Higher-Order Income Risk Over the Business Cycle*

Christopher Busch†   Alexander Ludwig‡

May 14, 2021

Abstract

We extend the canonical income process with persistent and transitory risk to cyclical shock distributions with left-skewness and excess kurtosis. We estimate our income process by GMM for US household data. We find countercyclical variance and procyclical skewness of persistent shocks. All shock distributions are highly leptokurtic. The tax and transfer system reduces dispersion and left-skewness. We then show that in a standard incomplete-markets life-cycle model, first, higher-order risk has sizable welfare implications, which depend on risk attitudes; second, it matters quantitatively for the welfare costs of cyclical idiosyncratic risk; third, it has non-trivial implications for self-insurance against shocks.

Keywords: Idiosyncratic Income Risk, Cyclical Income Risk, Life-Cycle Model

J.E.L. classification codes: D31, E24, E32, H31, J31

---

*We thank Helge Braun for numerous helpful discussions as well as Chris Carroll, Russell Cooper, Johannes Gierlinger, Fatih Guvenen, Daniel Harenberg, Greg Kaplan, Fatih Karahan, Magne Mogstad, Serdar Ozkan, Luigi Pistaferri, Luis Rojas, Raül Santaeulàlia-Llopis, Kjetil Storesletten, and seminar and conference participants at various places for insightful comments. We thank Rocío Madera for sharing her code to set up the PSID. Chris Busch gratefully acknowledges financial support from the Spanish Ministry of Economy and Competitiveness, through the Severo Ochoa Programme for Centres of Excellence in R&D and from ERC Advanced Grant "APMPAL-HET". Alex Ludwig gratefully acknowledges financial support by the Research Center SAFE, funded by the State of Hessen initiative for research LOEWE and from NORFACE Dynamics of Inequality across the Life-Course (TRISP) grant: 462-16-120.

†Universitat Autònoma de Barcelona, MOVE, and Barcelona GSE; Edifici B, Campus UAB, 08193 Bellaterra, Spain; chris.busch@movebarcelona.eu; www.chrisbusch.eu

‡Goethe University Frankfurt; ICIR; CEPR; House of Finance, Theodor-W.-Adorno-Platz 3, 60629 Frankfurt am Main; Germany; alexander.ludwig@econ.uni-frankfurt.de; www.alexander-ludwig.com

1 Introduction

The extent of idiosyncratic income risk matters for many macroeconomic questions. The first contribution of this paper is a novel parametric approach to estimate idiosyncratic income risk and its systematic variation over the business cycle within the canonical transitory-persistent decomposition (dating back at least to Gottschalk and Moffitt 1994). In our estimation framework we transparently identify skewness and kurtosis of both transitory and persistent shocks, together with their variance. The second contribution is that we systematically evaluate economic consequences of this higher-order risk. We find that, first, higher-order idiosyncratic risk has (economically relevant) implications for welfare. Second, cyclical higher-order idiosyncratic risk matters for the welfare costs of business cycles. Third, higher-order idiosyncratic risk matters for self-insurance through savings. Our moment-based approach allows for a clean decomposition into the role of the variance, skewness, and kurtosis of the shock distributions. In our analysis, we further transparently show which properties of preferences are relevant to understand the documented economic consequences.

We provide guidance for the empirical and quantitative analysis by first investigating a simple two-period model, in which agents face risky second period income. We compare a version of the model with higher-order risk to one without higher-order risk, but with the same dispersion of risky income. We show analytically that, first, larger higher-order risk (in particular: left-skewness) can have positive welfare implications (with log-utility). Second and related, the reaction of precautionary savings to larger higher-order risk is ambiguous. The utility and behavioral implications crucially depend on risk attitudes of households—and on the magnitude of (higher-order) risk. These results hold generally for shock distributions with the given moments.

Estimation of Higher-Order Risk. We characterize both transitory and persistent shocks by their second to fourth central moments, which in the case of the persistent shocks we allow to be state-contingent. We estimate these distribution moments using the second to fourth cross-sectional central moments and co-moments of incomes—while similar estimations traditionally are based solely on the variance-covariance matrix. Importantly, we do not impose any parametric distribution functions and estimate the moments of the shocks by the Generalized Method of Moments (GMM). Identification follows from the fact that the accumulated second to fourth central moments systematically differ across cohorts if these cohorts experience different macroeconomic histories—if the moments of shocks differ systematically over the business cycle. This identification idea was introduced in Storesletten et al. (2004) for the second moment and our extended estimator nests theirs as a special case. It is important to note that we include the third and fourth central moments in a way that does not affect the identification of the second moments or the persistence of the shocks: we proceed 'step-by-step', and first identify the second moments and persistence using only the variance-covariance moment conditions. We then hold persistence and second moments fixed and use the additional moment conditions only to identify the third and fourth central moments of the shocks.

While Storesletten et al. (2004) analyze household-level income including government transfers from the Panel Study of Income Dynamics (PSID) and find *countercyclical variance*[1] of persistent shocks, more recent evidence in Guvenen et al. (2014) suggests that the focus on the variance of log income changes alone misses the main characteristics of how individual risk varies with the aggregate state of the economy. Their findings based on administrative social security data (SSA) for individual males in the United States suggest that individual downside risk is larger in a contraction, while upside risk is smaller—this is reflected in a more pronounced left-skewness of the distribution of earnings changes. Related, Busch et al. (2020) conduct a non-parametric analysis of individual and household earnings dynamics in Germany, Sweden,

[1] This terminology has been introduced in the macroeconomic asset pricing literature, see Mankiw (1986), Constantinides and Duffie (1996), and Storesletten et al. (2007). Building on the conceptual framework of Storesletten et al. (2004), Bayer and Juessen (2012) focus on residual hourly wages (at the household level) and based on PSID data estimate countercyclical dispersion of persistent shocks in the United States.

France, and the US. They also find *procyclical skewness* of individual and household-level annual earnings changes.

Our estimation approach allows us to draw a richer image of income dynamics over the business cycle within the transitory-persistent framework and to thus bridge the previous analyses. Taking into account the second moment alone might lead to wrong conclusions if the change of the distribution is asymmetric, which is captured by a change of the third moment.[2]

We apply the estimation to survey data from the PSID, which allows us to control for a rich set of household-level information and to take into account several relevant transfer components. While being a smaller sample compared to administrative SSA data, it allows us to analyze features of earnings dynamics at the *household* level. Busch et al. (2020) show that the cyclical features of earnings changes at the individual level documented in Guvenen et al. (2014) are well reflected in the PSID. Also, De Nardi et al. (2020) show that many recently documented richer features of individual earnings dynamics carry over to the PSID.[3]

We estimate two separate income processes at the household level: one for joint labor income, and one for income after taxes and transfers. Comparison of the corresponding estimates is informative about the success of the existing tax and transfer scheme to dampen risk and its cyclicality. We find that both transitory and persistent shocks to pre-government earnings feature strong left-skewness, and that persistent shocks are significantly cyclical: in contractions, their distribution is more dispersed and more left-skewed. We also find that the existing tax and transfer system insures against both types of income shocks. The distribution of both shocks to post-government income (after taxes and

[2]As discussed in Busch et al. (2020), if the lower tail of a distribution expands by more than the upper tail collapses, then the distribution is more dispersed (an increase in the second moment) and more skewed to the left (a drop of the third moment).

[3]In follow-up work to Guvenen et al. (2014), Guvenen et al. (2016) document that, in a given year, most individuals experience very small earnings changes, while some workers experience very large changes of their earnings. This is summarized by a high kurtosis—relative to what the conventional assumption of log-normality implies. De Nardi et al. (2020) present similar evidence for the Netherlands, and Druedahl and Munk-Nielsen (2018) for Denmark.

transfers) is compressed relative to the respective shocks to pre-government income, but persistent shocks remain significantly cyclical. The magnitude of cyclicality of dispersion is in line with Storesletten et al. (2004). Finally, we find strong excess kurtosis of transitory and persistent shocks. It is higher for post- than for pre-government earnings suggesting that after redistribution more mass is concentrated in the center relative to the tails of the distribution. One related recent study of cyclical risk is Angelopoulos et al. (2019), who adapt a version of the GMM estimator developed in the present paper and document procyclical skewness of persistent shocks in Great Britain using data from the British Household Panel Study.

Implications of Higher-Order Risk. We next assess whether the estimated deviations from an income process with log-Normal shocks are economically significant. To this end we set up a standard incomplete-markets life-cycle model, in which households receive stochastic income following the estimated process throughout their working life, after which they enter a retirement phase and receive income through a pay-as-you-go pension system. We focus on ex-post heterogeneity, and thus the only source of inequality in the model is the risky idiosyncratic component of household income. The only means of self-insurance against the income risk explicitly present in the model is through private savings in a risk-free asset. We calibrate the model such that households face the income process estimated on post government household income, reflecting the view that it represents the amount of risk remaining after other channels of insurance against individual level risk—namely: within-household insurance and government taxes and transfers (cf. Blundell et al. 2008). We normalize all shocks in levels, and in this sense the economy does not feature aggregate risk. This reflects our interest in the role of cyclical changes in idiosyncratic risk, and in the relevance of higher-order risk. Agents have recursive preferences over consumption a la Epstein and Zin (1989, 1991), and Weil (1989), which we choose because it allows us to separately control the intertemporal elasticity of substitution and the coefficient of risk aversion. The latter also pins down the higher-order risk attitudes, and through this is a crucial determinant of the behavioral reaction to higher-order risk. To assess the implications of higher-order risk we compare model outcomes under the calibration with the estimated income process and under an alternative calibration where the process features the same dispersion of shocks, but with skewness and kurtosis of the Gaussian distribution (zero and three, respectively).

Our analysis delivers three main findings. First, evaluated from an ex-ante perspective higher-order risk has sizable negative welfare implications for strong risk attitudes: the consumption equivalent variation (CEV) that makes agents in the economy with log-Normal shocks indifferent to the economy with higher-order risk ranges between -0.4% (for a coefficient of relative risk aversion of 2) and -12.5% (relative risk aversion of 4). The dominant economic mechanism driving this welfare result is an expected reallocation of consumption over the life-cycle: when facing riskier income, risk-sensitive agents have more precautionary savings, and thus less consumption at young ages. With weak risk attitudes (specifically, for log utility), the welfare effect can be positive (CEV of 0.4%).[4]

Second, higher-order risk matters for the welfare costs of business cycles. Since Lucas (1987, 2003) argued that the gains of smoothing cycles beyond what the existing tax and transfer system does would be small, several studies have explored the role of both ex-ante and ex-post heterogeneity, wih Imrohoroglu (1989) being the first to emphasize the importance of idiosyncratic risk and incomplete markets. In a model similar to hers, Storesletten et al. (2001) allow for cyclical variance of persistent shocks as estimated in Storesletten et al. (2004). Following a similar strategy, we provide the first systematic assessment of the welfare consequences of cyclical higher-order risk as captured in a continuous distribution function, and thus bridge this approach to papers that explore cyclical downside risk in the form of unemployment (e.g.,

[4]It turns out that the mechanical relationship between the distribution of shocks in logs and the distribution of shocks in levels is important for the results: introducing left-skewness in logs (while holding the variance in logs constant) leads to a reduction of the variance in levels. In other words, the introduction of *third-order risk* (left-skewness) mechanically reduces *second-order risk* (variance) when characterizing the distribution in levels.

Krusell and Smith 1999, Krusell et al. 2009, Krebs, 2003, 2007, and Beaudry and Pages 2001). Under higher-order risk we find welfare costs (computed as CEV making households in the non-cyclical economy indifferent to the cyclical economy) which are 0.3%p (relative risk aversion of 2) to 6.4%p (relative risk aversion of 4) larger than under log-Normal shocks.

Third, we document that higher-order risk crucially matters for the degree of self-insurance against shocks. We employ a measure of self-insurance introduced in the literature by Blundell et al. (2008), who suggest to evaluate the degree of partial insurance against income shocks by estimating the pass-through of the identified transitory and permanent shocks to consumption changes. In the context of our model based analysis, we follow Kaplan and Violante (2010), who study how much of the empirically estimated partial insurance can be generated in a standard incomplete markets model. We find that incorporating higher-order risk leads to weaker pass-through of income shocks to consumption. However, this does not actually represent *better* insurance against negative shocks. In a scenario with higher-order risk agents have more precautionary savings (relative to a scenario in which they face log-Normal shocks), which implies that the consumption reaction to *positive* transitory and persistent shocks is weaker. *Negative* shocks actually translate stronger into negative consumption changes, because the higher savings do not suffice to smooth out shocks which are more pronounced relative to Normal shocks. Therefore, we caution against using only the insurance coefficient by Blundell et al. (2008) for the analysis of the degree of partial insurance against income risk.

