# Consumption Partial Insurance in the Presence of Tail Income Risk

ONLINE APPENDIX*

Anisha Ghosh Alexandros Theloudis

## Contents

**A Derivation of the consumption function** **2**
A.1 Taylor approximation to optimality conditions . . . . . 2
A.2 Taylor approximation to lifetime budget constraint . . . . . 3
A.3 Biennial consumption growth . . . . . 6

**B Identification details** **7**
B.1 Income process parameters . . . . . 8
B.2 Linear consumption function parameters . . . . . 9
B.3 Quadratic consumption function parameters . . . . . 10
B.4 Targeted moments . . . . . 14

**C Empirical details and additional results** **18**
C.1 Summary statistics in the PSID sample . . . . . 18
C.2 Comparison with administrative moments . . . . . 18
C.3 Sources of tail income shocks . . . . . 20
C.4 CEX data and imputation in the PSID . . . . . 20
C.5 Estimation using the original BPP data . . . . . 26
C.6 Additional results . . . . . 28

**D Some welfare costs of tail income risk** **38**

*Ghosh: Desautels Faculty of Management, McGill University; Theloudis (corresponding author): Department of Econometrics & Operations Research, Tilburg University, email: a.theloudis@gmail.com.

# A Derivation of the consumption function

This appendix shows how to derive the consumption function under the general AR(1) specification for log income in section 6 of the paper; the consumption function based on the baseline permanent-transitory income process is a special case of this. We follow similar derivations in Blundell et al. (2008), BPP hereafter, though our approach is a refinement of theirs given the second-order approximation we carry out herein and the relaxation of the unit root assumption for income. We show below how to derive the most general (quadratic) consumption function; the linear function is a special case of this. We let a time unit in the model correspond to one calendar year. However, data in the PSID in recent years are available only every second year, so we adapt the derivations to reflect this.

Let the utility function take the form $U(C_{it}; \mathbf{Z}_{it}) = \widetilde{U}(C_{it} \exp(-\mathbf{Z}'_{it}\boldsymbol{\alpha}))$, which simplifies the subsequent statements. The first-order conditions of the household problem (1) *s.t.* (2) in a generic period $t+s$ are

$$\begin{aligned} [C_{it+s}]: \quad & \widetilde{U}_C(\widetilde{C}_{it+s}) \exp(-\mathbf{Z}'_{it+s}\boldsymbol{\alpha}) = \lambda_{it+s} \\ [A_{it+s+1}]: \quad & \beta(1+r)\mathbb{E}_{t+s}\lambda_{it+s+1} = \lambda_{it+s}, \end{aligned}$$

where $\widetilde{U}_C$ is the marginal utility of consumption and $\widetilde{C}_{it} = C_{it} \exp(-\mathbf{Z}'_{it}\boldsymbol{\alpha})$. $\lambda_{it}$ is the Lagrange multiplier on the sequential budget constraint (the marginal utility of wealth).

## A.1 Taylor approximation to optimality conditions

**Approximation to static optimality condition.** Applying logs to the static condition and taking a first difference over time yields $\Delta \ln \widetilde{U}_C(\widetilde{C}_{it+s}) - \Delta(\mathbf{Z}'_{it+s}\boldsymbol{\alpha}) = \Delta \ln \lambda_{it+s}$. A first-order Taylor expansion of $\ln \widetilde{U}_C(\widetilde{C}_{it+s})$ around $\widetilde{C}_{it+s-1}$ yields

$$\begin{aligned} \Delta \ln \widetilde{U}_C(\widetilde{C}_{it+s}) &\approx \frac{\widetilde{U}_{CC}(\widetilde{C}_{it+s-1})}{\widetilde{U}_C(\widetilde{C}_{it+s-1})} \exp(-\mathbf{Z}'_{it+s-1}\boldsymbol{\alpha})(C_{it+s} - C_{it+s-1}) \\ &\approx \frac{\widetilde{U}_{CC}(\widetilde{C}_{it+s-1})}{\widetilde{U}_C(\widetilde{C}_{it+s-1})} C_{it+s-1} \exp(-\mathbf{Z}'_{it+s-1}\boldsymbol{\alpha}) \Delta \ln C_{it+s} = \theta^{-1}_{it+s-1} \Delta \ln C_{it+s}, \end{aligned}$$

where $\widetilde{U}_{CC}$ is the second derivative of the utility function and $\theta_{it} = \widetilde{U}^{-1}_{CC}(\widetilde{C}_{it})\widetilde{C}^{-1}_{it}\widetilde{U}_C(\widetilde{C}_{it})$ is the inverse of the consumption substitution elasticity. Plugging the last expression in the log-linearized first-order condition yields $\Delta \ln C_{it+s} - \theta_{it+s-1}\Delta(\mathbf{Z}'_{it+s}\boldsymbol{\alpha}) = \theta_{it+s-1}\Delta \ln \lambda_{it+s}$.

**Approximation to Euler equation.** The approximation to the intertemporal condition

involves future expectations. Let $\exp(\Gamma) = \beta^{-1}(1+r)^{-1}$ for some $\Gamma$. A second-order Taylor expansion of $\exp(\ln \lambda_{it+s+1})$ around $\ln \lambda_{it+s} + \Gamma$ yields

$$\lambda_{it+s+1} \approx \frac{\lambda_{it+s}}{\beta(1+r)} \left(1 + (\Delta \ln \lambda_{it+s+1} - \Gamma) + \frac{1}{2}(\Delta \ln \lambda_{it+s+1} - \Gamma)^2\right).$$

Taking expectations at time $t+s$ and noting that $\mathbb{E}_{t+s}\lambda_{it+s+1} = \lambda_{it+s}\beta^{-1}(1+r)^{-1}$, we obtain

$$\mathbb{E}_{t+s}\Delta \ln \lambda_{it+s+1} \approx \Gamma - \frac{1}{2}\mathbb{E}_{t+s}(\Delta \ln \lambda_{it+s+1} - \Gamma)^2 \Rightarrow \Delta \ln \lambda_{it+s+1} \approx \omega_{it+s+1} + \epsilon_{it+s+1}.$$

The first term $\omega_{it+s+1} = \Gamma - \frac{1}{2}\mathbb{E}_t(\Delta \ln \lambda_{it+s+1} - \Gamma)^2$ captures the impact of the interest rate, impatience, and precautionary motives on the growth of the marginal utility of wealth. To maintain tractability, we assume that $\omega$ is non-stochastic but possibly heterogeneous across households. The second term is an expectation error with $\mathbb{E}_t(\epsilon_{it+1}) = 0$; it captures idiosyncratic revisions to the marginal utility of wealth upon arrival of income shocks.

Combining the approximations to the optimality conditions, we obtain

$$\begin{aligned} \Delta \ln C_{it+s} - \theta_{it+s-1}\Delta(\mathbf{Z}'_{it+s}\boldsymbol{\alpha}) &= \theta_{it+s-1}(\omega_{it+s} + \epsilon_{it+s}) \\ \Delta \ln C_{it+s} - \theta_{it+s-1}\Delta(\mathbf{Z}'_{it+s}\boldsymbol{\alpha}) - \Xi_{it+s} &= \xi_{it+s} + \theta_{it+s-1}\epsilon_{it+s} \\ \Delta c_{it+s} &= \xi_{it+s} + \theta_{it+s-1}\epsilon_{it+s}, \end{aligned} \tag{A.1}$$

where $\Delta c_{it+s} = \Delta \ln C_{it+s} - \theta_{it+s-1}\Delta(\mathbf{Z}'_{it+s}\boldsymbol{\alpha}) - \Xi_{it+s}$. We have split $\theta_{it+s-1}\omega_{it+s}$ into $\Xi_{it+s}$ and $\xi_{it+s}$; the first term reflects the gradient of consumption due to impatience (discounting) and the interest rate, the second term reflects unobserved consumption taste heterogeneity. We let demographics and time fixed effects pick up the effect of taste shifters and the gradient in the consumption path. Therefore $\Delta c_{it+s}$ can be obtained empirically from a regression of consumption growth on appropriate observables (demographics, time fixed effects).

## A.2 Taylor approximation to lifetime budget constraint

Let $G(\boldsymbol{\xi}) = \ln \sum_{s=0}^{N} \exp \xi_s$ for $\boldsymbol{\xi} = (\xi_0, \xi_1, \ldots, \xi_N)'$. Applying a second-order Taylor expansion of $G(\boldsymbol{\xi})$ around a deterministic $\boldsymbol{\xi}^0$, and taking expectations given information $\mathcal{I}$, yields

$$\begin{aligned} \mathbb{E}_{\mathcal{I}} G(\boldsymbol{\xi}) \approx G(\boldsymbol{\xi}^0) &+ \sum_{s=0}^{N} \frac{\exp \xi_s^0}{\sum_{\kappa=0}^{N} \exp \xi_\kappa^0} \mathbb{E}_{\mathcal{I}}(\xi_s - \xi_s^0) \\ &+ \frac{1}{2}\sum_{s=0}^{N}\sum_{\ell=0}^{N} \frac{\exp \xi_s^0}{\sum_{\kappa=0}^{N} \exp \xi_\kappa^0} \left(\delta_{s\ell} - \frac{\exp \xi_\ell^0}{\sum_{\kappa=0}^{N} \exp \xi_\kappa^0}\right) \mathbb{E}_{\mathcal{I}}(\xi_s - \xi_s^0)(\xi_\ell - \xi_\ell^0), \end{aligned} \tag{A.2}$$

where $\delta_{s\ell}$ is the Kronecker delta ($\delta_{s\ell} = 1$ if $s = \ell$, $\delta_{s\ell} = 0$ otherwise). We apply this approximation to both sides of the intertemporal budget constraint.

Assuming expectations away, the logarithm of the *left* hand side of the budget constraint (2) in a generic period $t$ is $\ln\left(A_{it} + \sum_{s=0}^{T-t} \exp\left(\ln Y_{it+s} - s\ln(1+r)\right)\right)$. Letting

$$\xi_s = \begin{cases} \ln A_{it+s} & \text{for } s = 0 \\ \ln Y_{it+s-1} - (s-1)\ln(1+r) & \text{for } s = 1, \ldots, T-t+1 \end{cases}$$

$$\xi_s^0 = \begin{cases} \mathbb{E}_{t-2} \ln A_{it+s} & \text{for } s = 0 \\ \mathbb{E}_{t-2} \ln Y_{it+s-1} - (s-1)\ln(1+r) & \text{for } s = 1, \ldots, T-t+1, \end{cases}$$

and declaring the information set to contain time $t$ information, we follow (A.2) to write

$$\begin{aligned}
&\mathbb{E}_t \ln\left(A_{it} + \sum_{s=0}^{T-t} \frac{Y_{it+s}}{(1+r)^s}\right) \approx \ln\left(\exp(\mathbb{E}_{t-2}\ln A_{it}) + \sum_{s=0}^{T-t} \exp\left(\mathbb{E}_{t-2}\ln Y_{it+s} - s\ln(1+r)\right)\right) \\
&+ \pi_{it}\mathbb{E}_t(\ln A_{it} - \mathbb{E}_{t-2}\ln A_{it}) \\
&+ (1-\pi_{it}) \sum_{s=0}^{T-t} \vartheta_{it+s}^Y \mathbb{E}_t(\ln Y_{it+s} - \mathbb{E}_{t-2}\ln Y_{it+s}) \\
&+ \frac{1}{2}\pi_{it}(1-\pi_{it})\mathbb{E}_t(\ln A_{it} - \mathbb{E}_{t-2}\ln A_{it})^2 \\
&+ \frac{1}{2}(1-\pi_{it}) \sum_{s=0}^{T-t}\sum_{\ell=0}^{T-t} \vartheta_{it+s}^Y(\delta_{s\ell} - (1-\pi_{it})\vartheta_{it+\ell}^Y)\mathbb{E}_t(\ln Y_{it+s} - \mathbb{E}_{t-2}\ln Y_{it+s})(\ln Y_{it+\ell} - \mathbb{E}_{t-2}\ln Y_{it+\ell}) \\
&- \pi_{it}(1-\pi_{it}) \sum_{s=0}^{T-t} \vartheta_{it+s}^Y \mathbb{E}_t(\ln A_{it} - \mathbb{E}_{t-2}\ln A_{it})(\ln Y_{it+s} - \mathbb{E}_{t-2}\ln Y_{it+s}).
\end{aligned}$$

The notation is: $\pi_{it} = \frac{Q_{1t}}{Q_{1t}+Q_{2t}}$ is equal to the expected share of financial wealth $Q_{1t} = \exp(\mathbb{E}_{t-2}\ln A_{it})$ in the household's total financial and human wealth at $t$, the latter defined as $Q_{2t} = \sum_{\kappa=0}^{T-t} \exp\left(\mathbb{E}_{t-2}\ln Y_{it+\kappa} - \kappa\ln(1+r)\right)$, i.e., the household's expected lifetime income for the remaining of life; $\vartheta_{it+s}^Y = \exp\left(\mathbb{E}_{t-2}\ln Y_{it+s} - s\ln(1+r)\right)/Q_{2t}$ is an annuitization factor equal to the expected share of time $t+s$ household income in total lifetime income. $\pi_{it}$ and $\vartheta_{it+s}^Y$ pertain to expectations at $t-2$ so they are both known at $t-2$ and later.

Using the income process and assuming the deterministic observables away (they cancel out in a first difference in expectations at a later stage), we reach $\ln Y_{it+s} = \rho^{s+2}\ln Y_{it-2} + \sum_{\tau=-1}^{s} \rho^{s-\tau}\zeta_{it+\tau} + v_{it+s} - \rho^{s+2}v_{it-2}$, therefore $\ln Y_{it+s} - \mathbb{E}_{t-2}\ln Y_{it+s} = \sum_{\tau=-1}^{s} \rho^{s-\tau}\zeta_{it+\tau} + v_{it+s}$, given that $s \geq -1$. It follows that $\sum_{s=0}^{T-t} \vartheta_{it+s}^Y \mathbb{E}_t(\ln Y_{it+s} - \mathbb{E}_{t-2}\ln Y_{it+s}) = \Theta_{it}^Y(\zeta_{it} + \rho\zeta_{it-1}) + \vartheta_{it}^Y v_{it}$, where $\Theta_{it}^Y = \sum_{\kappa=0}^{T-t} \vartheta_{it+\kappa}^Y \rho^\kappa$, because $\mathbb{E}_t\zeta_{it+\kappa} = \mathbb{E}_t v_{it+\kappa} = 0$ for $\kappa > 0$. The second-order

term in income is more involved. This is given by

$$\begin{aligned}
&\sum_{s=0}^{T-t}\sum_{\ell=0}^{T-t}\vartheta^Y_{it+s}(\delta_{s\ell}-(1-\pi_{it})\vartheta^Y_{it+\ell})\mathbb{E}_t(\ln Y_{it+s}-\mathbb{E}_{t-2}\ln Y_{it+s})(\ln Y_{it+\ell}-\mathbb{E}_{t-2}\ln Y_{it+\ell})\\
&=\sum_{s=0}^{T-t}\sum_{\ell=0}^{T-t}\vartheta^Y_{it+s}(\delta_{s\ell}-(1-\pi_{it})\vartheta^Y_{it+\ell})\mathbb{E}_t\left(\rho^s\rho^\ell(\zeta_{it}+\rho\zeta_{it-1})^2+(\zeta_{it}+\rho\zeta_{it-1})(\rho^s v_{it+\ell}+\rho^\ell v_{it+s})+v_{it+s}v_{it+\ell}\right)\\
&=\widetilde{\pi}_{it}(\zeta_{it}+\rho\zeta_{it-1})^2+2\widetilde{\widetilde{\pi}}_{it}\vartheta^Y_{it}(\zeta_{it}+\rho\zeta_{it-1})v_{it}+\vartheta^Y_{it}(1-(1-\pi_{it})\vartheta^Y_{it})v^2_{it},
\end{aligned}$$

where $\widetilde{\pi}_{it}=\widetilde{\Theta}^Y_{it}-(1-\pi_{it})(\Theta^Y_{it})^2$, $\widetilde{\Theta}^Y_{it}=\sum_{\kappa=0}^{T-t}\vartheta^Y_{it+\kappa}\rho^{2\kappa}$, and $\widetilde{\widetilde{\pi}}_{it}=1-(1-\pi_{it})\Theta^Y_{it}$. We derive this noting that the expectation of future shocks is zero by construction.

Plugging these results in the approximation of the previous page, noting that assets at $t$ are 'beginning-of-period' and thus independent of the time $t$ income shocks, and taking a first difference in expectations between $t-2$ and $t$ yields

$$\begin{aligned}
\mathbb{E}_t\ln\left(A_{it}+\sum_{s=0}^{T-t}\frac{Y_{it+s}}{(1+r)^s}\right)-\mathbb{E}_{t-2}\ln\left(A_{it}+\sum_{s=0}^{T-t}\frac{Y_{it+s}}{(1+r)^s}\right)&\approx(1-\pi_{it})\Theta^Y_{it}(\zeta_{it}+\rho\zeta_{it-1})\\
&+(1-\pi_{it})\vartheta^Y_{it}v_{it}\\
&+\frac{1}{2}(1-\pi_{it})\widetilde{\pi}_{it}(\zeta_{it}+\rho\zeta_{it-1})^2\\
&+\frac{1}{2}(1-\pi_{it})\vartheta^Y_{it}(1-(1-\pi_{it})\vartheta^Y_{it})v^2_{it}\\
&+(1-\pi_{it})\widetilde{\widetilde{\pi}}_{it}\vartheta^Y_{it}(\zeta_{it}+\rho\zeta_{it-1})v_{it}.
\end{aligned}$$

The linear terms in the first two lines are the first-order approximation of BPP (adapted here for biennial data and income persistence); the quadratic terms in the next lines are the refinement from the second-order approximation.

Assuming expectations away, the logarithm of the *right* hand side of the budget constraint (2) is $\ln\sum_{s=0}^{T-t}\exp(\ln C_{it+s}-s\ln(1+r))$. Adopting the notation of (A.2), we let $\xi_s=\ln C_{it+s}-s\ln(1+r)$ and $\xi^0_s=\mathbb{E}_{t-2}\ln C_{it+s}-s\ln(1+r)$. Letting the information set contain time $t$ information, it follows that

$$\begin{aligned}
\mathbb{E}_t\ln\sum_{s=0}^{T-t}\frac{C_{it+s}}{(1+r)^s}&\approx\ln\sum_{s=0}^{T-t}\exp(\mathbb{E}_{t-2}\ln C_{it+s}-s\ln(1+r))\\
&+\sum_{s=0}^{T-t}\vartheta^C_{it+s}\mathbb{E}_t(\ln C_{it+s}-\mathbb{E}_{t-2}\ln C_{it+s})\\
&+\frac{1}{2}\sum_{s=0}^{T-t}\sum_{\ell=0}^{T-t}\vartheta^C_{it+s}(\delta_{s\ell}-\vartheta^C_{it+\ell})\mathbb{E}_t(\ln C_{it+s}-\mathbb{E}_{t-2}\ln C_{it+s})(\ln C_{it+\ell}-\mathbb{E}_{t-2}\ln C_{it+\ell}),
\end{aligned}$$

where $\vartheta^C_{it+s} = \exp\left(\mathbb{E}_{t-2}\ln C_{it+s} - s\ln(1+r)\right) / \sum_{\kappa=0}^{T-t}\exp\left(\mathbb{E}_{t-2}\ln C_{it+\kappa} - \kappa\ln(1+r)\right)$ is an annuitization factor equal to the expected share of time $t+s$ consumption in total lifetime consumption. $\vartheta^C_{it+s}$ pertains to expectations at $t-2$, so it is known at $t-2$ and later.

