stick to PSID terminology and call a male earner (husband) a household's head, unless it is a family with a female being the only earner (in this case, wife is the head). A natural alternative is to mark the top-earner within the family as its "head"; this, however, will make our exercise not directly comparable to previous studies based on PSID data, and we, therefore, opt to use a conventional definition instead.

[2] In order to make income definition comparable, we normalize both joint and post-government labor incomes by way of dividing them by 2 (equal split between spouses).

[3] Around 5% of family heads per year earn below the minimum threshold (see Table A.3 in the Appendix).

and spouse's income combined) plus government transfers (unemployment compensation, disability insurance and alike) minus taxes (federal and state). PSID provides imputed values of taxes for a subset of years (1978-1990), but we opt to use as many years of data as possible and, therefore, use TAXSIM to obtain our own estimates of state and federal government tax liabilities for the sample of households. TAXSIM, however, is capable of computing the taxes starting from 1978—no state-level tax regulations are available prior to that year—which forces us to restrict the sample to years 1978-2014. Detailed explanations of how the variables have been constructed are relegated to Appendix A.1.

PSID is a survey data which suffers from well-known issues, such as top-coding of labor income, potential misreporting of income and small sample size. We, however, argue that PSID is an appropriate source of data for our exercise for several reasons. First, we are interested in several types of labor income, which necessitates the knowledge of labor income separately for each spouse, transfers received, wages and hours. This information is typically not simultaneously available in other datasets. Second, while the top-coding problem is particularly acute for studying income inequality (especially at the right tail of distribution), in this paper we are primarily interested in general swings of income risk over the business cycle. We, therefore, do not expect this issue to affect our results considerably.

Finally, there is a technical issue with changing frequency: years 1969-1995 are covered with annual frequency, and 1996-2014 with biannual frequency. As it will become clear in Section 3, it is straightforward to handle gaps in the data using our methodology.

The rest of this section—Subsections 2.1 and 2.2—discusses sample selection and identification of business cycles.

## 2.1 Sample Selection

We broadly follow the sample selection strategy of Storesletten, Telmer and Yaron (2004). In PSID, the object of analysis is a family unit (FU). We track heads of FUs as follows: if FU contains a married couple, then the husband is arbitrarily assigned to be the head. A woman can be the head only if husband is missing. In our analysis, we treat split-off families as new independent families: that is, when the head of the household changes, we record it as a new family unit.

Next, we apply a series of selection criteria to construct our dataset. First, a FU is in our sample as long as the age of the head is between 23 and 60 years old. By doing so, we pick only those households where the head is likely to have finished the educational phase

Table 1: Summary Statistics

| | Wage | Head's LI | Head's LI (narrow) | Joint LI | Post-Govt LI |
|---|---|---|---|---|---|
| Min | 0.0 | 0.0 | 1554.3 | 0.0 | 2352.3 |
| Max | 1080.5 | 588716.5 | 588716.5 | 294925.4 | 197278.9 |
| Median | 23.0 | 52532.5 | 54199.9 | 37623.5 | 36822.2 |
| Std | 23.1 | 45419.2 | 44490.5 | 27252.4 | 19428.5 |
| Bottom 5% | 0.0 | 3684.4 | 14954.3 | 10035.3 | 13074.1 |
| Top 5% | 61.7 | 578327.4 | 578327.4 | 294358.3 | 195537.5 |

*Notes*: Table 1 reports the summary statistics for the PSID-based sample spanning the years 1978-2014 (see Appendix A.2 for details on data construction). The description of income variables (wages, head's labor income, head's labor income (narrow), joint labor income, post-government joint labor income) is in Appendix A.1. All nominal variables were deflated by CPI with 2010 being the base year. The statistics were calculated using PSID sample weights.

of his or her life and entered the labor market.

Second, we drop all families with zero or negative total labor income in any year. We also drop families with extreme labor income growth rates[4] (Meghir and Pistaferri, 2004). Observations with top-coded values are also dropped.[5]

Finally, we drop families which are part of the Survey of Economic Opportunity or Latino subsample.[6] This leaves us with approximately 55,000 observations. Table 1 provides the summary statistics for the final dataset.

Appendix A.2 contains more details on the process of sample selection. Table A.1 reports the number of observations retained at each step of data preparation.

## 2.2 Identifying Business Cycles

The is no unique way to classify years into "expansions" and "contractions". Even though PSID is the longest available panel data, its time span covers few recessionary periods as defined by NBER.[7] It has become a working standard in the literature to classify years into stages of business cycle based on whether the growth rate of some macro aggregate was above or below the long-run mean in that particular year: for example, Storesletten, Telmer

[4] That is, we keep household $i$ as long as

$$\Delta \ln(y_{it}) \in \left(\frac{1}{20}, 20\right) \ \forall t.$$

[5] In order to ensure that our estimates are not affected by extremely large and very small labor incomes, we drop the top and bottom percentiles of labor income distribution.

[6] Such families have 1968 interview ID in the interval $[5000, 7000]$ (Survey of Economic Opportunity), or above 7000 (Latino subsample).

[7] Recessionary years according to NBER are 1970, 1974-75, 1981-82, 1991, 2001, 2008-09.

and Yaron (2004) used GNP per capita and Lee and Mukoyama (2015) used real GDP. We opt to use the real GNP per capita growth rate as a determinant of economic expansions and contractions in our estimation exercise, leaving the analysis of alternatives to Appendix B. One of the reasons we prefer GNP per capita growth rate is that it keeps our exercise close to Storesletten, Telmer and Yaron (2004) which is important for comparison purposes. Appendix A.7 shows that classifications based on GDP and GNP per capita yield comparable results.

# 3 Estimation Methodology

We follow the estimation methodology proposed by Storesletten, Telmer and Yaron (2004) for three reasons. First, this is a parsimonious way to estimate the time-varying variance, and the estimates reported by the aforementioned study serve as a natural reference point. Second, the parametric assumptions this methodology relies on help mitigate small-sample size issues which are typical for easy-to-access datasets. Third, this methodology can be extended to allow for a time-varying skewness (see Section 5.3). Finally, the methodology can easily accommodate the change in the frequency of PSID from annual to biennial in the middle of the sample period (from 1996 onwards). In what follows, we give a brief summary of the estimation methodology in Subsection 3.1, and then provide an identification argument for this method in Subsection 3.2.

## 3.1 Overview of the Method

Let $y_{it}^h$ be the natural logarithm of labor income of household $i$ of age $h$ in year $t$. We first project log labor income on a set of observables:

$$y_{it}^h = g(x_{it}^h, Y_t) + u_{it}^h, \tag{1}$$

where $x_{it}^h$ is the deterministic component of household-specific income attributable to age, education and family size. $Y_t$ is a measure of aggregate conditions at time $t$, which picks up the business cycle component of individual labor income.

The residual $u_{it}^h$ is a random component which under standard assumptions satisfies the orthogonality condition

$$\mathbb{E}(u_{it}^h | Y_t, x_{it}^h) = 0 \quad \forall t.$$

Intuitively, the residual captures variation in labor income which cannot be attributed to personal characteristics (such as differences in education), and is not explained by aggregate conditions (information contained in $Y_t$).

Next, a certain structure is imposed on the $\{u_{it}^h\}$ process. In particular, it is assumed that the idiosyncratic earnings component $u_{it}^h$ follows the process:

$$\begin{aligned} u_{it} &= \alpha_i + z_{it} + \varepsilon_{it} \\ z_{it} &= \rho z_{i,t-1} + \eta_{it}. \end{aligned} \tag{2}$$

Here $\alpha_i$ is a time-invariant fixed effect which household $i$ draws at the beginning of its labor market life. Next, $\varepsilon_{it}$ is a purely transitory component, while $z_{it}$ is a persistent earnings component which follows an AR(1) process. Random variables $\alpha_i, \varepsilon_{it}$ and $\eta_{it}$ are independent across time and space, and are drawn from some distributions characterized by the mean, variance and skewness:

$$\begin{aligned} \alpha_i &\sim F^\alpha(\mu_1^\alpha, [\mu_2^\alpha]^2, \mu_3^\alpha), \\ \varepsilon_{it} &\sim F^\varepsilon(\mu_1^\varepsilon, [\mu_2^\varepsilon]^2, \mu_3^\varepsilon), \\ \eta_{it} &\sim F^\eta(\mu_1^\eta, [\mu_{2,t}^\eta]^2, \mu_{3,t}^\eta). \end{aligned}$$

The means of the corresponding distributions are set equal to zero: $\mu_1^\alpha = \mu_1^\varepsilon = \mu_1^\eta = 0$.

The model is capable of picking up the time-varying labor income risk, since it allows both the variance and skewness of innovations to the persistent component $\eta_{it}$ to be a function of the aggregate state:

$$\mu_{2,t}^\eta = \begin{cases} \mu_2^E \text{ if expansion at } t \\ \mu_2^C \text{ if contraction at } t, \end{cases} \tag{3}$$

and

$$\mu_{3,t}^\eta = \begin{cases} \mu_3^E \text{ if expansion at } t \\ \mu_3^C \text{ if contraction at } t. \end{cases} \tag{4}$$

Therefore, there are 7 parameters to estimate in total:

$$\Theta = \{\rho, \mu_2^\alpha, \mu_2^\varepsilon, \mu_2^E, \mu_2^C, \mu_3^E, \mu_3^C\}.$$

We estimate $\Theta$ by the Generalized Method of Moments (GMM), using the moment conditions that relate the cross-sectional variance and skewness of estimated residuals $\widehat{u}_{it}^h$ with the history of shocks households experienced throughout their labor market life.[8] For the rest of this section, we restrict ourselves to a discussion of a time-varying volatility, and will get back to the case of cyclical skewness in Section 5.

Using an independence assumption, we can express the variance of a labor income shock of family $i$ with the head aged $h$ in year $t$ as:

$$\begin{aligned}
\text{Var}[u_{it}^h] &= \text{Var}[\alpha_i + z_{it} + \varepsilon_{it}] \\
&= [\mu_2^\alpha]^2 + [\mu_2^\varepsilon]^2 + \text{Var}[\rho z_{it-1} + \eta_{it}] \\
&= [\mu_2^\alpha]^2 + [\mu_2^\varepsilon]^2 + \sum_{j=0}^{h-1} \rho^{2j} \left[ I_{t-j}[\mu_2^E]^2 + (1 - I_{t-j})[\mu_2^C]^2 \right]. \qquad (5)
\end{aligned}$$

In Equation (5), $I_t$ is an indicator of an aggregate expansion in year $t$. The sample analog of the population moment (5) takes the form:

$$\frac{1}{N_{ht}} \sum_{i=1}^{N_{ht}} \left\{ [u_{it}^h]^2 - ([\mu_2^\alpha]^2 + [\mu_2^\varepsilon]^2) - \sum_{j=0}^{h-1} \rho^{2j} [I_{t-j}[\mu_2^E]^2 + (1 - I_{t-j})[\mu_2^C]^2] \right\} = 0. \qquad (6)$$

Here, $N_{ht}$ is the number of families at time $t$ with a head aged $h$. Note that $\mu_2^\alpha$ and $\mu_2^\varepsilon$ are not identified separately.[9] In this paper, we are not interested in either of those parameters separately and thus we estimate their sum $\mu_2^\alpha + \mu_2^\varepsilon$.[10]

There are $H \times T$ moments of type (6) in total, with $H$ denoting the number of different ages in the data, and $T$—the number of years.[11] Furthermore, we aggregate moment conditions so that the number of observations in any $\{H, T\}$ cell does not fall below 100. To accomplish this, we break down all feasible ages 23-60 into 4 age groups indexed by $h \in \{25, 35, 45, 55\}$, and make each group contain ages $\pm 5$ from the mean age within the group.[12] These adjustments help us balance the two opposing forces: on the one hand, the more moments conditions we use, the more information we extract from the data. On the other hand, more moment conditions lead to some (age, year) cells being too small. Fol-

[8] We assume that individuals enter labor market at the age of 23.

