Appendix

—For Online Publication—

"Higher-Order Income Risk Over the Business Cycle"

(Christopher Busch and Alexander Ludwig)

A Analytical Appendix

A.1 Derivation of Equations (1)–(3)

Consider an exact Taylor series expansion of $u(c_1)$ around $c_1 = \mathbb{E}[y_1]$:

$$u(c_1) = \sum_{k=0}^{\infty} \frac{u^{(k)}(\mathbb{E}[y_1])}{k!}(y_1 - \mathbb{E}[y_1])^k.$$

Now take the expectation of this expression over y_1:

$$\mathbb{E}[u(c_1)] = \mathbb{E}\left[\sum_{k=0}^{\infty} \frac{u^{(k)}(\mathbb{E}[y_1])}{k!}(y_1 - \mathbb{E}[y_1])^k\right].$$

Linearity of the expectation operator implies that the expected value of the sum is the sum of the expected values and thus we obtain

$$\mathbb{E}[u(c_1)] = \sum_{k=0}^{\infty} \frac{u^{(k)}(\mathbb{E}[y_1])}{k!}\mathbb{E}[(y_1 - \mathbb{E}[y_1])^k].$$

With the definition of the k'th central moment as $\mu_k^{y_1} = \mathbb{E}[(y_1 - \mathbb{E}[y_1])^k]$ we thus obtain the expression used in Equation (1):

$$\mathbb{E}[u(c_1)] = \sum_{k=0}^{\infty} \frac{u^{(k)}(\mathbb{E}[y_1])}{k!}\mu_k^{y_1}.$$

In the same fashion, in Equation (2) we use a Taylor series expansion of $\mathbb{E}[u^{(1)}(c_1)]$ around $c_1 = \mathbb{E}[y_1] + a_1$. Note that this is done for a given choice of a_1, which hence is a constant when taking the Taylor series expansion. Going through the same steps as above, we first obtain:

$$u^{(1)}(c_1) = \sum_{k=0}^{\infty} \frac{u^{(k+1)}(\mathbb{E}[y_1] + a_1)}{k!}(y_1 + a_1 - (\mathbb{E}[y_1] + a_1))^k = \sum_{k=0}^{\infty} \frac{u^{(k+1)}(\mathbb{E}[y_1] + a_1)}{k!}(y_1 - \mathbb{E}[y_1])^k.$$

Then, take expectations to obtain the expression used in Equation (2):

$$\mathbb{E}[u^{(1)}(c_1)] = \mathbb{E}\left[\sum_{k=0}^{\infty} \frac{u^{(k+1)}(\mathbb{E}[y_1]+a_1)}{k!}(y_1-\mathbb{E}[y_1])^k\right] = \sum_{k=0}^{\infty} \frac{u^{(k+1)}(\mathbb{E}[y_1]+a_1)}{k!}\mu_k^{y_1}.$$

Next, turn to Equation (2) again, take the total derivative with respect to y_0, and set it equal to zero:

$$\begin{aligned}
\frac{df}{dy_0} =& \frac{d\left(u^{(1)}(y_0-a_1)-\mathbb{E}[u^{(1)}(y_1+a_1)]\right)}{dy_0} \stackrel{!}{=} 0 \\
\Leftrightarrow & u^{(2)}(y_0-a_1)\left(1-\frac{\partial a_1}{\partial y_0}\right) - \mathbb{E}\left[u^{(2)}(y_1+a_1)\right]\frac{\partial a_1}{\partial y_0} = 0 \\
\Leftrightarrow & u^{(2)}(y_0-a_1) = \frac{\partial a_1}{\partial y_0}\left(u^{(2)}(y_0-a_1)+\mathbb{E}\left[u^{(2)}(y_1+a_1)\right]\right) \\
\Leftrightarrow & 1 = \frac{\partial a_1}{\partial y_0}\left(1+\frac{\mathbb{E}\left[u^{(2)}(y_1+a_1)\right]}{u^{(2)}(y_0-a_1)}\right) \\
\Leftrightarrow & \frac{\partial a_1}{\partial y_0} = \left(1+\frac{\mathbb{E}\left[u^{(2)}(y_1+a_1)\right]}{u^{(2)}(y_0-a_1)}\right)^{-1}.
\end{aligned}$$

This gives the expression for the MPS used in Equation (3). Finally, in the discussion of Equation (3), we again take a Taylor series expansion:

$$u^{(2)}(y_1+a_1) = \sum_{k=0}^{\infty} \frac{u^{(k+2)}(\mathbb{E}[y_1]+a_1)}{k!}(y_1+a_1-(\mathbb{E}[y_1]+a_1))^k = \sum_{k=0}^{\infty} \frac{u^{(k+2)}(\mathbb{E}[y_1]+a_1)}{k!}(y_1-\mathbb{E}[y_1])^k.$$

Taking expectations we get:

$$\begin{aligned}
\mathbb{E}\left[u^{(2)}(y_1+a_1)\right] &= \mathbb{E}\left[\sum_{k=0}^{\infty} \frac{u^{(k+2)}(\mathbb{E}[y_1]+a_1)}{k!}(y_1-\mathbb{E}[y_1])^k\right] \\
&= \sum_{k=0}^{\infty} \frac{u^{(k+2)}(\mathbb{E}[y_1]+a_1)}{k!}\mathbb{E}\left[(y_1-\mathbb{E}[y_1])^k\right] \\
&= \sum_{k=0}^{\infty} \frac{u^{(k+2)}(\mathbb{E}[y_1]+a_1)}{k!}\mu_k^{y_1}.
\end{aligned}$$

A.2 Decomposition of Consumption Equivalent Variations

Recall that for a given shock scenario $i \in \{NORM, LKSW\}$, the welfare function is defined as the certainty equivalent of entering the economy with asset holdings a, and then optimally choosing consumption and savings in every period when facing shock scenario i.

Before splitting it up into different components, we briefly repeat the derivation of the CEV here. We denote the resulting stochastic consumption stream by $\boldsymbol{c}^i(a)$, which gives optimal consumption at every age for every possible history conditional on starting out with assets a. The welfare function is $W^i\left(\boldsymbol{c}^i(a)\right) = \nu\left(V_0^i(a, z, \varepsilon; s)\right)$, where

$$v(V_0(a, z, \varepsilon; s)) = \begin{cases} \left(\mathbb{E}_{-1}\left[V_0(a, z, \varepsilon; s)^{1-\theta}\right]\right)^{\frac{1}{1-\theta}} & \theta \neq 1 \\ \exp\left(\mathbb{E}_{-1}\left[\ln V_0(a, z, \varepsilon; s)\right]\right) & \text{otherwise,} \end{cases}$$

where the expectation is taken over $z = \eta, \varepsilon, s$. The overall consumption equivalent variation (CEV), g_c, that makes households in scenario NORM indifferent to scenario LKSW, is obtained by

$$W^{LKSW}\left(\boldsymbol{c}^{LKSW}(a = \bar{a}_0)\right) = W^{NORM}\left((1 + g_c)\boldsymbol{c}^{NORM}(a = \bar{a}_0)\right),$$

which gives

$$1 + g_c = \frac{W^{LKSW}\left(\boldsymbol{c}^{LKSW}(a = \bar{a}_0)\right)}{W^{NORM}\left(\boldsymbol{c}^{NORM}(a = \bar{a}_0)\right)}. \tag{A.1}$$

