Online Appendix: Cyclical labour income risk in Great Britain

Konstantinos Angelopoulos
University of Glasgow and CESifo

Spyridon Lazarakis
Lancaster University

James Malley
University of Glasgow and CESifo

May 20, 2021

Appendix A: Data

A.1 BHPS

The main dataset used in this paper is the British Household Panel Survey (BHPS), Institute for Social and Economic Research (2018). The BHPS is a comprehensive longitudinal study for the UK running from 1991 to 2008. As a panel data survey, the BHPS tracks individuals across households over time. In the first wave, the BHPS achieved a sample size of around 5,000 households (10,000 adult interviews) or a 65% response rate. After the first wave, due to sample attrition, the sample size shrank slightly. For example, in 2000 it achieved around 4,200 complete interviews or a 75% response rate (see Taylor *et al.* 2010).

Since the start of BHPS in 1991, several additional sub-samples have been added to the survey. For example, the European Community Household Panel Survey (ECHP) sub-sample started in 1997. This survey was added, mainly to include respondents from Northern Ireland and a low-income sample from the UK. Moreover, in 1999 two more additional boost samples, for Wales and Scotland, have been added. Since the focus is on GB, to maintain the longest possible time series dimension in our analysis, we only use the data starting in 1991, i.e. the original panel dataset. Finally, following Blundell and Etheridge (2010), we also make use of an auxiliary dataset called "Derived Current and Annual Net Household Income Variables" compiled by Bardasi *et al.* (2012).

The BHPS contains detailed information on key magnitudes of interest for this paper. In particular, earnings, hours worked and other income. Compared to other UK panel datasets for earnings, e.g. the New Earnings Survey (NES) for the period 1975-2002 and the Annual Survey of Hours and Earnings (ASHE), for the period 1997-2019, BHPS is much smaller in the cross-sectional dimension. The advantages of NES and ASHE are the accuracy and the sample size, which covers 1% of the total working population. However, these datasets do not provide information relating to (i) household physical and human capital, e.g. educational attainment; (ii) why individuals disappear from the survey, e.g. due to an injury, unemployment spell or move to self-employment; (iii) self-employed individuals, which are a considerable percentage of the working population (approximately 14%); (iv) individual *annual* earnings which are only available from 1999 onwards; and (v) household annual incomes (labour and net labour income).

In contrast to the NES and ASHE, the BHPS has information on both individual and household characteristics. Therefore, it allows the examination of compositional effects (i.e. differences between individuals and households) and thus issues relating to household insurance mechanisms. Furthermore, if the BHPS dataset is combined with the DCANHIV dataset, it allows the evaluation of social insurance mechanisms (as we do in this paper). Moreover, BHPS provides important human capital variables such as educational attainment. Furthermore, BHPS also covers the self-employed, the unemployed or even those who do not participate in the labour market for any reason. Finally, it provides a consistent measure of individual and household annual earnings/incomes over the whole period at hand.

The Wealth and Assets Survey (WAS) is one additional panel dataset for GB containing measures of annual earnings. However, it has some shortcomings for the methodology we follow in this paper. First, the WAS runs only after 2006 and collects data every two years. Hence, there are very few observations (6 waves or rounds in total). Second, only labour earnings are available for all 6 rounds whereas measures of net labour income are only available from round III (i.e. 2010). Moreover, from round VI there has been a change in the timing of the data collection resulting in some observations from round V being repeated in round VI.

A.2 Demographic and socioeconomic variables

1. **Head and relationship to head**: For each individual in the sample, BHPS reports the relationship to the head of household in any given wave. In our analysis, we focus on households whose head is married. Following Blundell and Etheridge (2010), the head of the household is defined as the oldest married (or living in partnership) male within the household. If there is no married couple within the household, the oldest working male is the head. If there is no working male, then the working female becomes the head.

2. **Education level**: BHPS includes information on educational attainment. For the BHPS we have used the variable wQFEDHI (where the prefix w denotes wave). To examine the potential heterogeneity of earnings risk in the main text, the sample is split into degree holders and non-degree holders. The former are the individuals who hold either a Higher Degree or 1st Degree, while the latter are the individuals who hold either Higher National Certificate/Diploma or teaching qualifications or A-levels/AS level/Highers or GCSE/O level/other qualification or they have no qualifications.

A.3 Income and hours variables

1. **Labour income:** is obtained from the Derived Current and Annual Net Household Income Variables dataset (Bardasi *et al.* 2012). Labour income is equal to total household annual labour income, wHHYRLG, plus annual private transfers income. Imputed values can be included in labour income only if they do not correspond to the head of the household earnings. Private transfers income totals all receipts from other transfers (including education grants, sickness insurance, maintenance, foster allowance and payments from TU/Friendly societies, from absent family members).

2. **Labour income − taxes − NI:** is obtained from Bardasi *et al.* (2012) and is equal to "Labour income" minus annual national insurance contributions, wYRNI, minus annual income tax after credits, wYRTAXNT, plus annual private transfers income, wHHYRT.

3. **Labour income + benefits:** is obtained from Bardasi *et al.* (2012) and is defined as "Labour income" plus annual social benefits income, wHHYRB. Social benefits income totals all receipts from state benefits including national insurance retirement pensions.

4. **Labour income + benefits − taxes − NI:** is obtained from Bardasi *et al.* (2012) and is defined as "Labour income − taxes − NI" plus annual social benefits income, wHHYRB.

A.4 Sample selection

For all of the measures discussed below, we use the original BHPS sample excluding the observations from the boost samples after 1997. We keep all the households in the BHPS where their heads are between 23-62 years old with no imputed earnings. Further, we discard the household where their head's earnings are less than half of the product between the minimum legal hourly wage times 520 hours. Also, the head must not be in the military and must not have missing values for earnings, region and educational attainment. For the remaining households, we only keep households who are in the sample for at least three consecutive periods.

Table A.1: Sample selection in steps

selection step	households (obs.)
1. Whole sample	130,974
2. Drop proxy & non-full interviews	128,348
3. Original sample	82,355
4. Full interview of all members in household	74,605
5. Drop if head in military	74,496
6. Drop if head's region missing	74,454
7. Drop if head's marital status is missing	74,448
8. Keep if head's earnings>threshold	39,170
9. Keep if heads' age $\geq$23, $\leq$62	36,259
10. Drop if no head's educational info	35,924
11. Keep if present at least 3 consecutive observations	30,117
ave. N obs per wave	1,673
N of unique households	3,475
ave. obs per household	8.7

Table A.2: Summary of Selected BHPS Data (1991-2008)

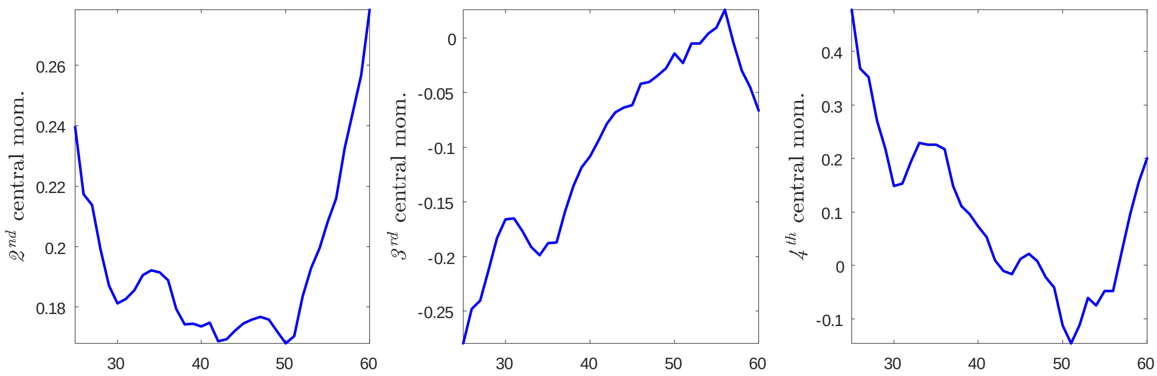
Variable	mean	s.d.	min	max
Head's age	41.03	9.87	23	62
Head's sex (male=0,female=1)	0.12	0.33	0	1
HH size	2.82	1.28	1	9
Marital Status (married/cohab.=1)	0.75	0.43	0	1
Head's earnings	30,644.8	21,997.7	997.8	714,581.7
Labour income	43,430.8	28,138.1	997.8	714,581.7
Labour income −taxes −NI	33,563.4	19,562.8	997.8	463,066.5
Labour income +benefits	45,255.7	27,687.9	1,275.7	714,581.7
Labour income +benefits −taxes −NI	35,388.3	19,154.9	1,275.7	464,572.4