[9] Storesletten, Telmer and Yaron (2004) use additional moment conditions for the autocovariances of $u_{it}^h$ in order to disentangle these two parameters. Specifically, they obtain $\widehat{\mu_2^\alpha + \mu_2^\varepsilon} = 0.316$ when they do not disentangle the two, and $\widehat{\mu_2^\alpha} = 0.201$ and $\widehat{\mu_2^\varepsilon} = 0.123$ when the two variances are separately identified.

[10] As we find below, our estimates of $\mu_2^\alpha + \mu_2^\varepsilon$ are close to what Storesletten, Telmer and Yaron (2004) report.

[11] This amounts to $37 \times 38 = 1,406$ moment conditions.

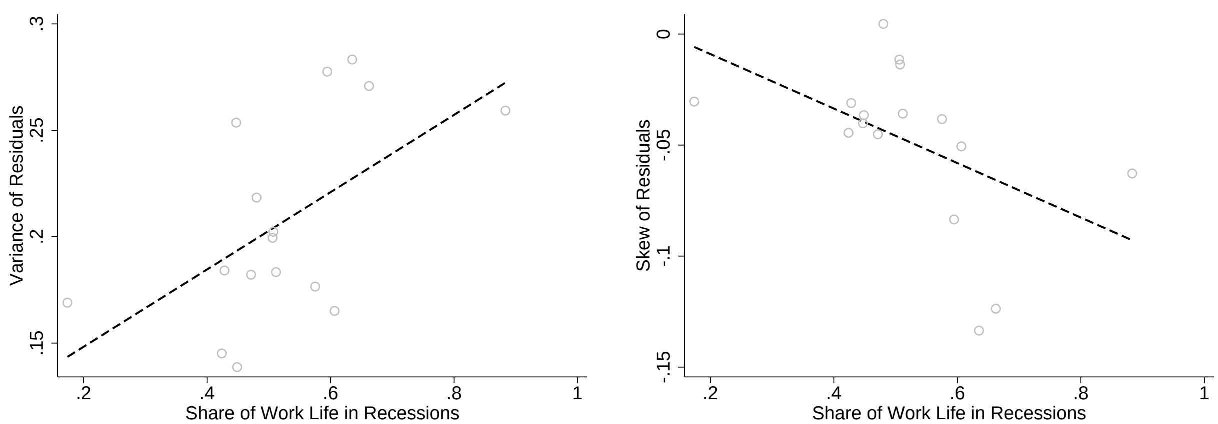
[12] Precise distribution of ages across 4 groups is as follows: group $h = 25$ contains ages 23-30, group $h = 35$ contains ages 31-40, group $h = 45$ encompasses ages $41 - 50$, and group $h = 55$ aggregates the remaining ages 51-60.

lowing Altonji and Segal (1996), we weight all moment conditions equally as it was shown that the identity weighting matrix dominates an asymptotically optimal weighting matrix in small samples.

## 3.2 Identification

The way estimation is set up in Section 3.1 highlights its benefits (see Equation 5): even though there are few families in the dataset whose working life we observe entirely—from the year when its head enters the labor market till the year when he or she retires—we can still incorporate the entire history of business cycle fluctuations that every household experienced over its lifetime into the estimation. In other words, the use of cross-sectional moments for identification allows us to exploit macroeconomic information that predates the micro panel, and thus include more business cycles in the analysis than covered by the sample.

Figure 1: Illustration of The Accumulation Argument




(A) Variance of residuals (B) Skew of residuals

*Notes*: Figure 1 is based on PSID data over the period 1978-2014. Panel A plots the variance of residuals $\widehat{u}$ against the share of working life spent in recessions. Panel B plots the skew (measured by the third central moment) of residuals. The statistics were aggregated within 4 age groups (23-30, 31-40, 41-50 and 51-60); the graphs depict the data corresponding to years 1980, 1990, 2000 and 2010.

The basic idea behind the entire approach is to exploit how the distribution of persistent idiosyncratic shocks accumulates over time: if the income process is persistent (values of $\rho$ are close to 1 in Equation (2)), then as a cohort ages, the cross-sectional income distribution at any age becomes a function of the sequence of shocks experienced by the cohort's members. If the variance of income shocks is higher in recessionary years than in expansionary ones, then a cohort that lived through more contractions will have a higher income variance at a given age than a cohort of the same age that lived through fewer contractions. Panel A in Figure 1 illustrates this intuition: it shows that the cross-sectional variance of $\widehat{u}$ tends to increase as the share of labor market life spent in recessions rises. Each circle corresponds to the variance of $\widehat{u}$ computed across households of a certain age in a specific year; the location of markers along the horizontal axis is determined by the share of working life a particular cohort spent in recessionary years. Figure A.4 in Appendix A.5 confirms that the aforementioned upward sloping relationship is present in all 4 age groups considered. The upward sloping pattern is more pronounced among younger cohorts, but the overall level of variance is higher for older households, reflecting their longer labor market history.

Our extension of that approach in Section 5.3 is based on the insight that a similar "accumulation" argument holds for skewness. If the probability of a large positive income shock is lower during an aggregate contraction, then the skewness of the shock in a recessionary period will be smaller (more negative) than in an expansion. Therefore, by way of comparing two cohorts of the same age, the distribution of residual income for the cohort that lived through more recessions will exhibit a smaller (more negative) cross-sectional skewness. Panel B in Figure 1 illustrates this logic: the skewness of income shocks decreases as the share of labor market life spent in recessions rises. Figure A.5 in Appendix A.6 additionally shows that this negative pattern is present in all 4 age groups.

# 4 Volatility of Idiosyncratic Labor Income Risk

In this section, we study how variance of labor income risk fluctuates across economic expansions and contractions. Conceptually, our exercise is reminiscent of Storesletten, Telmer and Yaron (2004) in that we estimate the same parameters using the same moment conditions, but we diverge from them in that we explore the nature of fluctuations in riskiness of several income definitions (Storesletten, Telmer and Yaron (2004) use joint labor income *after* transfers but *before* tax).

By studying different types of labor income, we are able to shed more light on the origins of income risk fluctuations. For example, by moving from hourly wage to head's labor income, we can speak to the quantitative importance of hours (including both intensive and extensive, i.e., employment and unemployment, adjustments) in shaping labor income risk. The intra-family insurance channel can be evaluated through the juxtaposition of risk between head's and joint (head and spouse) labor incomes. Finally, in order to quantify the role of government policy—including both taxes and transfers—in alleviating the cyclicality of labor income risk, we assess to what extent (pre-government) joint labor income is more volatile than post-government income.

We first conduct a graphical analysis in Subsection 4.1, before providing the estimates in Subsection 4.2.

## 4.1 Graphical Analysis

In order to shed light on the (counter)cyclical nature of idiosyncratic income shock volatility, we need to obtain the residuals $u_{it}^h$. We estimate Equation (1) by running a pooled regression.[13] The estimation results are provided in Table A.2 in Appendix A.3.

We consider the following specification of function $g(\cdot)$:

$$g(x_{it}^h, Y_t) = \theta_0 + \theta_1' \mathbf{D}(Y_t) + \theta_2' x_{it}^h,$$

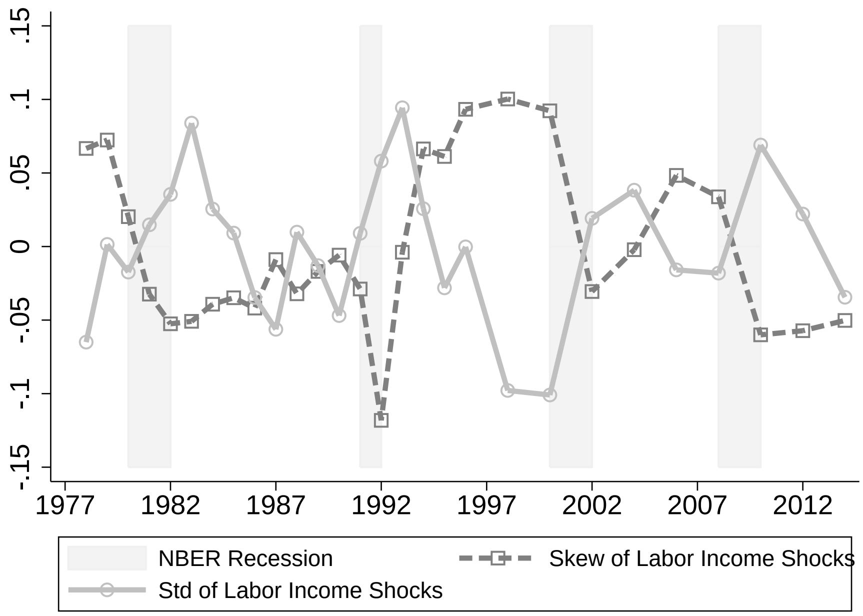
where $x_{it}^h$ includes the following list of observables: cubic polynomial in age, education of head, and the size of the family. Aggregate effects are absorbed by a full set of year fixed effects $\mathbf{D}(Y_t)$. Results in Table A.2 are broadly consistent with a wide body of literature: the earnings age profile is concave and increasing in education, large family sizes are associated with high labor incomes. All estimates are statistically significant and have the expected sign.

We subsequently retrieve $\widehat{u}_{it}^h$ as residuals from the estimated Equation (1). Figure 2 plots the time series of the second and third moments of $\widehat{u}_{it}^h$ for our sample years. The figure displays two strong patterns: the variance of labor income shocks is countercyclical, decreasing in expansions and increasing in recessions, while the skewness is strongly procyclical.

While Figure 2 indicates that the countercyclical income risk is a robust feature of the data, it is silent about the nature of this cyclicality. We, therefore, categorize every sample year into one of 3 bins, depending on the growth rate of real GNP per capita in that year: if GNP per capita grew by a lot (in top tercile of growth rate distribution), we place that year in bin 3. Conversely, if the growth was in a bottom tercile of the growth rate distribution, that year falls in bin 1. Table 2 reports the mean and median GNP growth rate for each tercile. Subsequently, we take the average (across years which are sorted in a particular bin) deviation of a corresponding statistic from its long-run mean, and do it for all 5 different income definitions (Figure 3).

[13] We also experimented with estimating a panel regression, but the results did not change significantly.

Figure 2: Second and Third Moments of $\widehat{u}_{it}^{h}$ over Sample Period




*Notes*: Figure 2 is based on PSID data over the period 1978-2014. The standard deviation of labor income shocks in year $t$ is a cross-sectional standard deviation of $\widehat{u}_{it}^{h}$—residuals from the estimated Equation (1). The skew of labor income shocks in year $t$ is a cross-sectional Kelley measure. We also subtract linear trends from the resulting series, which chiefly eliminates the long-run mean (the slope coefficient is nearly zero). Grey bars represent NBER recessions.