Our paper is part of a growing literature that explicitly analyzes the implications of new insights on cyclical skewness of persistent earnings shocks for macroeconomic questions. Golosov et al. (2016) allow for time-varying skewness of idiosyncratic risk in a study of optimal fiscal policy, Catherine (2019) analyzes the implications of procyclical skewness of idiosyncratic income risk for the equity premium, and McKay (2017) links procyclical skewness to aggregate consumption dynamics. Besides the particular economic outcomes of interest, our analysis differs by providing a transparent link between moments of the shock distribution and those outcomes, emphasizing the relevant properties of preferences. Our analysis is also related to work on the implications of rich earnings dynamics in general (without considering the cyclicality of risk). De Nardi et al. (2020) feed an income process a la Arellano et al. (2017) into an incomplete markets model and study the role of richer earnings dynamics for consumption insurance and the welfare costs of idiosyncratic risk. Their analysis focuses on non-linear features of the income process and corroborates results from Karahan and Ozkan (2013) regarding the role of age-dependent persistence and distributions of shocks. Civale et al. (2017) analyze implications of left-skewed and leptokurtic idiosyncratic shocks for the interest rate and aggregate savings in an otherwise standard Aiyagari economy.

The remainder of the paper is structured as follows. Section 2 provides guidance for the analysis by discussing the role of higher-order risk in a simple two-period model. Section 3 presents our empirical approach and discusses identification of the income process. Section 4 presents the results of applying our approach to US household earnings data from the PSID. Section 5 introduces the quantitative model to analyze the economic implications of higher-order income risk, Section 6 discusses the quantitative results, and Section 7 concludes.

2 Higher-Order Risk in a Two-Period Model

2.1 Setup

Endowments. A household lives for two periods denoted by $j \in \{0, 1\}$. At period 0 the household is endowed with an exogenous income of y_0. Period 1 income is risky, $y_1 = \exp(\varepsilon)$, for some random variable ε with distribution function $\Psi(\varepsilon)$, which features higher-order income risk. Households are born with zero assets and, in the general formulation of the model, have access to a risk-free savings technology with interest factor $R = 1$. Denoting by a_1 savings in period 1, the budget constraints in the two periods are

$$a_1 = y_0 - c_0, \qquad c_1 \leq a_1 + y_1.$$

Preferences. We consider additively separable preferences over consumption c_j in the two periods of life, $j \in \{0, 1\}$. The per period utility function takes the standard iso-elastic power utility form $u(c_j) = \frac{1}{1-\theta} c_j^{1-\theta}$, with concavity parameter θ. Thus preferences are given by

$$V = \begin{cases} \frac{1}{1-\theta}\left(c_0^{1-\theta} + \int c_1^{1-\theta} d\Psi(\varepsilon)\right) & \text{for } \theta \neq 1 \\ \ln(c_0) + \int \ln(c_0)\, d\Psi(\varepsilon) & \text{for } \theta = 1. \end{cases}$$

Notice that θ captures both risk attitudes as well as the inverse of the inter-temporal substitution elasticity. In the quantitative life-cycle model we use recursive preferences a la Epstein and Zin (1989, 1991), and Weil (1989) to distinguish the two aspects of preferences. In Appendix A.5 we show that the theoretical analysis presented in this section extends naturally to recursive preferences, and that the risk attitudes are the relevant component of preferences behind consumption reactions to higher-order income risk. In the following, we thus interpret θ as representing risk attitudes when appropriate.

Since we assume an interest rate of zero and no discounting of second-period utility, there is no life-cycle savings motive in this simple model.

2.2 Analysis

Hand-to-Mouth Consumers. We first analyze the role of higher-order risk for hand-to-mouth consumers by shutting down access to the savings technology through constraint $a_1 = 0$.

Consider a fourth-order Taylor series approximation of the objective function around the mean of second period consumption, $\mu_1^c = \mathbb{E}[c_1] = \int c_1 d\Psi(\varepsilon)$. After some transformations, cf. Appendix A.1 and in line with, e.g., Eeckhoudt and Schlesinger (2006), we find that

$$U \approx \frac{c_0^{1-\theta}}{1-\theta} + \left(\frac{1}{1-\theta} - \frac{\theta}{2}\mu_2^c + \frac{\theta(1+\theta)}{6}\mu_3^c - \frac{\theta(1+\theta)(2+\theta)}{24}\mu_4^c \right), \qquad (1)$$

where we impose the restriction $\mu_1^c = 1$ for expositional reasons (which is irrelevant for the results pertaining to second- to fourth-order risk discussed here). Note that under the assumption of the binding budget constraint,the central moments[5] of the level of consumption μ_k^c, $k = 1, \ldots, 4$ coincide with the respective moments $\mu_k^{\exp(\varepsilon)}$, $k = 1, \ldots, 4$, of second period income $\exp(\varepsilon)$.

We make the following observations using the expression in (1). First, consider changing one of the central moments of the distribution while holding the others constant. An increase of the variance, μ_2^c, or of the fourth central moment, μ_4^c, or a reduction of the third central moment, μ_3^c, leads to expected utility losses. Note that changing the third central moment while holding the variance fixed implies changing the shape of the distribution as summarized by the coefficient of *skewness*. Similarly, changing the fourth central moment while holding variance fixed implies changing the relative size of the center and tails of the distribution, as summarized by the coefficient of *kurtosis*. In the remainder of the analysis, whenever we speak of an increase of risk, we refer to a change of the distribution of shocks that entails at least one of these changes (increasing second or fourth central moments, or decreasing the third central moment). Second, the utility consequences of changes of risk are governed by relative risk attitudes,[6] which in case of the employed power utility function are all pinned down by θ. Stronger *relative risk aversion* θ implies stronger adverse effects of increasing variance; stronger *relative prudence* $1+\theta$ implies stronger adverse effects of increasing negative skewness; and stronger *relative temperance* $2+\theta$ implies stronger adverse effects of increasing kurtosis. Importantly, the role of higher-order risk increases exponentially in θ: the weight attributed to risk attitudes on the variance is θ, on the third moment

[5]The k^{th} central moment of variable x is given by $\mu_k^x = \mathbb{E}\left(x - \mu_1^x\right)^k$.

[6]The relative risk attitude of order n is given by $-\frac{u^n(c)}{u^{n-1}(c)}c$, where $u^n(c)$ denotes the n^{th} derivative of the per-period utility function $u(c)$.

is $\theta(1+\theta)$ and on the fourth moment is $\theta(1+\theta)(2+\theta)$. Third, for given θ the relative importance of risk decreases in the order of risk, which is captured by the weight terms of the Taylor approximation.

These observations play a crucial role for our quantitative evaluation. In particular, while our estimates presented in Section 4.2 imply a pronounced left-skewness and a strong excess kurtosis, which may lead to sizeable welfare losses, the overall effect depends crucially on the utility weight of this risk, and thus on the calibration of θ. Indeed, in the case of log utility ($\theta = 1$) the mean-preserving introduction of left-skewness generates welfare gains. We formally derive this implication, which at first glance may appear counterintuitive, in Appendix A.2. It turns out to be crucial that moments of shocks in *levels*, $\exp(\varepsilon)$, rather than of shocks in *logs*, ε, are relevant for utility.

Precautionary Savings. We now assume that households have access to a savings technology. Using the budget constraint in the utility function it is straightforward to derive the Euler equation of the maximization problem as (cf., e.g., Eeckhoudt and Schlesinger, 2008)

$$\begin{aligned}(y_0 - a_1)^{-\theta} = \mathbb{E}\left[(\exp(\varepsilon) + a_1)^{-\theta}\right] \approx (1 + a_1)^{-\theta} + \\ \frac{\theta(1+\theta)}{2}(1+a_1)^{-(2+\theta)}\mu_2^{\exp(\varepsilon)} - \frac{\theta(1+\theta)(2+\theta)}{6}(1+a_1)^{-(3+\theta)}\mu_3^{\exp(\varepsilon)} \\ + \frac{\theta(1+\theta)(2+\theta)(3+\theta)}{24}(1+a_1)^{-(4+\theta)}\mu_4^{\exp(\varepsilon)}. \quad (2)\end{aligned}$$

Notice that the LHS is increasing, and the RHS is decreasing in a_1 if $\mu_3^{\exp(\varepsilon)}$ is small enough relative to $\mu_2^{\exp(\varepsilon)}$ and $\mu_4^{\exp(\varepsilon)}$.[7] Consider the effect of an increase of risk of the income shock $\exp(\varepsilon)$ (through increasing the second or fourth central moment, or reducing the third central moment). An increase of the variance increases the RHS, scaled by the product of the measures of relative prudence and relative risk aversion $\theta \cdot (1+\theta)$. A reduction of the third central moment increases the RHS, additionally scaled by the measure of relative

[7] The RHS is deacreasing in a_1 iff $\mu_3^{\exp(\varepsilon)} \leq \frac{3}{(3+\theta)}(1+a_1)\mu_2^{\exp(\varepsilon)} + \frac{(4+\theta)}{4}(1+a_1)^{-1}\mu_4^{\exp(\varepsilon)}$.

temperance $(2+\theta)$. An increase of the fourth central moment increases the RHS, additionally scaled by the measure of *relative edginess* $(3+\theta)$.[8] Similar to what we saw in equation (1), the second to fourth moments are scaled by additional weight factors $\frac{1}{2(1+a_1)^{2+\theta}}$, $\frac{1}{6(1+a_1)^{3+\theta}}$, and $\frac{1}{24(1+a_1)^{4+\theta}}$, respectively.

Therefore, an increase of risk for a given a_1 increases the RHS, which is offset by an increase of savings a_1.[9] This result is very intuitive: ordinary and higher-order income risk increases precautionary savings, through which households reduce the adverse utility consequences of risk. The intensity of the behavioral reaction crucially depends on risk attitudes as governed by θ.

3 Income Process with Higher-Order Risk

3.1 The Income Process

Let log income of household i of age j in year t be

$$y_{ijt} = f\left(\mathbf{X}_{ijt}, Y_t\right) + \tilde{y}_{ijt}, \tag{3}$$

where $f\left(\mathbf{X}_{ijt}, Y_t\right)$ is the *deterministic* component of income, i.e., the part that can be explained by observable individual and aggregate characteristics, $\mathbf{X}_{ijt}$ and Y_t, respectively, and $\tilde{y}_{ijt}$ is the *residual* part of income, which is assumed to be orthogonal to $f\left(\mathbf{X}_{ijt}, Y_t\right)$. The deterministic component $f\left(\mathbf{X}_{ijt}, Y_t\right)$ is a linear combination of a cubic in age j, $f_{age}(j)$, the log of household size, year fixed effects, and an education premium $f_{EP}(t)$ for college education, which we allow to vary over years t:

$$f\left(\mathbf{X}_{ijt}, Y_t\right) = \beta_{0t} + f_{age}\left(j\right) + \mathbb{1}_{e_{it}=c} f_{EP}(t) + \beta^{size}\log\left(hhsize_{ijt}\right) \tag{4}$$

where $f_{age}\left(j\right) = \beta_1^{age} j + \beta_2^{age} j^2 + \beta_3^{age} j^3$, $f_{EP}(t) = \beta_0^{EP} + \beta_1^{EP} t + \beta_2^{EP} t^2$, and $\mathbb{1}_{e_{it}=c}$ is an indicator function that takes on value 1 for college-educated households.

[8]The term *edginess* was coined by Lajeri-Chaherli (2004).

[9]Formally, it is straightforward to show this by taking the total differential of (2), cf. Appendix A.4.

Residual income $\tilde{y}_{ijt}$ is the main object of interest in the analysis. We model $\tilde{y}_{ijt}$ as the sum of three components: a persistent component z_{ijt}, an i.i.d. transitory shock ε_{ijt}, and an idiosyncratic *fixed effect* χ_i. The idiosyncratic fixed effect is a shock drawn once upon entering the labor market from a distribution which is the same for every cohort.[10] The persistent component is modeled as an AR(1) process with innovation η_{ijt}:

$$\tilde{y}_{ijt} = \chi_i + z_{ijt} + \varepsilon_{ijt}, \text{ where } \varepsilon_{ijt} \underset{iid}{\sim} F_\varepsilon, \; \chi_i \underset{iid}{\sim} F_\chi \tag{5a}$$

$$z_{ijt} = \rho z_{ij-1t-1} + \eta_{ijt}, \text{ where } \eta_{ijt} \underset{id}{\sim} F_\eta\left(s\left(t\right)\right), \tag{5b}$$

where F_χ, F_ε, and $F_\eta\left(s\left(t\right)\right)$ denote the density functions of χ, ε_{ijt}, and η_{ijt}, respectively. We allow the density function of the persistent shock to depend on the aggregate state of the economy in period t, denoted by $s(t)$. This income process is exactly the canonical income process (e.g., Moffitt and Gottschalk, 2011). Unlike the canonical case, we do not (implicitly) assume that the shocks to the log income process are symmetric. Instead of only focussing on the variance of the shocks, we are interested in estimating the second to fourth central moments of the density functions, and denote those by μ_2^x, μ_3^x, and μ_4^x, for $x \in \{\chi, \varepsilon, \eta(s)\}$.[11]

As in Storesletten et al. (2004), the economy can be in one of two aggregate states, which we denote by E (expansion) and C (contraction). Thus, the central moments of the persistent shock $\mu_k^\eta\left(s\left(t\right)\right)$ are equal to $\mu_k^{\eta,E}$ if $s\left(t\right) = E$ and equal to $\mu_k^{\eta,C}$ if $s\left(t\right) = C$, for $k \in \{2, 3, 4\}$. Both empirical evidence (e.g., Blundell et al. 2008) and model-based analyses (e.g., Kaplan and Violante 2010) find that households can insure well against transitory shocks. We therefore follow Storesletten et al. (2004) and only consider the cyclicality of persistent income shocks, which have long-lasting effects in the context of a

[10]Thus, from the econometric perspective, we are estimating a random effects model.