Using the linearized consumption function in (A.1) to replace $\ln C_{it+s}$ recursively, we reach $\ln C_{it+s} = \ln C_{it-2} + \sum_{\tau=-1}^{s}\Psi_{it+\tau} + \sum_{\tau=-1}^{s}\xi_{it+\tau} + \sum_{\tau=-1}^{s}\theta_{it+\tau-1}\epsilon_{it+\tau}$, where $\Psi_{it+\tau} = \theta_{it+\tau-1}\Delta(\mathbf{Z}'_{it+\tau}\boldsymbol{\alpha}) + \Xi_{it+\tau}$, therefore $\ln C_{it+s} - \mathbb{E}_{t-2}\ln C_{it+s} = \sum_{\tau=-1}^{s}\theta_{it+\tau-1}\epsilon_{it+\tau}$ because $\Psi_{it}$ and $\xi_{it}$ are non-stochastic by assumption and they cancel out in the first difference. It follows that $\sum_{s=0}^{T-t}\vartheta^C_{it+s}\mathbb{E}_t(\ln C_{it+s} - \mathbb{E}_{t-2}\ln C_{it+s}) = \theta_{it-1}\epsilon_{it} + \theta_{it-2}\epsilon_{it-1}$ because $\mathbb{E}_t\epsilon_{it+\kappa} = 0$ for $\kappa > 0$ and $\sum_{s=0}^{T-t}\vartheta^C_{it+s} = 1$ by construction. The second-order term is more involved; it is given by

$$\begin{aligned}
&\sum_{s=0}^{T-t}\sum_{\ell=0}^{T-t}\vartheta^C_{it+s}(\delta_{s\ell} - \vartheta^C_{it+\ell})\mathbb{E}_t(\ln C_{it+s} - \mathbb{E}_{t-2}\ln C_{it+s})(\ln C_{it+\ell} - \mathbb{E}_{t-2}\ln C_{it+\ell}) \\
&= \sum_{s=0}^{T-t}\sum_{\ell=0}^{T-t}\vartheta^C_{it+s}(\delta_{s\ell} - \vartheta^C_{it+\ell})(\theta_{it-1}\epsilon_{it} + \theta_{it-2}\epsilon_{it-1})^2 = 0.
\end{aligned}$$

We derive this by noting that future expectations errors are zero by assumption, and that $\sum_{s=0}^{T-t}\sum_{\ell=0}^{T-t}\vartheta^C_{it+s}(\delta_{s\ell} - \vartheta^C_{it+\ell}) = \sum_{s=0}^{T-t}\vartheta^C_{it+s}\delta_{ss} + \sum_{s=0}^{T-t}\sum_{\ell\neq s}\vartheta^C_{it+s}\delta_{s\ell} - (\sum_{s=0}^{T-t}\vartheta^C_{it+s})(\sum_{\ell=0}^{T-t}\vartheta^C_{it+\ell}) = 0$. Plugging these results in the approximation above and taking a first difference in expectations between $t-2$ and $t$ yields

$$\mathbb{E}_t\ln\sum_{s=0}^{T-t}\frac{C_{it+s}}{(1+r)^s} - \mathbb{E}_{t-2}\ln\sum_{s=0}^{T-t}\frac{C_{it+s}}{(1+r)^s} \approx \theta_{it-1}\epsilon_{it} + \theta_{it-2}\epsilon_{it-1}.$$

## A.3 Biennial consumption growth

Given the biennial nature of our data, it follows from (A.1) that *observed* consumption growth is $\Delta^2 c_{it+s} = \xi_{it+s} + \xi_{it+s-1} + \theta_{it+s-1}\epsilon_{it+s} + \theta_{it+s-2}\epsilon_{it+s-1}$. The notation $\Delta^2 x_t = \Delta x_t + \Delta x_{t-1} = x_t - x_{t-2}$ indicates a difference between $t$ and $t-2$. Dropping the index $s$ and replacing the expectations errors with the quadratic expression in income shocks (the budget must balance so we bring the two sides of the budget constraint together), we obtain the final expression for consumption growth with biennial survey data and income persistence, given by

$$\begin{aligned}
\Delta^2 c_{it} \approx\ & \xi_{it} + \xi_{it-1} + (1-\pi_{it})\Theta^Y_{it}(\zeta_{it} + \rho\zeta_{it-1}) + (1-\pi_{it})\vartheta^Y_{it}v_{it} \\
&+ \frac{1}{2}(1-\pi_{it})\widetilde{\pi}_{it}(\zeta_{it} + \rho\zeta_{it-1})^2 + \frac{1}{2}(1-\pi_{it})\vartheta^Y_{it}(1-(1-\pi_{it})\vartheta^Y_{it})v^2_{it} \\
&+ (1-\pi_{it})\widetilde{\widetilde{\pi}}_{it}\vartheta^Y_{it}(\zeta_{it} + \rho\zeta_{it-1})v_{it}.
\end{aligned} \tag{A.3}$$

The first line of (A.3) corresponds to the linear consumption function, adapted here for

income persistence and the biennial nature of the data, and given compactly by

$$\Delta^2 c_{it} = \xi_{it} + \xi_{it-1} + \phi_{it}^{(1)} \left(\zeta_{it} + \rho\zeta_{it-1}\right) + \psi_{it}^{(1)} v_{it}.$$

The transmission parameters of income shocks are given by $\phi_{it}^{(1)} = (1-\pi_{it})\Theta_{it}^{Y}$ and $\psi_{it}^{(1)} = (1-\pi_{it})\vartheta_{it}^{Y}$. In the baseline, the permanent-transitory income process has $\rho = 1$ and further restricts $\phi_{it}^{(1)} = 1 - \pi_{it}$ because $\Theta_{it}^{Y} = \sum_{s=0}^{T-t} \vartheta_{it+s}^{Y} = 1$.

The quadratic terms in the next two lines of (A.3) reflect the refinement from the higher-order approximation. This leads to the quadratic consumption function, adapted here for income persistence and the biennial nature of the data, and given compactly by

$$\Delta^2 c_{it} = \xi_{it}+\xi_{it-1}+\phi_{it}^{(1)} \left(\zeta_{it} + \rho\zeta_{it-1}\right)+\psi_{it}^{(1)} v_{it}+\phi_{it}^{(2)} \left(\zeta_{it} + \rho\zeta_{it-1}\right)^2+\psi_{it}^{(2)} v_{it}^2+\omega_{it}^{(22)} \left(\zeta_{it} + \rho\zeta_{it-1}\right) v_{it}.$$

The additional transmission parameters of income shocks are given by $\phi_{it}^{(2)} = \frac{1}{2}(1-\pi_{it})\widetilde{\pi}_{it}$, $\psi_{it}^{(2)} = \frac{1}{2}(1-\pi_{it})\vartheta_{it}^{Y}(1-(1-\pi_{it})\vartheta_{it}^{Y})$, and $\omega_{it}^{(22)} = (1-\pi_{it})\widetilde{\widetilde{\pi}}_{it}\vartheta_{it}^{Y}$. The baseline permanent-transitory income process has $\rho = 1$ and further restricts $\phi_{it}^{(2)} = \frac{1}{2}(1-\pi_{it})\pi_{it}$ and $\omega_{it}^{(22)} = (1-\pi_{it})\pi_{it}\vartheta_{it}^{Y}$, because $\widetilde{\Theta}_{it}^{Y} = \Theta_{it}^{Y} = \sum_{s=0}^{T-t} \vartheta_{it+s}^{Y} = 1$.

# B Identification details

This appendix provides detailed identification statements based on our use of the recent biennial PSID data. There are three points of departure from section 3 in the paper.

First, we use the most general AR(1) specification for income, shown in section 6. The baseline permanent-transitory specification is a special case with $\rho = 1$. Second, we allow for classical measurement error. We let observed income $y_{it}^{*}$ and consumption $c_{it}^{*}$ be the sum of their corresponding true value and measurement error, namely

$$y_{it}^{*} = y_{it} + u_{it}^{y} \qquad \text{and} \qquad c_{it}^{*} = c_{it} + u_{it}^{c}.$$

The moments of income measurement error are not separately identifiable from the moments of the transitory shock. This is why we assume the income error to be Gaussian (but let the consumption error unrestricted), so we specify

$$\mathbb{E}((u_{it}^{y})^m) = \begin{cases} 0 & \text{for } m = 1 \\ \sigma_{u_t^y}^2 & \text{for } m = 2 \\ 0 & \text{for } m = 3 \\ 3(\sigma_{u_t^y}^2)^2 & \text{for } m = 4 \end{cases} \qquad \text{and} \qquad \mathbb{E}((u_{it}^{c})^m) = \begin{cases} 0 & \text{for } m = 1 \\ \sigma_{u_t^c}^2 & \text{for } m = 2 \\ \gamma_{u_t^c} & \text{for } m = 3 \\ \kappa_{u_t^c} & \text{for } m = 4. \end{cases}$$

We retrieve the variance of income measurement error from the validation study of the PSID in Bound et al. (1994), as we explain in section 4. We do not need similar assumptions for the consumption measurement error, whose moments we identify below. We further assume that income and consumption errors are mutually independent (but note the discussion of consumption imputation from the CEX in appendix C.4), independent over time, and independent of the income shocks and taste heterogeneity. Third, we recast the income and consumption processes to reflect the biennial frequency of the modern PSID.

These choices lead to the theoretical expressions that we bring to the data, namely the process for income growth, the quasi-difference in income, consumption growth in the linear specification, and consumption growth in the quadratic specification, given respectively by

$$\begin{aligned}
\Delta^2 y_{it} &= (\rho^2-1)P_{it-2} + \zeta_{it} + \rho\zeta_{it-1} + \Delta^2 v_{it} + \Delta^2 u^y_{it},\\
\Delta^2_\rho y_{it} &= \qquad\qquad\quad \zeta_{it} + \rho\zeta_{it-1} + \Delta^2_\rho v_{it} + \Delta^2_\rho u^y_{it},\\
\Delta^2 c_{it} &= \xi_{it} + \xi_{it-1} \quad + \phi^{(1)}_{it}(\zeta_{it} + \rho\zeta_{it-1}) + \psi^{(1)}_{it} v_{it} + \Delta^2 u^c_{it},\\
\Delta^2 c_{it} &= \xi_{it} + \xi_{it-1} \quad + \phi^{(1)}_{it}(\zeta_{it} + \rho\zeta_{it-1}) + \psi^{(1)}_{it} v_{it}\\
&\qquad\qquad + \phi^{(2)}_{it}(\zeta_{it} + \rho\zeta_{it-1})^2 + \psi^{(2)}_{it} v^2_{it} + \omega^{(22)}_{it}(\zeta_{it} + \rho\zeta_{it-1})v_{it} + \Delta^2 u^c_{it}.
\end{aligned}$$

The notation $\Delta^2 x_t = \Delta x_t + \Delta x_{t-1} = x_t - x_{t-2}$ indicates a first difference in $x$ between $t$ and $t-2$, i.e., the observed frequency. The quasi-difference is defined as $\Delta^2_\rho x_t = \Delta_\rho x_t + \rho\Delta_\rho x_{t-1} = x_t - \rho^2 x_{t-2}$, where $\Delta_\rho x_t = x_t - \rho x_{t-1}$ is the quasi-difference between $t$ and $t-1$.

Due to the biennial nature of the data, we cannot separately identify the moments of the *yearly* permanent shocks within a two-year period. We thus assume $\sigma^2_{\zeta_t} \approx \sigma^2_{\zeta_{t-1}}$, $\gamma_{\zeta_t} \approx \gamma_{\zeta_{t-1}}$, and $\kappa_{\zeta_t} \approx \kappa_{\zeta_{t-1}}$, which are exactly true if the underlying distributions are stationary. Similarly, we assume that $\sigma^2_{\xi_t} \approx \sigma^2_{\xi_{t-1}}$. For simplicity of the next statements, we also assume $\sigma^2_{\zeta_t} \approx \sigma^2_{\zeta_{t+1}} \approx \sigma^2_{\zeta_{t+2}}$, $\sigma^2_{v_t} \approx \sigma^2_{v_{t-2}} \approx \sigma^2_{v_{t+2}}$, $\sigma^2_{u^y_t} \approx \sigma^2_{u^y_{t-2}} \approx \sigma^2_{u^y_{t+2}}$, and $\kappa_{v_t} \approx \kappa_{v_{t-2}}$. These last restrictions are not needed for identification, but they simplify the illustration.

## B.1 Income process parameters

The second and third moments of shocks are given by

$$\sigma^2_{\zeta_t} = \frac{1}{1+\rho^2}\mathbb{E}(\Delta^2_\rho y_{it} \times \sum_{\kappa=\{-2,0,2\}} \rho^{-\kappa}\Delta^2_\rho y_{it+\kappa}) \qquad \text{and} \qquad \sigma^2_{v_t} = -\frac{1}{\rho^2}\mathbb{E}(\Delta^2_\rho y_{it} \times \Delta^2_\rho y_{it+2}) - \sigma^2_{u^y_t},$$

$$\gamma_{\zeta_t} = \frac{1}{1+\rho^3}\mathbb{E}((\Delta^2_\rho y_{it})^2 \times \sum_{\kappa=\{-2,0,2\}} \rho^{-\kappa}\Delta^2_\rho y_{it+\kappa}) \qquad \text{and} \qquad \gamma_{v_t} = -\frac{1}{\rho^2}\mathbb{E}((\Delta^2_\rho y_{it})^2 \times \Delta^2_\rho y_{it+2}).$$

The fourth moments are more involved and are given by

$$\kappa_{\zeta_t} = \frac{1}{1+\rho^4}\mathbb{E}((\Delta^2_\rho y_{it})^4) - \frac{1+\rho^8}{1+\rho^4}\kappa_{v_t} - 6\frac{\rho^2}{1+\rho^4}(\sigma^2_{\zeta_t})^2 - 6\frac{\rho^4}{1+\rho^4}(\sigma^2_{v_t})^2 - 3(1+\rho^4)(\sigma^2_{u^y_t})^2$$
$$- 6(1+\rho^2)\sigma^2_{\zeta_t}\sigma^2_{v_t} - 6(1+\rho^2)\sigma^2_{\zeta_t}\sigma^2_{u^y_t} - 6(1+\rho^4)\sigma^2_{v_t}\sigma^2_{u^y_t},$$
$$\kappa_{v_t} = \frac{1}{\rho^4}\mathbb{E}((\Delta^2_\rho y_{it})^2 \times (\Delta^2_\rho y_{it+2})^2) - \frac{(1+\rho^2)^2}{\rho^4}(\sigma^2_{\zeta_t})^2 - \frac{1+\rho^4+\rho^8}{\rho^4}(\sigma^2_{v_t})^2 - \frac{1+4\rho^4+\rho^8}{\rho^4}(\sigma^2_{u^y_t})^2$$
$$- 2(1+\rho^2)\frac{1+\rho^4}{\rho^4}\sigma^2_{\zeta_t}\sigma^2_{v_t} - 2(1+\rho^2)\frac{1+\rho^4}{\rho^4}\sigma^2_{\zeta_t}\sigma^2_{u^y_t} - \left(2\frac{(1+\rho^4)^2}{\rho^4} + 4\right)\sigma^2_{v_t}\sigma^2_{u^y_t}.$$

In all cases, either we need an external estimate for $\rho$ (e.g., $\rho = 1$ as in the permanent-transitory process) or treat $\rho$ as a free parameter to be jointly estimated with the moments of income shocks. In the latter case, $\rho$ is pinned down by

$$\mathbb{E}(\Delta^2_\rho y_{it} \times \Delta^2 y_{it-4}) = \mathbb{E}((y_{it} - \rho^2 y_{it-2}) \times (y_{it-4} - y_{it-6})) = 0,$$

namely a "long" intertemporal covariance. Note that when $\rho$ is a free parameter, five time period of data are needed at a minimum (i.e., $t-6$ through $t+2$ biennially) to identify both the variance of persistent shocks and $\rho$ itself, versus four periods when $\rho$ is set externally.

## B.2 Linear consumption function parameters

Given the general income process, the linear consumption function, and the properties of shocks, the transmission parameters are given by

$$\phi^{(1)}_t = \frac{1}{(1+\rho^2)\sigma^2_{\zeta_t}}\mathbb{E}(\Delta^2 c_{it} \times \sum_{\kappa=\{-2,0,2\}} \rho^{-\kappa}\Delta^2_\rho y_{it+\kappa}),$$
$$\psi^{(1)}_t = -\frac{1}{\rho^2\sigma^2_{v_t}}\mathbb{E}(\Delta^2 c_{it} \times \Delta^2_\rho y_{it+2}).$$

The variance of taste heterogeneity and consumption measurement error are identified as

$$\sigma^2_{\xi_t} = \frac{1}{2}\mathbb{E}(\Delta^2 c_{it} \times \sum_{\kappa=\{-2,0,2\}} \Delta^2 c_{it+\kappa}) - (\phi^{(1)}_t)^2\frac{1+\rho^2}{2}\sigma^2_{\zeta_t} - \frac{1}{2}(\psi^{(1)}_t)^2\sigma^2_{v_t},$$
$$\sigma^2_{u^c_t} = -\mathbb{E}(\Delta^2 c_{it} \times \Delta^2 c_{it+2}).$$

The third and fourth moments of taste heterogeneity and consumption error are identified by higher-order consumption moments through expressions analogous to those above.

## B.3 Quadratic consumption function parameters

Given the general income process, the quadratic consumption function, and the properties of shocks, the transmission parameters are given by

$$\begin{pmatrix} \phi_t^{(1)} \\ \psi_t^{(1)} \\ \phi_t^{(2)} \\ \psi_t^{(2)} \\ \omega_t^{(22)} \end{pmatrix} = \mathbf{A}^{-1} \begin{pmatrix} \mathbb{E}(\Delta^2 c_{it} \times \Delta_\rho^2 y_{it}) \\ \mathbb{E}(\Delta^2 c_{it} \times (\Delta_\rho^2 y_{it})^2) \\ \mathbb{E}(\Delta^2 c_{it} \times \Delta_\rho^2 y_{it+2}) \\ \mathbb{E}(\Delta^2 c_{it} \times (\Delta_\rho^2 y_{it+2})^2) \\ \mathbb{E}(\Delta^2 c_{it} \times \Delta_\rho^2 y_{it} \times \Delta_\rho^2 y_{it+2}) \end{pmatrix}, \tag{B.1}$$

where the matrix of coefficients $\mathbf{A}$ depends exclusively on the second, third, and fourth moments of income shocks, the AR(1) persistence parameter $\rho$, and the variance of income measurement error.[1] Its determinant[2] is given by

$$\det(\mathbf{A}) = -\rho^4(1+\rho^2)\sigma_{\zeta_t}^4\sigma_{v_t}^4 \times$$

[1] While the paper shows matrix $\mathbf{A}$ for the case of annual data, unit root in income, and no measurement error, the general form of $\mathbf{A} = [a_{ij}]_{5\times 5}$ with *biennial* data, income persistence, and measurement error has $a_{15} = a_{31} = a_{33} = a_{35} = a_{41} = a_{45} = a_{51} = 0$, and

$$\begin{aligned}
a_{11} &= (1+\rho^2)\sigma_{\zeta_t}^2, \\
a_{12} &= \sigma_{v_t}^2, \\
a_{13} &= (1+\rho^3)\gamma_{\zeta_t}, \\
a_{14} &= \gamma_{v_t}, \\
a_{21} &= (1+\rho^3)\gamma_{\zeta_t}, \\
a_{22} &= \gamma_{v_t}, \\
a_{23} &= (1+\rho^4)\kappa_{\zeta_t} + 6\rho^2(\sigma_{\zeta_t}^2)^2 + (1+\rho^2)(1+\rho^4)\sigma_{\zeta_t}^2(\sigma_{v_t}^2 + \sigma_{u_t^y}^2), \\
a_{24} &= \kappa_{v_t} + \rho^4(\sigma_{v_t}^2)^2 + (1+\rho^2)\sigma_{\zeta_t}^2\sigma_{v_t}^2 + (1+\rho^4)\sigma_{v_t}^2\sigma_{u_t^y}^2, \\
a_{25} &= 2(1+\rho^2)\sigma_{\zeta_t}^2\sigma_{v_t}^2, \\
a_{32} &= -\rho^2\sigma_{v_t}^2, \\
a_{34} &= -\rho^2\gamma_{v_t}, \\
a_{42} &= \rho^4\gamma_{v_t}, \\
a_{43} &= (1+\rho^2)^2(\sigma_{\zeta_t}^2)^2 + (1+\rho^2)(1+\rho^4)\sigma_{\zeta_t}^2(\sigma_{v_t}^2 + \sigma_{u_t^y}^2), \\
a_{44} &= \rho^4\kappa_{v_t} + (\sigma_{v_t}^2)^2 + (1+\rho^2)\sigma_{\zeta_t}^2\sigma_{v_t}^2 + (1+\rho^4)\sigma_{v_t}^2\sigma_{u_t^y}^2, \\
a_{52} &= -\rho^2\gamma_{v_t}, \\
a_{53} &= -\rho^2(1+\rho^2)\sigma_{\zeta_t}^2(\sigma_{v_t}^2 + \sigma_{u_t^y}^2), \\
a_{54} &= -\rho^2(\kappa_{v_t} + \sigma_{v_t}^2\sigma_{u_t^y}^2), \\
a_{55} &= -\rho^2(1+\rho^2)\sigma_{\zeta_t}^2\sigma_{v_t}^2.
\end{aligned}$$