Table 2: Mean and Median GNP per Capita Growth Rate in Each Bin

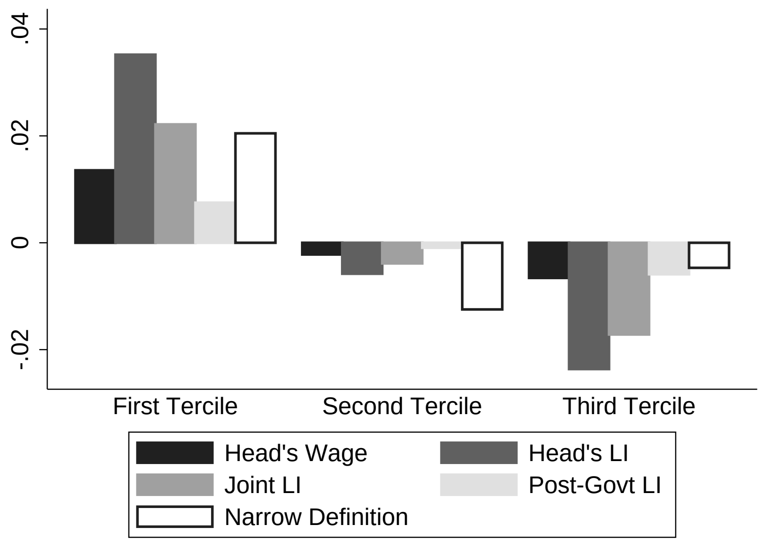
| | Tercile 1 | Tercile 2 | Tercile 3 |
|---|---|---|---|
| Mean growth rate, % | -2.8 | 1.9 | 6.5 |
| Median growth rate, % | -1.6 | 1.8 | 5.3 |

*Notes*: Table 2 reports the mean and median for each tercile of the distribution of GNP per capita growth rates.

Figure 3 anticipates several findings. At the very least, it shows that the volatility of income risk exhibits countercyclicality, and this pattern is robust across different labor income definitions. However, we can say more than that. First, consistently with findings of Guvenen, Ozkan and Song (2014), the narrowly-defined head's labor income exhibits relatively modest fluctuations in risk over the cycle. Second, head's wages are less volatile and less cyclical than head's labor income, pointing at the importance of hours in driving individual labor income risk fluctuations. Third, joint labor income, if anything, exhibits fluctuations in income risk which are smaller than those of the head's labor income. And, finally, cyclical patterns of post-government and narrowly-defined labor incomes are moderate.

These cyclical properties we observe here are robust to alternative definitions of the

Figure 3: Volatility of Idiosyncratic Income Risk by GNP per Capita Growth Tercile




*Notes*: Figure 3 is based on PSID data over the period 1978-2014. Each year from the period 1978-2014 is classified in one of 3 bins, depending on which tercile the growth rate of GNP per capita in that year falls into. Tercile 1 contains years with the lowest growth rate of GNP per capita, while tercile 3 contains years with the highest growth rates. The standard deviations shown are averages over years in each bin. Each tercile contains standard deviations for 5 measures of labor income: head's wage, head's labor income, joint labor income, post-government joint labor income, and head's labor income (narrow definition).

business cycle. In Appendix B, we provide analogous figures where we categorize years based on the NBER definition of recessions (B.1) and by real GDP growth rate (Section B.2).

## 4.2 Estimation Results

In Subsection 4.1, we provided suggestive evidence on the countercyclical nature of income shocks volatility. We now take a step forward and estimate the vector of structural parameters which govern the income process (2). As it has been discussed above, there are in total $H \times T$ moment conditions of type (6). In Section 3 we argued that we cannot use all of them, as the sample size of certain age-year cells becomes too small to obtain precise estimates. Instead, we focus on a subset of moment conditions which correspond to ages 25, 35,45 and 55. We check that there are at least 100 observations in each cell.

Table 3 provides GMM estimates for all 5 income definitions. Our results reconcile the findings of previous studies with seemingly conflicting results. On one hand, individual labor income exhibits a sizable countercyclical risk. Specifically, the estimated standard deviation

Table 3: GMM Estimation Results: Time-Varying Volatility

| | Head's wage | Head's LI | Joint LI | Post-govt LI | Narrow defn. |
|---|---|---|---|---|---|
| $\mu_2^E$ | $0.12^{***}$ | $0.12^{***}$ | $0.06^{***}$ | $0.05^{**}$ | $0.12^{***}$ |
| | (0.01) | (0.01) | (0.01) | (0.02) | (0.01) |
| $\mu_2^C$ | $0.11^{***}$ | $0.18^{***}$ | $0.11^{***}$ | $0.08^{***}$ | $0.12^{***}$ |
| | (0.01) | (0.01) | (0.01) | (0.03) | (0.01) |
| $\rho$ | $0.96^{***}$ | $0.96^{***}$ | $0.99^{***}$ | $0.90^{***}$ | $0.96^{***}$ |
| | (0.01) | (0.01) | (0.00) | (0.06) | (0.00) |
| $\sqrt{\mu_\alpha^2 + \mu_\varepsilon^2}$ | $0.52^{***}$ | $0.61^{***}$ | $0.48^{***}$ | $0.44^{***}$ | $0.54^{***}$ |
| | (0.01) | (0.01) | (0.01) | (0.01) | (0.01) |
| $\mu_2^C - \mu_2^E$ | -0.01 | 0.06 | 0.05 | 0.03 | 0.00 |

*Notes*: Table 3 reports the estimation results for $\Theta$ by GMM based on the moment conditions (6). Standard errors are in parentheses. *, **, *** denote statistical significance at 10, 5, and 1 percent levels, respectively.

is 0.12 in expansions and 0.18 in recessions. The ratio of our estimates (0.18/0.12=1.5) is somewhat lower than what Storesletten, Telmer and Yaron (2004) report (0.16/0.09=1.8), but still within the range of estimates they provide. This small difference might arise because of transfers: our definition of individual labor income corresponds to labor earnings before government transfers *and* taxes.[14]

On the other hand, the narrowly-defined head's labor income exhibits no countercyclicality of shock volatility. The ratio of the standard deviation in recessions and that in expansions is 1. This finding mirrors the results of Guvenen, Ozkan and Song (2014) who find that the second-moment of income risk is flat with respect to the business cycle. Remarkably, we obtain that result despite several differences in methodology, including the way we identified the income shock (residual from OLS regression, rather than income growth), the estimation approach (parametric, rather than non-parametric), and different data used (PSID vs. Social Security Administration records).

Wage rate also exhibits no countercyclicality, hinting towards an important quantitative role of hours (most likely, employment and unemployment). This finding mirrors the observation from Figure 3. Moving from head's labor income to joint labor income, we see that intra-family insurance channel through the added worker effect reduces the level of income risk both in expansions and recessions considerably. Finally, government taxes and transfers further mitigate both the level and cyclicality of income risk.

[14] Consistent with this logic, Heathcote, Perri and Violante (2010a) study the distributional effects of taxes and transfers and find that they compress the earnings inequality, especially at the bottom of the distribution.

# 5 Skewness of Idiosyncratic Labor Income Risk

In this section, we extend our analysis to allow for a time-varying skewness of income shocks. While countercyclical variance can tell us that tail events (large positive and negative shock realizations) become more likely during economic downturns, there is a growing body of literature highlighting the importance of the third moment (*inter alia,* Salgado, Guvenen and Bloom, 2016; Guvenen, Ozkan and Song, 2014; Busch, Domeij, Guvenen and Madera, 2018). Non-zero skewness implies that some extreme shock realizations are likely to be either positive or negative—depending on the sign of the coefficient of skewness. This also implies that constant skewness—something our analysis has implicitly assumed so far—can mask a rich heterogeneity between left- and right-tail events. We proceed as follows. First, in Subsection 5.1 we graphically visualize the presence of procyclical skewness for all 5 different definitions of labor income that we consider. In Subsection 5.2, we employ an alternative way of measuring income shocks, used by Guvenen, Ozkan and Song (2014), to further facilitate the comparison between our findings and theirs. Finally, in Subsection 5.3 we estimate the skewness of income shocks using GMM.

## 5.1 Graphical Analysis

Throughout our analysis of skewness, we consider the following two conventional measures:

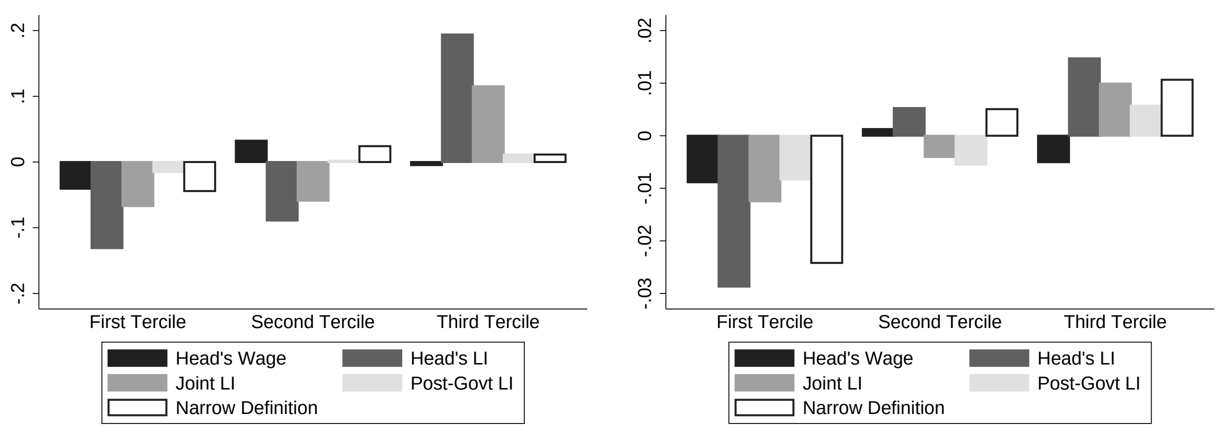
1. third central moment:
$$\mu_{th}^3 = \frac{\sum_i (\widehat{u}_{it}^h - \bar{\widehat{u}}_{it}^h)^3}{n_t},$$
2. Kelley measure:
$$\text{Kelley}_t = \frac{(P90_t - P50_t) - (P50_t - P10_t)}{P90_t - P10_t}.$$

The first statistic is a sample analog of the third central moment, while the second one—Kelley measure—is a function of several percentiles of $\widehat{u}_{it}^h$-distribution, which makes it robust to “extreme” observations (note that it is independent from the first and last deciles of the underlying distribution). Intuitively, Kelley measure computes the difference in inequality between the right ($P90 - P50$) and left ($P50 - P10$) tails, and relates it to the overall variation in the sample ($P90 - P10$). If the right tail is heavier than the left one (underlying distribution of $\widehat{u}_{it}^h$ has a positive skew), then the Kelley measure is positive. And the other way around, a heavier left tail makes the Kelley measure negative. Procyclical skewness implies that during economic upturns (downturns), the right (left) tail of income shocks thickens, leading to a disproportionate bigger fraction of large positive (negative) shocks. At the same time, the odds of receiving a large negative (positive) shock go down.

