// NO.20-022 | 05/2020

DISCUSSION PAPER

// CHRISTOPHER BUSCH
AND ALEXANDER LUDWIG

Higher-Order Income Risk over the Business Cycle

Leibniz Association

Higher-Order Income Risk over the Business Cycle*

Christopher Busch[†] Alexander Ludwig[‡]

March 27, 2020

Abstract

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

Keywords: Labor Income Risk, Business Cycle, GMM Estimation, Skewness, Persistent and Transitory Income Shocks, Risk Attitudes, 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, 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 letting us use some of her files 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; SAFE; CEPR; Netspar; MEA; ZEW; 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

It has long been established in the empirical macroeconomics literature that individual income risk varies with the aggregate state of the economy, and that this has important implications for the evaluation of many macroeconomic questions that pertain to the business cycle. The traditional way to capture cyclical changes of individual risk in macroeconomic analyses is to model idiosyncratic shocks with a larger variance in aggregate contractions. However, a growing body of recent empirical evidence challenges this focus on the variance, and emphasizes important deviations of the distribution of individual income changes from a (implicitly assumed) Gaussian distribution, namely non-zero skewness and high kurtosis. We refer to these deviations as capturing *higher-order income risk*.[1]

The first contribution of this paper is a novel parametric approach to estimate idiosyncratic labor income risk and its cyclicality. Within our estimation framework we can transparently identify skewness and kurtosis of both transitory and persistent shocks. To achieve this, we extend the canonical income process to account for higher-order risk.[2] We estimate the process for household level labor income and for post government income (after taxes and transfers) using household panel data from the United States.

The second contribution is that we systematically evaluate the role of higher-order risk for three fundamental, and related, questions that pertain to (cyclical) idiosyncratic risk. First, does higher-order idiosyncratic risk have (economically relevant) implications for welfare? Second, does cyclical higher-order idiosyncratic risk matter for the welfare costs of business cycles? Third, does higher-order idiosyncratic risk matter for self-insurance through savings? The answer to all three questions turns out to be *yes*. Our tool is a standard incomplete-markets life-cycle model, in which households face an exogenous income process estimated on post government income in the United States which features the estimated variance, skewness, and kurtosis of transitory and persistent income shocks.

In the estimation of our income process we do not impose any parametric distribution function on the transitory and persistent components. We characterize both shocks by their central moments and estimate those by the Generalized Method of Moments (GMM). Other than traditionally done in similar estimations, we do not base the estimation solely on the variance-covariance matrix of incomes. Instead, we use the second to fourth central moments and co-moments. This allows us to identify variance, skewness, and kurtosis of the distributions of the shock components. Through this we draw

[1]Of course, a stochastic income process does not necessarily measure *risk*. In our model analysis the estimated income process is exogenous to agents, and thus within the model the shocks of the stochastic process represent risk.

[2]Modelling individual (or household) income dynamics as a combination of transitory and persistent components dates back at least to Gottschalk and Moffitt (1994). It then became a standard input in life cycle household models of consumption and savings.

a richer image of income dynamics within the otherwise traditional transitory-persistent framework.

Our estimation procedure extends the approach taken in Storesletten et al. (2004), who estimate an income process with state-contingent variance of the persistent income shock. They analyze household-level income including government transfers from the Panel Study of Income Dynamics and find that the variance is higher in contractions, i.e., they find *countercyclical variance*.[3] Their identification of the state-dependent variance builds on the observation that persistent shocks accumulate over the life cycle such that the distribution of labor incomes observed for a given cohort widens as this cohort ages. This implies that cohorts that experienced different macroeconomic histories will feature different cross-sectional age-specific variances of labor incomes—if the variance of income shocks varies systematically over the business cycle.[4] In our extended version of the estimator, we allow the second to fourth central moments to be state-contingent. Identification follows from the fact that the accumulated second to fourth central moments differ across cohorts if these cohorts experience different macroeconomic histories—again, if these moments differ systematically over the business cycle.[5] 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 and the persistence of the shocks, which we identify 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.

We apply the estimation to survey data from the Panel Study of Income Dynamics (PSID). The survey allows us to control for a rich set of household-level information and to take into account several relevant transfer components. 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 transfers) is compressed relative to the respective shocks to pre-government income, but persistent shocks remain significantly cyclical. 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

[3] This terminology has been introduced in the macroeconomic asset pricing literature, see Mankiw (1986), Constantinides and Duffie (1996), and Storesletten, Telmer, and Yaron (2007).

[4] *History* here refers to a sequence of expansions and contractions.

[5] Note that we do not base identification on the stardardized moments (the coefficients of skewness and kurtosis), which unlike the central moments do not simply accumulate while a cohort ages.

more mass is concentrated in the center relative to the tails of the distribution. These findings are in line with recent empirical evidence as summarized below.

In our quantitative model, agents 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. The shocks of the income process are drawn from a parametric distribution function,[6] which we fit to the estimated central moments of the transitory and persistent shocks. The distribution of persistent shocks varies systematically over the business cycle as estimated in the data. We are interested in the role of cyclical changes in idiosyncratic risk, and in the relevance of higher-order risk, and thus we normalize all shocks in levels, which implies that the economy does not feature aggregate risk. The only means of self-insurance against income risk is a risk-free asset. Agents have recursive preferences over consumption a la Epstein and Zin (1989, 1991), and Weil (1989), which allows us to separately control the intertemporal elasticity of substitution and risk aversion of households. We then assess whether the estimated deviation of shocks from log-Normal shocks is relevant from a macroeconomic perspective.

First, does higher-order risk have economically relevant welfare implications relative to a world with log-Normal shocks?—*Yes.* We evaluate welfare from an ex ante perspective and show that the direction of welfare effects depends on relative risk attitudes of households (relative risk aversion, relative prudence, and relative temperance). When risk attitudes are strong, the introduction of higher-order risk has sizable negative welfare implications; when risk attitudes are weak (specifically, for log utility), the welfare effect can be positive. The dominant economic mechanism driving the welfare results is an expected reallocation of consumption over the life-cycle. When facing riskier income, risk-averse agents have more precautionary savings, and thus less consumption at young ages.[7]

Second, does higher-order risk matter for the welfare costs of business cycles?—*Yes.* Since Lucas (1987, 2003) argued in a representative agent framework that the gains of smoothing the business cycle beyond what the existing tax and transfer system does would be small, several studies (summarized in Section 2) emphasized the role of both ex-ante and ex-post heterogeneity for the welfare costs of business cycles. We follow up on this, and explore the implications of cyclical *higher-order* income risk. When we take into account excess kurtosis and skewness fluctuations, we find welfare costs of business cycles computed as a consumption equivalent variation which are $0.3\%p$ (for relative risk

[6]We use the *Flexible Generalized Lambda Distribution* developed by Freimer et al. (1988).

[7]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. Thus, parts of the results can be understood by the convention of characterizing risk by the distribution in logs.

aversion of 2) to 6.4%p (for relative risk aversion of 4) larger than under log-Normal shocks.

Third, does the presence of higher-order risk affect the degree of self-insurance?—*Yes.* We employ a measure of self-insurance motivated by Blundell et al. (2008), who suggest to evaluate the degree of partial insurance against income shocks by identifying transitory and permanent shocks to income and estimate the pass-through of the identified 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. Our results show that when incorporating higher-order risk, the model can be brought closer to the empirical estimates because the pass-through of income shocks to consumption is weaker. However, we also find that 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 Normal shocks), which implies that consumption reacts weaker to *positive* transitory and persistent shocks. *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 introduced in the literature by Blundell et al. (2008) for the analysis of the degree of partial insurance against income risk.

Before delving into the quantitative analysis, we analyze the effects of higher-order risk in a simple two-period model, in which agents face risky second period income. We analytically derive the implications of higher-order risk for life-time utility and precautionary savings. We explore in detail how risk attitudes of households matter delivering two main insights. First, larger higher-order risk (in particular: left-skewness) can have positive welfare implications (with log-utility), and second and related, the reaction of precautionary savings to larger higher-order risk is ambiguous. These results do not depend on a parametric assumption for the distribution of shocks and prove useful for interpreting the quantitative results.

The remainder of the paper is structured as follows. Section 2 places our analysis in the literature. Section 3 provides guidance for the analysis by discussing the role of higher-order risk in a simple two-period model. Section 4 first presents our empirical approach, and discusses identification of the income process. The remainder of the section presents the results of applying our approach to US earnings data from the PSID. Section 6 introduces the quantitative model to analyze the economic implications of higher-order income risk, Section 7 discusses the quantitative results, and Section 8 concludes.

2 Relation to the Literature

On the empirical side, many studies analyze (residual) income inequality over time. Prominent examples for the United States are Gottschalk and Moffitt (1994), Heathcote et al. (2010), and Moffitt and Gottschalk (2011) who document the development of residual inequality over the past decades. The focus of our study is on the systematic variation of the distribution of income changes over the business cycle. In a seminal contribution, Storesletten et al. (2004) estimate a countercyclical variance of persistent shocks to household-level income using PSID data. Building on the conceptual framework of Storesletten et al. (2004), Bayer and Juessen (2012) focus on residual hourly wages (at the household level) and estimate countercyclical dispersion of persistent shocks in the United States (PSID). Our empirical approach nests Storesletten et al. (2004) as a special case. Specifically, comparing our estimates to theirs, we find a similar magnitude of the cyclicality of dispersion.

Recently, Guvenen et al. (2014) stress that the focus on the variance of log income changes alone misses the main characteristic of how individual risk varies with the aggregate state of the economy. They use an extensive administrative dataset from US social security records for males. Their findings 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, while the variance is unchanged over the business cycle. Related, Busch et al. (2018) conduct a non-parametric analysis of individual and household earnings dynamics in Germany, Sweden, and the US. They find qualitatively the same dynamics as we do: individual and household-level earnings changes are more left-skewed in contractionary times, which suggests increased downside risk in contractions.

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. Druedahl and Munk-Nielsen (2018) use a regression tree approach to document similar dynamics for Danish males. Turning again to households, Arellano et al. (2017) document rich deviations from the canonical income process for household-level earnings in the United States (using survey data from the PSID) and Norway (using administrative data); to which De Nardi et al. (2019) add additional evidence for the Netherlands (using administrative data). Relative to those recent papers on income dynamics, we stick to the transitory-persistent decomposition of the canonical income process and extend it by considering the second to fourth moments of all shocks. Motivated by the recent empirical evidence, we allow the third moment of the persistent shocks to vary with the aggregate state of the economy, similar to Huggett and Kaplan (2016). One other recent paper is

Angelopoulos et al. (2019), who adapt a version of our GMM estimator to document procyclical skewness of persistent shocks in Great Britain using data from the British Household Panel Study.

Recently, the new evidence on richer earnings dynamics found its way into macroeconomic studies. For example, Golosov et al. (2016) allow for time-varying skewness of idiosyncratic risk in a study of optimal fiscal policy. Our paper is part of a growing literature that explicitly analyzes the implications of the new insights on richer earnings dynamics for macroeconomic questions. Catherine (2019) analyzes the implications of procyclical skewness of idiosyncratic income risk for the equity premium. McKay (2017) links procyclical skewness to aggregate consumption dynamics. 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. Closest to our paper is De Nardi et al. (2019), who apply the estimation of Arellano et al. (2017) to household-level income data from the PSID. They then feed this fitted income process into an otherwise standard incomplete markets model to study the role of richer earnings dynamics for consumption insurance and the welfare costs of idiosyncratic risk in comparison to a standard income process with log-normal shocks. Our analysis differs in two ways from theirs. First, they do not consider cyclicality of idiosyncratic risk, which is our main focus. Second, our analysis provides a transparent link of non-Gaussian moments of the distribution of shocks to macroeconomic implications.

Our analysis of the role of higher-order risk for the welfare costs of business cycles speaks to a rich literature that evolved after Lucas (1987). Imrohoroglu (1989) was the first study that analyzed the role of idiosyncratic risk and incomplete markets for the welfare costs of business cycles. Following up on her analysis, several studies emphasize in particular the role of unemployment risk (e.g., Krusell and Smith 1999, Krusell et al. 2009, Krebs, 2003, 2007, and Beaudry and Pages 2001). Dolmas (1998) and Epaulard and Pommeret (2003) both consider Epstein-Zin-Weil preferences, which is also the preference specification employed by us.[8] Closest to our paper is Storesletten et al. (2001), who analyze the welfare consequences of cyclical idiosyncratic risk in an incomplete markets model. They represent idiosyncratic risk by the income process with cyclical variance of persistent shocks as estimated in Storesletten et al. (2004). In the same fashion, we take our estimated income process as an exogenous income process in an incomplete markets model and assess the role of systematic changes of this risk over the business cycle.

[8]Epaulard and Pommeret (2003) study the relationship between cyclical variation and growth in an endogenous growth model, which is a strand of the literature we do not talk to. For an overview of studies that analyze the relationship between business cycles and growth see Barlevy (2005).

3 Higher-Order Risk in a Two-Period Model

3.1 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, $j \in \{0,1\}$. Period 0 consumption enters directly into utility, whereas risky period 1 consumption is transformed by function $v(c_1, \theta, \Psi)$, where θ parameterizes risk attitudes, and Ψ is the distribution function of income shocks ε.

Per-period utility takes a power form, and we adopt recursive preferences[9] a la Epstein and Zin (1989, 1991), and Weil (1989):[10]

$$U = \begin{cases} \frac{1}{1-\gamma}\left(c_0^{1-\frac{1}{\gamma}} + v\left(c_1, \theta, \Psi\right)^{1-\frac{1}{\gamma}}\right) & \text{for } \gamma \neq 1 \\ \ln\left(c_0\right) + \ln\left(v\left(c_1, \theta, \Psi\right)\right) & \text{for } \gamma = 1. \end{cases} \tag{1}$$

Thus, γ can be interpreted as the inter-temporal elasticity of substitution between c_0 and $v(\cdot)$, where $v(\cdot)$ represents the certainty equivalent from consumption in the second period, which is given by

$$v(c_1, \theta, \Psi) = \begin{cases} \left(\int c_1(\varepsilon)^{1-\theta} d\Psi(\varepsilon)\right)^{\frac{1}{1-\theta}} = \left(\mathbb{E}\left[c_1^{1-\theta}\right]\right)^{\frac{1}{1-\theta}} & \text{for } \theta \neq 1 \\ \exp\left(\int \ln(c_1(\varepsilon)) d\Psi(\varepsilon)\right) = \exp\left(\mathbb{E}\left[\ln(c_1)\right]\right) & \text{for } \theta = 1. \end{cases} \tag{2}$$

[9]In the two-period model the notion of *recursive preferences* is not strictly speaking correct. We introduce this terminology here as we adopt Epstein-Zin-Weil preferences in the quantitative life-cycle model.

[10]Notice that our representation of Epstein-Zin-Weil preferences, which goes back to Selden (1978, 1979), is a monotone transformation of the standard Epstein-Zin-Weil aggregator

$$V = \begin{cases} \left(c_0^{1-\frac{1}{\gamma}} + v(c_1, \theta, \Psi)^{1-\frac{1}{\gamma}}\right)^{\frac{1}{1-\frac{1}{\gamma}}} & \text{for } \gamma \neq 1 \\ c_0 \cdot v(c_1, \theta, \Psi) & \text{for } \gamma = 1, \end{cases}$$

where $U = \frac{1}{1-\frac{1}{\gamma}} V^{1-\frac{1}{\gamma}}$ if $\gamma \neq 1$ and $U = \ln(V)$ if $\gamma = 1$.

The specification of preferences gives standard CRRA (constant relative risk aversion) preferences if the measure of the IES γ and the measure of risk aversion θ are reciprocals: $\theta = \frac{1}{\gamma}$. Note that we assume an interest rate of zero and no discounting of second-period utility, which implies that there is no life-cycle savings motive in this model.

3.2 Analysis

3.2.1 Hand-to-Mouth Consumers

We first analyze the role of higher-order risk for hand-to-mouth consumers. To this end, we shut down access to the risk-free savings technology and introduce the additional constraint

$$a_1 = 0. \tag{3}$$

CRRA Preferences. We first focus on CRRA preferences and set $\theta = \frac{1}{\gamma}$. In the notation, we retain parameters γ and θ to separately illustrate the role of risk aversion. Consequently, the utility function simplifies to

$$U = \begin{cases} \frac{c_0^{1-\frac{1}{\gamma}}}{1-\frac{1}{\gamma}} + \frac{\mathbb{E}\left[c_1^{1-\theta}\right]}{1-\frac{1}{\gamma}} & \text{for } \theta = \frac{1}{\gamma} \neq 1 \\ \ln(c_0) + \mathbb{E}[\ln(c_1)] & \text{for } \theta = \gamma = 1. \end{cases}$$

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

$$U \approx \frac{c_0^{1-\frac{1}{\gamma}}}{1-\frac{1}{\gamma}} + \left(\frac{1}{1-\frac{1}{\gamma}} - \frac{\theta}{2}\mu_2^c + \frac{\theta(1+\theta)}{6}\mu_3^c - \frac{\theta(1+\theta)(2+\theta)}{24}\mu_4^c \right), \tag{4}$$

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 constraint (3), the central moments[11] 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 (4). 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 dis-

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

tribution 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,[12] 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 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 5.2 imply a pronounced left-skewness and a strong excess kurtosis, which may lead to sizeable welfare losses, the overall effect depends on the utility weight of this risk, which depends on the calibration of θ.

Logs vs. Levels. While the transformation from logs to levels is natural, it has non-trivial implications for the welfare effects of higher-order risk: the higher-order moments of the shocks in *levels*, $\exp(\varepsilon)$, rather than of the shocks in *logs*, ε, are relevant for utility consequences. Consider a mean preserving (thus $\mathbb{E}[\exp(\epsilon)] = 1$) change of idiosyncratic risk. When introducing left-skewness in logs, probability mass is shifted to the left, which reduces the variance of the shocks in levels. Without adjustment, the mean of the distribution in levels would be lower, so the distribution needs to be shifted up, which increases the mean in logs. Similarly, a higher variance or higher kurtosis of the distribution in logs increases the variance in levels. Without adjustment, the mean of the distribution in levels would be higher, so it needs to be shifted down. In the special case of log utility ($\theta = \frac{1}{\gamma} = 1$), what matters for expected lifetime utility is the mean of the distribution in logs: $U = \ln(c_1) + \mathbb{E}[\ln(c_2)]$. This gives the following

Proposition 1. *Suppose that the utility function is logarithmic ($\theta = \frac{1}{\gamma} = 1$) and that there is no savings technology (binding constraint (3)). Then a mean-preserving* reduction of skewness *('more negative skewness') leads to* utility gains*, whereas a mean-preserving* increase of variance or kurtosis *lead to* utility losses *in expectation.*

[12]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 $u(c)$.

Proof. The formal proof is given in Appendix A.2. □

Proposition 1 thus establishes that higher-order income risk in terms of the logs of the income process, specifically a reduction of skewness (increase of left-skewness), may in fact lead to welfare gains rather than losses. While this may appear counter-intuitive at first glance, the reason is the transformation of the shocks from logs, which are typically modelled and estimated, to levels, which eventually matters for welfare.[13] In Appendix B we provide a numerical illustration by considering a discrete three-point distribution. We show how changing moment μ_i^ϵ by holding other moments constant can be conceptualized and how this affects the conclusions on the welfare implications of higher-order risk.

Epstein-Zin-Weil Preferences. We now consider the general case where $\gamma \neq \frac{1}{\theta}$. By the analogous steps to the CRRA case we can approximate the certainty equivalent (2) as

$$\begin{aligned} v(c_1, \theta, \Psi) &= \left(\int c_1(\varepsilon)^{1-\theta} d\Psi(\varepsilon) \right)^{\frac{1}{1-\theta}} \\ &\approx \left(1 + (1-\theta) \left(-\frac{\theta}{2}\mu_2^c + \frac{\theta(1+\theta)}{6}\mu_3^c - \frac{\theta(1+\theta)(2+\theta)}{24}\mu_4^c \right) \right)^{\frac{1}{1-\theta}} . \end{aligned} \tag{5}$$

Since $v(g(c_1, \theta, \Psi))$, for $g(c_1, \theta, \Psi) = \int c_1(\varepsilon)^{1-\theta} d\Psi(\varepsilon)$ is decreasing in $g(\cdot)$ for $\theta > 1$ and increasing in $g(\cdot)$ for $\theta < 1$ we observe that an increase of risk of order $2-4$ reduces the certainty equivalent and thus the previous results for the CRRA case readily extend.

3.2.2 Precautionary Savings

We now relax constraint (3) and instead assume that households have access to a risk-free savings technology.

