Appendix

"Higher-Order Income Risk Over the Business Cycle"

(Christopher Busch and Alexander Ludwig)

A Analytical Appendix

A.1 Derivation of Equation (1)

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

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

Under a binding budget constraint 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 (1).

A.2 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, by Jensen's inequality for convex functions the mean of the distribution in levels would be lower, so to preserve the mean 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 fanning out of the support of shocks in logs increases the mean of the distribution in levels by Jensen's inequality for convex functions. In order to preserve the mean the distribution needs to be shifted down, which reduces the mean in logs. Since with log utility and in absence of a savings technology, expected life-time utility is $U = \ln(y_0) + \mathbb{E}[\ln(y_1)]$, solely the mean of the distribution in logs matters for life-time utility and thus a mean-preserving *reduction of skewness* leads to *utility gains*. Likewise, a mean-preserving *increase of variance or kurtosis* leads to *utility losses* in expectation. This is summarized in the following

Proposition 1. *Suppose that the utility function is logarithmic ($\theta = 1$) and that there is no savings technology ($a_1 = 0$). Then a mean-preserving* reduction of skewness *('more negative skewness') leads to* utility gains, *whereas a mean-preserving* increase of variance or kurtosis *leads to* utility losses *in expectation.*

Proof. Let $\mu_1^\varepsilon = E_\Psi[\varepsilon] = \int \varepsilon d\Psi$, $\mu_i^\varepsilon = \int (\varepsilon - \mu_1^\varepsilon)^i d\Psi$ for $i > 1$, and let $E_\Psi[\exp(\varepsilon)] = \int \exp(\varepsilon) d\Psi = 1$. Denote by $\tilde{\Psi}^{\delta_i}(\varepsilon)$ a mean preserving (constant μ_1^ε) distribution function that is obtained from $\Psi(\varepsilon)$ by changing central moment μ_i^ε holding all other moments $\mu_{\neg i}^\varepsilon$ for $i > 1$ constant. Also, define the random variable $\tilde{\varepsilon}^{\delta_i} = \varepsilon + \Delta^{\delta_i}$, which is obtained from ε by shifting all realizations by the constant Δ^{δ_i}. Let the normalization $E_{\tilde{\Psi}^{\delta_i}}[\exp(\tilde{\varepsilon}^{\delta_i})] = E_{\tilde{\Psi}^{\delta_i}}[\exp(\varepsilon + \Delta^{\delta_i})] = \int \exp(\varepsilon + \Delta^{\delta_i}) d\tilde{\Psi}^{\delta_i} = \exp(\Delta^{\delta_i}) \int \exp(\varepsilon) d\tilde{\Psi}^i = 1$ define the shift parameter $\Delta^{\delta_i} = -\ln\left(\int \exp(\varepsilon) d\tilde{\Psi}^i\right)$. Finally, observe that $E_{\tilde{\Psi}^{\delta_i}}[\varepsilon + \Delta^{\delta_i}] = E_{\tilde{\Psi}}[\varepsilon] + \Delta^{\delta_i} = E_\Psi[\varepsilon] + \Delta^{\delta_i}$ since μ_1^ε is held constant. With logarithmic utility and binding budget constraint, 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 $\mu_{\neg i}^\varepsilon$ for $i > 1$ increases $\int \exp(\varepsilon) d\tilde{\Psi}^i$ 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 $\mu_{\neg i}^\varepsilon$ for $i > 1$ decreases $\int \exp(\varepsilon) d\tilde{\Psi}^i$ below one which follows from Jensen's inequality for convex functions. Thus $\Delta^{\delta_i} > 0$.

□

While the finding in Proposition 1 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.[1] In Supplementary Appendix S.B we provide a numerical illustration by considering a discrete three-point distribution. We show how

[1] 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.

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.

A.3 Derivation of Equation (2)

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

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

A.4 Precautionary Savings

Rewrite the first-order condition, equation (2), as an implicit function

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

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

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

Note that since $\mu_2^{\exp(\varepsilon)} > 0, \mu_3^{\exp(\varepsilon)} < 0, \mu_4^{\exp(\varepsilon)} > 0$ 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 $\frac{\partial e(\cdot)}{\partial \mu_i^{\exp(\varepsilon)}} < 0$ for $i = 2, 4$ and $\frac{\partial e(\cdot)}{\partial \mu_3^{\exp(\varepsilon)}} > 0$. Thus, $\frac{da_1}{d\mu_i^{\exp(\varepsilon)}} > 0$ for $i = 2, 4$ and $\frac{da_1}{d\mu_3^{\exp(\varepsilon)}} < 0$.