We split this overall welfare effect into three components. First, we decompose the CEV into a *mean* and a *distribution* effect. 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, expected average consumption over the life cycle of some cohort is obtained as

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

for $i \in \{NORM, LKSW\}$, where $c_j^i(a_j, z_j, \varepsilon; s)$ denotes the consumption policy function in distribution i and $\Phi_j^i(a_j, z_j, \varepsilon_j; s)$ is the cross-sectional distribution. In the above equation and the exposition below we sometimes drop the explicit dependence of the stochastic consumption sequence on the initial assets $a = \bar{a}_0$ for notational convenience.

Denote by $\delta_c^{LKSW} = \frac{\mathbb{E}[\boldsymbol{c}^{LKSW}]}{\mathbb{E}[\boldsymbol{c}^{NORM}]} - 1$ the percentage difference of average consumption between scenarios NORM and LKSW. We then scale the consumption path $\boldsymbol{c}^{LKSW}$ by this average difference to obtain a consumption path adjusted for the mean difference: $\boldsymbol{c}_{adj}^{LKSW} = \boldsymbol{c}^{LKSW}/(1 + \delta_c^{LKSW})$. The resulting difference between the paths $\boldsymbol{c}_{adj}^{LKSW}$ and

$\boldsymbol{c}^{NORM}$ gives the welfare effect due to differences in the distribution around the mean:

$$W^{LKSW}\left(\boldsymbol{c}_{adj}^{LKSW}(a=\bar{a}_0)\right)=W^{NORM}\left((1+g_c^{distr.})\boldsymbol{c}^{NORM}(a=\bar{a}_0)\right),$$

which gives

$$W^{LKSW}\left(\boldsymbol{c}^{LKSW}(a=\bar{a}_0)/(1+\delta_c^{LKSW})\right)=W^{NORM}\left((1+g_c^{distr.})\boldsymbol{c}^{NORM}(a=\bar{a}_0)\right)$$

and thus we obtain

$$1+g_c^{distr.}=\frac{1}{1+\delta_c^{LKSW}}\frac{W^{LKSW}\left(\boldsymbol{c}^{LKSW}(a=\bar{a}_0)\right)}{W^{NORM}\left(\boldsymbol{c}^{NORM}(a=\bar{a}_0)\right)}=\frac{1+g_c}{1+\delta_c^{LKSW}}. \tag{A.2}$$

The mean effect is accordingly

$$g_c^{mean}=g_c-g_c^{distr.}=\frac{1+g_c}{1+\delta_c^{LKSW}}\delta_c^{LKSW}. \tag{A.3}$$

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 write $g_c^{distr.}$ as

$$g_c^{distr.}=g_c^{lcd}+g_c^{csd} \tag{A.4}$$

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

To obtain the cross-sectional distribution effect we correct the stochastic consumption path $\boldsymbol{c}^{LKSW}$ for the age-specific percentage difference to $\boldsymbol{c}^{NORM}$. Denote the latter by $\delta_{c,j}^{LKSW}=\frac{\mathbb{E}[\boldsymbol{c}^{LKSW}|j]}{\mathbb{E}[\boldsymbol{c}^{NORM}|j]}-1$. Adjusting the consumption plan accordingly we get $\boldsymbol{c}_{age-adj}^{LKSW}$ where for every age j the consumption plan $\boldsymbol{c}_{age-adj,j}^{LKSW}=\boldsymbol{c}_j^{LKSW}/(1+\delta_{c,j}^{LKSW})$. Accordingly, we get

$$1+g_c^{csd}=\frac{W^{LKSW}\left(\boldsymbol{c}_{age-adj}^{LKSW}(a=\bar{a}_0)\right)}{W^{NORM}\left(\boldsymbol{c}^{NORM}(a=\bar{a}_0)\right)}. \tag{A.5}$$

Once we have obtained g_c^{csd}, we directly get g_c^{lcd} from (A.4) as $g_c^{lcd}=g_c^{distr.}-g_c^{csd}$. Note that if $\delta_{c,j}^{LKSW}=\delta_c^{LKSW}\forall j$, i.e., if the behavioral response to higher-order risk yields a parallel shift of the mean consumption life-cycle profile, (A.5) implies that $g_c^{csd}=g_c^{distr.}$ (and accordingly $g_c^{lcd}=0$). Thus, the life-cycle distribution effect indeed reflects the welfare effect due to a tilting of the mean life-cycle consumption profile.

A.3 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* rather than of the shocks in *logs*. Consider a version of the two-period framework from Section 2.1 (agents live in periods 0 and 1 and face risky second period income) with log-utility as per-period utility function. Again, let y_0 and y_1 denote income in the two periods. Consider a mean preserving (thus $\mathbb{E}[y_1] = 1$) change of period 1 income 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. We summarize this below in Proposition 1. While the finding 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]

Proposition 1. *Suppose that the utility function is logarithmic 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.*

[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 y_1. Obviously, increasing the variance (or kurtosis) of y_1, while holding the mean constant at $\mathbb{E}[y_1] = 1$, has direct negative utility consequences. In this case utility is $U = \ln(y_0) + \mathbb{E}[\ln(y_1)]$, which for the maintained normalization $\mathbb{E}[y_1] = 1$ we could approximate as

$$U \approx \ln(y_0) - \frac{1}{2}\mu_2^{y_1} + \frac{1}{3}\mu_3^{y_1} - \frac{1}{4}\mu_4^{y_1}$$

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

Proof. Let $\mathbb{E}_{\Psi}[\ln(y_1)] = \int \ln(y_1) d\Psi$, $\mu_k^{\ln(y_1)} = \int \left(\ln(y_1) - \mathbb{E}_{\Psi}[\ln(y_1)]\right)^k d\Psi$ for $k > 1$, and let $\mathbb{E}_{\Psi}[y_1] = 1$, where $\Psi(\cdot)$ denotes some distribution function. Now consider the following three steps.

(i.) Denote by $\tilde{\Psi}^{\delta_k}(\ln(y_1))$ a distribution function that is obtained from $\Psi(\ln(y_1))$ by changing central moment $\mu_k^{y_1}$, while holding the mean constant, i.e., $\mathbb{E}_{\tilde{\Psi}^{\delta_k}}[\ln(y_1)] = \mathbb{E}_{\Psi}[\ln(y_1)]$.