CRRA Preferences. Using the budget constraint, we can write utility for $\theta = \frac{1}{\gamma} \neq 1$ as

$$U = \frac{(y_0 - a_1)^{1-\frac{1}{\gamma}}}{1-\frac{1}{\gamma}} + \frac{\mathbb{E}\left[(\exp(\varepsilon) + a_1)^{1-\theta} \right]}{1-\frac{1}{\gamma}}.$$

[13] Due to this re-transformation our findings are related to, but not the same, as first-order stochastic dominance, see Rothschild and Stiglitz (1970, 1971). Stochastic dominance refers to random variables in levels, in our case $\exp(\epsilon)$. Obviously, increasing the variance (or kurtosis) of $\exp(\epsilon)$, while holding the mean constant at $E[\exp(\epsilon)] = 1$, has direct negative utility consequences. In this case utility is $U = \ln(y_0) + \mathbb{E}[\ln(\exp(\epsilon))]$, which for the maintained normalization $E[\exp(\epsilon)] = 1$ we could approximate as

$$U \approx \ln(y_0) - \frac{1}{2}\mu_2^{\exp(\epsilon)} + \frac{1}{3}\mu_3^{\exp(\epsilon)} - \frac{1}{4}\mu_4^{\exp(\epsilon)}$$

from which the utility effects of increasing the variance or the kurtosis or decreasing the skewness are obviously all negative.

The Euler equation of the maximization problem is given by (cf. standard results from the literature on risk, e.g., Eeckhoudt and Schlesinger, 2008)

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

Notice that the LHS is increasing in a_1 and the RHS is decreasing in a_1 if $\mu_3^{\exp(\varepsilon)}$ is small enough relative to $\mu_2^{\exp(\varepsilon)}$ and $\mu_4^{\exp(\varepsilon)}$.[14] Now 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 temperance $(2+\theta)$. Finally, an increase of the fourth central moment increases the RHS, additionally scaled by the measure of *relative edginess* $(3+\theta)$.[15] Similar to what we saw in equation (4), 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 will for a given a_1 increase the RHS, which will be offset by an increase of savings a_1.[16] This result is very intuitive: ordinary and high-order income risk increases precautionary savings, through which households reduce the adverse utility consequences of this risk. The intensity of the behavioral reaction crucially depends on the risk attitudes governed by θ (cf. RHS), as well as on the inter-temporal elasticity of substitution γ (cf. LHS).

Epstein-Zin-Weil Preferences. In the general case where $\gamma \neq \frac{1}{\theta}$, we can use the resource constraint and write utility as

$$
U = \frac{(y_0 - a_1)^{1-\frac{1}{\gamma}}}{1-\frac{1}{\gamma}} + \frac{\left(\mathbb{E}\left[(\exp(\varepsilon) + a_1)^{1-\theta}\right]\right)^{\frac{1-\frac{1}{\gamma}}{1-\theta}}}{1-\frac{1}{\gamma}}.
$$

Observe that now the first-order condition is given by

$$
(y_0 - a_1)^{-\frac{1}{\gamma}} = v(c_1, \theta, \Psi)^{\theta - \frac{1}{\gamma}} \mathbb{E}\left[(\exp(\varepsilon) + a_1)^{-\theta}\right]. \tag{7}
$$

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

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

[16]Formally, it is straightforward to show this by taking the total differential of (6), cf. Appendix A.5.

In the sequel, we follow Kimball and Weil (2009) and assume that the marginal utility of saving, the RHS of (7), is a decreasing function of a_1 (just as earlier established for CRRA utility), which establishes uniqueness of the solution. With this assumption we obtain the next proposition, as in Kimball and Weil (2009):

Proposition 2. *For $\theta \neq \frac{1}{\gamma}$ an increase of (higher-order) risk leads to an increase of savings if $\gamma \leq 1$ or if $1 < \gamma \leq \frac{1}{\theta}$.*

Proof. Propositions 5 and 6 in Kimball and Weil (2009) and Appendix A.6. □

Thus, with a low IES ($\gamma \leq 1$), which since Hall (1988) most macroeconomists regard as a reasonable calibration, increasing risk leads to increasing savings. With a high IES ($\gamma > 1$), however, precautionary savings behavior *may not* arise if risk attitudes are also strong ($\gamma > \frac{1}{\theta}$). For a given degree of risk ($\mu_2^{\exp(\varepsilon)}, \mu_3^{\exp(\varepsilon)}, \mu_4^{\exp(\varepsilon)}$), the utility delivery from expected second period consumption as measured by the certainty equivalent is smaller, the stronger risk attitudes are. An increase of (higher-order) risk ($\mu_2^{\exp(\varepsilon)}, \mu_3^{\exp(\varepsilon)}, \mu_4^{\exp(\varepsilon)}$) implies a reduction of the certainty equivalent. This reduction is stronger if risk attitudes are stronger so that with a high IES the household may prefer to consume in the first period rather than to save for the second period and thus savings may decrease in response to the increase of risk.[17]

4 Canonical Income Process with Higher-Order Risk

4.1 The Income Process

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

$$y_{ijt} = f\left(\mathbf{X}_{ijt}, Y_t\right) + \tilde{y}_{ijt}, \tag{8}$$

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:

$$f\left(\mathbf{X}_{ijt}, Y_t\right) = \beta_{0t} + f_{age}\left(j\right) + \mathbb{1}_{e_{it}=c} f_{EP}(t) + \beta^{size} \log\left(hhsize_{ijt}\right) \tag{9}$$

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.

[17] Parts of this intuition is also discussed in Krueger and Ludwig (2019) for changes of second-order risk.

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 a idiosyncratic *fixed effect* χ_i. The idiosyncratic fixed effect is a shock drawn once upon entering the labor market from a distribution[18] which is the same for every cohort. The persistent component is modeled as an AR(1) process with innovation η_{ijt}:

$$\tilde{y}_{ijt} = \chi_i + z_{ijt} + \varepsilon_{ijt}, \text{ where } \varepsilon_{ijt} \underset{iid}{\sim} F_\varepsilon, \; \chi_i \underset{iid}{\sim} F_\chi \tag{10a}$$

$$z_{ijt} = \rho z_{ij-1t-1} + \eta_{ijt}, \text{ where } \eta_{ijt} \underset{id}{\sim} F_\eta\left(s\left(t\right)\right), \tag{10b}$$

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:[19]

$$\mu_k^x = E\left[\left(x - E\left[x\right]\right)^k\right] \text{ for } x \in \left\{\chi, \varepsilon, \eta(s)\right\}. \tag{11}$$

As in Storesletten et al. (2004), the economy can be in one of two aggregate states, indicated by $\mathbb{1}_{s(t)=E}$, which is 1 if year t is an expansion (denoted by E) and 0 if year t is a contraction (denoted by C). This gives central moment $k \in \{2,3,4\}$ as

$$\mu_k^\eta\left(s\left(t\right)\right) = \mathbb{1}_{s(t)=E}\mu_k^{\eta,E} + \left(1 - \mathbb{1}_{s(t)=E}\right)\mu_k^{\eta,C} \tag{12}$$

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 $\mu_k(z_{i0t}) = \mu_k^\eta(s(t))$.

4.2 GMM Approach to Estimation

We follow the common approach in the literature and estimate (8) and (10) in two steps. In the first step, we estimate (8), which yields residuals $\tilde{y}_{ijt}$. In the second step, we

[18] Thus, we are estimating a random effects model.

[19] 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.

estimate the parameters of the stochastic process (10) by fitting cross-sectional moments of the distribution of residual (log) income. As is standard, the variance terms of all components of (10) 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.[20] We denote central moments by $\mu_k(\cdot)$ and co-moments by $\mu_{kl}(\cdot)$, where

$$\mu_k(\tilde{y}_{ijt};\theta) = E\left[(\tilde{y}_{ijt} - E[\tilde{y}_{ijt}])^k \,|s^t\right] \tag{13a}$$

$$\mu_{kl}(\tilde{y}_{ijt}, \tilde{y}_{ij+1t+1};\theta) = E\left[(\tilde{y}_{ijt} - E[\tilde{y}_{ijt}])^k (\tilde{y}_{ij+1t+1} - E[\tilde{y}_{ij+1t+1}])^l \,|s^t\right]. \tag{13b}$$

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

$$\mu_2(\tilde{y}_{ijt};\theta) = \mu_2^{\chi} + \mu_2^{\varepsilon} + \mu_2(z_{ijt}) \tag{14a}$$

$$\mu_{11}(\tilde{y}_{ijt}, \tilde{y}_{ij+1t+1};\theta) = \mu_2^{\chi} + \rho\mu_2(z_{ijt}) \tag{14b}$$

$$\mu_3(\tilde{y}_{ijt};\theta) = \mu_3^{\chi} + \mu_3^{\varepsilon} + \mu_3(z_{ijt}) \tag{14c}$$

$$\mu_{21}(\tilde{y}_{ijt}, \tilde{y}_{ij+1t+1};\theta) = \mu_3^{\chi} + \rho\mu_3(z_{ijt}) \tag{14d}$$

$$\mu_4(\tilde{y}_{ijt};\theta) = \mu_4^{\chi} + \mu_4^{\varepsilon} + \mu_4(z_{ijt}) + 6\left(\mu_2^{\chi}\mu_2^{\varepsilon} + (\mu_2^{\chi} + \mu_2^{\varepsilon})\,\mu_2(z_{ijt})\right) \tag{14e}$$

$$\mu_{31}(\tilde{y}_{ijt}, \tilde{y}_{ij+1t+1};\theta) = \mu_4^{\chi} + \rho\mu_4(z_{ijt}) + 3\left(\mu_2^{\chi}\mu_2^{\varepsilon} + (\mu_2^{\chi} + \rho\,(\mu_2^{\chi} + \mu_2^{\varepsilon}))\,\mu_2(z_{ijt})\right), \tag{14f}$$

where the second to fourth central moments of z_{ijt} are given recursively by

$$\mu_2(z_{ijt}) = \rho^2\mu_2(z_{ij-1t-1}) + \mu_2^{\eta}(s(t)) \tag{15a}$$

$$\mu_3(z_{ijt}) = \rho^3\mu_3(z_{ij-1t-1}) + \mu_3^{\eta}(s(t)) \tag{15b}$$

$$\mu_4(z_{ijt}) = \rho^4\mu_4(z_{ij-1t-1}) + 6\rho^2\mu_2(z_{ij-1t-1})\mu_2^{\eta}(s(t)) + \mu_4^{\eta}(s(t)). \tag{15c}$$

[20]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.

Denote the empirical 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:

$$E\left[m_2\left(\tilde{y}_{ijt}\right) - \mu_2\left(\tilde{y}_{ijt};\theta\right)|s^t\right] = 0 \tag{16a}$$

$$E\left[m_{11}\left(\tilde{y}_{ijt}, \tilde{y}_{ij+1t+1}\right) - \mu_{11}\left(\tilde{y}_{ijt}, \tilde{y}_{ij+1t+1};\theta\right)|s^t\right] = 0 \tag{16b}$$

$$E\left[m_3\left(\tilde{y}_{ijt}\right) - \mu_3\left(\tilde{y}_{ijt};\theta\right)|s^t\right] = 0 \tag{16c}$$

$$E\left[m_{21}\left(\tilde{y}_{ijt}, \tilde{y}_{ij+1t+1}\right) - \mu_{21}\left(\tilde{y}_{ijt}, \tilde{y}_{ij+1t+1};\theta\right)|s^t\right] = 0 \tag{16d}$$

$$E\left[m_4\left(\tilde{y}_{ijt}\right) - \mu_4\left(\tilde{y}_{ijt};\theta\right)|s^t\right] = 0 \tag{16e}$$

$$E\left[m_{31}\left(\tilde{y}_{ijt}, \tilde{y}_{ij+1t+1}\right) - \mu_{31}\left(\tilde{y}_{ijt}, \tilde{y}_{ij+1t+1};\theta\right)|s^t\right] = 0. \tag{16f}$$

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.

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 (10b): the variance of the innovations accumulate as a cohort ages, as can be seen from the theoretical moment in equation (14a). 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 a similar accumulation holds for the other central moments, as seen in equations (14c) and (14e). Consider the third central moment. If for a given dispersion the probability of a large negative/positive income shock was higher/lower during a contractionary period, then the skewness of the shock in a contractionary period would be smaller (more negative) than in an expansion, i.e., $\mu_3^{\eta,C} < \mu_3^{\eta,E}$. 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 (14a), 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 (14c), which is identified via the age profile of the third central moment. Considering the sum in (14a), 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.[21]

Now consider the expressions for variance and covariance in equations (14a) and (14b). 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 (14c) and (14d), 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 (14e) and (14f) identify the fourth central moments μ_4^x for $x \in \{\chi, \varepsilon, \eta(s)\}$ in the same fashion as for the second and third central moments.

Instead of estimating all parameters simultaneously, we use moment conditions (16a) and (16b) to estimate the variance parameters and the persistence ρ. Given an estimate for ρ, we then use moment conditions (16c) and (16d) to estimate the third central moments. Likewise, given estimates for ρ and the variance parameters, we use moment conditions (16e) and (16f) to estimate the fourth central moments.

4.3 Discretization of the Estimated Process

For the quantitative evaluation in Section 6 we fit a discrete income process that features the estimated distributional characteristics of the shocks ε and $\eta(s)$. Given that this approach extends beyond our specific analysis in Section 6, we discuss the general procedure here, with details provided in Online Appendix A.

We proceed in two steps. First, for each shock we fit a parametric distribution function, the Flexible Generalized Lambda Distribution (FGLD) developed by Freimer et al. (1988), to match the estimated moments of the shock distribution. The FGLD is characterized by its quantile function $Q(p;\lambda)$, where λ is a vector of four parameters. For each shock $x \in \{\varepsilon, \eta(s)\}$, we follow Lakhany and Mausser (2000) and Su (2007) and fit these parameters such that the fitted FGLD matches the estimated first four central moments $\{\mu_i^x\}_{i=1}^4$ of the distributions F_ε and $F_\eta(s)$. The quantile function for location parameter λ_1, scale parameter λ_2 tail index parameters λ_3, λ_4[22] is given by

$$Q(p;\lambda) = F^{-1}(p;\lambda) = x = \lambda_1 + \frac{1}{\lambda_2}\left(\frac{p^{\lambda_3}-1}{\lambda_3} - \frac{(1-p)^{\lambda_4}-1}{\lambda_4}\right) \tag{17}$$

We choose λ_3, λ_4 jointly to fit the third and fourth central moments by solving

$$\min_{\lambda_3,\lambda_4} \sum_{i=3}^{4} \left(\mu_i(\lambda_3,\lambda_4) - \hat{\mu}_i\right)^2 \text{ s.t. } \min\{\lambda_3,\lambda_4\} > -\frac{1}{4}$$

[21]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.

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

where $\hat{\mu}_i$ is the point estimate of the i^{th} moment, and $\mu_i(\cdot)$ denotes the central moment of the FGLD. Next, we determine λ_2 to match the variance and λ_1 to match the mean, both in closed form. Second, we approximate the shocks 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 respective FGLD.

5 Estimation of the Income Process

5.1 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.[23] We estimate the extended canonical income process at the household level for both pre- and post-government household income. 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-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, which also smoothes the age profiles of these moments.

Defining Business Cycles. In order to implement the estimator we need to classify years as contractions or expansions. We initiate our definition on NBER peaks and trough data. Given the sluggish synchronization of labor market outcomes with the macroeconomic indicators that the NBER takes into account, we expand the dating based

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

on mean earnings of males in the the PSID. The relevant time period is 1942–2012. Given 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. All years that are not classified as contraction are classified as expansions. This gives the following years as contractions: 1945, 1949, 1953, 1957, 1960, 1970, 1974, 1980–83, 1990–91, 2001–02, 2008–10, and 2012.

5.2 Estimation Results: Cyclical Idiosyncratic Income Risk

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. 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 add a linear time trend to the third central moment of transitory shocks in order to accomodate for a low-frequency change of the cross-sectional distribution. We report here 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 500 bootstrap samples. Table 1 shows the estimates, and Figure 1 shows the fit over age and time of the estimated process for post government income, which we use in the quantitative analysis in Section 7.

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 the several 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 income process for post-government income 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-government and post-government income the estimated processes imply a countercyclical variance of persistent shocks: in aggregate downturns,

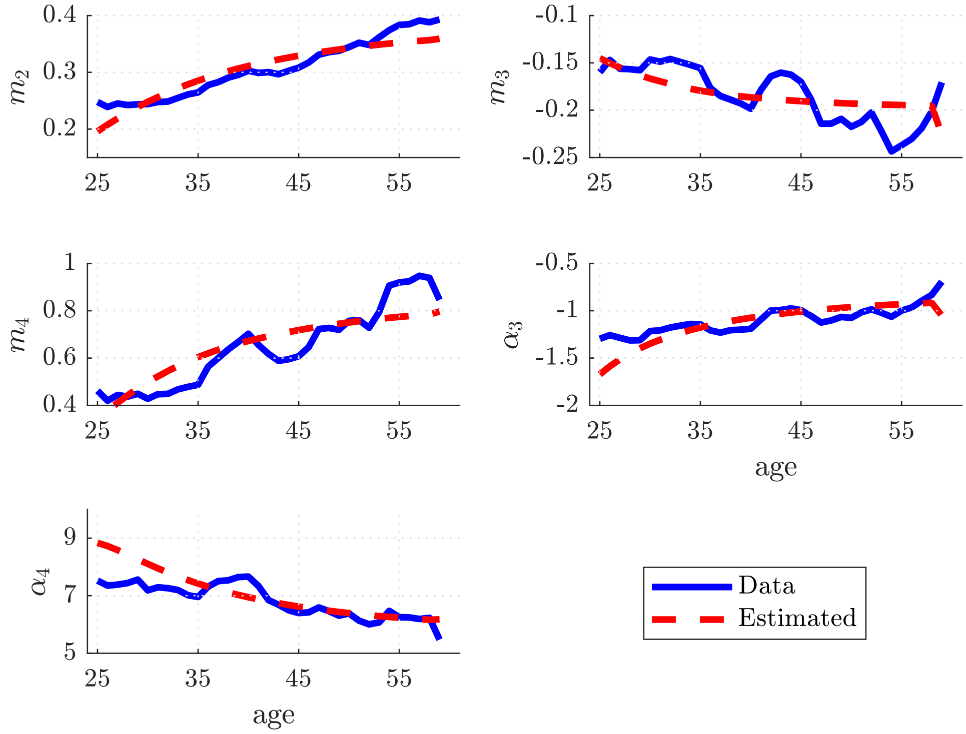
Table 1: Estimation Results for Pre- and Post Government Income

	Estimated Central Moments		Implied Standardized Moments	
	HH Pre	HH Post	HH Pre	HH Post
ρ	0.9602	0.9684		
	[0.9464; 0.9782]	[0.9632; 1.0000]		
μ_2^{χ}	0.1592	0.1077		
	[0.1405; 0.1821]	[0.0912; 0.1250]		
μ_2^{ε}	0.1043	0.0752		
	[0.0972; 0.1146]	[0.0716; 0.0856]		
$\mu_2^{\eta,C}$	0.0375	0.0223		
	[0.0273; 0.0477]	[0.0049; 0.0234]		
$\mu_2^{\eta,E}$	0.0152	0.0085		
	[0.0087; 0.0199]	[0.0046; 0.0120]		
μ_3^{χ}	−0.1130	−0.0522	−1.78	−1.48
	[−0.1293; −0.0472]	[−0.0546; −0.0019]		
μ_3^{ε}	−0.1516	−0.0866	−4.50	−4.20
	[−0.3055; −0.1891]	[−0.2206; −0.1433]		
$\mu_3^{\eta,C}$	−0.0331	−0.0164	−4.56	−4.93
	[−0.0622; −0.0302]	[−0.0314; −0.0101]		
$\mu_3^{\eta,E}$	−0.0046	−0.0012	−2.47	−1.51
	[−0.0055; 0.0102]	[−0.0024; 0.0054]		
μ_4^{χ}	0.0620	0.0180	2.45	1.55
	[0.0000; 0.0469]	[0; 0.0172]		
μ_4^{ε}	0.4244	0.2297	38.98	40.65
	[0.3724; 0.4992]	[0.2024; 0.2767]		
$\mu_4^{\eta,C}$	0.1357	0.0665	96.55	134.11
	[0.0614; 0.1447]	[0.0153; 0.0632]		
$\mu_4^{\eta,C*}$	0.0222	0.0097	96.55	134.11

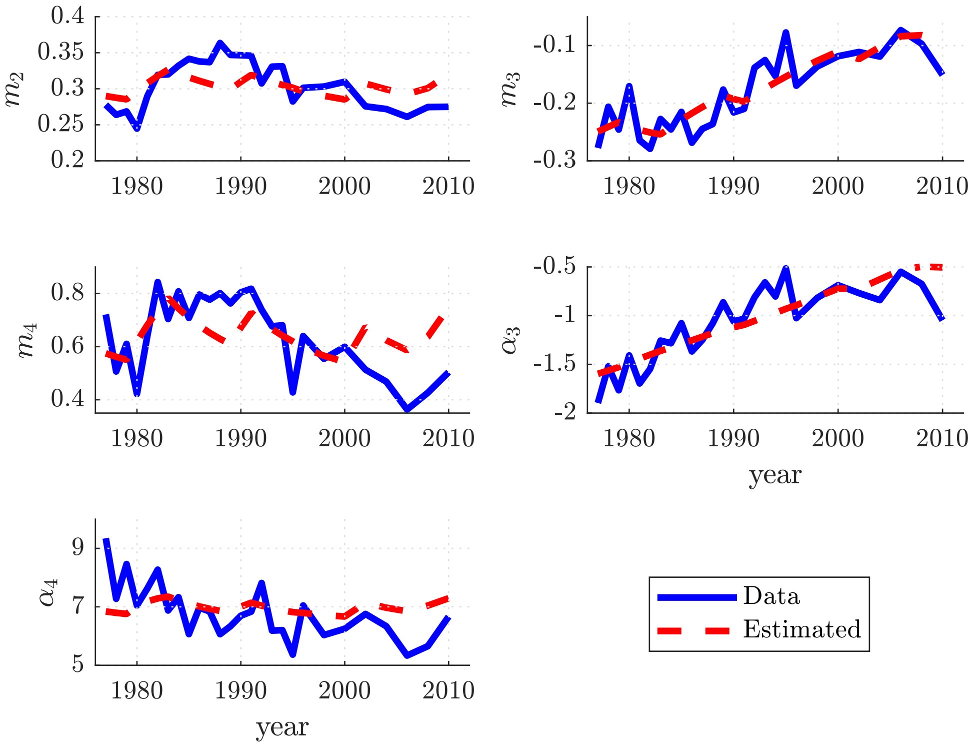
Notes: Table shows estimated central moments for household earnings (HH Pre) and household income after taxes and transfers (HH Post). Brackets show 5^{th} and 95^{th} percentiles of 500 bootstrap estimates. $^*\mu_4^{\eta,C}$ not separately estimated.