A.5 Extension of Two-Period Model: Recursive Preferences

Our results from the two-period analysis readily extend to a recursive preference specification. Of course, in the two-period model the notion of *recursive preferences* is not strictly speaking correct. We use this terminology here as we adopt Epstein-Zin-Weil preferences a la Epstein and Zin (1989, 1991), and Weil (1989) in the main analysis based on the quantitative life-cycle model. In the two-period context, with the alternative utility specification we can disentangle risk attitudes as parameterized by θ from the inter-temporal elasticity of substitution as parameterized by γ:[2]

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

Thus, γ is 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

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

The specification of preferences gives standard CRRA preferences considered in the main text if the measure of the IES γ and the measure of risk aversion θ are reciprocals: $\theta = \frac{1}{\gamma}$.

Hand-to-Mouth Consumers. By the analogous steps to the CRRA case we can approximate the certainty equivalent (2). To this purpose write (2) as

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

[2]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 $\gamma = 1$.

and 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

$$\begin{aligned} v(c_1, \theta, \Psi) &= \left(\int c_1(\varepsilon)^{1-\theta} d\Psi(\varepsilon)\right)^{\frac{1}{1-\theta}} \\ &\approx \left(1 + (1-\theta)\left(-\frac{\theta}{2}\mu_2^c + \frac{\theta(1+\theta)}{6}\mu_3^c - \frac{\theta(1+\theta)(2+\theta)}{24}\mu_4^c\right)\right)^{\frac{1}{1-\theta}}. \end{aligned} \tag{3}$$

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 results for the CRRA case readily extend.

Precautionary Savings. In the general case where $\gamma \neq \frac{1}{\theta}$, we can use the resource constraint and write utility as

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

The first-order condition is now given by

$$(y_0 - a_1)^{-\frac{1}{\gamma}} = v(c_1, \theta, \Psi)^{\theta - \frac{1}{\gamma}}\mathbb{E}\left[(\exp(\varepsilon) + a_1)^{-\theta}\right]. \tag{4}$$

In the sequel, we follow Kimball and Weil (2009) and assume that the marginal utility of saving, the RHS of (4), 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) (cf. Propositions 5 and 6):

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. Our proof of the proposition is adopted from Krueger and Ludwig (2019). Rewrite the RHS of the first-order condition in (4) as

$$RHS = v(c_1, \theta, \Psi)^{\theta - \frac{1}{\gamma}} f(c_1, \theta, \Psi) \tag{5}$$

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

$$= v(c_1, \theta, \Psi)^{1-\frac{1}{\gamma}} h(c_1, \theta, \Psi). \tag{6}$$

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

1. $\gamma = 1$: Then the RHS is simply from (6)

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

   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 $\exp(\varepsilon)$ 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 $\exp(\varepsilon)$, 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 (6), 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 (6), increasing precautionary savings.

3. $\gamma > 1$: We obtain the following case distinction from (5):

   (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 (5), increasing precautionary savings.

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

□

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.[3]

A.6 Recursive Representation of Persistent Income Component

The 2^{nd} to 4^{th} central moments of z_{ijt} are given recursively by

$$\mu_2(z_{ijt}) = \rho^2 \mu_2(z_{ij-1t-1}) + \mu_2^{\eta}(s(t)) \tag{7a}$$

$$\mu_3(z_{ijt}) = \rho^3 \mu_3(z_{ij-1t-1}) + \mu_3^{\eta}(s(t)) \tag{7b}$$

$$\mu_4(z_{ijt}) = \rho^4 \mu_4(z_{ij-1t-1}) + 6\rho^2 \mu_2(z_{ij-1t-1}) \mu_2^{\eta}(s(t)) + \mu_4^{\eta}(s(t)). \tag{7c}$$

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 Two-Period Model

We start with the decomposition for the two-period model of Section 2, which extends to the quantitative model in a straightforward fashion, as we show in the next subsection. Under the convenient transformation[4] of utility $V = \left[\left(1 - \frac{1}{\gamma}\right) U\right]^{\frac{1}{1-\frac{1}{\gamma}}}$ we compute

$$g_c^i = \frac{V(C^i)}{V(C^{NORM})} - 1 \tag{8}$$

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

[4] I.e., we retransform to the standard EZW functional, cf. Footnote 2.