[2] The simpler derivation for $\rho = 1$ appears in older versions of the paper, see arxiv.org/abs/2306.13208.

$$
\begin{aligned}
&\Bigg\{\Bigg((1+\rho^2)(1+\rho^4)(k_{\zeta_t}-\frac{\gamma_{\zeta_t}^2}{\sigma_{\zeta_t}^2}-\sigma_{\zeta_t}^4)+(\rho^2-\rho)^2\frac{\gamma_{\zeta_t}^2}{\sigma_{\zeta_t}^2}\\
&\quad+\frac{1+\rho^2}{\rho^4}(1+2\rho^2+2\rho^4+6\rho^6+\rho^8)\sigma_{\zeta_t}^4+\frac{(1+\rho^2)^2}{\rho^4}(1+\rho^8)\sigma_{\zeta_t}^2(\sigma_{v_t}^2+\sigma_{u_t^y}^2)\Bigg)\times\\
&\quad\Bigg(\rho^4(k_{v_t}-\frac{\gamma_{v_t}^2}{\sigma_{v_t}^2}-\sigma_{v_t}^4)+(1+\rho^4)\sigma_{v_t}^4+(1+\rho^2)\sigma_{\zeta_t}^2\sigma_{v_t}^2+(1+\rho^4)\sigma_{v_t}^2\sigma_{u_t^y}^2\Bigg)\\
&\quad-\frac{(1+\rho^2)^4}{\rho^4}(1+\rho^4)\sigma_{\zeta_t}^6\sigma_{v_t}^2-\frac{(1+\rho^2)^2}{\rho^4}(1+\rho^4)(1+\rho^8)\sigma_{\zeta_t}^2\sigma_{v_t}^6\\
&\quad-2\frac{(1+\rho^2)^3}{\rho^4}(1+\rho^4+\rho^8)\sigma_{\zeta_t}^4\sigma_{v_t}^4-\frac{(1+\rho^2)^2}{\rho^4}(1+\rho^4)(1+\rho^8)\sigma_{\zeta_t}^2\sigma_{v_t}^2\sigma_{u_t^y}^4\\
&\quad-2\frac{(1+\rho^2)^2}{\rho^4}(1+\rho^4)(1+\rho^8)\sigma_{\zeta_t}^2\sigma_{v_t}^4\sigma_{u_t^y}^2-2\frac{(1+\rho^2)^3}{\rho^4}(1+\rho^4+\rho^8)\sigma_{\zeta_t}^4\sigma_{v_t}^2\sigma_{u_t^y}^2\Bigg\}.
\end{aligned}
$$

Assuming that the shocks are *not* drawn from a Bernoulli distribution, kurtosis is bounded below by $\gamma^2/\sigma^2+\sigma^4$. Given that $0<\rho\leq 1$, it follows that

$$(1+\rho^2)(1+\rho^4)(k_{\zeta_t}-\gamma_{\zeta_t}^2/\sigma_{\zeta_t}^2-\sigma_{\zeta_t}^4)>0, \tag{B.2}$$

$$\rho^4(k_{v_t}-\gamma_{v_t}^2/\sigma_{v_t}^2-\sigma_{v_t}^4)>0. \tag{B.3}$$

Note also that

$$(\rho^2-\rho)^2\frac{\gamma_{\zeta_t}^2}{\sigma_{\zeta_t}^2}\geq 0, \tag{B.4}$$

$$\frac{1+\rho^2}{\rho^4}(1+2\rho^2+2\rho^4+6\rho^6+\rho^8)\sigma_{\zeta_t}^4+\frac{(1+\rho^2)^2}{\rho^4}(1+\rho^8)\sigma_{\zeta_t}^2(\sigma_{v_t}^2+\sigma_{u_t^y}^2)>0, \tag{B.5}$$

$$(1+\rho^4)\sigma_{v_t}^4+(1+\rho^2)\sigma_{\zeta_t}^2\sigma_{v_t}^2+(1+\rho^4)\sigma_{v_t}^2\sigma_{u_t^y}^2>0, \tag{B.6}$$

as the sum of mostly strictly positive terms. Adding the inequalities (B.2)+(B.4)+(B.5) and (B.3)+(B.6) by parts, it follows that the two long bracketed terms in the first lines of the determinant are strictly positive, so their product is also strictly positive. Given the lower bound on the kurtosis, the term inside the curly brackets is then strictly greater than

$$
\begin{aligned}
&\Bigg((\rho^2-\rho)^2\frac{\gamma_{\zeta_t}^2}{\sigma_{\zeta_t}^2}+\frac{1+\rho^2}{\rho^4}(1+2\rho^2+2\rho^4+6\rho^6+\rho^8)\sigma_{\zeta_t}^4+\frac{(1+\rho^2)^2}{\rho^4}(1+\rho^8)\sigma_{\zeta_t}^2(\sigma_{v_t}^2+\sigma_{u_t^y}^2)\Bigg)\times\\
&\qquad\Bigg((1+\rho^4)\sigma_{v_t}^4+(1+\rho^2)\sigma_{\zeta_t}^2\sigma_{v_t}^2+(1+\rho^4)\sigma_{v_t}^2\sigma_{u_t^y}^2\Bigg)\\
&\qquad-\frac{(1+\rho^2)^4}{\rho^4}(1+\rho^4)\sigma_{\zeta_t}^6\sigma_{v_t}^2-\frac{(1+\rho^2)^2}{\rho^4}(1+\rho^4)(1+\rho^8)\sigma_{\zeta_t}^2\sigma_{v_t}^6
\end{aligned}
$$

$$-2\frac{(1+\rho^2)^3}{\rho^4}(1+\rho^4+\rho^8)\sigma^4_{\zeta_t}\sigma^4_{v_t} - \frac{(1+\rho^2)^2}{\rho^4}(1+\rho^4)(1+\rho^8)\sigma^2_{\zeta_t}\sigma^2_{v_t}\sigma^4_{u^y_t}$$
$$-2\frac{(1+\rho^2)^2}{\rho^4}(1+\rho^4)(1+\rho^8)\sigma^2_{\zeta_t}\sigma^4_{v_t}\sigma^2_{u^y_t} - 2\frac{(1+\rho^2)^3}{\rho^4}(1+\rho^4+\rho^8)\sigma^4_{\zeta_t}\sigma^2_{v_t}\sigma^2_{u^y_t} =$$
$$(\rho^2-\rho)^2\frac{\gamma^2_{\zeta_t}}{\sigma^2_{\zeta_t}}\left((1+\rho^4)\sigma^4_{v_t} + (1+\rho^2)\sigma^2_{\zeta_t}\sigma^2_{v_t} + (1+\rho^4)\sigma^2_{v_t}\sigma^2_{u^y_t}\right)$$
$$+4\rho^2(1+\rho^2)^2\sigma^6_{\zeta_t}\sigma^2_{v_t} + 4\rho^2(1+\rho^2)(1+\rho^4)\sigma^4_{\zeta_t}\sigma^4_{v_t} + 4\rho^2(1+\rho^2)(1+\rho^4)\sigma^4_{\zeta_t}\sigma^2_{v_t}\sigma^2_{u^y_t} > 0.$$

The last inequality follows from $0 < \rho \leq 1$ and the fact that we are adding mostly strictly positive terms (and one weakly positive, the term involving $\gamma^2_{\zeta_t}$). The expression inside the curly brackets in the determinant is thus strictly positive, so $\det(\mathbf{A}) < 0$. $\mathbf{A}$ is thus full rank and invertible for *any* distribution of permanent & transitory shocks except Bernoulli.[3]

How do different levels of skewness and kurtosis influence the transmission parameters? Consider the transmission of permanent shocks, abstracting from transitory shocks, i.e., $\psi^{(1)}_t = \psi^{(2)}_t = \omega^{(22)}_t = 0$, and let $\rho = 1$. $\phi^{(1)}_t$ and $\phi^{(2)}_t$ are pinned down by the first equation in (B.1), namely $2\sigma^2_{\zeta_t}\phi^{(1)}_t + 2\gamma_{\zeta_t}\phi^{(2)}_t = \mathbb{E}(\Delta^2 c_{it} \times \Delta^2_\rho y_{it})$. Absent transmission of transitory shocks, the right hand side reflects the strength of the covariance between consumption growth and permanent shocks. Suppose $\phi^{(1)}_t$ is fixed. We highlight three cases:

- $\mathbb{E}(\Delta^2 c_{it} \times \Delta^2_\rho y_{it}) - 2\sigma^2_{\zeta_t}\phi^{(1)}_t = 0$. Letting $\gamma_{\zeta_t} < 0$ as in the data, this equation restricts $\phi^{(2)}_t = 0$. Tail shocks thus transmit similarly to non-tail ones, i.e., this configuration identifies a *linear* consumption process.

- $\mathbb{E}(\Delta^2 c_{it} \times \Delta^2_\rho y_{it}) - 2\sigma^2_{\zeta_t}\phi^{(1)}_t > 0$. Consumption growth and permanent shocks co-vary more strongly than the linear consumption process suggests, i.e., there is excess co-movement. For $\gamma_{\zeta_t} < 0$, this configuration identifies $\phi^{(2)}_t < 0$ and implies that negative shocks transmit more strongly than other shocks. Negative shocks are often the *largest* shocks (due to left skewness), which explains the excess co-movement. Ceteris paribus, a *less skewed* distribution ($\gamma_{\zeta_t}$ drops in magnitude) implies a more negative $\phi^{(2)}_t$, because a more symmetric distribution associated with fewer large shocks requires a stronger transmission of said shocks to maintain a given excess co-movement. If $\gamma_{\zeta_t} > 0$, this configuration identifies $\phi^{(2)}_t > 0$ and analogous arguments apply to positive shocks.

- $\mathbb{E}(\Delta^2 c_{it} \times \Delta^2_\rho y_{it}) - 2\sigma^2_{\zeta_t}\phi^{(1)}_t < 0$. Consumption growth and permanent shocks co-move less strongly than the linear process suggests, i.e., there is excess smoothness. For $\gamma_{\zeta_t} < 0$, this identifies $\phi^{(2)}_t > 0$ and suggests a muted pass-through of negative shocks.

[3] If the distributions of both permanent and transitory shocks were Bernoulli, kurtosis would equal its lower bound $\gamma^2/\sigma^2 + \sigma^4$ and the determinant may not always be strictly away from zero.

These are the largest shocks to income, so muting their pass-through suppresses the covariance between consumption and income growth. The opposite is true if $\gamma_{\zeta_t} > 0$.

Another equation that pins down $\phi_t^{(1)}$ and $\phi_t^{(2)}$ is $2\gamma_{\zeta_t}\phi_t^{(1)} + 2(\kappa_{\zeta_t} + \sigma_{\zeta_t}^2(3\sigma_{\zeta_t}^2 + 2\sigma_{v_t}^2))\phi_t^{(2)} = \mathbb{E}(\Delta^2 c_{it} \times (\Delta_\rho^2 y_{it})^2)$, which, abstracting from income measurement error, corresponds to the second line in (B.1). Absent transmission of transitory shocks, the right hand side reflects the strength of the covariance between consumption growth and squared shocks so it captures the pass-through of shocks of large magnitude. We highlight again three cases (for $\phi_t^{(1)}$ fixed):

- $\mathbb{E}(\Delta^2 c_{it} \times (\Delta_\rho^2 y_{it})^2) - 2\gamma_{\zeta_t}\phi_t^{(1)} = 0$. This identifies $\phi_t^{(2)} = 0$, so tail shocks transmit similarly to non-tail ones. In fact, $\mathbb{E}(\Delta^2 c_{it} \times (\Delta_\rho^2 y_{it})^2) = 2\gamma_{\zeta_t}\phi_t^{(1)}$ is implied by $\Delta^2 c_{it} = \phi_t^{(1)}(\zeta_{it} + \rho\zeta_{it-1})$, so this configuration identifies the *linear* consumption process.

- $\mathbb{E}(\Delta^2 c_{it} \times (\Delta_\rho^2 y_{it})^2) - 2\gamma_{\zeta_t}\phi_t^{(1)} < 0$, which, for $\gamma_{\zeta_t} < 0$, is true when $\mathbb{E}(\Delta^2 c_{it} \times (\Delta_\rho^2 y_{it})^2)$ is negative and large. This suggests that large shocks (of which negative shocks are the most likely) transmit strongly with a negative coefficient. It thus identifies $\phi_t^{(2)} < 0$. Ceteris paribus, a *less leptokurtic* income distribution (lower $\kappa_{\zeta_t}$) turns $\phi_t^{(2)}$ more negative; a bell-shaped distribution is associated with fewer large shocks, thus necessitating a stronger transmission of said shocks to maintain a given excess co-movement.

- $\mathbb{E}(\Delta^2 c_{it} \times (\Delta_\rho^2 y_{it})^2) - 2\gamma_{\zeta_t}\phi_t^{(1)} > 0$. This identifies $\phi_t^{(2)} > 0$ for analogous reasons. For given excess co-movement, the smaller the kurtosis of permanent shocks, the larger the value of $\phi_t^{(2)}$ must be to achieve said excess co-movement.

In practice, $\phi_t^{(1)}$ and $\phi_t^{(2)}$ are jointly pinned down by these and other equations of system (B.1), and they are also determined jointly with the transmission parameters of transitory shocks, to which analogous statements apply – skipped here for brevity. The sensitivity analysis is thus more nuanced, so simulations can prove useful. Starting from the baseline estimates of the income process (table 2), reducing the left skewness of permanent shocks by half increases the magnitude of $\phi_t^{(2)}$ by almost 10% (more negative). Independently reducing the kurtosis of permanent shocks by half increases the magnitude of $\phi_t^{(2)}$ by 60% (more negative). The response of $\phi_t^{(1)}$ is much smaller, confirming that the values of skewness and kurtosis are crucial for identifying the pass-through of *tail* shocks.

An interesting special case arises when $\rho = 1$, the distributions of permanent and transitory shocks are normal and have the same variance, and there is no measurement error. In this case, $\det(\mathbf{A}) = -288\sigma_t^{16}$, $\phi_t^{(1)} = (1/(2\sigma_t^2))\mathbb{E}\left(\Delta^2 c_{it} \times \Delta_\rho^2 y_{it}\right) + (1/(2\sigma_t^2))\mathbb{E}\left(\Delta^2 c_{it} \times \Delta_\rho^2 y_{it+2}\right)$ and $\phi_t^{(2)} = (1/(12\sigma_t^4))\mathbb{E}\left(\Delta^2 c_{it} \times (\Delta_\rho^2 y_{it})^2\right) + (1/(6\sigma_t^4))\mathbb{E}\left(\Delta^2 c_{it} \times \Delta_\rho^2 y_{it} \times \Delta_\rho^2 y_{it+2}\right)$, where $\sigma_t$ is the dispersion of the hypothetical distribution. Even in such a special environment, $\phi_t^{(2)}$, the transmission parameter of tail shocks, is separately identified from $\phi_t^{(1)}$. Identification

Table B.1: Targeted moments of income

| *Moments:* | $2^{\text{nd}}$ | $2^{\text{nd}}$, $3^{\text{rd}}$ | $2^{\text{nd}}$, $3^{\text{rd}}$, $4^{\text{th}}$ |
|---|---|---|---|
| 1 | $\mathbb{E}((\Delta^2_\rho y_{it})^2)$ | $\mathbb{E}((\Delta^2_\rho y_{it})^2)$ | $\mathbb{E}((\Delta^2_\rho y_{it})^2)$ |
| 2 | $\mathbb{E}(\Delta^2_\rho y_{it} \times \Delta^2_\rho y_{it-2})$ | $\mathbb{E}(\Delta^2_\rho y_{it} \times \Delta^2_\rho y_{it-2})$ | $\mathbb{E}(\Delta^2_\rho y_{it} \times \Delta^2_\rho y_{it-2})$ |
| 3 | $\mathbb{E}(\Delta^2_\rho y_{it} \times \Delta^2 y_{it-4})$ | $\mathbb{E}(\Delta^2_\rho y_{it} \times \Delta^2 y_{it-4})$ | $\mathbb{E}(\Delta^2_\rho y_{it} \times \Delta^2 y_{it-4})$ |
| 4 | | $\mathbb{E}((\Delta^2_\rho y_{it})^3)$ | $\mathbb{E}((\Delta^2_\rho y_{it})^3)$ |
| 5 | | $\mathbb{E}((\Delta^2_\rho y_{it})^2 \times \Delta^2_\rho y_{it-2})$ | $\mathbb{E}((\Delta^2_\rho y_{it})^2 \times \Delta^2_\rho y_{it-2})$ |
| 6 | | $\mathbb{E}(\Delta^2_\rho y_{it} \times (\Delta^2_\rho y_{it-2})^2)$ | $\mathbb{E}(\Delta^2_\rho y_{it} \times (\Delta^2_\rho y_{it-2})^2)$ |
| 7 | | | $\mathbb{E}((\Delta^2_\rho y_{it})^4)$ |
| 8 | | | $\mathbb{E}((\Delta^2_\rho y_{it})^2 \times (\Delta^2_\rho y_{it-2})^2)$ |
| 9 | | | $\mathbb{E}((\Delta^2_\rho y_{it})^3 \times \Delta^2_\rho y_{it-2})$ |
| 10 | | | $\mathbb{E}(\Delta^2_\rho y_{it} \times (\Delta^2_\rho y_{it-2})^3)$ |

*Notes:* The table lists the moments targeted in the estimation of the various specifications of the income process. $\Delta^2 y_t = y_t - y_{t-2}$ indicates a first difference between periods $t$ and $t-2$; $\Delta^2_\rho y_t = y_t - \rho^2 y_{t-2}$ indicates the analogous quasi-difference. Note that $\Delta^2 y_t = \Delta^2_\rho y_t$ if $\rho = 1$ (baseline permanent-transitory process).

does *not* require shocks to be non-Gaussian, but skewness and kurtosis (both defined also in the Gaussian case) are still crucial for identifying the pass-through of tail shocks. Despite normality, shocks in the (normal) tails are associated with a different pass-through from that of average shocks. Analogous expressions separately identify $\psi_t^{(1)}$ and $\psi_t^{(2)}$.

Upon identification of the transmission parameters, $\mathbb{E}(\Delta^2 c_{it} \times \Delta^2 c_{it+2})$ identifies $\sigma^2_{u^c_t}$, while $\mathbb{E}((\Delta^2 c_{it})^2)$ identifies $\sigma^2_{\xi_t}$. Higher moments of the consumption error or of the taste shock involve higher than fourth-order moments of income, which we do not model.

## B.4 Targeted moments

The estimation targets several over-identifying moments. Table B.1 lists the target income moments, while table B.2 lists the moments targeted in the various specifications of the consumption model. For simplicity, we illustrate these moments assuming stationarity in the distributions of income shocks, taste heterogeneity, and measurement error, and time-invariance of the partial insurance parameters. These are *not* identifying assumptions but they make the illustration more compact by allowing us to pool common terms together.