Figure 1: Fit of Estimated Process for Post-Government Earnings

(a) Age profiles of moments of the cross-sectional distribution



(b) Year profiles of moments of the cross-sectional distribution



Notes: Shows empirical profiles of the cross-sectional distribution moments ("Data") and the counterpart of these moments implied by the estimated income process ("Estimated"). The displayed age profiles for each age show the average moment over years; the year profiles for each year show the average moment over age.

the cross-sectional distribution of shocks is more dispersed. The countercyclicality we estimate for post-government income is quantitatively similar to the one estimated by Storesletten et al. (2004): the estimated standard deviation of persistent shocks is 62% 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., $\alpha_3^\eta(s(t) = C) = \alpha_3^\eta(s(t) = E)$. 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 2.5 for pre-government income, and 1.55 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 distribution. Note that while these estimates of kurtosis seem extreme at first glance, they imply a good fit of the cross-sectional distribution over age and over years as shown in Figure 1. Furthermore, the estimated income process is in line with the average kurtosis of income changes.

6 A Quantitative Model

6.1 The Economy

We now set up a quantitative version of the simple two-period model of Section 3 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.

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. Thus, 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

$$y(z, \varepsilon, j; s) = e_j \cdot \exp(z(s) + \varepsilon), \tag{18}$$

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

$$z'(s') = \begin{cases} \rho z + \eta', & \text{where } \eta' \underset{iid}{\sim} \tilde{F}_\eta(s') \text{ for } j < j_r \\ z & \text{for } j \geq j_r, \end{cases} \tag{19}$$

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 the initial draws are $\exp(\varepsilon_0) = \exp(z_0) = 1$. In retirement, $j \in \{j_r, \ldots, J\}$, households earn a fixed earnings related pension income $y_j = b(z_j)$. Thus, pension payments are contingent on the last income state before retirement.[24] 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

$$a'(z, \varepsilon, j; s) = a(1 + r) + y(z, \varepsilon, j; s) - c \geq 0. \tag{20}$$

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

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 economy s. The recursive problem of households is

$$V_j(a, z, \varepsilon; s) = \max_{c,a'} \begin{cases} \left((1-\tilde{\beta})c^{1-\frac{1}{\gamma}} + \tilde{\beta}\left(v(V_{j+1}(a', z', \varepsilon'; s'))\right)^{1-\frac{1}{\gamma}}\right)^{\frac{1}{1-\frac{1}{\gamma}}} & \gamma \neq 1 \\ \exp\left\{(1-\tilde{\beta})\ln c + \tilde{\beta}\ln\left(v(V_{j+1}(a', z', \varepsilon'; s'))\right)\right\} & \text{otherwise} \end{cases}$$
$$\text{s.t. } (18), (19), \text{and } (20),$$

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 utility from age j consumption c_j and the certainty equivalent from the continuation utility $v(V_{j+1}(\cdot))$. Given γ, parameter θ pins down the relative risk attitudes of households as discussed in Section 3. 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 the household policy and value functions using the method of endogenous gridpoints. We aggregate using explicit aggregation characterizing and 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 of the distribution $G_j(a_j, z_j, \varepsilon_j; s)$. The latter is induced by the exogenous laws of motion for z, s, the exogenous distribution of ε, and the endogenous transitions $a'_j(a_j, z_j, \varepsilon_j; s)$.

6.2 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. This gives the aggregate transition matrix for $s \in \{C, E\}$:

$$\pi(s' \mid s) = \begin{bmatrix} 0.388 & 1-0.388 \\ 1-0.769 & 0.769 \end{bmatrix}$$

with associated stationary invariant distribution $\Pi_s = [0.274, 0.726]'$. Thus the persistence of staying in a contraction and the corresponding unconditional probability of contractions is much lower than for expansions.

Age Bins and Age Productivity. Each model period corresponds to one life year. Consistent with our empirical specification, households are assumed to 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 C.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 in the context of our analysis is the specification of the distribution functions of the idosyncratic shocks. We use the Flexible Generalized Lambda Distribution (FGLD) to form a discrete approximation of the estimated transitory and persistent shocks as discussed in Section 4.3. We consider three alternative parameterizations of the FGLD with the following restrictions on central moments (and parameters) to which we refer to as *distribution scenarios*:[25]

1. NORM: FGLD with moments of Normal distribution; $\hat{\mu}_3 = 0$, $\frac{\hat{\mu}_4}{\hat{\mu}_2^2} = 3$, $\lambda_3 = \lambda_4$
2. LK: FGLD with excess kurtosis; $\hat{\mu}_3 = 0$, $\lambda_3 = \lambda_4$
3. LKSW: FGLD with excess kurtosis and left-skewness; no restrictions apply.

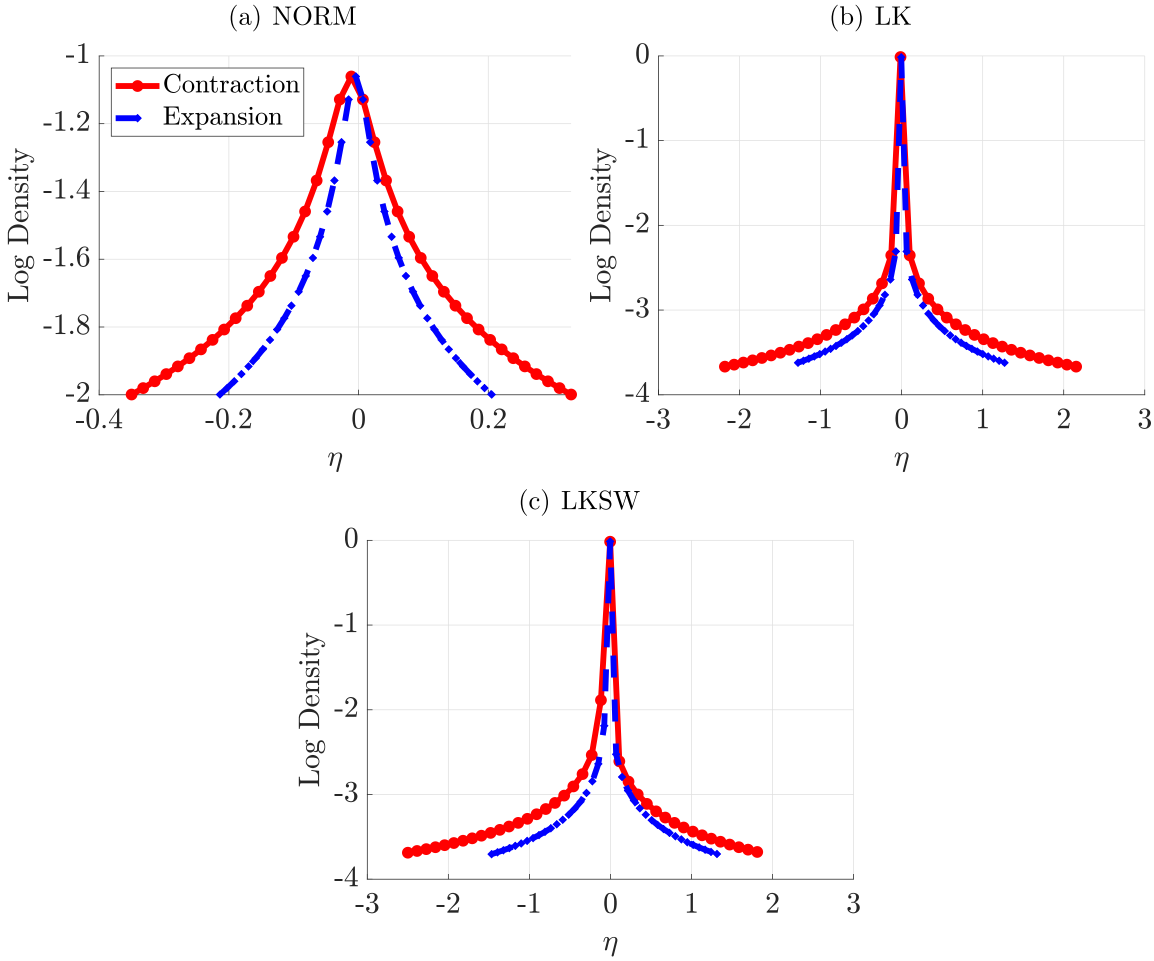
Thus, scenario NORM features symmetric shock distributions with the estimated variance and a kurtosis of 3,[26] scenario LK features leptokurtic shock distributions with the estimated second and fourth moments, and scenario LKSW adds the estimated negative third moments to LK. Figure 2 shows the log distribution functions for the distributions of the persistent shock $\eta(s)$. Panel (a) shows the distribution in scenario NORM in contractions and expansions, illustrating the counter-cyclical variance. Panels (b) and (c) show

[25] 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, say, leptokurtic 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.

[26] One apparent drawback of the FGLD is that it does not nest the normal distribution. In our computations we also consider a scenario in which we draw shocks from the normal distribution and discretize it using standard Gaussian Quadrature methods. Results are numerically almost identical to those obtained for FGLD distribution NORM. This is documented in Appendix D.1.

the distributions in scenarios LK and LKSW, respectively. Relative to scenario NORM, the distributions in scenarios LK and LKSW have more mass in the center and are more spread out in the tails. The comparison between the distributions in scenario LKSW in Panel (c) and scenario LK in Panel (b) further illustrates the effects of left-skewness and the increasing left-skewness in contractions. Appendix C.1 reports the estimated, fitted, and discretized moments, as well as the parameter vectors λ for all shocks under the three scenarios NORM, LK, LKSW. In all three distribution scenarios we scale down the transitory shocks because part of the estimated variance is likely due to measurement error.[27]

Figure 2: Discretized Log Distribution Functions: Persistent Shock



Notes: Discretized log distribution functions for the persistent shock η. NORM: FGLD with moments of the normal distribution, LK: FGLD with excess kurtosis, LKSW: FGLD with additional left-skewness (in logs). Markers denote the grid points used in the discretized distribution. Log density is the base 10 logarithm of the PDF.

[27]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.

In Section 3 we emphasize that it is crucially important to see how earnings processes estimated in logs translate into levels of the earnings distribution. Appendix C.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 C.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 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 is normalized to one, since the means of the shocks z_j, ϵ_j are equal to one, and since the total population in age group j is also 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 C.5 in Appendix C.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).

6.2.1 Preferences

As we show in Section 3, 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 $\gamma = 1$.[28] For each $\theta \in \{1,2,3,4\}$, we determine endogenously the discount factor β to match life-cycle asset profiles scaled by 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 across distribution scenarios, i.e., in the scenarios NORM and LK, 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.

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 C.5
Average tax rate	$\tau[\%] = 16.8\%$
Aggregate shocks	$\pi(s' = c \mid s = c) = 0.38, \pi(s' = e \mid s = e) = 0.77$
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.

[28] 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.

7 The Quantitative Role of Higher-Order Income Risk

7.1 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 $\varepsilon = 0$, weighted over aggregate states s, where the Pareto weight is the stationary invariant distribution Π_s:

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

We then quantify the welfare gain of being born into the world with higher-order risk by calculating 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 (distribution scenarios LK and LKSW, respectively). Given the homotheticity of the utility function, the CEV is

$$g_c^i = \frac{W^i}{W^{NORM}} - 1.$$

We distinguish between three different channels through which idiosyncratic risk translates into utility consequences evaluated from this ex-ante perspective. 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 across age and the cross-section. 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}. We accordingly decompose the total CEV of moving from scenario NORM to scenario $i \in \{LK, LKSW\}$ as (cf. Appendix A.7 for explicit expressions)

$$g_c^i = g_c^{i,mean} + g_c^{i,lcd} + g_c^{i,csd}.$$

Table 3 summarizes the welfare implications of higher-order income risk by showing the CEV and its decomposition. Scenario LK leads to welfare losses because of the high variance and kurtosis of the earnings distribution in levels, cf. Table C.6, whereas scenario LKSW leads to welfare gains when risk attitudes are weak. This is consistent with our analytical findings in Proposition 1. With stronger risk attitudes, however, welfare losses show up for scenario LKSW, because the increasing variance and the high kurtosis dominate the welfare effects.

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

CEV	g_c	g_c^{mean}	g_c^{lcd}	g_c^{csd}
Risk Aversion, $\theta = 1$				
LK	-1.084	0.400	-1.474	-0.010
LKSW	0.371	-0.154	0.506	0.019
Risk Aversion, $\theta = 2$				
LK	-1.595	0.450	-2.035	-0.010
LKSW	-0.386	-0.161	-0.256	0.031
Risk Aversion, $\theta = 3$				
LK	-3.919	0.682	-4.557	-0.044
LKSW	-4.488	0.318	-4.751	-0.055
Risk Aversion, $\theta = 4$				
LK	-9.399	1.247	-10.472	-0.174
LKSW	-12.474	1.211	-13.392	-0.294

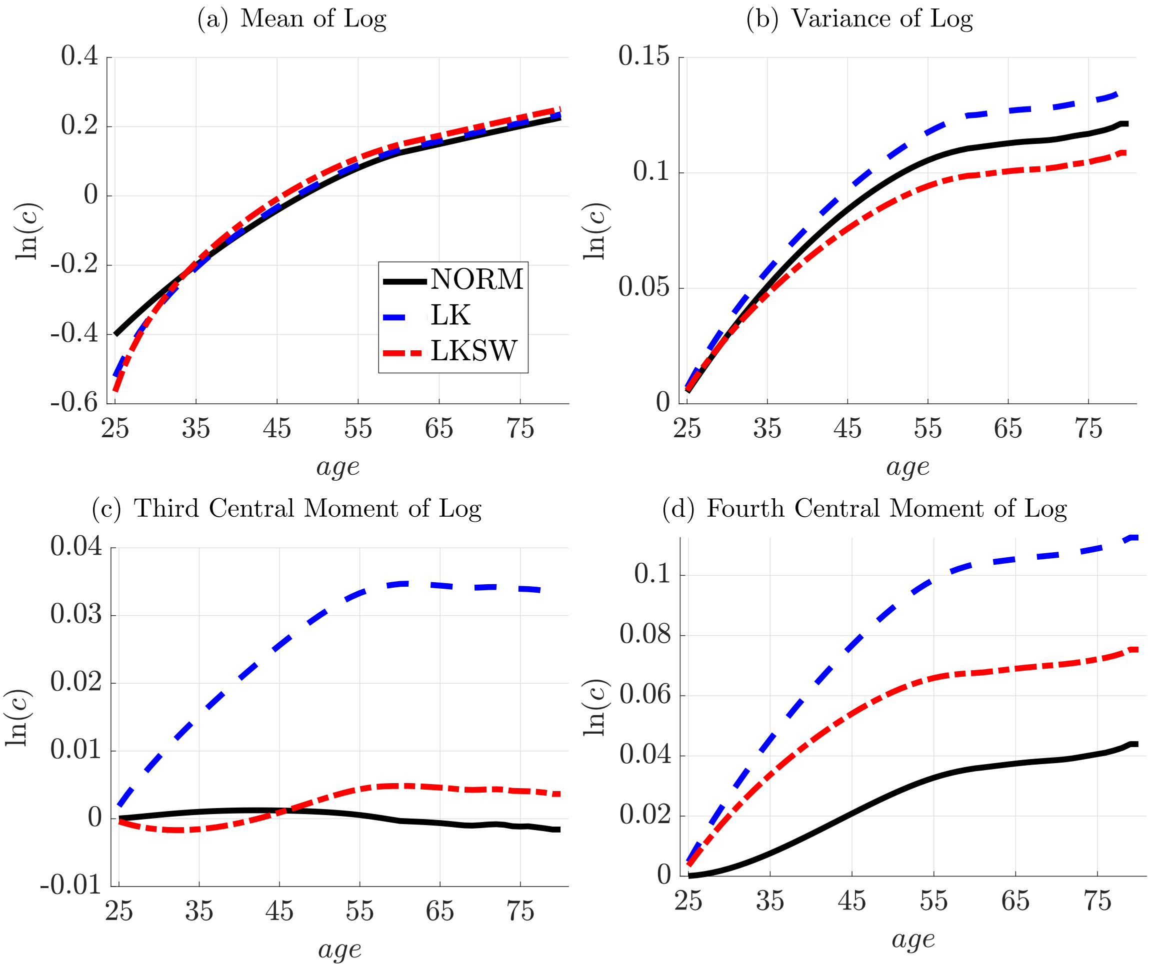
Notes: Welfare gains (positive numbers) and losses (negative numbers) of higher-order income risk, expressed as a Consumption Equivalent Variation (CEV) in scenario NORM that makes households indifferent to scenarios LK and LKSW, respectively. g_c: total CEV, g_c^{mean}: CEV from changes of mean consumption,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

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 3, which displays mean log consumption over the life-cycle.[29] Consumption in scenarios LK and 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 (exceptions are results for LKSW 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 3 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

[29] 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, they are the same in the other risk aversion calibrations. Note that consumption is monotonically increasing over the life-cycle and thus does not display the typical hump-shaped profile. One reason is that our model with no mortality risk is not a good model for life-cycle consumption behavior in retirement.

Figure 3: Central Moments of Log Consumption by Age: Strong Risk Attitudes, $\theta = 4$



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

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.[30] The lower 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.

7.2 Welfare Costs of Cyclical Idiosyncratic Risk

Next, we quantify the utility consequences of *cyclical idiosyncratic risk*. To this end, for each scenario we evaluate the welfare implications for households of facing the *actual*

[30] The Gini coefficient for assets features the same ranking as the variance of log consumption. E.g., for a risk aversion of 4 it is at 0.35 in scenario NORM, at 0.37 in scenario LK and at 0.34 in scenario LKSW. Throughout it is substantially lower than in the data. One reason is that we ignore ex-ante heterogeneity in productivity (fixed productivity types), as well as heterogeneity in initial asset holdings at labor market entry.

cyclical income process relative to a *counterfactual* income process in which we shut down the cyclical variation of the distribution. We then calculate the CEV necessary in the non-cyclical scenario to make households indifferent to the cyclical scenario. 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.[31]

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 next compute for all scenarios $i \in \{NORM, LK, LKSW\}$:

$$g_c^{i,cr} = \frac{W^i}{W^{i,ncr}} - 1.$$

As with the role of higher order income risk per se, we further decompose the total CEV from cyclical risk into its components, i.e., we compute for scenario $i \in \{NORM, LK, LKSW\}$

$$g_c^{i,cr} = g_c^{i,cr,mean} + g_c^{i,cr,lcd} + g_c^{i,cr,csd}.$$

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 a weighted average of shock distributions for expansions and contractions.[32] There are two reasons for this. First and more importantly, it is conceptually not clear what characterizes an "average" distribution, once other moments than the variance are taken into account. Second, we avoid any potential inaccuracies that would arise from our discretization methods. 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.[33] However, we are mainly interested in the difference of welfare costs of cyclical income risk across scenarios, i.e., the "difference in difference" comparison between $g_c^{i,cr}$ and $g_c^{NORM,cr}$, i.e., $\Delta g_c^{i,cr} = g_c^{i,cr} - g_c^{NORM,cr}$ for $i \in \{LK, LKSW\}$. Thus, our approach to "normalize" the economy without cyclical idiosyncratic risk is of second order importance as long as it is consistent across scenarios.

[31]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 to focus on the role of the idiosyncratic shock distribution.

[32]When using log-Normal distributions of shocks, a typical 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).

[33]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 below, we also consider CRRA preferences with $\theta = 2$. In this case we obtain welfare costs of about 2.6%. 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.

