CONSUMPTION INEQUALITY AND FAMILY LABOR SUPPLY: Online Appendix, Not for Publication*

Appendix 1: Approximation of the First Order Conditions and Intertemporal Budget Constraint

This appendix explains the derivation of the equations linking consumption and labor supply to the transitory and permanent wage shocks of the two earners.

We start by showing how to derive equations (7) and (8) in the text using approximations of the first order conditions of the life cycle problem. For discount rate δ and interest rate r, the first order condition for assets gives:

$$E_t\left(\lambda_{i,t+1}\right) = \frac{1+\delta}{1+r}\lambda_{i,t}$$

Define ρ such that $e^{\rho} = \frac{1+\delta}{1+r}$, and apply a second order Taylor approximation around $\ln\lambda_{i,t} + \rho$:

$$\lambda_{i,t+1} \simeq \lambda_{i,t}e^{\rho}\left[1 + \left(\Delta\ln\lambda_{i,t+1} - \rho\right) + \frac{1}{2}\left(\Delta\ln\lambda_{i,t+1} - \rho\right)^2\right]$$

Taking expectations with respect to the information set at time t yields:

$$E_t\left[\Delta\ln\lambda_{i,t+1}\right] \simeq \rho - \frac{1}{2}E_t\left(\Delta\ln\lambda_{i,t+1} - \rho\right)^2$$

which in turn can be written as

$$\Delta\ln\lambda_{i,t+1} = \omega_{t+1} + \varepsilon_{i,t+1} \tag{A1.1}$$

*Blundell: Department of Economics, University College London, and Institute for Fiscal Studies, Gower Street, London WC1E 6BT, UK (email: r.blundell@ucl.ac.uk); Pistaferri: Department of Economics, Stanford University, Stanford, CA 94305 (email: pista@stanford.edu); Saporta-Eksten: The Eitan Berglas School of Economics, Tel Aviv, Israel, and Department of Economics, University College London (email: itaysap@post.tau.ac.il). We thank Mark Aguiar, Nick Bloom, Olympia Bover, Martin Browning, Chris Carroll, Jon Levin, Tom MaCurdy, Alessandra Voena, Gianluca Violante and participants at various conferences and workshops in the US and Europe for comments. Thanks to Kerwin Charles for initial help with the new consumption data from the PSID. The authors gratefully acknowledge financial support from the UK Economic and Social Research Council (Blundell), and the ERC starting grant 284024 (Pistaferri). All errors are ours.

where $\omega_{t+1} = \rho - \frac{1}{2}E_t\left(\Delta \ln \lambda_{i,t+1} - \rho\right)^2$ is assumed to be fixed in the cross section and $E_t\left(\varepsilon_{i,t+1}\right) = 0$ by definition of an expectation error.

Apply a first order Taylor approximation for $u_c\left(C_{i,t+1}, H_{i,1,t+1}, H_{i,2,t+1}\right)$ around $c_t \equiv \ln\left(C_t\right), h_{1,t} \equiv \ln\left(H_{1,t}\right)$ and $h_{2,t} \equiv \ln\left(H_{2,t}\right)$ (omitting the household subscripts i for simplicity):

$$\Delta u_c\left(t+1\right) \simeq \frac{C_t u_{cc}\left(t\right)}{u_c\left(t\right)}\Delta c_{t+1} + \frac{H_{1,t}u_{ch_1}\left(t\right)}{u_c\left(t\right)}\Delta h_{1,t+1} + \frac{H_{2,t}u_{ch_2}\left(t\right)}{u_c\left(t\right)}\Delta h_{2,t+1} \tag{A1.2}$$

where $\Delta x_{t+1} = x_{t+1} - x_t$, and we replace the arguments in the derivative functions with timing for brevity. We can similarly approximate $u_{h_1}\left(C_{i,t+1}, H_{i,1,t+1}, H_{i,2,t+1}\right)$ and $u_{h_2}\left(C_{i,t+1}, H_{i,1,t+1}, H_{i,2,t+1}\right)$.

The first order conditions for consumption and hours of the two earners are given by:

$$\begin{aligned} u_c\left(t\right) &= \lambda_t \\ -u_{h_1}\left(t\right) &= \lambda_t\left(1-\chi_t\right)\left(1-\mu_t\right)Y_t^{-\mu_t}W_{1,t} \\ -u_{h_2}\left(t\right) &= \lambda_t\left(1-\chi_t\right)\left(1-\mu_t\right)Y_t^{-\mu_t}W_{2,t}. \end{aligned}$$

Substituting these first order conditions into (A1.2) and re-arranging we get equations (7) and (8) in the text. In matrix notation (and omitting the subscript t from μ_t) these are:

$$\begin{pmatrix} \Delta c_{t+1} \\ \Delta h_{1,t+1} \\ \Delta h_{2,t+1} \end{pmatrix} = \begin{pmatrix} -\eta_{c,p} & \eta_{c,w_1} & \eta_{c,w_2} \\ \eta_{h_1,p} & \eta_{h_1,w_1} & \eta_{h_1,w_2} \\ \eta_{h_2,p} & \eta_{h_2,w_1} & \eta_{h_2,w_2} \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & -\mu \\ 1 & 0 & 1 & -\mu \end{pmatrix} \begin{pmatrix} \Delta \ln \lambda_{t+1} \\ \Delta w_{1,t+1} \\ \Delta w_{2,t+1} \\ \Delta y_{t+1} \end{pmatrix}$$

where the $\eta_{l,m}$ are Frisch elasticities, defined as:

$$\begin{aligned} \eta_{c,p} &= -\frac{1}{C_t}\frac{1}{|\mathcal{H}|}u_c\left(u_{h_1h_1}u_{h_2h_2} - u_{h_1h_2}^2\right) \\ \eta_{c,w_1} &= \frac{1}{C_t}\frac{1}{|\mathcal{H}|}u_{h_1}\left(u_{ch_2}u_{h_2h_1} - u_{ch_1}u_{h_2h_2}\right) \\ \eta_{c,w_2} &= \frac{1}{C_t}\frac{1}{|\mathcal{H}|}u_{h_2}\left(u_{ch_1}u_{h_1h_2} - u_{h_1h_1}u_{ch_2}\right) \\ \eta_{h_1,p} &= \frac{1}{H_{1,t}}\frac{1}{|\mathcal{H}|}u_c\left(u_{h_1h_2}u_{h_2c} - u_{h_1c}u_{h_2h_2}\right) \\ \eta_{h_1,w_1} &= \frac{1}{H_{1,t}}\frac{1}{|\mathcal{H}|}u_{h_1}\left(u_{h_2h_2}u_{cc} - u_{ch_2}^2\right) \\ \eta_{h_1,w_2} &= \frac{1}{H_{1,t}}\frac{1}{|\mathcal{H}|}u_{h_2}\left(u_{ch_1}u_{ch_2} - u_{cc}u_{h_1h_2}\right) \\ \eta_{h_2,p} &= \frac{1}{H_{2,t}}\frac{1}{|\mathcal{H}|}u_c\left(u_{ch_1}u_{h_2h_1} - u_{ch_2}u_{h_1h_1}\right) \\ \eta_{h_2,w_1} &= \frac{1}{H_{2,t}}\frac{1}{|\mathcal{H}|}u_{h_1}\left(u_{ch_2}u_{ch_1} - u_{cc}u_{h_2h_1}\right) \\ \eta_{h_2,w_2} &= \frac{1}{H_{2,t}}\frac{1}{|\mathcal{H}|}u_{h_2}\left(u_{h_1h_1}u_{cc} - u_{ch_1}^2\right) \end{aligned}$$

Note that the negative sign on $\eta_{c,p}$ is to keep the convention that $\eta_{c,p} > 0$, and $|\mathcal{H}|$ is the determinant of the Hessian of the utility function:

$$|\mathcal{H}| = u_{cc}u_{h_1h_1}u_{h_2h_2} - u_{h_2h_2}u_{ch_1}^2 - u_{h_1h_1}u_{ch_2}^2 - u_{cc}u_{h_1h_2}^2 + 2u_{ch_1}u_{ch_2}u_{h_1h_2}$$

We next need to express Δy_{t+1} as changes in wages and hours of the two earners. Define $q_t = \frac{Y_{1,t}}{Y_t}$ (the share of income from the head's earnings at time t). Then:

$$\begin{aligned} \ln(Y_{t+1}) &\simeq \ln(Y_t) + \frac{Y_{1,t}}{Y_t}\Delta \ln Y_{1,t+1} + \frac{Y_{2,t}}{Y_t}\Delta \ln Y_{2,t+1} && \text{(A1.3)} \\ &\Rightarrow \Delta y_{t+1} \simeq q_t \Delta y_{1,t+1} + (1-q_t)\Delta y_{2,t+1} && \text{(A1.4)} \end{aligned}$$

where q_t is predetermined (by the nature of the approximation). Since $\Delta y = \Delta h + \Delta w$, one can substitute (A1.4) into equations (7) and (8) in the main text to yield:

$$\begin{pmatrix} \Delta c_{t+1} \\ \Delta h_{1,t+1} \\ \Delta h_{2,t+1} \end{pmatrix} = \underbrace{\begin{pmatrix} \psi_{c,\lambda} & \psi_{c,w_1} & \psi_{c,w_2} \\ \psi_{h_1,\lambda} & \psi_{h_1,w_1} & \psi_{h_1,w_2} \\ \psi_{h_2,\lambda} & \psi_{h_2,w_1} & \psi_{h_2,w_2} \end{pmatrix}}_{\Psi} \begin{pmatrix} \Delta \ln \lambda_{t+1} \\ \Delta w_{1,t+1} \\ \Delta w_{2,t+1} \end{pmatrix} \tag{A1.5}$$