(ii.) Consider a mean-preserving change of the distribution, which combines the change from $\Psi(\cdot)$ to $\tilde{\Psi}^{\delta_k}(\cdot)$ with a shift Δ^{δ_k} such that $\mathbb{E}_{\tilde{\Psi}^{\delta_k}}[\exp(\ln(y_1)+\Delta^{\delta_k})] = \mathbb{E}_{\Psi}[\exp(\ln(y_1))] = 1$. The normalization $1 = \mathbb{E}_{\tilde{\Psi}^{\delta_k}}[\exp(\ln(y_1)+\Delta^{\delta_k})] = \int \exp(\ln(y_1)+\Delta^{\delta_k}) d\tilde{\Psi}^{\delta_k} = \exp(\Delta^{\delta_k}) \int y_1 d\tilde{\Psi}^{\delta_k}$ implicitly defines $\Delta^{\delta_k} = -\ln\left(\int y_1 d\tilde{\Psi}^{\delta_k}\right)$.

(iii.) Thus, when facing distribution $\Psi(\ln(y_1))$, the expected income in log terms is $\mathbb{E}_{\Psi}[\ln(y_1)]$, and when facing distribution $\tilde{\Psi}^{\delta_k}(\ln(y_1))$ together with the mean-preserving shift Δ^{δ_k}, the expected income in log terms is $\mathbb{E}_{\tilde{\Psi}^{\delta_k}}[\ln(y_1) + \Delta^{\delta_k}] = \mathbb{E}_{\tilde{\Psi}^{\delta_k}}[\ln(y_1)] + \Delta^{\delta_k} = \mathbb{E}_{\Psi}[\ln(y_1)] + \Delta^{\delta_k}$. The last equality follows from (i.). With logarithmic utility and binding budget constraint, the expected utility difference across distributions Ψ and $\tilde{\Psi}^{\delta_k}$ is thus $\Delta U = (U \mid \tilde{\Psi}^{\delta_k}) - (U \mid \Psi) = \Delta^{\delta_k}$.

We then get the following:

- Shifting probability mass from the center to the tails, either by increasing the variance ($k = 2$) or kurtosis ($k = 4$), increases $\int y_1 d\tilde{\Psi}^{\delta_k}$ above one, which follows from Jensen's inequality for convex functions. Thus $\Delta^{\delta_k} = -\ln\left(\int y_1 d\tilde{\Psi}^{\delta_k}\right) < 0$.

- Shifting probability mass from the right tail to the left tail decreasing the skewness ($k = 3$)—i.e., making the distribution more left-skewed—decreases $\int y_1 d\tilde{\Psi}^{\delta_k}$ below one, which follows from Jensen's inequality for convex functions. Thus $\Delta^{\delta_k} = -\ln\left(\int y_1 d\tilde{\Psi}^{\delta_k}\right) > 0$.

□

B Additional Empirics

B.1 Additional Moments and Pre-Government Income

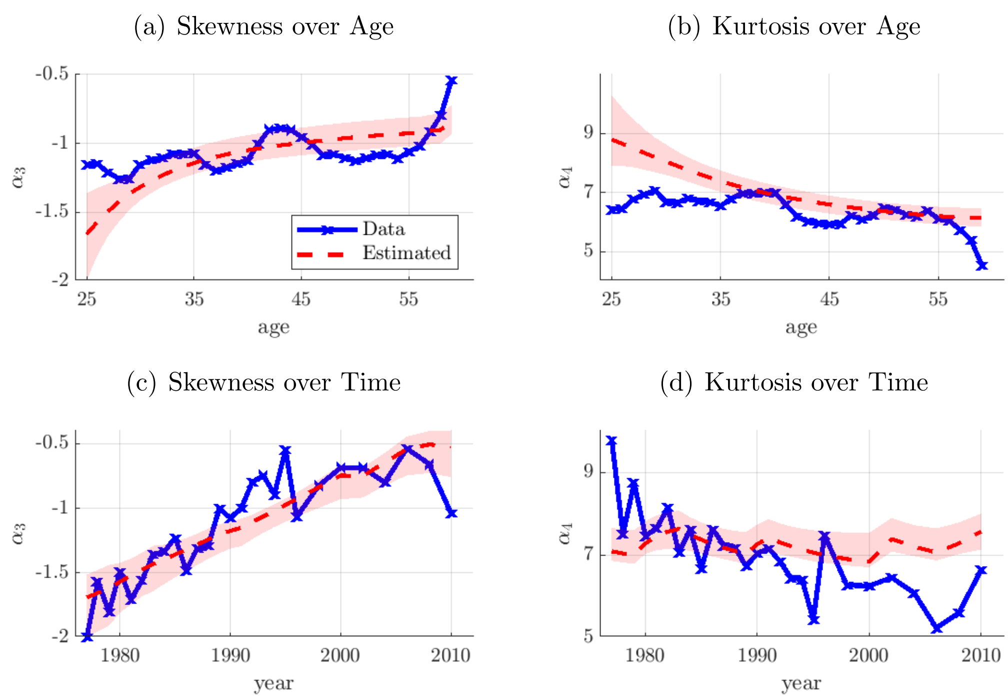
Table B.1 shows the standardized moments implied by the estimates in Table 1. We denote the i standardized moment of shock $x \in \{\varepsilon, \eta(s)\}$ by α_i^x, where $\alpha_i^x = \frac{\mu_i^x}{\mu_2^{x/2}}$. Figure B.1 displays age and year profiles of the standardized third and fourth moments, i.e., of the coefficients of skewness and kurtosis, implied by the estimated theoretical moments for post-government income and their empirical counterparts. Table B.2 shows moments of income changes in the data and implied by the estimated income process.

Table B.1: Estimation Results for Household Net Income—Standardized Moments

Standardized Moments					
μ_2^χ	0.1076	α_3^χ	-1.47	α_4^χ	1.50
	[0.0897; 0.1237]		[−2.32; −0.78]		[0.00; 5.38]
μ_2^ε	0.0752	α_3^ε	-4.20	α_4^ε	40.62
	[0.0677; 0.0816]		[−4.79; −3.73]		[36.27; 47.65]
$\mu_2^{\eta,C}$	0.0223	$\alpha_3^{\eta,C}$	-4.95	$\alpha_4^{\eta,C}$	134.47
	[0.0152; 0.0291]		[−7.80; −2.18]		[82.02; 191.34]
$\mu_2^{\eta,E}$	0.0085	$\alpha_3^{\eta,E}$	-1.54	$\alpha_4^{\eta,E}$	134.47
	[0.0044; 0.0153]		[−7.97; 9.09]		[82.02; 191.34]

Notes: Table shows standardized moments for household income after taxes and transfers implied by estimates in Table 1. Brackets show 5^{th} and 95^{th} percentiles of 1,000 bootstrap estimates (998 of the bootstrap iterations converge).

Figure B.1: Fit of Estimated Process for Household Net Income: 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.