### B.4.1 Moments of income

$$\begin{aligned}
\mathbb{E}((\Delta^2_\rho y_{it})^2) &= (1+\rho^2)\sigma^2_\zeta + (1+\rho^4)\sigma^2_v + (1+\rho^4)\sigma^2_{u^y} \\
\mathbb{E}(\Delta^2_\rho y_{it} \times \Delta^2_\rho y_{it-2}) &= -\rho^2\sigma^2_v - \rho^2\sigma^2_{u^y} \\
\mathbb{E}(\Delta^2_\rho y_{it} \times \Delta^2 y_{it-4}) &= 0
\end{aligned}$$

$$
\begin{aligned}
\mathbb{E}((\Delta^2_\rho y_{it})^3) &= (1+\rho^3)\gamma_\zeta + (1-\rho^6)\gamma_v \\
\mathbb{E}((\Delta^2_\rho y_{it})^2 \times \Delta^2_\rho y_{it-2}) &= \rho^4\gamma_v \\
\mathbb{E}(\Delta^2_\rho y_{it} \times (\Delta^2_\rho y_{it-2})^2) &= -\rho^2\gamma_v \\
\mathbb{E}((\Delta^2_\rho y_{it})^4) &= (1+\rho^4)\kappa_\zeta + (1+\rho^8)\kappa_v + 6\rho^2(\sigma^2_\zeta)^2 + 6\rho^4(\sigma^2_v)^2 + 3(1+\rho^4)^2(\sigma^2_{u^y})^2 \\
&\quad + 6(1+\rho^4)\Big((1+\rho^2)\sigma^2_\zeta\sigma^2_v + (1+\rho^2)\sigma^2_\zeta\sigma^2_{u^y} + (1+\rho^4)\sigma^2_v\sigma^2_{u^y}\Big) \\
\mathbb{E}((\Delta^2_\rho y_{it})^2 \times (\Delta^2_\rho y_{it-2})^2) &= \rho^4\kappa_v + (1+\rho^2)^2(\sigma^2_\zeta)^2 + (1+\rho^4+\rho^8)(\sigma^2_v)^2 + (1+4\rho^4+\rho^8)(\sigma^2_{u^y})^2 \\
&\quad + 2(1+\rho^2)(1+\rho^4)\sigma^2_\zeta\sigma^2_v + 2(1+\rho^2)(1+\rho^4)\sigma^2_\zeta\sigma^2_{u^y} + (2(1+\rho^4)^2+4\rho^4)\sigma^2_v\sigma^2_{u^y} \\
\mathbb{E}((\Delta^2_\rho y_{it})^3 \times \Delta^2_\rho y_{it-2}) &= -\rho^6\kappa_v - 3\rho^2(\sigma^2_v)^2 - 3\rho^2(1+\rho^4)(\sigma^2_{u^y})^2 \\
&\quad - 3\rho^2(1+\rho^2)\sigma^2_\zeta\sigma^2_v - 3\rho^2(1+\rho^2)\sigma^2_\zeta\sigma^2_{u^y} - 6\rho^2(1+\rho^4)\sigma^2_v\sigma^2_{u^y} \\
\mathbb{E}(\Delta^2_\rho y_{it} \times (\Delta^2_\rho y_{it-2})^3) &= -\rho^2\kappa_v - 3\rho^6(\sigma^2_v)^2 - 3\rho^2(1+\rho^4)(\sigma^2_{u^y})^2 \\
&\quad - 3\rho^2(1+\rho^2)\sigma^2_\zeta\sigma^2_v - 3\rho^2(1+\rho^2)\sigma^2_\zeta\sigma^2_{u^y} - 6\rho^2(1+\rho^4)\sigma^2_v\sigma^2_{u^y}
\end{aligned}
$$

### B.4.2 Moments of consumption – linear function

$$
\begin{aligned}
\mathbb{E}((\Delta^2 c_{it})^2) &= 2\sigma^2_\xi + (\phi^{(1)})^2(1+\rho^2)\sigma^2_\zeta + (\psi^{(1)})^2\sigma^2_v + 2\sigma^2_{u^c} \\
\mathbb{E}(\Delta^2 c_{it} \times \Delta^2 c_{it-2}) &= -\sigma^2_{u^c} \\
\mathbb{E}(\Delta^2_\rho y_{it} \times \Delta^2 c_{it}) &= \phi^{(1)}(1+\rho^2)\sigma^2_\zeta + \psi^{(1)}\sigma^2_v \\
\mathbb{E}(\Delta^2_\rho y_{it} \times \Delta^2 c_{it-2}) &= -\psi^{(1)}\rho^2\sigma^2_v \\
\mathbb{E}((\Delta^2 c_{it})^3) &= 2\gamma_\xi + (\phi^{(1)})^3(1+\rho^3)\gamma_\zeta + (\psi^{(1)})^3\gamma_v \\
\mathbb{E}((\Delta^2 c_{it})^2 \times \Delta^2 c_{it-2}) &= \gamma_{u^c} \\
\mathbb{E}(\Delta^2 c_{it} \times (\Delta^2 c_{it-2})^2) &= -\gamma_{u^c} \\
\mathbb{E}((\Delta^2_\rho y_{it})^2 \times \Delta^2 c_{it}) &= \phi^{(1)}(1+\rho^3)\gamma_\zeta + \psi^{(1)}\gamma_v \\
\mathbb{E}(\Delta^2_\rho y_{it} \times (\Delta^2 c_{it})^2) &= (\phi^{(1)})^2(1+\rho^3)\gamma_\zeta + (\psi^{(1)})^2\gamma_v \\
\mathbb{E}((\Delta^2_\rho y_{it})^2 \times \Delta^2 c_{it-2}) &= \psi^{(1)}\rho^4\gamma_v \\
\mathbb{E}(\Delta^2_\rho y_{it} \times (\Delta^2 c_{it-2})^2) &= -(\psi^{(1)})^2\rho^2\gamma_v \\
\mathbb{E}((\Delta^2_\rho y_{it-2})^2 \times \Delta^2 c_{it}) &= 0 \\
\mathbb{E}(\Delta^2_\rho y_{it} \times \Delta^2_\rho y_{it+2} \times \Delta^2 c_{it}) &= -\psi^{(1)}\rho^2\gamma_v \\
\mathbb{E}(\Delta^2_\rho y_{it} \times \Delta^2_\rho y_{it+2} \times \Delta^2 c_{it-2}) &= 0 \\
\mathbb{E}(\Delta^2_\rho y_{it} \times \Delta^2_\rho y_{it+2} \times \Delta^2 c_{it+2}) &= 0 \\
\mathbb{E}((\Delta^2 c_{it})^4) &= 2\kappa_\xi + (\phi^{(1)})^4(1+\rho^4)\kappa_\zeta + (\psi^{(1)})^4\kappa_v + 2\kappa_{u^c} + 6(\sigma^2_\xi)^2
\end{aligned}
$$

Table B.2: Targeted joint moments of consumption and income

| *Cons. fn.:* | **Linear** | | | **Quadratic** |
|---|---|---|---|---|
| *Moments:* | $2^{\text{nd}}$ | $2^{\text{nd}}$, $3^{\text{rd}}$ | $2^{\text{nd}}$, $3^{\text{rd}}$, $4^{\text{th}}$ | $2^{\text{nd}}$, $3^{\text{rd}}$, $4^{\text{th}}$ (income shocks) |
| 1 | $\mathbb{E}((\Delta^2 c_{it})^2)$ | $\mathbb{E}((\Delta^2 c_{it})^2)$ | $\mathbb{E}((\Delta^2 c_{it})^2)$ | $\mathbb{E}((\Delta^2 c_{it})^2)$ |
| 2 | $\mathbb{E}(\Delta^2 c_{it} \times \Delta^2 c_{it-2})$ | $\mathbb{E}(\Delta^2 c_{it} \times \Delta^2 c_{it-2})$ | $\mathbb{E}(\Delta^2 c_{it} \times \Delta^2 c_{it-2})$ | $\mathbb{E}(\Delta^2 c_{it} \times \Delta^2 c_{it-2})$ |
| 3 | $\mathbb{E}(\Delta^2_\rho y_{it} \times \Delta^2 c_{it})$ | $\mathbb{E}(\Delta^2_\rho y_{it} \times \Delta^2 c_{it})$ | $\mathbb{E}(\Delta^2_\rho y_{it} \times \Delta^2 c_{it})$ | $\mathbb{E}(\Delta^2_\rho y_{it} \times \Delta^2 c_{it})$ |
| 4 | $\mathbb{E}(\Delta^2_\rho y_{it} \times \Delta^2 c_{it-2})$ | $\mathbb{E}(\Delta^2_\rho y_{it} \times \Delta^2 c_{it-2})$ | $\mathbb{E}(\Delta^2_\rho y_{it} \times \Delta^2 c_{it-2})$ | $\mathbb{E}(\Delta^2_\rho y_{it} \times \Delta^2 c_{it-2})$ |
| 5 | | $\mathbb{E}((\Delta^2 c_{it})^3)$ | $\mathbb{E}((\Delta^2 c_{it})^3)$ | |
| 6 | | $\mathbb{E}((\Delta^2 c_{it})^2 \times \Delta^2 c_{it-2})$ | $\mathbb{E}((\Delta^2 c_{it})^2 \times \Delta^2 c_{it-2})$ | |
| 7 | | $\mathbb{E}(\Delta^2 c_{it} \times (\Delta^2 c_{it-2})^2)$ | $\mathbb{E}(\Delta^2 c_{it} \times (\Delta^2 c_{it-2})^2)$ | |
| 8 | | $\mathbb{E}((\Delta^2_\rho y_{it})^2 \times \Delta^2 c_{it})$ | $\mathbb{E}((\Delta^2_\rho y_{it})^2 \times \Delta^2 c_{it})$ | $\mathbb{E}((\Delta^2_\rho y_{it})^2 \times \Delta^2 c_{it})$ |
| 9 | | $\mathbb{E}(\Delta^2_\rho y_{it} \times (\Delta^2 c_{it})^2)$ | $\mathbb{E}(\Delta^2_\rho y_{it} \times (\Delta^2 c_{it})^2)$ | |
| 10 | | $\mathbb{E}((\Delta^2_\rho y_{it})^2 \times \Delta^2 c_{it-2})$ | $\mathbb{E}((\Delta^2_\rho y_{it})^2 \times \Delta^2 c_{it-2})$ | $\mathbb{E}((\Delta^2_\rho y_{it})^2 \times \Delta^2 c_{it-2})$ |
| 11 | | $\mathbb{E}(\Delta^2_\rho y_{it} \times (\Delta c_{it-2})^2)$ | $\mathbb{E}(\Delta^2_\rho y_{it} \times (\Delta^2 c_{it-2})^2)$ | |
| 12 | | $\mathbb{E}((\Delta^2_\rho y_{it-2})^2 \times \Delta^2 c_{it})$ | $\mathbb{E}((\Delta^2_\rho y_{it-2})^2 \times \Delta^2 c_{it})$ | $\mathbb{E}((\Delta^2_\rho y_{it-2})^2 \times \Delta^2 c_{it})$ |
| 13 | | $\mathbb{E}(\Delta^2_\rho y_{it} \times \Delta^2_\rho y_{it+2} \times \Delta^2 c_{it})$ | $\mathbb{E}(\Delta^2_\rho y_{it} \times \Delta^2_\rho y_{it+2} \times \Delta^2 c_{it})$ | $\mathbb{E}(\Delta^2_\rho y_{it} \times \Delta^2_\rho y_{it+2} \times \Delta^2 c_{it})$ |
| 14 | | $\mathbb{E}(\Delta^2_\rho y_{it} \times \Delta^2_\rho y_{it+2} \times \Delta^2 c_{it-2})$ | $\mathbb{E}(\Delta^2_\rho y_{it} \times \Delta^2_\rho y_{it+2} \times \Delta^2 c_{it-2})$ | $\mathbb{E}(\Delta^2_\rho y_{it} \times \Delta^2_\rho y_{it+2} \times \Delta^2 c_{it-2})$ |
| 15 | | $\mathbb{E}(\Delta^2_\rho y_{it} \times \Delta^2_\rho y_{it+2} \times \Delta^2 c_{it+2})$ | $\mathbb{E}(\Delta^2_\rho y_{it} \times \Delta^2_\rho y_{it+2} \times \Delta^2 c_{it+2})$ | $\mathbb{E}(\Delta^2_\rho y_{it} \times \Delta^2_\rho y_{it+2} \times \Delta^2 c_{it+2})$ |
| 16 | | | $\mathbb{E}((\Delta^2 c_{it})^4)$ | |
| 17 | | | $\mathbb{E}((\Delta^2 c_{it})^2 \times (\Delta^2 c_{it-2})^2)$ | |
| 18 | | | $\mathbb{E}((\Delta^2 c_{it})^3 \times \Delta^2 c_{it-2})$ | |
| 19 | | | $\mathbb{E}(\Delta^2 c_{it} \times (\Delta^2 c_{it-2})^3)$ | |
| 20 | | | $\mathbb{E}((\Delta^2_\rho y_{it})^2 \times (\Delta^2 c_{it})^2)$ | |
| 21 | | | $\mathbb{E}((\Delta^2_\rho y_{it})^3 \times \Delta^2 c_{it})$ | |
| 22 | | | $\mathbb{E}(\Delta^2_\rho y_{it} \times (\Delta^2 c_{it})^3)$ | |
| 23 | | | $\mathbb{E}((\Delta^2_\rho y_{it})^2 \times (\Delta^2 c_{it-2})^2)$ | |
| 24 | | | $\mathbb{E}((\Delta^2_\rho y_{it-2})^2 \times (\Delta^2 c_{it})^2)$ | |
| 25 | | | $\mathbb{E}((\Delta^2_\rho y_{it})^3 \times \Delta^2 c_{it-2})$ | |
| 26 | | | $\mathbb{E}(\Delta^2_\rho y_{it} \times (\Delta^2 c_{it-2})^3)$ | |

*Notes:* The table lists the moments targeted in the estimation of the different specifications of the consumption process. The notation $\Delta^2 x_t = x_t - x_{t-2}$ indicates a first difference of the generic variable $x$ between periods $t$ and $t-2$; $\Delta^2_\rho x_t = x_t - \rho^2 x_{t-2}$ indicates the analogous quasi-difference.

$$
\begin{aligned}
&+ 6(\phi^{(1)})^4\rho^2(\sigma_\zeta^2)^2 + 6(\sigma_{u^c}^2)^2 + 12(\phi^{(1)})^2(1+\rho^2)\sigma_\xi^2\sigma_\zeta^2 \\
&+ 12(\psi^{(1)})^2\sigma_\xi^2\sigma_v^2 + 24\sigma_\xi^2\sigma_{u^c}^2 + 6(\phi^{(1)})^2(\psi^{(1)})^2(1+\rho^2)\sigma_\zeta^2\sigma_v^2 \\
&+ 12(\phi^{(1)})^2(1+\rho^2)\sigma_\zeta^2\sigma_{u^c}^2 + 12(\psi^{(1)})^2\sigma_v^2\sigma_{u^c}^2 \\
\mathbb{E}((\Delta^2 c_{it})^2 \times (\Delta^2 c_{it-2})^2) &= 4(\sigma_\xi^2)^2 + 4(\phi^{(1)})^2(1+\rho^2)\sigma_\xi^2\sigma_\zeta^2 + 4(\psi^{(1)})^2\sigma_\xi^2\sigma_v^2 + 8\sigma_\xi^2\sigma_{u^c}^2 \\
&+ (\phi^{(1)})^4(1+\rho^2)^2(\sigma_\zeta^2)^2 + 3(\sigma_{u^c}^2)^2 + 2(\phi^{(1)})^2(\psi^{(1)})^2(1+\rho^2)\sigma_\zeta^2\sigma_v^2 \\
&+ 4(\phi^{(1)})^2(1+\rho^2)\sigma_\zeta^2\sigma_{u^c}^2 + (\psi^{(1)})^4(\sigma_v^2)^2 + 4(\psi^{(1)})^2\sigma_v^2\sigma_{u^c}^2 + \kappa_{u^c} \\
\mathbb{E}((\Delta^2 c_{it})^3 \times \Delta^2 c_{it-2}) &= -\kappa_{u^c} - 3(\sigma_{u^c}^2)^2 - 6\sigma_\xi^2\sigma_{u^c}^2 - 3(\phi^{(1)})^2(1+\rho^2)\sigma_\zeta^2\sigma_{u^c}^2 - 3(\psi^{(1)})^2\sigma_v^2\sigma_{u^c}^2 \\
\mathbb{E}(\Delta^2 c_{it} \times (\Delta^2 c_{it-2})^3) &= -\kappa_{u^c} - 3(\sigma_{u^c}^2)^2 - 6\sigma_\xi^2\sigma_{u^c}^2 - 3(\phi^{(1)})^2(1+\rho^2)\sigma_\zeta^2\sigma_{u^c}^2 - 3(\psi^{(1)})^2\sigma_v^2\sigma_{u^c}^2 \\
\mathbb{E}((\Delta_\rho^2 y_{it})^2 \times (\Delta^2 c_{it})^2) &= 2(1+\rho^2)\sigma_\zeta^2\sigma_\xi^2 + (\phi^{(1)})^2(1+\rho^4)\kappa_\zeta + 6(\phi^{(1)})^2\rho^2(\sigma_\zeta^2)^2 \\
&+ (\psi^{(1)})^2(1+\rho^2)\sigma_\zeta^2\sigma_v^2 + 2(1+\rho^2)\sigma_\zeta^2\sigma_{u^c}^2 + 2(1+\rho^4)\sigma_v^2\sigma_\xi^2 \\
&+ (\phi^{(1)})^2(1+\rho^2)(1+\rho^4)\sigma_v^2\sigma_\zeta^2 + (\psi^{(1)})^2\kappa_v + (\psi^{(1)})^2\rho^4(\sigma_v^2)^2 \\
&+ 2(1+\rho^4)\sigma_v^2\sigma_{u^c}^2 + 2(1+\rho^4)\sigma_{u^y}^2\sigma_\xi^2 + (\phi^{(1)})^2(1+\rho^2)(1+\rho^4)\sigma_{u^y}^2\sigma_\zeta^2 \\
&+ (\psi^{(1)})^2(1+\rho^4)\sigma_{u^y}^2\sigma_v^2 + 2(1+\rho^4)\sigma_{u^y}^2\sigma_{u^c}^2 + 4\phi^{(1)}\psi^{(1)}(1+\rho^2)\sigma_\zeta^2\sigma_v^2 \\
\mathbb{E}((\Delta_\rho^2 y_{it})^3 \times \Delta^2 c_{it}) &= \phi^{(1)}(1+\rho^4)\kappa_\zeta + 6\phi^{(1)}\rho^2(\sigma_\zeta^2)^2 + \psi^{(1)}\kappa_v + 3\psi^{(1)}\rho^4(\sigma_v^2)^2 \\
&+ 3\psi^{(1)}(1+\rho^2)\sigma_\zeta^2\sigma_v^2 + 3\phi^{(1)}(1+\rho^2)(1+\rho^4)\sigma_\zeta^2\sigma_v^2 \\
&+ 3\phi^{(1)}(1+\rho^2)(1+\rho^4)\sigma_\zeta^2\sigma_{u^y}^2 + 3\psi^{(1)}(1+\rho^4)\sigma_v^2\sigma_{u^y}^2 \\
\mathbb{E}(\Delta_\rho^2 y_{it} \times (\Delta^2 c_{it})^3) &= (\phi^{(1)})^3(1+\rho^4)\kappa_\zeta + 6(\phi^{(1)})^3\rho^2(\sigma_\zeta^2)^2 + (\psi^{(1)})^3\kappa_v \\
&+ 6\phi^{(1)}(1+\rho^2)\sigma_\xi^2\sigma_\zeta^2 + 6\psi^{(1)}\sigma_\xi^2\sigma_v^2 + 3\phi^{(1)}(\psi^{(1)})^2(1+\rho^2)\sigma_\zeta^2\sigma_v^2 \\
&+ 3(\phi^{(1)})^2\psi^{(1)}(1+\rho^2)\sigma_\zeta^2\sigma_v^2 + 6\phi^{(1)}(1+\rho^2)\sigma_\zeta^2\sigma_{u^c}^2 + 6\psi^{(1)}\sigma_v^2\sigma_{u^c}^2 \\
\mathbb{E}((\Delta_\rho^2 y_{it})^2 \times (\Delta^2 c_{it-2})^2) &= 2(1+\rho^2)\sigma_\zeta^2\sigma_\xi^2 + (\phi^{(1)})^2(1+\rho^2)^2(\sigma_\zeta^2)^2 + (\psi^{(1)})^2(1+\rho^2)\sigma_\zeta^2\sigma_v^2 \\
&+ 2(1+\rho^2)\sigma_\zeta^2\sigma_{u^c}^2 + 2(1+\rho^4)\sigma_v^2\sigma_\xi^2 + (\phi^{(1)})^2(1+\rho^2)(1+\rho^4)\sigma_v^2\sigma_\zeta^2 \\
&+ (\psi^{(1)})^2\rho^4\kappa_v + (\psi^{(1)})^2(\sigma_v^2)^2 + 2(1+\rho^4)\sigma_v^2\sigma_{u^c}^2 + 2(1+\rho^4)\sigma_{u^y}^2\sigma_\xi^2 \\
&+ (\phi^{(1)})^2(1+\rho^2)(1+\rho^4)\sigma_{u^y}^2\sigma_\zeta^2 + (\psi^{(1)})^2(1+\rho^4)\sigma_{u^y}^2\sigma_v^2 + 2(1+\rho^4)\sigma_{u^y}^2\sigma_{u^c}^2 \\
\mathbb{E}((\Delta_\rho^2 y_{it-2})^2 \times (\Delta^2 c_{it})^2) &= 2(1+\rho^2)\sigma_\zeta^2\sigma_\xi^2 + (\phi^{(1)})^2(1+\rho^2)^2(\sigma_\zeta^2)^2 + (\psi^{(1)})^2(1+\rho^2)\sigma_\zeta^2\sigma_v^2 \\
&+ 2(1+\rho^2)\sigma_\zeta^2\sigma_{u^c}^2 + 2(1+\rho^4)\sigma_v^2\sigma_\xi^2 + (\phi^{(1)})^2(1+\rho^2)(1+\rho^4)\sigma_v^2\sigma_\zeta^2 \\
&+ (\psi^{(1)})^2(1+\rho^4)(\sigma_v^2)^2 + 2(1+\rho^4)\sigma_v^2\sigma_{u^c}^2 + 2(1+\rho^4)\sigma_{u^y}^2\sigma_\xi^2 \\
&+ (\phi^{(1)})^2(1+\rho^2)(1+\rho^4)\sigma_{u^y}^2\sigma_\zeta^2 + (\psi^{(1)})^2(1+\rho^4)\sigma_{u^y}^2\sigma_v^2 + 2(1+\rho^4)\sigma_{u^y}^2\sigma_{u^c}^2 \\
\mathbb{E}((\Delta_\rho^2 y_{it})^3 \times \Delta^2 c_{it-2}) &= -\psi^{(1)}\rho^6\kappa_v - 3\psi^{(1)}\rho^2(\sigma_v^2)^2 - 3\psi^{(1)}(1+\rho^2)\rho^2\sigma_\zeta^2\sigma_v^2 - 3\psi^{(1)}(1+\rho^4)\rho^2\sigma_v^2\sigma_{u^y}^2 \\
\mathbb{E}(\Delta_\rho^2 y_{it} \times (\Delta^2 c_{it-2})^3) &= -(\psi^{(1)})^3\rho^2\kappa_v - 6\psi^{(1)}\rho^2\sigma_v^2\sigma_\xi^2 - 3(\phi^{(1)})^2\psi^{(1)}(1+\rho^2)\rho^2\sigma_v^2\sigma_\zeta^2 - 6\psi^{(1)}\rho^2\sigma_v^2\sigma_{u^c}^2
\end{aligned}
$$