where

$$\begin{aligned} \Psi &= \Upsilon^{-1}\varrho \\ \Upsilon &= \begin{pmatrix} 1 & \mu q_t\left(\eta_{c,w_1}+\eta_{c,w_2}\right) & \mu(1-q_t)\left(\eta_{c,w_1}+\eta_{c,w_2}\right) \\ 0 & 1+\mu q_t\left(\eta_{h_1,w_1}+\eta_{h_1,w_2}\right) & \mu(1-q_t)\left(\eta_{h_1,w_1}+\eta_{h_1,w_2}\right) \\ 0 & \mu q_t\left(\eta_{h_2,w_1}+\eta_{h_2,w_2}\right) & 1+\mu(1-q_t)\left(\eta_{h_2,w_1}+\eta_{h_2,w_2}\right) \end{pmatrix} \\ \varrho &= \begin{pmatrix} -\eta_{c,p}+\eta_{c,w_1}+\eta_{c,w_2} & \eta_{c,w_1}-\mu q_t\left(\eta_{c,w_1}+\eta_{c,w_2}\right) & \eta_{c,w_2}-\mu(1-q_t)\left(\eta_{c,w_1}+\eta_{c,w_2}\right) \\ \eta_{h_1,p}+\eta_{h_1,w_1}+\eta_{h_1,w_2} & \eta_{h_1,w_1}-\mu q_t\left(\eta_{h_1,w_1}+\eta_{h_1,w_2}\right) & \eta_{h_1,w_2}-\mu(1-q_t)\left(\eta_{h_1,w_1}+\eta_{h_1,w_2}\right) \\ \eta_{h_2,p}+\eta_{h_2,w_1}+\eta_{h_2,w_2} & \eta_{h_2,w_1}-\mu q_t\left(\eta_{h_2,w_1}+\eta_{h_2,w_2}\right) & \eta_{h_2,w_2}-\mu(1-q_t)\left(\eta_{h_2,w_1}+\eta_{h_2,w_2}\right) \end{pmatrix} \end{aligned}$$

The system in (A1.5) provides the dynamics of consumption and hours as a function of the change in wages and the changes in the marginal utility of wealth ($\Delta \ln \lambda_{t+1}$). Note that with proportional taxation ($\mu = 0$), Υ is an identity matrix, and hence Ψ reduces to:

$$\Psi = \begin{pmatrix} -\eta_{c,p}+\eta_{c,w_1}+\eta_{c,w_2} & \eta_{c,w_1} & \eta_{c,w_2} \\ \eta_{h_1,p}+\eta_{h_1,w_1}+\eta_{h_1,w_2} & \eta_{h_1,w_1} & \eta_{h_1,w_2} \\ \eta_{h_2,p}+\eta_{h_2,w_1}+\eta_{h_2,w_2} & \eta_{h_2,w_1} & \eta_{h_2,w_2} \end{pmatrix}$$

The final step is to eliminate the change in the marginal utility of wealth from this system (as it is unobserved), and write the dynamics of consumption and hours as functions of the wage shocks. To do that, we approximate the intertemporal budget constraint:

$$\sum_{s=0}^{T-t} \frac{C_{i,t+s}}{(1+r)^s} = A_t + \sum_{s=0}^{T-t} \frac{(1-\chi)\,Y_{i,t+s}^{1-\mu}}{(1+r)^s}$$

where $Y_{i,t} = H_{i,1,t}W_{i,1,t} + H_{i,2,t}W_{i,2,t}$.

The general approximation rule (see for example Blundell, Low and Preston, 2013) is :

$$\begin{aligned} E_I\left[\ln\sum_{k=0}^{T-t}\exp\xi_k\right] &= \ln\sum_{i=0}^{T-t}\exp\xi_k^0 + \sum_{i=0}^{T-t}\frac{\exp\xi_k^0}{\sum_{i=0}^{T-t}\exp\xi_k^0}\left(E_I\xi_k - \xi_k^0\right) \\ &+\frac{1}{2}\sum_{k=0}^{T-t}\sum_{l=0}^{T-t}E_I\left(\frac{\partial^2}{\xi_k\xi_l}\left[\ln\sum_{k=0}^{T-t}\exp\tilde{\xi}_k\right]\left(\xi_k-\xi_k^0\right)\left(\xi_l-\xi_l^0\right)\right) \end{aligned}$$

where ξ_k is the series we wish to approximate, ξ_k^0 the series we are approximating around, and $\tilde{\xi}$ a vector chosen such that the Taylor expansion is accurate. We neglect the second order term from now on. For an accurate derivation of the order of magnitude of this term see Blundell, Low and Preston (2013).

Plug the left hand side of the budget constraint into the approximation by defining $\xi_k = \ln C_{t+k} - k\ln(1+r)$ and $\xi_k^0 = E_{t-1}\ln C_{t+k} - k\ln(1+r)$ to get:

$$\begin{aligned} E_I\left[\ln\sum_{k=0}^{T-t}\frac{C_{t+k}}{(1+r)^k}\right] &= \ln\sum_{k=0}^{T-t}\exp\left[E_{t-1}\ln C_{t+k} - k\ln(1+r)\right] \\ &+\sum_{k=0}^{T-t}\frac{\exp\left[E_{t-1}\ln C_{t+k} - k\ln(1+r)\right]}{\sum_{j=0}^{T-t}\exp\left[E_{t-1}\ln C_{t+j} - j\ln(1+r)\right]}\left(E_I\ln C_{t+k} - E_{t-1}\ln C_{t+k}\right) \end{aligned}$$

Defining $\theta_{t+k} = \frac{\exp[E_{t-1}\ln C_{t+k} - k\ln(1+r)]}{\sum_{j=0}^{T-t}\exp[E_{t-1}\ln C_{t+j} - j\ln(1+r)]}$, we can use the more compact notation:

$$\begin{aligned} E_I\left[\ln\sum_{k=0}^{T-t}\frac{C_{t+k}}{(1+r)^k}\right] &\simeq \ln\sum_{k=0}^{T-t}\exp\left[E_{t-1}\ln C_{t+k} - k\ln(1+r)\right] \\ &+\sum_{k=0}^{T-t}\theta_{t+k}\left(E_I\ln C_{t+k} - E_{t-1}\ln C_{t+k}\right) \end{aligned}$$

Note that θ_{t+k} is known at t (since we are taking expectations dated $t-1$). To proceed, note that from (A1.1) and (A1.5):

$$\ln C_{t+k} \simeq \ln C_{t+k-1} + \psi_{c,\lambda}\left(\omega_{t+k} + \varepsilon_{i,t+k}\right) + \psi_{c,w_1}\left(v_{1,t+k} + \Delta u_{1,t+k}\right) + \psi_{c,w_2}\left(v_{2,t+k} + \Delta u_{2,t+k}\right)$$

Therefore, defining $I = t$:[1]

$$\begin{aligned} &\sum_{k=0}^{T-t}\theta_{t+k}\left(E_t\ln C_{t+k} - E_{t-1}\ln C_{t+k}\right) \\ \simeq\ &\psi_{c,\lambda}\varepsilon_{i,t} + \psi_{c,w_1}v_{1,t} + \psi_{c,w_2}v_{2,t} + \theta_t\left(\psi_{c,w_1}u_{1,t} + \psi_{c,w_2}u_{2,t}\right) \end{aligned}$$

where the last equality comes from the identity $\left(\sum_{k=0}^{T-t}\theta_{t+k}\right) = 1$. Now assume that θ_t (consumption today as a share of remaining lifetime consumption) is small and can be neglected. Take the difference:

$$E_t\left[\ln\sum_{k=0}^{T-t}\frac{C_{t+k}}{(1+r)^k}\right] - E_{t-1}\left[\ln\sum_{k=0}^{T-t}\frac{C_{t+k}}{(1+r)^k}\right] \simeq \psi_{c,\lambda}\varepsilon_{i,t} + \psi_{c,w_1}v_{1,t} + \psi_{c,w_2}v_{2,t} \qquad \text{(A1.6)}$$

[1] We do not need to keep track of ω_t as it does not contain any stochastic elements.