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

$$g_c^{distr^i} = \frac{V\left(\frac{C^i}{1+\delta_c^i}\right)}{V(C^{NORM})} - 1 = \frac{1+g_c^i}{1+\delta_c^i} - 1. \tag{9}$$

The corresponding mean effect is accordingly

$$g_c^{mean^i} = g_c^i - g_c^{distr^i} = \frac{1+g_c^i}{1+\delta_c^i}\delta_c^i \approx \delta_c^i. \tag{10}$$

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

$$g_c^{distr^i} = g_c^{lcd^i} + g_c^{csd^i} \tag{11}$$

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 change (from $NORM$ to i) 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 change 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

$$g_c^{csd^i} = \frac{\tilde{V}^i}{V(C^{NORM})} - 1 = \frac{1+g_c^i}{1+\delta_c^i} - 1 \tag{12}$$

and thus the life-cycle distribution effect follows as

$$g_c^{lcd^i} = g_c^{distr^i} - g_c^{csd^i}. \tag{13}$$

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 change 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

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

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 to

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

and thus

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

Continuing along these lines we get

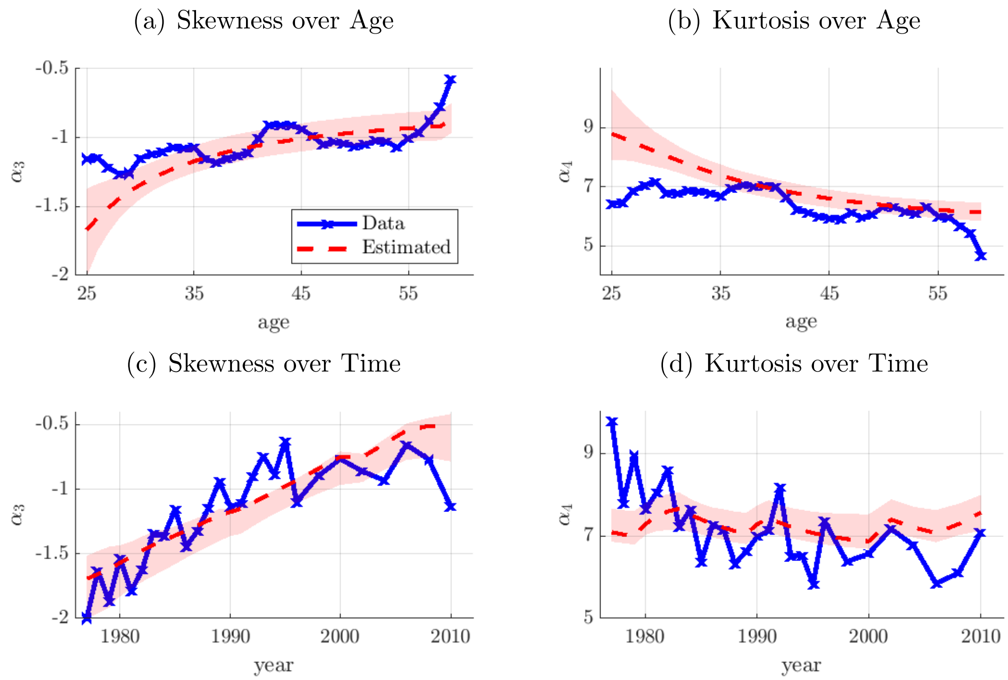
$$\tilde{V}_0^i = \left(\prod_{j=0}^{J} \left(\frac{1}{\delta_j^{c^i}} \right)^{\tilde{\beta}^j} \right)^{1-\tilde{\beta}} V_0^i .$$

With this construction we can now decompose the CEV into the cross-sectional and the life-cycle distribution effects using (12) and (13).

B Fit of Estimated Process

Figure B.1 displays age and year profiles of the standardized third and fourth moments, i.e., of the coefficients of skewness and kurtosis, implied by the estimated theoretical moments and their empirical counterparts.

Figure B.1: Fit of Estimated Process for Post-Government Earnings: Standardized Moments



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

C Calibration Appendix

In this appendix we present details regarding the calibration of the exogenous income profile and shock process. We discuss the two distribution scenarios from the main text, i.e., the full higher-order risk scenario LKSW with leptokurtic and left-skewed shocks, and the reference scenario NORM with Gaussian shocks. Alongside these two scenarios we also show the parameterization of an alternative scenario LK. This is a counterfactual scenario which features leptokurtic shocks that are symmetric in logs, i.e., they are not skewed.

C.1 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

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

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 $\tilde{\tilde{x}}_i = \frac{\tilde{x}_{i+1}+\tilde{x}_i}{2}$.