[11]One potential disadvantage of using central moments to characterize the shocks in the income process is that they are hard to interpret by themselves. However, in the samples we use, the central moments of the cross-sectional income distribution are strongly correlated with percentile-based counterparts to those moments. We are thus confident that the estimated central moments—and the implied standardized moments *skewness* and *kurtosis*—do capture the salient features of the distribution.

life-cycle decision making problem. We still do capture skewness and kurtosis of the (acyclical) transitory component and explore its quantitative role.

We assume that upon entering the labor market, in addition to drawing the fixed effect χ_i, each worker draws the first realizations of transitory and persistent shocks, ε_{it} and η_{it}, from the distributions F_ε and $F_\eta\left(s\left(t\right)\right)$, respectively. Thus, the moments of the distribution of the persistent component for the cohort entering in year t at age $j=0$ are $\mu_k(z_{i0t}) = \mu_k^\eta(s(t))$.

3.2 GMM Approach to Estimation

We follow the common approach in the literature and estimate (3) and (5) in two steps. In the first step, we estimate (3), which yields residuals $\tilde{y}_{ijt}$. In the second step, we estimate the parameters of the stochastic process (5) by fitting cross-sectional moments of the distribution of residual (log) income. As is standard, the variance terms of all components of (5) can be identified by the variance-covariance matrix. Similarly, the third and fourth central moments can be identified by third and fourth central moments and co-moments. Let $\theta = \left(\rho, \mu_2^\chi, \mu_2^\varepsilon, \mu_2^{\eta,E}, \mu_2^{\eta,C}, \mu_3^\chi, \mu_3^\varepsilon, \mu_3^{\eta,E}, \mu_3^{\eta,C}, \mu_4^\chi, \mu_4^\varepsilon, \mu_4^{\eta,E}, \mu_4^{\eta,C}\right)$ be the vector of second-stage parameters, and let s^t summarize the history of aggregate states up to year t.[12] We denote central moments by $\mu_k\left(\cdot\right)$ and co-moments by $\mu_{kl}\left(\cdot\right)$, where

$$\mu_k\left(\tilde{y}_{ijt};\theta\right) = E\left[\left(\tilde{y}_{ijt} - E\left[\tilde{y}_{ijt}\right]\right)^k |s^t\right] \tag{6a}$$

$$\mu_{kl}\left(\tilde{y}_{ijt}, \tilde{y}_{ij+1t+1};\theta\right) = E\left[\left(\tilde{y}_{ijt} - E\left[\tilde{y}_{ijt}\right]\right)^k \left(\tilde{y}_{ij+1t+1} - E\left[\tilde{y}_{ij+1t+1}\right]\right)^l |s^t\right]. \tag{6b}$$

[12] Note that we need to condition only on s^t, not on s^{t+1}, because period $t+1$ shocks are uncorrelated with all shocks accumulated up to period t.

The imposed process implies the following moments of the distribution of residual income at age j in year t:

$$\mu_2(\tilde{y}_{ijt};\theta) = \mu_2^{\chi} + \mu_2^{\varepsilon} + \mu_2(z_{ijt}) \tag{7a}$$

$$\mu_{11}\left(\tilde{y}_{ijt}, \tilde{y}_{ij+1t+1};\theta\right) = \mu_2^{\chi} + \rho\mu_2(z_{ijt}) \tag{7b}$$

$$\mu_3\left(\tilde{y}_{ijt};\theta\right) = \mu_3^{\chi} + \mu_3^{\varepsilon} + \mu_3(z_{ijt}) \tag{7c}$$

$$\mu_{21}\left(\tilde{y}_{ijt}, \tilde{y}_{ij+1t+1};\theta\right) = \mu_3^{\chi} + \rho\mu_3(z_{ijt}) \tag{7d}$$

$$\mu_4\left(\tilde{y}_{ijt};\theta\right) = \mu_4^{\chi} + \mu_4^{\varepsilon} + \mu_4(z_{ijt}) + 6\left(\mu_2^{\chi}\mu_2^{\varepsilon} + \left(\mu_2^{\chi} + \mu_2^{\varepsilon}\right)\mu_2(z_{ijt})\right) \tag{7e}$$

$$\mu_{31}\left(\tilde{y}_{ijt}, \tilde{y}_{ij+1t+1};\theta\right) = \mu_4^{\chi} + \rho\mu_4(z_{ijt}) + 3\left(\mu_2^{\chi}\mu_2^{\varepsilon} + \left(\mu_2^{\chi} + \rho\left(\mu_2^{\chi} + \mu_2^{\varepsilon}\right)\right)\mu_2(z_{ijt})\right), \tag{7f}$$

where $\mu_k(z_{ijt})$, for $k = 2, 3, 4$ is shown in Appendix A.6.

A crucial implication of equations (7c) and (7e) is that the cross-sectional distribution of $\tilde{y}_{ijt}$ does not converge to a Normal distribution, as the third and fourth central moments of the shocks accumulate over age. This allows us to identify these higher-order moments of the shock distributions based on cross-sectional moments as outlined below. Denote the empirical counterparts of the moments by $m_2(\cdot)$, $m_3(\cdot)$, $m_4(\cdot)$, $m_{11}(\cdot)$, $m_{21}(\cdot)$, and $m_{31}(\cdot)$. This gives the following set of moment conditions employed in the GMM estimation:

$$E\left[m_2\left(\tilde{y}_{ijt}\right) - \mu_2\left(\tilde{y}_{ijt};\theta\right)|s^t\right] = 0 \tag{8a}$$

$$E\left[m_{11}\left(\tilde{y}_{ijt}, \tilde{y}_{ij+1t+1}\right) - \mu_{11}\left(\tilde{y}_{ijt}, \tilde{y}_{ij+1t+1};\theta\right)|s^t\right] = 0 \tag{8b}$$

$$E\left[m_3\left(\tilde{y}_{ijt}\right) - \mu_3\left(\tilde{y}_{ijt};\theta\right)|s^t\right] = 0 \tag{8c}$$

$$E\left[m_{21}\left(\tilde{y}_{ijt}, \tilde{y}_{ij+1t+1}\right) - \mu_{21}\left(\tilde{y}_{ijt}, \tilde{y}_{ij+1t+1};\theta\right)|s^t\right] = 0 \tag{8d}$$

$$E\left[m_4\left(\tilde{y}_{ijt}\right) - \mu_4\left(\tilde{y}_{ijt};\theta\right)|s^t\right] = 0 \tag{8e}$$

$$E\left[m_{31}\left(\tilde{y}_{ijt}, \tilde{y}_{ij+1t+1}\right) - \mu_{31}\left(\tilde{y}_{ijt}, \tilde{y}_{ij+1t+1};\theta\right)|s^t\right] = 0. \tag{8f}$$

Huggett and Kaplan (2016) use a similar strategy based on second and third central moments and co-moments, without resorting to pre-sample aggregate information in the spirit of Storesletten et al. (2004) as we do. We use moment conditions (8a) and (8b) to estimate the variance parameters and the persistence ρ. Given an estimate for ρ, we then use moment conditions (8c) and (8d) to estimate the third central moments. Likewise, given estimates for ρ and the variance parameters, we use moment conditions (8e) and (8f) to estimate the fourth central moments.

Identification. The use of cross-sectional moments for identification allows us to exploit macroeconomic information that predates the micro panel, thereby incorporating more business cycles in the analysis than covered by the sample, as pointed out by Storesletten et al. (2004). Consider the persistent component of the income process in equation (5b): the variance of the innovations accumulate as a cohort ages, as can be seen from the theoretical moment in equation (7a). If the innovation variance is higher in contractionary years, then a cohort that lived through more contractions will have a higher income variance at a given age than a cohort at the same age that lived through fewer contractions, if the persistence is high.

Our extension of Storesletten et al. (2004) is based on the insight that other central moments accumulate in a similar fashion, as seen in equations (7c) and (7e). Consider the third central moment. If the probability of a large negative income shock was higher (or that of a large positive shock lower) during a contractionary period, then this would translate into the third central moment of the shock being smaller (more negative) than in an expansion, i.e., $\mu_3^{\eta,C} < \mu_3^{\eta,E}$. For a given dispersion this implies a reduction of skewness (a more left-skewed distribution). Comparing again two cohorts when they reach a certain age, this would imply a more negative cross-sectional third central moment for the cohort that worked through more contractions.

As seen in (7a), the sum $(\mu_2^{\chi} + \mu_2^{\varepsilon})$ is identified as the intercept of the variance profile over age. The same holds for $(\mu_3^{\chi} + \mu_3^{\varepsilon})$ in (7c), which is identified via the age profile of the third central moment. Considering the sum in (7a), we see that the magnitude of the increase of the cross-sectional variance over age identifies the variance of persistent shocks. The difference between $\mu_2^{\eta,C}$ and $\mu_2^{\eta,E}$ is identified by the difference of the cross-sectional variance of different cohorts of the same age. Likewise, the difference between

$\mu_3^{\eta,C}$ and $\mu_3^{\eta,E}$ is identified by the difference of the cross-sectional third central moment of different cohorts. Note that by restricting the transitory shocks to not vary over the business cycle we do not bias the estimated cyclicality of persistent shocks, which is identified via accumulated shock distributions.

Now consider the expressions for variance and covariance in equations (7a) and (7b). The difference between the two expressions identifies μ_2^χ separately from μ_2^ε. Likewise, the difference between the expressions for the third central moment and co-moment, equations (7c) and (7d), identifies μ_3^χ separately from μ_3^ε. Given ρ and the variance parameters μ_2^x for $x \in \{\chi, \varepsilon, \eta(s)\}$, equations (7e) and (7f) identify the fourth central moments μ_4^x for $x \in \{\chi, \varepsilon, \eta(s)\}$ in the same way as for the second and third central moments.

4 Estimation of the Income Process

4.1 Data and Sample Selection

We use data from the Panel Study of Income Dynamics (PSID), which interviews households in the United States annually from 1968 to 1997 and every other year since then. The representative core sample consists of about 2,000 households in each wave, and we use data from 1977–2012.[13] We estimate the income process at the household level for both pre- and post-government household income. De Nardi et al. (2020) show that at the individual level, the PSID sample captures well the salient features of earnings dynamics documented in administrative social security data by Guvenen et al. (2016), and Busch et al. (2020) document that the cyclical changes of the distribution of annual earnings changes in the PSID reflect the dynamics in social security data documented by Guvenen et al. (2014). Similarly, Arellano et al. (2017) estimate a rich earnings process using the PSID.

Household pre-government income is defined as labor income before taxes, which we calculate as the sum of head and spouse annual labor income. Post-

[13]We do not use earlier waves because of poor coverage of income transfers before the 1977 wave.

government income is defined as household labor income plus transfers minus taxes. As measure of labor income we use annual total labor income which includes income from wages and salaries, bonuses, and the labor part of self-employment income. We impute taxes using Taxsim, and add 50% of the estimated payroll taxes to the sum of head and spouse labor incomes to obtain pre-government income. We aggregate transfers to the household level and include measures of unemployment benefits, workers' compensation, combined old-age social security and disability insurance (OASI), supplemental security income, aid to families with dependent children (AFDC), food stamps, and other welfare.

We deflate all nominal values with the annual CPI, and select households if the household head is between 25 and 60 years of age. The minimum of household pre- and post-government income needs to be above a constant threshold, which is defined as the income from working 520 hours at half the minimum wage. Central moments (especially of higher order) are imprecisely estimated in small samples. We therefore estimate the moments for a given year and age group based on a sample from a five-year window over age within the year, which also smoothes the age profiles of these moments.