Table B.3 shows estimated central moments for pre-government household income, and table B.4 shows the implied standardized moments. Comparison to the estimates for post government incomes shows that the estimated process captures intuitive features: the distributions of shocks to post-government income are more compressed in comparison to the

Table B.2: Moments of Income Changes

	m_2		m_3		m_4	
	Data	Estimated	Data	Estimated	Data	Estimated
$\Delta_1 y$	0.1366	*0.1635*	-0.0029	*-0.0004*	0.4006	*0.4877*
$\Delta_2 y$	0.1708	*0.1762*	-0.0047	*-0.0006*	0.4715	*0.5145*
$\Delta_3 y$	0.1935	*0.1884*	0.0030	*-0.0003*	0.5287	*0.5392*
$\Delta_4 y$	0.2146	*0.2004*	0.0005	*0.0002*	0.6223	*0.5631*
$\Delta_5 y$	0.2300	*0.2096*	0.0174	*0.0036*	0.7124	*0.5758*

Notes: Table shows the averages of second, third, and fourth central moments of 1–5 year income changes in data, and implied by estimates.

Table B.3: Estimation Results for Household Pre-Government Income

Estimated Central Moments					
ρ	0.9601				
	[0.9412; 0.9756]				
μ_2^χ	0.1591	μ_3^χ	−0.1089	μ_4^χ	0.0607
	[0.1361; 0.1786]		[−0.1497; −0.0689]		[0.0000; 0.1508]
μ_2^ε	0.1045	μ_3^ε	−0.1513	μ_4^ε	0.4250
	[0.0948; 0.1133]		[−0.1621; −0.1373]		[0.3630; 0.4867]
$\mu_2^{\eta,C}$	0.0375	$\mu_3^{\eta,C}$	−0.0340	$\mu_4^{\eta,C}$	0.1359
	[0.0263; 0.0477]		[−0.0485; −0.0180]		[0.0856; 0.1719]
$\mu_2^{\eta,E}$	0.0152	$\mu_3^{\eta,E}$	−0.0052	$\mu_4^{\eta,E*}$	0.0225
	[0.0099; 0.0229]		[−0.0136; 0.0029]		[0.0089; 0.0488]

Notes: Estimated central moments for household income before taxes and transfers. Brackets show 5^{th} and 95^{th} percentiles of 1,000 bootstrap estimates. $^*\mu_4^{\eta,E}$ not separately estimated.

estimated shocks to pre-government income. As expected from this reduced dispersion, the third central moments are smaller in magnitude for post-government income. The cyclical pattern is qualitatively the same in that dispersion is countercyclical and skewness is procyclical. Estimates of the kurtosis reveal that the distribution of post-government income shocks is more concentrated in the center, while some households experience shocks that are more extreme relative to the overall more compressed (in comparison to pre-government income) distribution.

Table B.4: Estimation Results for Household Pre-Government Income—Standardized Moments

Standardized Moments					
μ_2^{χ}	0.1591	α_3^{χ}	-1.72	α_4^{χ}	2.40
	[0.1361; 0.1786]		[−2.41; −1.15]		[0.00; 5.39]
μ_2^{ε}	0.1045	α_3^{ε}	-4.48	α_4^{ε}	38.94
	[0.0948; 0.1133]		[−5.05; −3.97]		[34.52; 44.92]
$\mu_2^{\eta,C}$	0.0375	$\alpha_3^{\eta,C}$	-4.69	$\alpha_4^{\eta,C}$	96.85
	[0.0263; 0.0477]		[−6.85; −2.69]		[61.15; 141.97]
$\mu_2^{\eta,E}$	0.0152	$\alpha_3^{\eta,E}$	-2.76	$\alpha_4^{\eta,E*}$	96.85
	[0.0099; 0.0229]		[−6.97; 2.03]		[61.15; 141.97]

Notes: Table shows standardized moments for household income before taxes and transfers implied by estimates in Table B.3. Brackets show 5^{th} and 95^{th} percentiles of 1,000 bootstrap estimates.

B.2 Robustness Checks on Estimates

Table B.5 shows a set of robustness specifications for the estimation. The overall insight is that the core parameters that go into our model analysis—the second to fourth central moments of the transitory and persistent shocks—are robust to various alternatives. We now explain the different robustness checks step-by-step in the order in which they are presented in Table B.5. Column (0) repeats the parameters from the baseline specification (either the estimates or the implied parameters in case of restrictions).

(1) joint: In the baseline specification, we estimate the sets of second, third, and fourth central moments of shocks separately. In column *joint* we report instead the estimates implied when estimating all parameters jointly. All parameters lie are quantitatively very close to the baseline estimates, and all lie within the confidence bands.

(2)-(3) W Δ: In the baseline specification, we use a weight of 10% for the moments of income changes. In the two columns *W Δ5%* and *W Δ15%*, we present estimates when instead using either 5% or 15%, respectively. The estimates for the core parameters are not significantly different.

(4) cyc ε: In the baseline specification, we do not allow the transitory shock distribution to vary over the cycle. In alternative specification *cyc* ε, we remove this restriction. As expected, the point estimates of moments of ε in the baseline specification lie in between the two estimated state-specific moments. Importantly, the estimates of the persistent component are unaffected by this different specification, in line with the fact that those moments are identified from accumulated differences from contractions and expansions.

(5) cyc α_4: In the baseline specification, we do not allow the kurtosis of the persistent shock to vary over the cycle. In alternative specification *cyc* α_4, we remove this restriction. Of course, estimates for second and third moments (the first two panels) are unaffected. Turning to the fourth moments, as expected the transitory shock moments are unaffected, while the (now unrestricted) estimates for the persistent component are not significantly different from the baseline (restricted) specification.

(6) χ norm: In the baseline specification, we do not restrict the moments of the fixed effect, χ. In the baseline, the confidence band is at the lower limit, and in two of the preceding robustness checks, the point estimates are at the lower limit. In other words, the moment is not precisely estimated. We thus consider a restricted estimation, in which we do assume the third and fourth moments of a Normal distribution, i.e., we restrict the skewness and third central moment of χ to zero, and the kurtosis to three. Importantly, this restriction does not significantly affect the estimates of the core parameters (third and fourth moments of transitory and persistent shocks).

(7)-(8) trend μ_3: In the baseline specification, we introduce a linear trend into the third central moment of the transitory shock, m_3^ε, and report the time-series average as the estimate of m_3^ε. Reason for the introduction of this trend component is that the time-series displays a trend, while the model we are estimating is stationary, and would thus, obviously, not stand a good chance of matching the data moments. In two alternative specifications, we instead control for the trend in the data by introducing either, (i.), a linear trend over cohorts into the distribution of χ, or by introducing, (ii.), a linear trend over time in the persistent shocks. In (i.), the reported μ_3^χ is the implied raw average over cohorts. Importantly, this does barely affect the estimated third moments of the transitory and persistent shocks. The third moment of transitory shocks is almost identical. Similarly, the third moment of persistent shocks in contractions is very close and not statistically different from baseline. The third moment of persistent shocks in expansions, at 0.0080, is slightly above the confidence band for the baseline, which goes until 0.0040. In (ii.), we add a linear trend to the persistent component starting in 1977, the first year of our sample. The reported $\mu_3^{\eta,C}$ and $\mu_3^{\eta,E}$ show the estimated moments net of the trend component. The estimates are almost identical to the baseline. The additive trend component has an estimated slope of 0.0007. Again, the transitory shocks are basically unaffected. The persistent shocks are estimated to be more skewed to the left. The cyclical pattern is unaffected.