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

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.2 Moments of the FGLD Distribution

Table C.1 summarizes the moments for distributions NORM, LK, and LKSW, and Table C.2 contains the corresponding parameters of λ of the fitted FGLD distributions.

Table C.1: Moments in Three Distribution Scenarios

Moment	$\hat{\mu}_2$	$\hat{\mu}_3$	$\hat{\mu}_4$	$\hat{\mu}_2$	$\hat{\mu}_3$	$\hat{\mu}_4$	$\hat{\mu}_2$	$\hat{\mu}_3$	$\hat{\mu}_4$
	NORM			LK			LKSW		
Transitory Shock:									
target	0.05	0	0.008	0.05	0	0.219	0.05	-0.047	0.102
fitted	0.05	0	0.008	0.05	0	0.219	0.05	-0.047	0.102
discrete	0.05	0	0.008	0.05	0	0.219	0.051	-0.051	0.107
Persistent Shock—Contraction:									
target	0.022	0	0.001	0.022	0	0.061	0.022	-0.016	0.067
fitted	0.022	0	0.001	0.022	0	0.061	0.022	-0.016	0.067
discrete	0.022	0	0.001	0.022	0	0.061	0.023	-0.02	0.07
Persistent Shock—Expansion:									
target	0.009	0	0	0.009	0	0.008	0.009	-0.001	0.01
fitted	0.009	0	0	0.009	0	0.008	0.009	-0.001	0.01
discrete	0.009	0	0	0.009	0	0.008	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 three distribution scenarios, cf. Section 5.2.

Table C.2: 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 5.2.

C.3 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.3 contains the calibrated values of the pension indexation factor ϱ, which is required to clear the budget of the pension system.

Table C.3: 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.4 Moments of the Earnings Process

Table C.4 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.[5] 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 (counterfactual) distribution LK are extremely high. Left-skewness in logs in distribution LKSW substantially reduces both moments.

Table C.4: Moments of the Earnings Distribution in Logs and Levels

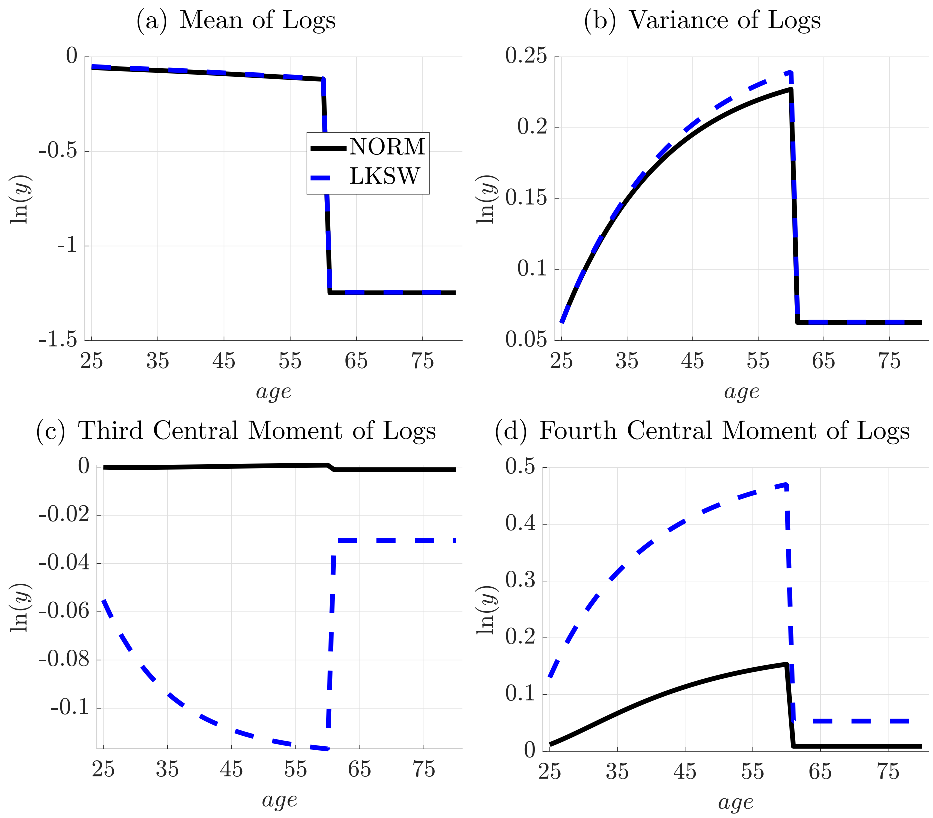
	Logs			Levels		
	Age 25 ($j=0$)					
	NORM	LK	LKSW	NORM	LK	LKSW
μ_2	0.06	0.06	0.06	0.06	0.36	0.05
μ_3	0	0	-0.06	0.01	4.44	0.09
μ_4	0.01	0.24	0.13	0.01	129.43	0.41
	Age 60 ($j=35$)					
μ_2	0.23	0.24	0.24	0.25	0.86	0.3
μ_3	0	0	-0.12	0.21	27.52	1.12
μ_4	0.15	0.56	0.47	0.5	27889.82	27.85

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).