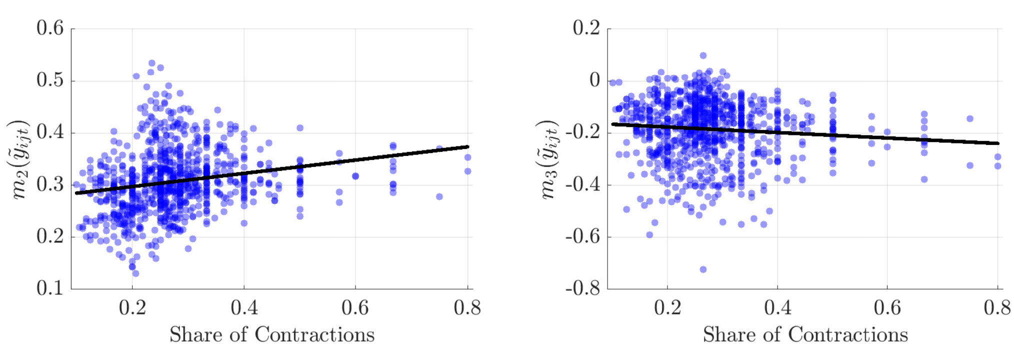
Defining Business Cycles. In order to implement the estimator we classify years as contractions or expansions. We initiate our definition on NBER peaks and trough data. The relevant time period is 1942–2012. Starting from the dating of peaks and troughs, we classify a year as a contraction if (i) it completely is in a contractionary period, which is defined as the time from peak to trough, (ii) if the peak is in the first half of the year and the contraction continues into the next year, (iii) if a contraction started before the year and the trough is in the second half of the year. Given the sluggish synchronization of labor market outcomes with the macroeconomic indicators that the NBER takes into account (cf. Guvenen et al. 2014, Huggett and Kaplan 2016), we expand the dating based on mean earnings of males in the PSID. We classify the following years as contractions: 1945, 1949, 1953, 1957, 1960, 1970, 1974, 1980–83, 1990–91, 2001–02, 2008–10. All years that are not classified as contractions are classified as expansions.

4.2 Estimation Results: Cyclical Idiosyncratic Risk

Illustration of Identification. Before turning to the estimation, in Figure 1 we plot the cross-sectional second and third central moments of residual income (for post-government income) used in the estimation, i.e., each marker denotes a moment for households of some age j in some year t, $m_k(\tilde{y}_{ijt})$ for $k = 2, 3$. The moments are plotted against the share of years classified as contractions out of all years a cohort went through since age 25 once reaching the given year. The pattern that emerges is that a higher share of contractionary years correlates positively with the cross-sectional second moments, and correlates negatively with the cross-sectional third moments. These correlations identify the cyclicality of the moments of the shocks in the esimated income process.

Figure 1: Cross-Sectional Moments by Aggregate History

(a) Second Cross-Sectional Moment (b) Third Cross-Sectional Moment



Notes: Cross-sectional moments of residual income are net of age effects. Share of contractions for a given moment is the fraction of years classified as contraction since age 25. The slopes of the fitted lines are 0.13 and –0.11 for m_2 and m_3, respectively. Moments for shares of 0 or 1 are not displayed here for visualization reasons (they are used in the GMM estimation).

Estimation. We now turn to the estimation results for household pre-government labor income (before taxes and transfers) and household post-government labor income (after taxes and transfers). We use the number of observations that contribute to an empirical moment as weights for the moment conditions, and this way assign more weight to those moments that are themselves estimated more reliably in the data. As additional moment conditions we add the averages over years of the second to fourth central moments of 1-5 year income changes. This ensures that the estimated income process is consistent both with moments of the cross-sectional distribution and with moments of income changes. We give a collective weight of 10% to the average moments of changes. In addition to the structure imposed so far, we hold the kurtosis of η fixed over the business cycle. Let α_i denote the i^{th} standardized moment: $\alpha_i = \mu_i/\mu_2^{i/2}$. Assuming $\alpha_4^{\eta}(s(t)) = \alpha_4^{\eta}$ implies $\mu_4^{\eta,C} = \alpha_4^{\eta}\left(\mu_2^{\eta,C}\right)^2$ and $\mu_4^{\eta,E} = \alpha_4^{\eta}\left(\mu_2^{\eta,E}\right)^2$. This leaves us with 12 parameters that need to be estimated. We use moment conditions (8a) and (8b) to estimate the variance parameters and the persistence ρ. Given an estimate for ρ, we then use moment conditions (8c) and (8d) to estimate the third central moments. Likewise, given estimates for ρ and the variance parameters, we use moment conditions (8e) and (8f) to estimate the fourth central moments. The third central moment of the cross-sectional distribution features a low-frequency change (see Panel (e) of Figure 2). In order to accommodate this in the estimation, and to not confound the estimated cyclicality, we add a linear trend to the third central moment of transitory shocks. We report the time average of the implied moment. For inference, we apply a block bootstrap procedure and resample households, which preserves the autocorrelation structure of the original sample. We draw 1,000 bootstrap samples. Table 1 shows the estimates, and Figure 2 illustrates the fit over age and time of the estimated process for post government income, the income variable we use in the quantitative analysis in Section 6 (see Appendix B for the implied standardized moments).

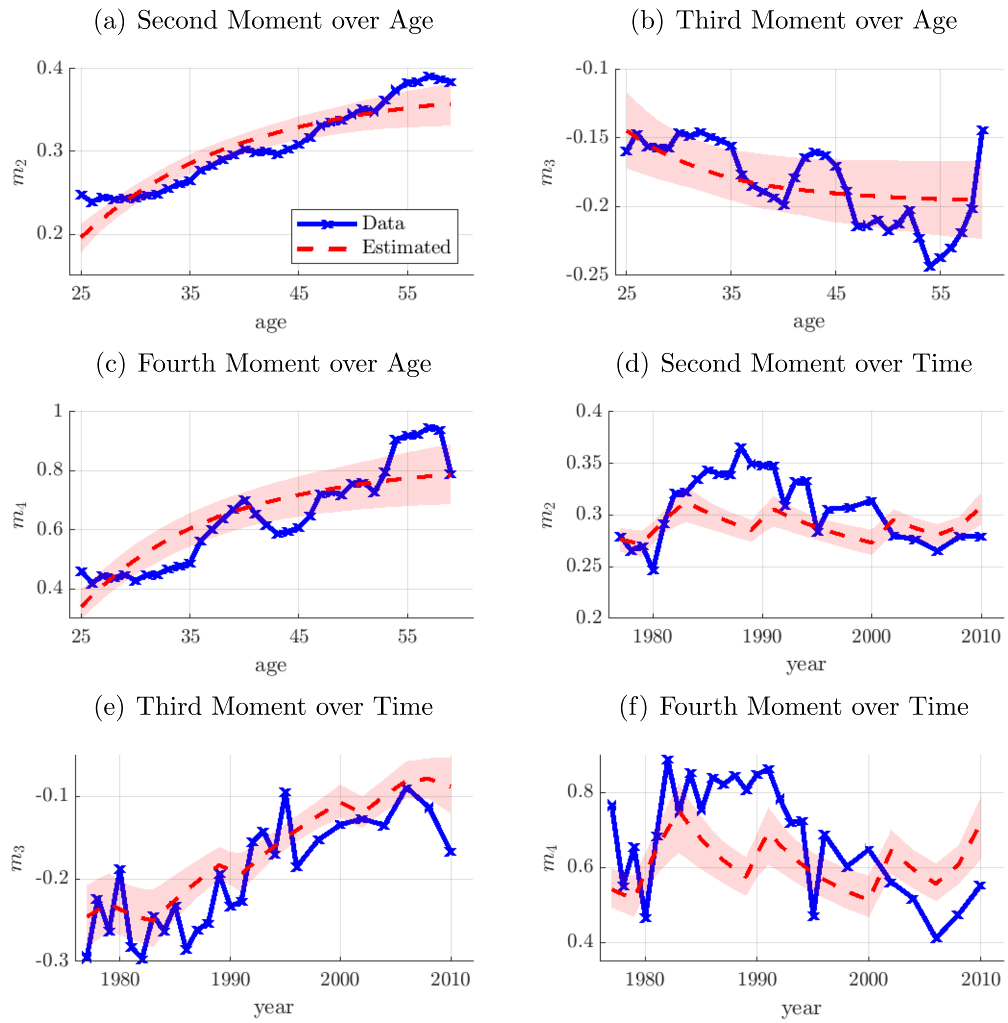
Cyclical Dispersion. The first panel of Table 1 reports the persistence of the AR(1) component of income along with the estimates of the variances of

Table 1: Estimation Results for Pre- and Post Government Income

	Estimated Central Moments		Implied Standardized Moments	
	HH Pre	HH Post	HH Pre	HH Post
ρ	0.9601	0.9683		
	[0.9412; 0.9756]	[0.9463; 0.9841]		
μ_2^{χ}	0.1591	0.1076		
	[0.1361; 0.1786]	[0.0897; 0.1237]		
μ_2^{ε}	0.1045	0.0752		
	[0.0948; 0.1133]	[0.0677; 0.0816]		
$\mu_2^{\eta,C}$	0.0375	0.0223		
	[0.0263; 0.0477]	[0.0152; 0.0291]		
$\mu_2^{\eta,E}$	0.0152	0.0085		
	[0.0099; 0.0229]	[0.0044; 0.0153]		
μ_3^{χ}	−0.1126	−0.0520	−1.77	−1.47
	[−0.1530; −0.0727]	[−0.0785; −0.0253]	[−2.47; −1.21]	[−2.32; −0.78]
μ_3^{ε}	−0.1516	−0.0866	−4.49	−4.20
	[−0.1623; −0.1374]	[−0.0935; −0.0772]	[−5.06; −3.97]	[−4.79; −3.73]
$\mu_3^{\eta,C}$	−0.0332	−0.0164	−4.59	−4.95
	[−0.0474; −0.0175]	[−0.0263; −0.0061]	[−6.73; −2.61]	[−7.80; −2.18]
$\mu_3^{\eta,E}$	−0.0047	−0.0012	−2.49	−1.54
	[−0.0128; 0.0035]	[−0.0069; 0.0040]	[−6.73; 2.26]	[−7.97; 9.09]
μ_4^{χ}	0.0607	0.0173	2.40	1.50
	[0.0000; 0.1508]	[0.0000; 0.0741]	[0.00; 5.39]	[0.00; 5.38]
μ_4^{ε}	0.4250	0.2300	38.94	40.62
	[0.3630; 0.4867]	[0.1927; 0.2664]	[34.52; 44.92]	[36.27; 47.65]
$\mu_4^{\eta,C}$	0.1359	0.0666	96.85	134.47
	[0.0856; 0.1719]	[0.0363; 0.0847]	[61.15; 141.97]	[82.02; 191.34]
$\mu_4^{\eta,E*}$	0.0225	0.0098	96.85	134.47
	[0.0089; 0.0488]	[0.0022; 0.0272]	[61.15; 141.97]	[82.02; 191.34]

Notes: Table shows estimated central moments for household earnings (HH Pre) and household income after taxes and transfers (HH Post). Brackets show 5^{th} and 95^{th} percentiles of 1,000 bootstrap estimates (in the case of post government income, 998 of the bootstrap iterations converge). $^*\mu_4^{\eta,E}$ not separately estimated.

Figure 2: Fit of Estimated Process for Post-Government Earnings



(a) Second Moment over Age

(b) Third Moment over Age

(c) Fourth Moment over Age

(d) Second Moment over Time

(e) Third Moment over Time

(f) Fourth Moment over Time

Notes: Moments are cross-sectional central moments. For each moment, age and year profiles are based on a regression of the moment on a set of age and year dummies. Blue lines: empirical moments; red dashed lines: theoretical moments implied by point estimates; shaded area denotes a 90% confidence band based on the bootstrap iterations.

the components of the income process estimated jointly. We estimate persistence parameters (ρ) of .96 and .97 for pre and post government income, respectively. The estimated variances of all components of the post-government income process are smaller than their counterparts for pre-government income. This is consistent with an interpretation that the existing tax and transfer system effectively dampens the idiosyncratic risk faced by households. Both for pre- and post-government income the estimates imply a countercyclical variance of persistent shocks: in aggregate downturns, the cross-sectional distribution of shocks is more dispersed. Our estimate of countercyclicality for post-government income is quantitatively similar to the one estimated by Storesletten et al. (2004): the estimated standard deviation of persistent shocks is 61% higher in aggregate contractions.

Cyclical Skewness. The second panel of Table 1 reports the third central moments. We find that all shock components estimated for pre-government and post-government income processes have negative third central moments, implying negative skewness of shocks. Comparing the post-government income process to the pre-government income process, the third central moments are smaller in magnitude, as expected from the reduced dispersion. For both pre and post government income, the third central moment of persistent shocks is significantly negative in contractions; point estimates of the third central moments of persistent shocks in expansions are also negative, however not statistically different from zero. The second and third central moments together translate into the third standardized moment, the coefficient of skewness, which is informative about the shape of the distribution and shown in the last two columns of Table 1. The cyclicality of the third central moment is stronger relative to the cyclicality of the second moment, which translates into the standardized moment displaying pro-cyclicality. Thus, aggregate contractions are periods in which negative persistent shocks become relatively more pronounced.