Note: All monetary values are expressed in 2012 prices using the RPI deflator. The summary statistics refer to sample selection step 11 in Table A.1.

A.5 Supplement to Section 2.2

Figure A.1 complements Figure 1 by presenting additional plots of age-level inequality, partialling out cohort effects. We plot the age coefficients against age (see also Figure 1a in Storesletten *et al.* (2004) for the US and Blundell *et al.* (2015) for Norway, for similar life-cycle variance plots). In particular, the plots show that age effects on the variance of labour income and fourth central moment are U-shaped, while they have an inverted U-shape regarding the third central moment. The U-shaped life cycle pattern of labour income variance is broadly similar to that in the U.S., based on 1960 Census data (see Ch 6 in Mincer (1974)) and in Norway, see Blundell *et al.* (2015). However, they are dissimilar to the results until 1993 for the US using the PSID, reported in Storesletten *et al.* (2004). In this study, the variance starts to increase quickly after the first few working years. Compared with Storesletten *et al.* (2004), the pattern here provides an explanation for why our estimates of the persistence parameter in labour income (see Table 1 in the main text) are lower than their estimate of 0.99. Note also that as Blundell *et al.* (2015) point out, a non-monotonic pattern in life-cycle inequality is not inconsistent with an age-invariant persistence parameter in the labour income dynamics model.

Figure A.1: Age effects on the distribution of residual household labour income



Note: Figure A.1 plots the coefficients of age-specific dummies in a regression of central moments of residual household labour income on cohort and age dummies since the age of 25.

Appendix B: Parameter estimation

B.1 Theoretical moments

The theoretical moments of $\upsilon_{i,h,t}$ are:

$$E\left[\upsilon_{i,h,t}\right] = 0, \tag{1}$$

$$E\left[\upsilon_{i,h,t}^2\right] = m_2^{\chi} + b_2^{\chi}\gamma_g + m_2^{\varepsilon,f(t)} + E\left[z_{i,h,t}^2\right], \tag{2}$$

$$E\left[\upsilon_{i,h,t}^3\right] = m_3^{\chi} + b_3^{\chi}\gamma_g + m_3^{\varepsilon,f(t)} + E\left[z_{i,h,t}^3\right], \tag{3}$$

$$E\left(\upsilon_{i,h,t}^4\right) = \left\{\begin{array}{c} m_4^{\chi} + b_4^{\chi}\gamma_g + m_4^{\varepsilon,f(t)} + 6\left(m_2^{\chi} + b_2^{\chi}\gamma_g\right) m_2^{\varepsilon,f(t)} \\ +E\left[z_{i,h,t}^4\right] + 6\left(m_2^{\chi} + b_2^{\chi}\gamma_g + m_2^{\varepsilon,f(t)}\right) E\left[z_{i,h,t}^2\right] \end{array}\right\}, \tag{4}$$

$$\begin{aligned} Cov(\upsilon_{i,h,t}, \upsilon_{i,h+\kappa,t+\kappa}) &= E\left[\upsilon_{i,h,t}\upsilon_{i,h+\kappa,t+\kappa}\right] \\ &= m_2^{\chi} + b_2^{\chi}\gamma_g + E\left[z_{i,h,t}z_{i,h+\kappa,t+\kappa}\right], \end{aligned} \tag{5}$$

$$\begin{aligned} CoSk(\upsilon_{i,h,t}, \upsilon_{i,h+\kappa,t+\kappa}) &= E\left[\upsilon_{i,h,t}^2\upsilon_{i,h+\kappa,t+\kappa}\right] \\ &= m_3^{\chi} + b_3^{\chi}\gamma_g + E\left[z_{i,h,t}^2 z_{i,h+\kappa,t+\kappa}\right], \end{aligned} \tag{6}$$

$$\begin{aligned} CoKurt(\upsilon_{i,h,t}, \upsilon_{i,h+\kappa,t+\kappa}) &= E\left[\upsilon_{i,h,t}^2\upsilon_{i,h+\kappa,t+\kappa}^2\right] \\ &= \left\{\begin{array}{c} m_4^{\chi} + b_4^{\chi}\gamma_g + \left(m_2^{\varepsilon,f(t)} + m_2^{\varepsilon,f(t+\kappa)}\right)\left(m_2^{\chi} + b_2^{\chi}\gamma_g\right) \\ +m_2^{\varepsilon,f(t)} m_2^{\varepsilon,f(t+\kappa)} + E\left[z_{i,h,t}^2 z_{i,h,t+\kappa}^2\right] \\ +\left(m_2^{\chi} + b_2^{\chi}\gamma_g\right)\left(E\left[z_{i,h,t}^2\right] + E\left[z_{i,h+\kappa,t+\kappa}^2\right] + 4E\left[z_{i,h,t}z_{i,h+\kappa,t+\kappa}\right]\right) \\ +m_2^{\varepsilon,f(t+\kappa)} E\left[z_{i,h,t}^2\right] + m_2^{\varepsilon,f(t)} E\left[z_{i,h+\kappa,t+\kappa}^2\right] \end{array}\right\}, \end{aligned} \tag{7}$$

where

$$E\left[z_{i,h,t}^2\right] = \sum_{j=0}^{h-1}\rho^{2j} m_2^{\eta,f(t-j)},\ E\left[z_{i,h,t}z_{i,h+\kappa,t+\kappa}\right] = \rho^{\kappa} E\left[z_{i,h,t}^2\right], \tag{8}$$

$$E\left[z_{i,h,t}^3\right] = \sum_{j=0}^{h-1}\rho^{3j} m_3^{\eta,f(t-j)},\ E\left[z_{i,h,t}^2 z_{i,h+\kappa,t+\kappa}\right] = \rho^{\kappa} E\left[z_{i,h,t}^3\right], \tag{9}$$

$$E\left[z_{i,h,t}^4\right] = \left[\begin{array}{c} m_4^{\eta,f(t)} + \rho^4 E\left[z_{i,h-1,t-1}^4\right] \\ +\mathbf{1}_{[h>1]}\left[6\sum_{j=1}^{h-1}\rho^{2j} m_2^{\eta,f(t-j)} m_2^{\eta,f(t)}\right] \end{array}\right], \tag{10}$$

$$E\left[z_{i,h,t}^2 z_{i,h+\kappa,t+\kappa}^2\right] = \left[\begin{array}{c} \rho^2 E\left[z_{i,h,t}^2 z_{i,h+\kappa-1,t+\kappa-1}^2\right] \\ +\sum_{j=0}^{h-1}\rho^{2j} m_2^{\eta,f(t-j)} m_2^{\eta,f(t+\kappa)} \end{array}\right],\ \kappa > 0. \tag{11}$$