We now need to apply a similar approximation to the right hand side of the budget constraint. Define:

$$
\begin{aligned}
Q_1 &= \sum_{k=0}^{T-t} \exp\left[E_{t-1} \ln (1-\chi) Y_{i,t+k}^{1-\mu} - k \ln (1+r)\right] \\
Q_2 &= \exp E_{t-1} \ln A_t
\end{aligned}
$$

where Q_1 is approximately human wealth at time t. Apply the approximation above to the RHS of the intertemporal budget constraint:

$$
\begin{aligned}
& E_t \left[\ln\left(\sum_{k=0}^{T-t} \frac{(1-\chi) Y_{i,t+k}^{1-\mu}}{(1+r)^k} + A_t\right)\right] = \qquad \text{(A1.7)} \\
\simeq\ & \ln\left[\sum_{k=0}^{T-t} \exp\left[E_{t-1}(1-\chi) Y_{i,t+k}^{1-\mu} - k\ln(1+r)\right] + \exp E_{t-1} \ln A_t\right] \\
& + (1-\mu) \sum_{k=0}^{T-t} \frac{\exp\left[E_{t-1} \ln\left[(1-\chi) Y_{i,t+k}^{1-\mu}\right] - k \ln (1+r)\right]}{Q_1 + Q_2} \left[E_t \ln (Y_{1,t+k} + Y_{2,t+k}) - E_{t-1} \ln (Y_{1,t+k} + Y_{2,t+k})\right] \\
& + \frac{Q_2}{Q_1+Q_2} \left(E_I \ln A_t - E_{t-1} \ln A_t\right)
\end{aligned}
$$

We now approximate the $\ln Y_{t+k}$ term to obtain:

$$
\begin{aligned}
\ln(Y_{1,t+k} + Y_{2,t+k}) \simeq\ & \ln(E_{t-1}Y_{1,t+k-1} + E_{t-1}Y_{2,t+k-1}) \\
& + (\ln Y_{1,t+k} - \ln E_{t-1}Y_{1,t+k-1}) \frac{E_{t-1}Y_{1,t+k-1}}{E_{t-1}Y_{1,t+k-1} + E_{t-1}Y_{2,t+k-1}} \\
& + (\ln Y_{2,t+k} - \ln E_{t-1}Y_{2,t+k-1}) \frac{E_{t-1}Y_{2,t+k-1}}{E_{t-1}Y_{1,t+k-1} + E_{t-1}Y_{2,t+k-1}} + \xi
\end{aligned}
$$

and define $\tilde{q}_{t+k-1} = \frac{E_{t-1}Y_{1,t+k-1}}{E_{t-1}Y_{1,t+k-1}+E_{t-1}Y_{2,t+k-1}}$. Note that $\tilde{q}$ is determined at time $t-1$ for all periods because of the expectations. Define:

$$
\begin{aligned}
\alpha_{t+k} &= \frac{\exp\left[E_{t-1} \ln (1-\chi) Y_{i,t+k}^{1-\mu} - k \ln (1+r)\right]}{Q_1} \\
\pi_t &= \frac{\exp\left[E_{t-1} \ln A_t\right]}{Q_1 + Q_2}
\end{aligned}
$$

and (A1.7) can be written as:

$$
\begin{aligned}
& E_t\left[\ln\left(\sum_{k=0}^{T-t}\frac{(1-\chi)\,Y_{i,t+k}^{1-\mu}}{(1+r)^k}+A_t\right)\right] \\
\simeq\quad & \ln\left[\sum_{k=0}^{T-t}\exp\left[E_{t-1}(1-\chi)(Y_{t+k})^{1-\mu}-k\ln(1+r)\right]+\exp E_{t-1}\ln A_t\right] \\
& +(1-\mu)(1-\pi_t)\sum_{k=0}^{T-t}\alpha_{t+k}\tilde{q}_{t+k-1}\left(E_t\left[\ln Y_{1,t+k}\right]-E_{t-1}\left[\ln Y_{1,t+k}\right]\right) \\
& +(1-\mu)(1-\pi_t)\sum_{k=0}^{T-t}\alpha_{t+k}\left(1-\tilde{q}_{t+k-1}\right)\left(E_t\left[\ln Y_{2,t+k}\right]-E_{t-1}\left[\ln Y_{2,t+k}\right]\right) \\
& +\pi_t\left(E_I\ln A_t-E_{t-1}\ln A_t\right)
\end{aligned}
$$

There are four sums to solve:

$$\sum_{k=0}^{T-t}\alpha_{t+k}\tilde{q}_{t+k-1}\left(E_t\left[\ln W_{1,t+k}\right]-E_{t-1}\left[\ln W_{1,t+k}\right]\right) \tag{A1.8}$$

$$\sum_{k=0}^{T-t}\alpha_{t+k}\left(1-\tilde{q}_{t+k-1}\right)\left(E_t\left[\ln W_{2,t+k}\right]-E_{t-1}\left[\ln W_{2,t+k}\right]\right) \tag{A1.9}$$

$$\sum_{k=0}^{T-t}\alpha_{t+k}\tilde{q}_{t+k-1}\left(E_t\left[\ln H_{1,t+k}\right]-E_{t-1}\left[\ln H_{1,t+k}\right]\right) \tag{A1.10}$$

$$\sum_{k=0}^{T-t}\alpha_{t+k}\left(1-\tilde{q}_{t+k-1}\right)\left(E_t\left[\ln H_{2,t+k}\right]-E_{t-1}\left[\ln H_{2,t+k}\right]\right) \tag{A1.11}$$

where the first two (A1.8 and A1.9) are easier. Hence we focus on the solution for the third summation (A1.10):

$$
\begin{aligned}
& \sum_{k=0}^{T-t}\alpha_{t+k}\tilde{q}_{t+k-1}\left(E_t\left[\ln H_{1,i,t+k}\right]-E_{t-1}\left[\ln H_{1,i,t+k}\right]\right) \\
=\quad & \sum_{k=0}^{T-t}\alpha_{t+k}\tilde{q}_{t+k-1}\psi_{h_1,\lambda}\varepsilon_t+\sum_{k=0}^{T-t}\alpha_{t+k}\tilde{q}_{t+k-1}\psi_{h_1,w_1}v_{1,t}+\sum_{k=0}^{T-t}\alpha_{t+k}\tilde{q}_{t+k-1}\psi_{h_1,w_2}v_{2,t} \\
& +\alpha_t\tilde{q}_{t-1}\left(\psi_{h_1,w_1}u_{1,t}+\psi_{h_1,w_2}u_{2,t}\right)
\end{aligned}
$$

and as with θ_t, we assume that α_t is small. Hence, we can neglect $\alpha_t\tilde{q}_{t-1}\left(\psi_{h_1,w_1}u_{1,t}+\psi_{h_1,w_2}u_{2,t}\right)$. We further define:

$$s_t=\sum_{k=0}^{T-t}\alpha_{t+k}\tilde{q}_{t+k-1}$$

which allows us to simplify sum (A1.10) and similarly obtain sum (A1.11) below:

$$\begin{aligned}\sum_{k=0}^{T-t}\alpha_{t+k}\tilde{q}_{t+k-1}\left(E_t\left[\ln H_{1,i,t+k}\right]-E_{t-1}\left[\ln H_{1,i,t+k}\right]\right) &= s_t\left(\psi_{h_1,\lambda}\varepsilon_t+\psi_{h_1,w_1}v_{1,t}+\psi_{h_1,w_2}v_{2,t}\right)\\ \sum_{k=0}^{T-t}\alpha_{t+k}\left(1-\tilde{q}_{t+k-1}\right)\left(E_t\left[\ln H_{2,t+k}\right]-E_{t-1}\left[\ln H_{2,t+k}\right]\right) &= \left(1-s_t\right)\left(\psi_{h_2,\lambda}\varepsilon_t+\psi_{h_2,w_1}v_{1,t}+\psi_{h_2,w_2}v_{2,t}\right)\end{aligned}$$

In a similar way we can also solve for the first two summations (A1.8) and (A1.9):

$$\begin{aligned}\sum_{k=0}^{T-t}\alpha_{t+k}\tilde{q}_{t+k-1}\left(E_t\left[\ln W_{1,i,t+k}\right]-E_{t-1}\left[\ln W_{1,i,t+k}\right]\right) &= s_t v_{1,t}\\ \sum_{k=0}^{T-t}\alpha_{t+k}\left(1-\tilde{q}_{t+k-1}\right)\left(E_t\left[\ln W_{2,t+k}\right]-E_{t-1}\left[\ln W_{2,t+k}\right]\right) &= \left(1-s_t\right)v_{2,t}\end{aligned}$$

Similar to the approximation of the LHS, take the difference in expectations between $t-1$ and t:

$$\begin{aligned}&E_t\left[\ln\left(\sum_{k=0}^{T-t}\frac{(1-\chi)\,Y_{i,t+k}^{1-\mu}}{(1+r)^k}+A_t\right)\right]-E_{t-1}\left[\ln\left(\sum_{k=0}^{T-t}\frac{(1-\chi)\,Y_{i,t+k}^{1-\mu}}{(1+r)^k}+A_t\right)\right]\\ &= (1-\mu)\left(1-\pi_t\right)\left[\begin{array}{c}\left(s_t\psi_{h_1,\lambda}+\left(1-s_t\right)\psi_{h_2,\lambda}\right)\varepsilon_t\\ +\left(s_t\left(1+\psi_{h_1,w_1}\right)+\left(1-s_t\right)\psi_{h_2,w_1}\right)v_{1,t}\\ +\left(s_t\psi_{h_1,w_2}+\left(1-s_t\right)\left(1+\psi_{h_2,w_2}\right)\right)v_{2,t}\end{array}\right]\end{aligned}$$

We can now define:

$$\begin{aligned}\tilde{\psi}_{h,\lambda} &= s_t\psi_{h_1,\lambda}+\left(1-s_t\right)\psi_{h_2,\lambda}\\ \tilde{\psi}_{h,w_1} &= s_t\psi_{h_1,w_1}+\left(1-s_t\right)\psi_{h_2,w_1}\\ \tilde{\psi}_{h,w_2} &= s_t\psi_{h_1,w_2}+\left(1-s_t\right)\psi_{h_2,w_2},\end{aligned}$$

and combine with the approximation to the LHS of the intertemporal BC (A1.6), to get

$$\psi_{c,\lambda}\varepsilon_t+\psi_{c,w_1}v_1+\psi_{c,w_2}v_2=(1-\mu)\left(1-\pi_t\right)\left[\tilde{\psi}_{h,\lambda}\varepsilon_t+\left(s_t+\tilde{\psi}_{h,w_1}\right)v_{1,t}+\left(\left(1-s_t\right)+\tilde{\psi}_{h,w_2}\right)v_{2,t}\right]$$

from which one can calculate ε_t, the innovation in the marginal utility of wealth:

$$\varepsilon_t=\frac{1}{\psi_{c,\lambda}-(1-\mu)\left(1-\pi_t\right)\tilde{\psi}_{h,\lambda}}\left\{\begin{array}{c}\left[(1-\mu)\left(1-\pi_t\right)\left(s_t+\tilde{\psi}_{h,w_1}\right)-\psi_{c,w_1}\right]v_{1,t}\\ +\left[(1-\mu)\left(1-\pi_t\right)\left(\left(1-s_t\right)+\tilde{\psi}_{h,w_2}\right)-\psi_{c,w_2}\right]v_{2,t}\end{array}\right\}\tag{A1.12}$$