Excess Kurtosis. The third panel of Table 1 reports the fourth central moments. We restrict the kurtosis of persistent shocks to not vary with the aggregate state of the economy, i.e., $\alpha_4^\eta(s(t) = C) = \alpha_4^\eta(s(t) = E)$. Again, the last two columns of Table 1 list the implied standardized fourth moments (coefficients of kurtosis). The fixed effects are very imprecisely estimated; the point estimates imply relatively flat distributions (compared to a Normal distribtion, which has a kurtosis of 3): the implied kurtosis coefficient at the point estimates is 2.4 for pre-government income, and 1.5 for post-government income. The transitory and persistent shocks are estimated to display very pronounced excess kurtosis of about 39 and 97 for pre-government earnings, and about 41 and 134 for post-government earnings. These estimates imply that the distribution of post-government income shocks is more concentrated in the center, while some households experience shocks that are more extreme *relative* to the overall more compressed (in comparison to pre-government income) distribution. Note that while these estimates of kurtosis seem very high at first glance, they imply a good fit of the cross-sectional distribution over age and over years as shown in Figure 2. Furthermore, the estimated income process is in line with the average kurtosis of income changes.

5 A Quantitative Model

5.1 The Economy

We now set up a quantitative version of the simple two-period model of Section 2 by extending it to a standard multi-period life-cycle model with a stochastic earnings process, a zero borrowing constraint, a fixed retirement age, and an earnings-related retirement income. To calibrate higher-order risk attitudes separately from the inter-temporal elasticity of substitution, we take Epstein-Zin-Weil preferences a la Epstein and Zin (1989, 1991), and Weil (1989).

Endowments. Households earnings are exogenous and consist of a deterministic age profile and a stochastic income component with transitory and persistent shocks. The distribution of persistent shocks varies with the aggregate state $s \in \{C, E\}$, which follows a Markov process with time-invariant transition matrix Π_s. We abstract from the aggregate effects of fluctuations on wages and interest rates by holding both constant. In this sense there is *no aggregate risk*, but *cyclical idiosyncratic risk*.

Households live from age $j = 0$ to age $j = J$. They retire at the exogenously given retirement age j_r. Labor income net of taxes and transfers at age $j \in \{0, \ldots, j_r - 1\}$ in aggregate state s is given by

$$y(z, \varepsilon, j; s) = e_j \cdot \exp(z(s) + \varepsilon), \tag{9}$$

where e_j is the deterministic age profile, ε is the transitory income shock, drawn iid from distribution $\tilde{F}_\varepsilon$, and $z(s)$ is the persistent income component which obeys

$$z'(s') = \begin{cases} \rho z + \eta', & \text{where } \eta' \underset{iid}{\sim} \tilde{F}_\eta(s') \text{ for } j < j_r \\ z & \text{for } j \geq j_r, \end{cases} \tag{10}$$

where ρ is the autocorrelation coefficient and η' is the persistent income shock, drawn from distribution $\tilde{F}_\eta(s')$ that depends on aggregate state s. We assume that $\exp(\varepsilon_0) = \exp(z_0) = 1$. In retirement, $j \in \{j_r, \ldots, J\}$, households earn a fixed earnings related pension income contingent on the last income state before retirement $y_j = b(z_j)$.[14] Households have access to a risk-free savings technology with rate of return r, and face a zero borrowing constraint. Thus, the dynamic budget constraint is

$$a'(z, \varepsilon, j; s) = a(1 + r) + y(z, \varepsilon, j; s) - c \geq 0. \tag{11}$$

Preferences and Household Problem. Households born into the economy at history s^t, date t maximize recursive utility by solving a consumption-savings problem every period. They discount the future at factor $\beta > 0$. The state variables of the household's problem are age j, asset holdings a, the persistent income state z, the transitory shock ε, and the aggregate state of the

[14]With this specification we approximate the average indexed monthly earnings (AIME) of the US pension system.

economy s. The recursive problem of households is

$$V_j(a,z,\varepsilon;s) = \max_{c,a'} \begin{cases} \left((1-\tilde{\beta})c^{1-\frac{1}{\gamma}} + \tilde{\beta}\left(v(V_{j+1}(a',z',\varepsilon';s'))\right)^{1-\frac{1}{\gamma}}\right)^{\frac{1}{1-\frac{1}{\gamma}}} & \gamma \neq 1 \\ \exp\left\{(1-\tilde{\beta})\ln c + \tilde{\beta}\ln\left(v(V_{j+1}(a',z',\varepsilon';s'))\right)\right\} & \text{otherwise} \end{cases}$$

$$\text{s.t. } (9), (10), \text{and } (11),$$

where $\tilde{\beta} = \frac{\beta}{1+\beta}$ denotes the relative utility weight on the certainty equivalent $v(V_{j+1})$ from next period's continuation utility $V_{j+1}(\cdot)$, which is

$$v(V_{j+1}(a',z',\varepsilon';s')) = \begin{cases} \left(\mathbb{E}_j\left[V_{j+1}(a',z',\varepsilon';s')^{1-\theta}\right]\right)^{\frac{1}{1-\theta}} & \theta \neq 1 \\ \exp\left(\mathbb{E}_j\left[\ln V_{j+1}(a',z',\varepsilon';s')\right]\right) & \text{otherwise.} \end{cases}$$

Parameter γ denotes the inter-temporal elasticity of substitution between instantaneous utility from consumption and the certainty equivalent of the continuation utility $v(V_{j+1}(\cdot))$. Given γ, parameter θ pins down the relative risk attitudes of households as discussed in Section 2, respectively in Appendix A. Conditional expectations are defined with respect to the realization of next period's aggregate state of the economy s', transitory income shock ε', and persistent income shock η'.

We solve for household policy and value functions using the method of endogenous gridpoints. We aggregate by explicit aggregation iterating forward on the cross-sectional distribution $\Phi_j(a_j, z_j, \varepsilon; s)$, which follows from the initial distribution $\Phi_0(a_0, z_0, \varepsilon_0; s)$ and the transition function $G_j(a_j, z_j, \varepsilon_j; s)$. The latter is induced by the exogenous laws of motion of z, s, the exogenous distribution of ε, and the endogenous transitions $a'_j(a_j, z_j, \varepsilon_j; s)$.

5.2 Calibration

Aggregate Shock Process. Based on our classification of time periods as contractions and expansions for the US economy, we estimate a Markov transition process on this data. We estimate $\pi(E|E) = 0.788$ and $\pi(C|C) = 0.389$, implying the stationary invariant distribution $\Pi_s = [0.257, 0.743]'$.

Age Bins and Age Productivity. Each model period corresponds to one life year. Consistent with our empirical specification, households start working at age 25 (model age $j = 0$) and retire at age 60 (model age $j = 35$).

In the economic model, we abstract from heterogeneity along the dimensions of education, labor market experience, or household size. We calibrate the age productivity process e_j by the fitted age polynomial $f_{age}(j)$ of the first stage estimation of the earnings process for household post government earnings. We take the weighted average of college and non-college age earnings profiles that display the usual hump-shaped pattern, cf. Appendix D.3, and normalize it such that average productivity is equal to one, $\frac{1}{j_r}\sum_{j=0}^{j_r-1} e_j = 1$.

Idiosyncratic Shock Processes. The most important element of the calibration is the specification of the distribution functions of the idosyncratic shocks. The goal of our approach is to directly assess the economic consequences of distributional aspects of these shocks that are summarized in the central moments—and to thus extend the illustrative analysis from Section 2, which does not need to make any (parametric) distributional assumptions, to a quantitative framework. For a given shock, our approach can be summarized in two steps. First, we use a parametric continuous distribution function, which we parameterize such that its first four central moments fit the ones estimated. Second, we discretize this distribution function. Thus, our approach allows us to translate the estimated central moments directly into the model's shock distributions without having to simulate the income process.

As distribution function we choose the Flexible Generalized Lambda Distribution (FGLD) developed by Freimer et al. (1988), which is characterized by its quantile function

$$Q(p;\lambda) = F^{-1}(p;\lambda) = x = \lambda_1 + \frac{1}{\lambda_2}\left(\frac{p^{\lambda_3}-1}{\lambda_3} - \frac{(1-p)^{\lambda_4}-1}{\lambda_4}\right), \qquad (12)$$

where λ is a vector of four parameters with location parameter λ_1, scale parameter λ_2, and tail index parameters λ_3, λ_4.[15] For each shock $x \in \{\varepsilon, \eta(s)\}$,

[15] The parametric constraints are $\lambda_2 > 0$, and $\min\{\lambda_3, \lambda_4\} > -\frac{1}{4}$.

we fit these parameters such that the FGLD matches the estimated central moments $\{\hat{\mu}_i^x\}_{i=1}^4$ of distributions F_ε and $F_\eta(s)$ (as in Lakhany and Mausser 2000 and Su 2007). We numerically solve for λ_3 and λ_4 jointly to fit the third and fourth central moments.[16] Next, we determine λ_2 to match the variance and λ_1 to match the mean, both in closed form. We then discretize by spanning equidistant grids for the respective random variable $x \in \{\varepsilon, \eta(s)\}$ and by assigning to each grid point probabilities from the integrated probability density function of the distribution (details in Appendix C).

We consider two alternative parameterizations of the FGLD to which we refer as *distribution scenarios*. The first scenario features symmetric shock distributions ($\hat{\mu}_3 = 0$) with the estimated variance and a kurtosis of $\frac{\hat{\mu}_4}{\hat{\mu}_2^2} = 3$. The parameter restriction on the FGLD is that $\lambda_3 = \lambda_4$. We refer to this scenario as NORM, reflecting that it features the first four central (as well as standardized) moments of the Normal distribution. The second scenario, to which we refer as LKSW, features leptokurtic and left-skewed shock distributions with the estimated second, third, and fourth central moments; no restrictions apply to the FGLD parameters.[17]

Figure 3 shows the log density functions of the persistent shock $\eta(s)$ in contractions and expansions. Panel (a) shows the distributions in scenario LKSW which features the estimated countercyclical variance, procyclical third, and countercyclical fourth moments. For comparison, Panel (b) shows Gaussian distributions featuring the same countercyclical variance. The FGLD distribution does not nest the Gaussian distribution, and thus, while FGLD distribution NORM features the same first four central moments, it does not display the same inverse quadratic log density function. In both distributions, all odd

[16] Specifically, we solve the minimization problem $\min_{\lambda_3,\lambda_4} \sum_{i=3}^4 (\mu_i(\lambda_3,\lambda_4) - \hat{\mu}_i)^2$ s.t. $\min\{\lambda_3,\lambda_4\} > -\frac{1}{4}$, where $\hat{\mu}_i$ is the point estimate of the i^{th} moment, and $\mu_i(\cdot)$ denotes the central moment of the FGLD.

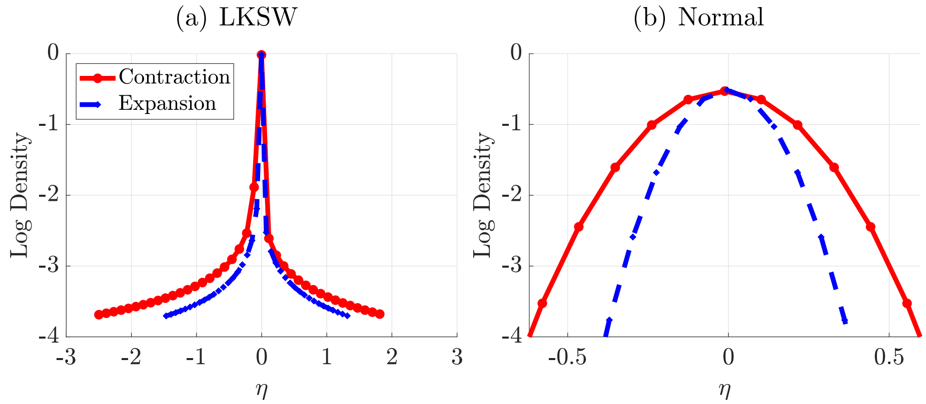
[17] We also impose a minimum post-government household income that remains unchanged across scenarios, i.e., when moving from the scenario with normally distributed shocks to the scenario with leptokurtic and left-skewed shocks, the lowest level of income that households can reach is by construction unchanged. This minimum income is expressed relative to average income. We then adjust incomes such that average income (before multiplying with the age profile) remains 1.

moments are zero, and thus the differences between FGLD distribution NORM and its Gaussian counterpart are captured by the even central moments higher than the fourth (i.e., the sixth, eighth, tenth, etc.). Following our analytical analysis in Section 2, these distributional differences mechanically become less and less important the higher the order is, and eventually their implications for utility (and consumption behavior) depend on risk attitudes. We thus also consider a scenario in which we draw shocks from the Normal distribution and discretize it using standard Gaussian quadrature methods. As documented in Appendix E.1, all quantitative results are numerically almost identical to those obtained for FGLD distribution NORM. We therefore use the latter as our benchmark to which we compare FGLD distribution LKSW. Appendix D.1 reports the estimated, fitted, and discretized moments, as well as the parameter vectors λ for all shocks under the two scenarios NORM and LKSW. In both distribution scenarios we scale down the transitory shocks because part of the estimated variance is likely due to measurement error.[18] Appendix D.3 shows central moments 2-4 in logs and levels that result from our parametrization.