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.

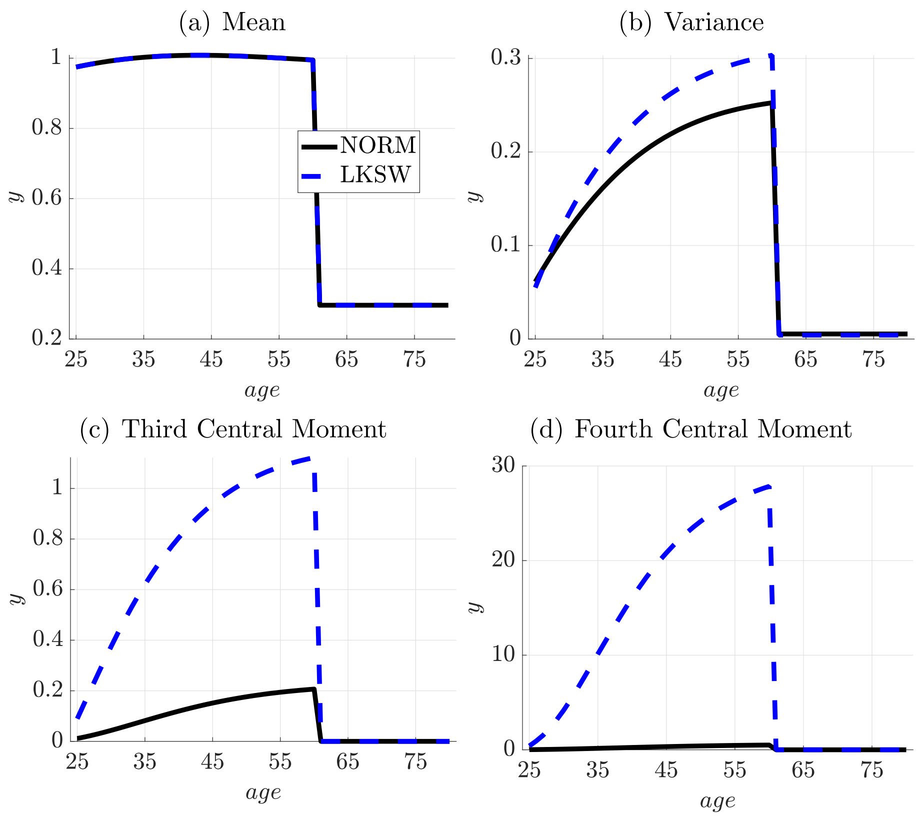
[5] By construction, the variance of the log earnings distribution is the same across distribution scenarios. The difference of 0.01 showing up at age 60 is due to numerical inaccuracies of coarse grids for assets a and the persistent income state z.

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, LKSW: FGLD with excess kurtosis and left-skewness (in logs).

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



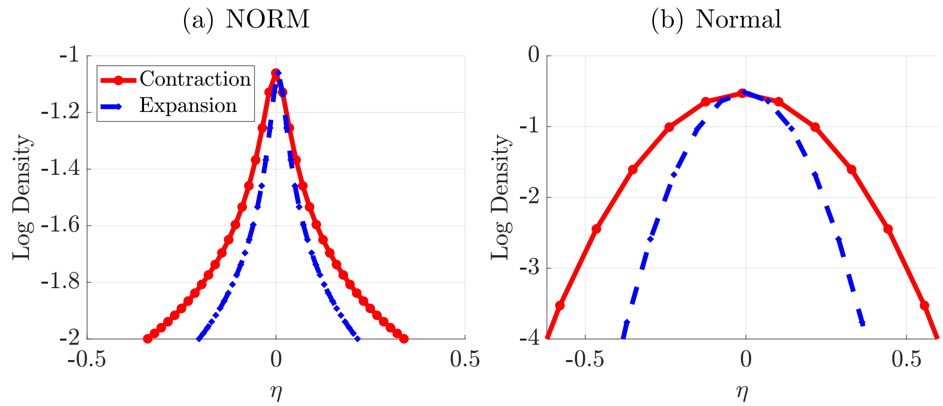
Notes: Figures show moments of cross-sectional distribution of earnings over the life-cycle for each scenario of shock distributions. NORM: FGLD with moments of the normal distribution, LKSW: FGLD with excess kurtosis and left-skewness (in logs).