### B.4.3 Moments of consumption – quadratic function

$$
\begin{aligned}
\mathbb{E}((\Delta^2 c_{it})^2) &= 2\sigma_\xi^2 + (\phi^{(1)})^2(1+\rho^2)\sigma_\zeta^2 + (\psi^{(1)})^2\sigma_v^2 + (\phi^{(2)})^2\Big((1+\rho^4)\kappa_\zeta + 6\rho^2(\sigma_\zeta^2)^2\Big) \\
&\quad + (\psi^{(2)})^2\kappa_v + (\omega^{(22)})^2(1+\rho^2)\sigma_\zeta^2\sigma_v^2 + 2\sigma_{u^c}^2 \\
&\quad + 2\phi^{(1)}\phi^{(2)}(1+\rho^3)\gamma_\zeta + 2\psi^{(1)}\psi^{(2)}\gamma_v + 2\phi^{(2)}\psi^{(2)}(1+\rho^2)\sigma_\zeta^2\sigma_v^2 \\
\mathbb{E}(\Delta^2 c_{it} \times \Delta^2 c_{it-2}) &= (\phi^{(2)})^2(1+\rho^2)^2(\sigma_\zeta^2)^2 + 2\phi^{(2)}\psi^{(2)}(1+\rho^2)\sigma_\zeta^2\sigma_v^2 + (\psi^{(2)})^2(\sigma_v^2)^2 - \sigma_{u^c}^2 \\
\mathbb{E}(\Delta_\rho^2 y_{it} \times \Delta^2 c_{it}) &= \phi^{(1)}(1+\rho^2)\sigma_\zeta^2 + \psi^{(1)}\sigma_v^2 + \phi^{(2)}(1+\rho^3)\gamma_\zeta + \psi^{(2)}\gamma_v \\
\mathbb{E}(\Delta_\rho^2 y_{it} \times \Delta^2 c_{it-2}) &= -\psi^{(1)}\rho^2\sigma_v^2 - \psi^{(2)}\rho^2\gamma_v \\
\mathbb{E}((\Delta_\rho^2 y_{it})^2 \times \Delta^2 c_{it}) &= \phi^{(1)}(1+\rho^3)\gamma_\zeta + \psi^{(1)}\gamma_v \\
&\quad + \phi^{(2)}\Big((1+\rho^4)\kappa_\zeta + 6\rho^2(\sigma_\zeta^2)^2 + (1+\rho^2)(1+\rho^4)\sigma_\zeta^2(\sigma_v^2+\sigma_{u^y}^2)\Big) \\
&\quad + \psi^{(2)}\Big(\kappa_v + \rho^4(\sigma_v^2)^2 + (1+\rho^2)\sigma_\zeta^2\sigma_v^2 + (1+\rho^4)\sigma_v^2\sigma_{u^y}^2\Big) + 2\omega^{(22)}(1+\rho^2)\sigma_\zeta^2\sigma_v^2 \\
\mathbb{E}((\Delta_\rho^2 y_{it})^2 \times \Delta^2 c_{it-2}) &= \psi^{(1)}\rho^4\gamma_v + \phi^{(2)}(1+\rho^2)\sigma_\zeta^2\Big((1+\rho^2)\sigma_\zeta^2 + (1+\rho^4)(\sigma_v^2+\sigma_{u^y}^2)\Big) \\
&\quad + \psi^{(2)}\Big(\rho^4\kappa_v + (\sigma_v^2)^2 + (1+\rho^2)\sigma_\zeta^2\sigma_v^2 + (1+\rho^4)\sigma_v^2\sigma_{u^y}^2\Big) \\
\mathbb{E}((\Delta_\rho^2 y_{it-2})^2 \times \Delta^2 c_{it}) &= \phi^{(2)}(1+\rho^2)\sigma_\zeta^2\Big((1+\rho^2)\sigma_\zeta^2 + (1+\rho^4)(\sigma_v^2+\sigma_{u^y}^2)\Big) \\
&\quad + \psi^{(2)}\sigma_v^2\Big((1+\rho^2)\sigma_\zeta^2 + (1+\rho^4)(\sigma_v^2+\sigma_{u^y}^2)\Big) \\
\mathbb{E}(\Delta_\rho^2 y_{it} \times \Delta_\rho^2 y_{it+2} \times \Delta^2 c_{it}) &= -\psi^{(1)}\rho^2\gamma_v - \phi^{(2)}\rho^2(1+\rho^2)\sigma_\zeta^2(\sigma_v^2+\sigma_{u^y}^2) \\
&\quad - \psi^{(2)}\rho^2(\kappa_v + \sigma_v^2\sigma_{u^y}^2) - \omega^{(22)}\rho^2(1+\rho^2)\sigma_\zeta^2\sigma_v^2 \\
\mathbb{E}(\Delta_\rho^2 y_{it} \times \Delta_\rho^2 y_{it+2} \times \Delta^2 c_{it-2}) &= -\phi^{(2)}\rho^2(1+\rho^2)\sigma_\zeta^2(\sigma_v^2+\sigma_{u^y}^2) - \psi^{(2)}\rho^2\sigma_v^2(\sigma_v^2+\sigma_{u^y}^2) \\
\mathbb{E}(\Delta_\rho^2 y_{it} \times \Delta_\rho^2 y_{it+2} \times \Delta^2 c_{it+2}) &= -\phi^{(2)}\rho^2(1+\rho^2)\sigma_\zeta^2(\sigma_v^2+\sigma_{u^y}^2) - \psi^{(2)}\rho^2\sigma_v^2(\sigma_v^2+\sigma_{u^y}^2)
\end{aligned}
$$

# C Empirical details and additional results

## C.1 Summary statistics in the PSID sample

Table C.1 presents summary statistics in the PSID sample over the period 1999–2019 (columns 1–2), as well as in subsamples sorted by wealth and education (columns 3–6).

## C.2 Comparison with administrative moments

Table C.2 reports the empirical second, third, and fourth moments of the cross-sectional distribution of unexplained growth in male earnings (panel A) and household earnings (panel B), as opposed to the moments of household disposable income in table 1. Column 1 presents

Table C.1: Descriptive statistics, PSID

| | Baseline sample | | Low wealth | | High wealth | |
|---|---|---|---|---|---|---|
| | | | No college | Some college | No college | Some college |
| | mean | median | mean | mean | mean | mean |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| *Consumption* | | | | | | |
| Nondurable+services | 26,766 | 22,895 | 21,280 | 24,566 | 23,675 | 31,479 |
| Food at home | 8,256 | 7,609 | 7,438 | 8,210 | 7,505 | 8,870 |
| Utilities | 3,515 | 3,238 | 3,351 | 3,205 | 3,471 | 3,787 |
| Food out | 3,318 | 2,511 | 2,387 | 2,928 | 3,109 | 4,037 |
| Health | 1,614 | 863 | 1,321 | 1,445 | 1,623 | 1,850 |
| Transportation | 5,866 | 4,900 | 5,283 | 5,489 | 5,711 | 6,406 |
| Education | 3,202 | 0 | 957 | 1,986 | 1,810 | 5,352 |
| Childcare | 977 | 0 | 546 | 1,259 | 537 | 1,136 |
| Rent (or rent eq.) | 18,105 | 13,821 | 8,910 | 12,631 | 16,449 | 26,020 |
| Home insurance | 894 | 747 | 585 | 712 | 908 | 1,142 |
| *Income* | | | | | | |
| Disposable income | 142,910 | 113,581 | 84,414 | 110,176 | 120,931 | 195,158 |
| Taxable+transfer inc. | 137,277 | 109,003 | 77,707 | 106,276 | 113,789 | 189,446 |
| Earnings | 118,423 | 97,154 | 71,504 | 99,798 | 90,950 | 158,616 |
| Male | 81,857 | 62,185 | 46,025 | 65,725 | 57,519 | 114,545 |
| Female | 36,567 | 29,183 | 25,479 | 34,074 | 33,430 | 44,070 |
| *Wealth* | 547,761 | 191,193 | 68,561 | 70,665 | 670,948 | 1,024,577 |
| *Demographics* | | | | | | |
| % working (male) | 0.89 | 1.00 | 0.89 | 0.94 | 0.82 | 0.89 |
| % working (female) | 0.79 | 1.00 | 0.79 | 0.82 | 0.78 | 0.78 |
| Age (male) | 46 | 45 | 44 | 42 | 50 | 48 |
| % some college (male) | 0.68 | 1.00 | 0.00 | 1.00 | 0.00 | 1.00 |
| % some college (female) | 0.71 | 1.00 | 0.42 | 0.79 | 0.48 | 0.87 |
| Periods per household | 5.63 | 5.00 | 4.90 | 4.88 | 5.86 | 6.62 |
| Observations | 20,866 | | 4,258 | 5,159 | 2,428 | 9,021 |

*Notes:* The table reports summary statistics for consumption, income, wealth, and demographics in our baseline sample in the PSID and in subsamples sorted by wealth and education, over 1999–2019. Monetary amounts are expressed in \$2018. Rent for owners is calculated at 6% of the self-reported value of the house.

moments in the PSID sample, while column 2 presents the corresponding statistics, wherever available, from the U.S. Social Security Administration data of Guvenen et al. (2021).

Table C.2: Empirical moments, male earnings and household earnings

| | Baseline sample | | Guvenen et al. (2021) |
|---|---|---|---|
| *Income data:* | PSID | | Social Security Admin. |
| | (1) | | (2) |
| **Panel A. Moments of male earnings growth** | | | |
| $\text{Var}(\Delta y_{it})$ | 0.271 | (0.010) | 0.323 |
| $\text{Cov}(\Delta y_{it}, \Delta y_{it+1})$ | -0.065 | (0.004) | – |
| $\text{Skew}(\Delta y_{it})$ | -0.569 | (0.128) | -1.039 |
| $\text{Cov}((\Delta y_{it})^2, \Delta y_{it+1})$ | 0.087 | (0.010) | – |
| $\text{Kurt}(\Delta y_{it})$ | 13.711 | (0.551) | 13.494 |
| $\text{Cov}((\Delta y_{it})^2, (\Delta y_{it+1})^2)$ | 0.253 | (0.022) | – |
| **Panel B. Moments of household earnings growth** | | | |
| $\text{Var}(\Delta y_{it})$ | 0.233 | (0.009) | – |
| $\text{Cov}(\Delta y_{it}, \Delta y_{it+1})$ | -0.056 | (0.004) | – |
| $\text{Skew}(\Delta y_{it})$ | -0.728 | (0.142) | – |
| $\text{Cov}((\Delta y_{it})^2, \Delta y_{it+1})$ | 0.045 | (0.008) | – |
| $\text{Kurt}(\Delta y_{it})$ | 14.728 | (0.822) | – |
| $\text{Cov}((\Delta y_{it})^2, (\Delta y_{it+1})^2)$ | 0.176 | (0.017) | – |

*Notes:* The table presents the second, third, and fourth moments of earnings growth. Skewness and kurtosis correspond to the third and fourth standardized moments respectively. Column 1 reports moments of biennial earnings growth in the PSID (we maintain the notation $\Delta x_t$ for the first difference of variable $x$ over time, noting that, given the biennial nature of the PSID, this corresponds to a difference over two calendar years), while column 2 reports moments of male earnings growth from the U.S. Social Security Administration data. The latter concern the one-year male earnings growth, unconditional on past earnings, and correspond to averages over ages 30-55 of, respectively, dispersion squared, skewness, and kurtosis, as reported in appendix tables C23-C25 in Guvenen et al. (2021). Block bootstrap standard errors are in parentheses.

## C.3 Sources of tail income shocks

Table C.3 presents evidence on the underlying sources of large income reductions that drive the left skewness and excess kurtosis of income growth. Columns 1 & 2 report average unemployment, disability, bad health, etc, in the bottom 10% and top 90% of the distribution of residual income changes $\Delta y_{it}$, respectively. Column 3 reports the $p$-value of the difference between the two; in most cases, the bottom 10% (effectively households facing large income reductions) experience unemployment, worsening health, large drops in hours, and precarious housing at much higher rates than the rest.

## C.4 CEX data and imputation in the PSID

The Consumer Expenditure Survey (CEX) accounts for about 95% of all household expenditures from a highly disaggregated list of consumption goods and services (see Meyer and

Table C.3: Sources of tail income shocks

| | Disposable income growth | | |
|---|---|---|---|
| | Bottom 10% | Top 90% | |
| | mean | mean | $p$-value $\Delta$mean |
| | (1) | (2) | (3) |
| % unemployed $t$ (male) | 0.082 | 0.034 | 0.000 |
| % unemployed $t-1$ (male) | 0.084 | 0.036 | 0.000 |
| % unemployed $t$ (female) | 0.081 | 0.037 | 0.000 |
| % unemployed $t-1$ (female) | 0.081 | 0.041 | 0.000 |
| % main job laid off (male) | 0.054 | 0.039 | 0.003 |
| % second job laid off (male) | 0.153 | 0.117 | 0.000 |
| % main job laid off (female) | 0.073 | 0.059 | 0.023 |
| % second job laid off (female) | 0.096 | 0.095 | 0.921 |
| % physical disability (male) | 0.130 | 0.103 | 0.001 |
| % physical disability (female) | 0.121 | 0.120 | 0.840 |
| % physical disability (both) | 0.032 | 0.025 | 0.077 |
| % overall health got worse (male) | 0.129 | 0.110 | 0.036 |
| % overall health got worse (female) | 0.121 | 0.104 | 0.052 |
| % eviction from home | 0.049 | 0.031 | 0.000 |
| Hours growth (male) | -0.309 | 0.031 | 0.000 |
| Hours growth (female) | -0.228 | 0.023 | 0.000 |

*Notes:* The table reports the sample mean (mostly empirical probabilities) of the various events listed in the rows of the table in the bottom 10% and top 90% of the distribution of residual income growth $\Delta y_{it}$. Given the biennial nature of the PSID, $\Delta y_{it}$ corresponds to a difference over two calendar years.

Sullivan, 2023, for a recent overview). The CEX is run by the Census Bureau and the Bureau of Labor Statistics and forms the basis for the calculation of the consumer's price index.

**Sample selection and variables.** To parallel the design of the PSID sample, we use CEX interview data between 1999 and 2019. We select a sample of households that mimics closely our baseline selection in the PSID.[4] Specifically, we select continuously married couples with the male spouse aged 30 to 65. We require non-missing data on expenditure and basic demographics, and we drop those with zero food expenditure given the nature of the imputation subsequently. As in the PSID, we define consumption as the sum of real expenditure on nondurable goods and services, namely food (at home and outside), utilities, out-of-pocket health expenses, public transport, vehicle expenses, education, and daycare. As the CEX collects quarterly data over four rolling quarters, we generate annual consumption

[4]To improve the quality of the food demand estimation subsequently, we keep *all* calendar years between 1999 and 2019, i.e., not just the odd ones as in the PSID.

Table C.4: Comparison of sample means, PSID and CEX

| | 1999 | | 2009 | | 2019 | |
|---|---|---|---|---|---|---|
| | PSID | CEX | PSID | CEX | PSID | CEX |
| Earnings | 72,829 | 59,011 | 109,047 | 84,383 | 130,488 | 106,316 |
| Consumption | 16,115 | 10,631 | 22,317 | 15,454 | 27,315 | 18,592 |
| Food | 7,434 | 4,461 | 9,306 | 6,090 | 12,575 | 7,636 |
| Age | 44.70 | 45.76 | 46.40 | 47.59 | 46.69 | 48.40 |
| Family size | 3.39 | 3.54 | 3.24 | 3.48 | 3.42 | 3.44 |
| # children | 1.13 | 1.21 | 1.01 | 1.11 | 1.19 | 1.07 |
| % white | 0.93 | 0.88 | 0.91 | 0.86 | 0.91 | 0.84 |
| % high school dropout | 0.09 | 0.12 | 0.06 | 0.12 | 0.07 | 0.10 |
| % high school graduate | 0.28 | 0.27 | 0.25 | 0.25 | 0.22 | 0.21 |
| % college dropout | 0.62 | 0.62 | 0.69 | 0.63 | 0.71 | 0.69 |
| % Northeast | 0.17 | 0.18 | 0.18 | 0.19 | 0.16 | 0.17 |
| % Midwest | 0.32 | 0.24 | 0.30 | 0.25 | 0.31 | 0.21 |
| % South | 0.33 | 0.32 | 0.32 | 0.34 | 0.33 | 0.33 |
| % West | 0.18 | 0.25 | 0.20 | 0.22 | 0.20 | 0.28 |
| % working (male) | 0.89 | 0.88 | 0.91 | 0.93 | 0.90 | 0.92 |
| % working (female) | 0.81 | 0.77 | 0.79 | 0.81 | 0.79 | 0.79 |

*Notes:* The table presents a comparison of means in the PSID and CEX samples in 1999, 2009, and 2019.

as the sum of consumption over four quarters. We assign the final observation to the calendar year to which the underlying quarterly data mostly correspond to. Our final sample consists of 31,751 observations.

Table C.4 presents a comparison of means between the PSID and CEX samples for a host of household characteristics over the period 1999–2019. The two samples are very close with respect to age, family size, number of children, race, education, region of residence, and labor market participation. Household earnings are higher in the PSID than in the CEX; BPP noted this also for the earlier period (1980–1992) and argued that it is due to the more comprehensive definition of income in the PSID. Food expenditure in the PSID is higher, on average, by about 60% than food expenditure in the CEX. This is in contrast to BPP who find that food expenditure is similar across the two surveys over 1980–1992. This results in consumption being about 25% higher in the PSID than in the CEX.

Figure C.1 plots the variance, skewness, and kurtosis of log consumption in the PSID (blue solid line) and CEX samples (black dotted line). The patterns for the variance and kurtosis are quite similar in the two surveys over time, although consumption in the PSID is less volatile but slightly more leptokurtic. For skewness, there is disagreement both in the trends over time and the sign. The PSID features positive while the CEX weakly negative

skewness. Overall, these statistics (in particular skewness) suggest that consumption across otherwise similarly selected samples in the PSID and CEX exhibits some notable differences.

**Consumption imputation from CEX into PSID.** The previous comparison is not directly informative about the moments of consumption *growth*, which is the relevant variable in our model. The CEX lacks the panel dimension needed for our analysis, so it cannot be used to obtain measures of annual (or biennial) consumption growth.[5] We thus follow BPP and impute nondurable consumption from the CEX into the PSID.