Table 4 reports the results on the welfare costs of cyclical idiosyncratic risk in scenarios NORM, LK, and LKSW. First, note that consistent with our theoretical analysis of Section 3 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 previously documented result (and with our theoretical analysis of Section 3) that with logarithmic utility the total welfare effect from higher-order income risk is negative for scenario LK and positive for scenario LKSW, we now correspondingly find that welfare losses from cyclical idiosyncratic risk are about 0.25%p higher in scenario LK than in scenario NORM and 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.

We can thus 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).

7.3 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:

$$1 - \phi_j^x(s) = \frac{cov(\Delta \ln (c_{j+1}(s' \mid s)), x_{j+1}(s'))}{var(x_{j+1}(s'))}, \tag{21}$$

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
LK	-1.966	0.579	-2.506	-0.039	-0.246
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
LK	-3.534	0.912	-4.337	-0.109	-0.271
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
LK	-5.968	1.288	-7.06	-0.196	-1.361
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
LK	-9.738	1.731	-11.146	-0.322	-3.980
LKSW	-12.171	1.944	-13.686	-0.429	-6.413

Notes: Welfare gains (positive numbers) and losses (negative numbers) of cyclical idiosyncratic risk expressed as Consumption Equivalent Variation (CEV) in the non-cyclical scenario that makes households indifferent to the cyclical scenario. Displayed for scenarios NORM, LK, and 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 $g_c = g_c^{mean} + g_c^{lcd} + g_c^{csd}$. $\Delta g_c = g_c^i - g_c^{NORM}$, for $i \in \{LK, LKSW\}$: difference in percentage points relative to scenario NORM.

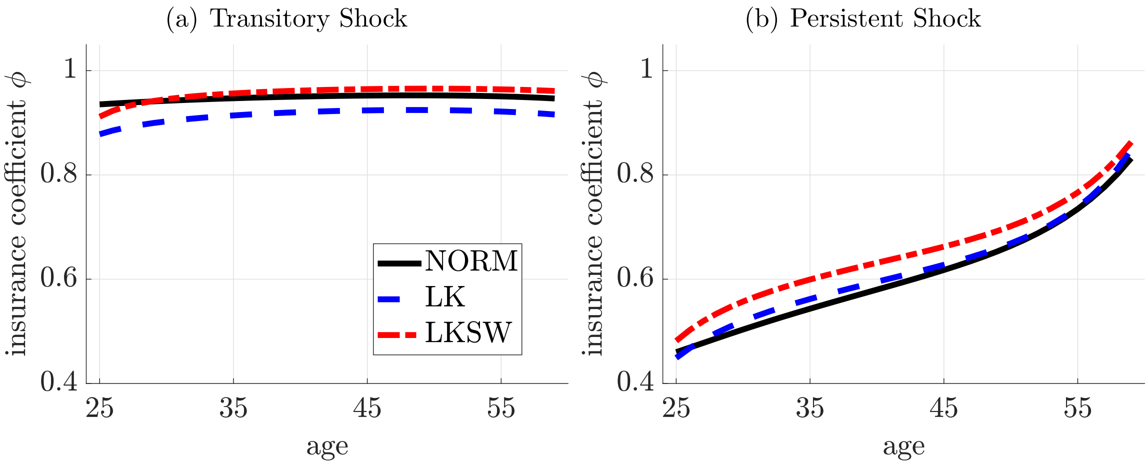
for $\Delta \ln (c_j(s' \mid s)) = \ln (c_{j+1}(s' \mid s)) - \ln (c_j(s))$.

Figure 4 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[34] 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. Insurance in scenario LK is worse (coefficients for transitory shocks shift down by about $4\%p$ compared to distribution NORM): the shocks (in levels) are more dispersed and insurance through increased precautionary savings is not strong enough to offset the direct consequences of higher risk.

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

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

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



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

of the pass-through of shocks to consumption changes in equation (21). Consider the aggregate (integrating over age and averaging over states s) pass-through coefficient for shock $x \in \eta, \varepsilon$:

$$
\begin{aligned}
1 - \phi^x =& \frac{E\left[\Delta \ln(c(\cdot))x\right] - E\left[\Delta \ln(c(\cdot))\right] E\left[x\right]}{var(x)} \\
=& \frac{E\left[\Delta \ln(c(\cdot))x | \, x > 0\right]}{var(x)} + \frac{E\left[\Delta \ln(c(\cdot))x | \, x < 0\right]}{var(x)} - \frac{E\left[\Delta \ln(c(\cdot))\right] E\left[x\right]}{var(x)}.
\end{aligned} \tag{22}
$$

The first two components of the sum in equation (22) 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 4, 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.

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

Transitory Shock	$1-\phi^{\varepsilon}$	$E[\Delta^c \cdot \epsilon, \epsilon < 0]$	$E[\Delta^c \cdot \epsilon, \epsilon > 0]$	$-E[\Delta^c] \cdot E[\epsilon]$
NORM	0.055	-0.034	0.898	0.136
LKSW	0.047	0.302	0.525	0.173
Persistent Shock	$1-\phi^{\eta}$	$E[\Delta^c \cdot \eta, \eta < 0]$	$E[\Delta^c \cdot \eta, \eta > 0]$	$-E[\Delta^c] \cdot E[\eta]$
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 22. Values are expressed as shares of total pass-through. $\Delta^c = \Delta \ln(c(\cdot))$.

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 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.

7.4 Sensitivity Analysis

7.4.1 General Equilibrium

In the analyses presented so far we consider a partial equilibrium framework where interest rates and average wages are constant. The increased precautionary savings from higher-order risk documented above may lead to a higher capital stock which in a general equilibrium would increase wages and lower returns on savings. To investigate the robustness of our findings with respect to this feedback, we consider a general equilibrium variant of our model, where we treat scenario LKSW with cyclical risk as a baseline for each level of risk aversion when (re)calibrating the model in general equilibrium.

In this baseline net wages are normalized to one and the interest rate is calibrated to $r[\%] = 4.2\%$. As a first step, we make this choice consistent with a standard static representative firm problem in general equilibrium and accordingly compute the implied parameters of the aggregate production function. As a second step, we hold constant these parameters and compute the equilibrium interest rate and wage rate for each considered scenario. A detailed description of this procedure is provided in Appendix C.4. The equilibrium wage and return rates vary little across the different scenarios. The reason for the modest differences lies in the life cycle structure of the economy. Consider moving from scenario NORM with cyclical risk to scenario LKSW with cyclical risk. While young agents have higher precautionary savings when facing higher-order risk, these savings will be dis-saved at old age. The aggregate savings of the economy will thus not change strongly.

For brevity of the exposition of the sensitivity analysis, we focus on the welfare costs of cyclical idiosyncratic risk, and on the comparison of scenario NORM to scenario LKSW. Column 2 of Table 6 shows the welfare costs of cyclical risk next to the baseline results, which are repeated in column 1. The percentage point difference of the CEV between scenario LKSW and scenario NORM is almost identical in the general equilibrium version of the model. For instance, for $\theta = 4$ the difference now stands at $-6.00\%p$, compared to $-6.41\%p$. We therefore conclude that our main findings are robust in general equilibrium.

A detailed summary of the results in a format corresponding to Table 4 is contained in Appendix D.2. Again, the main negative welfare effect comes from the life cycle distribution effect g_c^{lcd}. However, the effect is weaker, because the higher wage rate and the lower interest rate in the cyclical economy mute the precautionary savings response. This also implies a weaker (positive) mean effect in general equilibrium: the mean consumption difference between the cyclical and non-cyclical worlds is smaller compared to the partial equilibrium benchmark.

Table 6: Total CEV g_c of Cyclical Idiosyncratic Risk: Sensitivity Analyses

	Baseline	GE	CRRA	BC	IR
Risk Aversion, $\theta = 1$					
NORM	-1.720	-1.018	-1.720	-1.893	-1.905
LKSW	-1.443	-0.884	-1.443	-1.612	-1.611
Risk Aversion, $\theta = 2$					
NORM	-3.263	-2.000	-2.552	-3.609	-3.627
LKSW	-3.516	-2.367	-2.564	-4.293	-3.972
Risk Aversion, $\theta = 3$					
NORM	-4.607	-2.872	-3.335	-5.113	-5.123
LKSW	-7.177	-5.282	-4.456	-9.725	-8.253
Risk Aversion, $\theta = 4$					
NORM	-5.758	-3.611	-4.072	-6.404	-6.399
LKSW	-12.171	-9.619	-7.530	-17.14	-14.283

Notes: Total welfare gains (positive numbers) and losses (negative numbers) of cyclical idiosyncratic risk expressed as Consumption Equivalent Variation (CEV) g_c in the distribution scenario NORM and the leptokurtic and left-skewed scenario LKSW. CRRA: CRRA utility, BC: “borrowing constraints”, IR: interest rate, GE: general equilibrium.

7.4.2 Other Sensitivity Analyses

We last consider the sensitivity of our results with respect to selected modeling and calibration assumptions. Specifically, we consider an 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. Again, Table 6 summarizes the results and further details are contained in Appendix D.2.

CRRA Utility. Assuming CRRA preferences with $\theta = \frac{1}{\gamma}$ we conduct experiments for $\theta \in \{2, 3, 4\}$, since for $\theta = 1$ results are of course as before. As in our previous baseline analysis, we recalibrate discount factor β for each value of θ. For $\theta \in \{2, 3, 4\}$ we obtain $\beta \in \{0.982, 0.990, 0.995\}$ and thus, in contrast to our experiments with EZW utility, the calibrated discount factor is increasing in θ. With increasing risk attitudes θ the precautionary savings motive is strengthened, while the simultaneous reduction of the IES $\gamma = \frac{1}{\theta}$ reduces life-cycle savings. The second effect turns out to dominate so that calibration calls for less impatience in order to hold the average asset accumulation unchanged.

Column 3 of Table 6 summarizes the results on the welfare effects of cyclical idiosyncratic risk for this alternative choice of preferences. In comparison to Table 4 we observe a lower increase of welfare losses from cyclical idiosyncratic risk when risk aversion is increased (the IES is decreased). Likewise, our difference in difference comparison to scenario NORM shows that higher-order income risk still substantially matters for the welfare costs of cyclical idiosyncratic risk, but less than with EZW preferences. The reason is that with a lower IES the overall consumption profile is smoother and thus reacts less to changes in risk. Thus, the simultaneous reduction of the IES when relative risk attitudes are strengthened confounds the welfare analysis.

The Role of Borrowing Constraints. In our baseline calibration households start their economic life with positive assets and calibrated impatience is relatively strong. As a consequence, very few households are borrowing constrained (numerically, the fraction is basically zero in all scenarios). We now investigate the sensitivity of our results with regard to the role of the borrowing constraint by setting initial assets to 0. In this experiment, we do not recalibrate because we aim at disentangling the role of the constraint.

As a consequence of zero initial assets, the fraction of borrowing constrained hand-to-mouth consumers increases strongly. For $\theta = 1$, roughly 6.6% of all households are constrained in scenario NORM and 4.0% in scenario LKSW. Column 4 of Table 6 shows that this leads to higher overall welfare losses from cyclical idiosyncratic risk and an increasing importance for higher-order risk. For $\theta = 4$ the difference in the CEV between scenarios LKSW and NORM is about $-10.7\%p$, compared to $-6.4\%p$ reported in Table 4.

Thus, borrowing constraints increase the role played by higher-order income risk for the welfare losses from cyclical idiosyncratic risk.

Lower Interest Rate. Next, rather than assuming an annual interest rate of 4.2% we reduce it to 2%. We recalibrate the discount factor β in all four experiments for $\theta \in \{1, 2, 3, 4\}$, which gives $\beta \in \{0.990, 0.988, 0.986, 0.983\}$ and thus the discount factors are higher because lower returns reduce life-cycle savings which is offset in calibration by stronger patience. While the role played by higher-order income risk for the welfare losses from cyclical idiosyncratic risk is slightly increased, the difference to the baseline calibration is modest.

8 Conclusion

In this paper 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 the salient 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 the shocks can be estimated. We apply our method to household-level earnings income from the Panel Study of Income Dynamics. Our estimates 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. We then estimate the process for post-government household income. The estimates imply that both transitory and persistent income shocks are dampened and cyclicality is reduced by the existing tax and transfer system.

In the second part of the paper we show that the identified deviation from log-Normal shocks, i.e., *higher-order risk*, has important macroeconomic implications. We set up a standard quantitative life-cycle model in which households face an exogenous income process which features transitory and persistent shocks. Households can self-insure by means of saving in a risk-free asset. In the calibration of the income process, we use a parametric distribution function (the Flexible Generalized Lambda distribution) to implement shocks with higher-order risk, which we fit to the estimates of the central moments. We then discretize the obtained shock distributions.

We find that, first, higher-order risk has relevant implications for welfare. 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, because households increase their precautionary savings, and thus the pass-through of income shocks to consumption is reduced. However, a decomposition of the pass-through coefficient reveals that this does not imply better insurance in the presence of higher-order risk: increased savings do not suffice to insure against increased downside risk, and therefore the pass-through of negative shocks is actually stronger than under Normal shocks with the same dispersion. We therefore caution against using only the insurance coefficient for the analysis of the degree of partial insurance against income risk and view it as an interesting avenue for future research to dig deeper into this finding by combining our analysis with consumption data.

A Analytical Appendix

A.1 Derivation of Equation (4)

Take a fourth order Taylor series expansion of the age 1 subperiod utility function around $c_1 = \mu_1^c$ to get

$$U \approx \frac{c_0^{1-\frac{1}{\gamma}}}{1-\frac{1}{\gamma}} + \frac{1}{1-\frac{1}{\gamma}} \left(\mu_1^{c^{1-\theta}} + \mathbb{E}\left[(1-\theta)\mu_1^{c^{-\theta}} (c_1 - \mu_1^c) - \frac{(1-\theta)\theta}{2} \mu_1^{c^{-(1+\theta)}} (c_1-\mu_1^c)^2 \right.\right.$$
$$\left.\left. + \frac{(1-\theta)\theta(1+\theta)}{6} \mu_1^{c^{-(2+\theta)}} (c_1-\mu_1^c)^3 - \frac{(1-\theta)\theta(1+\theta)(2+\theta)}{24} \mu_1^{c^{-(3+\theta)}} (c_1-\mu_1^c)^4 \right]\right)$$

Under constraint (3) and the additional assumption that $\mathbb{E}[\exp(\varepsilon)] = 1$ we obtain $\mu_1^c = 1$. Also impose that $\theta = \frac{1}{\gamma}$. Using these conditions in the above we obtain (4).

A.2 Formal Proof of Proposition 1

Proof. Let $\mu_1^\epsilon = E_\Psi[\epsilon] = \int \epsilon d\Psi$, $\mu_i^\epsilon = \int (\epsilon - \mu_1^\epsilon)^i \, d\Psi$ for $i > 1$, and let $E_\Psi[\exp(\epsilon)] = \int \exp(\epsilon) d\Psi = 1$. Denote by $\tilde{\Psi}^{\delta_i}(\epsilon)$ a mean preserving (constant μ_1^ϵ) distribution function that is obtained from $\Psi(\epsilon)$ by changing central moment μ_i^ϵ holding all other moments $\mu_{\neg i}^\epsilon$ for $i > 1$ constant. Also, define the random variable $\tilde{\epsilon}^{\delta_i} = \epsilon + \Delta^{\delta_i}$, which is obtained from ϵ by shifting all realizations by the constant Δ^{δ_i}. Let the normalization $E_{\tilde{\Psi}^{\delta_i}}[\exp(\tilde{\epsilon}^{\delta_i})] = E_{\tilde{\Psi}^{\delta_i}}[\exp(\epsilon + \Delta^{\delta_i})] = \int \exp(\epsilon + \Delta^{\delta_i}) d\tilde{\Psi}^{\delta_i} = \exp(\Delta^{\delta_i}) \int \exp(\epsilon) d\tilde{\Psi}^i = 1$ define the shift parameter $\Delta^{\delta_i} = -\ln\left(\int \exp(\epsilon) d\tilde{\Psi}^i\right)$. Finally, observe that $E_{\tilde{\Psi}^{\delta_i}}[\epsilon + \Delta^{\delta_i}] = E_{\tilde{\Psi}}[\epsilon] + \Delta^{\delta_i} = E_\Psi[\epsilon] + \Delta^{\delta_i}$ since μ_1^ϵ is held constant. With logarithmic utility and binding constraint (3), the expected utility difference across distributions Ψ and $\tilde{\Psi}^{\delta_i}$ is thus $\Delta U = (U \mid \Psi) - (U \mid \tilde{\Psi}) = \Delta^{\delta_i}$ and thus exclusively driven by the shift parameter. We then get the following:

- Shifting probability mass from the center to the tails, either by increasing the variance ($i = 2$) or kurtosis ($i = 4$) holding constant all $\mu_{\neg i}^\epsilon$ for $i > 1$ increases $\int \exp(\epsilon) d\tilde{\Psi}^i$ above one which follows from Jensen's inequality for convex functions. Thus $\Delta^{\delta_i} < 0$.

- Shifting probability mass from the right tail to the left tail decreasing the skewness ($i = 3$) (i.e., making the distribution more left-skewed), holding constant all $\mu_{\neg i}^\epsilon$ for $i > 1$ decreases $\int \exp(\epsilon) d\tilde{\Psi}^i$ below one which follows from Jensen's inequality for convex functions. Thus $\Delta^{\delta_i} > 0$.

$\square$

A.3 Derivation of Equation (5)

Rewrite (2) as

$$v(c_1,\theta,\Psi)=\left(\int \tilde{g}(c_1(\varepsilon))d\Psi(\varepsilon)\right)^{\frac{1}{1-\theta}},\quad \text{where}\quad \tilde{g}(c_1(\varepsilon))=c_1(\varepsilon)^{1-\theta}.$$

Take a fourth order Taylor series expansion of $\tilde{g}(c_1(\varepsilon))$ around μ_1^c, noticing that $c_1=\exp(\varepsilon)$ and $E[\exp(\varepsilon)]=1$ to get

$$\mathbb{E}[\tilde{g}(c_1(\varepsilon))]\approx 1+(1-\theta)\left(-\frac{1}{2}\theta\mu_2^{\exp(\varepsilon)}+\frac{1}{6}\theta(1+\theta)\mu_3^{\exp(\varepsilon)}-\frac{1}{24}\theta(1+\theta)(2+\theta)\mu_4^{\exp(\varepsilon)}\right)$$

and thus the certainty equivalent is approximated as in equation (5).

A.4 Derivation of Equation (6)

Take a fourth order Taylor series expansion of the RHS of the first-order condition around $\mathbb{E}[\exp(\varepsilon)]=1$ to get

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

A.5 Precautionary Savings for CRRA Utility

Rewrite the first-order condition, equation (6), as an implicit function

$$\begin{aligned}
e\left(a_1,\mu_i^{\exp(\epsilon)}\right)=\ &(y_0-a_1)^{-\frac{1}{\gamma}}-(1+a_1)^{-\theta}-\frac{\theta(1+\theta)}{2}\,(1+a_1)^{-(2+\theta)}\,\mu_2^{\exp(\varepsilon)}\\
&+\frac{\theta(1+\theta)(2+\theta)}{6}\,(1+a_1)^{-(3+\theta)}\,\mu_3^{\exp(\varepsilon)}\\
&-\frac{\theta(1+\theta)(2+\theta)(3+\theta)}{24}\,(1+a_1)^{-(4+\theta)}\,\mu_4^{\exp(\varepsilon)}=0
\end{aligned}$$

and from the total differential of $e(\cdot)$ note that

$$\frac{da_1}{d\mu_i^{\exp(\epsilon)}}=-\frac{\frac{\partial e(\cdot)}{\partial \mu_i^{\exp(\epsilon)}}}{\frac{\partial e(\cdot)}{\partial a_1}}$$

Note that since $\mu_2^{\exp(\epsilon)} > 0, \mu_3^{\exp(\epsilon)} < 0, \mu_4^{\exp(\epsilon)} > 0$ we have $\frac{\partial e(\cdot)}{\partial a_1} > 0$, which reflects that the marginal utility of savings is decreasing in a_1. Also note that $\frac{\partial e(\cdot)}{\partial \mu_i^{\exp(\epsilon)}} < 0$ for $i = 2, 4$ and $\frac{\partial e(\cdot)}{\partial \mu_3^{\exp(\epsilon)}} > 0$. Thus, $\frac{da_1}{d\mu_i^{\exp(\epsilon)}} > 0$ for $i = 2, 4$ and $\frac{da_1}{d\mu_3^{\exp(\epsilon)}} < 0$.

A.6 Proof of Proposition 2

We proof the proposition using an alternative strategy to Kimball and Weil (2009). As in their analysis, we assume that the marginal utility of saving, the RHS of (7), is a decreasing function of a_1. Our proof is adopted from Krueger and Ludwig (2019).