D Additional Results

D.1 Comparison of FGLD with Normal Distribution

In the application in the main text, we compare the FGLD distribution with left skewness and excess kurtosis (LKSW) to the FGLD with zero skewness and kurtosis of three (NORM). Figure D.1 shows the distribution against the Normal distribution using Gaussian quadrature. The second to fourth central (and standardized) moments of the two distributions are the same—the visual differences are captured by the even moments of higher order. It turns out that these higher-order differences are quantitatively irrelevant in our application: Table D.1 documents the CEV under distribution scenario NORM in comparison to one where shocks are drawn from a Normal distribution. Thus, for the preferences used the differences of moments are not crucial in the calibrated version of the model, and therefore we choose the FGLD distribution NORM as the benchmark.

Figure D.1: Discretized Log Distribution Functions: Persistent Shock



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

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	-1.823	0.36	-2.144	-0.039
NORMAL	-1.825	0.36	-2.146	-0.04
Risk Aversion, $\theta = 2$				
NORM	-3.466	0.652	-4.004	-0.114
NORMAL	-3.471	0.652	-4.009	-0.114
Risk Aversion, $\theta = 3$				
NORM	-4.895	0.896	-5.606	-0.185
NORMAL	-4.903	0.897	-5.614	-0.185
Risk Aversion, $\theta = 4$				
NORM	-6.117	1.111	-6.98	-0.248
NORMAL	-6.127	1.112	-6.989	-0.25

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 and Decomposition

In this appendix, we start out with the baseline calibration and vary it to consider, first, an expected utility formulation with CRRA preferences where we restrict $\theta = \frac{1}{\gamma}$, second, we give households zero initial assets, and third, we recalibrate the discount rate under scenario NORM. We then start out with the KV calibration and vary it by giving households (positive) average assets at the start of the life-cycle as under the baseline. Table D.2 summarizes the results.

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 rate ρ for each value of θ. For $\theta \in \{2, 3, 4\}$ we obtain $\rho \in \{0.009, 0.002, -0.0005\}$ and thus, in contrast to our experiments with EZW utility, the calibrated discount rate is decreasing in θ. For stronger risk attitudes θ the precautionary savings motive is stronger, while the simultaneous lower IES $\gamma = \frac{1}{\theta}$ implies smaller life-cycle savings. The second effect turns out to dominate so that calibration calls for less impatience in order to deliver the same asset profile and the calibrated discount rate even turns negative for $\theta = 4$.

Table D.2: Total CEV g_c of Cyclical Idiosyncratic Risk: Sensitivity Analyses

	Baseline	CRRA	ASS=0	RECAL	KV	KV,ASS>0	GE
				Risk Aversion, $\theta = 1$			
NORM	-1.823	-1.823	-1.975	-1.826	-1.369	-2.067	-1.093
LKSW	-1.536	-1.536	-1.692	-1.536	-1.44	-1.897	-0.952
				Risk Aversion, $\theta = 2$			
NORM	-3.466	-2.704	-3.771	-3.470	-2.407	-3.845	-2.141
LKSW	-3.767	-2.731	-4.465	-3.767	-4.413	-5.127	-2.544
				Risk Aversion, $\theta = 3$			
NORM	-4.895	-3.533	-5.342	-4.884	-2.836	-4.994	-3.070
LKSW	-7.755	-4.752	-10.117	-7.755	-10.588	-10.131	-5.708
				Risk Aversion, $\theta = 4$			
NORM	-6.117	-4.314	-6.686	-6.065	-2.987	-5.242	-3.85
LKSW	-13.272	-8.069	-17.925	-13.272	-17.159	-14.304	-10.487

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; ASS=0: zero initial assets; RECAL: recalibration of discount rate in scenario NORM; GE: general equilibrium; KV: target aggregate capital/income ratio + zero initial assets; KV,ASS>0: KV with positive initial assets.

Column 2 of Table D.2 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 for stronger risk attitudes (lower IES). 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.

Decomposing Baseline vs. KV-Calibration. 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.