The imputation involves two distinct steps. In the first step, we estimate food demand in the CEX. Specifically, we let real food expenditure be a log-linear function of the log of nondurable expenditure $c_{it}$, relative prices $\mathbf{p}_t$, and demographics $\mathbf{W}_{it}$, namely

$$f_{it} = \beta(D_{it})c_{it} + \mathbf{p}_t'\boldsymbol{\gamma} + \mathbf{W}_{it}'\boldsymbol{\mu} + e_{it}, \tag{C.1}$$

where $f_{it}$ is the log of real annual food expenditure for household $i$ in the CEX in year $t$, $\beta$ is the budget elasticity of food demand, which we allow to shift with time and household characteristics $D_{it}$, and $e_{it}$ captures unobserved heterogeneity in food demand.[6] To address possible endogeneity of nondurable expenditure, we follow an instrumental variables approach using as instruments the mean and standard deviation of household disposable income and the standard deviation of spousal wages by cohort, education, year, the former interacted with time, education, and number of children dummies.[7]

In the second step, we invert (C.1) under the assumption that food demand is monotonic in nondurable expenditure. We can then use the inverted equation to predict $c_{it}$ for a given level of food expenditure, given prices and demographics. Since the CEX and PSID samples are representative of the same underlying population, and since food expenditure, prices, and demographics are available in both surveys, we can use the inverted (C.1) to impute consumption into the PSID. This allows us to obtain an alternative, external measure of consumption in the PSID, and thus of consumption growth.

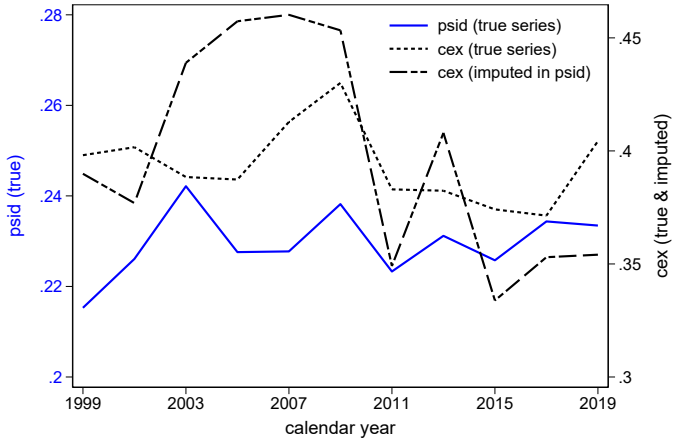
Figure C.1 plots the variance, skewness, and kurtosis of log consumption imputed from the CEX (black dashed-dotted line). The magnitudes of the various moments are quite

[5]The CEX consists of repeated cross sections, in which households are interviewed over 4 quarters. There is also an initial training quarter, which is rarely used in empirical analyses.

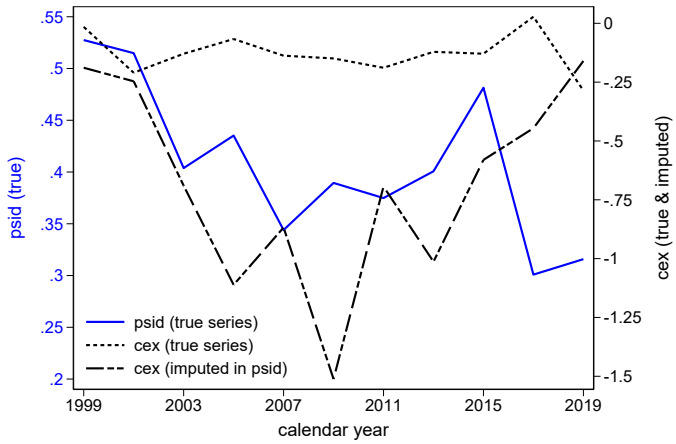
[6]Prices $\mathbf{p}_t$ include the relative prices of food, public transportation, utilities, medical services, and childcare. Demographics $\mathbf{W}_{it}$ include a quadratic polynomial in age and dummy variables for the number of children, family size, education, region of residence, race, and cohort. $D_{it}$ includes year dummies and dummies for education and the number of children.

[7]These instruments are different from those used in BPP, namely the cohort-education-year specific averages of male and female hourly wages, interacted with time, education, and kids dummies. The use of their instruments over our sample period often resulted in convergence problems.

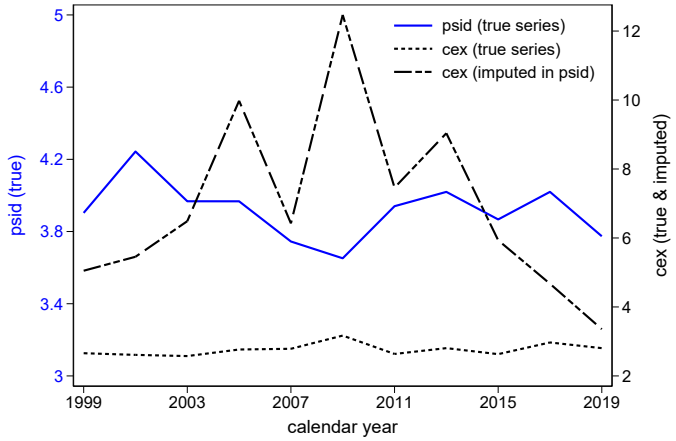
Figure C.1: Moments of log consumption in the PSID and CEX




(a) Variance of log consumption




(b) Skewness of log consumption




(c) Kurtosis of log consumption

*Notes:* The figure plots the moments of log consumption in the PSID (blue solid line; left), the CEX (black dotted line; right), and imputed from the CEX into the PSID (black dashed-dotted line; right axis).
*Alt text:* Three line graphs of log-consumption moments, 1999-2019, for PSID, CEX, and imputed CEX: variance shows similar trends across datasets, remaining roughly 0.2-0.4; skewness declines in PSID from about 0.52 to 0.3, is flat in CEX around -0.1, and is U-shaped in imputed CEX, with a high near -0.2 and a low near -1.5; kurtosis is flat in PSID and CEX at roughly 2–4, but imputed CEX displays an inverse U-shape, peaking around 12 and bottoming near 3.

different between the true PSID and the imputed series, as are also the trends. Imputed consumption is more volatile than the true PSID, skewness is negative (as opposed to positive in the PSID) and peaks down in 2007–2011, kurtosis is often more than twice as high as kurtosis in the PSID and peaks upwards in 2007–2011.

Table 1 in the text reports the second and higher-moments of consumption growth in the imputed series, residualized through similar first-stage regressions as in the true data. The main qualitative features of the distributions remain similar across the true PSID data and the imputed series, although the moments of imputed consumption are several-fold accentuated relative to the true data. The imputation likely imparts substantial measurement/imputation error to consumption, which casts doubt on the reliability of the partial insurance estimates derived from imputed consumption data from the CEX.

**An explanation for why the imputation inflates $\phi_t^{(1)}$.** Our empirical findings suggest that estimation of the linear specification using imputed consumption data inflates $\phi_t^{(1)}$ relative to estimation using the true PSID series. One reason behind this is neglected correlation between measurement error in income and expenditure.

Suppose that consumption and income measurement errors are positively correlated, i.e., $\text{Cov}(u_{it}^c, u_{it}^y) = \sigma_{u_t} > 0$. The positive correlation may be the result of, e.g., recall bias that affects similarly all monetary amounts. This is something we have not allowed for in the paper, and neither have BPP. While $\phi_t^{(1)}$ is identified via $\mathbb{E}(\Delta c_{it}(\Delta y_{it-1}+\Delta y_{it}+\Delta y_{it+1}))/\sigma_{\zeta_t}^2$ in the baseline permanent-transitory income process, in practice BPP and virtually all follow-up papers estimate it through $(\mathbb{E}(\Delta c_{it}\Delta y_{it}) + \mathbb{E}(\Delta c_{it}\Delta y_{it+1}))/\sigma_{\zeta_t}^2$, assuming that $\sigma_{u_t} = 0$ (uncorrelated errors). This is what we also do in our baseline estimation of $\phi_t^{(1)}$ in the linear specification. However, if the errors *are* correlated, then this target moment in fact identifies $\phi_t^{(1)\text{psid}^*} = \phi_t^{(1)} + \sigma_{u_t}/\sigma_{\zeta_t}^2$. Clearly, correlation between consumption and income errors causes positive bias as $\phi_t^{(1)\text{psid}^*} > \phi_t^{(1)}$. This is true in the baseline, i.e., without the use of imputed consumption, but the bias becomes multiplicatively larger with imputed consumption, as we show below.

In the simplest form without covariates, imputed log consumption in the PSID is given by $c_{it}^{\text{imp}} = \kappa_t(f_{it} - \pi_t)$, where $f_{it}$ is log food expenditure in the PSID, $\pi_t = \mathbf{p}_t'\boldsymbol{\gamma}$ is an intercept that depends on prices, and $\kappa_t = \beta(D_{it})^{-1}$ is the inverse budget elasticity of food consumption. Suppose that $\pi_t = 0$ for simplicity, so $c_{it}^{\text{imp}} = \kappa_t f_{it}$. As $\beta(D_{it})$ is typically less than 1 (e.g., Blundell et al., 2004, BPP), it follows that $\kappa_t > 1$. Blundell et al. (2004) discuss extensively the demand estimation in the CEX, and provide an in-depth characterization of the bias induced by endogenous expenditure or non-classical measurement error in the CEX. However, they do not discuss measurement error in food expenditure *in the PSID*, and in

particular the case when such error may correlate with error in income.

Suppose that food expenditure in the PSID is measured with error, so imputed consumption is $c_{it}^{\text{imp}^*} = \kappa_t(f_{it} + u_{it}^f)$. Suppose that food consumption and income error are positively correlated, i.e., $\text{Cov}(u_{it}^f, u_{it}^y) = \sigma_{u_t} > 0$. Because $\kappa_t > 1$, the consumption imputation magnifies the role of the consumption measurement error. One can show that, with imputed consumption, the moment used to identify $\phi_t^{(1)}$ in fact identifies $\phi_t^{(1)\text{imp}^*} = \phi_t^{(1)} + \kappa_t \sigma_{u_t}/\sigma_{\zeta_t}^2$. As a result, $\phi_t^{(1)\text{imp}^*} > \phi_t^{(1)\text{psid}^*}$ and the imputation inflates the transmission parameter of permanent shocks vis-à-vis the true consumption series.

## C.5 Estimation using the original BPP data

Table C.5 uses the original BPP data over 1980–1992 (taken from the replication package of Blundell et al., 2008, available online) and does three things.

First, we replicate BPP over the original data/period of time, using our code but BPP's *annual* growth rates of income and consumption. This is column 1 in table C.5. We obtain numerically similar results to BPP, namely $\phi^{(1)} = 0.64$ while $\psi^{(1)}$ is statistically zero.

Second, we calculate *biennial* growth rates of income and consumption and re-estimate the linear model. This is column 2 in table C.5; this estimation is analogous to the linear function estimation in column 1 of table 3, albeit using the original BPP sample. Note that despite the use of biennial data, $\phi^{(1)}$ and $\psi^{(1)}$ still express the pass-through at annual rates (see appendix B). $\phi^{(1)}$ drops by 20% to 0.518, while $\psi^{(1)}$ remains small but turns statistically significant. With biennial data, we only identify the pass-though of an 'aggregate' permanent shock over two years. Some higher frequency shocks are muted at the observed lower frequency, given that we can only observe an average of two one-year permanent shocks over the two-year period. In other words, the biennial data allow us to measure only an 'aggregate' or average pass-through within a given period, and some higher-frequency shocks are inevitably smoothed out at this observed lower frequency.

Third, we re-estimate the linear model assuming that the target second-order income and consumption moments are time-invariant. This differs from the original BPP exercise where the moments vary with time, but it is otherwise similar to our main exercise in the paper given that we cannot precisely estimate higher-order moments on a year-by-year basis. The results are in columns 3-4 in table C.5. Time-invariance further reduces the transmission parameter of permanent shocks ($\phi^{(1)}$ drops to 0.404), bringing the pass-through closer to the value we estimate in the modern data. Taking averages of moments over time attenuates the largest income-consumption co-movements; consequently, the model can fit the data using, ceteris paribus, smaller transmission parameters. This is similar to Chatterjee et al.

Table C.5: Replication of BPP, original BPP data 1980-1992

| *Consumption fn.:* | **Linear** | | | |
|---|---|---|---|---|
| *Income moments:* | $2^{nd}$ | $2^{nd}$ | $2^{nd}$ | $2^{nd}$ |
| *Consumption data:* | CEX | CEX | CEX | CEX |
| *Time variability:* | **Time-varying moments** | | **Time-invariant moments** | |
| *Growth rates:* | annual (BPP*) | biennial | annual | biennial |
| | (1) | (2) | (3) | (4) |
| $\phi^{(1)}$ | 0.644 | 0.518 | 0.504 | 0.404 |
| | (0.079) | (0.068) | (0.076) | (0.057) |
| $\psi^{(1)}$ | 0.024 | 0.117 | 0.043 | 0.158 |
| | (0.044) | (0.050) | (0.044) | (0.051) |
| $\theta^{\#}$ | 0.112 | 0 | 0.115 | 0 |
| | (0.025) | (not identified) | (0.026) | (not identified) |

*Notes:* The table presents the estimates of the parameters of the linear consumption function, using the original BPP data over 1980–1992 (PSID income data, consumption data imputed from the CEX). The data are available through Blundell et al. (2008)'s online replication package. Columns 1-2 allow the target second-order income and consumption moments to vary with calendar time, while columns 3-4 target said moments imposing time-invariance. Columns 1 and 3 estimate the model using annual growth rates of income and consumption, while columns 2 and 4 use biennial rates. In line with BPP, estimation is done via diagonally weighted GMM; GMM standard errors are in parentheses.

*The small unimportant numerical differences between column 1 herein and the original BPP estimates are likely due to different software and optimization algorithms (our code uses Matlab and `fmincon` while BPP use Gauss).

#$\theta$ is the moving average parameter of the transitory shock. BPP assume that the transitory component of income is given by $v_{it} = \epsilon_{it} + \theta\epsilon_{it-1}$, where $\epsilon_{it}$ is the transitory shock. Given their use of annual growth rates, $\theta$ is identified by the first-order income autocovariance. Our use of biennial data in this paper rules out identification of $\theta$, which we thus set to 0.

(2021) who also find smaller pass-through rates (higher degree of partial insurance) when they impose time-invariance in BPP's target moments.

There are some caveats comparing the exercise in table C.5 and our results from the recent PSID in table 3. First, our main exercise uses consumption data internally available in the PSID, while table C.5 uses consumption imputed from the CEX. We showed earlier that the imputation imparts substantial imputation error. Second, while the estimation in table C.5 does not correct the income moments for measurement error (in line with BPP), our main exercise in table 3 does so (in line with Blundell et al., 2016). Finally, the use of biennial growth rates does not allow the identification of $\theta$, the moving average parameter of the transitory shock, so the transitory shock is $MA(0)$ in our baseline ($\theta = 0$), while it is $MA(1)$ in the original BPP exercise.

## C.6 Additional results

Tables C.6 and C.7 present the estimates of the parameters of the income process and the consumption function using the empirical covariance matrix of the target moments as weighting matrix (optimal GMM). As such, these tables accompany tables 2 and 3 in the paper, which report results from equally weighted GMM.

Table C.8 presents the estimates of the parameters of the consumption function using imputed consumption from the CEX. As such, this table accompanies table 3 in the paper, which reports results using consumption internally available in the PSID. In the linear specification, $\phi^{(1)}$ is consistently higher than that obtained from the consumption data internally available in the PSID, regardless of whether we target only second- or also higher-order moments. This echoes our earlier discussion about the imputation inflating the pass-through of shocks. We also find that $\psi^{(1)}$ is negative in all cases. Moreover, when up to fourth-order moments are targeted (column 3), the point estimate is particularly large in magnitude ($-0.418$). This is counterintuitive, raising further concerns about the reliability of the imputation procedure, particularly when higher-order moments are involved. We draw similar conclusions about $\phi^{(1)}$ and $\psi^{(1)}$ in the context of the quadratic function. $\phi^{(2)} < 0$ in that case, with a magnitude similar to our baseline results from the PSID ($-0.036$ versus $-0.04$; albeit not significant). Once again, this is indicative of the asymmetric pass-through of negative versus positive permanent shocks. Similar to our baseline, we find $\psi^{(2)} > 0$.

Table C.9 presents the estimates of the income process by age and the household's position in the income distribution. Skewness of all shocks is more negative at older ages. Skewness of permanent shocks is weakly positive in the bottom income tertile but turns very negative in the top tertile. The opposite is true for transitory shocks; skewness is negative in the bottom income tertile but strongly positive in the top.

Tables C.10 and C.11 present results from the linear consumption function by age and the household's position in the income distribution, respectively. $\phi^{(1)}$ exhibits a non-monotonic pattern over the three age brackets – decreasing from 0.154 over ages 30–40 to 0.126 over 41–50 before increasing to 0.180 over ages 51–65. A similar non-monotonic pattern appears in column 2, where third moments are also targeted, and is further accentuated in column 3 where all moments up to fourth-order are targeted. This is indicative of a misspecification in the linear function, in the sense that its theoretical restrictions (higher wealth insures income fluctuations better) are not supported by the empirical results.[8] By contrast, $\psi^{(1)}$ is statistically indistinguishable from zero in all three age brackets. $\phi^{(1)}$ is uniformly lower in

[8]This conclusion is consistent with BPP, who find a decreasing pattern with age only when they impose a linear age trend on the transmission parameters. By contrast, when they generalize this to a quadratic trend, they report a worsening of their results without including further details.

Table C.6: Estimates of the income process, optimal weighting

| *Income moments:* | | | $2^{nd}$ | | $2^{nd}$, $3^{rd}$ | | $2^{nd}$, $3^{rd}$, $4^{th}$ | |
|---|---|---|---|---|---|---|---|---|
| *Income data:* | | | PSID | | PSID | | PSID | |
| | | | (1) | | (2) | | (3) | |
| **Panel A. Permanent shocks** | | | | | | | | |
| Variance | $\sigma_\zeta^2$ | | 0.027 | (0.002) | 0.027 | (0.002) | 0.025 | (0.002) |
| Skewness | $\gamma_\zeta$ | central | | | -0.004 | (0.003) | -0.004 | (0.003) |
| | | standardized | | | -0.895 | (0.633) | -0.997 | (0.760) |
| Kurtosis | $\kappa_\zeta$ | central | | | | | 0.037 | (0.007) |
| | | standardized | | | | | 61.636 | (14.018) |
| **Panel B. Transitory shocks** | | | | | | | | |
| Variance | $\sigma_v^2$ | | 0.033 | (0.002) | 0.033 | (0.002) | 0.034 | (0.002) |
| Skewness | $\gamma_v$ | central | | | -0.008 | (0.003) | -0.008 | (0.003) |
| | | standardized | | | -1.351 | (0.517) | -1.324 | (0.485) |
| Kurtosis | $\kappa_v$ | central | | | | | 0.049 | (0.005) |
| | | standardized | | | | | 41.579 | (7.052) |

*Notes:* The table presents the estimates of the parameters of the income process, imposing stationarity over time. Column 1 targets only the second moments of income growth; column 2 targets also third moments; column 3 targets all moments up to fourth-order. Estimation is via GMM using the optimal weighting matrix; block bootstrap standard errors are in parentheses. This table accompanies table 2 in the paper, which shows the baseline results from equally weighted GMM.

the top income tertile vis-à-vis the bottom tertile, suggesting that the income-rich, who are also wealthier, have more insurance against permanent income shocks than the poor.

Table C.12 presents estimation results for the linear consumption function in the subsamples formed by wealth and education. The odd-numbered columns present parameter estimates when only second-order moments are targeted. $\phi^{(1)}$ becomes monotonically smaller as we move from group 1 (low wealth, no college) to group 4 (high wealth, some college), reflecting the insurance role of wealth, which increases monotonically across groups. This is similar to what we find in the quadratic specification. However, as soon as we target up to fourth-order moments in the even-numbered columns, this monotonic pattern is lost. Transitory shocks transmit strongly in group 1, and with a negative coefficient.

Table C.13 presents the estimates of the income process, when permanent income follows an AR(1) process, i.e., it features linear persistence. Table C.14 presents the associated estimates of the parameters of the linear and quadratic consumption functions.