Proof. Rewrite the RHS of the first-order condition in (7) as

$$\begin{aligned} RHS &= v(c_1, \theta, \Psi)^{\theta - \frac{1}{\gamma}} f(c_1, \theta, \Psi) & (23) \\ &= v(c_1, \theta, \Psi)^{1 - \frac{1}{\gamma}} \frac{\mathbb{E}\left[\left(\exp(\varepsilon) + a_1\right)^{-\theta} \right]}{\mathbb{E}\left[\left(\exp(\varepsilon) + a_1\right)^{1-\theta} \right]} \\ &= v(c_1, \theta, \Psi)^{1 - \frac{1}{\gamma}} h(c_1, \theta, \Psi). & (24) \end{aligned}$$

where $f(c_1, \theta, \Psi) = \mathbb{E}\left[\left(\exp(\varepsilon) + a_1\right)^{-\theta} \right]$ and $h(c_1, \theta, \Psi) = \frac{f(c_1, \theta, \Psi)}{g(c_1, \theta, \Psi)}$, where $g(c_1, \theta, \Psi) = \mathbb{E}\left[\left(\exp(\varepsilon) + a_1\right)^{1-\theta} \right]$. Consider the following case distinction:

1. $\gamma = 1$: Then the RHS is simply from (24)

$$RHS = h(c_1, \theta, \Psi)$$

giving rise to the following case distinction with respect to θ (throughout, we assume that $\theta > 0$, $\theta < \infty$):

   (a) $\theta \in (0, 1]$: $h(\cdot)$ is the ratio of function $f(\cdot)$ which is strictly convex in $\exp(\epsilon)$ in the numerator and function $g(\cdot)$ which is concave in $\exp(\varepsilon)$ in the denominator (the denominator equals 1 for $\theta = 1$). Thus, an increase of (higher-order) risk increases $h(\cdot)$.

   (b) $\theta > 1$: $h(\cdot)$ is the ratio of two strictly convex functions $f(\cdot), g(\cdot)$ in $\exp(\epsilon)$, where the degree of convexity is stronger in the numerator than in the denominator (the exponent in the numerator is θ and in the denominator it is $1 - \theta$). Thus, an increase of (higher-order) risk increases $h(\cdot)$.

   Thus, an increase of (higher-order) risk unambiguously increases the RHS in (24), increasing precautionary savings.

2. $\gamma < 1$: For the behavior of $h(\cdot)$ the same logic as in item 1 applies. Furthermore, an increase of risk decreases $v(\cdot)$, which, for $\gamma < 1$, increases $v(\cdot)^{1-\frac{1}{\gamma}}$, since $1 - \frac{1}{\gamma} <$

0. Thus, an increase of risk unambiguously increases the RHS in (24), increasing precautionary savings.

3. $\gamma > 1$: We obtain the following case distinction from (23):

   (a) $\theta \leq \frac{1}{\gamma}$: An increase of risk increases $v(\cdot)^{\theta-\frac{1}{\gamma}}$ (respectively leaves it unchanged at 1 if $\theta = \frac{1}{\gamma}$), so that an increase of risk unambiguously increases the RHS in (23), increasing precautionary savings.

   (b) $\theta > \frac{1}{\gamma}$: the overall effect is ambiguous.

□

A.7 Decomposition of Consumption Equivalent Variations

We evaluate the welfare implications of higher-order risk by computing the consumption equivalent variation (CEV) that makes households that live in the world with shock distributions NORM indifferent to live with shock distributions $i \in \{LK, LKSW\}$.

A.7.1 Decomposition in the 2-Period Model

We start with the decomposition for the two-period model of Section 3, which extends to the quantitative model in a straightforward fashion, as we show in the next subsection. Under the convenient transformation[35] of utility $V = \left[\left(1 - \frac{1}{\gamma}\right) U\right]^{\frac{1}{1-\frac{1}{\gamma}}}$ we compute

$$g_c^i = \frac{V(C^i)}{V(C^{NORM})} - 1 \tag{25}$$

and thus the respective CEVs are defined as the percentage consumption loss in each period from the respective distribution with higher order risk relative to the distribution NORM.

We further decompose the CEV into *mean* and *distribution* effects. The mean effect is the welfare effect stemming from changes in average consumption and the distribution effect captures changes in the distribution of consumption. Formally, let $\mathbb{E}[C^i] = \frac{1}{2}\left(c_0^i + \int c_1^i(\varepsilon) d\Psi^i(\varepsilon)\right)$ for $i \in \{NORM, LK, LKSW\}$. Denote by $\delta_c^i = \frac{\mathbb{E}[C^i]}{\mathbb{E}[C^{NORM}]} - 1$ the percent change of consumption for $i \in \{LK, LKSW\}$. Then, the distribution effect corrects for the percentage change of mean consumption and is thus given by

$$g_c^{distr^i} = \frac{V\left(\frac{C^i}{1+\delta_c^i}\right)}{V(C^{NORM})} - 1 = \frac{1+g_c^i}{1+\delta_c^i} - 1. \tag{26}$$

[35] I.e., we retransform to the standard EZW functional, cf. Footnote 10.

The corresponding mean effect is accordingly

$$g_c^{mean^i} = g_c^i - g_c^{distr^i} = \frac{1+g_c^i}{1+\delta_c^i}\delta_c^i \approx \delta_c^i. \tag{27}$$

The distribution effect itself captures two changes. The first reflects the utility difference stemming from the change of the average life-cycle consumption profile, which we refer to as the *life-cycle distribution* effect. The second captures the utility change stemming from the change of the cross-sectional distribution of stochastic second period consumption, which we accordingly refer to as the *cross-sectional distribution* effect. Thus, we can rewrite $g_c^{distr^i}$ as

$$g_c^{distr^i} = g_c^{lcd^i} + g_c^{csd^i} \tag{28}$$

for the CEV stemming from the life-cycle redistribution (lcd) and cross-sectional distribution (csd) effect.

To compute the $g_c^{csd^i}$, first let $\mathbb{E}[C^i \mid j]$ denote the age j specific mean consumption, i.e., $\mathbb{E}[C^i \mid j=0] = c_0^i$ and $\mathbb{E}[C^i \mid j=1] = \int c_1^i(\varepsilon)d\Psi^i(\varepsilon)$. Next compute the age j specific consumption growth rate as $\delta_j^{c^i} = \frac{\mathbb{E}[C^i|j]}{\mathbb{E}[C^{NORM}|j]}$ for $i \in \{LK, LKSW\}$. Then compute the utility in distribution scenario $i \in \{LK, LKSW\}$ after correcting for mean consumption growth as

$$\tilde{V}^i = \left(\left(\frac{1}{1+\delta_0^{c^i}}\right)^{1-\frac{1}{\gamma}} c_0^{i\,1-\frac{1}{\gamma}} + \left(\frac{1}{1+\delta_1^{c^i}}\right)^{1-\frac{1}{\gamma}} v(c_1^i,\theta,\Psi^i)^{1-\frac{1}{\gamma}}\right)^{\frac{1}{1-\frac{1}{\gamma}}},$$

which for $\gamma = 1$ simplifies to

$$\tilde{V}^i = \frac{1}{1+\delta_0^{c^i}}\frac{1}{1+\delta_1^{c^i}} \cdot c_0^i \cdot v(c_1^i,\theta,\Psi^i) = \frac{1}{1+\delta_0^{c^i}}\frac{1}{1+\delta_1^{c^i}} V^i.$$

Having corrected for the percent change of age-specific mean consumption, the CEV from the cross-sectional distribution effect is then

$$g_c^{csd^i} = \frac{\tilde{V}^i}{V(C^{NORM})} - 1 = \frac{1+g_c^i}{1+\delta_c^i} - 1 \tag{29}$$

and thus the life-cycle distribution effect follows as

$$g_c^{lcd^i} = g_c^{distr^i} - g_c^{csd^i}. \tag{30}$$

A.7.2 Decomposition in the Full Life Cycle Model

The decomposition into the mean and distribution effect is analogous to the two-period model, where average consumption is given by

$$\mathbb{E}[C^i] = \frac{1}{J+1}\sum_{j=0}^{J}\int c_j^i(a_j, z_j; s)d\Psi_j^i(a_j, z_j; s)$$

for $i \in \{NORM, LK, LKSW\}$, where $c_j^i(a_j, z_j; s)$ is the consumption policy function in distribution i and $\Phi_j^i(a_j, z_j; s)$ is the cross-sectional distribution.

To compute the cross-sectional distribution effect, let, as above, the age j specific consumption growth rate be $\delta_j^{c^i} = \frac{\mathbb{E}[C^i|j]}{\mathbb{E}[C^{NORM}|j]}$ for $i \in \{LK, LKSW\}$, where now $\mathbb{E}[C^i \mid j] = \int c_j^i(a_j, z_j; s)d\Phi_j^i(a_j, z_j; s)$. Next, observe that

$$\begin{aligned}\tilde{V}_J^i &= \left((1-\hat{\beta})\left(\frac{c_J^i}{\delta_J^{c^i}}\right)^{1-\frac{1}{\gamma}}\right)^{\frac{1}{1-\frac{1}{\gamma}}} = \frac{1}{\delta_J^{c^i}}V_J^B \\ v\left(\tilde{V}_J^i\right) &= \frac{1}{\delta_J^{c^i}}v\left(V_J^i\right).\end{aligned}$$

and thus

$$\tilde{V}_{J-1}^i = \left((1-\tilde{\beta})\left(\frac{1}{\delta_{J-1}^{c^i}}\right)^{1-\frac{1}{\gamma}}\left(c_{J-1}^i\right)^{1-\frac{1}{\gamma}} + \tilde{\beta}\left(v\left(\tilde{V}_J^i\right)\right)^{1-\frac{1}{\gamma}}\right)^{\frac{1}{1-\frac{1}{\gamma}}}$$

which extends to any period j as

$$\tilde{V}_j^i = \left((1-\tilde{\beta})\left(\frac{c_j^i}{\delta_j^{c^i}}\right)^{1-\frac{1}{\gamma}} + \tilde{\beta}\left(v\left(\tilde{V}_{j+1}^i\right)\right)^{1-\frac{1}{\gamma}}\right)^{\frac{1}{1-\frac{1}{\gamma}}}.$$

With the parametric restriction $\gamma = 1$ the decomposition simplifies. For $\gamma = 1$ we get

$$\tilde{V}_J^i = \exp\left((1-\tilde{\beta})\ln\left(\frac{c_J^i}{\delta_J^{c^i}}\right)\right) = \left(\frac{1}{\delta_J^{c^i}}\right)^{1-\tilde{\beta}}V_J^i$$

and thus

$$\begin{aligned}
\tilde{V}^i_{J-1} &= \exp\left((1-\tilde{\beta})\ln\left(\frac{c^i_{J-1}}{\delta^{c^i}_{J-1}}\right) + \tilde{\beta}\ln\left(v\left(\tilde{V}^i_J\right)\right)\right) \\
&= \exp\left((1-\tilde{\beta})\ln\left(\frac{1}{\delta^{c^i}_{J-1}}\right) + (1-\tilde{\beta})\ln\left(c^i_{J-1}\right) + \tilde{\beta}(1-\tilde{\beta})\ln\left(\frac{1}{\delta^{c^i}_J}\right) + \tilde{\beta}\ln\left(v\left(V^i_J\right)\right)\right) \\
&= \left(\left(\frac{1}{\delta^{c^i}_{J-1}}\right)\left(\frac{1}{\delta^{c^i}_J}\right)^{\tilde{\beta}}\right)^{1-\tilde{\beta}} V^i_J
\end{aligned}$$

Continuing along these lines we get

$$\tilde{V}^i_0 = \left(\prod_{j=0}^{J}\left(\frac{1}{\delta^{c^i}_j}\right)^{\tilde{\beta}^j}\right)^{1-\tilde{\beta}} V^i_0.$$

With this construction we can now decompose the CEV into the cross-sectional and the life-cycle distribution effects using (29) and (30).

B Discretization of the FGLD

For each Flexible Generalized Lambda Distribution (FGLD) our discretization procedure is as follows:

1. Determine the endpoints of a grid $\mathcal{G}^{\tilde{x}}$ from the quantile function of the FGLD for a small probability $\tilde{\pi}_1 = \varepsilon$ such that

$$\begin{aligned}
\tilde{x}_1 &= Q(\tilde{\pi}_1) \\
\tilde{x}_n &= Q(1-\tilde{\pi}_1).
\end{aligned}$$

2. Build grid $\mathcal{G}^{\tilde{x}}$ by drawing n equidistant nodes on the interval $[\tilde{x}_1, \tilde{x}_n]$.

3. For $\tilde{x}_i \in \mathcal{G}^{\tilde{x}}$, $i = 1, n-1$ compute auxiliary gridpoint $\tilde{\bar{x}}_i = \frac{\tilde{x}_{i+1}+\tilde{x}_i}{2}$.

4. On all $\tilde{\bar{x}}_i$ compute cumulative probability p_i from the quantile function of the FGLD. Since the quantile function of the FGLD maps $\tilde{\bar{x}}_i = Q(p_i)$, this requires a numerical solver to compute $p_i = Q^{-1}(\tilde{\bar{x}}_i)$.

5. Now assign to gridpoint $\tilde{x}_1$ the probability $\pi_1 = p_1$ and to all gridpoints i, $i = 2, \ldots, n-1$, the probability $\pi_i = p_i - p_{i-1}$ and to gridpoint $\tilde{x}_n$ the probability $1-p_{n-1}$.

C Calibration Appendix

C.1 Moments of the FGLD Distribution

Tables C.1– C.3 summarize the moments for distributions NORM, LK, and LKSW, and Table C.4 contains the corresponding parameters of λ of the fitted FGLD distributions.

Table C.1: Moments: Distribution NORM

Moment	$\hat{\mu}_2$	$\hat{\mu}_3$	$\hat{\mu}_4$
Transitory Shock:			
target	0.05	0	0.008
fitted	0.05	0	0.008
discrete	0.05	0	0.008
Persistent Shock—Contraction:			
target	0.022	0	0.001
fitted	0.022	0	0.001
discrete	0.022	0	0.001
Persistent Shock—Expansion:			
target	0.009	0	0
fitted	0.009	0	0
discrete	0.009	0	0

Notes: This table shows the target central moment together with the central moment of the fitted FGLD, and of the discretized FGLD for the distribution NORM, cf. Section 6.2.

Table C.2: Moments: Distribution LK

Moment	$\hat{\mu}_2$	$\hat{\mu}_3$	$\hat{\mu}_4$
Transitory Shock:			
target	0.05	0	0.219
fitted	0.05	0	0.219
discrete	0.05	0	0.219
Persistent Shock—Contraction:			
target	0.022	0	0.061
fitted	0.022	0	0.061
discrete	0.022	0	0.061
Persistent Shock—Expansion:			
target	0.009	0	0.008
fitted	0.009	0	0.008
discrete	0.009	0	0.008

Notes: This table shows the etarget central moment together with the central moment of the fitted FGLD, and of the discretized FGLD for the distribution LK, cf. Section 6.2.

Table C.3: Moments: Distribution LKSW

Moment	$\hat{\mu}_2$	$\hat{\mu}_3$	$\hat{\mu}_4$
Transitory Shock:			
target	0.05	-0.047	0.102
fitted	0.05	-0.047	0.102
discrete	0.051	-0.051	0.107
Persistent Shock—Contraction:			
target	0.022	-0.016	0.066
fitted	0.022	-0.016	0.066
discrete	0.023	-0.02	0.07
Persistent Shock—Expansion:			
target	0.009	-0.001	0.01
fitted	0.009	-0.001	0.01
discrete	0.009	-0.002	0.01

Notes: This table shows the target central moment together with the central moment of the fitted FGLD, and of the discretized FGLD for the distribution LKSW, cf. Section 6.2.

Table C.4: Fitted Parameters of FGLD

Parameter	$\hat{\lambda}_1$	$\hat{\lambda}_2$	$\hat{\lambda}_3$	$\hat{\lambda}_4$
		NORM		
Transitory:	1.000	0.359	5.203	5.203
Pers.—Contraction:	1.000	0.539	5.203	5.203
Pers.—Expansion:	1.000	0.871	5.203	5.203
		LK		
Transitory:	1.000	0.002	173.309	173.309
Pers.—Contraction:	1.000	0.002	244.954	244.954
Pers.—Expansion:	1.000	0.003	220.344	220.344
		LKSW		
Transitory:	0.197	0.008	92.959	57.755
Pers.—Contraction:	0.425	0.002	289.898	225.714
Pers.—Expansion:	0.894	0.003	275.612	256.735

Notes: This table shows the estimated λ-values for the fitted FGLD for distributions NORM, LK and LKSW, cf. Section 6.2.

C.2 The Bend Point Formula and the Pension Indexation Factor

Approximating the AIME with the last income state before entering into retirement z_{j_r-1} the primary insurance amount according to the bend point formula is determined as follows:

$$
p(z_{j_r-1}) = \begin{cases} s_1 z_{j_r-1} & \text{for } z_{j_r-1} < b_1 \\ s_1 b_1 + s_2 \left(z_{j_r-1} - b_1\right) & \text{for } b_1 \leq z_{j_r-1} < b_2 \\ s_1 b_1 + s_2 \left(b_2 - b_1\right) + s_3 \left(z_{h_r-1} - b_2\right) & \text{for } b_2 \leq z_{j_r-1} < b_3 \\ s_1 b_1 + s_2 \left(b_2 - b_1\right) + s_3 \left(b_3 - b_2\right) & \text{for } z_{j_r-1} \geq b_3 \end{cases}
$$

Table C.5 contains the calibrated values of the pension indexation factor ϱ, which is required to clear the budget of the pension system.

Table C.5: Pension Indexation Factor ϱ

	CR	NCR
NORM	0.6817	0.6692
LK	0.7007	0.6787
LKSW	0.6866	0.6758

Notes: Calibrated pension benefit level ϱ under a balanced budget. CR: cyclical risk, NCR: no cyclical risk.

C.3 Moments of the Earnings Process

Table C.6 shows cross-sectional central moments of the earnings distribution in logs and levels at labor market entry (age 25) and exit (age 60). We observe that all distributions are skewed to the right in levels and that, despite left skewness in logs, right skewness of distribution LKSW is higher in levels than of distribution NORM. Furthermore, the variance is initially lower in distribution LKSW than in distribution NORM.[36] Both features constitute a source of welfare gains from higher-order income risk, whereas the higher kurtosis in levels and the increasing variance work against it. Finally, skewness and in particular kurtosis in levels under distribution LK are extremely high. Left-skewness in logs in distribution LKSW substantially reduces both moments.

Figures C.1 and C.2 summarize the calibration of the earnings process during the working period and the pension income in retirement for central moments 1-4 of the earnings distribution in levels and logs, respectively.

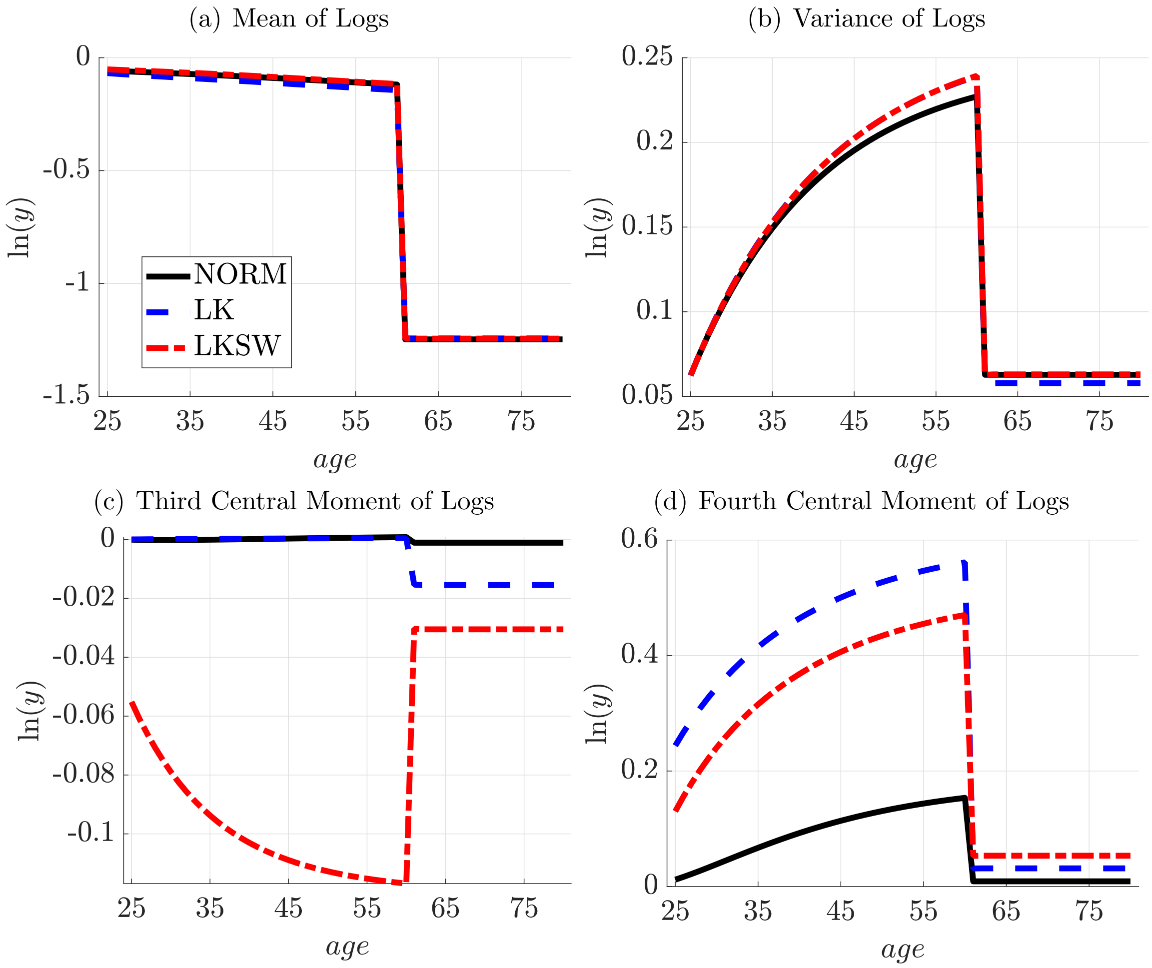
[36] By construction, the variance of the log earnings distribution is the same across distribution scenarios. The difference of 0.01 showing up at age 60 is due to numerical inaccuracies of coarse grids for assets a and the persistent income state z.