Zero initial assets imply that the fraction of borrowing constrained hand-to-mouth consumers increases. For $\theta = 1$, initially roughly 3% of households are constrained in scenario NORM and 2.4% in scenario LKSW. Column 3 of Table D.2 shows that this leads to higher overall welfare losses from cyclical idiosyncratic risk and an increasing importance of higher-order risk. For $\theta = 4$ the difference in the CEV between scenarios LKSW and NORM is about $-11.2\%p$, compared to $-7.2\%p$ under the baseline calibration. Thus, a larger fraction of households at the borrowing constraint increases the role played by higher-order income risk for the welfare losses from cyclical idiosyncratic risk. This explains part of the difference between the baseline calibration and the alternative KV calibration.

Starting from the other end, i.e., from the KV-calibration where $\rho > r$, column 6 reports the CEV of cyclical idiosyncratic risk if households start with positive assets. The difference between LKSW and NORM is about $9.1\%p$ compared to $14.2\%p$ under the KV-calibration. Combining the two steps of the decomposition, zero versus positive initial assets are quantitatively somewhat more relevant for the difference between the welfare costs of cyclical risk under the two alternative calibrations.

Recalibration under scenario NORM. Column 4 reports results under the baseline calibration, where for scenario NORM, we recalibrate the discount rate. Thus, we give the model with Normal shocks its best chance to match the asset profile. Comparing the CEVs under the baseline calibration with the ones for the recalibrated discount rates for the different values of θ reveals that this recalibration is numerically almost irrelevant.

Separating the Role of Kurtosis. Table D.3 provides additional insights on the roles of the different components, i.e., here the excess kurtosis in isolation, of higher-order risk for the high risk aversion calibration with $\theta = 4$. Welfare costs of cyclical risk are about $4\%p$ higher in this distribution scenario than in scenario NORM.

Table D.3: The Welfare Effects of Cyclical Idiosyncratic Risk for Distribution Scenario LK

CEV	g_c	g_c^{mean}	g_c^{lcd}	g_c^{csd}	Δg_c
			Risk Aversion, $\theta = 4$		
LK	-10.595	1.302	-11.582	-0.316	-4.478

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 scenario LK. 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 $g_c = g_c^{mean} + g_c^{lcd} + g_c^{csd}$. $\Delta g_c = g_c^{LK} - g_c^{NORM}$: difference in percentage points relative to scenario NORM.

D.3 General Equilibrium

Overview. In the analyses presented in the main text we consider a partial equilibrium framework where interest rates and average wages are constant. Increased precautionary savings from higher-order risk may lead to a higher capital stock, which in general equilibrium increases wages and lowers 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. We focus on the role of higher-order risk itself, and deliver an approximation regarding the cyclical nature of risk: when in aggregate state $s \in \{E, C\}$ agents make choices assuming that the state of the world will remain unchanged. We then calculate ex-ante expected life-time utility by weighting the two possible states with the probabilities according to the stationary distribution of the aggregate Markov transition matrix.

As a first step, in the baseline scenario we take a normalization such that the net wage rate is one and calibrate the model to an interest rate of $r[\%] = 3\%$ consistent with a standard static representative firm problem in general equilibrium, and accordingly compute the implied parameters of the aggregate production function, with all calibration details provided below.

As a second step, we hold constant these parameters and compute the equilibrium interest rate and wage rate for each considered scenario, cf. Table D.4. Overall, in the economy with cyclical risk, the additional precautionary savings in general equilibrium increase the capital stock increasing wages and decreasing returns. 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. However, these changes are not large, which is a consequence of 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. Aggregate savings of the economy will thus not change strongly.

Column 5 of Table D.2 shows the resulting welfare costs of cyclical risk next to the baseline results. 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 $-6.64\%p$, compared to $-7.16\%p$. We therefore conclude that our main findings are robust in general equilibrium.

Calibration. We close this discussion on the GE extension by providing the details of the calibration. Assuming Cobb-Douglas production with capital elasticity α and a technology

Table D.4: Aggregate Prices in General Equilibrium Model

Variable	w^n		r	
	CR	NCR	CR	NCR
	Risk Aversion, $\theta = 1$			
NORM	1.0009	0.998	0.0299	0.0302
LKSW	1	0.9978	0.03	0.0303
	Risk Aversion, $\theta = 4$			
NORM	0.9977	0.9903	0.0303	0.0312
LKSW	1	0.991	0.03	0.0311

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.

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

$$r = \Upsilon \alpha k^{\alpha-1} - \delta \tag{14a}$$

$$w = \Upsilon(1-\alpha)k^{\alpha}, \tag{14b}$$

which also implies that

$$\frac{w}{r+\delta} = \frac{1-\alpha}{\alpha}k. \tag{15}$$

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 (15) 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 (14a) as

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

Table D.5 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 3%.