Pension System. Social security benefits follow a fixed replacement schedule that approximates the current US bend point formula. We approximate average indexed monthly earnings (AIME) by the realization of the persistent income shock before entering into retirement z_{j_r-1}. We then apply the bend point formula contained in Appendix D.2 and denote the according model equivalent to the primary insurance amount (PIA) by $p(z_{j_r-1})$. To achieve budget clearing of the pension system, pension payments are further scaled by the aggregate indexation factor ϱ so that individual pension income is $b(z_{j_r-1}) = \varrho \cdot p(z_{j_r-1})$. As to contributions to the pension system, we compute the average contribution rate from the data giving $\tau^p = 11.7\%$ (which is close to the current legislation featuring a marginal contribution rate

[18]Following Huggett and Kaplan (2016) we assume that one third of the estimated variance of the transitory shock is measurement error and reduce the targeted variance accordingly. We assume that this measurement error is symmetric and accordingly adjust the third and fourth central moments such that the implied coefficients of skewness and kurtosis are unchanged.

Figure 3: Discretized Log Distribution Functions: Persistent Shock



Notes: Discretized log distribution functions for the persistent shock η. LKSW: FGLD with estimated variance, skewness, and kurtosis. Markers denote the grid points used in the discretized distribution. Normal: Normal distribution with estimated variance discretized using Gaussian quadrature method. Log density is the base 10 logarithm of the PDF.

of $\tau^p = 12.4\%$). The base for pension contributions in our model is average gross earnings. Since earnings processes in the model are based on net wages—net of all taxes and transfers—and since we normalize average net wages to one, average gross wages are $\frac{1}{1-\tau^p-\tau}$, where τ is some average labor income tax rate (including transfers). We compute τ from the data giving $\tau = 16.88\%$.

Since average labor productivity, the means of the shocks z_j, ϵ_j as well as the total population in age group j are all normalized to one, efficiency weighted aggregate labor in the economy is equal to $j_r - 1$. The number of pensioners is $J - j_r + 1$. The pension budget is therefore given by

$$\tau^p \cdot \frac{1}{1-\tau-\tau^p} \cdot (j_r - 1) = \varrho \cdot \int p(z_{j_r-1}) d\Phi(z_{j_r-1}) \cdot (J - j_r + 1) \,.$$

We calibrate ϱ in each distribution scenario so that the pension budget clears. Since contributions obey a linear tax schedule and by our normalization of income, aggregate contributions are constant across all scenarios. Recalibrating ϱ therefore implies that also average pension income is the same across all scenarios. Table D.5 in Appendix D.2 provides the accordingly calibrated values of ϱ.

Initial Assets and Interest Rate. For simplicity, we assume that all households are born with the same initial assets $a_0 = \bar{a}_0$. We compute those from the average asset to net earnings data at age 25, which we calculate from PSID data as 0.89. We set the annual interest rate of the risk-free asset to $r = 4.2\%$, based on Siegel (2002).

Table 2: Calibrated Parameters

Working period	25 ($j = 0$) to 60 ($j = j_r - 1$)
Maximum age	80
IES	$\gamma = 1$
RA	$\theta \in \{1, 2, 3, 4\}$
Discount factor (2^{nd} stage)	$\beta \in \{0.971, 0.970, 0.967, 0.965\}$
Interest rate	$r = 0.042$
Pension contribution rate	$\tau^p[\%] = 11.7\%$
Pension benefit level	See Table D.5
Average tax rate	$\tau[\%] = 16.8\%$
Aggregate shocks	$\pi(s' = c \mid s = c) = 0.39, \pi(s' = e \mid s = e) = 0.77$
Initial ass. / inc.	$\bar{a}_0 = 0.89$

Notes: Calibration parameters. IES: inter-temporal elasticity of substitution, RA: coefficient of risk aversion. The discount factor β is calibrated endogenously to match asset to income data from the PSID. The pension benefit level parameter ϱ is calibrated such that the pension budget clears.

Preferences. As we show in Section 2, risk attitudes play a crucial role for the welfare effects of higher-order income risk and for the precautionary savings motive. For each model variant we therefore consider four alternative parameterizations and vary $\theta \in \{1, 2, 3, 4\}$. Throughout, we consider risk-sensitive preferences (Tallarini 2000) and accordingly set the inter-temporal elasticity of substitution to $\gamma = 1$.[19] For each $\theta \in \{1, 2, 3, 4\}$, we determine endogenously the discount factor β to match life-cycle asset profiles scaled by

[19]Cooper and Zhu (2016) estimate a portfolio choice model where agents have Epstein-Zin-Weil preferences, and face the canonical income process with log Normal shocks. They estimate a risk aversion of 4.4 and an IES of 0.6. We choose an IES of 1 as a natural benchmark. This is also very convenient when we decompose the welfare effects as described in Appendix A.7.

net earnings, which we compute from PSID data. Since our model is not designed to match saving patterns in retirement (there is neither survival risk nor a bequest motive), we match assets for ages 25-60, the working period in our model. This calibration is done for distribution scenario LKSW, and we then hold the calibrated discount factor constant when moving to scenario NORM, for each calibration of θ.

Calibrated discount factors range from 0.971 for $\theta = 1$ to 0.965 for $\theta = 4$, see Table 2, which summarizes the calibration of the model. The reason for the decline of the calibrated discount factor in θ is that increasing θ leads to higher precautionary savings which is offset in the calibration by lowering β so that the life-cycle savings motive is less potent.

6 Quantitative Role of Higher-Order Risk

6.1 Welfare Implications of Higher-Order Income Risk

In order to assess the welfare implications of higher-order income risk, we ask which world households would prefer to be born into. Taking this ex-ante perspective, we accordingly define the Utilitarian social welfare function as the expected life-time utility function of households born with initial assets $a_0 = \bar{a}_0$, idiosyncratic persistent income state $z_0 = 0$, and transitory shock $\varepsilon = 0$. Corresponding with our notion of an ex-ante perspective we aggregate expected life-time utilities of newborns in the stationary invariant distribution of the economy. Since the transition probabilities over aggregate states are encoded in the value functions and since aggregate fluctuations in our partial equilibrium model do not affect relative prices, evaluating welfare in the stationary invariant distribution of the economy is equivalent to aggregating newborns' value functions with the stationary invariant distribution of the Markov chain process, Π_s. Accordingly, welfare is given by

$$W = \sum_s \Pi_s V_0(a_0 = \bar{a}_0, z_0 = 0, \varepsilon = 0; s).$$

We then calculate the consumption equivalent variation (CEV) that households need to receive in the world without higher-order risk (distribution scenario NORM) in order to be indifferent to a world with higher-order risk as parameterized by the distribution scenario LKSW. Given the homotheticity of the utility function, the CEV is $g_c = W^{LKSW}/W^{NORM} - 1$.

We distinguish between three different channels through which idiosyncratic risk translates into utility consequences evaluated from this ex-ante perspective, and express the CEV as the sum of three components: $g_c = g_c^{mean} + g_c^{lcd} + g_c^{csd}$ (cf. Appendix A.7 for explicit expressions). While we hold mean income constant, consumption is endogenous. When facing different (distribution) scenarios, households make different savings decisions, and thus realize different mean consumption, i.e., consumption averaged cross-sectionally and over age. We call the welfare consequence of this change of mean consumption the *mean effect*, g_c^{mean}, which is proportional to changes in mean consumption. We in turn refer to utility consequences of changes in the distribution around mean consumption as the *distribution effect*, g_c^{distr}, which we decompose into two components: the utility consequences of, first, the change of the distribution of mean consumption over the life-cycle, the *life-cycle distribution effect*, g_c^{lcd}, and, second, the change of the cross-sectional distribution of consumption around the mean life-cycle profile, the *cross-sectional distribution effect*, g_c^{csd}.

Table 3 summarizes the welfare implications of higher-order income risk by showing the CEV and its decomposition. Consistent with our analytical findings in Proposition 1 (see Appendix A.2) higher-order risk leads to welfare gains when risk attitudes are weak. With stronger risk attitudes, however, welfare losses show up, because the increasing variance and the high kurtosis dominate the welfare effects.

The main force for the welfare results is the redistribution of consumption over the life-cycle reflected in g_c^{lcd}. This is a consequence of increased precautionary savings as reflected in Panel (a) of Figure 4, which displays

Table 3: Welfare Implications of Higher-Order Income Risk: CEV in %

Risk Aversion / CEV	g_c	g_c^{mean}	g_c^{lcd}	g_c^{csd}
$\theta = 1$	0.371	-0.154	0.506	0.019
$\theta = 2$	-0.386	-0.161	-0.256	0.031
$\theta = 3$	-4.488	0.318	-4.751	-0.055
$\theta = 4$	-12.474	1.211	-13.392	-0.294

Notes: Welfare gains (positive numbers) and losses (negative numbers) of higher-order income risk, expressed as a Consumption Equivalent Variation (CEV) in scenario NORM that makes households indifferent to the higher-order income risk scenario LKSW. g_c: total CEV, g_c^{mean}: CEV from changes of mean consumption, g_c^{lcd}: CEV from changes in the distribution of consumption over the life-cycle, g_c^{csd}: CEV from changes in the cross-sectional distribution of consumption, where $g_c = g_c^{mean} + g_c^{lcd} + g_c^{csd}$.

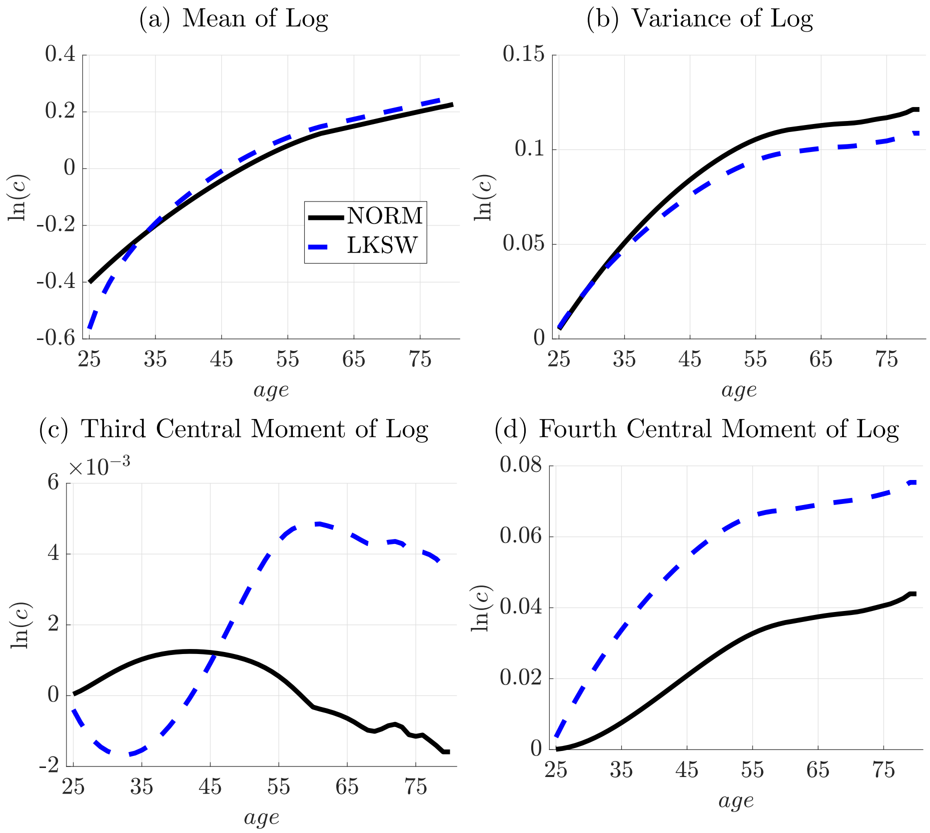
mean log consumption over the life-cycle.[20] Consumption in the higher-order income risk scenario LKSW is lower when young and higher when old compared to scenario NORM. In welfare terms lower consumption when young dominates higher consumption when old due to discounting. The mean effect g_c^{mean} instead is mostly positive because the increased consumption when old dominates (except for $\theta = 1$ and $\theta = 2$). In sum, total welfare losses for scenario LKSW range from about 0.4% (i.e., small gains) for $\theta = 1$ to -12.5% for strong risk attitudes with $\theta = 4$.