(9) pers. ≥20: In the baseline specification, we assume that a cohort enters the labor market at age 25, and then accumulates shocks. In this alternative specification, we instead assume that a cohort enters at age 20 and accordingly the persistent component accumulates from this earlier age on. For the oldest cohort in the sample (those aged 60 in the first wave), this implies accumulation from 1937 on (instead of 1942). Based on NBER business cycle dating, we classify year 1937 as contraction, and the other four years as expansions. We estimate it on the same data moments as the baseline. Again, estimates are not significantly affected. Note that the variance of χ, m_2^{χ}, is smaller because (together with m_2^{ε}) it now captures variance at age 20 as opposed to age 25.

(10) hh dum: In the baseline specification, we control for the log of household size in the first stage regression. In this alternative specification, we instead control for household size by introducing dummies for household size, and then use the resulting residuals in the second stage estimation of the income process. The estimates are virtually unchanged relative to the baseline estimation.

Table B.5: Estimation Results for Robustness Specifications

	(0) Baseline	(1) Joint	(2) W $\Delta 5\%$	(3) W $\Delta 15\%$	(4) cyc ε	(5) cyc α_4	(6) χ norm.	(7) trend μ_3^χ	(8) trend μ_3^η	(9) bc ≥ 20	(10) hh dum
ρ	0.9683 [0.9463; 0.9841]	0.9780	0.9788	0.9576	0.9673	0.9683	0.9683	0.9683	0.9683	0.9649	0.9682
μ_2^χ	0.1076 [0.0897; 0.1237]	0.0755	0.1207	0.0954	0.1066	0.1076	0.1076	0.1076	0.1076	0.0295	0.1085
$\mu_2^{\varepsilon,C}$	0.0752 [0.0677; 0.0816]	0.0720	0.0786	0.0704	0.0565	0.0752	0.0752	0.0752	0.0752	0.0662	0.0755
$\mu_2^{\varepsilon,E}$	0.0752 [0.0677; 0.0816]	0.0720	0.0786	0.0704	0.0834	0.0752	0.0752	0.0752	0.0752	0.0662	0.0755
$\mu_2^{\eta,C}$	0.0223 [0.0152; 0.0291]	0.0178	0.0183	0.0260	0.0248	0.0223	0.0223	0.0223	0.0223	0.0301	0.0222
$\mu_2^{\eta,E}$	0.0085 [0.0044; 0.0153]	0.0103	0.0056	0.0124	0.0080	0.0085	0.0085	0.0085	0.0085	0.0146	0.0084
μ_3^χ	−0.0508 [−0.0780; −0.0253]	−0.0577	−0.0491	−0.0561	−0.0513	−0.0508	0	−0.1163	−0.0830	−0.0378	−0.0504
$\mu_3^{\varepsilon,C}$	−0.0866 [−0.0935; −0.0771]	−0.0822	−0.0871	−0.0863	−0.0835	−0.0866	−0.0990	−0.0837	−0.0830	−0.0858	−0.0865
$\mu_3^{\varepsilon,E}$	−0.0866 [−0.0935; −0.0771]	−0.0822	−0.0871	−0.0864	−0.0880	−0.0866	−0.0990	−0.0837	−0.0830	−0.0858	−0.0865
$\mu_3^{\eta,C}$	−0.0167 [−0.0266; −0.0062]	−0.0146	−0.0153	−0.0170	−0.0172	−0.0167	−0.0217	−0.0105	−0.0172	−0.0168	−0.0167
$\mu_3^{\eta,E}$	−0.0013 [−0.0073; 0.0040]	0.0001	−0.0006	−0.0016	−0.0012	−0.0013	−0.0065	0.0080	−0.0062	−0.0030	−0.0014
μ_4^χ	0.0173 [0.0000; 0.0741]	0.0579	0.0455	0.0000	0.0018	0.0000	0.0347	0.0173	0.0173	0.0000	0.0181
$\mu_4^{\varepsilon,C}$	0.2300 [0.1927; 0.2664]	0.2462	0.2399	0.2215	0.1249	0.2259	0.2326	0.2300	0.2300	0.2290	0.2300
$\mu_4^{\varepsilon,E}$	0.2300 [0.1927; 0.2664]	0.2462	0.2399	0.2215	0.2716	0.2259	0.2326	0.2300	0.2300	0.2290	0.2300
$\mu_4^{\eta,C}$	0.0666 [0.0363; 0.0847]	0.0294	0.0510	0.0785	0.0823	0.0492	0.0602	0.0666	0.0666	0.0544	0.0667
$\mu_4^{\eta,E}$	0.0098 [0.0022; 0.0272]	0.0098	0.0047	0.0179	0.0085	0.0222	0.0089	0.0098	0.0098	0.0129	0.0097

Notes: Table shows estimated central moments for household income after taxes and transfers. Column 1 shows baseline estimates with 5^{th} and 95^{th} percentiles of bootstrap estimates in brackets. Other columns show point estimates for different specifications.

C Calibration Appendix

In this appendix we present details regarding the calibration of the exogenous income profile and shock process. First, in Section C.1, we discuss the stochastic working life income process in 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 left-skewed. The quantitative results based on LK are shown in appendix D.3. Second, in Section C.2, we discuss the pension system which pins down the (non-stochastic) old-age income. Further, in Section C.3 we show the asset to income ratio.

C.1 The Income Process

Discretization of the FGLD

We numerically solve for λ_3 and λ_4 jointly to fit the third and fourth central moments.[2] Next, we determine λ_2 to match the variance and λ_1 to match the mean, both in closed form. In the distribution NORM the parameter restriction on the FGLD is that $\lambda_3 = \lambda_4$.

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

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 $p_i = Q^{-1}(\tilde{\tilde{x}}_i)$.

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

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 $1 - p_{n-1}$.

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. We choose $n = 41$.

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	0.050	0	0.008	0.050	0	0.219	0.050	-0.047	0.102
fitted	0.050	0	0.008	0.050	0	0.219	0.050	-0.047	0.102
discrete	0.050	0	0.008	0.050	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	-0.017	0.066
fitted	0.022	0	0.001	0.022	0	0.061	0.022	-0.016	0.066
discrete	0.022	0	0.001	0.022	0	0.061	0.023	-0.020	0.070
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: Table shows the target central moment together with the central moment of the fitted FGLD, and of the discretized FGLD for three distribution scenarios.