Table D.5: Technology Level and Depreciation Rate in General Equilibrium Variant

Parameter		
θ	δ	Υ
1	0.0296	0.9322
2	0.03	0.9338
3	0.0303	0.9356
4	0.0306	0.9374

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

Having determined the parameters δ, Υ in the economy CR/LKSW for each level of risk aversion as summarized in Table D.5 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 given the interest rate from (15) and (14b) as

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

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

References

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.

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

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

Krueger, D. and A. Ludwig (2019). Optimal Taxes on Capital in the OLG Model with Uninsurable Idiosyncratic Income Risk. Working Paper.

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×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.

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

Supplementary Appendix

(Not for Publication)

"Higher-Order Income Risk Over the Business Cycle"

(Christopher Busch and Alexander Ludwig)

S.A Fitting Moments of the FGLD

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

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

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

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

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

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

Lakhany and Mausser (2000) and Su (2007) describe how to estimate the parameters of (S.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

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

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 (S.A.2) giving $f(x) = \frac{dp}{dQ(p)}$ to rewrite

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

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 (S.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 (S.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 (S.A.4) to get

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

We can therefore write explicitly

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

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 $k = 1, \ldots, 4$ 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.

S.B A Numerical Example of the Two-Period Model

In this supplementary 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 S.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.

S.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 S.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 **?**), 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 S.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 S.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 (1) and (2).

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 (1) and (2).

Table S.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.

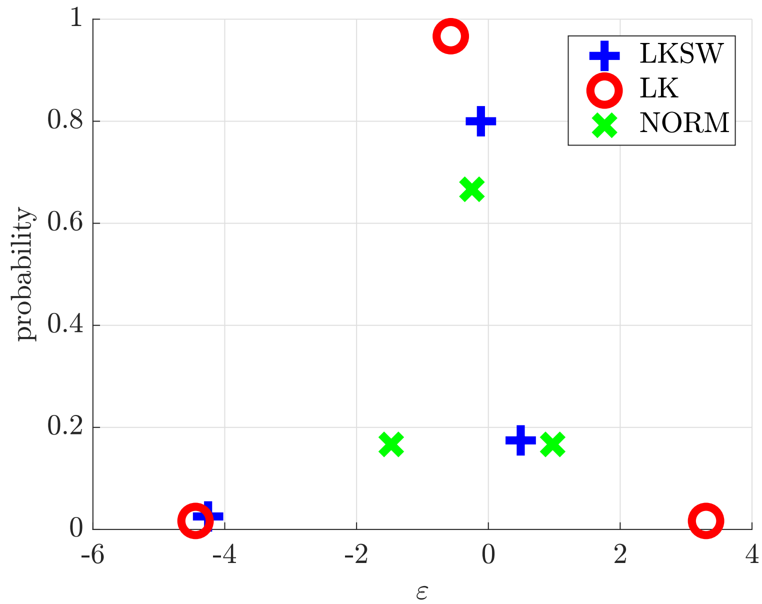
Table S.B.2: 2-Period Model: Shocks, central moments

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 S.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 S.B.1: Distribution of ε



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

S.B.2 Allocations

Table S.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 S.B.4 displays the welfare consequence if there is no access to a savings technology under a binding budget constraint 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 (1). Third, both gains and losses decrease in scenario ST compared to scenario NST. The rea-

Table S.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 S.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.

S.B.3 Decomposition of Consumption Equivalent Variations

Table S.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 S.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.

S.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 S.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 S.B.5 shows that now the mean effect is non-zero. The reason is that savings are inter-temporally 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.

S.B.5 Proof of Proposition S.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

$$\mu_2 = (1-p)q(1-q)\Delta_\varepsilon^2, \tag{S.B.6a}$$

$$\alpha_3 = \frac{2q-1}{\sqrt{(1-p)q(1-q)}}, \tag{S.B.6b}$$

$$\alpha_4 = \frac{3q^2-3q+1}{(1-p)q(1-q)}. \tag{S.B.6c}$$

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

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

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 (S.B.6a) and (S.B.6c) as

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

   and therefore

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

   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 (S.B.6a) we get

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

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

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

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

Now use (S.B.7a) in (S.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

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

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:

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

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 (S.B.7a) that

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

and from (S.B.6a) that

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

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

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

(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

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

Thus condition (S.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 $\alpha_3 < 0$

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

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

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

and following the steps above we thus get

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

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

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

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

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

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

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

□