We can now plug (A1.12) into (A1.5) and simplify to get:

$$\begin{aligned}
\Delta c_t &\simeq \psi_{c,w_1}\Delta u_1 + \psi_{c,w_2}\Delta u_2 + \left(\psi_{c,w_1} + \frac{\psi_{c,\lambda}\left[(1-\mu)(1-\pi_t)\left(s_t+\tilde{\psi}_{h,w_1}\right)-\psi_{c,w_1}\right]}{\psi_{c,\lambda}-(1-\mu)(1-\pi_t)\tilde{\psi}_{h,\lambda}}\right)v_{1,t} \quad \text{(A1.13)}\\
&\quad + \left(\psi_{c,w_2} + \frac{\psi_{c,\lambda}\left[(1-\mu)(1-\pi_t)\left((1-s_t)+\tilde{\psi}_{h,w_2}\right)-\psi_{c,w_2}\right]}{\psi_{c,\lambda}-(1-\mu)(1-\pi_t)\tilde{\psi}_{h,\lambda}}\right)v_{2,t}\\
\Delta h_{1,t} &\simeq \psi_{h_1,w_1}\Delta u_1 + \psi_{h_1,w_2}\Delta u_2 + \left(\psi_{h_1,w_1} + \frac{\psi_{h_1,\lambda}\left[(1-\mu)(1-\pi_t)\left(s_t+\tilde{\psi}_{h,w_1}\right)-\psi_{c,w_1}\right]}{\psi_{c,\lambda}-(1-\mu)(1-\pi_t)\tilde{\psi}_{h,\lambda}}\right)v_1 \quad \text{(A1.14)}\\
&\quad + \left(\psi_{h_1,w_2} + \frac{\psi_{h_1,\lambda}\left[(1-\mu)(1-\pi_t)\left((1-s_t)+\tilde{\psi}_{h,w_2}\right)-\psi_{c,w_2}\right]}{\psi_{c,\lambda}-(1-\mu)(1-\pi_t)\tilde{\psi}_{h,\lambda}}\right)v_2\\
\Delta h_{2,t} &\simeq \psi_{h_2,w_1}\Delta u_1 + \psi_{h_2,w_2}\Delta u_2 + \left(\psi_{h_2,w_1} + \frac{\psi_{h_2,\lambda}\left[(1-\mu)(1-\pi_t)\left(s_t+\tilde{\psi}_{h,w_1}\right)-\psi_{c,w_1}\right]}{\psi_{c,\lambda}-(1-\mu)(1-\pi_t)\tilde{\psi}_{h,\lambda}}\right)v_1 \quad \text{(A1.15)}\\
&\quad + \left(\psi_{h_2,w_2} + \frac{\psi_{h_2,\lambda}\left[(1-\mu)(1-\pi_t)\left((1-s_t)+\tilde{\psi}_{h,w_2}\right)-\psi_{c,w_2}\right]}{\psi_{c,\lambda}-(1-\mu)(1-\pi_t)\tilde{\psi}_{h,\lambda}}\right)v_2
\end{aligned}$$

or more compactly as in equation (9) in the paper:

$$\begin{aligned}
\Delta c_t &\simeq \kappa_{c,u_1}\Delta u_{1,t} + \kappa_{c,u_2}\Delta u_{2,t} + \kappa_{c,v_1}v_{1,t} + \kappa_{c,v_2}v_{2,t}\\
\Delta h_{1,t} &\simeq \kappa_{h_1,u_1}\Delta u_{1,t} + \kappa_{h_1,u_2}\Delta u_{2,t} + \kappa_{h_1,v_1}v_{1,t} + \kappa_{h_1,v_2}v_{2,t}\\
\Delta h_{2,t} &\simeq \kappa_{h_2,u_1}\Delta u_{1,t} + \kappa_{h_2,u_2}\Delta u_{2,t} + \kappa_{h_2,v_1}v_{1,t} + \kappa_{h_2,v_2}v_{2,t}
\end{aligned}$$

with the κ's defined as the relevant coefficients in equations (A1.13), (A1.14) and (A1.15). Note that the κ's have both subscript t and subscript i (omitted here for brevity).

From the estimation point of view there are two things to note. First, the estimation is conducted using earnings data, where earning growth is the sum of hours and wage growth implying $\kappa_{y_j,u_j} = \kappa_{h_j,u_j} + 1$ and $\kappa_{y_j,v_j} = \kappa_{h_j,v_j} + 1$ (all other coefficients are unchanged). Second, while we allow for κ's to change over households and over time, we make the simplifying assumption that the κ's are approximately the same over two adjacent periods within the household, which significantly simplifies the expressions for the moment conditions.

Finally, using the definition of the elements of Ψ, it is easy to show that with linear taxes (i.e. $\mu = 0$, so that Υ becomes an identity matrix), the coefficients on the transitory shocks are the (own- and cross-) Frisch elasticities of labour supply:

$$\kappa_{c,u_j} = \eta_{c,w_j}, \kappa_{h_j,u_j} = \eta_{h_j,w_j}, \kappa_{h_j,u_{-j}} = \eta_{h_j,w_{-j}} \quad \text{(A1.16)}$$

and the mapping with the other parameters is:[2]

[2] To make notation more compact, we use s_j rather than s. With this notation, $s_1 = s$ and $s_2 = 1 - s$.

$$\kappa_{c,v_j} = \eta_{c,w_j} + \frac{\left(\eta_{c,p} - \eta_{c,w_1} - \eta_{c,w_2}\right)\left[(1-\pi_t)\left(s_j + \overline{\eta_{h,w_j}}\right) - \eta_{c,w_j}\right]}{\left(\eta_{c,p} - \eta_{c,w_1} - \eta_{c,w_2}\right) + (1-\pi_t)\left(\overline{\eta_{h,w_1}} + \overline{\eta_{h,w_2}} + \overline{\eta_{h,p}}\right)} \tag{A1.17}$$

$$\kappa_{h_j,v_j} = \eta_{h_j,w_j} - \frac{\left(\eta_{h_j,p} + \eta_{h_j,w_j} + \eta_{h_j,w_{-j}}\right)\left[(1-\pi_t)\left(s_j + \overline{\eta_{h,w_j}}\right) - \eta_{c,w_j}\right]}{\left(\eta_{c,p} - \eta_{c,w_1} - \eta_{c,w_2}\right) + (1-\pi_t)\left(\overline{\eta_{h,w_1}} + \overline{\eta_{h,w_2}} + \overline{\eta_{h,p}}\right)} \tag{A1.18}$$

$$\kappa_{h_j,v_{-j}} = \eta_{h_j,w_{-j}} - \frac{\left(\eta_{h_j,p} + \eta_{h_j,w_j} + \eta_{h_j,w_{-j}}\right)\left[(1-\pi_t)\left((1-s_t) + \overline{\eta_{h,w_{-j}}}\right) - \eta_{c,w_{-j}}\right]}{\left(\eta_{c,p} - \eta_{c,w_1} - \eta_{c,w_2}\right) + (1-\pi_t)\left(\overline{\eta_{h,w_1}} + \overline{\eta_{h,w_2}} + \overline{\eta_{h,p}}\right)} \tag{A1.19}$$

for $j = \{1,2\}$, and where $\overline{\eta_{h,w_j}} = s\eta_{h_1,w_j} + (1-s)\,\eta_{h_2,w_j}$, $\overline{\eta_{h,w_{-j}}} = s\eta_{h_1,w_{-j}} + (1-s)\,\eta_{h_2,w_{-j}}$, and $\overline{\eta_{h,p}} = s\eta_{h_1,p} + (1-s)\,\eta_{h_2,p}$.

Some Special Cases

Above we have derived a closed form solution for the transmission coefficients for consumption and hours in the case of linear taxes. Here we consider special cases to show analytically some of the mechanisms driving the transmission of shocks to consumption and hours which were discussed in the main text. First, we focus on the separable case without taxes to illustrate under which conditions labor supply is used as an insurance mechanism. Second, we use a simple single earner case with taxes to illustrate the potential bias that stems from ignoring taxes in estimation.