Table C.7: Estimates of the consumption function, optimal weighting

| *Consumption fn.:* | **Linear** | | | | **Quadratic** |
|---|---|---|---|---|---|
| *Income moments:* | $2^{\text{nd}}$ | $2^{\text{nd}}$ | $2^{\text{nd}}$, $3^{\text{rd}}$ | $2^{\text{nd}}$, $3^{\text{rd}}$, $4^{\text{th}}$ | $2^{\text{nd}}$, $3^{\text{rd}}$, $4^{\text{th}}$ |
| *Consumption data:* | PSID | CEX | PSID | PSID | PSID |
| | (1) | (2) | (3) | (4) | (5) |
| $\phi^{(1)}$ | 0.152 | 0.288 | 0.149 | 0.146 | 0.169 |
| | (0.028) | (0.044) | (0.025) | (0.029) | (0.030) |
| $\psi^{(1)}$ | -0.006 | -0.109 | -0.015 | -0.004 | -0.018 |
| | (0.045) | (0.078) | (0.037) | (0.041) | (0.047) |
| $\phi^{(2)}$ | | | | | -0.037 |
| | | | | | (0.022) |
| $\psi^{(2)}$ | | | | | -0.007 |
| | | | | | (0.030) |
| $\omega^{(22)}$ | | | | | 0.740 |
| | | | | | (0.731) |
| $\sigma^2_\xi$ | 0.019 | 0.019 | 0.019 | 0.018 | 0.018 |
| | (0.001) | (0.004) | (0.001) | (0.001) | (0.001) |
| $\gamma_\xi$ | | | -0.001 | -0.001 | |
| | | | (0.001) | (0.001) | |
| $\kappa_\xi$ | | | | 0.005 | |
| | | | | (0.001) | |
| $\sigma^2_{u_c}$ | 0.044 | 0.140 | 0.043 | 0.043 | 0.044 |
| | (0.002) | (0.009) | (0.001) | (0.001) | (0.001) |
| $\gamma_{u_c}$ | | | 0.002 | 0.002 | |
| | | | (0.001) | (0.001) | |
| $\kappa_{u_c}$ | | | | 0.012 | |
| | | | | (0.001) | |

*Notes:* The table presents the estimates of the parameters of the consumption function, imposing homogeneous transmission parameters over the lifecycle/in the cross-section and stationarity of taste heterogeneity and consumption measurement error. Columns 1-4 present parameter estimates in the linear function, while column 5 presents estimates in the quadratic case; the order of moments targeted in each case is shown at the top of the table. Except for column 2 where we use imputed data from the CEX, all other columns use consumption data internally available in the PSID. Estimation is via GMM using the optimal weighting matrix; block bootstrap standard errors are in parentheses. This table accompanies table 3 in the paper, which shows the baseline results from equally weighted GMM.

Table C.8: Estimates of the consumption function, imputed consumption

| *Consumption fn.:* | **Linear** | | | **Quadratic** |
|---|---|---|---|---|
| *Income moments:* | $2^{\text{nd}}$ | $2^{\text{nd}}$, $3^{\text{rd}}$ | $2^{\text{nd}}$, $3^{\text{rd}}$, $4^{\text{th}}$ | $2^{\text{nd}}$, $3^{\text{rd}}$, $4^{\text{th}}$ |
| *Consumption data:* | CEX | CEX | CEX | CEX |
| | (1) | (2) | (3) | (4) |
| $\phi^{(1)}$ | 0.288 | 0.339 | 0.404 | 0.260 |
| | (0.044) | (0.074) | (0.118) | (0.043) |
| $\psi^{(1)}$ | -0.109 | -0.155 | -0.418 | -0.108 |
| | (0.078) | (0.098) | (0.247) | (0.086) |
| $\phi^{(2)}$ | | | | -0.036 |
| | | | | (0.038) |
| $\psi^{(2)}$ | | | | 0.025 |
| | | | | (0.045) |
| $\omega^{(22)}$ | | | | -0.157 |
| | | | | (1.195) |
| $\sigma^2_\xi$ | 0.019 | 0.018 | 0.000 | 0.019 |
| | (0.004) | (0.004) | (0.004) | (0.004) |
| $\gamma_\xi$ | | -0.016 | -0.016 | |
| | | (0.007) | (0.007) | |
| $\kappa_\xi$ | | | 0.000 | |
| | | | (0.000) | |
| $\sigma^2_{u_c}$ | 0.140 | 0.140 | 0.152 | 0.140 |
| | (0.009) | (0.009) | (0.012) | (0.009) |
| $\gamma_{u_c}$ | | -0.185 | -0.185 | |
| | | (0.037) | (0.037) | |
| $\kappa_{u_c}$ | | | 0.816 | |
| | | | (0.183) | |

*Notes:* The table presents the estimates of the parameters of the consumption function, imposing homogeneous transmission parameters over the lifecycle/in the cross-section and stationarity of taste heterogeneity and consumption measurement error. Columns 1-3 present parameter estimates in the linear function, while column 4 presents estimates in the quadratic case; the order of moments targeted in each case is shown at the top of the table. All columns use consumption imputed from the CEX (see appendix C.4). Estimation is via equally weighted GMM; block bootstrap standard errors are in parentheses. This table accompanies table 3 in the paper, which shows the baseline results using consumption internally available in the PSID (column 2 in table 3 is the same as column 1 in the present table).

Table C.9: Estimates of the income process, by age and position in income distribution

| | | | Age | | | Position in income distribution $y_{it-1}$ | |
|---|---|---|---|---|---|---|---|
| | | | 30–40 | 41–50 | 51–65 | Bottom tertile | Top tertile |
| *Income moments:* | | | $2^{nd}$, $3^{rd}$, $4^{th}$ | $2^{nd}$, $3^{rd}$, $4^{th}$ | $2^{nd}$, $3^{rd}$, $4^{th}$ | $2^{nd}$, $3^{rd}$, $4^{th}$ | $2^{nd}$, $3^{rd}$, $4^{th}$ |
| *Income data:* | | | PSID | PSID | PSID | PSID | PSID |
| | | | (1) | (2) | (3) | (4) | (5) |
| **Panel A. Permanent shocks** | | | | | | | |
| Variance | $\sigma^2_\zeta$ | | 0.028 | 0.026 | 0.034 | 0.037 | 0.032 |
| | | | (0.004) | (0.004) | (0.005) | (0.004) | (0.003) |
| Skewness | $\gamma_\zeta$ | central | -0.003 | 0.004 | -0.015 | 0.010 | -0.028 |
| | | | (0.006) | (0.005) | (0.006) | (0.005) | (0.004) |
| | | standardized | -0.653 | 0.926 | -2.386 | 1.344 | -4.851 |
| | | | (1.291) | (1.114) | (1.167) | (0.783) | (0.987) |
| Kurtosis | $\kappa_\zeta$ | central | 0.043 | 0.030 | 0.058 | 0.052 | 0.039 |
| | | | (0.014) | (0.013) | (0.016) | (0.013) | (0.008) |
| | | standardized | 56.275 | 45.537 | 50.945 | 37.627 | 36.974 |
| | | | (23.787) | (22.709) | (21.605) | (13.058) | (10.586) |
| **Panel B. Transitory shocks** | | | | | | | |
| Variance | $\sigma^2_v$ | | 0.033 | 0.027 | 0.036 | 0.020 | 0.020 |
| | | | (0.005) | (0.004) | (0.005) | (0.004) | (0.003) |
| Skewness | $\gamma_v$ | central | -0.008 | -0.005 | -0.010 | -0.024 | 0.021 |
| | | | (0.006) | (0.005) | (0.005) | (0.005) | (0.004) |
| | | standardized | -1.427 | -1.178 | -1.453 | -8.246 | 7.333 |
| | | | (1.057) | (1.088) | (0.782) | (3.074) | (2.235) |
| Kurtosis | $\kappa_v$ | central | 0.056 | 0.040 | 0.050 | 0.014 | 0.011 |
| | | | (0.013) | (0.007) | (0.008) | (0.008) | (0.006) |
| | | standardized | 52.863 | 53.730 | 38.832 | 33.974 | 27.286 |
| | | | (19.670) | (16.619) | (12.566) | (24.531) | (16.935) |

*Notes:* The table presents the estimates of the parameters of the income process, allowing them to vary over ages 30–40, 41–50, and 51–65 of the male spouse, and, separately, by the household's position in the income distribution (bottom and top tertiles of the distribution of lag income $y_{it-1}$). We target all income moments up to fourth-order; results when targeting only lower order moments are available upon request. Estimation is via equally weighted GMM; GMM standard errors clustered at the household level are in parentheses. This table accompanies tables 4 and 5 in the paper, as well as the following tables C.10 and C.11.

Table C.10: Estimates of the linear consumption function, by age

| *Consumption fn.:* | | **Linear** | | | | | |
|---|---|---|---|---|---|---|---|
| *Income moments:* | | $2^{nd}$ | | $2^{nd}$, $3^{rd}$ | | $2^{nd}$, $3^{rd}$, $4^{th}$ | |
| *Consumption data:* | | PSID | | PSID | | PSID | |
| | | (1) | | (2) | | (3) | |
| $\phi^{(1)}$ | 30-40 | 0.154 | (0.045) | 0.164 | (0.051) | 0.089 | (0.058) |
| | 41-50 | 0.126 | (0.051) | 0.146 | (0.055) | 0.079 | (0.089) |
| | 51-65 | 0.180 | (0.049) | 0.175 | (0.051) | 0.136 | (0.061) |
| $\psi^{(1)}$ | 30-40 | 0.094 | (0.079) | 0.083 | (0.073) | 0.095 | (0.107) |
| | 41-50 | 0.010 | (0.079) | -0.016 | (0.072) | -0.075 | (0.061) |
| | 51-65 | -0.046 | (0.070) | -0.044 | (0.066) | 0.037 | (0.101) |
| $\sigma^2_\xi$ | 30-40 | 0.014 | (0.002) | 0.014 | (0.002) | 0.016 | (0.002) |
| | 41-50 | 0.022 | (0.002) | 0.022 | (0.002) | 0.022 | (0.002) |
| | 51-65 | 0.021 | (0.002) | 0.021 | (0.002) | 0.023 | (0.002) |
| $\gamma_\xi$ | 30-40 | | | -0.001 | (0.001) | -0.001 | (0.001) |
| | 41-50 | | | 0.006 | (0.001) | 0.006 | (0.002) |
| | 51-65 | | | -0.006 | (0.002) | -0.006 | (0.002) |
| $\kappa_\xi$ | 30-40 | | | | | 0.004 | (0.002) |
| | 41-50 | | | | | 0.007 | (0.002) |
| | 51-65 | | | | | 0.006 | (0.002) |
| $\sigma^2_{u_c}$ | 30-40 | 0.046 | (0.003) | 0.046 | (0.003) | 0.046 | (0.003) |
| | 41-50 | 0.040 | (0.002) | 0.040 | (0.002) | 0.040 | (0.002) |
| | 51-65 | 0.043 | (0.002) | 0.043 | (0.002) | 0.043 | (0.002) |
| $\gamma_{u_c}$ | 30-40 | | | 0.000 | (0.002) | 0.000 | (0.002) |
| | 41-50 | | | 0.002 | (0.002) | 0.002 | (0.002) |
| | 51-65 | | | 0.004 | (0.002) | 0.004 | (0.002) |
| $\kappa_{u_c}$ | 30-40 | | | | | 0.014 | (0.003) |
| | 41-50 | | | | | 0.012 | (0.002) |
| | 51-65 | | | | | 0.012 | (0.002) |

*Notes:* The table presents the estimates of the parameters of the linear consumption function, allowing them to vary over ages 30–40, 41–50, and 51–65 of the male spouse. The underlying income moments also vary over these age brackets (table C.9). All columns use consumption data internally available in the PSID. Estimation is via equally weighted GMM; block bootstrap standard errors are in parentheses. This table accompanies table 4 in the paper, which shows results from the quadratic specification.

Table C.11: Estimates of the linear consumption function, over position in income distribution

| *Consumption fn.:* | **Linear** | | | | | |
|---|---|---|---|---|---|---|
| *Lag income* $y_{it-1}$*:* | Bottom tertile | | | Top tertile | | |
| *Income moments:* | $2^{\text{nd}}$ | $2^{\text{nd}}$, $3^{\text{rd}}$ | $2^{\text{nd}}$, $3^{\text{rd}}$, $4^{\text{th}}$ | $2^{\text{nd}}$ | $2^{\text{nd}}$, $3^{\text{rd}}$ | $2^{\text{nd}}$, $3^{\text{rd}}$, $4^{\text{th}}$ |
| *Consumption data:* | PSID | PSID | PSID | PSID | PSID | PSID |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| $\phi^{(1)}$ | 0.103 | 0.083 | 0.063 | 0.102 | 0.060 | 0.036 |
| | (0.031) | (0.032) | (0.035) | (0.036) | (0.034) | (0.041) |
| $\psi^{(1)}$ | -0.045 | 0.013 | 0.025 | -0.004 | 0.101 | 0.071 |
| | (0.086) | (0.038) | (0.054) | (0.093) | (0.048) | (0.073) |
| $\sigma^2_\xi$ | 0.014 | 0.014 | 0.014 | 0.025 | 0.025 | 0.026 |
| | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) |
| $\gamma_\xi$ | | -0.001 | -0.001 | | 0.002 | 0.002 |
| | | (0.001) | (0.001) | | (0.001) | (0.001) |
| $\kappa_\xi$ | | | 0.003 | | | 0.009 |
| | | | (0.002) | | | (0.002) |
| $\sigma^2_{u_c}$ | 0.044 | 0.044 | 0.044 | 0.041 | 0.041 | 0.041 |
| | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) |
| $\gamma_{u_c}$ | | 0.003 | 0.003 | | 0.003 | 0.003 |
| | | (0.002) | (0.002) | | (0.002) | (0.002) |
| $\kappa_{u_c}$ | | | 0.011 | | | 0.010 |
| | | | (0.002) | | | (0.002) |

*Notes:* The table presents the estimates of the parameters of the linear consumption function in the bottom and top tertiles of the distribution of lag income $y_{it-1}$. The underlying income moments also vary over the income distribution (table C.9). All columns use consumption data internally available in the PSID. Estimation is via equally weighted GMM; GMM standard errors clustered at the household level are in parentheses. This table accompanies table 5 in the paper, which shows results from the quadratic specification.

Table C.12: Estimates of the linear consumption function, by wealth and education

| *Consumption fn.:* | **Linear** | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Low wealth | | | | High wealth | | | |
| *Subsample:* | No college | | Some college | | No college | | Some college | |
| *Income moments:* | $2^{nd}$ | $2^{nd}$, $3^{rd}$, $4^{th}$ | $2^{nd}$ | $2^{nd}$, $3^{rd}$, $4^{th}$ | $2^{nd}$ | $2^{nd}$, $3^{rd}$, $4^{th}$ | $2^{nd}$ | $2^{nd}$, $3^{rd}$, $4^{th}$ |
| *Consumption data:* | PSID | PSID | PSID | PSID | PSID | PSID | PSID | PSID |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| $\phi^{(1)}$ | 0.242 | 0.143 | 0.198 | 0.153 | 0.171 | 0.028 | 0.099 | 0.106 |
| | (0.053) | (0.042) | (0.053) | (0.056) | (0.071) | (0.089) | (0.041) | (0.055) |
| $\psi^{(1)}$ | -0.178 | -0.164 | 0.015 | 0.073 | 0.056 | 0.142 | 0.003 | 0.004 |
| | (0.107) | (0.077) | (0.089) | (0.091) | (0.089) | (0.131) | (0.059) | (0.071) |
| $\sigma^2_\xi$ | 0.014 | 0.016 | 0.013 | 0.014 | 0.014 | 0.015 | 0.025 | 0.026 |
| | (0.003) | (0.003) | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) |
| $\gamma_\xi$ | | -0.003 | | -0.001 | | -0.003 | | 0.000 |
| | | (0.002) | | (0.001) | | (0.002) | | (0.001) |
| $\kappa_\xi$ | | 0.005 | | 0.000 | | 0.000 | | 0.009 |
| | | (0.002) | | (0.002) | | (0.002) | | (0.002) |
| $\sigma^2_{u_c}$ | 0.045 | 0.045 | 0.045 | 0.045 | 0.038 | 0.038 | 0.043 | 0.043 |
| | (0.003) | (0.003) | (0.003) | (0.003) | (0.004) | (0.004) | (0.002) | (0.002) |
| $\gamma_{u_c}$ | | -0.001 | | 0.004 | | 0.010 | | 0.002 |
| | | (0.003) | | (0.002) | | (0.005) | | (0.001) |
| $\kappa_{u_c}$ | | 0.013 | | 0.013 | | 0.018 | | 0.011 |
| | | (0.003) | | (0.002) | | (0.006) | | (0.002) |

*Notes:* The table presents the estimates of the parameters of the linear consumption function, allowing them to vary by wealth and education. The underlying income moments also vary by wealth and education (results not shown for brevity). We do not report results when up to third-order moments of income and consumption are targeted, but these are available upon request. Low (high) wealth is defined on the basis of household wealth being less (more) than median real wealth in the sample over 1999–2019. No college versus some college is defined on the basis of the highest level of education attained by the male spouse (cutoff is 12 years of schooling). All columns use consumption data internally available in the PSID. Estimation is via equally weighted GMM; GMM standard errors clustered at the household level are in parentheses. This table accompanies the quadratic specification estimates in table 6 in the paper.

Table C.13: Estimates of the income process, income persistence

| *Income moments:* | | | $2^{nd}$ | | $2^{nd}$, $3^{rd}$ | | $2^{nd}$, $3^{rd}$, $4^{th}$ | |
|---|---|---|---|---|---|---|---|---|
| *Income data:* | | | PSID | | PSID | | PSID | |
| | | | (1) | | (2) | | (3) | |
| **Panel A. Permanent shocks** | | | | | | | | |
| Variance | $\sigma_\zeta^2$ | | 0.058 | (0.016) | 0.056 | (0.016) | 0.036 | (0.006) |
| Skewness | $\gamma_\zeta$ | central | | | -0.007 | (0.003) | -0.006 | (0.003) |
| | | standardized | | | -0.503 | (0.337) | -0.910 | (0.512) |
| Kurtosis | $\kappa_\zeta$ | central | | | | | 0.039 | (0.007) |
| | | standardized | | | | | 29.891 | (11.459) |
| **Panel B. Transitory shocks** | | | | | | | | |
| Variance | $\sigma_v^2$ | | 0.011 | (0.015) | 0.014 | (0.014) | 0.030 | (0.005) |
| Skewness | $\gamma_v$ | central | | | -0.005 | (0.003) | -0.006 | (0.003) |
| | | standardized | | | -3.272 | (5.415) | -1.178 | (0.694) |
| Kurtosis | $\kappa_v$ | central | | | | | 0.047 | (0.006) |
| | | standardized | | | | | 52.091 | (19.426) |
| **Panel C. Linear persistence** | | | | | | | | |
| AR(1) | $\rho$ | | 0.846 | (0.038) | 0.854 | (0.038) | 0.931 | (0.028) |

*Notes:* The table presents the estimates of the parameters of the AR(1) income process. Column 1 targets only the second moments of income growth; column 2 targets also third moments; column 3 targets all moments up to fourth-order. Estimation is via equally weighted GMM; block bootstrap standard errors are in parentheses. This table accompanies table 7 in the paper, as well as table C.14 that follows.

Table C.14: Estimates of the consumption function, income persistence

| *Consumption fn.:* | **Linear** | | | **Quadratic** |
|---|---|---|---|---|
| *Income moments:* | $2^{\text{nd}}$ | $2^{\text{nd}}$, $3^{\text{rd}}$ | $2^{\text{nd}}$, $3^{\text{rd}}$, $4^{\text{th}}$ | $2^{\text{nd}}$, $3^{\text{rd}}$, $4^{\text{th}}$ |
| *Consumption data:* | PSID | PSID | PSID | PSID |
| | (1) | (2) | (3) | (4) |
| $\phi^{(1)}$ | 0.089 | 0.090 | 0.110 | 0.112 |
| | (0.029) | (0.029) | (0.041) | (0.024) |
| $\psi^{(1)}$ | -0.139 | -0.077 | -0.013 | -0.010 |
| | (0.265) | (0.083) | (0.077) | (0.066) |
| $\phi^{(2)}$ | | | | -0.049 |
| | | | | (0.026) |
| $\psi^{(2)}$ | | | | 0.019 |
| | | | | (0.037) |
| $\omega^{(22)}$ | | | | 0.430 |
| | | | | (0.847) |
| $\sigma^2_\xi$ | 0.019 | 0.019 | 0.020 | 0.019 |
| | (0.001) | (0.001) | (0.001) | (0.001) |
| $\gamma_\xi$ | | -0.001 | -0.001 | |
| | | (0.001) | (0.001) | |
| $\kappa_\xi$ | | | 0.031 | |
| | | | (0.002) | |
| $\sigma^2_{u_c}$ | 0.044 | 0.044 | 0.044 | 0.044 |
| | (0.002) | (0.002) | (0.001) | (0.002) |
| $\gamma_{u_c}$ | | 0.002 | 0.002 | |
| | | (0.001) | (0.001) | |
| $\kappa_{u_c}$ | | | 0.013 | |
| | | | (0.001) | |

*Notes:* The table presents the estimates of the parameters of the consumption function, when income features linear persistence. Columns 1-3 present parameter estimates in the linear function, while column 4 presents estimates in the quadratic case; the order of moments targeted in each case is shown at the top of the table. Estimation is via equally weighted GMM; block bootstrap standard errors are in parentheses. This table accompanies table 7 in the paper.