Table C.6: Moments of the Earnings Distribution in Logs and Levels

	Logs			Levels		
	Age 25 ($j = 0$)					
	NORM	LK	LKSW	NORM	LK	LKSW
μ_2	0.06	0.06	0.06	0.06	0.36	0.05
μ_3	0	0	-0.06	0.01	4.44	0.09
μ_4	0.01	0.24	0.13	0.01	129.43	0.41
	Age 60 ($j = 35$)					
μ_2	0.23	0.24	0.24	0.25	0.86	0.3
μ_3	0	0	-0.12	0.21	27.52	1.12
μ_4	0.15	0.56	0.47	0.5	27889.82	27.85

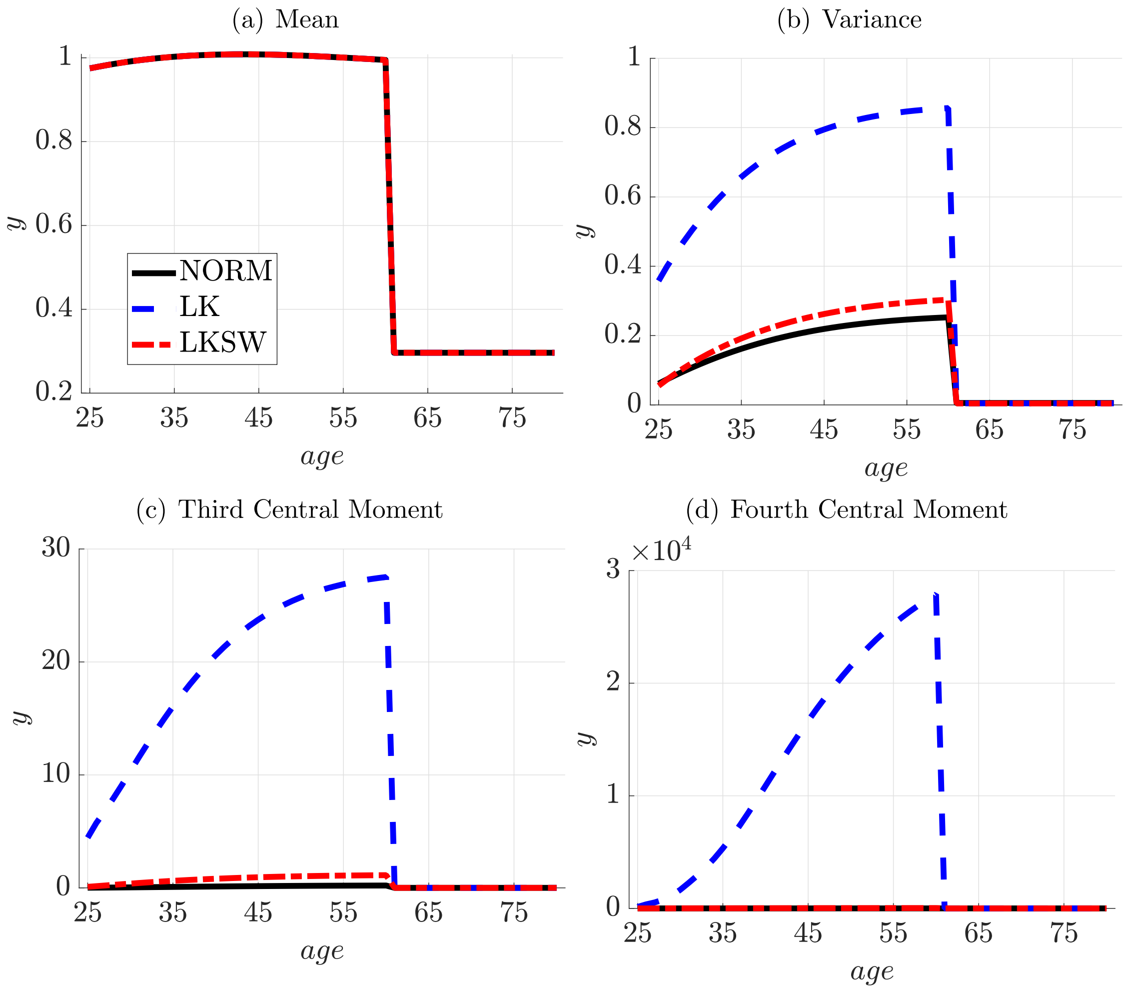
Notes: Moments of cross-sectional distribution of log earnings and earnings at ages 25 ($j = 0$) and 60 ($j = 35$) for each scenario of shock distributions. NORM: FGLD with moments of the normal distribution, LK: FGLD with excess kurtosis, LKSW: FGLD with excess kurtosis and left-skewness (in logs).

Figure C.1: Moments of Life-Cycle Earnings by Age: Logs



Notes: Figures show moments of cross-sectional distribution of log earnings over the life-cycle for each scenario of shock distributions. NORM: FGLD with moments of the normal distribution, LK: FGLD with excess kurtosis, LKSW: FGLD with excess kurtosis and left-skewness (in logs).

Figure C.2: Moments of Life-Cycle Earnings by Age: Levels



Notes: Figures show moments of cross-sectional distribution of earnings over the life-cycle for each scenario of shock distributions. NORM: FGLD with moments of the normal distribution, LK: FGLD with excess kurtosis (in logs), LKSW: FGLD with excess kurtosis and left-skewness (in logs).

C.4 Calibration in the General Equilibrium Variant of the Model

First, we determine parameters of the production function of a representative firm such that the baseline model (distribution LKSW and risk scenario CR) in which the aggregate net wage level is normalized to $w^n = 1$ and the rate of return of $r[\%] = 4.2\%$ is consistent with a general equilibrium interpretation. Second, we hold constant these parameters in all other scenarios and determine wages and interest rates endogenously.

Assuming Cobb-Douglas production with capital elasticity α and a technology level Υ output of the representative firm is

$$Y = \Upsilon K^{\alpha} L^{1-\alpha}.$$

Denoting by $k = \frac{K}{L}$ the capital intensity and assuming a constant depreciation rate of δ the first-order conditions are given by

$$r = \Upsilon \alpha k^{\alpha-1} - \delta \tag{31a}$$

$$w = \Upsilon(1-\alpha)k^{\alpha}, \tag{31b}$$

which also implies that

$$\frac{w}{r+\delta} = \frac{1-\alpha}{\alpha}k. \tag{32}$$

Assuming capital market clearing in a closed economy so that aggregate assets are equal to the capital stock $K = A$, and knowing that aggregate efficient labor in our economy is normalized to $L = h_r - 1$, we can compute $k = \frac{A}{h_r - 1}$, and given prices r and $w = \frac{1}{1-\tau-\tau^p}$ (since net wages $w^n = 1$) the implied depreciation rate follows from using this in (32) as

$$\delta = \frac{w}{\frac{1-\alpha}{\alpha}k} - r = \frac{\frac{1}{1-\tau-\tau^p}}{\frac{1-\alpha}{\alpha}\frac{A}{h_r-1}} - r$$

as well as the implied technology level follows from (31a) as

$$\Upsilon = \frac{r+\delta}{\alpha}k^{1-\alpha} = \frac{r+\delta}{\alpha}\left(\frac{A}{h_r-1}\right)^{1-\alpha}.$$

Table C.7 summarizes this calibration. The calibrated depreciation rate is low, which is not surprising giving our target of a wealth to income ratio from the data and an interest rate of 4.2%.

Having determined the parameters δ, Υ in the economy CR/LKSW for each level of risk aversion as summarized in Table C.7 we then hold constant δ, Υ in all other economies and iterate on the interest rate until market clearing. In each iteration, we compute wages

Table C.7: Technology Level and Depreciation Rate in General Equilibrium Variant

Parameter		
θ	δ	Υ
1	0.0159	0.9234
2	0.0163	0.9252
3	0.0167	0.9271
4	0.0171	0.9293

Notes: Calibrated depreciation rate δ and technology level Υ in the general equilibrium variant of the model.

given the interest rate from (32) and (31b) as

$$w = \Upsilon^{\frac{1}{1-\alpha}}(1-\alpha)\left(\frac{\alpha}{r+\delta}\right)^{\frac{\alpha}{1-\alpha}}$$

and net wages as $w^n = (1 - \tau^p - \tau)w$.

Table C.8 shows the general equilibrium prices for each distribution scenario, and for the two versions with and without cyclical idiosyncratic risk. In the economy with cyclical risk, there are additional precautionary savings which in general equilibrium increases the capital stock increasing wages and decreasing returns. Thus wage rates are higher, and returns are lower. Furthermore, this difference in net wages and returns between increases in risk aversion θ because the precautionary savings reaction is stronger with higher risk aversion.

Table C.8: Aggregate Prices in General Equilibrium Model

Variable	w^n		r	
	CR	NCR	CR	NCR
	Risk Aversion, $\theta = 1$			
NORM	1.0009	0.9979	0.0419	0.0422
LK	1.0033	1	0.0416	0.042
LKSW	1	0.9976	0.042	0.0423
	Risk Aversion, $\theta = 4$			
NORM	0.9911	0.9815	0.0431	0.0443
LK	1	0.9876	0.042	0.0435
LKSW	1	0.9856	0.042	0.0438

Notes: Net wage w^n and return r in the general equilibrium variants of the model. CR: cyclical idiosyncratic risk, NCR: no cyclical idiosyncratic risk.

D Additional Results

D.1 Comparison to the Normal Distribution

Table D.1 documents the CEV in distribution NORM (an FGLD with zero skewness and a kurtosis of 3) to one where shocks are drawn from a normal distribution using standard Gaussian Quadrature methods. Differences are very small.

Table D.1: Welfare Effects of Cyclical Idiosyncratic Risk: FGLD(NORM) versus Normal Distribution

CEV	g_c	g_c^{mean}	g_c^{lcd}	g_c^{csd}
Risk Aversion, $\theta = 1$				
NORM	-1.72	0.499	-2.175	-0.044
NORMAL	-1.722	0.5	-2.176	-0.045
Risk Aversion, $\theta = 2$				
NORM	-3.263	0.898	-4.038	-0.123
NORMAL	-3.268	0.898	-4.043	-0.123
Risk Aversion, $\theta = 3$				
NORM	-4.607	1.229	-5.638	-0.198
NORMAL	-4.615	1.23	-5.646	-0.199
Risk Aversion, $\theta = 4$				
NORM	-5.758	1.515	-7.009	-0.264
NORMAL	-5.767	1.516	-7.018	-0.265

Notes: Welfare gains (positive numbers) and losses (negative numbers) of cyclical idiosyncratic risk expressed as consumption equivalent variation (CEV) for FGLD distribution NORM and the normal distribution, NORMAL. 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 $g_c = g_c^{mean} + g_c^{lcd} + g_c^{csd}$.

D.2 Sensitivity Analyses

The main text summarizes the results of our sensitivity analyses in Table 6. The subsequent tables D.2 to D.5 contain further details. This underscores the robustness of our findings also with respect to the dominant role played by the life-cycle distribution effect g_c^{lcd}.

Table D.2: Welfare Effects of Cyclical Idiosyncratic Risk: General Equilibrium Model

CEV	g_c	g_c^{mean}	g_c^{lcd}	g_c^{csd}	Δg_c
			Risk Aversion, $\theta = 1$		
NORM	-1.018	0.226	-1.2	-0.044	0
LK	-1.196	0.281	-1.439	-0.038	-0.178
LKSW	-0.884	0.18	-1.03	-0.035	0.134
			Risk Aversion, $\theta = 2$		
NORM	-2	0.416	-2.294	-0.122	0
LK	-2.257	0.426	-2.575	-0.108	-0.257
LKSW	-2.367	0.381	-2.639	-0.11	-0.367
			Risk Aversion, $\theta = 3$		
NORM	-2.872	0.582	-3.258	-0.196	0
LK	-4.191	0.619	-4.616	-0.194	-1.319
LKSW	-5.282	0.664	-5.704	-0.242	-2.41
			Risk Aversion, $\theta = 4$		
NORM	-3.611	0.735	-4.082	-0.264	0
LK	-7.394	0.861	-7.934	-0.321	-3.783
LKSW	-9.619	0.984	-10.17	-0.433	-6.008

Notes: Welfare gains (positive numbers) and losses (negative numbers) in general equilibrium of cyclical idiosyncratic risk expressed as Consumption Equivalent Variation (CEV) in the distribution with normal moments NORM, the leptokurtic distribution LK and the leptokurtic and left-skewed distribution 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 $g_c = g_c^{mean} + g_c^{lcd} + g_c^{csd}$. $\Delta g_c = g_c^i - g_c^{NORM}$, for $i \in \{LK, LKSW\}$: change in percentage points relative to distribution NORM.

Table D.3: Welfare Effects of Cyclical Idiosyncratic Risk: CRRA Preferences

CEV	g_c	g_c^{mean}	g_c^{lcd}	g_c^{csd}	Δg_c
			Risk Aversion, $\theta = 2$		
NORM	-2.552	0.709	-3.139	-0.122	0
LK	-2.813	0.756	-3.456	-0.113	-0.261
LKSW	-2.564	0.626	-3.074	-0.115	-0.012
			Risk Aversion, $\theta = 3$		
NORM	-3.335	0.901	-4.035	-0.201	0
LK	-4.055	0.966	-4.803	-0.218	-0.72
LKSW	-4.456	0.937	-5.121	-0.272	-1.121
			Risk Aversion, $\theta = 4$		
NORM	-4.072	1.081	-4.876	-0.277	0
LK	-6.081	1.243	-6.91	-0.414	-2.009
LKSW	-7.53	1.337	-8.267	-0.6	-3.458

Notes: Welfare gains (positive numbers) and losses (negative numbers) for CRRA utility where $\theta = \frac{1}{\rho}$ of cyclical idiosyncratic risk expressed as Consumption Equivalent Variation (CEV) in the distribution with normal moments NORM, the leptokurtic distribution LK and the leptokurtic and left-skewed distribution 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 $g_c = g_c^{mean} + g_c^{lcd} + g_c^{csd}$. $\Delta g_c = g_c^i - g_c^{NORM}$, for $i \in \{LK, LKSW\}$: change in percentage points relative to distribution NORM.

Table D.4: Welfare Effects of Cyclical Idiosyncratic Risk: Zero Initial Assets

CEV	g_c	g_c^{mean}	g_c^{lcd}	g_c^{csd}	Δg_c
			Risk Aversion, $\theta = 1$		
NORM	-1.893	0.54	-2.317	-0.116	0
LK	-2.139	0.622	-2.706	-0.055	-0.246
LKSW	-1.612	0.439	-2.001	-0.05	0.281
			Risk Aversion, $\theta = 2$		
NORM	-3.609	0.965	-4.326	-0.248	0
LK	-4.091	0.999	-4.938	-0.152	-0.482
LKSW	-4.293	0.932	-5.035	-0.19	-0.684
			Risk Aversion, $\theta = 3$		
NORM	-5.113	1.314	-6.06	-0.367	0
LK	-7.685	1.445	-8.83	-0.299	-2.572
LKSW	-9.725	1.573	-10.812	-0.486	-4.612
			Risk Aversion, $\theta = 4$		
NORM	-6.404	1.61	-7.528	-0.485	0
LK	-13.586	1.914	-14.924	-0.577	-7.182
LKSW	-17.14	2.112	-18.314	-0.939	-10.736

Notes: Welfare gains (positive numbers) and losses (negative numbers) with zero initial assets of cyclical idiosyncratic risk expressed as Consumption Equivalent Variation (CEV) in the distribution with normal moments NORM, the leptokurtic distribution LK and the leptokurtic and left-skewed distribution 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 $g_c = g_c^{mean} + g_c^{lcd} + g_c^{csd}$. $\Delta g_c = g_c^i - g_c^{NORM}$, for $i \in \{LK, LKSW\}$: change in percentage points relative to distribution NORM.

Table D.5: Welfare Effects of Cyclical Idiosyncratic Risk: Lower Rate of Return

CEV	g_c	g_c^{mean}	g_c^{lcd}	g_c^{csd}	Δg_c
			Risk Aversion, $\theta = 1$		
NORM	-1.905	0.24	-2.11	-0.035	0
LK	-2.158	0.274	-2.406	-0.025	-0.253
LKSW	-1.611	0.193	-1.777	-0.027	0.294
			Risk Aversion, $\theta = 2$		
NORM	-3.627	0.436	-3.958	-0.105	0
LK	-3.945	0.445	-4.3	-0.09	-0.318
LKSW	-3.972	0.411	-4.285	-0.098	-0.345
			Risk Aversion, $\theta = 3$		
NORM	-5.123	0.603	-5.554	-0.172	0
LK	-6.786	0.65	-7.26	-0.177	-1.663
LKSW	-8.253	0.711	-8.732	-0.232	-3.13
			Risk Aversion, $\theta = 4$		
NORM	-6.399	0.751	-6.919	-0.231	0
LK	-11.362	0.904	-11.955	-0.311	-4.963
LKSW	-14.283	1.027	-14.863	-0.446	-7.884

Notes: Welfare gains (positive numbers) and losses (negative numbers) with a lower rate of return of $r[\%] = 2\%$ of cyclical idiosyncratic risk expressed as Consumption Equivalent Variation (CEV) in the distribution with normal moments NORM, the leptokurtic distribution LK and the leptokurtic and left-skewed distribution 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 $g_c = g_c^{mean} + g_c^{lcd} + g_c^{csd}$. $\Delta g_c = g_c^i - g_c^{NORM}$, for $i \in \{LK, LKSW\}$: change in percentage points relative to distribution NORM.

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.

Barlevy, G. (2005). The cost of business cycles and the benefits of stabilization. *Economic Perspectives 2005*(Q I), 32–49.

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, August). 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.

Busch, C., D. Domeij, F. Guvenen, and R. Madera (2018). Asymmetric Business-Cycle Risk and Social Insurance. Working Paper Series 1031, Barcelona GSE.

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 (2019). Family and Government Insurance: Wage, Earnings, and Income Risks in the Netherlands and the U.S. Working Paper 25832, NBER.

De Nardi, M., G. Fella, and G. Paz-Pardo (2019, 03). Nonlinear Household Earnings Dynamics, Self-Insurance, and Welfare. *Journal of the European Economic Association*.

Dolmas, J. (1998). Risk Preferences and the Welfare Costs of Business Cycles. *Review of Economic Dynamics 1*, 646–676.

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.

Epaulard, A. and A. Pommeret (2003). Recursive utility, endogenous growth, and the welfare cost of volatility. *Review of Economic Dynamics 6*, 672–684.

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? Working Paper.

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

Hall, R. E. (1988). Intertemporal Substitution in Consumption. *Journal of Political Economy 96*, 339–357.

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

Heathcote, J., F. Perri, and G. Violante (2010). Unequal we stand: an empirical analysis of economic inequality in the united states, 1967-2006. *Review of Economic Dynamics 13*, 15–51.

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.

Kimball, M. and P. Weil (2009). Precautionary Saving and Consumption Smoothing. *Journal of Money, Credit and Banking 41*(2-3), 245–284.

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 (2019). Optimal Taxes on Capital in the OLG Model with Uninsurable Idiosyncratic Income Risk. Working Paper.

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, January). 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, February). 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.

Rothschild, M. and J. E. Stiglitz (1970). Increasing Risk {I}: A Definition. *Journal of Economic Theory 2*, 225–244.

Rothschild, M. and J. E. Stiglitz (1971). Increasing Risk {II}: Its Economic Consequences. *Journal of Economic Theory 3*, 66–84.

Selden, L. (1978). A New Representation of Preferences over "Certain x Uncertain" Consumption Pairs: The "Ordinal Certainty Equivalent" Hypothesis. *Econometrica 46*(5), 1045–1060.

Selden, L. (1979). An OCE Analysis of the Effect of Uncertainty on Saving Under Risk Preference Independence. *The Review of Economic Studies 46*(1), 73–82.

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.