In order to shed light on how skewness of income shocks moves over the business cycle, in Figure 4 we plot the average skewness of income shocks for expansions and contractions defined in accordance with GNP per capita growth terciles. Panel A plots the third central moment, and Panel B plots the Kelley measure. The figure confirms that the skewness is procyclical, with Kelley measure exhibiting a somewhat stronger cyclical pattern. That means that during economic downturns, a large negative income shock is more likely than an equally large positive one.

A closer inspection of Figure 4 reveals that the skewness of post-government joint labor income barely changes over the cycle. It implies that the odds of getting a very negative shock for that income definition co-move with the odds of getting a very positive shock. Both measures of skewness rank post-government income as the one with the most stable skewness.

Figure 4: Skew of Idiosyncratic Labor Income Risk, by GNP per Capita Growth Tercile




(A) Third central moment

(B) Kelley measure

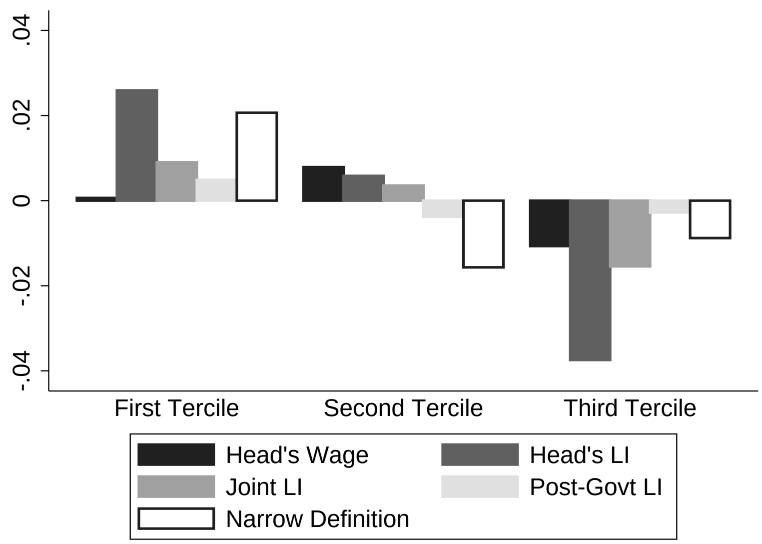
*Notes*: Figure 4 is based on PSID data over the period 1978-2014. Panel A plots the third central moment, Panel B plots the Kelley measure. Each year from the period 1978-2014 is classified in one of 3 bins, depending on which tercile the growth rate of GNP per capita in that year falls into. Tercile 1 contains years with the lowest growth rate of GNP per capita, while tercile 3 contains years with the highest growth rates. The measures of skewness shown are averages over years in each bin. Each tercile contains skewness measures for 5 different types of labor income: head's wage, head's labor income, joint labor income, post-government joint labor income and head's labor income (narrow definition).

When we look at household's income without transfers and taxes (joint labor income), we find that the cyclicality of skewness becomes stronger: the probability of getting a large negative income shock increases by more than the odds of getting a large positive shock during economic contractions. Head's labor income exhibits the strongest fluctuations in skewness. This observation might reflect intra-family insurance through an added worker effect: during economic downturns the probability of getting laid off increases, and spouse can step in and compensate for the head's job loss (by working more hours, getting an extra job, etc.). The head's wage exhibits relatively moderate fluctuations in skewness. The narrowly defined labor income shows sizable shifts in skewness.

Figures A.7 and A.9 in Appendix B display similar patterns and confirm that the above observations are broadly robust to alternative ways of business cycle identification.

## 5.2 Labor Income Risk As in Guvenen, Ozkan and Song (2014)

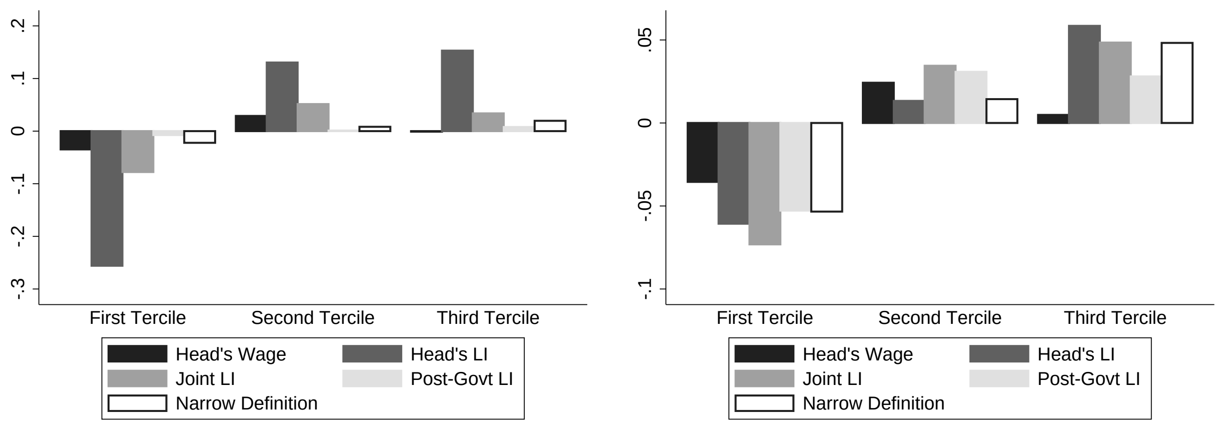
Figure 5: Labor Income Risk as in Guvenen et al. (2014): Volatility of Labor Income Risk by GNP per Capita Growth Tercile




*Notes*: Figure 5 is based on PSID data over the period 1978-2014. Each year from the period 1978-2014 is classified in one of 3 bins, depending on which tercile the growth rate of GNP per capita in that year falls into. Tercile 1 contains years with the lowest growth rate of GNP per capita, while tercile 3 contains years with the highest growth rates. The standard deviations shown are averages over years in each bin. Each tetcile contains standard deviations for 5 measures of labor income: head's wage, head's labor income, joint labor income, post-government joint labor income, and head's labor income (narrow definition).

This section explores the cyclical nature of shocks when those are identified as the growth rate of income (Guvenen, Ozkan and Song, 2014; Busch, Domeij, Guvenen and Madera, 2018). Their approach is non-parametric, and allows to study fluctuations in risk with few identifying assumptions.[15] While in this paper we opt to follow Storesletten, Telmer and Yaron (2004) and use the parametric approach given the size and changing frequency of the dataset, it is important to establish the connection between these two approaches. Taking into account the fact that PSID became biannual starting from 1996, we take the 2-year growth rate as our alternative measure of income shocks.[16]

Figure 6: Labor Income Risk as in Guvenen et al. (2014): Skew of Labor Income Risk, by GNP per capita Growth Tercile




(A) Third central moment

(B) Kelley measure

*Notes*: Figure 6 is based on PSID data over the period 1978-2014. Panel A plots the third central moment, Panel B plots the Kelley measure. Each year from the period 1978-2014 is classified in one of 3 bins, depending on which tercile the growth rate of GNP per capita in that year falls into. Tercile 1 contains years with the lowest growth rate of GNP per capita, while tercile 3 contains years with the highest growth rates. The measures of skewness shown are averages over years in each bin. Each tercile contains skewness measures for 5 different types of labor income: head's wage, head's labor income, joint labor income, post-government joint labor income and head's labor income (narrow definition).

First, regarding the cyclical nature of volatility of idiosyncratic labor income shocks, Figure 5 shows that the countercyclicality of the volatility carries over to this alternative way of shock identification. By way of comparing 5 definitions of labor income, one can observe that the head's labor income exhibits the strongest cyclical movement of volatility, while the narrowly-defined head's labor income fluctuates less. This is consistent with our benchmark result using the parametric approach. Moreover, post-government joint labor income exhibits little cyclicality, especially when compared with sizable cyclicality of individual labor income. This indicates the role public insurance plays in lowering income volatility.

[15] Guvenen, Ozkan and Song (2014) differentiate between the transitory and persistent components of income. *Transitory* component is measured as a 1-year growth rate ($\Delta \log y_{it}$). *Persistent* component is a 5-year growth rate ($\Delta_5 \log y_{it}$).

[16] Busch, Domeij, Guvenen and Madera (2018), who also use the PSID, use 1-year growth rate before 1996 and 2-year growth rate starting from 1996 to measure income shocks.

Second, Figure 6 confirms that procyclicality of skewness, measured as the third central moment (Panel A) and Kelley measure (Panel B), is also preserved when shocks are measured as growth rates of income. Both measures of skewness exhibit strong procyclicality, declining sharply during contractions. Overall, our key qualitative results are robust to the non-parametric way of income risk measurement; for various definitions of labor income, variance of income risk is countercyclical, while skewness is procyclical, but the cyclicality of variance is dampened when labor income is narrowly defined.

### 5.3 Joint Estimation of Cyclical Volatility and Skewness

Motivated by evidence from Subsection 5.1 that different income categories exhibit systematic procyclical skewness, in this section we extend the methodology laid out in Section 3 to handle the time-varying skewness of income innovations.

Conceptually, our insight is based on the accumulation argument we graphically shown in Figure 1 Panel B: the distribution of residual income for cohorts that lived through more recessions exhibits a smaller (more negative) cross-sectional skewness than cohorts which lived through fewer recessionary episodes. Therefore, we identify the time-varying skewness of income shocks by relating the skewness of labor incomes of cohorts with different aggregate histories.

Theoretically, assuming non-skewed distributions for $\alpha$ and $\varepsilon$, we can express the skewness of the residual labor income of cohort aged $h$ in year $t$ as

$$\mu_{3,t}^{h} = \sum_{j=0}^{h-1} \rho^{3j} \left[ I_{t-j}\mu_3^E + (1 - I_{t-j})\mu_3^C \right]. \tag{7}$$

where $\mu_3^E$ and $\mu_3^C$ are coefficients of skewness for the persistent innovation $\eta$ in expansions are recessions, respectively. The sample analog of (7) is given by

$$\frac{1}{N_{ht}} \sum_{i=1}^{N_{ht}} \left\{ \mu_{3,t}^{h} - \sum_{j=0}^{h-1} \rho^{3j} [I_{t-j}\mu_2^E + (1 - I_{t-j})\mu_2^C] \right\} = 0. \tag{8}$$

Table 4 reports the results of GMM estimation. There are several observations. First, the skewness appears to be procyclical, confirming graphical analysis in Subsection 5.1. Second, the skewness of wages is similar across expansions are recessions, while there is a large procyclical skewness in case of individual labor income; this points at the importance of hours worked—rather than wages—in accounting for a large increase in probability of labor

Table 4: GMM Estimation Results: Time-Varying Skewness

| | Head's wage | Head's LI | Joint LI | Post-govt LI | Narrow defn. |
|---|---|---|---|---|---|
| $\mu_3^E$ | -0.01*** | -0.09*** | -0.01*** | -0.00*** | -0.07*** |
| | (0.00) | (0.00) | (0.00) | (0.00) | (0.00) |
| $\mu_3^C$ | -0.01*** | -0.13*** | -0.03*** | -0.01*** | -0.10*** |
| | (0.00) | (0.00) | (0.00) | (0.00) | (0.00) |
| $\mu_3^C - \mu_3^E$ | 0.00 | 0.04 | 0.02 | 0.01 | 0.03 |

*Notes*: Table 4 reports the estimation results for $\Theta$ by GMM based on the moment conditions (8). Standard errors are in parentheses. *, **, *** denote statistical significance at 10, 5, and 1 percent levels, respectively.

income drops during economic downturns.