The Additive Separability/No Tax Case

Assume additive separable preferences ($\eta_{c,w_j} = \eta_{h_j,p} = 0$ for $j = \{1,2\}$ and $\eta_{h_1,w_2} = \eta_{h_2,w_1} = 0$) and no taxes ($\mu_{i,t} = \chi_{i,t} = 0$).[3] In this case, we get the following equations for consumption growth and for the growth of hours of the two earners:

$$\begin{pmatrix} \Delta c_{i,t} \\ \Delta h_{i,1,t} \\ \Delta h_{i,2,t} \end{pmatrix} \simeq \begin{pmatrix} 0 & 0 & \kappa^S_{c,v_1} & \kappa^S_{c,v_2} \\ \kappa_{h_1,u_1} & 0 & \kappa^S_{h_1,v_1} & \kappa^S_{h_1,v_2} \\ 0 & \kappa^S_{h_2,u_2} & \kappa^S_{h_2,v_1} & \kappa^S_{h_2,v_2} \end{pmatrix} \begin{pmatrix} \Delta u_{i,1,t} \\ \Delta u_{i,2,t} \\ v_{i,1,t} \\ v_{i,2,t} \end{pmatrix} \tag{A2.20}$$

where the transmission coefficients κ^S (the superscript S standing for "Separability") have now simpler analytical interpretations (maintaining the assumption of negligible wealth effect for transitory shocks, and omitting the i subscripts):

[3] A proportional tax rate ($\chi_{i,t} \neq 0, \mu_{i,t} = 0$) will deliver similar expressions for the κ's, but would imply a different path for the accumulation of human wealth, therefore a different $\pi_{i,t}$.

$$\kappa^S_{c,v_j} = \frac{\eta_{c,p}\left(1-\pi_t\right)s_{j,t}\left(1+\eta_{h_j,w_j}\right)}{\eta_{c,p}+\left(1-\pi_t\right)\overline{\eta_{h,w}}} \tag{A2.21}$$

$$\kappa^S_{h_j,u_j} = \eta_{h_j,w_j} \tag{A2.22}$$

$$\kappa^S_{h_j,v_j} = \eta_{h_j,w_j}\left(1-\frac{\left(1-\pi_t\right)s_{j,t}\left(1+\eta_{h_j,w_j}\right)}{\eta_{c,p}+\left(1-\pi_t\right)\overline{\eta_{h,w}}}\right) \tag{A2.23}$$

$$\kappa^S_{h_j,v_{-j}} = -\frac{\eta_{h_j,w_j}\left(1-\pi_t\right)\left(1-s_{j,t}\right)\left(1+\eta_{h_{-j},w_{-j}}\right)}{\eta_{c,p}+\left(1-\pi_t\right)\overline{\eta_{h,w}}}, \tag{A2.24}$$

with $\overline{\eta_{h,w}} = s_{j,t}\eta_{h_j,w_j} + \left(1-s_{j,t}\right)\eta_{h_{-j},w_{-j}}$.

We now show under what conditions labor supply is used as consumption smoothing device, i.e. $\kappa^S_{h_j,v_j} < 0$. Assume for simplicity that there is only one earner ($s_{j,t} = 1$). In this case, the condition that ensures that own labor supply is used as a consumption smoothing device is:

$$\left(1-\pi_t\right) - \eta_{c,p} > 0$$

which is more likely to hold as $\pi_t \to 0$ or as $\eta_{c,p} \to 0$.

In the more general case with multiple earners $\kappa^S_{h_j,v_j} < 0$ if:

$$\left(1-\pi_t\right)\left[s_{j,t}-\left(1-s_{j,t}\right)\eta_{h_{-j},w_{-j}}\right] - \eta_{c,p} > 0$$

For this case, as long as $\left[s_{j,t}-\left(1-s_{j,t}\right)\eta_{h-j,w_{-j}}\right] > 0$, labor supply of the primary earner is more likely to be used to smooth consumption if the secondary earner counts little in the balance of life time earnings ($(1-s_{i,t})$ is low) or if her labor supply is relatively inelastic (η_{h_2,w_2} is small).

Note that for this case, unambiguously $\kappa^S_{h_j,v_{-j}} < 0$.

The Additive Separability Case with Progressive Tax

We show how ignoring progressive taxation leads to biased estimates of the Frisch elasticities. Consider the response of hours to a transitory (or λ-constant) shock to the *before-tax* hourly wage. In the one-earner case one can use (A1.15) to show that:

$$\kappa^T_{h_j,u_j} = \frac{\eta_{h_j,w_j}\left(1-\mu\right)}{1+\mu\eta_{h_j,w_j}} \tag{A2.25}$$

where $\kappa^T_{h_j,u_j}$ is, as before, the elasticity of labor supply with respect to *before-tax* wage changes, $0 \leq \mu \leq 1$, and we use the superscript T to indicate the "Progressive Tax" model. Note that this parameter no longer coincides with the Frisch elasticity η_{h_j,w_j} (the curvature of the utility function with respect to leisure). In particular, when μ increases (i.e., when the tax system becomes more redistributive), the labor supply response to a before-tax wage change is dampened relative to the no-tax or flat-tax case, because any labor supply increase induced by an exogenous increase in before-tax wages is attenuated by a decrease in the return to work as people cross tax brackets. Note also that the Frisch elasticity that is estimated in the model that neglects progressive taxation is downward biased, as researchers attribute a low response of hours to wage changes to high tastes for leisure, while in fact it may reflect the disincentive to work induced by taxes.

Researchers interested in the effect of taxes on labor supply may want to distinguish between the elasticity of labor supply with respect to before-tax changes in wages (κ_{BT}) and the elasticity of labor supply with respect to after-tax changes in wages (κ_{AT}) (MaCurdy, 1983). The responses of hours and consumption to wage shocks captured by the κ's in equation (9) in the paper are equivalent to the former. However, we can also back out the latter for both the Frisch and the Marshallian case. It is straightforward to recover Frisch elasticities with respect to after-tax wage changes as they are simply the preference parameters we estimate (i.e., the η's). As for Marshallian elasticities, we use the preference parameters estimated in the progressive tax case to re-calculate the κ's with respect to tax neutral permanent shocks to wages ($\kappa_{c,v_j}, \kappa_{h_j,v_j}$). For the simple single-earner case, the responses to before- and after-tax wage changes are linked through the relationship $\frac{\kappa_{BT}}{1-\mu(1+\kappa_{BT})} = \kappa_{AT}$.[4]

In the consumption case, one can calculate that the response of consumption to a before-tax transitory wage shock is no longer η_{c,w_j} (which was identifying the extent of non-separability between consumption and leisure). In the single earner case, for example, it equals $\frac{\eta_{c,w_1}(1-\mu)}{1+\mu\eta_{h_1,w_1}}$, and hence it is also dampened (in absolute value). The reason is that this coefficient captures the extent of consumption co-movement with hours, but in the case with taxes hours move less and this lower sensitivity of hours to wage shocks spills over to a lower sensitivity of consumption to wage shocks induced by preference non-separability.

[4] To see the intuition for this expression, consider the relationship between after-tax (ω) and before-tax wage (W):

$$\omega_{i,j,t} = (1 - m_{i,j,t})\, W_{i,j,t}$$

where m is the marginal tax rate. In the one-earner case taxes paid are $M_{i,t} = Y_{i,t} - (1-\chi_t)\, Y_{i,t}^{1-\mu_t}$, from which we derive the marginal tax rate $m_{i,j,t} = \frac{\partial M_{i,j,t}}{\partial Y_{i,t}} = 1 - (1-\chi_t)(1-\mu_t)\, Y_{i,t}^{-\mu_t}$. Hence, the after-tax wage and the before-tax wage are linked by the relationship:

$$\log \omega_{i,j,t} = \log(1-\chi_t) + \log(1-\mu_t) - \mu_t \log H_{i,j,t} + (1-\mu_t)\log W_{i,j,t}$$

Taking total derivatives, we obtain:

$$\frac{\kappa_{BT}}{1-\mu(1+\kappa_{BT})} = \kappa_{AT}$$

where $\kappa_{BT} = d\log H_{i,j,t}/d\log W_{i,j,t}$ and $\kappa_{AT} = d\log H_{i,j,t}/d\log \omega_{i,j,t}$. By evaluating these two elasticities in λ-constant or λ-varying scenarios, one gets Frisch and Marshallian elasticities with respect to before-tax and after-tax wage changes, respectively. For example, in the Frisch case: $(d\log H_{i,j,t}/d\log W_{i,j,t})_{d\lambda=0} = \kappa_{h_j,u_j} = \frac{\eta_{h_j,w_j}(1-\mu)}{1+\mu\eta_{h_j,w_j}}$ and $(d\log H_{i,j,t}/d\log \omega_{i,j,t})_{d\lambda=0} = \eta_{h_j,w_j}$. For the two earners case, this calculation has to include the effect of changing one earner's hours on the other earner. Our approximation procedure takes this into account (see Appendix 1).

Appendix 2: Identification

Here we discuss how the parameters of interest are identified. We discuss the identification of measurement error parameters in Appendix 3. To keep the discussion tractable, we focus on the nonseparable case with proportional taxes (i.e., where the mapping between structural parameters and transmission coefficients is given by equations (A1.16)-(A1.19)). Identification arguments are similar in the case with taxes, but the notation is much more cumbersome. Furthermore, while we present the identification arguments in one-year log-differences, the biennial nature of the PSID requires that the actual moment conditions are written as a two-year difference. The change to two-year differences is straightforward.

We estimate 6 wage parameters $(\sigma^2_{v_1}, \sigma^2_{v_2}, \sigma^2_{u_1}, \sigma^2_{u_2}, \sigma_{v_1,v_2}, \sigma_{u_1,u_2})$ for five age groups (a total of 30 parameters) and 10 preference and insurance parameters $(\eta_{c,p}, \eta_{h_1,w_1}, \eta_{h_2,w_2}, \eta_{h_1,w_2}, \eta_{h_2,w_1}, \eta_{c,w_1}, \eta_{c,w_2}, \eta_{h_1,p}, \eta_{h_2,p}, \beta)$.[5] It is worth stressing that while we provide here formal arguments for the identification of these parameters, our system is heavily over-identified, implying that other moments (besides the ones mentioned below) contribute to identification.