Online Appendix

(Not for Publication)

"Higher-Order Income Risk over the Business Cycle"

(Christopher Busch and Alexander Ludwig)

A Fitting Moments of the FGLD

This online appendix describes how we fit the Flexible Generalized Lambda Distribution (FGLD). The quantile function is

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

where λ_1 is a location and λ_2 is a scale parameter, λ_3, λ_4 in turn are tail index parameters.[1]

We will need to use the relationship between the quantile function and the probability density function (PDF). Noticing that $x = F^{-1}(p) = Q(p)$ and $F(x) = p$ we can derive the PDF $f(x)$ from the quantile function $Q(p)$ by

$$f(x) = f(Q(p)) = \frac{\partial F(x)}{\partial x} = \frac{\partial p}{\partial Q(p)} = \frac{1}{\frac{\partial Q(p)}{\partial p}}. \tag{A.2}$$

Differentiating (A.1) we therefore find the PDF to be

$$f(x) = f(Q(p)) = \frac{\lambda_2}{p^{\lambda_3-1} + (1-p)^{\lambda_4-1}}. \tag{A.3}$$

Lakhany and Mausser (2000) and Su (2007) describe how to estimate the parameters of (A.1) using moments of the distribution. The kth raw moment of a random variable X is given as

$$E\left[X^k\right] = \int_{-\infty}^{\infty} x^k f(x)dx, \quad k \geq 1$$

where $f(x)$ is the distribution function. Setting $k = 1$ gives the expected value $\mu_1 = E[X]$.

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

The kth central moment is defined as

$$E\left[(X-\mu_1)^k\right] = \int_{-\infty}^{\infty} (x-\mu_1)^k f(x)dx, \quad k \geq 1.$$

We can use binomial expansion to write central moments in terms of raw moments as

$$E\left[(X-\mu_1)^k\right] = E\left[\sum_{j=0}^{k} \binom{k}{j} (-1)^j (X)^{k-j} \mu_1^j\right] \tag{A.4}$$

where $\binom{k}{j}$ are binomial coefficients.

Now apply the same logic to evaluate the kth raw moment of a percentile function. Use variable substitution $p = Q^{-1}(p) = F(x)$, noticing that $Q^{-1}(-\infty) = 0$ and $Q^{-1}(\infty) = 1$ so that the integration bounds change. Furthermore, use (A.2) giving $f(x) = \frac{dp}{dQ(p)}$ to rewrite

$$\int_{-\infty}^{\infty} x^k f(x)dx = \int_0^1 Q(p)^k \frac{dp}{dQ(p)} dQ(p) = \int_0^1 Q(p)^k dp. \tag{A.5}$$

Hence the kth raw moment using quantile functions is given by

$$E\left[X^k\right] = \int_0^1 Q(p)^k dp.$$

Next, observe that (A.1) can be rewritten as

$$\begin{aligned} Q(p) = F^{-1}(p) = x &= \lambda_1 - \frac{1}{\lambda_2\lambda_3} + \frac{1}{\lambda_2\lambda_4} + \frac{1}{\lambda_2}\left(\frac{p^{\lambda_3}}{\lambda_3} - \frac{(1-p)^{\lambda_4}}{\lambda_4}\right) \\ &= a + b\tilde{Q}(p). \end{aligned}$$

Let X be the random variable with quantile function $Q(p)$ and let Y be the random variable with quantile function $\tilde{Q}(p)$. We then have

$$\begin{aligned} E[X] &= a + bE[Y], \quad k = 1 \\ E\left[(X - E[X])^k\right] &= b^k E\left[(Y - E[Y])^k\right], \quad k > 1 \end{aligned}$$

for the kth central moments. In what follows, we denote the raw moments of Y by ν, hence $\nu_k = EY^k$. Using (A.4) we thus get for the first four central moments (recalling that $\binom{n}{k} = \frac{n!}{k!(n-k)!}$, with $\binom{n}{n} = \binom{n}{0} = 1$):

$$\begin{aligned}\mu_1 = E[X] &= a + bE[Y] = a + b\nu_1 \\ &= \lambda_1 - \frac{1}{\lambda_2\lambda_3} + \frac{1}{\lambda_2\lambda_4} + \frac{1}{\lambda_2}\nu_1.\end{aligned}$$

For the remaining moments, we rewrite (A.4) to get

$$\begin{aligned}E\left[(Y - E[Y])^k\right] &= E\left[\sum_{j=0}^{k} \binom{k}{j} (-1)^j (Y)^{k-j} \nu(1)^j\right] \\ &= \left[\sum_{j=0}^{k} \binom{k}{j} (-1)^j E\left[(Y)^{k-j}\right] \nu(1)^j\right]\end{aligned}$$

We can therefore write explicitly

$$\begin{aligned}\mu_2 &= b^2 \left(E[Y^2] - (E[Y])^2\right) = \frac{1}{\lambda_2^2}(\nu_2 - \nu_1^2) \\ \mu_3 &= b^3 E\left[\sum_{j=0}^{3} \binom{3}{j} (-1)^j (Y)^{3-j} (\nu_1)^j\right] \\ &= b^3 E\left[Y^3 - 3Y^2\nu_1 + 3Y\nu_1^2 - \nu_1^3\right] \\ &= \frac{1}{\lambda_2^3}\left(\nu_3 - 3\nu_1\nu_2 + 2\nu_1^3\right] \\ \mu_4 &= b^4 E\left[\sum_{j=0}^{4} \binom{4}{j} (-1)^j (Y)^{4-j} (\nu_1)^j\right] \\ &= b^4 E\left[Y^4 - 4Y^3\nu_1 + 6Y^2\nu_1^2 - 4Y\nu_1^3 + \nu_1^4\right] \\ &= \frac{1}{\lambda_2^4}\left(\nu_4 - 4\nu_1\nu_3 + 6\nu_1^2\nu_2 - 3\nu_1^4\right).\end{aligned}$$

Finally, we need to determine expressions for the raw moments of Y. To this end, we have to evaluate

$$E\left[Y^k\right] = \nu_k = \int_0^1 \tilde{Q}(p)^k dp = \int_0^1 \left(\frac{p^{\lambda_3}}{\lambda_3} - \frac{(1-p)^{\lambda_4}}{\lambda_4}\right)^k dp$$

Again using binomial expansion, we can rewrite this integral as

$$
\begin{aligned}
\nu_k &= \int_0^1 \sum_{j=0}^{k} \binom{k}{j} (-1)^j \left(\frac{p^{\lambda_3}}{\lambda_3}\right)^{k-j} - \left(\frac{(1-p)^{\lambda_4}}{\lambda_4}\right)^j dp \\
&= \sum_{j=0}^{k} \binom{k}{j} \frac{(-1)^j}{\lambda_3^{k-j}\lambda_4^j} \int_0^1 \left(p^{\lambda_3(k-j)} - (1-p)^{\lambda_4 j}\right) dp \\
&= \sum_{j=0}^{k} \binom{k}{j} \frac{(-1)^j}{\lambda_3^{k-j}\lambda_4^j} \beta\left(\lambda_3(k-j)+1, \lambda_4 j + 1\right),
\end{aligned}
$$

where $\beta(\cdot,\cdot)$ is the β-function. Observe that the β-function is only well defined if all arguments are positive. This requires that

$$
\lambda_3(k-j)+1>0 \quad \text{and} \quad \lambda_4 j + 1 > 0
$$

for all k, j. This equality can only be binding if $\lambda_3, \lambda_4 < 0$. Since $j \leq k$ we can rewrite the above inequality as

$$
\min(\lambda_3, \lambda_4) > -\frac{1}{k}.
$$

Observe that the RHS in the above is decreasing in k. Therefore, if we target at matching moments up to $k = 4$, the constraint reads as $\min(\lambda_3, \lambda_4) > -\frac{1}{4}$.

We can also write out ν_k, for $k = 1, \dots, 4$ explicitly as functions of λ_3, λ_4 as:

$$
\begin{aligned}
\nu_1 &= \sum_{j=0}^{1} \binom{1}{j} \frac{(-1)^j}{\lambda_3^{1-j}\lambda_4^j} \beta\left(\lambda_3(1-j)+1, \lambda_4 j+1\right) \\
&= \frac{1}{\lambda_3}\beta\left(\lambda_3+1, 1\right) - \frac{1}{\lambda_4}\beta\left(1, \lambda_4+1\right) \\
&= \frac{1}{\lambda_3(\lambda_3+1)} - \frac{1}{\lambda_4(\lambda_4+1)} \\
\nu_2 &= \sum_{j=0}^{2} \binom{2}{j} \frac{(-1)^j}{\lambda_3^{2-j}\lambda_4^j} \beta\left(\lambda_3(2-j)+1, \lambda_4 j+1\right) = \nu_1(\lambda_3, \lambda_4) \\
&= \frac{1}{\lambda_3^2}\beta\left(2\lambda_3+1, 1\right) - 2\frac{1}{\lambda_3\lambda_4}\beta\left(\lambda_3+1, \lambda_4+1\right) + \frac{1}{\lambda_4^2}\beta\left(1, 2\lambda_4+1\right) \\
&= \frac{1}{\lambda_3^2\left(2\lambda_3+1\right)} + \frac{1}{\lambda_4^2\left(2\lambda_4+1\right)} - 2\frac{1}{\lambda_3\lambda_4}\beta\left(\lambda_3+1, \lambda_4+1\right) = \nu_2(\lambda_3, \lambda_4) \\
\nu_3 &= \sum_{j=0}^{3} \binom{3}{j} \frac{(-1)^j}{\lambda_3^{3-j}\lambda_4^j} \beta\left(\lambda_3(3-j)+1, \lambda_4 j+1\right) \\
&= \frac{1}{\lambda_3^3}\beta\left(3\lambda_3+1, 1\right) - \frac{3}{\lambda_3^2\lambda_4}\beta\left(2\lambda_3+1, \lambda_4+1\right) + \frac{3}{\lambda_3\lambda_4^2}\beta\left(\lambda_3+1, 2\lambda_4+1\right) - \frac{1}{\lambda_4^3}\beta\left(1, 3\lambda_4+1\right) \\
&= \frac{1}{\lambda_3^3(3\lambda_3+1)} - \frac{1}{\lambda_4^3(3\lambda_4+1)} - \frac{3}{\lambda_3^2\lambda_4}\beta\left(2\lambda_3+1, \lambda_4+1\right) + \frac{3}{\lambda_3\lambda_4^2}\beta\left(\lambda_3+1, 2\lambda_4+1\right) = \nu_3(\lambda_3, \lambda_4) \\
\nu_4 &= \sum_{j=0}^{4} \binom{4}{j} \frac{(-1)^j}{\lambda_3^{4-j}\lambda_4^j} \beta\left(\lambda_3(4-j)+1, \lambda_4 j+1\right) \\
&= \frac{1}{\lambda_3^4}\beta\left(4\lambda_3+1, 1\right) - \frac{4}{\lambda_3^3\lambda_4}\beta\left(3\lambda_3+1, 2\lambda_4+1\right) + \frac{6}{\lambda_3^2\lambda_4^2}\beta\left(2\lambda_3+1, 2\lambda_4+1\right) - \frac{4}{\lambda_3\lambda_4^3}\beta\left(\lambda_3+1, 3\lambda_4+1\right) + \\
&\quad \frac{1}{\lambda_4^4}\beta\left(1, 4\lambda_4+1\right) \\
&= \frac{1}{\lambda_3^4(4\lambda_3+1)} + \frac{1}{\lambda_4^4(4\lambda_4+1)} - \frac{4}{\lambda_3^3\lambda_4}\beta\left(3\lambda_3+1, 2\lambda_4+1\right) - \frac{4}{\lambda_3\lambda_4^3}\beta\left(\lambda_3+1, 3\lambda_4+1\right) + \\
&\quad \frac{6}{\lambda_3^2\lambda_4^2}\beta\left(2\lambda_3+1, 2\lambda_4+1\right) = \nu_4(\lambda_3, \lambda_4).
\end{aligned}
$$

From the above observe that the third and fourth central moments μ_3, μ_4 of random variable X are only functions of λ_3, λ_4. Therefore, the procedure is to determine λ_3, λ_4 jointly to target μ_3, μ_4 under the parameter restriction $\min(\lambda_3, \lambda_4) > -\frac{1}{4}$. Next, we can successively determine λ_2 from targeting μ_2 and, finally, λ_1 by targeting μ_1.

B A Numerical Example of the Two-Period Model

In this online appendix, we present a quantitative illustration of the two-period model in order to show that higher-order income risk (in logs) may indeed lead to lower precautionary savings and utility gains. Specifically, we consider three different parameterizations of discrete PDFs $\Psi(\varepsilon)$ based on Proposition B.1: NORM is a symmetric distribution with a kurtosis of $\alpha_4 = 3$ as for a normal distribution. Distribution LK is also symmetric but strongly leptokurtic with a kurtosis of $\alpha_4 = 30$, and distribution LKSW additionally introduces left-skewness of $\alpha_3 = -5$. For all distributions we set the variance $\mu_2^\varepsilon = 0.5$. Throughout we normalize such that $\mathbb{E}[\exp(\varepsilon)] = 1$. To investigate the role of higher-order risk attitudes we consider two parametrizations with $\theta \in \{1, 4\}$. Throughout, we set the IES γ equal to 1, thus we focus on risk sensitive preferences.

B.1 Shocks

The shock ε in this two-period model is taken to be discrete. Specifically, we consider a simple lottery such that $\varepsilon \in \{\varepsilon_l, \varepsilon_0, \varepsilon_h\}$ with $\varepsilon_l < \varepsilon_0 < \varepsilon_h$ and respective probabilities $\{(1-p) \cdot q, p, (1-p) \cdot (1-q)\}$. This simple structure enables us to derive a parametrization with a closed form representation for the variance, skewness and kurtosis of the shock process, as stated in the following proposition:[2]

Proposition B.0 Let $\varepsilon \in \{\varepsilon_l, \varepsilon_0, \varepsilon_h\}$, drawn with respective probabilities $\{(1-p) \cdot q, p, (1-p) \cdot (1-q)\}$. Then, if and only if $\alpha_4 > 1$ and, for $\alpha_3 \neq 0$ in addition

1. either $\alpha_3 \in (0, \sqrt{\alpha_4 - 1})$

2. or $\alpha_3 \in (-\sqrt{\alpha_4 - 1}, 0)$,

[2] Our approach extends Ebert (2015), who analyzes skewness using a two-point distribution, to the fourth moment.

we match $\mu_2, \alpha_3, \alpha_4$, with the normalization $E[\exp(\varepsilon)] = 1$ by choosing

$$q = \frac{1}{2} \begin{cases} +\frac{1}{2}\sqrt{1 - \frac{4\frac{\alpha_4}{\alpha_3^2} - 4}{4\frac{\alpha_4}{\alpha_3^2} - 3}} & \text{if } \alpha_3 > 0 \\ -\frac{1}{2}\sqrt{1 - \frac{4\frac{\alpha_4}{\alpha_3^2} - 4}{4\frac{\alpha_4}{\alpha_3^2} - 3}} & \text{if } \alpha_3 < 0 \\ 0.5 & \text{if } \alpha_3 = 0 \end{cases}$$

$$p = \begin{cases} 1 - \frac{(2q-1)^2}{q(1-q)\alpha_3^2} & \text{if } \alpha_3 \neq 0 \\ 1 - \frac{1}{\alpha_4} & \text{if } \alpha_3 = 0 \end{cases}$$

$$\Delta_\varepsilon = \begin{cases} \frac{\sqrt{\mu_2}\alpha_3}{2q-1} & \text{if } \alpha_3 \neq 0 \\ 2\sqrt{\mu_2}\sqrt{\alpha_4} & \text{if } \alpha_3 = 0, \end{cases}$$

and

$$\begin{aligned} \varepsilon_l &= -\ln\left[p \exp\left((1-q)\Delta_\varepsilon\right) + (1-p)\left(q + (1-q)\exp(\Delta_\varepsilon)\right)\right] \\ \varepsilon_0 &= \varepsilon_l + (1-q)\Delta_\varepsilon \\ \varepsilon_h &= \varepsilon_l + \Delta_\varepsilon. \end{aligned}$$

Proof. See Section B.5. □

This representation of risk is useful because it enables us to transparently illustrate how higher-order income risk affects the distribution using a very simple structure with a closed-form solution from payoffs to the respective moments of higher-order income risk.

The upper part of Table B.1 summarizes the moments for the calibration of ε for these three distributions. The lower part shows how this translates into respective moments in level of the innovation, $\exp(\varepsilon)$. Going from distribution NORM to distribution LK we observe that not only the kurtosis increases strongly but also the variance. Simultaneously, the distribution becomes more skewed to the right. Thus, whether the higher kurtosis of the innovation ε also leads to welfare losses (or a strong increase in precautionary savings) depends on whether the effects on the variance and kurtosis dominate those on the skewness, cf. equations (4) and (6).

In turn, going from distribution NORM to distribution LKSW we observe that the distribution is now more skewed to the left and features a higher kurtosis. However, at the same time, the variance goes down quite strongly. Thus, whether the simultaneously higher kurtosis and lower skewness (or: increased left-skewness) of the innovation ε relative to distribution NORM lead to welfare losses (or a strong increase in precautionary savings) depends on whether the effects on the skewness and kurtosis dominate those on the variance, again see equations (4) and (6).

Table B.1: 2-Period Model: Shocks, standardized moments

Moments of Innovation in Logs, ε			
	μ_2^{ε}	α_3^{ε}	α_4^{ε}
NORM	0.5	0	3
LK	0.5	0	30
LKSW	0.5	-5	30
Moments of Innovation in Levels, $\exp(\varepsilon)$			
	$\mu_2^{\exp(\varepsilon)}$	$\alpha_3^{\exp(\varepsilon)}$	$\alpha_4^{\exp(\varepsilon)}$
NORM	0.5868	1.4885	3.7882
LK	11.6316	7.5458	57.9669
LKSW	0.1039	0.5684	4.8371

Notes: Standardized moments of the discrete shock distribution.

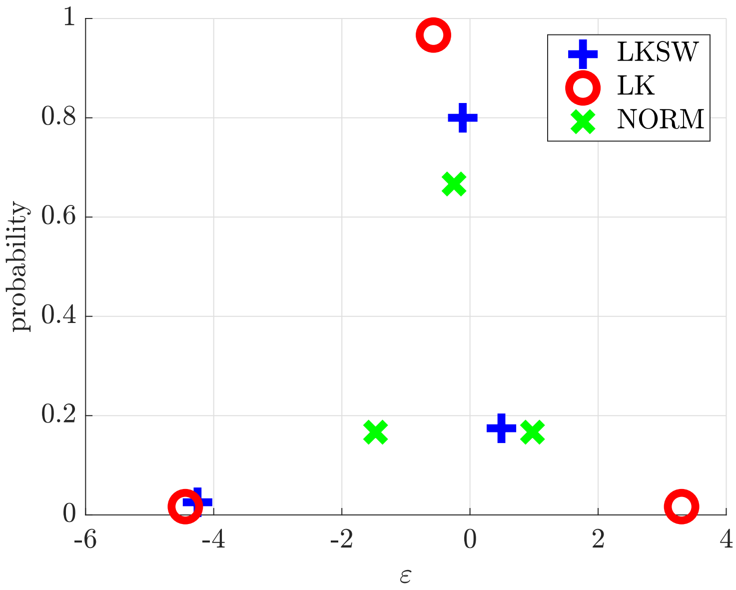
Table B.2: 2-Period Model: Shocks, central moments

Moments of Innovation in Logs, ε			
	μ_2^{ε}	μ_3^{ε}	μ_4^{ε}
NORM	0.5	0	0.75
LK	0.5	0	7.5
LKSW	0.5	-1.7678	7.5
Moments of Innovation in Levels, $\exp(\varepsilon)$			
	$\mu_2^{\exp(\varepsilon)}$	$\mu_3^{\exp(\varepsilon)}$	$\mu_4^{\exp(\varepsilon)}$
NORM	0.5868	0.6691	1.3045
LK	11.6316	299.3406	7842.5727
LKSW	0.1039	0.0190	0.0523

Notes: Central moments of the discrete shock distribution.

Figure B.1 plots the the corresponding PDFs $\Psi(\varepsilon)$. Relative to NORM, the distribution LK leads to a fanning out of the shocks. As can be seen for the realization of $\exp(\varepsilon_0)$ this induces a shift of the shock realizations to the left such that $\mathbb{E}[\varepsilon]$ is reduced from -0.24 to -0.57. Moving from distribution LK to distribution LKSW by additionally introducing skewness shifts the probability mass to the left tail such that $\mathbb{E}[\varepsilon]$ increases to -0.11. From this observation we know from Proposition 1 that with logarithmic utility ($\theta = 1$), we have welfare losses from the symmetric and leptokurtic distribution LK and welfare gains for the additionally left skewed distribution LKSW if households do not have access to a savings technology.

Figure B.1: Distribution of ε



Notes: Distribution function of the discrete shock with three points as in Proposition B.1 under the three scenarios NORM, LK, and LKSW.