Panels (b) to (d) of Figure 4 show the second to fourth central moments of the consumption distribution over the life-cycle, which are relevant for the cross-sectional distribution effect g_c^{csd}. To interpret it observe that the variance of log consumption is lower in scenario LKSW than in scenario NORM for most ages, whereas the third central moment is initially negative and the kurtosis of the log consumption distribution is higher at all ages.[21] The lower

[20]Here we show the profile for a high risk aversion parameter of $\theta = 4$, because in this calibration the effects are most evident visually. Qualitatively, effects are the same in the other risk attitude calibrations. Note that consumption is monotonically increasing over the life-cycle and thus does not display the typical hump-shaped profile, because we do not model life-cycle consumption behavior in retirement.

[21]The Gini coefficient for assets for a risk aversion of 4 is at 0.35 in scenario NORM, and at 0.34 in scenario LKSW. Thus, introducing higher-order income risk does not increase the Gini coefficient in a quantitative model such as ours. Also, note that the Gini coefficient in our calibrated model is substantially lower than in the data and also lower than what

Figure 4: Central Moments of Log Consumption by Age ($\theta = 4$)



Notes: Moments of cross-sectional distribution of log consumption over the life-cycle. NORM: FGLD with moments of the normal distribution, LKSW: FGLD with excess kurtosis and left-skewness (in logs).

variance contributes positively to g_c^{csd}, which dominates for low risk aversion, whereas the negative skewness and the excess kurtosis contribute negatively, and dominate for strong risk attitudes.

is typically found in quantitative work; e.g., Krueger and Ludwig (2016) compute a Gini coefficient of assets of 0.55 in an overlapping generations model calibrated to the US economy. The key reason for the relatively modest asset inequality lies in our focus on ex-post heterogeneity, i.e., the only source of heterogeneity is income risk faced throughout the life cycle.

6.2 Welfare Costs of Cyclical Idiosyncratic Risk

Next, we quantify the utility consequences of *cyclical idiosyncratic risk*. To this end, for each of the two distribution scenarios NORM and LKSW we evaluate the welfare implications for households of facing the *actual* cyclical income process relative to a *counterfactual* income process in which we shut down the cyclical variation of the distribution. By holding mean wages and interest rates constant over the cycle, the welfare effects of cyclical risk we report constitute a lower bound for each scenario.[22]

As before, W^i denotes the social welfare function in the *cyclical risk scenario*, while $W^{i,ncr}$ denotes the social welfare function in the *no cyclical risk scenario*. We then compute the CEV necessary in the scenario with no cyclical risk to be indifferent to the scenario with cyclical risk, $g_c^{i,cr} = W^i/W^{i,ncr} - 1$, and decompose the total CEV from cyclical risk into its components, i.e., $g_c^{i,cr} = g_c^{i,cr,mean} + g_c^{i,cr,lcd} + g_c^{i,cr,csd}$, for $i \in \{NORM, LKSW\}$.

When computing welfare in the non-cyclical scenario $W^{i,ncr}$ we assume that households always draw from the "expansion-distribution" of the scenario rather than taking some weighted average of shock distributions for expansions and contractions. When using log-Normal distributions of shocks, one approach in the literature is to consider an *average* distribution, which features the average of expansion and contraction variances, see for example Storesletten et al. (2001). This approach is not applicable in our analysis as it is conceptually not clear what characterizes an "average" distribution once other moments than the variance are taken into account. To the extent that some average distribution represents a better non-cyclical counterfactual scenario, the pure effect of cyclical idiosyncratic risk is overstated in our analysis.[23] However, we are mainly interested in the difference of welfare costs of

[22]Note that the direct effect of business cycles is typically found to be small. For example, Storesletten et al. (2001) find the direct effect to be an order of magnitude smaller than the role of cyclical variation in idiosyncratic risk. However, there can be indirect utility "interactions" between aggregate and idiosyncratic risk, which may be large (Harenberg and Ludwig 2019), and which we abstract from here in order to focus on the role of the idiosyncratic shock distribution.

[23]Indeed, Storesletten et al. (2001) find welfare costs of cyclical risk of about 1.3%. They consider CRRA preferences with $\theta = 2$. In one of our sensitivity checks in Appendix E.2,

cyclical income risk across scenarios, i.e., the "difference in difference" comparison between $g_c^{LKSW,cr}$ and $g_c^{NORM,cr}$, i.e., $\Delta g_c^{cr} = g_c^{LKSW,cr} - g_c^{NORM,cr}$. Thus, our approach to "normalize" the economy without cyclical idiosyncratic risk is of second order importance as it is consistent across scenarios.[24]

Table 4 reports the results on the welfare costs of cyclical idiosyncratic risk in scenarios NORM and LKSW. First, note that consistent with our theoretical analysis of Section 2 in each scenario the welfare costs of business cycles increase monotonically in θ. Second, as for the welfare costs of higher-order risk, the main contributor to the welfare consequences is the redistribution of consumption over the life-cycle as quantified by g_c^{lcd}. Third, mean effects are positive. Recall that a negative g_c^{lcd} is a consequence of the counter-clockwise tilting of the consumption profile because of increased precautionary savings. Higher savings increase consumption in the middle of the life-cycle, which pushes up mean consumption. As previously, on average over the life-cycle this second effect dominates.

Consistent with the afore documented result (and with our theoretical analysis of Section 2) that with logarithmic utility the total welfare effect from higher-order income risk is positive for scenario LKSW, we now correspondingly find that welfare losses from cyclical idiosyncratic risk are about 0.28%p lower in scenario LKSW (last column in first panel of Table 4). Similarly, with moderate risk attitudes (risk aversion of 2), the welfare implications of cyclical income risk in scenario LKSW are only mildly higher than those obtained in scenario NORM. With strong risk attitudes ($\theta = 4$), the welfare losses compared to scenario NORM are significantly higher: They are about 6.4%p higher in scenario LKSW.

In Appendix E.2 we investigate the sensitivity of our results with respect to selected modeling and calibration assumptions. Specifically, we consider an

---

we also consider CRRA preferences with $\theta = 2$. In this case we obtain welfare costs of about 2.6% in scenario NORM. Besides other differences between our model and theirs, one reason for the higher welfare costs in our analysis lies in the different approach to characterizing the non-cyclical scenario.

[24]One alternative is to follow the "integrating out" principle (see Krusell and Smith 1999 and Krusell et al. 2009), which first isolates a true idiosyncratic component of the shock, and then integrates over the probability distribution of the aggregate state.

Table 4: Welfare Effects of Cyclical Idiosyncratic Risk

CEV	g_c	g_c^{mean}	g_c^{lcd}	g_c^{csd}	Δg_c
		Risk Aversion, $\theta = 1$			
NORM	-1.720	0.499	-2.175	-0.044	0
LKSW	-1.443	0.398	-1.806	-0.035	0.277
		Risk Aversion, $\theta = 2$			
NORM	-3.263	0.898	-4.038	-0.123	0
LKSW	-3.516	0.823	-4.228	-0.111	-0.253
		Risk Aversion, $\theta = 3$			
NORM	-4.607	1.229	-5.638	-0.198	0
LKSW	-7.177	1.379	-8.313	-0.243	-2.570
		Risk Aversion, $\theta = 4$			
NORM	-5.758	1.515	-7.009	-0.264	0
LKSW	-12.171	1.944	-13.686	-0.429	-6.413

Notes: Consumption Equivalent Variation in the non-cyclical scenario that makes households indifferent to the cyclical scenario. $g_c = g_c^{mean} + g_c^{lcd} + g_c^{csd}$, where g_c^{mean}: CEV from change of mean consumption, g_c^{lcd}: CEV from changes in the distribution of consumption over the life-cycle, g_c^{csd}: CEV from changes in the cross-sectional distribution. $\Delta g_c = g_c^{LKSW} - g_c^{NORM}$.

expected utility formulation with CRRA preferences where we restrict $\theta = \frac{1}{\gamma}$, we analyze the role of borrowing constraints in the model, and we investigate how results are affected by our choice of the interest rate. The results shown in the appendix underscore the robustness of our findings also with respect to the dominant role played by the life-cycle distribution effect g_c^{lcd}.

Furthermore, in Appendix E.3 we analyze an alternative distribution scenario, which features shocks that have excess kurtosis, but are symmetric (in logs). In the calibration with $\theta = 4$, welfare costs of cyclical risk are about 4% higher in this distribution scenario than in scenario NORM (see table E.3). Combined with the lower part of table 4 we thus find that of the differential welfare losses from higher-order risk approximately 62%($\approx 3.98/6.41 \cdot 100\%$) are due to the excess kurtosis and the remaining 38% are due to the left-skewness of shocks.

Finally, in Appendix E.4 we discuss a general equilibrium extension of our model, in which we take into account the general equilibrium effect of higher-order idiosyncratic risk on wages and interest rates, that works through the equilibrium capital allocation. We still abstract from aggregate productivity risk. We consider a solution that fully reflects higher-order risk, while being an approximation regarding its cyclicality. The equilibrium feedback effect turns out to be weak, which is explained by the life-cycle structure of the economy and the consumption profile of Figure 4. When facing scenario LKSW instead of scenario NORM, young agents have higher precautionary savings, however these savings will be dis-saved at old age. Thus, aggregate savings of the economy do not differ strongly.

Summing up, we can conclude that the welfare effects of cyclical risk are strongly underestimated in conventional approaches based on Gaussian distributions of innovations if risk attitudes are strong (levels of θ of 3 or 4) and that both features of higher-order income risk—excess kurtosis and left skewness—are quantitatively important for this finding, whereby about 60% of the effect can be attributed to the excess kurtosis.

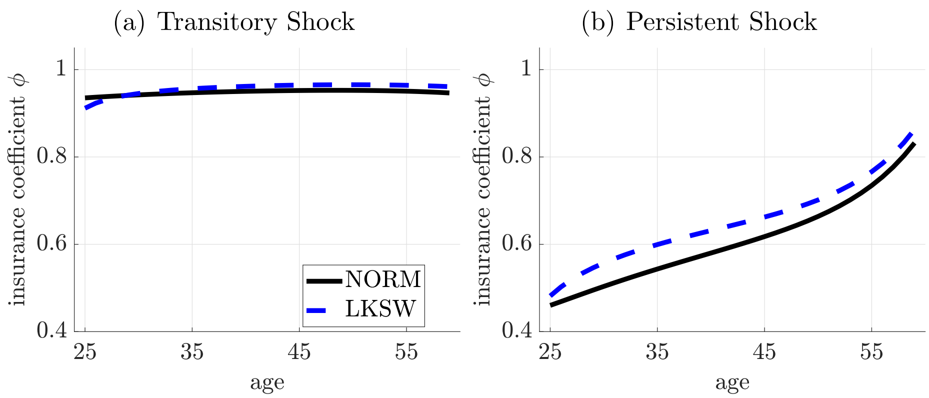
6.3 Insurance Against Idiosyncratic Risk

Finally, we adopt concepts developed in the literature on consumption insurance (Blundell et al. 2008; Kaplan and Violante 2010) to ask how households are self-insured against income shocks $x_j(s) \in \{\varepsilon_j, \eta_j(s)\}$ and how this insurance varies across scenarios. In the model, the transitory and persistent shocks are directly observed and thus we adopt the measure of Kaplan and Violante (2010) to our setting with cyclical risk. Conditional on today's aggregate state s, the insurance coefficient $\phi_j^x(s)$ is given as the share of the variance of next period's shock $x_{j+1}(s')$ that does not translate into consumption growth, and thus the pass-through coefficient $1 - \phi_j^x(s)$ is the coefficient of a linear regression of consumption growth on shock x, which captures how strongly the shock translates into consumption:

$$1 - \phi_j^x(s) = \frac{cov(\Delta \ln (c_{j+1}(s' \mid s)), x_{j+1}(s'))}{var(x_{j+1}(s'))}, \tag{13}$$

for $\Delta \ln(c_j(s' \mid s)) = \ln(c_{j+1}(s' \mid s)) - \ln(c_j(s))$.

Figure 5: Insurance Coefficients: Strong Risk Attitudes, $\theta = 4$



Notes: Figures show the degree of consumption insurance against transitory and persistent shocks separately by age.