The moment conditions employed in the estimation are:

$$E\left[\widehat{\upsilon}_{i,h,t}^{2} - E\left[\upsilon_{i,h,t}^{2} \mid \boldsymbol{\theta}\right]\right] = 0, \tag{12}$$

$$E\left[\widehat{\upsilon}_{i,h,t}^{3} - E\left[\upsilon_{i,h,t}^{3} \mid \boldsymbol{\theta}\right]\right] = 0, \tag{13}$$

$$E\left[\widehat{\upsilon}_{i,h,t}^{4} - E\left[\upsilon_{i,h,t}^{4} \mid \boldsymbol{\theta}\right]\right] = 0, \tag{14}$$

$$E\left[\widehat{\upsilon}_{i,h,t}\widehat{\upsilon}_{i,h+\kappa,t+\kappa} - E\left[\upsilon_{i,h,t}\upsilon_{i,h+\kappa,t+\kappa} \mid \boldsymbol{\theta}\right]\right] = 0, \tag{15}$$

$$E\left[\widehat{\upsilon}_{i,h,t}^{2}\widehat{\upsilon}_{i,h+\kappa,t+\kappa} - E\left[\upsilon_{i,h,t}^{2}\upsilon_{i,h+\kappa,t+\kappa} \mid \boldsymbol{\theta}\right]\right] = 0, \tag{16}$$

$$E\left[\widehat{\upsilon}_{i,h,t}^{2}\widehat{\upsilon}_{i,h+\kappa,t+\kappa}^{2} - E\left[\upsilon_{i,h,t}^{2}\upsilon_{i,h+\kappa,t+\kappa}^{2} \mid \boldsymbol{\theta}\right]\right] = 0, \tag{17}$$

where $\boldsymbol{\theta}$ is the vector of parameters to be estimated:

$$\boldsymbol{\theta} = \left\{ \begin{array}{c} \rho, m_2^{\chi}, m_3^{\chi}, m_4^{\chi}, b_2^{\chi}, b_3^{\chi}, b_4^{\chi}, \\ m_2^{\varepsilon,c}, m_2^{\varepsilon,e}, m_2^{\eta,c}, m_2^{\eta,e}, \\ m_3^{\varepsilon,c}, m_3^{\varepsilon,e}, m_3^{\eta,c}, m_3^{\eta,e}, \\ m_4^{\varepsilon,c}, m_4^{\varepsilon,e}, m_4^{\eta,c}, m_4^{\eta,e} \end{array} \right\}.$$

We also need to impose more restrictions on the parameters of interest with respect to the fourth moments. First, we restrict all variances and kurtosis to be positive by applying exponential transformations. Second, for each density function we need to impose the following relation between the third and the fourth moments:

$$\frac{m_4^{k,f(t)}}{\left(m_2^{k,f(t)}\right)^2} \geq \left[\frac{m_3^{k,f(t)}}{\left(m_2^{k,f(t)}\right)^{3/2}}\right]^2 + 1, \tag{18}$$

for $k = \eta, \varepsilon, \chi$ so that the fourth moments exist (see e.g. Sen (2012)).

B.2 Estimation

The system given by (12)-(17) is over-identified. We estimate the parameters in two steps (see also Blundell *et al.* (2015) and Busch and Ludwig (2020) for other applications of estimation of such systems in steps). First, we estimate: (i) the persistence parameter ρ; (ii) the 2^{nd} and 3^{rd} moments of the distribution of the initial conditions, m_2^{χ} and m_3^{χ} as well as their trends, b_2^{χ} and b_3^{χ}; (iii) the time-dependent 2^{nd} and 3^{rd} moments of the distribution of the transitory shocks, $m_2^{\varepsilon,f(t)}$ and $m_3^{\varepsilon,f(t)}$; and (iv) the time-dependent higher moments for innovations to the persistent component i.e. $m_2^{\eta,f(t)}$ and $m_3^{\eta,f(t)}$. These parameters can be identified using the empirical moments in conjunction with the theoretical conditions in (1), (2), (3), (5) and (6). Given these estimates, we estimate: (v) the 4^{th} moments of the distribution of the initial conditions, m_4^{χ} as well as their trends, b_4^{χ}; (vi) the 4^{th} moments of the distribution of the transitory shocks, $m_4^{\varepsilon,f(t)}$; and (vii) the 4^{th} moments of the distribution of the transitory shocks, $m_4^{\eta,f(t)}$. The second set of parameters are identified by the set of equations (4) and (7). The advantage of estimating the parameters in steps is that it makes the system numerically more stable. Note that ρ and variances enter in a highly nonlinear manner in the fourth moment conditions. We denote the two subsets of the parameter vector θ as $\boldsymbol{\theta}_1$ and $\boldsymbol{\theta}_2$.

Let $\mathbf{m}$ be the vector with all the available empirical moments constructed as above and $\mathbf{G}(\boldsymbol{\theta})$ the vector of the respective theoretical moments. We partition vector $\mathbf{m}$ and $\mathbf{G}(\boldsymbol{\theta})$ into two parts, one containing the second and the third moments, and the other with the fourth moments. The goal is to estimate a model for $\mathbf{m}$ as:

$$\begin{bmatrix}\mathbf{m}_2\\ \mathbf{m}_3\end{bmatrix} = \mathbf{G}_1(\boldsymbol{\theta}_1) + \boldsymbol{\Upsilon}_1, \tag{19}$$

$$\mathbf{m}_4 = \mathbf{G}_2(\boldsymbol{\theta}_2 \mid \boldsymbol{\theta}_1) + \boldsymbol{\Upsilon}_2, \tag{20}$$

where $\boldsymbol{\Upsilon}_1$ and $\boldsymbol{\Upsilon}_2$ capture sampling variability. For the estimation, we minimize the distance between the empirical and the theoretical moments. Formally, we numerically minimize sequentially the following two objective functions:

$$Q_1(\boldsymbol{\theta}_1) = \min_{\theta_1} \left(\begin{bmatrix}\mathbf{m}_2\\ \mathbf{m}_3\end{bmatrix} - \mathbf{G}_1(\boldsymbol{\theta}_1) \right)' \mathcal{W}_1 \left(\begin{bmatrix}\mathbf{m}_2\\ \mathbf{m}_3\end{bmatrix} - \mathbf{G}_1(\boldsymbol{\theta}_1) \right), \tag{21}$$

$$Q_2(\boldsymbol{\theta}_2 \mid \boldsymbol{\theta}_1) = \min_{\theta_2} \left(\mathbf{m}_4 - \mathbf{G}_2(\boldsymbol{\theta}_2 \mid \boldsymbol{\theta}_1)\right)' \mathcal{W}_2 \left(\mathbf{m}_4 - \mathbf{G}_2(\boldsymbol{\theta}_2 \mid \boldsymbol{\theta}_1)\right), \tag{22}$$

where $\mathcal{W}_1$ and $\mathcal{W}_2$ are weighting matrices. Following Altonji and Segal (1996), the typical choice of a weighting matrix in the literature is the identity matrix. However, notice that each moment is calculated by a different number of observations. Moreover, since we are calculating higher moments, it is well known that bigger samples give more accurate results. Hence, we weight each moment equation by the number of observations used to calculate its empirical part since the panel is unbalanced, i.e. the diagonal elements of $\mathcal{W}_1$ and $\mathcal{W}_2$ are not ones but the relative weights.[1]

To compute the confidence intervals, we follow MaCurdy (2007) and use the block bootstrap procedure for 3,000 replications. For each replication, we sequentially estimate the parameters of the two stages so that the confidence intervals for the second step estimates are correct. The resulting confidence intervals account for serial correlation of arbitrary form, heteroskedasticity and that we use pre-estimated residuals.[2] The bootstrap confidence intervals are calculated as $\mathrm{CI}_{1-\alpha} = [2 \times \widehat{\boldsymbol{\theta}} - (\boldsymbol{\theta}^*)_{1-\alpha/2}, 2 \times \widehat{\boldsymbol{\theta}} - (\boldsymbol{\theta}^*)_{\alpha/2}]$, where α is the significance level, and $(\boldsymbol{\theta}^*)_{\alpha/2}$ and $(\boldsymbol{\theta}^*)_{1-\alpha/2}$ denote the $(\alpha/2)^{th}$ and $(1-\alpha/2)^{th}$ percentiles of the bootstrap distribution of $\boldsymbol{\theta}^*$'s respectively (see Davidson and Hinkley (1997)).