# D Some welfare costs of tail income risk

Given our empirical, quasi reduced-form approach to measuring partial insurance, we calculate the welfare costs of income risk outside of a fully specified structural model. Specifically, we draw a sequence of income shocks over the lifetime, we use the different consumption specifications to simulate responses to those shocks, and we use parametric utility to calculate welfare based on those consumption responses. Our focus is on the costs of non-Gaussian relative to (counterfactual) Gaussian income risk, and we are particularly interested in contrasting the implications of linear versus nonlinear transmission.

Normal mixtures offer a flexible way of modeling non-Gaussian distributions and have by now become a standard choice in quantitative income dynamics (Guvenen et al., 2021, 2024). We draw permanent and transitory income shocks for 50,000 households from a mixture of two normal distributions, namely

$$\zeta_{it} \sim \begin{cases} \mathcal{N}(\mu_{\zeta,1}, \sigma_{\zeta,1}) \text{ with prob. } p_\zeta \\ \mathcal{N}(\mu_{\zeta,2}, \sigma_{\zeta,2}) \text{ with prob. } 1-p_\zeta \end{cases} \quad \text{and} \quad v_{it} \sim \begin{cases} \mathcal{N}(\mu_{v,1}, \sigma_{v,1}) \text{ with prob. } p_v \\ \mathcal{N}(\mu_{v,2}, \sigma_{v,2}) \text{ with prob. } 1-p_v. \end{cases}$$

There are five parameters per shock $(p, \mu_1, \sigma_1, \mu_2, \sigma_2)$, which are estimated using the method of simulated moments to match the first four moments of shocks in the PSID (column 3, table 2). We assume $\mu_1 < 0$ for identification. Table D.1 reports the parameter estimates and the (excellent) model fit. In the counterfactual scenario with Gaussian risk, we fix $p = 0$, $\mu_2 = 0$, and set $\sigma_2^2$ equal to the empirical variance of each shock.

We simulate alternative consumption profiles in response to the income shocks using the linear and nonlinear specifications

$$\begin{aligned} \Delta \ln C_{it} &= \mu_t^c + \phi_{it}^{(1)} \zeta_{it} + \psi_{it}^{(1)} v_{it}, \\ \Delta \ln C_{it} &= \mu_t^c + \phi_{it}^{(1)} \zeta_{it} + \psi_{it}^{(1)} v_{it} + \phi_{it}^{(2)} \zeta_{it}^2 + \psi_{it}^{(2)} v_{it}^2 + \omega_{it}^{(22)} \zeta_{it} v_{it}. \end{aligned}$$

These reflect the linear and quadratic consumption functions (4) and (5) in the paper, with $\mu_t^c = f(\Delta \mathbf{Z}_{it})$ (deterministic profile) and $\xi_{it} = 0$ (mean taste shock). In practice, we let $\mu_t^c$ equal the average (non-residualized) consumption growth by year, which we estimate in the data. We simulate consumption over a working life of 35 years, starting at age 30.[9] In the baseline, we use the time-invariant homogeneous partial insurance parameters from table 3, but we subsequently relax time-invariance and homogeneity across households.

[9] We also simulate a retirement period of 20 years, starting at age 65, in which absence of income shocks implies $\Delta \ln C_{it} = \mu_t^c$. We fix the initial condition at age 30 ($\ln C_{it=30}$) to its mean value in the sample. In all the simulations, we restrict $\omega_{it}^{(22)} = 0$ as this is never precisely estimated in any cut of the sample.

Table D.1: Normal mixtures, parameter estimates and model fit

| | | | | | |
|---|---|---|---|---|---|
| **Panel A. Parameter estimates** | | | | | |
| Permanent | $p_\zeta$ | $\mu_{\zeta,1}$ | $\sigma_{\zeta,1}$ | $\mu_{\zeta,2}$ | $\sigma_{\zeta,2}$ |
| | 0.058 | -0.049 | 0.683 | 0.003 | 0.053 |
| Transitory | $p_v$ | $\mu_{v,1}$ | $\sigma_{v,1}$ | $\mu_{v,2}$ | $\sigma_{v,2}$ |
| | 0.059 | -0.093 | 0.713 | 0.006 | 0.020 |
| **Panel B. Model fit permanent shocks** | | | | | |
| | | | | ***Model*** | ***Data*** |
| Variance | $\sigma^2_\zeta$ | | | 0.030 | 0.030 |
| | | | | | (0.002) |
| Skewness | $\gamma_\zeta$ | central | | -0.004 | -0.004 |
| | | | | | (0.003) |
| | | standardized | | -0.757 | -0.831 |
| | | | | | (0.571) |
| Kurtosis | $\kappa_\zeta$ | central | | 0.038 | 0.039 |
| | | | | | (0.007) |
| | | standardized | | 43.146 | 44.393 |
| | | | | | (11.080) |
| **Panel C. Model fit transitory shocks** | | | | | |
| | | | | ***Model*** | ***Data*** |
| Variance | $\sigma^2_v$ | | | 0.031 | 0.031 |
| | | | | | (0.003) |
| Skewness | $\gamma_v$ | central | | -0.008 | -0.008 |
| | | | | | (0.003) |
| | | standardized | | -1.520 | -1.431 |
| | | | | | (0.566) |
| Kurtosis | $\kappa_v$ | central | | 0.047 | 0.048 |
| | | | | | (0.005) |
| | | standardized | | 49.630 | 49.226 |
| | | | | | (10.119) |

*Notes:* The table reports the parameter estimates (panel A) and the model fit (panels B & C) of the quantitative model for income shocks (mixture of normal distributions). Estimation is done via the method of simulated moments, based on 50,000 random draws for each type of income shock.

We calculate the (ex-ante) welfare implied by a given consumption profile $\mathcal{C}_i = \{C_{it}\}_{t=0}^T$ using constant coefficient of relative risk aversion utility $\mathcal{U}(\mathcal{C}_i) = \sum_{t=0}^T \beta^t \frac{C_{it}^{1-\gamma}-1}{1-\gamma}$ if $\gamma \neq 1$, and logarithmic utility $\mathcal{U}(\mathcal{C}_i) = \sum_{t=0}^T \beta^t \log(C_{it})$ if $\gamma = 1$. We then calculate the welfare cost of tail risk as the percentage reduction in lifetime consumption that a household who faces Gaussian income risk is willing to accept in order to avoid tail (namely, non-Gaussian) risk. This is given by $\alpha_i$ for a given household $i$, which, similar to the welfare calculations in Wu and Krueger (2021), satisfies the indifference equation $\mathcal{U}\big((1-\alpha_i)\mathcal{C}_i^*\big) = \mathcal{U}\big(\mathcal{C}_i\big)$. Here, $\mathcal{C}_i^*$ and

Table D.2: Welfare costs of non-Gaussian relative to Gaussian income risk

| *Specification:* | *% households*$^{\#}$ | Average welfare costs $\mathbb{E}(\alpha_i)$, $\gamma = 2$ | | |
|---|---|---|---|---|
| | | $\mathcal{C}_i^*$ linear<br>$\mathcal{C}_i$ linear<br>(1) | $\mathcal{C}_i^*$ quadratic<br>$\mathcal{C}_i$ quadratic<br>(2) | $\mathcal{C}_i^*$ linear<br>$\mathcal{C}_i$ quadratic<br>(3) |
| All households | 100% | -0.81 | -0.88 | 0.66 |
| $\min\{\zeta_{it}\} < -0.50$ | 40.31% | 4.80 | 6.39 | 7.82 |
| $\min\{\zeta_{it}\} < -0.70$ | 29.11% | 6.36 | 8.65 | 10.04 |
| $\min\{\zeta_{it}\} < -0.95$ | 17.07% | 8.54 | 11.89 | 13.23 |

*Notes:* The table reports the welfare costs of non-Gaussian relative to Gaussian income risk, for $\gamma = 2$. We simulate consumption for 50,000 households using the time-invariant transmission parameters of table 3. The first row presents the average costs across all households. The next rows present the average costs over subsets of households defined by the magnitude of the worst permanent shock they experience. $^{\#}$% of households who face a shock at least as bad, drawn from the normal mixture, at least once in their lifetime.

$\mathcal{C}_i$ are alternative consumption profiles that reflect *less* (Gaussian) and *more* (non-Gaussian) income risk, respectively. We consider three alternative scenarios over which we calculate the welfare costs of non-Gaussian income risk.

In a first scenario, both $\mathcal{C}_i^*$ and $\mathcal{C}_i$ are simulated from the *linear* specification in response to Gaussian and non-Gaussian income shocks, respectively. The estimate of $\alpha_i$ in this case reflects the welfare cost of tail risk in the context of linear transmission. Tail risk is still costly in this context because the probability of very negative shocks (reflecting the long left fat tail) is higher relative to a Gaussian density.

In a second scenario, both $\mathcal{C}_i^*$ and $\mathcal{C}_i$ are simulated from the *quadratic* specification in response to Gaussian and non-Gaussian income shocks, respectively. The estimate of $\alpha_i$ here reflects the welfare cost of tail risk in the presence of nonlinear transmission. Given the evidence that larger negative shocks have a stronger impact on consumption when nonlinear transmission is allowed for, we expect larger welfare costs of tail risk in this scenario.

In a third scenario, the less risky profile $\mathcal{C}_i^*$ reflects the *linear* specification in response to Gaussian risk, while the riskier profile $\mathcal{C}_i$ is simulated from the *quadratic* specification in response to non-Gaussian shocks. Here $\alpha_i$ reflects both the welfare cost of tail risk *and* the bias induced by neglecting the possibility of nonlinear transmission of such risk.

Table D.2 averages $\alpha_i$ across households, for $\gamma = 2$. The first row presents the unconditional average cost across all households. The following rows present the average costs over subsets of households defined by the magnitude of the worst permanent shock they experience. We also count the fraction of households who face a shock at least as bad, drawn from the normal mixture, *at least once* in their lifetime. Two main points stand out.

First, the unconditional cost is small but negative in both the first (linear) and second (nonlinear) scenarios, suggesting households may in fact prefer tail *over* Gaussian risk. While this may seem counterintuitive, table D.1 shows that tail risk entails a small probability (about 5%) of shocks from a distribution with negative mean and large variance, and a high probability of shocks from one with a small positive mean and very low variance. Most households never actually experience shocks from the former; instead, they face the latter, whose variance is much smaller than that of Gaussian risk. Counterfactually, Gaussian risk implies larger variability for many, so households are *on average* willing to pay to move to the non-Gaussian setting. However, the average cost turns positive in scenario 3, which evaluates the consumption loss of tail risk and the bias from a misspecified consumption function. That is, neglecting non-Gaussian features and nonlinear transmission leads to an overestimation of consumer welfare by 0.66% of lifetime consumption.

Second, the welfare costs become substantial for households who actually experience large negative shocks at least once in life – those drawing from the first distribution in the mixture. Households facing a permanent shock at least $3\sigma$ below the mean ($\min\{\zeta_{it}\} < -0.50$; 40.31% of households experience this at least once) give up 4.8% of lifetime consumption to avoid tail risk when risk is transmitted linearly (column 1, second row). This rises to 6.39% in the presence of nonlinear transmission (column 2, second row), as bad shocks transmit more strongly in this case. For worse events, costs increase further to almost 1/8 or 11.89% of lifetime consumption among households with at least one $5\sigma$ shock below the mean (column 2, fourth row; $\min\{\zeta_{it}\} < -0.95$; 17.07% of households experience this at least once), and are higher still in scenario 3 where we evaluate the welfare cost of tail risk and the bias from a misspecified consumption function. They are also (much) higher under a counterfactual *zero risk* scenario, like that explored in Guvenen et al. (2024).

These findings recur when we account for heterogeneity in income risk, consumption transmission, or wealth. Table D.3 calculates the welfare costs allowing for age-specific risk and consumption transmission; the costs are in the same ballpark as the baseline described above. Table D.4 further calculates the welfare costs by wealth. The least wealthy have costs almost twice as large as the baseline, while the wealthier have smaller costs. This gradient is explained by the different degrees of consumption insurance between the two groups. As the likelihood that a borrowing constraint binds depends on wealth and age, the wealth- and age-specific results are indirectly reflective also of such constraints. Without a structural model, we cannot allow households to switch wealth groups endogenously (e.g., in response to an income shock), yet the baseline costs (table D.2) lie in the middle of the wealth-specific costs (table D.4), so exogenous wealth is unlikely to have caused a serious bias in our baseline calculations. The welfare estimates are also robust to small variations in $\gamma$ (table D.5).

Table D.3: Welfare costs of non-Gaussian relative to Gaussian income risk, age-specific income risk and transmission parameters

| *Specification:* | *% households*[#] | Average welfare costs $\mathbb{E}(\alpha_i)$, $\gamma = 2$ | | |
|---|---|---|---|---|
| | | $\mathcal{C}_i^*$ linear<br>$\mathcal{C}_i$ linear<br>(1) | $\mathcal{C}_i^*$ quadratic<br>$\mathcal{C}_i$ quadratic<br>(2) | $\mathcal{C}_i^*$ linear<br>$\mathcal{C}_i$ quadratic<br>(3) |
| All households | 100% | -1.02 | -1.06 | 0.31 |
| $\min\{\zeta_{it}\} < -0.50$ | 39.18% | 3.87 | 6.07 | 7.35 |
| $\min\{\zeta_{it}\} < -0.70$ | 29.23% | 5.14 | 8.29 | 9.54 |
| $\min\{\zeta_{it}\} < -0.95$ | 18.28% | 6.79 | 11.55 | 12.75 |

*Notes:* The table reports the welfare costs of non-Gaussian relative to Gaussian income risk, for $\gamma = 2$. We simulate consumption for 50,000 households using age-specific income parameters (table C.9) and consumption transmission parameters (table 4). The first row presents the average costs across all households. The next rows present the average costs over subsets of households defined by the magnitude of the worst permanent shock they experience. [#]% of households who face a shock at least as bad, drawn from the normal mixture, at least once in their lifetime.

Table D.4: Welfare costs of non-Gaussian relative to Gaussian income risk, by wealth

| *Specification:* | *% households*[#] | Average welfare costs $\mathbb{E}(\alpha_i)$, $\gamma = 2$ | | |
|---|---|---|---|---|
| | | $\mathcal{C}_i^*$ linear<br>$\mathcal{C}_i$ linear<br>(1) | $\mathcal{C}_i^*$ quadratic<br>$\mathcal{C}_i$ quadratic<br>(2) | $\mathcal{C}_i^*$ linear<br>$\mathcal{C}_i$ quadratic<br>(3) |
| **Panel A. Low wealth** | | | | |
| All households | 100% | -1.58 | -1.80 | 6.14 |
| $\min\{\zeta_{it}\} < -0.50$ | 40.96% | 6.78 | 7.43 | 14.64 |
| $\min\{\zeta_{it}\} < -0.70$ | 26.87% | 9.61 | 10.63 | 17.60 |
| $\min\{\zeta_{it}\} < -0.95$ | 12.99% | 13.12 | 14.72 | 21.36 |
| **Panel B. High wealth** | | | | |
| All households | 100% | -0.62 | -0.69 | -0.16 |
| $\min\{\zeta_{it}\} < -0.50$ | 47.18% | 3.06 | 4.59 | 5.09 |
| $\min\{\zeta_{it}\} < -0.70$ | 35.12% | 4.18 | 6.42 | 6.91 |
| $\min\{\zeta_{it}\} < -0.95$ | 21.91% | 5.73 | 9.06 | 9.54 |

*Notes:* The table reports the welfare costs of non-Gaussian relative to Gaussian income risk, for $\gamma = 2$ and low (panel A) versus high wealth (panel B). Low (high) wealth is defined on the basis of household wealth being less (more) than median real wealth over 1999–2019. We simulate consumption for 50,000 households using wealth-specific income and consumption transmission parameters, which, for reasons of brevity, are not reported in the main text. The consumption parameters can be proxied by table 6, which reports wealth-and-education-specific results, i.e., an even finer cut of the sample. The first row presents the average costs across all households in the wealth group. The next rows present the average costs over subsets of households defined by the magnitude of the worst permanent shock they experience. [#]% of households who face a shock at least as bad, drawn from the normal mixture, at least once in their lifetime.

Table D.5: Welfare costs of non-Gaussian relative to Gaussian income risk, varying $\gamma$

| *Specification:* | *% households*$^{\#}$ | Average welfare costs $\mathbb{E}(\alpha_i)$ $\mathcal{C}_i^*$ linear $\mathcal{C}_i$ linear (1) | $\mathcal{C}_i^*$ quadratic $\mathcal{C}_i$ quadratic (2) | $\mathcal{C}_i^*$ linear $\mathcal{C}_i$ quadratic (3) |
|---|---|---|---|---|
| **Panel A. $\gamma = 1$** | | | | |
| All households | 100% | -0.86 | -0.95 | 0.64 |
| $\min\{\zeta_{it}\} < -0.50$ | 40.31% | 4.93 | 6.46 | 7.94 |
| $\min\{\zeta_{it}\} < -0.70$ | 29.11% | 6.50 | 8.71 | 10.15 |
| $\min\{\zeta_{it}\} < -0.95$ | 17.07% | 8.69 | 11.90 | 13.29 |
| **Panel B. $\gamma = 3$** | | | | |
| All households | 100% | -0.76 | -0.81 | 0.66 |
| $\min\{\zeta_{it}\} < -0.50$ | 40.31% | 4.68 | 6.32 | 7.69 |
| $\min\{\zeta_{it}\} < -0.70$ | 29.11% | 6.22 | 8.59 | 9.93 |
| $\min\{\zeta_{it}\} < -0.95$ | 17.07% | 8.40 | 11.89 | 13.17 |

*Notes:* The table reports the welfare costs of non-Gaussian relative to Gaussian income risk, for $\gamma = 1$ (panel A) and $\gamma = 3$ (panel B). We simulate consumption for 50,000 households using the time-invariant consumption transmission parameters of table 3. The first row presents the average costs across all households. The next rows present the average costs over subsets of households defined by the magnitude of the worst permanent shock they experience. $^{\#}$% of households who face a shock at least as bad, drawn from the normal mixture, at least once in their lifetime.

# References


Blundell, R., L. Pistaferri, and I. Preston (2004). Imputing consumption in the PSID using food demand estimates from the CEX. Technical report, IFS Working Papers.

Blundell, R., L. Pistaferri, and I. Preston (2008). Consumption inequality and partial insurance. *American Economic Review 98*(5), 1887–1921.

Blundell, R., L. Pistaferri, and I. Saporta-Eksten (2016). Consumption inequality and family labor supply. *American Economic Review 106*(2), 387–435.

Bound, J., C. Brown, G. J. Duncan, and W. L. Rodgers (1994). Evidence on the validity of cross-sectional and longitudinal labor market data. *Journal of Labor Economics 12*(3), 345–368.

Chatterjee, A., J. Morley, and A. Singh (2021). Estimating household consumption insurance. *Journal of Applied Econometrics 36*(5), 628–635.

Guvenen, F., F. Karahan, S. Ozkan, and J. Song (2021). What do data on millions of U.S. workers reveal about lifecycle earnings dynamics? *Econometrica 89*(5), 2303–2339.

Guvenen, F., S. Ozkan, and R. Madera (2024). Consumption dynamics and welfare under non-Gaussian earnings risk. NBER WP 32298.

Meyer, B. D. and J. X. Sullivan (2023). Consumption and income inequality in the United States since the 1960s. *Journal of Political Economy 131*(2), 247–284.

Wu, C. and D. Krueger (2021). Consumption insurance against wage risk: Family labor supply and optimal progressive income taxation. *American Economic Journal: Macroeconomics 13*(1), 79–113.