Furthermore, moving from individual to joint labor income, we find that the latter exhibits twice smaller movements in skewness over the cycle (0.04 and 0.02, respectively). Remarkably, the level of skewness is also substantially smaller in case of joint labor income. This observation signals about potentially large private (within family) insurance channels during adverse aggregate conditions.

Labor income after taxes and transfers shows very modest shifts in skewness (-0.01 in recessions and 0.00 in expansions), reflecting a quantitatively important role government policy plays in reducing the downside income risk during contractions. Finally, the narrowly defined labor income shows sizable movements in skewness (-0.10 in recessions and -0.07 in expansions); thus, it is not surprising that Guvenen, Ozkan and Song (2014) found procyclical skewness along with acyclical variance.

While the estimates of skewness and variance are suggestive, they are not directly informative about the income risks households face at different stages of the business cycle. The key reason is that both the volatility *and* skewness of income shocks simultaneously determine the probability of tail events. In the next section, we explore the economic implications of the reported estimates.

# 6 Economic Interpretation

In this section, we provide economic interpretations for our estimates. First, we argue that the skew normal distribution is a reasonable representation of income shocks (Subsection 6.1), and subsequently we graphically show how shock distributions change depending on the aggregate economic state (Subsection 6.2). Third, in Section 6.3, we quantify the changes in probabilities of positive and negative income shocks over the business cycle. Our results imply a substantial heterogeneity in probabilities of tail events across income types. Finally, in Section 6.4, we discuss how to interpret the magnitude of cyclical changes in labor income risk.

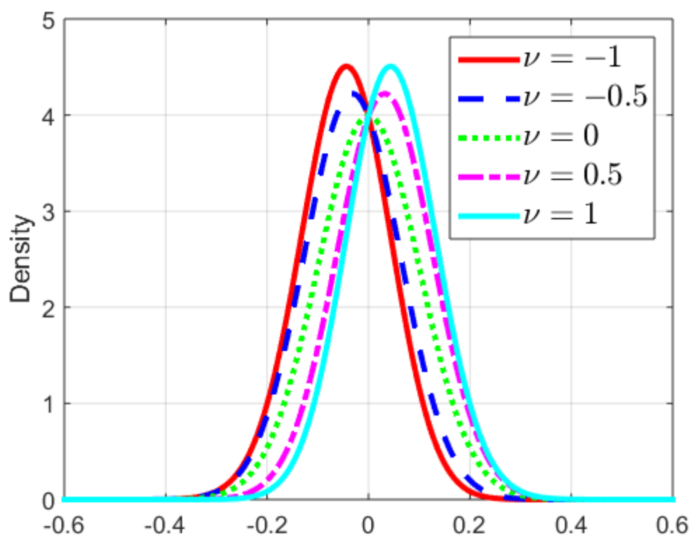
## 6.1 Skew Normal Distribution

We assume that innovations to a persistent component $\eta_{it}$ are drawn from a *skew normal distribution*, which is a generalization of a normal distribution to the case with a non-zero skewness:

$$\eta_{it} \sim \mathcal{SN}(\zeta, \omega_t, \nu_t). \tag{9}$$

The skew normal distribution is a family of probability distributions governed by 3 parameters: location ($\zeta \in \mathbb{R}$), scale ($\omega \in \mathbb{R}_{++}$), and shape ($\nu \in \mathbb{R}$).[17] We assume that the location parameter is business cycle invariant, and we normalize it to 0.

FIGURE 7: SKEW NORMAL DENSITY FOR DIFFERENT VALUES OF $\nu$




*Notes*: Figure 7 plots the skew normal density for 5 values of $\nu \in \{-1, -0.5, 0, 0.5, 1\}$. The shape parameter $\omega$ is 0.1, and location parameter $\zeta$ is set to 0.

Figure 7 shows how the shape parameter $\nu$ governs the skewness: the corresponding density tends to be skewed towards more positive values (positive skew) if $\nu > 0$, and towards negative values (negative skew) if $\nu < 0$.

[17]The p.d.f. of the skew normal distribution is $f(x) = 2\phi(x)\Phi(\nu x)$, where $\phi(x)$ and $\Phi(x)$ are p.d.f. and c.d.f. of the standard normal distribution, respectively. In case of symmetric distribution ($\nu = 0$), the formula collapses to a standard normal p.d.f.

We make the variance and skewness be state-dependent by allowing the shock structure to change between expansions and contractions. In particular, the support of both $\omega_t$ and $\nu_t$ consists of two points:

$$\omega_t = \begin{cases} \omega_E \text{ if expansion at } t \\ \omega_C \text{ if contraction at } t \end{cases}$$

and

$$\nu_t = \begin{cases} \nu_E \text{ if expansion at } t \\ \nu_C \text{ if contraction at } t. \end{cases}$$

Given our estimates of variance and skewness (Tables 3 and 4), we recover the vector of scale and shape parameters $\{\omega_E, \omega_C, \nu_E, \nu_C\}$ for each income category.[18]

Note that while the skew normal distribution has bounded skewness[19], at no point in our estimation step we imposed that restriction. Rather, we verify ex post that the skewness never goes outside that interval, and thus the skew normal distribution is a reasonable choice for the representation of persistent income shocks.

## 6.2 Implied Distributions of Income Shocks

In this subsection, we plot the implied distributions of persistent income shocks $\eta_{it}$ using the estimates from Tables 3 and 4. Figure 8 shows distributions of shocks $\eta_{it}$ in expansions (solid red lines) and in recessions (dashed blue lines), for 5 labor income definitions. The most striking difference between expansion and contraction distributions of income shocks is observed in case of the head's labor income (Panel B): while the solid red line (expansion) is fairly symmetric around 0 with a small variance, the blue dashed line (contraction) is substantially more dispersed with a heavy left tail. This observation implies that individuals are more likely to be hit by a large negative (rather than positive) shock during economic contractions. However, when the economy is expanding, tail events, both positive and negative, become substantially less likely.

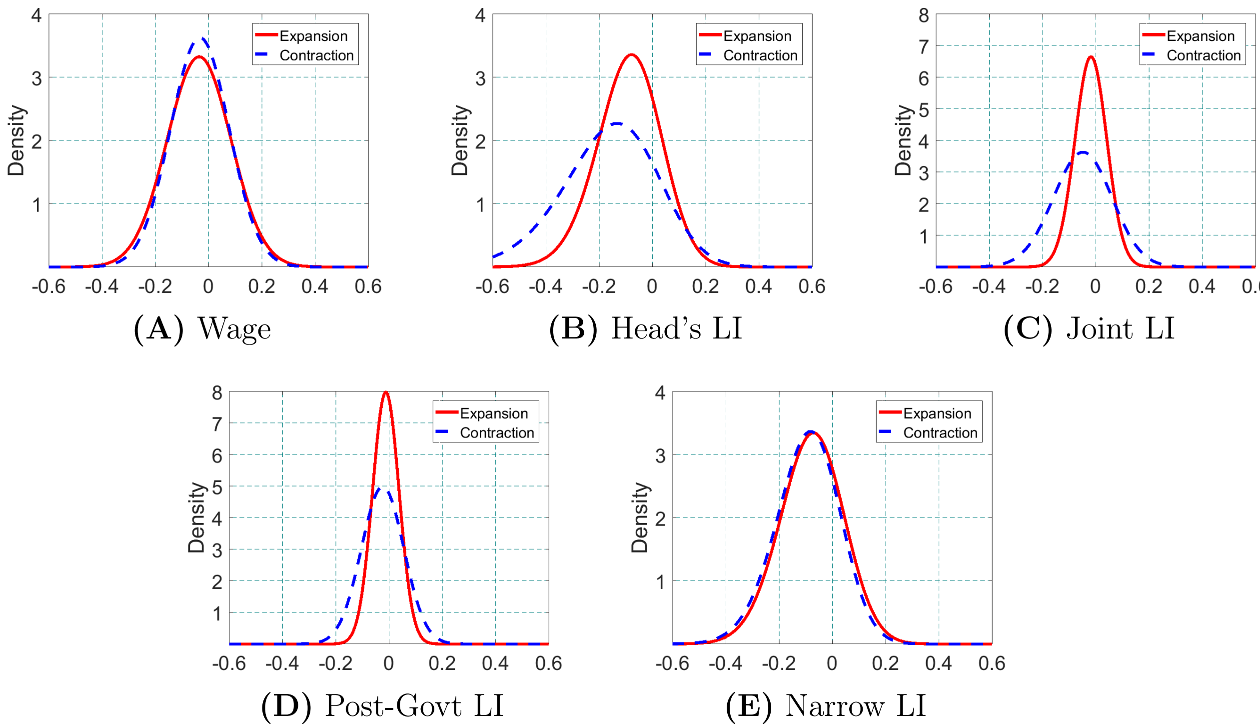
Panel A of Figure 8 confirms again that it is the number of hours worked rather than hourly wages that drives a significant portion of countercyclical head's labor income risk. In

[18] The variance of a skew normal random variable with parameters $\zeta, \omega$ and $\nu$ is $\omega^2\left[1-\frac{2\delta^2}{\pi}\right]$, where $\delta = \frac{\nu}{1+\nu^2}$. The skewness is given by $\frac{4-\pi}{2}\frac{\left(\delta\sqrt{2/\pi}\right)^3}{(1-2\delta^2/\pi)^{\frac{3}{2}}}$. Therefore, we recover parameters $\omega$ and $\nu$ by equating the estimated variance and skewness to these expressions, and simultaneously solving the resulting system of two equations.

[19] The skewness of a skew normal random variable lies in the interval $[-1, 1]$.

particular, the recessionary and expansionary distributions of wage shocks nearly coincide thus attributing all the fluctuations in labor income risk to hours.

Figure 8: Implied Distributions of Persistent Income Shocks: Expansions and Contractions




*Notes*: Figure 8 plots the estimated distributions of persistent income shocks $\eta_{it}$ for economic expansions (solid red) and contractions (dashed blue). The parameters of those distributions are taken from Tables 3 and 4.

The role of intra-family insurance channels is visible in Panel C of Figure 8—as compared to head's labor income, variance of shocks is substantially lower both in expansions and recessions, and probability of large negative shocks during downturns is reduced. Government transfers and taxes (Panel D) smooth out countercyclical risk and procyclical skewness significantly, even after the spousal channel has been taken into account. Finally, labor income risk for people with a strong market attachment (narrow definition) exhibits more moderate swings in volatility and skewness over the cycle as compared with individual labor income (Panels E vs. B).

## 6.3 Positive and Negative Income Shocks

In this subsection, we quantify the importance of income risk cyclicality by evaluating how the probability of positive and negative income shocks changes between recessions and expansions

Table 5: Change In Probability of Positive and Negative Income Shocks (Recessions vs. Expansions), %

| **Income Category** | $\Delta\%\mathbf{P}[\eta > 0]$ | $\Delta\%\mathbf{P}[\eta < 0]$ |
|---|---|---|
| Head's Wage | 0.0 | 0.0 |
| Head's LI | -21.9 | 4.8 |
| Joint LI | -14.4 | 8.8 |
| Post-Govt Joint LI | -5.8 | 3.9 |
| Head LI (narrow) | -15.5 | 5.4 |

*Notes*: Table 5 reports the percentage change in probabilities of positive and negative income shocks in recessions as compared to expansions. The table is based on the estimates reported in Tables 3 and 4.

for all 5 different income definitions.