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
Persistent—Contraction:	1.000	0.539	5.203	5.203
Persistent—Expansion:	1.000	0.871	5.203	5.203
		LK		
Transitory:	1.000	0.002	173.309	173.309
Persistent—Contraction:	1.000	0.002	244.954	244.954
Persistent—Expansion:	1.000	0.003	220.344	220.344
		LKSW		
Transitory:	0.197	0.008	92.959	57.755
Persistent—Contraction:	0.425	0.002	289.898	225.714
Persistent—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 4.1.

Moments of the Earnings Process

Table C.3 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.[3] 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.

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.

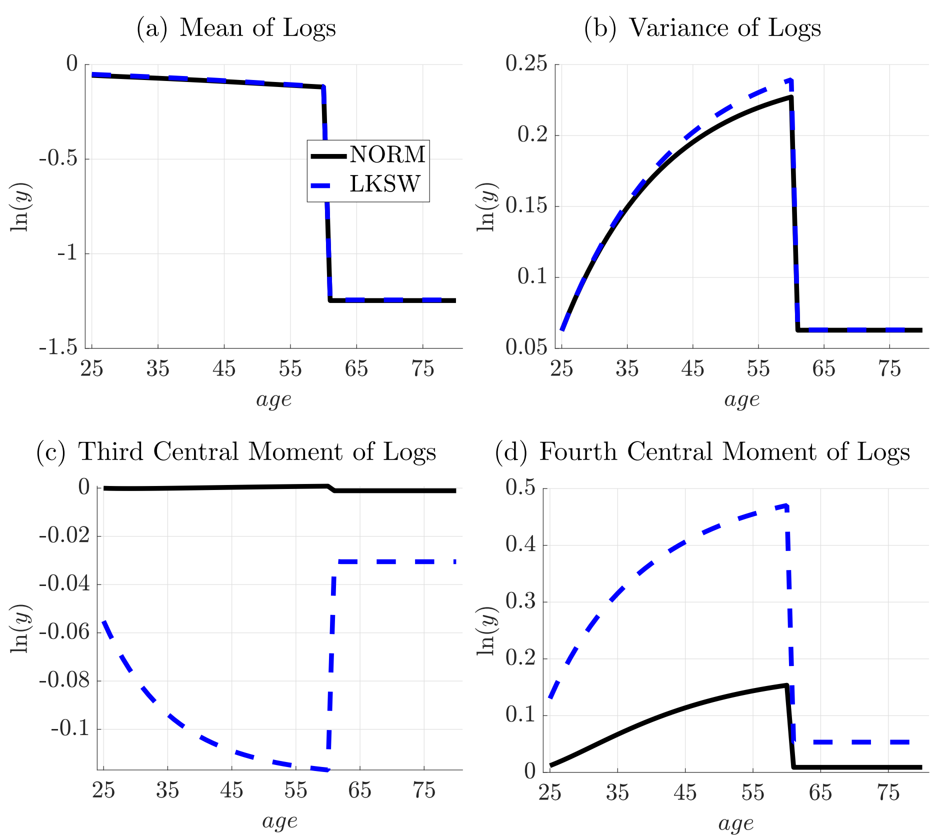
[3]By construction, the variance of the log earnings distribution is the same across distribution scenarios. The difference of 0.01 showing up at age 60 is due to numerical inaccuracies of coarse grids for assets a and the persistent income state z.

Table C.3: 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

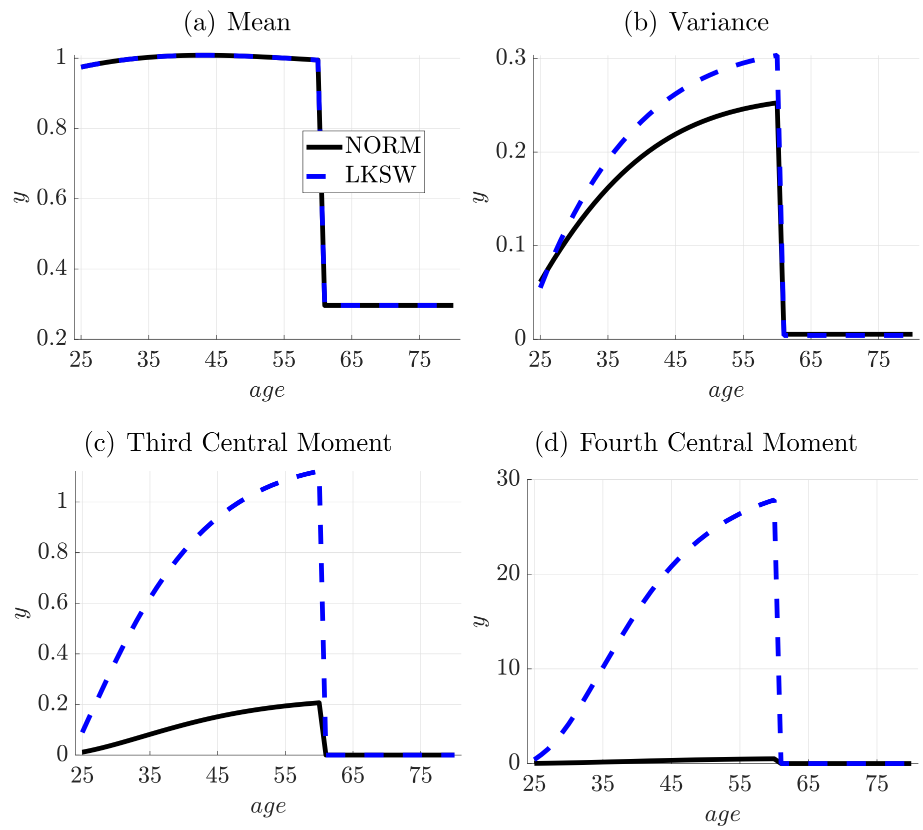
Notes: Moments of cross-sectional distribution of log earnings and earnings at ages 25 ($j = 0$) and 60 ($j = 35$) for each scenario of shock distributions. NORM: FGLD with moments of the normal distribution, LK: FGLD with excess kurtosis, LKSW: FGLD with excess kurtosis and left-skewness (in logs).

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, 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).

C.2 The Pension System

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:

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

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

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

Table C.4 contains the calibrated values of the pension indexation factor ϱ, which is required to clear the budget of the pension system.