Identification of the wage parameters is similar to Meghir and Pistaferri (2004). Consider the stochastic component of wage growth of earner j:

$$\Delta w_{j,t} = \Delta u_{j,t} + v_{j,t}$$

It can be shown that the wage parameters are identified using:

$$\begin{aligned}
\sigma^2_{u_j} &= -E\left(\Delta w_{j,t}\Delta w_{j,t+1}\right) \\
\sigma^2_{v_j} &= E\left(\Delta w_{j,t}\left(\Delta w_{j,t+1} + \Delta w_{j,t} + \Delta w_{j,t-1}\right)\right) \\
\sigma_{u_1,u_2} &= -E\left(\Delta w_{1,t}\Delta w_{2,t+1}\right) \\
\sigma_{v_1,v_2} &= E\left(\Delta w_{1,t}\left(\Delta w_{2,t+1} + \Delta w_{2,t} + \Delta w_{2,t-1}\right)\right)
\end{aligned}$$

Identification of $\sigma^2_{u_j}$ rests on the idea that wage growth rates are autocorrelated due to mean reversion caused by the transitory component (the permanent component is subject to i.i.d. shocks). Identification of $\sigma^2_{u_1 u_2}$ is an extension of this idea - between-period and between-earner wage growth correlation reflects the correlation of the mean-reverting components. Identification of $\sigma^2_{v_j}$ rests on the idea that the variance of wage growth $(E\left(\Delta w_{j,t}\Delta w_{j,t}\right))$, subtracted the contribution of the mean reverting component $(E\left(\Delta w_{j,t}\Delta w_{j,t-1}\right) + E\left(\Delta w_{j,t}\Delta w_{j,t+1}\right))$, identifies the variance of innovations to the permanent component. Identification of $\sigma_{v_j v_{-j}}$ follows a similar logic.

Next we discuss the identification of hours elasticity parameters (own and cross-elasticities). Consider

[5] π and s are backed out directly from data so we do not discuss them here.

the following moments:

$$\begin{aligned} m_1 &= E\left(\Delta w_{1,t} \Delta y_{1,t+1}\right) = -\left(1+\eta_{h_1,w_1}\right)\sigma^2_{u_1} - \eta_{h_1,w_2}\sigma_{u_1,u_2} \\ m_2 &= E\left(\Delta w_{2,t} \Delta y_{1,t+1}\right) = -\left(1+\eta_{h_1,w_1}\right)\sigma_{u_1,u_2} - \eta_{h_1,w_2}\sigma^2_{u_2} \\ m_3 &= E\left(\Delta w_{1,t} \Delta w_{1,t+1}\right) = -\sigma^2_{u_1} \\ m_4 &= E\left(\Delta w_{2,t} \Delta w_{2,t+1}\right) = -\sigma^2_{u_2} \\ m_5 &= E\left(\Delta w_{2,t} \Delta w_{1,t+1}\right) = -\sigma_{u_1,u_2} \end{aligned}$$

Combination of these moments gives

$$\begin{aligned} \eta_{h_1,w_1} &= \frac{m_1 m_4 - m_2 m_5}{m_3 m_4 - m_5^2} - 1 \\ \eta_{h_1,w_2} &= \frac{m_2 m_3 - m_1 m_5}{m_3 m_4 - m_5^2} \end{aligned}$$

which shows identification of the parameters η_{h_1,w_1} and η_{h_1,w_2}. The identification of the parameters η_{h_2,w_2} and η_{h_2,w_1} is symmetric and hence omitted.

The idea behind the identification of η_{c,w_1} and η_{c,w_2} is very similar. In particular, consider the following two moments:

$$\begin{aligned} m_6 &= E\left(\Delta w_{1,t} \Delta c_{i,t+1}\right) = -\eta_{c,w_1}\sigma^2_{u_1} - \eta_{c,w_2}\sigma_{u_1,u_2} \\ m_7 &= E\left(\Delta w_{2,t} \Delta c_{i,t+1}\right) = -\eta_{c,w_1}\sigma_{u_1,u_2} - \eta_{c,w_2}\sigma^2_{u_2} \end{aligned}$$

We can now show that:

$$\begin{aligned} \eta_{c,w_1} &= \frac{m_6 m_4 - m_7 m_5}{m_3 m_4 - m_5^2} \\ \eta_{c,w_2} &= \frac{m_7 m_3 - m_6 m_5}{m_3 m_4 - m_5^2} \end{aligned}$$

While the covariance between consumption growth and lagged wage growth plays an important role in identifying the extent of non-separability between consumption and hours, the model is over-identified. Hence there are other moments (besides m_6 and m_7) playing a role in identifying these parameters, and more so because the covariance between consumption growth and lagged wage growth is rather noisy in the data (see Table 1). For example, $E\left(y_{2,t}^2\right)$ plays an important role in the identification of η_{c,w_1}. To see why, suppose that consumption and labor supply are Frisch substitutes, i.e. $\eta_{c,w_1} < 0$. When η_{c,w_1} becomes higher (less negative), the household is willing to tolerate larger responses of consumption to permanent shocks to wages (because of a shift towards consumption-hours complementarity, or less substitutability). This implies that there is lower demand for insurance through added worker effect within the household (i.e., less responsiveness of the wife to permanent shocks faced by the husband). The outcome is that increasing η_{c,w_1} results in a decrease in the volatility of labor supply and earnings of the wife, i.e., a smaller $E\left(y_{2,t}^2\right)$. Negative η_{c,w_1} therefore helps matching the high volatility of wife's earnings observed in the data. There is a symmetric argument linking η_{c,w_2} to $E\left(y_{1,t}^2\right)$.

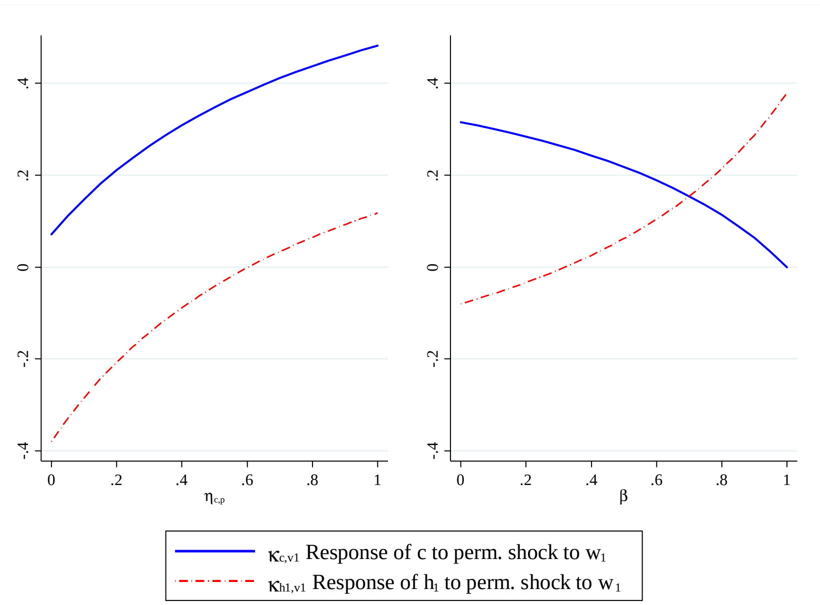

Figure 1: Identification of $\eta_{c,p}$ and β

Given that η_{c,w_1} and η_{c,w_2} are identified, the symmetry of the Frisch substitution matrix can be applied to derive estimates for $\eta_{h_1,p}$, $\eta_{h_2,p}$ (see Appendix 5 below).

Finally, while the identification of $\eta_{c,p}$ is discussed in the text, we discuss here the joint identification of $\eta_{c,p}$ and β for the specification that is estimated in column 5 of Table 4. These parameters are identified separately through the different effect that they have on the responsiveness of consumption and earnings to permanent shocks. When $\eta_{c,p}$ is high, individuals are more tolerant of (or less averse to) intertemporal fluctuations in their consumption induced by permanent wage shocks. When β is high, individuals have more outside insurance, so can smooth consumption more easily, implying that a high β has a similar effect on the volatility of consumption as a low $\eta_{c,p}$. Consider now the response of hours to permanent wage shocks. As we discuss in the main text, the sign of the relationship depends on whether substitution effect dominates the wealth effect, or vice versa. The wealth effect is more likely to dominate (and hence hours move in the opposite direction of a wage change, implying that it is used as a consumption smoothing device) when $\eta_{c,p}$ is low (consumers prefer cutting leisure rather than goods when their wages decline), or when consumers have no access to insurance (low β), implying that labor supply is the only insurance device available. This intuition is illustrated in Figure 1.