Table 5 reports the percentage change in probabilities of positive and negative income shocks in recessions as compared to expansions using the implied skew normal distributions discussed in Subsection 6.1.

Several observations are in order. First, shocks to wages are drawn from similar across expansions and recessions distributions (Panels A in Figure 8): this results in insignificant changes in probabilities of both positive and negative shocks between booms and busts.

Other income types exhibit sizable procyclical skewness, which implies that probability of negative shocks increases in recessions and probability of positive shocks declines. Individual labor income is 22% less likely to increase during aggregate contractions, and at the same time is 5% more likely to go down. Joint and post-government labor incomes also exhibit asymmetry over the business cycle which is somewhat less pronounced as compared to head's labor income. Narrowly-defined labor income has a time-invariant but relatively high variance, which exacerbates changes in probabilities of positive and negative shocks over the cycle.

## 6.4 Evaluating Large Negative Events

The discussion so far was set in terms of *residual income*, but it is important to understand how large income shocks are relative to *labor income*. This subsection attempts to shed light on the quantitative relevance of income shocks.

In particular, we will be looking at the percentage change in labor income when household is hit by a large negative income shock. In order to avoid complicated simulations and to focus on the quantitative magnitude of income shocks, we assume that both persistent and transitory components of residual income ($z_{it}$ and $\varepsilon_{it}$) are zero, and only innovations to a persistent component $\eta_{it}$ are allowed to change. Furthermore, to mimic a large negative income shock, we set $\eta_{it} = -3\mu$, where $\mu$ is a standard deviation of the corresponding shock distribution (see Table 3).

Table 6 reports the results. For each income category, we calculate the percentage change in income following a large negative income shock which occurred in recession (first line) or in expansion (second line). In order to facilitate the comparison of different income categories, we focus on a prime age male (40 years old) with a spouse, who is a college graduate and has no children. We obtain labor income corresponding to such a household by substituting estimates from Table A.2 into Equation (1).

Table 6: Evaluating Magnitude of Income Shocks

| | Wage | Head's LI | Joint LI | Post-Govt LI | Head's LI (narrow) |
|---|---|---|---|---|---|
| Recession | -26% | -37% | -30% | -23% | -32% |
| Expansion | -26% | -30% | -21% | -19% | -28% |

*Notes*: Table 6 reports the percentage change in labor income following an extreme (3 standard deviations) negative labor income shock during aggregate expansion and recession. See text for more details.

Our results imply that the cyclicality of income shocks is quantitatively pronounced: a reduction in labor income following an extreme negative shock during aggregate expansion is sizably weaker as compared to when the shock hits in recession. In line with earlier discussion, the starkest differences are found in case of head's and joint labor incomes, while cyclical differences are weaker for wage and post-government labor income.

# 7 Conclusion

In this paper, we analyze volatility and skewness of labor income risk over the business cycle. We systematically apply our analysis to 5 definitions of labor income, which allows us to disentangle the role of hourly wage, hours, second earners, government taxes and transfers, and labor market attachment. We make several findings. First, for head's labor income, both variance and right skewness increase in contractions. Second, head's hourly wage is less cyclical than head's labor income in both variance and skewness, implying that changes in hours, possibly due to unemployment, are behind the cyclicality of head's labor income risk. Third, we find that existence of the second earner (private insurance) lowers the cyclicality of skewness, and also compresses volatility. On the other hand, government taxes and transfers (public insurance) are found to lower both the level and the cyclicality of labor income volatility. Both channels help mitigating increasing risk of negative tail events in recessions. Finally, among heads with strong labor market attachment, cyclicality of labor income volatility becomes much weaker, while cyclicality of skewness remains. This implies that whether one focuses on heads with strong labor market attachment or not helps reconcile seemingly conflicting evidence about cyclicality of labor income risk in the recent literature.

We see several fruitful avenues for future research. The first one is to quantify macroeconomic impact of countercyclical labor income risk using a quantitative macro model. Another potential line of research can explore the opposite direction of causality and identify macro shocks that generate the observed fluctuations in income risk. Finally, more research—both empirical and theoretical—is needed on the intra-family insurance channel (through the second earner), given that our results suggest its quantitative soundness.

# References

**Abowd, John and David Card**, "On the Covariance Structure of Earnings and Hours Changes," *Econometrica*, 1989, *57*, 411–445.

**Altonji, Joseph and Lewis Segal**, "Samll Sample Bias in GMM Estimation of Covariance Structures," *Journal of Business and Economic Statistics*, 1996, *14*, 353–366.

**Busch, Christopher, David Domeij, Fatih Guvenen, and Rocio Madera**, "Asymmetric Business Cycle Risk and Social Insurance," *manuscript*, 2018.

**Constantinides, George and Darrell Duffie**, "Asset Pricing with Heterogeneous Consumers," *Journal of Political Economy*, 1996, *104*, 219–240.

**Deaton, Angus**, *The Analysis of Household Surveys: A Microeconometric Approach to Development Policy*, The World Bank, 1997.

**Guvenen, Fatih, Fatih Karahan, Serdar Ozkan, and Jae Song**, "What Do Data on Millions of U.S. Workers Reveal about Life-Cycle Earnings Dynamics?," *manuscript*, 2019.

**_ , Serdar Ozkan, and Jae Song**, "The Nature of Countercyclical Income Risk," *Journal of Political Economy*, 2014, *122*, 621–660.

**Heathcote, Jonathan, Fabrizio Perri, and Giovanni L. Violante**, "Unequal We Stand: An Empirical Analysis of Economic Inequality In the United States 1967-2006," *Review of Economic Dynamics*, 2010, *13*, 15–51.

**_ , Kjetil Storesletten, and Giovanni L. Violante**, "The Macroeconomic Implications of Rising Wage Inequality in the United States," *Journal of Political Economy*, 2010, *118*, 681–722.

**Kaplan, Greg and Giovanni L. Violante**, "How Much Consumption Insurance beyond Self-Insurance?," *American Economic Journal: Macroeconomics*, 2010, *2*, 53–87.

**Lee, Yoonsoo and Toshihiko Mukoyama**, "Entry and Exit of Manufacturing Plants over the Business Cycle," *European Economic Review*, 2015, *77*, 20–27.

**Meghir, Costas and Luigi Pistaferri**, "Income Variance Dynamics and Heterogeneity," *Econometrica*, 2004, *72*, 1–32.

**Salgado, Sergio, Fatih Guvenen, and Nicholas Bloom**, "Skewed Business Cycles," *manuscript*, 2016.

**Storesletten, Kjetil, Christopher Telmer, and Amir Yaron**, "The Welfare Cost of Business Cycles Revisited: Finite Lives and Cyclical Variation in Idiosyncratic Risk," *European Economic Review*, 2001, *45*, 1311–1339.

**_ , _ , and _** , "Cyclical Dynamics in Idiosyncratic Labor Market Risk," *Journal of Political Economy*, 2004, *112*, 695–717.

**_ , _ , and _** , "Asset pricing with idiosyncratic risk and overlapping generations," *Review of Economic Dynamics*, 2007, *10*, 519–548.

# Table of Contents for Appendix

# A Data

In this section, we first describe the way we construct variables from PSID data (Subsection A.1), and then we discuss the process of sample selection (Subsection A.2).

## A.1 Definition of Variables

We break down all the variables into 2 categories: demographic (Section A.1.1) and income-related (Section A.1.2).

### A.1.1 Demographic and Socioeconomic Variables

- **Head** We identify current heads as those individuals within the family unit with Sequence Number equal to 1. In the PSID, the male is labeled as the household's head and the female as his spouse conditional on this family being full (married couple). A female is assigned to be the household's head only if she is unmarried. We select a new head in case of split-off families.

- **Age** Prior to 1996, PSID interviews were conducted annually (and biannually since then). However, the interview dates were not exactly a year apart, and, therefore, it could be the case that individuals report either the same age or numbers 2 years apart in consecutive waves. We create a consistent age variable by taking the age reported in the first year a particular individual appears in the survey, and add 1 to this variable in each subsequent year (2 for when the survey became biannual).

- **Education** This paper is focused on people in the labor market stage of their lives (age 23-60); hence, it is natural to assume that individuals are typically done with their education by the time they are first interviewed. Our measure of education is equal to the number of complete years of education. This variable is, however, not reported consistently over years (sometimes, only bracketed information is available). We, therefore, reconstruct this variable by taking the maximum number reported for each individuals over the years he/she was in the sample.

- **Family size** We found that this variable is consistently reported throughout the waves.

### A.1.2 Income Variables

- **Head's labor income** This is among the most consistently reported income-related variables in PSID, available throughout all waves. However, this variable is bracketed—and, therefore, useless for our analysis—in waves 1968 and 1969. The PSID reporting standards are wave-specific: if, for example, head's labor income in waves 1970-1982 is bounded by $99,999, in subsequent 10 waves (1983-1992) this variable is capped at $999,999. We make sure to drop all "capped" observations. Besides, another issue associated with this variable is that starting from 1994 wave, PSID stopped including labor portions of farm and business incomes into head's labor income. We correct for this by adding those income sources (labor part of Farm income and labor part of Business income) for waves 1994 onwards.

- **Wife's labor income** We follow similar (to the head's labor income) steps to construct this variable. The only difference is that PSID stopped including labor portions of farm and business incomes for spouses starting from wave 1993 (not 1994). We correct for that inconsistency.

- **Joint labor income** This variable is a sum of the head's and wife's labor incomes.

- **Post-government labor income** We add family-wide transfers and subtract taxes from the spousal labor income to get the labor-related portion of the family-wide disposable income. Transfers and taxes are reported poorly in PSID; we discuss them next.

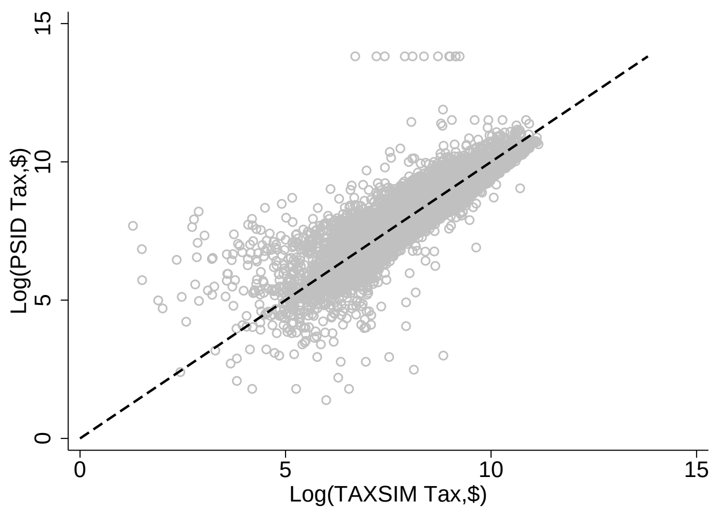
- **Transfers** Transfer data is reported inconsistently across the years. In general, we consider the old-age, survivor, and disability insurance (routinely abbreviated as OASDI), unemployment insurance (coming from the household's head, spouse and from other family members (OFUMs)), food stamps, as well as some other minor categories (bonuses, miscellaneous transfers, transfers received by OFUMS).