Table C.4: Pension Indexation Factor ϱ

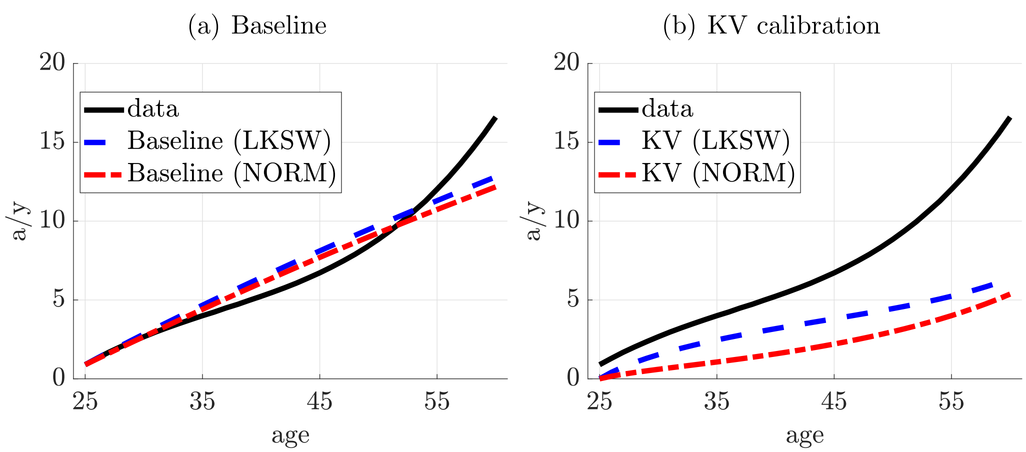
	CR	NCR
NORM	0.6817	0.6692
LK	0.7007	0.6787
LKSW	0.6866	0.6758

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

C.3 Asset Profile

Figure C.3 shows the profile of the asset to income ratio over the working life obtained from our PSID sample (the black solid line in both panels). Panels (a) and (b) show the corresponding model profiles under two distribution scenarios in the baseline and KV calibrations, respectively. In panel (a), the blue dashed line corresponds to scenario LKSW, where the discount rate is chosen to match the life cycle profile given θ. The aggregate asset to income ratio obtained in the calibration is 5.4. The red line shows the profile under scenario NORM (given the LKSW calibration of model parameters). In panel (b), the blue dashed line again shows the profile under scenario LKSW, where in the KV calibration the discount rate is chosen to match an aggregate asset-to-income ratio of 2.5. Again, the red line shows the corresponding profile in scenario NORM (given the LKSW calibration of model parameters). We show the model profiles for $\theta = 4$; the other ones are virtually the same.

Figure C.3: Asset to Income Ratio over the Life Cycle



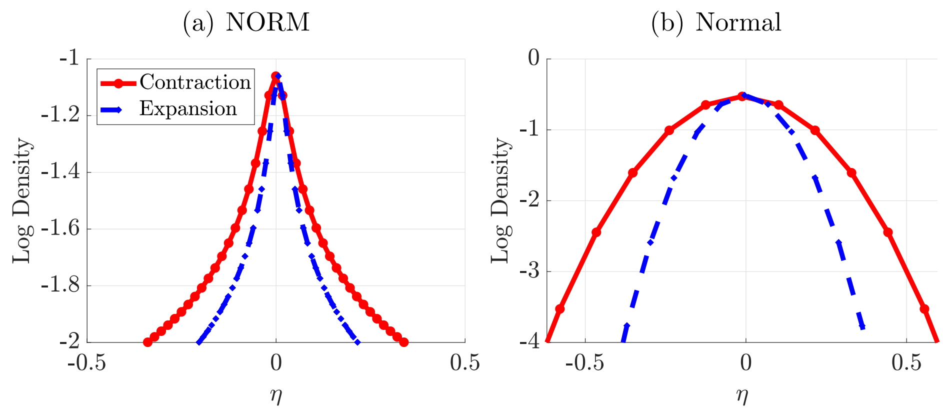
Notes: Shows empirical asset to income ratio over the working life together with the corresponding model profiles implied under the two scenarios LKSW and NORM for the baseline calibration (panel (a)) and the KV calibration (panel (b)) for $\theta = 4$.

D Additional Quantitative 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

	g_c	g_c^{mean}	g_c^{lcd}	g_c^{csd}
Risk Aversion, $\theta = 1$				
NORM: cyclical risk	-1.918	0.36	-2.142	-0.136
NORMAL: cyclical risk	-1.920	0.36	-2.144	-0.136
Risk Aversion, $\theta = 2$				
NORM: cyclical risk	-3.622	0.651	-3.998	-0.274
NORMAL: cyclical risk	-3.628	0.651	-4.003	-0.276
Risk Aversion, $\theta = 3$				
NORM: cyclical risk	-5.106	0.894	-5.594	-0.406
NORMAL: cyclical risk	-5.116	0.895	-5.602	-0.409
Risk Aversion, $\theta = 4$				
NORM: cyclical risk	-6.373	1.108	-6.961	-0.520
NORMAL: cyclical risk	-6.386	1.109	-6.970	-0.525

Notes: Welfare gains (positive numbers) and losses (negative numbers) of cyclical idiosyncratic risk expressed as consumption equivalent variation (CEV) in percentages 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 Separating the Role of Kurtosis

Tables D.2 and D.3 provide 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 4.9 percentage points higher in this distribution scenario than in scenario NORM. Table D.3 shows the pass-through and its decompostion under scenario LK.

Table D.2: The Welfare Effects of Cyclical Idiosyncratic Risk for Distribution Scenario LK

	g_c	g_c^{mean}	g_c^{lcd}	g_c^{csd}	Δg_c
		Risk Aversion, $\theta = 1$			
NORM→LK	-1.145	0.260	-1.302	-0.103	-
LK: cyclical risk	-2.218	0.414	-2.454	-0.178	-0.300
		Risk Aversion, $\theta = 2$			
NORM→LK	-1.733	0.294	-1.858	-0.169	-
LK: cyclical risk	-3.992	0.662	-4.318	-0.336	-0.37
		Risk Aversion, $\theta = 3$			
NORM→LK	-4.438	0.478	-4.509	-0.407	-
LK: cyclical risk	-6.797	0.949	-7.152	-0.595	-1.691
		Risk Aversion, $\theta = 4$			
NORM→LK	-10.909	0.925	-10.787	-1.047	-
LK: cyclical risk	-11.258	1.293	-11.496	-1.055	-4.885

Notes: Welfare gains (positive numbers) and losses (negative numbers) of cyclical idiosyncratic risk expressed as Consumption Equivalent Variation (CEV) in percentages 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.