B.3 Variation in the aggregate state

The set of theoretical moments for $\upsilon_{i,h,t}$ in (1)-(7) are a function of past moments of innovations to the persistent component, via (8)-(11). Therefore, the estimation of the parameters of interest requires knowledge of the specific history of the aggregate state for $h-1$ years before those in the observed sample of the households, which here is for 1956, as this will determine which moments enter, in which period, in the summation operators in (8)-(11). Thus, we use the variation in the

[1] For similar treatment see Heathcote *et al.* (2010), Domeij and Floden (2010) and Busch and Ludwig (2020).

[2] See also Hall and Horowitz (1996) and Horowitz (2003).

aggregate state since 1956 in the estimation. Naturally, a higher variation in the aggregate state increases the estimation accuracy of the relevant moments separately for periods of expansion and contraction.

B.4 Identification

In this section, we show that under sufficient variation in the history of the aggregate state, if we observe four time periods and four age groups in the panel dataset used in the estimation, the parameters in the theoretical model are identified. In turn, this implies that the parameters are identified in our estimation since we have 18 time periods and 36 age groups.

To keep the exposition tractable, we assume that $t \in T = \{1, 2, 3, 4\}$ and $h \in H = \{1, 2, 3, 4\}$. This restriction implies a total of 90 empirical moments to identify 19 parameters, ρ, m_2^{χ}, b_2^{χ}, m_3^{χ}, b_3^{χ}, m_4^{χ}, b_4^{χ}, $m_2^{\varepsilon,c}$, $m_3^{\varepsilon,c}$, $m_4^{\varepsilon,c}$, $m_2^{\varepsilon,e}$, $m_3^{\varepsilon,e}$, $m_4^{\varepsilon,e}$, $m_2^{\eta,c}$, $m_2^{\eta,e}$, $m_3^{\eta,c}$, $m_3^{\eta,e}$, $m_4^{\eta,c}$, and $m_4^{\eta,e}$. To illustrate identification using a specific history of the aggregate state, assume the following sequence of time periods classified into expansions and contractions:

$$\begin{bmatrix} \text{-2:} & \text{expansion} \\ \text{-1:} & \text{contraction} \\ \text{0:} & \text{expansion} \\ \text{1:} & \text{contraction} \\ \text{2:} & \text{expansion} \\ \text{3:} & \text{contraction} \\ \text{4:} & \text{expansion} \end{bmatrix},$$

which implies:

$$\begin{bmatrix} m_2^{\eta}(f(-2)) = m_2^{\eta,e} & m_3^{\eta}(f(-2)) = m_3^{\eta,e} & m_4^{\eta}(f(-2)) = m_4^{\eta,e} \\ m_2^{\eta}(f(-1)) = m_2^{\eta,c} & m_3^{\eta}(f(-1)) = m_3^{\eta,c} & m_4^{\eta}(f(-1)) = m_4^{\eta,c} \\ m_2^{\eta}(f(0)) = m_2^{\eta,e} & m_3^{\eta}(f(0)) = m_3^{\eta,e} & m_4^{\eta}(f(0)) = m_4^{\eta,e} \\ m_2^{\eta}(f(1)) = m_2^{\eta,c} & m_3^{\eta}(f(1)) = m_3^{\eta,c} & m_4^{\eta}(f(1)) = m_4^{\eta,c} \\ m_2^{\eta}(f(2)) = m_2^{\eta,e} & m_3^{\eta}(f(2)) = m_3^{\eta,e} & m_4^{\eta}(f(2)) = m_4^{\eta,e} \\ m_2^{\eta}(f(3)) = m_2^{\eta,c} & m_3^{\eta}(f(3)) = m_3^{\eta,c} & m_4^{\eta}(f(3)) = m_4^{\eta,c} \\ m_2^{\eta}(f(4)) = m_2^{\eta,e} & m_3^{\eta}(f(4)) = m_3^{\eta,e} & m_4^{\eta}(f(4)) = m_4^{\eta,e} \end{bmatrix}$$

Note that the periods 0,-1 and -2, i.e. past periods appear in the table above. The reason is that since an agent's income has a persistent component, then she is accumulating shocks. In turn, this means that some of the agents in the sample "bring" with them these past shocks, and thus, the central moments of these past shocks appear in our theoretical moments. Consequently, we have extra information which we exploit to get more accurate estimates for $m_2^{\eta,c}$, $m_2^{\eta,c}$, $m_3^{\eta,c}$, $m_3^{\eta,e}$, $m_4^{\eta,c}$, and $m_4^{\eta,e}$. Moreover, we have seven cohorts, and hence we denote the cohort effects as $m_{\tau}^{\chi}(\gamma_g) = m_{\tau}^{\chi} + \gamma_g b_{\tau}^{\chi}$, where $\tau = 2, 3, 4$ and $\gamma_g = 1, 2, ..., 7$.

B.4.1 Persistence

The persistence parameter ρ is identified by employing (5) for period $t = 1$, ages $h = 1$ and $\kappa = 1, 2, 3$:

$$\rho = \left[\frac{Cov\left(\upsilon_{i,1,1}, \upsilon_{i,4,4}\right) - Cov\left(\upsilon_{i,1,1}, \upsilon_{i,3,3}\right)}{Cov\left(\upsilon_{i,1,1}, \upsilon_{i,3,3}\right) - Cov\left(\upsilon_{i,1,1}, \upsilon_{i,2,2}\right)}\right]. \tag{23}$$

B.4.2 Second and third moments for transitory shocks

Using (2) and (5) for periods $t = 1, 2$, age $h = 1$ and $\kappa = 1$, $m_2^{\varepsilon,c}$ and $m_2^{\varepsilon,e}$ are identified by:

$$m_2^{\varepsilon,c} = E\left(\upsilon_{i,1,1}^2\right) - \rho Cov\left(\upsilon_{i,1,1}, \upsilon_{i,2,2}\right), \tag{24}$$

$$m_2^{\varepsilon,e} = E\left(\upsilon_{i,1,2}^2\right) - \rho Cov\left(\upsilon_{i,1,2}, \upsilon_{i,2,3}\right), \tag{25}$$

and likewise employing equations (3) and (6), $m_3^{\varepsilon,c}$ and $m_3^{\varepsilon,e}$ are identified by :

$$m_3^{\varepsilon,c} = E\left(\upsilon_{i,1,1}^3\right) - \rho CoSk\left(\upsilon_{i,1,1}, \upsilon_{i,2,2}\right), \tag{26}$$

$$m_3^{\varepsilon,e} = E\left(\upsilon_{i,1,2}^3\right) - \rho CoSk\left(\upsilon_{i,1,2}, \upsilon_{i,2,3}\right). \tag{27}$$