Figure 5 reports the insurance coefficients ϕ_j^x for all ages $j \in \{0, \ldots, J\}$, as a weighted average of the coefficients in contractions and expansions[25] for the transitory shock ε in Panel (a) and for the persistent shock $\eta(s)$ in Panel (b). Results are quantitatively similar for different values of risk attitudes, so we discuss only the numbers for $\theta = 4$. For scenario LKSW, consumption insurance against both transitory and persistent shocks is improved relative to scenario NORM as measured by the ϕ-coefficients. This is a direct consequence of increased precautionary savings, which lead to shocks translating less into consumption.

Do the higher insurance coefficients in scenario LKSW really represent *better insurance*, though? Arguably, better insurance would mean that negative shocks translate less into consumption. This is not the case as can be illustrated by one simple decomposition of the pass-through of shocks to consumption changes in equation (13). Consider the aggregate (integrating over

[25]We weigh with the stationary invariant distribution Π_s.

age and averaging over states s) pass-through coefficient for shock $x \in \eta, \varepsilon$:

$$\begin{aligned} 1-\phi^x = & \frac{E\left[\Delta \ln(c(\cdot))x\right] - E\left[\Delta \ln(c(\cdot))\right]E\left[x\right]}{var(x)} \\ = & \frac{E\left[\Delta \ln(c(\cdot))x \middle| x>0\right]}{var(x)} + \frac{E\left[\Delta \ln(c(\cdot))x \middle| x<0\right]}{var(x)} - \frac{E\left[\Delta \ln(c(\cdot))\right]E\left[x\right]}{var(x)}. \end{aligned} \tag{14}$$

The first two components of the sum in equation (14) give the contribution to the overall pass-through coefficient of comovements of consumption with positive and negative shocks, respectively. Table 5 shows the aggregate pass-through coefficient of the economy along with the contributions of its components. As already learned from Figure 5, the aggregate pass-through of both transitory and persistent shocks is smaller in scenario LKSW (insurance coefficient is larger). Now consider the contribution of positive and negative shocks to the aggregate pass-through coefficient. In scenario NORM, negative transitory shocks do not translate into negative consumption changes: comovements with negative realizations of ε contribute -3.4% to the pass-through coefficient. In scenario LKSW, the (negative) consumption reaction to negative shocks is important: 30.2% of the pass-through coefficient are accounted for by negative transitory shocks leading to negative consumption adjustments. At the same time, consumption reacts less strongly to positive changes. Thus, the fact that the aggregate pass-through is smaller (the insurance coefficient is larger) is indeed explained by increased precautionary savings. However, built-up savings do not suffice to smooth out the negative shocks in scenario LKSW as well as they do in scenario NORM.

For persistent shocks, the same mechanics are at work. In scenario NORM, about 40% of the pass-through coefficient is generated by consumption reductions with negative shocks, while about 59% come from consumption increases with positive shocks. In scenario LKSW, negative shocks pass-through more (51% of overall), and positive shocks pass-through less (46%). So for both transitory and persistent shocks, the reduction of the pass-through (increase of insurance coefficient) when moving from scenario NORM to scenario LKSW

Table 5: Aggregate Pass-Through and its Decomposition, $\theta = 4$

Transitory	$1-\phi^{\varepsilon}$	$E[\Delta^c \cdot \epsilon, \epsilon < 0]$	$E[\Delta^c \cdot \epsilon, \epsilon > 0]$	$-E[\Delta^c] \cdot E[\epsilon]$
NORM	0.055	-0.034	0.898	0.136
LKSW	0.047	0.302	0.525	0.173
Persistent	$1-\phi^{\eta}$	$E[\Delta^c \cdot \eta, \eta < 0]$	$E[\Delta^c \cdot \eta, \eta > 0]$	$-E[\Delta^c] \cdot E[\eta]$
NORM	0.395	0.395	0.586	0.019
LKSW	0.353	0.514	0.458	0.028

Notes: Table shows aggregate consumption pass-through coefficient (1-insurance coefficient), and its decomposition into components according to equation 14. Values are expressed as shares of total pass-through. $\Delta^c = \Delta \ln(c(\cdot))$.

is driven by an increased propensity to save, while at the same time negative shocks actually translate more into consumption.

We can thus conclude that in an economy with higher-order income risk aggregate insurance (or pass-through) coefficients are imprecise measures of insurance against risk, if one plausibly has in mind that better insurance means that negative shocks translate less into consumption.

7 Conclusion

We first develop a novel Generalized Method of Moments estimator of higher-order income risk, that starts out with the canonical income process, which captures key features of labor income risk as a combination of persistent and transitory income shocks. We show how the second to fourth central moments of the distributions of shocks can be estimated. Our estimates on PSID household-level earnings imply that the distribution of persistent income shocks exhibits strong cyclicality: the variance is countercyclical, while the third central moment is procyclical. All shock components exhibit strong excess kurtosis. The existing tax and transfer system dampens both transitory and persistent income shocks and reduces cyclicality.

In the second part of the paper we analyze the relevance of the identified *higher-order risk* from an economic perspective. Within an otherwise standard partial equilibrium life-cycle model with incomplete markets, we find that, first, higher-order risk has important welfare consequences—relative to a world with log-Normal shocks. Second, the presence of higher-order risk matters for the welfare costs of business cycles. Third, higher-order income risk affects the degree of consumption self-insurance.

References

Angelopoulos, K., S. Lazarakis, and J. Malley (2019). Cyclical Income Risk in Great Britain. Technical report, University of Glasgow.

Arellano, M., R. Blundell, and S. Bonhomme (2017). Earnings and Consumption Dynamics: A Nonlinear Panel Data Framework. *Econometrica 85*(3), 693–734.

Bayer, C. and F. Juessen (2012). The Life-Cycle and the Business-Cycle of Wage Risk. Cross-Country Comparisons. *Economics Letters 117*, 831–833.

Beaudry, P. and C. Pages (2001). The Cost of Business Cycles and the Stabilization Value of Unemployment Insurance. *European Economic Review 45*(8), 1545–1572.

Blundell, R., L. Pistaferri, and I. Preston (2008). Consumption Inequality and Partial Insurance. *American Economic Review 98*(5), 1887–1921.

Busch, C., D. Domeij, F. Guvenen, and R. Madera (2020). Skewed Idiosyncratic Income Risk over the Business Cycle: Sources and Insurance. *American Economic Journal: Macroeconomics (forthcoming)*.

Catherine, S. (2019). Countercyclical Labor Income Risk and Portfolio Choices over the Life-Cycle. Technical report, University of Pennsylvania.

Civale, S., L. Diéz-Catalán, and F. Fazilet (2017). Discretizing a Process with Non-zero Skewness and High Kurtosis. Technical report, University of Minnesota.

Constantinides, G. M. and D. Duffie (1996). Asset Pricing with Heterogeneous Consumers. *Journal of Political Economy 104*(2), 219–240.

Cooper, R. and G. Zhu (2016, April). Household Finance Over the Life-Cycle: What Does Education Contribute? *Review of Economic Dynamics 20*(1), 63–89.

De Nardi, M., G. Fella, M. Knoef, G. Paz-Pardo, and R. Van Ooijen (2020). Family and Government Insurance: Wage, Earnings, and Income Risks in the Netherlands and the U.S. *Journal of Public Economics (forthcoming)*.

De Nardi, M., G. Fella, and G. Paz-Pardo (2020). Nonlinear Household Earnings Dynamics, Self-Insurance, and Welfare. *Journal of the European Economic Association 18*(2), 890–926.

Druedahl, J. and A. Munk-Nielsen (2018). Higher-order Income Dynamics with Linked Regression Trees. Technical report, University of Copenhagen.

Eeckhoudt, L. and H. Schlesinger (2006). Putting Risk in its Proper Place. *Amercian Economic Review 96*(1), 280–289.

Eeckhoudt, L. and H. Schlesinger (2008). Changes in Risk and the Demand for Saving. *Journal of Monetary Economics 55*(7), 1329–1336.

Epstein, L. G. and S. Zin (1989). Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework. *Econometrica 57*(4), 937–969.

Epstein, L. G. L. and S. Zin (1991). Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: An Empirical Analysis. *Journal of Political Economy 99*(4), 263–286.

Freimer, M., C. T. Lin, and G. S. Mudholkar (1988). A Study Of The Generalized Tukey Lambda Family. *Communications in Statistics - Theory and Methods 17*(10), 3547–3567.

Golosov, M., M. Troshkin, and A. Tsyvinski (2016). Redistribution and Social Insurance. *American Economic Review 106*(2), 359–386.

Gottschalk, P. and R. Moffitt (1994). The Growth of Earnings Instability in the U.S. Labor Market. *Brookings Papers on Economic Activity 25*(2), 217–272.

Guvenen, F., F. Karahan, S. Ozkan, and J. Song (2016). What Do Data on Millions of U.S. Workers Reveal about Life-Cycle Earnings Dynamics?

Guvenen, F., S. Ozkan, and J. Song (2014). The Nature of Countercyclical Income Risk. *Journal of Political Economy 122*(3), 1–59.

Harenberg, D. and A. Ludwig (2019). Idiosyncratic Risk, Aggregate Risk, and the Welfare Effects of Social Security. *forthcoming: Internation Economic Review*.

Huggett, M. and G. Kaplan (2016). How Large is the Stock Component of Human Capital? *Review of Economic Dynamics 22*, 21–51.

Imrohoroglu, A. (1989). Cost of Business Cycles With Indivisibilities and Liquidity Constraints. *Journal of Political Economy 97*(6), 1364–1383.

Kaplan, G. and G. L. Violante (2010). How much consumption insurance beyond self-insurance? *American Economic Journal: Macroeconomics 2*(4), 53–87.

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

Krebs, T. (2003). Growth and Welfare Effects of Business Cycles in Economies with Idiosyncratic Human Capital Risk. *Review of Economic Dynamics 6*(4), 846–868.

Krebs, T. (2007). Job Displacement Risk and the Cost of Business Cycles. *The American Economic Review 97*(3), 664–686.

Krueger, D. and A. Ludwig (2016). On the Optimal Provision of Social Insurance: Progressive Taxation versus Education Subsidies in General Equilibrium. *forthcoming: Journal of Monetary Economics*.

Krusell, P., T. Mukoyama, A. Sahin, and A. A. Smith, Jr. (2009, July). Revisiting the Welfare Effects of Eliminating Business Cycles. *Review of Economic Dynamics 12*(3), 393–402.

Krusell, P. and A. A. Smith, Jr. (1999). On the Welfare Effects of Eliminating Business Cycles. *Review of Economic Dynamics 2*(1), 245–272.

Lajeri-Chaherli, F. (2004). Proper Prudence, Standard Prudence and Precautionary Vulnerability. *Economics Letters 82*(1), 29–34.

Lakhany, A. and H. Mausser (2000). Estimating the Parameters of the Generalized Lambda Distribution. *ALGO Research Quarterly 3*(3), 47–58.

Lucas, R. E. (1987). *Models of business cycles*. New York: Basil Blackwell.

Lucas, R. E. (2003). Macroeconomic Priorities. *American Economic Review 93*(1), 1–14.

Mankiw, N. G. (1986). The Equity Premium and the Concentration of Aggregate Shocks. *Journal of Financial Economics 17*(1), 211–219.

McKay, A. (2017). Time-Varying Idiosyncratic Risk and Aggregate Consumption Dynamics. *Journal of Monetary Economics 88*, 1 – 14.

Moffitt, R. and P. Gottschalk (2011). Trends in the Transitory Variance of Male Earnings in the U.S., 1970-2004. Working Paper 16833, NBER.

Panel Study of Income Dynamics (PSID) (2015). Public Use Dataset, Produced and Distributed by the Survey Research Center, Institute for Social Research, University of Michigan, Ann Arbor, MI.

Siegel, J. J. (2002). *Stocks for the Long Run : The Definitive Guide to Financial Market Returns and Long-Term Investment Strategies*. New York: McGraw-Hill.

Storesletten, K., C. Telmer, and A. Yaron (2007). Asset Pricing with Ideoyncratic Risk and Overlapping Generations. *Review of Economic Dynamics 10(4)*, 519–548.

Storesletten, K., C. I. Telmer, and A. Yaron (2001). The Welfare Cost of Business Cycles Revisited: Finite Lives and Cyclical Variation in Idiosyncratic Risk. *European Economic Review 45*(7), 1311 – 1339.

Storesletten, K., C. I. Telmer, and A. Yaron (2004). Cyclical Dynamics in Idiosyncratic Labor Market Risk. *The Journal of Political Economy 112*(3), pp. 695–717.

Su, S. (2007). Numerical Maximum Log Likelihood Estimation for Generalized Lambda Distributions. *Computational Statistics and Data Analysis 51*(8), 3983–3998.

Tallarini, T. D. J. (2000). Risk-Sensitive Real Business Cycles. *Journal of Monetary Economics 45*(3), 507–532.

Weil, P. (1989). The Equity Premium Puzzle and the Risk-Free Rate Puzzle. *Journal of Monetary Economics 24*(3), 401–421.