- **Taxes** Taxes are imputed by PSID analysts for waves 1978-1990. We want to use as many waves as possible, and so we need to construct a measure of household-wide taxes which will be consistent throughout the years. We use TAXSIM system to impute the federal- and state-level liabilities for individuals in our sample.[20] In particular, we feed

[20] TAXSIM is able to compute state taxes for years 1978 onwards; therefore, we drop years prior to 1978 in our final dataset.

in the following information: primary earner's (head) labor income, secondary earner's (spouse) labor income, state (needed to compute state-level tax liabilities) and number of dependents (family size minus 1 or 2 depending on the family composition). In order to check how consistent our measure of tax liabilities is with the tax data reported in PSID, we plot a scatter plot for these 2 measures pooled across years when both of them are available (1978-1990).

Figure A.1: PSID Tax Against the TAXSIM Data, 1978-1990




*Notes*: Figure A.1 depicts a scatter plot of PSID tax data and the tax liabilities imputed through the TAXSIM. Every circle represents a particular family-year pair, with years covered being 1978-1990.

Figure A.1 shows that the data imputed through the TAXSIM does a reasonable job, as the majority of observations lie on or close to the 45-degree line.

– **Head's labor income (narrow)** is obtained from the head's labor income data, when we drop observations which are below the minimum threshold. The threshold is defined as half of the current minimum hourly wage multiplied by 520 hours (13 weeks, 40 hours in each). The data on the minimum hourly wage comes from Federal Reserve Economic Data.[21]

## A.2 Sample Selection

Our baseline sample is based on PSID Family files and spans the time period 1969-2014. We use Individual Files to track individuals over time, and subsequently use information

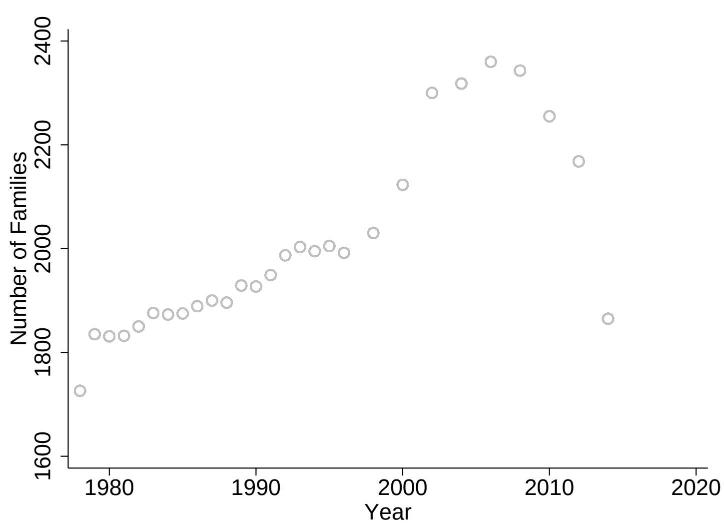
[21] https://fred.stlouisfed.org

from the Family Files to obtain family-wide variables. We track only heads of households; therefore, any "split-off" family we treat as a new family. In what follows, we provide a step-by-step algorithm of data preparation. Table A.1 shows the number of observations retained at each step of the preparation process.

1. We start off with downloading the 1968-2014 PSID files;
2. families which are part of SEO along with a Latino subsample are dropped;
3. only households' heads are tracked;
4. observations with missing or non-positive head's labor incomes are dropped;
5. observations with negative spousal labor incomes are dropped;
6. years prior to 1978 are dropped;
7. heads aged between 23 to 60 are considered;
8. trim the top and bottom 1% of the household labor income;
9. drop households with income growth anomalies (annual log growth rate must be between $\frac{1}{20}$ and 20).

Figure A.2 illustrates the number of families in the final sample across years.

Figure A.2: Number of Families in the Final Sample




*Notes*: Figure A.2 is based on PSID data over the period 1978-2014. Each dot indicates the number of families in a given year, after selection criteria outlined in Appendix A.2 have been applied.

Table A.1: Number of Observations Kept At Each Step

| **Step** | Observations Retained |
|---|---|
| Start | 3,011,697 |
| Only heads | 278,119 |
| No CEO, Latino | 98,478 |
| Working age | 78,148 |
| Missing/negative income | 78,144 |
| Years ≥1978 | 97,400 |
| No outliers | 62,903 |
| No growth anomalies | 54,744 |

*Notes*: Table A.1 reports the total number of remaining observations after each step of data preparation (see Appendix A.2).

## A.3 Estimation of Equation 1

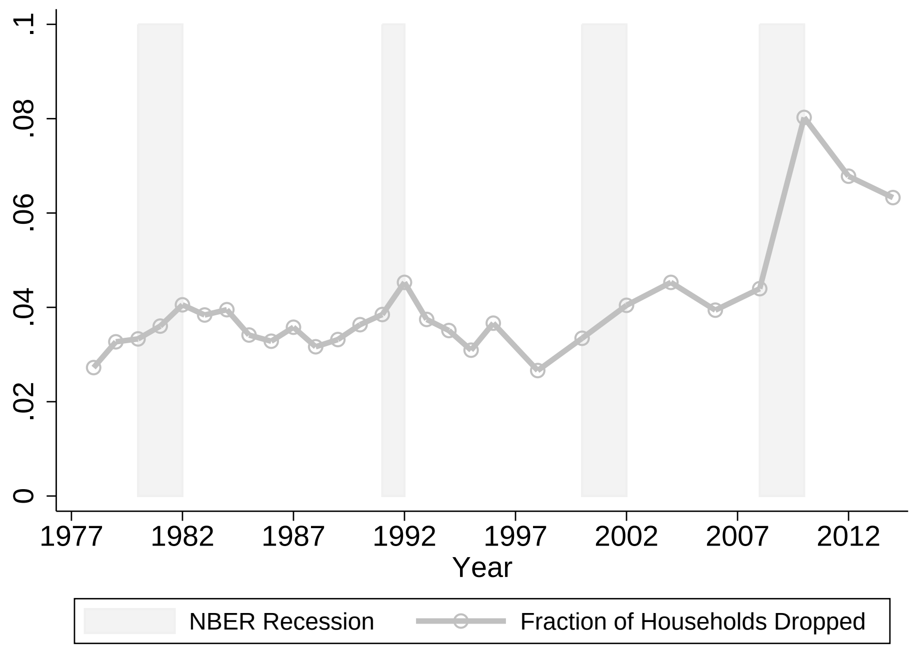
Table A.2: Estimation of Equation 1

| | Dependent variable $\log(y_{it})$ | | | | |
|---|---|---|---|---|---|
| | Post-govt LI | Joint LI | Head's LI | Wage | Head's LI (narrow) |
| Age | 0.1269*** | 0.1105*** | 0.1318*** | 0.1414*** | 0.1479*** |
| | (0.012) | (0.016) | (0.019) | (0.016) | (0.017) |
| Age$^2$ | -0.0020*** | -0.0012*** | -0.0019*** | -0.0024*** | -0.0023*** |
| | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| Age$^3$ | 0.0000*** | -0.0000 | 0.0000* | 0.0000*** | 0.0000*** |
| | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| log(Educ.) | 1.0392*** | 1.3021*** | 1.2959*** | 1.1498*** | 1.2667*** |
| | (0.013) | (0.018) | (0.020) | (0.017) | (0.018) |
| log(F. size) | -0.0326*** | -0.0786*** | 0.0430*** | 0.0286*** | 0.0497*** |
| | (0.004) | (0.006) | (0.007) | (0.006) | (0.006) |
| $N$ | 55675 | 55139 | 53705 | 51453 | 53432 |
| Adj. $R^2$ | 0.24 | 0.20 | 0.15 | 0.16 | 0.17 |

*Notes*: Table A.2 reports the results of OLS estimation, and is based on PSID data over the period 1978-2014. `Age` is the age of a household's head, `Education` is a number of completed (by the head) years of education. `Fam.size` is a number family members in a family unit. Standard errors are in parentheses. *, **, *** denote statistical significance at 10, 5, and 1 percent levels, respectively.

## A.4 Share of Heads with Weak Labor Market Attachment

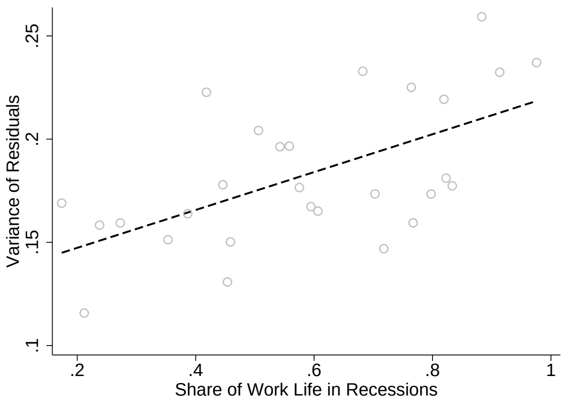
Figure A.3: Share of Heads with Weak Labor Market Attachment




*Notes*: Figure A.3 is based on PSID data over the period 1978-2014. The line plots the share of heads with weak labor market attachment (see Section 2 for details). Grey bars represent NBER recession dates.

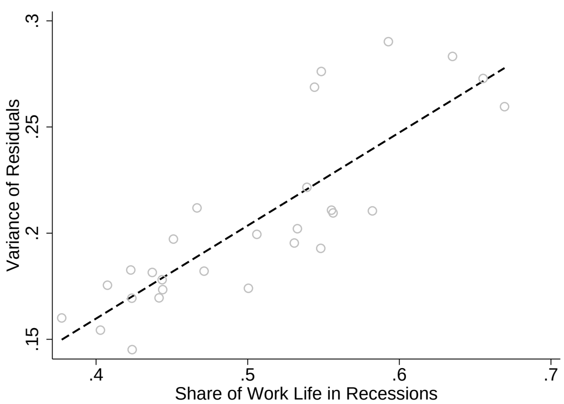
## A.5 Variance of Residuals as Function of Work History

Figure A.4: Variance of Residuals as Function of Work History




(a) Ages 23-29


(b) Ages 30-39



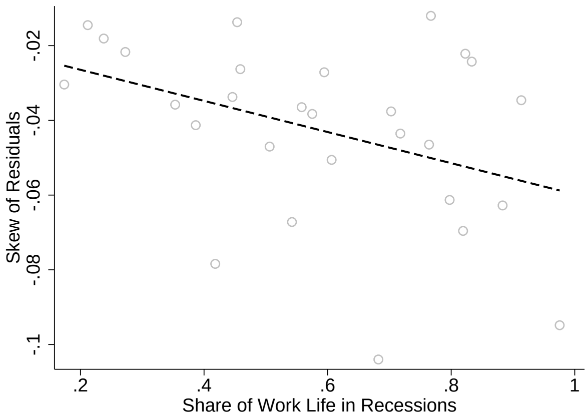

(c) Ages 40-49


(d) Ages 50-60

*Notes*: Figure A.4 plots the variance of residuals $\widehat{u}$ against the share of working life spent in recessions for 4 age groups: 23-29, 30-39, 40-49, and 50-60. The graphs depict the data corresponding to years 1980, 1990, 2000 and 2010. Share of work history spent in recessions is a continuous variable bounded between 0 and 1; it represents a share of (working) life each cohort spent in recessions.