B.4.3 Second and third moments for fixed effects

Using (5) for periods $t = 1, 2$, ages $h = 1$ and $\kappa = 1, 2$, $m_2^{\chi}(4)$ and $m_2^{\chi}(5)$ are identified by:

$$m_2^{\chi}(4) = \frac{\rho Cov\left(\upsilon_{i,1,1}, \upsilon_{i,2,2}\right) - Cov\left(\upsilon_{i,1,1}, \upsilon_{i,3,3}\right)}{\rho - 1}, \tag{28}$$

$$m_2^{\chi}(5) = \frac{\rho Cov\left(\upsilon_{i,1,2}, \upsilon_{i,2,3}\right) - Cov\left(\upsilon_{i,1,2}, \upsilon_{i,3,4}\right)}{\rho - 1}, \tag{29}$$

and likewise using (6), $m_3^{\chi}(4)$ and $m_3^{\chi}(5)$ are given by:

$$m_3^{\chi}(4) = \frac{\rho CoSk\left(\upsilon_{i,1,1}, \upsilon_{i,2,2}\right) - CoSk\left(\upsilon_{i,1,1}, \upsilon_{i,3,3}\right)}{\rho - 1}, \tag{30}$$

$$m_3^{\chi}(5) = \frac{\rho CoSk\left(\upsilon_{i,1,2}, \upsilon_{i,2,3}\right) - CoSk\left(\upsilon_{i,1,2}, \upsilon_{i,3,4}\right)}{\rho - 1}. \tag{31}$$

Therefore, $m_2^{\chi}(4)$ and $m_2^{\chi}(5)$ will identify m_2^{χ} and b_2^{χ} while $m_3^{\chi}(4)$ and $m_3^{\chi}(5)$ will identify m_3^{χ} and b_3^{χ}.

B.4.4 Moments for innovations to the persistent component

Using (5) for periods $t = 1, 2$, ages $h = 1$ and $\kappa = 1, 2$, $m_2^{\eta,c}$ and $m_2^{\eta,e}$ are identified by:

$$m_2^{\eta,c} = \frac{Cov\left(\upsilon_{i,1,1}, \upsilon_{i,3,3}\right) - Cov\left(\upsilon_{i,1,1}, \upsilon_{i,2,2}\right)}{\rho\left(\rho - 1\right)}, \tag{32}$$

$$m_2^{\eta,e} = \frac{Cov\left(\upsilon_{i,1,2}, \upsilon_{i,3,4}\right) - Cov\left(\upsilon_{i,1,2}, \upsilon_{i,2,3}\right)}{\rho\left(\rho - 1\right)}. \tag{33}$$

Likewise, using equation (6) for the same t, h and κ, $m_3^{\eta,c}$ and $m_3^{\eta,e}$ are given by:

$$m_3^{\eta,c} = \frac{CoSk\left(\upsilon_{i,1,1}, \upsilon_{i,3,3}\right) - CoSk\left(\upsilon_{i,1,1}, \upsilon_{i,2,2}\right)}{\rho\left(\rho-1\right)}, \tag{34}$$

$$m_3^{\eta,e} = \frac{CoSk\left(\upsilon_{i,1,2}, \upsilon_{i,3,4}\right) - CoSk\left(\upsilon_{i,1,2}, \upsilon_{i,2,3}\right)}{\rho\left(\rho-1\right)}. \tag{35}$$

B.4.5 Fourth moments

For the fourth moments we can use (4) for periods $t = 1, 2$, ages $h = 1$ and (7) for periods $t = 1, 2$, ages $h = 1$ and $\kappa = 1, 2$ to get:

$$\begin{aligned} E\left(\upsilon_{i,1,1}^4\right) &= \left\{ \begin{array}{c} m_4^{\chi}(4) + m_4^{\varepsilon,c} + 6m_2^{\chi}(4)m_2^{\varepsilon,c} \\ +m_4^{\eta,c} + 6\left(m_2^{\chi}(4) + m_2^{\varepsilon,c}\right)m_2^{\eta,c} \end{array} \right\}, \\ &= m_4^{\chi}(4) + m_4^{\varepsilon,c} + m_4^{\eta,c} + \omega_1, \end{aligned} \tag{36}$$

$$\begin{aligned} E\left(\upsilon_{i,1,2}^4\right) &= \left\{ \begin{array}{c} m_4^{\chi}(5) + m_4^{\varepsilon,e} + 6m_2^{\chi}(5)m_2^{\varepsilon,c} \\ +m_4^{\eta,e} + 6\left(m_2^{\chi}(5) + m_2^{\varepsilon,e}\right)m_2^{\eta,e} \end{array} \right\}, \\ &= m_4^{\chi}(5) + m_4^{\varepsilon,e} + m_4^{\eta,e} + \omega_2, \end{aligned} \tag{37}$$

$$\begin{aligned} E(\upsilon_{i,1,1}^2, \upsilon_{i,2,2}^2) &= \left\{ \begin{array}{c} m_4^{\chi}(4) + m_2^{\chi}(4)\left(m_2^{\varepsilon,c} + m_2^{\varepsilon,e}\right) + m_2^{\varepsilon,c}m_2^{\varepsilon,e} \\ +\rho^2 m_4^{\eta,c} + m_2^{\eta,c}m_2^{\eta,e} \\ +m_2^{\chi}(4)\left\{m_2^{\eta,c} + \rho^2 m_2^{\eta,c} + m_2^{\eta,e} + 4\rho m_2^{\eta,c}\right\} \\ m_2^{\varepsilon,e}m_2^{\eta,c} + m_2^{\varepsilon,c}\left(\rho^2 m_2^{\eta,c} + m_2^{\eta,e}\right) \end{array} \right\}, \\ &= m_4^{\chi}(4) + \rho^2 m_4^{\eta,c} + \omega_3 \end{aligned} \tag{38}$$

$$\begin{aligned} E(\upsilon_{i,1,2}^2, \upsilon_{i,2,3}^2) &= \left\{ \begin{array}{c} m_4^{\chi}(5) + m_2^{\chi}(5)\left(m_2^{\varepsilon,c} + m_2^{\varepsilon,e}\right) + m_2^{\varepsilon,c}m_2^{\varepsilon,e} \\ +\rho^2 m_4^{\eta}\left(e\right) + m_2^{\eta,c}m_2^{\eta,e} \\ +m_2^{\chi}(5)\left\{m_2^{\eta,e} + \rho^2 m_2^{\eta,e} + m_2^{\eta,c} + 4\rho m_2^{\eta,e}\right\} \\ m_2^{\varepsilon,c}m_2^{\eta,e} + m_2^{\varepsilon,e}\left(\rho^2 m_2^{\eta,e} + m_2^{\eta,c}\right) \end{array} \right\}, \\ &= m_4^{\chi}(5) + \rho^2 m_4^{\eta}\left(e\right) + \omega_4, \end{aligned} \tag{39}$$

$$E(\upsilon_{i,1,1}^2,\upsilon_{i,3,3}^2)=\left\{\begin{array}{c} m_4^\chi(4)+m_2^\chi(4)\left(m_2^{\varepsilon,c}+m_2^{\varepsilon,c}\right)+m_2^{\varepsilon,c}m_2^{\varepsilon,c} \\ +\rho^4 m_4^{\eta,c}+\rho^2 m_2^{\eta,c}m_2^{\eta,e}+m_2^{\eta,c}m_2^{\eta,c} \\ +m_2^\chi(4)\left\{m_2^{\eta,c}+\rho^2 m_2^{\eta,e}+\rho^4 m_2^{\eta,c}+4\rho^2 m_2^{\eta,c}\right\} \\ m_2^{\varepsilon,c}m_2^{\eta,c}+m_2^{\varepsilon,c}\left(m_2^{\eta,c}+\rho^2 m_2^{\eta,e}+\rho^4 m_2^{\eta,c}\right) \end{array}\right\},$$