For the separable case, it is easy to show that the assumption of positive own Frisch elasticities ($\eta_{h_1,w_1} >$

$0, \eta_{h_2,w_2} > 0$ and $\eta_{c,p} > 0$) is sufficient for the following sign restrictions:[6]

$$\frac{\partial \kappa_{c,v_j}}{\partial \eta_{c,p}} > 0 \quad \text{(A2.1)}$$

$$\frac{\partial \kappa_{y_j,v_j}}{\partial \eta_{c,p}} > 0 \quad \text{(A2.2)}$$

$$\frac{\partial \kappa_{c,v_j}}{\partial \pi} < 0 \quad \text{(A2.3)}$$

$$\frac{\partial \kappa_{y_j,v_j}}{\partial \pi} > 0 \quad \text{(A2.4)}$$

Appendix 3: Measurement Error

We rewrite the equations for wage growth, consumption growth and earnings growth to allow for measurement errors as:[7]

$$\begin{pmatrix} \Delta w_{i,1,t} \\ \Delta w_{i,2,t} \\ \Delta c_{i,t} \\ \Delta y_{i,1,t} \\ \Delta y_{i,2,t} \end{pmatrix} \simeq \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ \kappa_{c,u_1} & \kappa_{c,u_2} & \kappa_{c,v_1} & \kappa_{c,v_2} \\ \kappa_{y_1,u_1} & \kappa_{y_1,u_2} & \kappa_{y_1,v_1} & \kappa_{y_1,v_2} \\ \kappa_{y_2,u_1} & \kappa_{y_2,u_2} & \kappa_{y_2,v_1} & \kappa_{y_2,v_2} \end{pmatrix} \begin{pmatrix} \Delta u_{i,1,t} \\ \Delta u_{i,2,t} \\ v_{i,1,t} \\ v_{i,2,t} \end{pmatrix} + \begin{pmatrix} \Delta \xi^w_{i,1,t} \\ \Delta \xi^w_{i,2,t} \\ \Delta \xi^c_{i,t} \\ \Delta \xi^y_{i,1,t} \\ \Delta \xi^y_{i,2,t} \end{pmatrix}$$

where $\xi^w_{i,j,t}$, $\xi^c_{i,t}$ and $\xi^y_{i,j,t}$ are measurement errors in log wages of earner j, log consumption, and log earnings of earner j.[8]

Ignoring the variance of measurement error in wages or earnings is problematic since it has a direct effect on the estimates of the structural parameters. We thus follow Meghir and Pistaferri (2004) and use findings from validation studies to set *a priori* the amount of wage variability that can be attributed to error. We use the estimates of Bound et al. (1994), who estimate the share of variance associated with measurement error using a validation study for the PSID (which is the data set we are using). While the validation study they use covers only a small fraction of the PSID sample, they extrapolate their findings to estimate the share of measurement errors in representative samples. We adopt their estimates for the share of measurement error in log earnings ($var(\xi^y) = 0.04 var(y)$). For log hourly wages, their estimates range from 0.072 to 0.162. We use an estimate in the middle of this range ($var(\xi^w) = 0.13 var(w)$). Finally, for log hours they report $var\left(\xi^h\right) = 0.23 var(h)$. Note that these estimates can be used to correct all "own" moments

[6] While it is clear that the volatility of consumption and earnings (and hence the transmission coefficients) respond differently to $\eta_{c,p}$ and to β also in the non-separable case, more restrictions on the parameters are required for the exact relations in A2.1 to A2.4 to hold.

[7] Since we use earnings data, we discuss the measurement error in the context of the system that uses earnings instead of hours.

[8] Formally, we should write $\Delta\widetilde{x} = \Delta x + \Delta\xi^x$ for $x = w, c, y$, with $\widetilde{x}$ and x being the observed and true value of x, respectively, but this would just make notation harder to follow, and so we omit it.

(such as $E\left((\Delta y_{i,j,t})^2\right)$, $E\left(\Delta y_{i,j,t}\Delta y_{i,j,t+1}\right)$, etc.) with the only assumption (not entirely uncontroversial, see Bound and Krueger, 1991) that measurement error is not correlated over time. However, in order to use these estimates to correct cross equations moments (such as $E\left(\Delta w_{i,j,t}\Delta y_{i,j,t}\right)$), we need to calculate the covariance between the various measurement errors.[9] By definition this covariance is non-zero, since in our data set wage=earnings/hours. We can write the relationship between errors for log earnings, hours and wages as

$$var\left(\xi^w\right) = var\left(\xi^y\right) + var\left(\xi^h\right) - 2cov(\xi^y, \xi^h)$$

and given the share of measurement error for earnings, hours and wages, we can back out the covariance between measurement error of wages and measurement error in earnings as:

$$cov(\xi^y, \xi^w) = cov(\xi^y, \left(\xi^y - \xi^h\right)) = var\left(\xi^y\right) - \frac{1}{2}\left[var\left(\xi^y\right) + var\left(\xi^h\right) - var\left(\xi^w\right)\right]$$

Finally, for separable utility, log consumption is a martingale. Hence, the variance of the measurement error in consumption is directly identified from the moment $E\left(\Delta c_{i,t}\Delta c_{i,t+1}\right) = -var\left(\xi^c\right)$. We keep this exact identification also for the non-separable case, where a sufficient condition for $-E\left(\Delta c_{i,t}\Delta c_{i,t+1}\right)$ to be an upper bound on the measurement error is that $sign\left(\kappa_{c,u_1}\right) = sign\left(\kappa_{c,u_2}\right)$. In the absence of further (and stronger) assumptions, we will make no attempt to distinguish measurement errors in consumption from stochastic changes in preferences or shocks to higher moments of the distribution of wages.

Appendix 4: Selection Into Work by the Second Earner

To illustrate the strategy we adopt, suppose that the participation decision $P_{i,2,t} = \{0, 1\}$ of the secondary earner depends on some latent variable $I^*_{i,2,t}$, which can be written as

$$\begin{aligned} I^*_{i,2,t} &= \mathbf{m}'_{i,t}\varsigma + \tau_{i,t} \\ P_{i,2,t} &= \begin{cases} 1 & \text{if } I^*_{i,t} \geq 0 \\ 0 & \text{otherwise} \end{cases} \end{aligned}$$

where $I^*_{i,2,t}$ is a latent variable, $P_{i,2,t}$ the observed choice, and $\mathbf{m}_{i,t}$ a vector of observed characteristics. $P_{i,2,t} = 1$ for participation (employment) and zero otherwise. Assuming that the shocks are normally distributed $\Pr\left(P_{i,2,t} = 1\right) = \Phi\left(m'_{i,t}\varsigma\right)$ and assuming

$$\begin{aligned} \mathbb{E}\left(u_{i,j,t}\tau_{i,s}\right) &= \begin{cases} \sigma_{u_j,\tau} & \text{if } s = t,\ j = \{1,2\} \\ 0 & \text{otherwise} \end{cases} \\ \mathbb{E}\left(v_{i,j,t}\tau_{i,s}\right) &= \begin{cases} \sigma_{v_j,\tau} & \text{if } s = t,\ j = \{1,2\} \\ 0 & \text{otherwise} \end{cases} \end{aligned}$$

[9] Assuming that measurement error is not correlated across earners, the only four relevant cross moments that need to be corrected are $\mathbb{E}\left[\Delta w_{i,1,t}\Delta y_{i,1,t}\right]$, $\mathbb{E}\left[\Delta w_{i,2,t}\Delta y_{i,2,t}\right]$, $\mathbb{E}\left[\Delta w_{i,1,t}\Delta y_{i,1,t-1}\right]$, $\mathbb{E}\left[\Delta w_{i,1,t-1}\Delta y_{i,1,t}\right]$, $\mathbb{E}\left[\Delta w_{i,2,t}\Delta y_{i,2,t-1}\right]$ and $\mathbb{E}\left[\Delta w_{i,2,t-1}\Delta y_{i,2,t}\right]$.

and the variance of τ is normalized to be 1. Assuming the four correlations $\rho_{v_1,\tau}, \rho_{v_2,\tau}, \rho_{u_1,\tau}$ and $\rho_{u_2,\tau}$ can be identified, the identification strategy presented in Appendix 3 remains unchanged. However, the moment conditions are corrected for sample selection using formulae for the multinomial truncated normal case (see Tallis, 1961). To see how this correction works, consider for example the product of second earner's wage growth with lagged wage growth:[10]

$$\mathbb{E}\left(\Delta w_{i,2,t}\Delta w_{i,2,t-1}|P_{i,2,t}=P_{i,2,t-1}=1\right)=-\sigma_{u_2}^2\left(1-\rho_{u_2,\tau}^2\Lambda_{2,t-1}\mathbf{m}_{2,t-1}'\varsigma\right)$$

where $\Lambda_{i,t}=\frac{\phi\left(m_{i,t}'\varsigma\right)}{\Phi\left(m_{i,t}'\varsigma\right)}$. The other moments can be corrected in a similar fashion.

Appendix 5: Quasi-Concavity of the Utility Function

Following Phlips (1974) and others, one can show that without taxes, the "fundamental" substitution effects matrix (evaluated at $d\lambda=0$ - i.e., considering λ-constant demand functions) is:

$$\begin{aligned}\begin{pmatrix}\frac{dc}{dp} & \frac{dc}{dw_1} & \frac{dc}{dw_2}\\ \frac{dl_1}{dp} & \frac{dl_1}{dw_1} & \frac{dl_1}{dw_2}\\ \frac{dl_2}{dp} & \frac{dl_2}{dw_1} & \frac{dl_2}{dw_2}\end{pmatrix} &= \lambda\begin{pmatrix}\frac{d^2u}{dc^2} & \frac{d^2u}{dcdl_1} & \frac{d^2u}{dcdl_2}\\ \frac{d^2u}{dl_1dc} & \frac{d^2u}{dl_1^2} & \frac{d^2u}{dl_1dl_2}\\ \frac{d^2u}{dl_2dc} & \frac{d^2u}{dl_2dl_1} & \frac{d^2u}{dl_2^2}\end{pmatrix}^{-1}\\ \mathbf{A} &= \lambda\mathbf{U}^{-1}\end{aligned} \quad (A5.1)$$

where all subscripts have been omitted for simplicity. In terms of the consumption/hours elasticities (and noting that $\eta_{lx}=-\eta_{hx}\frac{h}{l}$ for a generic price x), the matrix of behavioral responses can be written as:

$$\begin{pmatrix}\frac{dc}{dp} & \frac{dc}{dw_1} & \frac{dc}{dw_2}\\ \frac{dl_1}{dp} & \frac{dl_1}{dw_1} & \frac{dl_1}{dw_2}\\ \frac{dl_2}{dp} & \frac{dl_2}{dw_1} & \frac{dl_2}{dw_2}\end{pmatrix}=\begin{pmatrix}\eta_{cp}\frac{c}{p} & \eta_{cw_1}\frac{c}{w_1} & \eta_{cw_2}\frac{c}{w_2}\\ -\eta_{h_1p}\frac{h_1}{p} & -\eta_{h_1w_1}\frac{h_1}{w_1} & -\eta_{h_1w_2}\frac{h_1}{w_2}\\ -\eta_{h_2p}\frac{h_2}{p} & -\eta_{h_2w_1}\frac{h_2}{w_1} & -\eta_{h_2w_2}\frac{h_2}{w_2}\end{pmatrix}$$