Table D.3: Aggregate Pass-Through and its Decomposition for Scenario LK, $\theta = 4$

		components of $1-\phi$*:*			*conditional* $1-\phi$*:*	
Transitory:	$1-\phi^{\varepsilon}$	$\frac{\pi^- A^{\varepsilon}}{1-\phi^{\varepsilon}}$	$\frac{\pi^+ B^{\varepsilon}}{1-\phi^{\varepsilon}}$	$\frac{C^{\varepsilon}}{1-\phi^{\varepsilon}}$	$1-\phi^{\varepsilon}\|\varepsilon^-$	$1-\phi^{\varepsilon}\|\varepsilon^+$
LK	0.086	-0.029	0.886	0.143	0.027	0.172
Persistent:	$1-\phi^{\eta}$	$\frac{\pi^- A^{\eta}}{1-\phi^{\eta}}$	$\frac{\pi^+ B^{\eta}}{1-\phi^{\eta}}$	$\frac{C^{\eta}}{1-\phi^{\eta}}$	$1-\phi^{\eta}\|\eta^-$	$1-\phi^{\eta}\|\eta^+$
LK	0.438	0.381	0.596	0.023	0.368	0.539

Notes: Column 1 shows the aggregate consumption pass-through coefficient, columns 2-4 its decomposition into components according to equation (17), expressed as shares of total pass-through. π^- and π^+ are short for the probabilities of the shock being negative or positive. Columns 5 and 6 show conditional pass-through coefficients for negative and positive shocks.

D.3 Sensitivity Analyses

Calibration of θ. In Section 4.4 we report the analysis of the insurance coefficient for the benchmark calibration of $\theta = 4$. Results for other calibrations of θ turn out to deliver quantitatively virtually the same results. In Table D.4, we report the aggregate pass-through to consumption of transitory and persistent shocks, and the decomposition into the relative importance of positive and negative shocks for $\theta = 2$.

Table D.4: Aggregate Pass-Through and its Decomposition, $\theta = 2$

Transitory:	$1-\phi^{\varepsilon}$	*components of* $1-\phi$: $\frac{\pi^- A^{\varepsilon}}{1-\phi^{\varepsilon}}$	$\frac{\pi^+ B^{\varepsilon}}{1-\phi^{\varepsilon}}$	$\frac{C^{\varepsilon}}{1-\phi^{\varepsilon}}$	*conditional* $1-\phi$: $1-\phi^{\varepsilon}\vert\varepsilon^-$	$1-\phi^{\varepsilon}\vert\varepsilon^+$
NORM	0.046	-0.07	0.925	0.145	0.037	0.057
LKSW	0.035	0.322	0.524	0.154	0.023	0.074
Persistent:	$1-\phi^{\eta}$	$\pi^- A^{\eta}$	$\pi^+ B^{\eta}$	C^{η}	$1-\phi^{\eta}\vert\eta^-$	$1-\phi^{\eta}\vert\eta^+$
NORM	0.455	0.411	0.574	0.015	0.472	0.500
LKSW	0.426	0.540	0.444	0.015	0.374	0.537

Notes: Column 1 shows the aggregate consumption pass-through coefficient, columns 2-4 its decomposition into components according to equation (17), expressed as shares of total pass-through. π^- and π^+ are short for the probabilities of the shock being negative or positive. Columns 5 and 6 show conditional pass-through coefficients for negative and positive shocks.

Alternative Calibrations. Turning to more substantial calibration differences, 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.5 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 $\rho \in \{0.0093, 0.0025, -0.0005\}$ 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 $\theta = 4$.

Column 2 of Table D.5 summarizes the results on the welfare effects of cyclical idiosyncratic risk for this alternative choice of preferences. In comparison to Table 3 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

Table D.5: Welfare Analysis: Sensitivity Analyses

	Baseline	CRRA	ASS=0	RECAL	KV	KV,ASS>0
			Risk Aversion, $\theta = 1$			
NORM→LKSW	0.384	-0.629	-0.629	-	-0.328	0.058
NORM: cyclical risk	-1.918	-1.918	-2.029	-1.920	-1.506	-2.153
LKSW: cyclical risk	-1.620	-1.620	-1.780	-1.620	-1.562	-1.979
			Risk Aversion, $\theta = 2$			
NORM→LKSW	-0.086	-2.547	-2.547	-	-4.008	-2.628
NORM: cyclical risk	-3.622	-2.863	-3.858	-3.626	-2.472	-3.961
LKSW: cyclical risk	-3.943	-2.902	-4.663	-3.943	-4.397	-5.289
			Risk Aversion, $\theta = 3$			
NORM→LKSW	-2.183	-10.185	-10.185	-	-13.537	-10.846
NORM: cyclical risk	-5.106	-3.754	-5.459	-5.095	-2.949	-5.143
LKSW: cyclical risk	-8.128	-5.108	-10.617	-8.128	-10.479	-10.466
			Risk Aversion, $\theta = 4$			
NORM→LKSW	-6.976	-24.265	-24.265	-	-25.553	-21.028
NORM: cyclical risk	-6.373	-4.594	-6.800	-6.322	-3.081	-5.480
LKSW: cyclical risk	-13.976	-8.843	-19.232	-13.976	-16.958	-14.881

Notes: Welfare gains (positive numbers) and losses (negative numbers) of higher-order income risk, expressed as a Consumption Equivalent Variation (CEV) in percentages in scenario NORM that makes households indifferent to the higher-order income risk scenario LKSW. Also: Consumption Equivalent Variation in the non-cyclical scenario that makes households indifferent to the cyclical scenario. 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.

the welfare analysis.

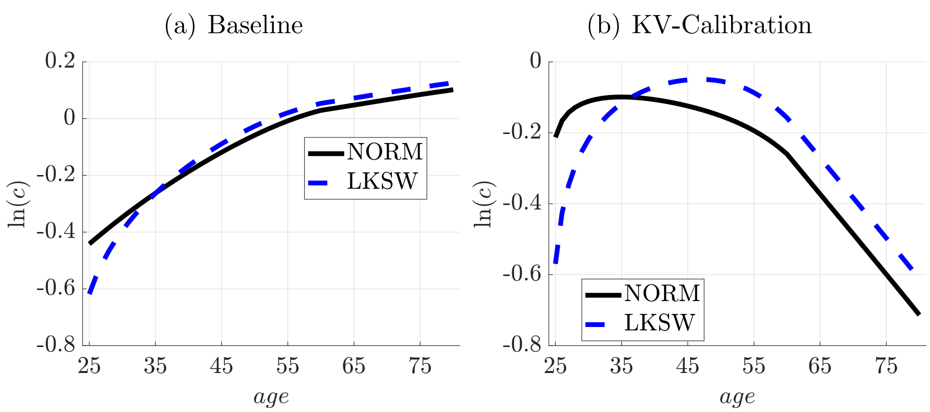
Decomposing Baseline vs. KV-Calibration. In our baseline calibration households start their economic life with positive assets and calibrated impatience is relatively strong. As a consequence, very few households are borrowing constrained (numerically, the fraction is basically zero in all scenarios). 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.5 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 -12.4 percentage points, compared to -7.6 percentage points 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 9.4 percentage points compared to 13.9 percentage points 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.

Figure D.2 shows the mean consumption profile in the KV calibration and compares it to the one from the baseline calibration.

Figure D.2: Age Profile of Consumption: Baseline vs. KV-Calibration ($\theta = 4$)



Notes: Shows life cycle profile of average log consumption under 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.

References

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.