$$= m_4^\chi(4)+\rho^4 m_4^{\eta,c}+\omega_5, \tag{40}$$

$$E(\upsilon_{i,1,2}^2,\upsilon_{i,3,4}^2)=\left\{\begin{array}{c} m_4^\chi(5)+m_2^\chi(5)\left(m_2^{\varepsilon,e}+m_2^{\varepsilon,e}\right)+m_2^{\varepsilon,e}m_2^{\varepsilon,e} \\ +\rho^4 m_4^{\eta,e}+\rho^2 m_2^{\eta,c}m_2^{\eta,e}+m_2^{\eta,e}m_2^{\eta,e} \\ +m_2^\chi(5)\left\{m_2^{\eta,e}+\rho^2 m_2^{\eta,c}+\rho^4 m_2^{\eta,e}+4\rho^2 m_2^{\eta,e}\right\} \\ m_2^{\varepsilon,e}m_2^{\eta,e}+m_2^{\varepsilon,e}\left(m_2^{\eta,e}+\rho^2 m_2^{\eta,c}+\rho^4 m_2^{\eta,e}\right) \end{array}\right\},$$

$$= m_4^\chi(5)+\rho^4 m_4^{\eta,e}+\omega_6, \tag{41}$$

Thus we have to solve the following linear system:

$$\begin{aligned}
E\left(\upsilon_{i,1,1}^4\right) &= m_4^\chi+4b_4^\chi+m_4^{\varepsilon,c}+m_4^{\eta,c}+\omega_1,\\
E\left(\upsilon_{i,1,2}^4\right) &= m_4^\chi+5b_4^\chi+m_4^{\varepsilon,e}+m_4^{\eta,e}+\omega_2,\\
E(\upsilon_{i,1,1}^2,\upsilon_{i,2,2}^2) &= m_4^\chi+4b_4^\chi+\rho^2 m_4^{\eta,c}+\omega_3,\\
E(\upsilon_{i,1,2}^2,\upsilon_{i,2,3}^2) &= m_4^\chi+5b_4^\chi+\rho^2 m_4^{\eta,e}+\omega_4,\\
E(\upsilon_{i,1,1}^2,\upsilon_{i,3,3}^2) &= m_4^\chi+4b_4^\chi+\rho^4 m_4^{\eta,c}+\omega_5,\\
E(\upsilon_{i,1,2}^2,\upsilon_{i,3,4}^2) &= m_4^\chi+5b_4^\chi+\rho^4 m_4^{\eta,e}+\omega_6,
\end{aligned}$$

where $\omega_1-\omega_6$ are real numbers constructed by the parameters we have already identified. So, we have 6 equations in 6 unknowns, m_4^χ, b_4^χ, $m_4^{\eta,c}$, $m_4^{\eta,e}$, $m_4^{\varepsilon,c}$ and $m_4^{\varepsilon,e}$ and therefore the 6-by-6 system above identifies m_4^χ, b_4^χ, $m_4^{\eta,c}$, $m_4^{\eta,e}$, $m_4^{\varepsilon,c}$ and $m_4^{\varepsilon,e}$.

Finally, note that when $size(T)=size(H)=4$, we have 19 parameters to identify and a total of 90 moment conditions. However, in demonstrating identification we have used exactly 19 moments and 19 identifying equations, (23)-(41). However, many parameters of the statistical model are already over-identified even with $size(T)=size(H)=4$. Clearly the parameters will be even more over-identified as $size(T)$ and $size(H)$ grow.

B.4.6 History dependence

The additional information regarding the moments of the innovations to the persistent component is exploited using the 2nd, 3rd and 4th moments of each age-time group. In particular, given the identification of the cohort effects and transitory component, we can take advantage of extra information due to history dependence by using moments such as:

$$\begin{aligned}
E(\upsilon_{i,2,1}^2) &= m_2^\chi(3)+m_2^{\varepsilon,c}+m_2^{\eta,c}+\rho^2 m_2^{\eta,e},\\
&= m_2^{\eta,c}+\rho^2 m_2^{\eta,e}+\xi_1,\\
E(\upsilon_{i,3,1}^2) &= m_2^\chi(2)+m_2^{\varepsilon,c}+m_2^{\eta,c}+\rho^2 m_2^{\eta,e}+\rho^4 m_2^{\eta,c},\\
&= m_2^{\eta,c}+\rho^2 m_2^{\eta,e}+\rho^4 m_2^{\eta,c}+\xi_2,
\end{aligned}$$

$$
\begin{aligned}
E(\upsilon^3_{i,2,1}) &= m^{\chi}_3(3) + m^{\varepsilon,c}_3 + m^{\eta,c}_3 + \rho^3 m^{\eta,e}_3, \\
&= m^{\eta,c}_3 + \rho^3 m^{\eta,e}_3 + \xi_3, \\
E(\upsilon^3_{i,3,1}) &= m^{\chi}_3(2) + m^{\varepsilon,c}_3 + m^{\eta,c}_3 + \rho^3 m^{\eta,e}_3 + \rho^6 m^{\eta,c}_3, \\
&= m^{\eta,c}_3 + \rho^3 m^{\eta,e}_3 + \rho^6 m^{\eta,c}_3 + \xi_4
\end{aligned}
$$

$$
\begin{aligned}
E(\upsilon^4_{i,2,1}) &= \left\{ \begin{array}{c} m^{\chi}_4(3) + m^{\varepsilon,c}_4 + 6m^{\chi}_2(3)m^{\varepsilon,c}_2 \\ +\rho^4 m^{\eta,e}_4 + m^{\eta,c}_4 + 6\rho^2 m^{\eta,e}_2 m^{\eta,c}_2 \\ +6\left(m^{\chi}_2(3) + m^{\varepsilon,c}_2\right)\left(m^{\eta,c}_2 + \rho^2 m^{\eta,e}_2\right) \end{array} \right\}, \\
&= \rho^4 m^{\eta,e}_4 + m^{\eta,c}_4 + \xi_5, \\
E(\upsilon^4_{i,3,1}) &= \left\{ \begin{array}{c} m^{\chi}_4(2) + m^{\varepsilon,c}_4 + 6m^{\chi}_2(2)m^{\varepsilon,c}_2 \\ +\rho^8 m^{\eta,c}_4 + \rho^4 m^{\eta,e}_4 + m^{\eta,c}_4 \\ +6\left(\rho^6 m^{\eta,e}_2 m^{\eta,c}_2 + \rho^4 m^{\eta,c}_2 m^{\eta,c}_2 + \rho^2 m^{\eta,e}_2 m^{\eta,c}_2\right) \\ +6\left(m^{\chi}_2(2) + m^{\varepsilon,c}_2\right)\left(m^{\eta,c}_2 + \rho^2 m^{\eta,e}_2\right) \end{array} \right\}, \\
&= \rho^8 m^{\eta,c}_4 + \rho^4 m^{\eta,e}_4 + m^{\eta,c}_4 + \xi_6.
\end{aligned}
$$

where $\xi_1 - \xi_6$ are real numbers constructed by the parameters we have already identified.