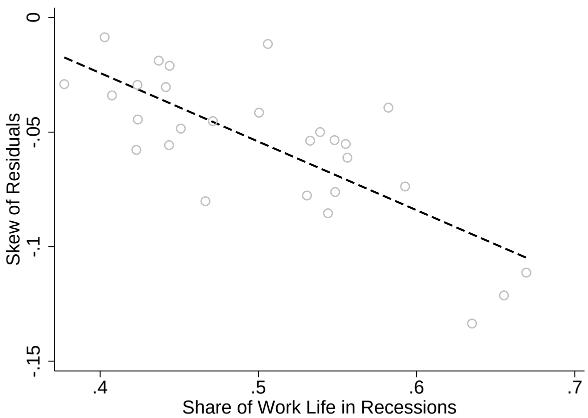
## A.6 Skew of Residuals as a Function of Work History

Figure A.5: Skew of Residuals as Function of Work History

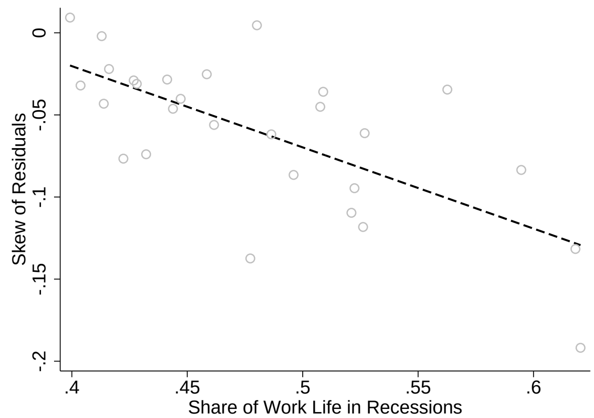



(a) Ages 23-29


(b) Ages 30-39




(c) Ages 40-49




(d) Ages 50-60

*Notes*: Figure A.5 plots the skewness (third central moment) of residuals $\widehat{u}$ against the share of working life spent in recessions for 4 age groups: 23-29, 30-39, 40-49, and 50-60. The graphs depict the data corresponding to years 1980, 1990, 2000 and 2010. Share of work history spent in recessions is a continuous variable bounded between 0 and 1; it represents a share of (working) life each cohort spent in recessions.

## A.7 Recessionary and Expansionary Years: A Comparison

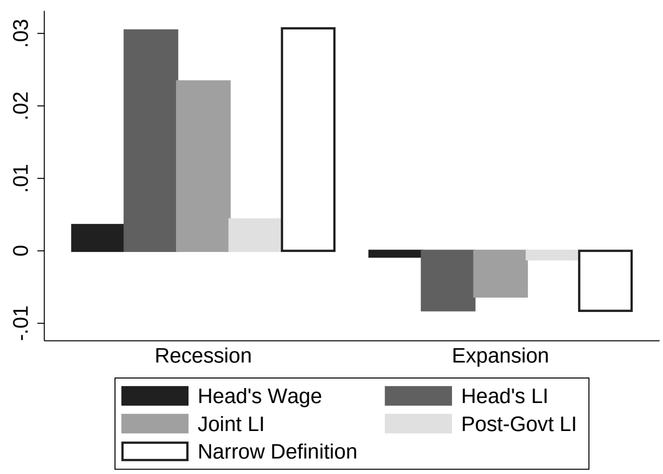
| Year | NBER | GDP | GNP per capita |
|---|---|---|---|
| 1978 | | | |
| 1979 | | ✓ | ✓ |
| 1980 | | ✓ | ✓ |
| 1981 | ✓ | ✓ | ✓ |
| 1982 | ✓ | ✓ | ✓ |
| 1983 | | | |
| 1984 | | | |
| 1985 | | | |
| 1986 | | | |
| 1987 | | | ✓ |
| 1988 | | | |
| 1989 | | ✓ | ✓ |
| 1990 | | ✓ | ✓ |
| 1991 | ✓ | ✓ | ✓ |
| 1992 | | ✓ | ✓ |
| 1993 | | | ✓ |
| 1994 | | | |
| 1995 | | ✓ | ✓ |
| 1996 | | ✓ | ✓ |
| 1997 | | | |
| 1998 | | | |
| 1999 | | | |
| 2000 | | | |
| 2001 | ✓ | ✓ | ✓ |
| 2002 | | ✓ | ✓ |
| 2003 | | ✓ | ✓ |
| 2004 | | | |
| 2005 | | | |
| 2006 | | | |
| 2007 | | ✓ | ✓ |
| 2008 | ✓ | ✓ | ✓ |
| 2009 | ✓ | ✓ | ✓ |
| 2010 | | ✓ | ✓ |
| 2011 | | | ✓ |
| 2012 | | | ✓ |
| 2013 | | | ✓ |
| 2014 | | | |

*Notes*: The table reports the classification of sample years (1978-2014) into recessions and expansions. Recessionary years are marked with a checkmark. The table provides 3 classification: based on real GDP, real GNP per capita, and NBER recessionary years.

# B Alternative Measures of Business Cycle

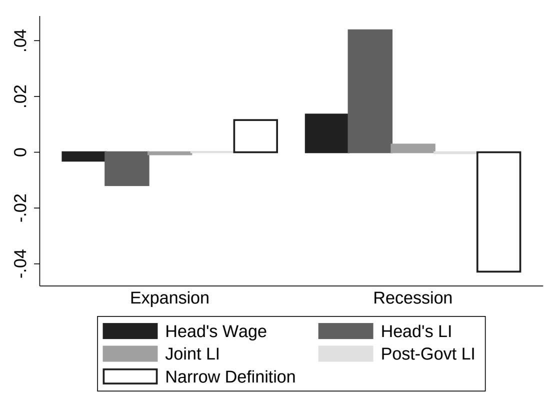
## B.1 NBER Recession Dates

Figure A.6: Volatility of Idiosyncratic Labor Income Risk by NBER Recession Dates

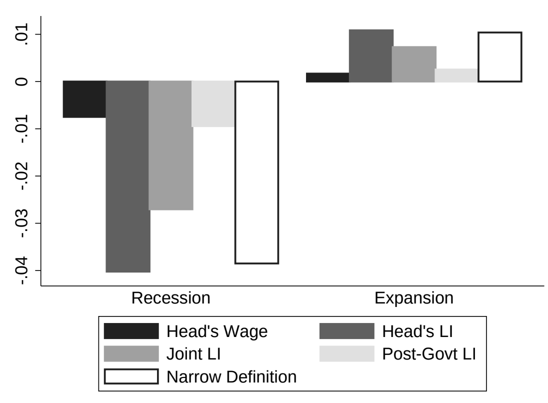



*Notes*: Figure A.6 is based on PSID data over the period 1978-2014. Each year from the period 1978-2014 is classified as a NBER recession or expansion. The standard deviations shown are averages over years in each bin. Each bin contains standard deviations for 5 measures of labor income: head's wage, head's labor income, joint labor income, post-government joint labor labor income, and head's labor income (narrow definition).

Figure A.7: Skew of Idiosyncratic Labor Income Risk, by NBER Recession Dates




(a) Third central moment

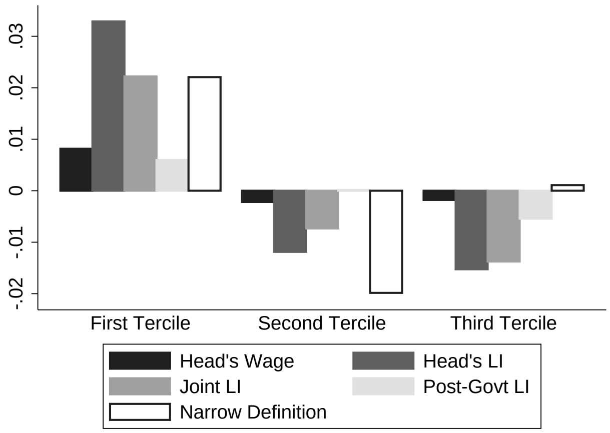



(b) Kelley measure

*Notes*: Figure A.7 is based on PSID data over the period 1978-2014. Panel A plots the third central moment, Panel B plots the Kelley measure. Each year from the period 1978-2014 is classified as a NBER recession or expansion. The measures of skewness shown are averages over years in each bin. Each bin contains skewness measures for 5 different types of labor income: head's wage, head's labor income, joint labor income, post-government joint labor labor income and head's labor income (narrow definition).

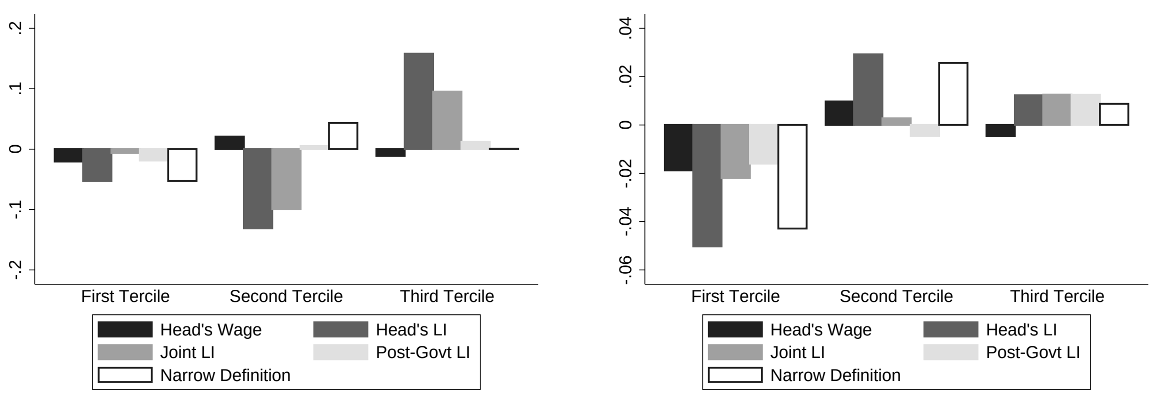
## B.2 Real GDP Growth

Figure A.8: Volatility of Idiosyncratic Labor Income Risk by Real GDP Growth Tercile




*Notes*: Figure A.8 is based on PSID data over the period 1978-2014. Each year from the period 1978-2014 is classified in of 3 bins, depending on which tercile the growth rate of real GDP in that year falls into. Tercile 1 contains years with the lowest growth rate of GDP, while tercile 3 contains years with the highest growth rates. The standard deviations shown are averages over years in each bin. Each tercile contains standard deviations for 5 measures of labor income: head's wage, head's labor income, joint labor income, post-government joint labor income, and head's labor income (narrow definition).

Figure A.9: Skew of Idiosyncratic Labor Income Risk, By Real GDP Growth Tercile




(a) Third central moment

(b) Kelley measure

*Notes*: Figure A.9 is based on PSID data over the period 1978-2014. Panel A plots the third central moment, Panel B plots the Kelley measure. Each year from the period 1978-2014 is classified in one of 3 bins, depending on which tercile the growth rate of real GDP in that year falls into. Tercile 1 contains years with the lowest growth rate of GDP, while tercile 3 contains years with the highest growth rates. The measures of skewness shown are averages over years in each bin. Each tercile contains skewness measures for 5 different types of labor income: head's wage, head's labor income, joint labor income, post-government joint labor income and head's labor income (narrow definition).