B.2 Allocations

Table B.3 reports results on allocations, assuming that households have access to a savings technology. Increasing risk attitude coefficient θ leads to more precautionary savings and reduces the differences in precautionary savings across scenarios. Holding θ constant, compared to the distribution NORM we observe more precautionary savings for distribution LK and thus the effects of increased variance and kurtosis dominate the effects of higher skewness. In contrast, with θ constant we observe less precautionary savings for distribution LKSW and thus the effects of the lower variance dominate the effects of higher kurtosis and left-skewness.

Table B.4 displays the welfare consequence if there is no access to a savings technology under binding constraint (3) in column NST and with access in column ST. First, with $\theta = 1$, the distribution LKSW leads to utility gains. Thus, for our shock parametrization, the positive welfare effects of lower skewness dominate the losses of an increased kurtosis. This is true for both scenarios NST, cf. Proposition 1, as well as for scenario ST. Second, under NST utility consequences are strongly increasing in θ, as we learned from equation (4). Third, both gains and losses decrease in scenario ST compared to scenario NST. The rea-

Table B.3: Results from 2-Period Model: Allocations

	c_0	$\mathbb{E}[c_1]$	a_1
Risk Aversion, $\theta = 1$			
NORM	0.837	1.162	0.162
LK	0.773	1.226	0.226
LKSW	0.895	1.104	0.104
Risk Aversion, $\theta = 4$			
NORM	0.671	1.328	0.328
LK	0.662	1.337	0.337
LKSW	0.614	1.385	0.385

Notes: Allocations in the two-period model.

son is the precautionary savings response, which reduces utility losses from risk in both the denominator and the numerator of the CEV calculation. Fourth, as a consequence of the precautionary savings response, absolute values of the CEV are lower with higher risk aversion in scenario ST. This shows that the utility consequences of higher-order risk, expressed in terms of CEVs, may be non-monotonic in the degree of risk aversion.

Table B.4: Results from 2-Period Model: CEV

	NST	ST
Risk Aversion, $\theta = 1$		
LK	-14.82%	-11.75%
LKSW	7.03%	6.76%
Risk Aversion, $\theta = 4$		
LK	-66.20%	-3.22%
LKSW	-65.35%	5.66%

Notes: CEV relative to NORM. NST: no access to savings technology. ST: assess to savings technology.

B.3 Decomposition of Consumption Equivalent Variations

Table B.5 reports the results for the decomposition of the CEV, for sake of brevity only for $\theta = 1$ and with access to a savings technology (ST). With this calibration, most of the changes appear in the cross-sectional distribution effect.

Table B.5: Results from 2-Period Model: Decomposition of CEV for Log Utility

CEV	g_c	g_c^{mean}	g_c^{lcd}	g_c^{csd}
Baseline				
LK	-11.75%	0	-2.35%	-9.40%
LKSW	6.76%	0	2.16%	4.59%
Impatience				
LK	-11.04%	0	-9.82%	-1.22%
LKSW	-4.10%	0	-10.50%	6.40%
Positive Interest Rate				
LK	-5.56%	2.65%	-4.92%	-3.30%
LKSW	1.70%	2.63%	-4.85%	3.93%
Borrowing Constraint				
LK	-5.04%	0.34%	-0.65%	-4.73%
LKSW	2.26%	0.13%	-0.26%	2.38%

Notes: CEV relative to NORM for $\theta = 1, \rho = 1$ for scenario ST. LK: leptokurtik distribution, LKSW: leptokurtik and skewed distribution.

B.4 Additional Model Elements

For the remaining exercises we add step by step model elements included in the quantitative model. Throughout, we take $\theta = \frac{1}{\rho} = 1$ and only analyze the welfare consequences in terms of the consumption equivalent variation. Results are contained in the remaining rows of Table B.5 .

Impatience. We first add a period discount factor β of 0.96, such that the discount factor accounting for the 40-year periodicity is $0.96^{40} \approx 0.19$. This introduces a life-cycle savings motive into the model and preferences now write as (for $\rho \neq 1$)

$$U = \frac{1}{1-\rho}\left((1-\tilde{\beta})c_0^{1-\frac{1}{\rho}} + \tilde{\beta}v\left(c_1, \theta, \Psi\right)^{1-\frac{1}{\rho}}\right),$$

where $\tilde{\beta} = \frac{\beta}{1+\beta}$ and β is the raw time discount factor. As a consequence of discounting, the life-cycle distribution effect becomes more potent. Households now take on debt to finance consumption when young. Given the riskiness of second period consumption, borrowing is much lower in distributions LK and LKSW than in distribution NORM. Therefore, the life-cycle distribution effect is strongly negative.

Positive Returns. Next, we also assume a positive interest rate on savings with an annual raw interest rate of 2%. Given the length of each model period of 40 real life years, this corresponds to $R = 1.02^{40} \approx 2.2$. Thus, the budget constraints now write as

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

Table B.5 shows that now the mean effect is non-zero. The reason is that savings are intertemporally shifted at a non-zero rate so that average consumption increases. Results also show that the aforementioned life-cycle effects are muted. Still the life-cycle distribution effects are negative.

Borrowing Constraints. Next, we add occasionally binding borrowing constraints at zero borrowing, i.e., we add the constraint

$$a_1 \geq 0.$$

For the chosen parametrization this constraint turns out to be binding only in scenario NORM. Since households are thus worse off in NORM relative to the other scenarios, welfare losses in distribution LK decrease and gains in distribution LKSW increase.

Throughout all these scenarios, we observe that the cross-sectional distribution effect is negative in scenario LK, and positive in scenario LKSW.

B.5 Proof of Proposition B.1

Proof. Take $\varepsilon_0 = \mu_1$, thus

$$\begin{aligned} \mu_1 &= p\varepsilon_0 + (1-p)\left(q\varepsilon_l + (1-q)\varepsilon_h\right) \\ &= p\mu_1 + (1-p)\left(q\varepsilon_l + (1-q)\varepsilon_h\right) \\ \Leftrightarrow \quad \mu_1 &= q\varepsilon_l + (1-q)\varepsilon_h. \end{aligned}$$

Now, let $\varepsilon_h = \varepsilon_l + \Delta_\varepsilon$ to get

$$\begin{aligned} \mu_1 &= q\varepsilon_l + (1-q)\left(\varepsilon_l + \Delta_\varepsilon\right) \\ &= \varepsilon_l + (1-q)\Delta_\varepsilon. \end{aligned}$$

For the variance we get

$$\begin{aligned}\mu_2 &= (1-p)\left(q(\varepsilon_l-\mu_1)^2+(1-q)(\varepsilon_h-\mu_1)^2\right)\\&= (1-p)\left(q\left(\varepsilon_l-(\varepsilon_l+(1-q)\Delta_\varepsilon)\right)^2+(1-q)\left(\varepsilon_h-(\varepsilon_l+(1-q)\Delta_\varepsilon)\right)^2\right)\\&= (1-p)\left(q(1-q)^2+(1-q)q^2\right)\Delta_\varepsilon^2\\&= (1-p)q(1-q)\Delta_\varepsilon^2.\end{aligned}$$

For the third central moment μ_3 we get

$$\begin{aligned}\mu_3 &= (1-p)\left(q(\varepsilon_l-\mu_1)^3+(1-q)(\varepsilon_h-\mu_1)^3\right)\\&= (1-p)\left(q\left(\varepsilon_l-(\varepsilon_l+(1-q)\Delta_\varepsilon)\right)^3+(1-q)\left(\varepsilon_h-(\varepsilon_l+(1-q)\Delta_\varepsilon)\right)^3\right)\\&= (1-p)\left(-q(1-q)^3+(1-q)q^3\right)\Delta_\varepsilon^3\\&= (1-p)q(1-q)\left(-(1-q)^2+q^2\right)\Delta_\varepsilon^3\\&= (1-p)q(1-q)(2q-1)\Delta_\varepsilon^3\end{aligned}$$

and we can thus write the skewness α_3 as

$$\alpha_3 = \frac{\mu_3}{\sqrt{\mu_2}^3} = \frac{2q-1}{\sqrt{(1-p)q(1-q)}}.$$

For the fourth central moment μ_4 we get

$$\begin{aligned}\mu_4 &= (1-p)\left(q(\varepsilon_l-\mu_1)^4+(1-q)(\varepsilon_h-\mu_1)^4\right)\\&= (1-p)\left(q\left(\varepsilon_l-(\varepsilon_l+(1-q)\Delta_\varepsilon)\right)^4+(1-q)\left(\varepsilon_h-(\varepsilon_l+(1-q)\Delta_\varepsilon)\right)^4\right)\\&= (1-p)\left(q(1-q)^4+(1-q)q^4\right)\Delta_\varepsilon^4\\&= (1-p)q(1-q)\left((1-q)^3+q^3\right)\Delta_\varepsilon^4\\&= (1-p)q(1-q)\left((1-2q+q^2)(1-q)+q^3\right)\Delta_\varepsilon^4\\&= (1-p)q(1-q)\left(1-3q+3q^2\right)\Delta_\varepsilon^4\end{aligned}$$

and can therefore write the kurtosis as

$$\alpha_4 = \frac{\mu_4}{\mu_2^2} = \frac{3q^2-3q+1}{(1-p)q(1-q)}.$$

To summarize, the terms we seek to match are

$$\mu_2 = (1-p)q(1-q)\Delta_\varepsilon^2, \tag{B.6a}$$
$$\alpha_3 = \frac{2q-1}{\sqrt{(1-p)q(1-q)}}, \tag{B.6b}$$
$$\alpha_4 = \frac{3q^2-3q+1}{(1-p)q(1-q)}. \tag{B.6c}$$

To obtain $\alpha_4 > 0$ we require $p \in (0,1)$, $q \in (0,1)$ and

$$\begin{aligned} & q^2 - q + \frac{1}{3} > 0 \\ \Leftrightarrow \quad & \left(q - \frac{1}{2}\right)^2 > -\frac{1}{12} \end{aligned}$$

which always holds.

Let us next characterize the solution according to the following case distinction:

1. $\alpha_3 = 0$. Then we obviously have $q = 1 - q = 0.5$. We can accordingly rewrite (B.6a) and (B.6c) as

$$\begin{aligned} \mu_2 &= (1-p)\frac{1}{4}\Delta_\varepsilon^2, \\ \alpha_4 &= \frac{1}{(1-p)}, \end{aligned}$$

and therefore

$$\begin{aligned} q &= \frac{1}{2} \\ p &= 1 - \frac{1}{\alpha_4} \\ \Delta_\varepsilon &= 2\sqrt{\mu_2}\sqrt{\alpha_4} \end{aligned}$$

characterizes the solution. Notice that $\alpha_4 > 0$ and thus $p < 1$. To get $p > 0$ we require

$$1 - \frac{1}{\alpha_4} > 0 \qquad \Leftrightarrow \qquad \alpha_4 > 1.$$

2. $\alpha_3 \neq 0$. From (B.6a) we get

$$(1-p)q(1-q) = \frac{\mu_2}{\Delta_\varepsilon^2}$$

Using this in (B.6b) and (B.6c) we get

$$\alpha_3 = \frac{(2q-1)\Delta_\varepsilon}{\sqrt{\mu_2}}, \tag{B.7a}$$

$$\alpha_4 = \frac{(3q^2-3q+1)\Delta_\varepsilon^2}{\mu_2}. \tag{B.7b}$$

Now use (B.7a) in (B.7b) to get

$$\begin{aligned}
& \frac{(3q^2-3q+1)}{(2q-1)^2} = \frac{\alpha_4}{\alpha_3^2} \\
\Leftrightarrow \quad & (3q^2-3q+1) = \frac{\alpha_4}{\alpha_3^2}\left(4q^2-4q+1\right) \\
\Leftrightarrow \quad & q^2\left(4\frac{\alpha_4}{\alpha_3^2}-3\right) - q\left(4\frac{\alpha_4}{\alpha_3^2}-3\right) + \frac{\alpha_4}{\alpha_3^2} - 1 = 0 \\
\Leftrightarrow \quad & q^2 - q + \frac{\frac{\alpha_4}{\alpha_3^2}-1}{4\frac{\alpha_4}{\alpha_3^2}-3} = 0
\end{aligned}$$

and thus

$$q_\pm = \frac{1}{2} \pm \frac{1}{2}\sqrt{\underbrace{1 - \frac{4\frac{\alpha_4}{\alpha_3^2}-4}{4\frac{\alpha_4}{\alpha_3^2}-3}}_{=\Psi}} \tag{B.8}$$

Thus, the first restriction for $q_\pm \in (0,1)$ is that $\Psi > 0$. Consider the following case distinction:

(a) $4\frac{\alpha_4}{\alpha_3^2} - 3 > 0 \Leftrightarrow \frac{\alpha_4}{\alpha_3^2} > \frac{3}{4}$: Then

$$\begin{aligned}
& 1 - \frac{4\frac{\alpha_4}{\alpha_3^2}-4}{4\frac{\alpha_4}{\alpha_3^2}-3} > 0 \\
\Leftrightarrow \quad & 4\frac{\alpha_4}{\alpha_3^2} - 3 > 4\frac{\alpha_4}{\alpha_3^2} - 4 \\
\Leftrightarrow \quad & 4 > 3
\end{aligned}$$

and thus for $\frac{\alpha_4}{\alpha_3^2} > \frac{3}{4}$ we get $\Psi > 0$.

(b) $4\frac{\alpha_4}{\alpha_3^2} - 3 < 0 \Leftrightarrow \frac{\alpha_4}{\alpha_3^2} < \frac{3}{4}$ then we obviously get a contradiction.

Thus, we require $\alpha_4 > \frac{3}{4}\alpha_3^2$.

Next, for both the positive and the negative root, we further require $\Psi < 1$. Again investigate the case $\alpha_4 > \frac{3}{4}\alpha_3^2$. We get

$$
\begin{aligned}
& 1 - \frac{4\frac{\alpha_4}{\alpha_3^2} - 4}{4\frac{\alpha_4}{\alpha_3^2} - 3} < 1 \\
\Leftrightarrow \quad & \frac{4\frac{\alpha_4}{\alpha_3^2} - 4}{4\frac{\alpha_4}{\alpha_3^2} - 3} > 0 \\
\Leftrightarrow \quad & 4\frac{\alpha_4}{\alpha_3^2} - 4 > 0 \\
\Leftrightarrow \quad & \frac{\alpha_4}{\alpha_3^2} > 1
\end{aligned}
$$

and thus a necessary and sufficient condition for $q_\pm \in (0,1)$ is:

$$\alpha_4 > \alpha_3^2. \tag{B.9}$$

Since $\alpha_3 = \frac{(2q-1)\Delta_\varepsilon}{\sqrt{\mu_2}}$ and since $\Delta_\varepsilon > 0$ (by construction) and $\sqrt{\mu_2} > 0$ we choose the positive root $q^\star = q_+$ for a right-skewed distribution with $\alpha_3 > 0$ and the negative root $q^\star = q_-$ to model a left-skewed with $\alpha_3 < 0$.

We next get from (B.7a) that

$$\Delta_\varepsilon = \frac{\sqrt{\mu_2}\alpha_3}{2q^\star - 1}$$

and from (B.6a) that

$$p = 1 - \frac{\mu_2}{q^\star(1-q^\star)\Delta_\varepsilon^2} = 1 - \frac{(2q^\star - 1)^2}{q^\star(1-q^\star)\alpha_3^2}. \tag{B.10}$$

We have already established that under condition (B.9) $q^\star \in (0,1)$. Next, we need to establish conditions such that $p \in (0,1)$. From (B.10) we observe that $q^\star \in (0,1)$ gives $p < 1$. Also observe that $p > 0$ is equivalent to

$$\alpha_3^2 > \frac{(2q^\star - 1)^2}{q^\star(1-q^\star)} \tag{B.11}$$

(a) Case $\alpha_3 < 0$: Recall that for this case we take the negative root $q_-^\star$, where

$$q_-^\star = \frac{1}{2} - \frac{1}{2}\sqrt{\Psi} > 0.$$

for $\Psi \in (0,1)$ iff $\alpha_4 > \alpha_3^2$. Thus the case $\alpha_3 < 0$ implies that $\alpha_3 > -\sqrt{\alpha_4}$. Next observe that

$$(2q^\star - 1)^2 = (1 - \sqrt{\Psi} - 1)^2 = \Psi$$

and

$$\begin{aligned} q^\star(1-q^\star) &= \left(\frac{1}{2} - \frac{1}{2}\sqrt{\Psi}\right)\left(\frac{1}{2} + \frac{1}{2}\sqrt{\Psi}\right) \\ &= \frac{1}{4} - \frac{1}{4}\Psi = \frac{1}{4}(1-\Psi). \end{aligned}$$

Thus condition (B.11) can be rewritten as

$$\begin{aligned} & \alpha_3^2 > \frac{(2q^\star-1)^2}{q^\star(1-q^\star)} = \frac{4\Psi}{1-\Psi} \\ \Leftrightarrow \quad & \alpha_3^2(1-\Psi) > 4\Psi \\ \Leftrightarrow \quad & \alpha_3^2 \frac{\frac{4\alpha_4}{\alpha_3^2} - 4}{4\frac{\alpha_4}{\alpha_3^2} - 3} > 4\left(1 - \frac{\frac{4\alpha_4}{\alpha_3^2} - 4}{4\frac{\alpha_4}{\alpha_3^2} - 3}\right) \\ \Leftrightarrow \quad & \alpha_3^2\left(\frac{\alpha_4}{\alpha_3^2} - 1\right) > 4\frac{\alpha_4}{\alpha_3^2} - 3 - \left(4\frac{\alpha_4}{\alpha_3^2} - 4\right) \\ \Leftrightarrow \quad & \alpha_4 - \alpha_3^2 > 1 \\ \Leftrightarrow \quad & \alpha_3 > -\sqrt{\alpha_4 - 1}, \text{ since } \alpha_3 < 0 \end{aligned}$$

which also implies that we require $\alpha_4 > 1$. Since $-\sqrt{\alpha_4 - 1} > -\sqrt{\alpha_4}$ we thus obtain as a necessary and sufficient condition for the case $\alpha_3 < 0$

$$\alpha_4 > 1 \text{ and } \alpha_3 > -\sqrt{\alpha_4 - 1} \tag{B.12}$$

to get $q \in (0, \frac{1}{2})$, $p \in (0,1)$ and $\Delta\epsilon > 0$.

(b) Case $\alpha_3 > 0$: Recall that for this case we take the positive root $q_+^\star$ where

$$q_+^\star = \frac{1}{2} + \frac{1}{2}\sqrt{\Psi} > 0.$$

for $\Psi \in (0,1)$ iff $\alpha_4 > \alpha_3^2$ and thus $\alpha_3 < \sqrt{\alpha_4}$. Thus

$$(2q^\star - 1)^2 = \Psi$$

and

$$
\begin{aligned}
q^\star(1-q^\star) &= \left(\frac{1}{2}+\frac{1}{2}\sqrt{\Psi}\right)\left(\frac{1}{2}-\frac{1}{2}\sqrt{\Psi}\right) \\
&= \frac{1}{4}-\frac{1}{4}\Psi = \frac{1}{4}\left(1-\Psi\right).
\end{aligned}
$$

and following the steps above we thus get

$$
\begin{aligned}
& \alpha_4 - \alpha_3^2 > 1 \\
\Leftrightarrow \qquad & \alpha_3 < \sqrt{\alpha_4 - 1},
\end{aligned}
$$

Since $\sqrt{\alpha_4 - 1} < \sqrt{\alpha_4}$ we thus obtain as a necessary and sufficient condition for the case $\alpha_3 > 0$

$$
\alpha_4 > 1 \;\text{ and }\; \alpha_3 < \sqrt{\alpha_4 - 1} \tag{B.13}
$$

to get $q \in (\frac{1}{2}, 1)$, $p \in (0, 1)$ and $\Delta\epsilon > 0$.

Finally, for ε_l given, the mean of the exponent of the random variable x is given by

$$
\begin{aligned}
E\left[\exp(x)\right] &= p\exp\left(\varepsilon_l + (1-q)\Delta_\varepsilon\right) + (1-p)\left(q\exp\left(\varepsilon_l\right) + (1-q)\exp\left(\varepsilon_l + \Delta_\varepsilon\right)\right) \\
&= \exp(\varepsilon_l)\left[p\exp\left((1-q)\Delta_\varepsilon\right) + (1-p)\left(q + (1-q)\exp\left(\Delta_\varepsilon\right)\right)\right].
\end{aligned}
$$

Normalizing $E\left[\exp(x)\right] = 1$ we thus get

$$
\varepsilon_l = -\ln\left[p\exp\left((1-q)\Delta_\varepsilon\right) + (1-p)\left(q + (1-q)\exp\left(\Delta_\varepsilon\right)\right)\right].
$$

$\square$

References

Ebert, S. (2015). On Skewed Risks in Economic Models and Experiments. *Journal of Economic Behavior & Organization 112*(C), 85–97.

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

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

//
IMPRINT

ZEW – Leibniz-Zentrum für Europäische Wirtschaftsforschung GmbH Mannheim

ZEW – Leibniz Centre for European Economic Research

L 7,1 · 68161 Mannheim · Germany
Phone +49 621 1235-01
info@zew.de · zew.de