We impose symmetry of elasticities in the following sense: $\eta_{h_jp}=-\eta_{cw_j}\frac{pc}{w_jh_j}$ $(j=1,2)$ and $\eta_{h_2w_1}=\eta_{h_1w_2}\frac{w_1h_1}{w_2h_2}$. In estimation we replace w_1h_1, w_2h_2 and pc with the mean of the two earners' earnings and with the mean of consumption respectively. Note that for the case with progressive taxation, wages need to be adjusted by $\frac{1}{(1-\chi)(1-\mu)Y^{-\mu}}$, so that: $\eta_{h_jp}=-\eta_{c,w_1}\frac{1}{(1-\chi)(1-\mu)Y^{-\mu}}\frac{pc}{w_jh_j}$ $(j=1,2)$ and $\frac{1}{(1-\chi)(1-\mu)Y^{-\mu}}$ is replaced by its sample mean.

To next show that our empirical estimates imply quasi-concavity of preferences, it suffices to show that the Hessian of the utility function is negative semi-definite (n.s.d. henceforth). Since the inverse of an n.s.d. matrix is n.s.d., it is sufficient to show that the $\mathbf{A}$ is n.s.d. ($\lambda\geq 0$ by definition, so it plays no role in this). We perform this exercise by considering a value of $\mathbf{A}$ obtained using our estimates of the various elasticities, and evaluating the consumption, hours, and wages terms at their sample medians (we normalize $p=1$).

[10] This strategy ignores the fact that since the truncated shocks are no longer mean-zero, the expectation of the product of any of these shocks is not zero as well. We assume that the expectation of these products can be neglected.

We find that matrix A is estimated to be:

$$\widehat{\mathbf{A}} = 10^4 \times \begin{pmatrix} -1.637 & -0.022 & -0.010 \\ -0.029 & -0.005 & -0.002 \\ -0.013 & -0.002 & -0.008 \end{pmatrix}$$

with eigenvalues given by: $10^4 \times \begin{pmatrix} -1.637 & -0.009 & -0.004 \end{pmatrix}'$. It follows that concavity of the utility function holds empirically. The signs on the Hessian are as expected (note that utility is written in terms of leisure and not hours, and therefore the cross derivatives of consumption are positive):

$$\widehat{\mathbf{H}} = \begin{pmatrix} -0.0001 & 0.0003 & 0.0000 \\ 0.0003 & -0.0231 & 0.0045 \\ 0.0000 & 0.0045 & -0.0131 \end{pmatrix}$$

References

[1] Blundell, Richard, Hamish Low and Ian Preston (2013). “Decomposing Changes in Income Risk Using Consumption Data”, Quantitative Economics, 4, 1–37, March.

[2] Bound John, Charles Brown, Greg J. Duncan, and Willard L. Rodgers (1994), “Evidence on the Validity of Cross-Sectional and Longitudinal Labor Market Data.” *Journal of Labor Economics*, 12(3), 345-368.

[3] Bound John, and Alan B. Krueger (1991). "The Extent of Measurement Error in Longitudinal Earnings Data: Do Two Wrongs Make a Right?", Journal of Labor Economics, 9(1), 1-24

[4] Phlips, Louis (1974), Applied consumption analysis, Amsterdam: North Holland.

[5] Tallis, G.M. (1961). “The moment generating function of the truncated multi-normal distribution.” Journal of the Royal Statistical Society. Series B (Methodological), 223–229.

Table 1: **Moment values in the data and in the model**

Moment	Model	Data	S.E. Data	t-stat Data
$E\left(\Delta w_{1,t}^2\right)$	0.2444	0.2432	0.0099	24.57
$E\left(\Delta w_{2,t}^2\right)$	0.2050	0.2044	0.0074	27.62
$E\left(\Delta w_{1,t}\Delta w_{2,t}\right)$	0.0184	0.0169	0.0035	4.83
$E\left(\Delta w_{1,t}\Delta w_{1,t-2}\right)$	-0.0918	-0.0913	0.0060	15.22
$E\left(\Delta w_{2,t}\Delta w_{2,t-2}\right)$	-0.0638	-0.0640	0.0052	12.31
$E\left(\Delta w_{1,t}\Delta w_{2,t-2}\right)$	-0.0059	-0.0027	0.0036	0.75
$E\left(\Delta w_{2,t}\Delta w_{1,t-2}\right)$	-0.0115	-0.0089	0.0035	2.54
$E\left(\Delta c_t^2\right)$	0.0774	0.0919	0.0022	41.77
$E\left(\Delta y_{1,t}^2\right)$	0.2415	0.2481	0.0091	27.26
$E\left(\Delta y_{2,t}^2\right)$	0.3408	0.3468	0.0169	20.52
$E\left(\Delta w_{1,t}\Delta c_t\right)$	0.0123	0.0017	0.0020	0.85
$E\left(\Delta w_{2,t}\Delta c_t\right)$	0.0135	0.0031	0.0018	1.72
$E\left(\Delta y_{1,t}\Delta c_t\right)$	0.0018	0.0065	0.0020	3.25
$E\left(\Delta y_{2,t}\Delta c_t\right)$	0.0005	0.0050	0.0025	2.00
$E\left(\Delta w_{1,t}\Delta y_{1,t}\right)$	0.2019	0.1865	0.0073	25.55
$E\left(\Delta w_{1,t}\Delta y_{2,t}\right)$	-0.0033	0.0020	0.0041	0.49
$E\left(\Delta w_{2,t}\Delta y_{1,t}\right)$	0.0109	0.0147	0.0032	4.59
$E\left(\Delta w_{2,t}\Delta y_{2,t}\right)$	0.1468	0.1312	0.0073	17.97
$E\left(\Delta y_{1,t}\Delta y_{2,t}\right)$	0.0025	0.0028	0.0042	0.67
$E\left(\Delta c_t\Delta c_{t-2}\right)$	-0.0338	-0.0332	0.0017	19.53
$E\left(\Delta w_{1,t}\Delta c_{t-2}\right)$	0.0037	0.0022	0.0022	1.00
$E\left(\Delta w_{2,t}\Delta c_{t-2}\right)$	0.0013	0.0011	0.0021	0.52
$E\left(\Delta c_t\Delta w_{1,t-2}\right)$	0.0037	0.0004	0.0022	0.18
$E\left(\Delta c_t\Delta w_{2,t-2}\right)$	0.0013	0.0010	0.0021	0.48
$E\left(\Delta y_{1,t}\Delta c_{t-2}\right)$	0.0061	0.0015	0.0022	0.68
$E\left(\Delta y_{2,t}\Delta c_{t-2}\right)$	0.0031	0.0002	0.0026	0.08
$E\left(\Delta c_t\Delta y_{1,t-2}\right)$	0.0061	-0.0044	0.0021	2.10
$E\left(\Delta c_t\Delta y_{2,t-2}\right)$	0.0031	0.0013	0.0025	0.52
$E\left(\Delta y_{1,t}\Delta y_{1,t-2}\right)$	-0.0907	-0.0830	0.0062	13.39
$E\left(\Delta y_{2,t}\Delta y_{2,t-2}\right)$	-0.0807	-0.0669	0.0078	8.58
$E\left(\Delta w_{1,t}\Delta y_{1,t-2}\right)$	-0.0719	-0.0657	0.0058	11.33
$E\left(\Delta w_{2,t}\Delta y_{2,t-2}\right)$	-0.0215	-0.0300	0.0055	5.45
$E\left(\Delta y_{1,t}\Delta w_{1,t-2}\right)$	-0.0719	-0.0709	0.0057	12.44
$E\left(\Delta y_{2,t}\Delta w_{2,t-2}\right)$	-0.0215	-0.0337	0.0055	6.13
$E\left(\Delta w_{1,t}\Delta y_{2,t-2}\right)$	-0.0158	0.0075	0.0044	1.70
$E\left(\Delta w_{2,t}\Delta y_{1,t-2}\right)$	-0.0106	-0.0080	0.0034	2.35
$E\left(\Delta y_{2,t}\Delta w_{1,t-2}\right)$	-0.0158	-0.0122	0.0039	3.13
$E\left(\Delta y_{1,t}\Delta w_{2,t-2}\right)$	-0.0106	-0.0061	0.0038	1.61
$E\left(\Delta h_{1,t}^2\right)$	0.0822	0.1315	0.0079	16.65
$E\left(\Delta h_{2,t}^2\right)$	0.2524	0.3517	0.0189	18.61
$E\left(\Delta h_{1,t}\Delta h_{2,t}\right)$	0.0133	0.0023	0.0027	0.85
$E\left(\Delta h_{1,t}\Delta h_{1,t-2}\right)$	-0.0387	-0.0503	0.0058	8.67
$E\left(\Delta h_{2,t}\Delta h_{2,t-2}\right)$	-0.1015	-0.1152	0.0140	8.23