Appendix C: Supplement to Section 4

Table C.1: Transitory income process

[1]	[2]	[3]	[4]	[5]	[6]
$m_2^{\varepsilon,e}$	$m_2^{\varepsilon,c}$	$m_3^{\varepsilon,e}$	$m_3^{\varepsilon,c}$	$m_4^{\varepsilon,e}$	$m_4^{\varepsilon,c}$
0.0250	0.0102	-0.0491***	0.0513**	0.1689***	0.2578

The H$_0$'s $m_2^{\varepsilon,j}=0$, $m_3^{\varepsilon,j}=0$ for $j \in(\{c\}, \{e\})$ are rejected at significance level 1% (***), 5% (**) or 10% (*), based on a two-tail confidence interval procedure. For m_4, the H$_0$ is $m_4^{\varepsilon,j}=3(m_2^{\varepsilon,j})^2$, where $j \in(\{c\}, \{e\})$ and H$_A$ is $m_4^{\varepsilon,e}\geq 3(m_2^{\varepsilon,e})^2$, and is rejected at the same significance levels based on a one-tailed confidence interval. All the tests are implemented using a block bootstrap with 3,000 replications.

Table C.2: Initial conditions

[1]	[2]	[3]	[4]	[5]	[6]
m_2^{χ}	m_3^{χ}	m_4^{χ}	b_2^{χ}	b_3^{χ}	b_4^{χ}
0.1837***	-0.0505*	0.0476	-0.0026***	0.0011	0.0026

The H$_0$'s $m_\tau^{\chi}=0$, $b_\tau^{\chi}=0$ for $\tau=2,3,4$ are rejected at significance level 1% (***), 5% (**) or 10% (*), based on a two-tail confidence interval procedure implemented using a block bootstrap with 3,000 replications.

C.1 Computation of shock distributions

We assume that η_{it}'s follow a mixing of Gaussian distributions:

$$\eta_{i,t} = \begin{cases} \eta_{i,t}^1 \sim N(\mu_{1,f(t)}, \sigma_1^2) \text{ with probability } p_{f(t)} \\ \eta_{i,t}^2 \sim N(\mu_{2,f(t)}, \sigma_2^2) \text{ with probability } 1-p_{f(t)} \end{cases}$$

where $0<\sigma_1^2, \sigma_2^2<\infty$ and $0 \leq p_{f(t)} \leq 1$. For a given aggregate state, the four first theoretical central moments are specified by:

$$E(\eta_{it}) = p_{f(t)}\mu_{1,f(t)} + (1-p)\,\mu_{2,f(t)} = 0,$$
$$E(\eta_{it}^2) = p_{f(t)}\left(\left(\mu_{1,f(t)}\right)^2+\sigma_1^2\right) + \left(1-p_{f(t)}\right)\left(\left(\mu_{2,f(t)}\right)^2+\sigma_2^2\right),$$
$$E(\eta_{it}^3) = p_{f(t)}\left(\left(\mu_{1,f(t)}\right)^3+3\mu_{1,f(t)}\sigma_1^2\right) + \left(1-p_{f(t)}\right)\left(\left(\mu_{2,f(t)}\right)^3+3\mu_{2,f(t)}\sigma_2^2\right),$$
$$E(\eta_{it}^4) = p_{f(t)}\left(\left(\mu_{1,f(t)}\right)^4+6\left(\mu_{1,f(t)}\right)^2\sigma_1^2+3\sigma_1^4\right) + \left(1-p_{f(t)}\right)\left(\left(\mu_{2,f(t)}\right)^4+6\left(\mu_{2,f(t)}\right)^2\sigma_2^2+3\sigma_2^4\right).$$

Note that these moments are equivalent to the moments we have estimated with our estimation procedure. To calibrate the parameters, we solve the following nonlinear least squares problem:

$$\min_{p_e, p_c, \mu_{1e}, \mu_{2e}, \mu_{1c}, \mu_{2c}, \sigma_1^2, \sigma_2^2} \mathbf{D}'\mathbf{D},$$

where

$$\mathbf{D} = \begin{bmatrix} E(\eta_{i,c}) - 0 \\ E(\eta_{ic}^2) - m_2^{\eta,c} \\ E(\eta_{ic}^3) - m_3^{\eta,c} \\ E(\eta_{ic}^4) - m_4^{\eta,c} \\ E(\eta_{ie}) - 0 \\ E(\eta_{ie}^2) - m_2^{\eta,e} \\ E(\eta_{ie}^3) - m_3^{\eta,e} \\ E(\eta_{ie}^4) - m_4^{\eta,e} \end{bmatrix}.$$

The parameters that minimise the objective function are the following:

Table C.4: Parameter calibration

	$p_{f(t)}$	$\mu_{1,f(t)}$	$\mu_{2,f(t)}$	σ_1^2	σ_2^2
expansions	0.7201	0.1486	-0.3823	0.0004	0.0314
recessions	0.9685	0.0474	-1.4582	0.0004	0.0314

Appendix D: Supplement to Section 5

Table D.1: Persistent income process

	[1]	[2]	[3]	[4]	[5]	[6]	[7]
	ρ	$m_2^{\eta,e}$	$m_2^{\eta,c}$	$m_3^{\eta,e}$	$m_3^{\eta,c}$	$m_4^{\eta,e}$	$m_4^{\eta,c}$
Y	0.7391^{***}	0.0654^{***}	0.0701^{***}	-0.0241^{**}	-0.1025^{***}	0.0149	0.1549^{*}
$Y-T$	0.7265^{***}	0.0566^{***}	0.0574^{***}	-0.0175^{***}	-0.0760^{***}	0.0086	0.1040
$Y+bn$	0.7250^{***}	0.0486^{***}	0.0337^{***}	-0.0008	-0.0100	0.0024	0.0041
$Y+bn-T$	0.7162^{***}	0.0408^{***}	0.0270^{***}	-0.0002	-0.0025	0.0017	0.0010

Y, bn and T(=t+ni) refer to labour income, benefits and taxes and national insurance respectively. The H_0's $m_2^{\eta,j}=0$, $m_3^{\eta,j}=0$ for $j\in(\{c\},\{e\})$ are rejected at significance level 1% (***), 5% (**) or 10% (*), based on a two-tail confidence interval procedure. For m_4, the $H_0$is $m_4^{\eta,j}=3(m_2^{\eta,j})^2$, where $j\in(\{c\},\{e\})$ and H_A is $m_4^{\eta,e}\geq 3(m_2^{\eta,e})^2$, and is rejected at the same significance levels based on a one-tailed confidence interval. All the tests are implemented using a block bootstrap with 3,000 replications.

Table D.2: Transitory income process

	[1]	[2]	[3]	[4]	[5]	[6]
	$m_2^{\varepsilon,e}$	$m_2^{\varepsilon,c}$	$m_3^{\varepsilon,e}$	$m_3^{\varepsilon,c}$	$m_4^{\varepsilon,e}$	$m_4^{\varepsilon,c}$
Y	0.0250	0.0102	-0.0491^{***}	0.0513^{**}	0.1689^{**}	0.2578
$Y-T$	0.0180	0.0112	-0.0338^{***}	0.0374^{*}	0.0789	0.1252
$Y+bn$	0.0156	0.0235	-0.0202^{***}	0.0073	0.0266	0.0028
$Y+bn-T$	0.0118^{**}	0.0213^{**}	-0.0124^{**}	0.0014	0.0131	0.0006

Y, bn and T(=t+ni) refer to labour income, benefits and taxes and national insurance respectively. For more details see notes of Table C1.

Table D.3: Initial conditions

	[1]	[2]	[3]	[4]	[5]	[6]
	m_2^{χ}	m_3^{χ}	m_4^{χ}	b_2^{χ}	b_3^{χ}	b_4^{χ}
Y	0.1837^{***}	-0.0505^{*}	0.0476	-0.0026^{***}	0.0011	0.0026
$Y-T$	0.1543^{***}	-0.0403^{**}	0.0343	-0.0022^{***}	0.0009^{*}	0.0029
$Y+bn$	0.1656^{***}	-0.0230	0.0306	-0.0026^{***}	0.0007	0.0092
$Y+bn-T$	0.1343^{***}	-0.0179^{**}	0.0204	-0.0022^{***}	0.0006^{**}	0.0115

Y, bn and T(=t+ni) refer to labour income, benefits and taxes and national insurance respectively. For more details see notes of Table C2.

Appendix E: Robustness of co-*kurtosis*

In contrast to how we have defined co-kurtosis between two random variables $\upsilon_{i,h,t}$ and $\upsilon_{i,h+\kappa,t+\kappa}$, i.e. $CoKurt(\upsilon_{i,h,t}, \upsilon_{i,h+\kappa,t+\kappa}) = E\left[\upsilon_{i,h,t}^2 \upsilon_{i,h+\kappa,t+\kappa}^2\right]$, another measure used in the literature is $CoKurt(\upsilon_{i,h,t}, \upsilon_{i,h+\kappa,t+\kappa})' = E\left[\upsilon_{i,h,t}^3 \upsilon_{i,h+\kappa,t+\kappa}\right]$ (see, e.g. Busch and Ludwig (2020)) where:

$$E\left[\upsilon_{i,h,t}^3 \upsilon_{i,h+\kappa,t+\kappa}\right] = \left\{ \begin{array}{c} m_4^{\chi} + b_4^{\chi}\gamma_g \\ +3\left(m_2^{\chi} + b_2^{\chi}\gamma_g\right)\left(E\left[z_{i,h,t} z_{i,h,t+\kappa}\right] + E\left[z_{i,h,t}^2\right] + m_2^{\varepsilon,f(t)}\right) \\ +E\left[z_{i,h,t}^3 z_{i,h,t+\kappa}\right] + m_2^{\varepsilon,f(t+\kappa)} E\left[z_{i,h,t} z_{i,h+\kappa,t+\kappa}\right] \end{array} \right\}.$$

To assess the robustness of our results, we apply this alternative definition and find that both sets of results are very similar in every respect (see Tables E.1-E.4 below).

E.1 Alternative definition of co-kurtosis

Table E.1: Persistent income process

[1]	[2]	[3]	[4]	[5]	[6]	[7]
ρ	$m_2^{\eta,e}$	$m_2^{\eta,c}$	$m_3^{\eta,e}$	$m_3^{\eta,c}$	$m_4^{\eta,e}$	$m_4^{\eta,c}$
0.7391***	0.0654***	0.0701***	-0.0241**	-0.1025***	0.0424*	0.1549*

The H$_0$'s $\rho = 0$, $m_2^{\eta,j} = 0$, $m_3^{\eta,j} = 0$ for $j \in(\{c\}, \{e\})$ are rejected at significance level 1% (***), 5% (**) or 10% (*), based on a two-tail confidence interval procedure. For m_4, the H$_0$ is $m_4^{\eta,j} = 3\left(m_2^{\eta,j}\right)^2$, where $j \in(\{c\}, \{e\})$ and H$_A$ is $m_4^{\eta,e} \geq 3\left(m_2^{\eta,j}\right)^2$, and is rejected at the same significance levels based on a one-tailed confidence interval. All the tests are implemented using a block bootstrap with 3,000 replications.

Table E.2: Tests of procyclicality

[1]	[2]	[3]
$m_2^{\eta,e} - m_2^{\eta,c}$	$m_3^{\eta,e} - m_3^{\eta,c}$	$m_4^{\eta,e} - m_4^{\eta,c}$
-0.0047	0.0784***	-0.1125

The quantities in the table have been calculated by using the estimates in Table E.1. The H$_0$'s $m_2^{\eta,e} - m_2^{\eta,c} \geq 0$, $m_3^{\eta,e} - m_3^{\eta,c} \leq 0$, $m_4^{\eta,e} - m_4^{\eta,c} \geq 0$ are rejected at significance level 1% (***), 5% (**) or 10% (*), based on a one-tailed confidence interval procedure, implemented using a block bootstrap with 3,000 replications.

Table E.3: Transitory income process

[1]	[2]	[3]	[4]	[5]	[6]
$m_2^{\varepsilon,e}$	$m_2^{\varepsilon,c}$	$m_3^{\varepsilon,e}$	$m_3^{\varepsilon,c}$	$m_4^{\varepsilon,e}$	$m_4^{\varepsilon,c}$
0.0250	0.0102	-0.0491***	0.0513**	0.1025	0.2578

See notes of Table E1.

Table E.4: Initial conditions

[1]	[2]	[3]	[4]	[5]	[6]
m_2^χ	m_3^χ	m_4^χ	b_2^χ	b_3^χ	b_4^χ
0.1837***	-0.0505*	0.1093	-0.0026***	0.0011	0.0014

The H_0's $m_\tau^\chi = 0$, $b_\tau^\chi = 0$ for $\tau = 2, 3, 4$ are rejected at significance level 1% (***), 5% (**) or 10% (*), based on a two-tail confidence interval procedure implemented using a block bootstrap with 3,000 replications.

References

[1] Altonji, J. G., and L. M. Segal (1996). Small-sample bias in GMM estimation of covariance structures, *Journal of Business & Economic Statistics*, 14, 353-366.

[2] Bardasi, E., Jenkins, S., Sutherland, H., Levy, H., and F. Zantomio (2012). British Household Panel Survey derived current and annual net household income variables, Waves 1–18, 1991–2009. University of Essex, Institute for Social and Economic Research, Colchester, Essex: U.K. Data Archive, SN, 3909.

[3] Blundell, R., and B. Etheridge (2010). Consumption, income and earnings inequality in Britain, *Review of Economic Dynamics*, 13, 76–102.

[4] Blundell, R., Graber, M., and M. Mogstad (2015). Labor income dynamics and the insurance from taxes, transfers, and the family, *Journal of Public Economics*, 127, 58-73.

[5] Busch, C., and A. Ludwig (2020). Higher-Order Income Risk over the Business Cycle, *Working Paper* 1159, Barcelona GSE.

[6] Davidson, A. C., and D. V. Hinkley (1997). *Bootstrap methods and their application*, Cambridge University Press.

[7] Domeij, D. and M. Floden (2010). Inequality trends in Sweden 1978–2004, *Review of Economic Dynamics*, 13, 179-208.

[8] Heathcote, J., Perri, F., and G. L. Violante (2010). Unequal we stand: An empirical analysis of economic inequality in the United States, 1967–2006. *Review of Economic Dynamics*, 13, 15-51.

[9] Horowitz, J. L. (2003). Bootstrap methods for Markov processes, *Econometrica*, 71, 1049-1082.

[10] Institute for Social and Economic Research (2018). British Household Panel Survey: Waves 1-18, 1991-2009, 8th Edition, UK Data Service. SN: 5151.

[11] MaCurdy, T. (2007). A practitioner's approach to estimating intertemporal relationships using longitudinal data: lessons from applications in wage dynamics, *Handbook of Econometrics*, 6, 4057-4167.

[12] Mincer, J. (1974). *Schooling, Experience, and Earnings*, Human Behavior & Social Institutions No. 2.

[13] Sen, A. (2012). On the interrelation between the sample mean and the sample variance, *The American Statistician*, 66, 112-117.

[14] Storesletten, K., Telmer, C., and A. Yaron (2004). Cyclical dynamics in idiosyncratic labor market risk, *Journal of Political Economy*, 112, 695-717.

[15] Taylor, M., Brice J., Buck N., and E. Prentice-Lane (2010). British Household Panel Survey User Manual Volume A: Introduction, Technical Report and Appendices. Colchester: University of Essex.