1 Baseline Model

1.1 The Income Process

For each education group and birth cohort, we write log-income net of predictable components[1]of individual i at age $a \in [a_{\min}, ..., a_{\max}]$ and as

$$y_{i,a} = \alpha_i + v_{i,a} + \tau_{i,a}, \tag{1}$$

Individual-specific differences in levels are given by the mean-zero parameter α_i with variance $var(\alpha_i)$. Transitory income $\tau_{i,a}$ is assumed to follow a mean-reverting MA(1) process,

$$\tau_{i,a} = \varepsilon_{i,a} + \theta\varepsilon_{i,a-1},$$

while the permanent component $v_{i,a}$ is assumed to follow an AR(1) process,

$$v_{i,a} = \rho v_{i,a-1} + u_{i,a},$$

with persistence parameter ρ. The shocks to income $\varepsilon_{i,a}$ and $u_{i,a}$ are assumed to be independent of each other, serially uncorrelated, independent of α_i, with zero-mean and variances ω_a^2 and σ_a^2 respectively. The quasi-difference $\Delta^\rho y_{i,a} \equiv y_{i,a} - \rho y_{i,a-1}$ can be written as

$$\Delta^\rho y_{i,a} = \alpha_i(1-\rho) + u_{i,a} + \Delta^\rho\varepsilon_{i,a} + \theta\Delta^\rho\varepsilon_{i,a-1}, \quad a = a_{\min}+1, ..., a_{\max}, \tag{2}$$

The autocovariance is given by

$$cov(\Delta^\rho y_{i,a}, \Delta^\rho y_{i,a+s}) = (1-\rho)^2\, var(\alpha_i) \begin{cases} +\sigma_a^2 + \omega_a^2 + (\theta-\rho)^2\,\omega_{a-1}^2 + \theta^2\rho^2\omega_{a-2}^2 & \text{if } s=0 \\ +(\theta-\rho)\left(\omega_a^2 - \theta\rho\omega_{a-1}^2\right) & \text{if } s=1 \\ -\theta\rho\omega_a^2 & \text{if } s=2 \\ +0 & \text{if } s>2 \end{cases}$$

1st order autocorrelation coefficients. The 1st order autocorrelation coefficient for the income process (1) at age a is given by

$$\hat{\rho}_{1,a} = \frac{cov(y_{i,a}, y_{i,a+1})}{\sqrt{var(y_{i,a})}\sqrt{var(y_{i,a+1})}}.$$

The autocovariance is given by

$$\begin{aligned} cov(y_{i,a}, y_{i,a+1}) &= cov(\alpha_i + v_{i,a} + \tau_{i,a}, \alpha_i + \rho v_{i,a} + u_{i,a+1} + \tau_{i,a+1}) \\ &= var(\alpha_i) + \rho var(v_{i,a}) + cov(\tau_{i,a}, \tau_{i,a+1}) \end{aligned}$$

[1]Observable income characteristics including a second polynomial in age, dummies for education, region, family size and marital status, the interaction of family size and marital status and the interaction of the polynomial in age and the education dummy. Family income is given by total household income (NOTE: I removed the equivalence scale adjustment, because it did not matter for our results in the previous version and to be consistent with the analysis of Magne's β in case we include it)

and note that with $var(y_{i,a}) \backsimeq var(y_{i,a+1})$ the first order autocorrelation coefficient becomes

$$\hat{\rho}_{1,a} \simeq \frac{var(\alpha_i) + \rho var(v_{i,a}) + cov(\tau_{i,a}, \tau_{i,a+1})}{var(\alpha_i) + var(v_{i,a}) + var(\tau_{i,a})}.$$

Since

$$\begin{aligned} cov(\tau_{i,a}, \tau_{i,a+1}) &= \theta var(\varepsilon_{i,a}) \\ var(\tau_{i,a}) &= var(\varepsilon_{i,a}) + \theta^2 var(\varepsilon_{i,a-1}) \\ var(v_{i,a}) &= \Sigma_{s=0}^{a} \rho^{2s} var(u_{i,a-s}) \end{aligned}$$

the first order autocorrelation coefficient can also be written as

$$\hat{\rho}_{1,a} \simeq \frac{var(\alpha_i) + \rho \Sigma_{s=0}^{a} \rho^{2s} var(u_{i,a-s}) + \theta var(\varepsilon_{i,a})}{var(\alpha_i) + \Sigma_{s=0}^{a} \rho^{2s} var(u_{i,a-s}) + var(\varepsilon_{i,a}) + \theta^2 var(\varepsilon_{i,a-1})}.$$

1.2 Estimation

For each value of ρ on a grid we calculate the $A \times A$ autocovariance matrix for the male member of the household for each cohort and each education group in our baseline sample, where $A \equiv a_{\max} - a_{\min}$, with $a_{\min}$and $a_{\max}$the minimum and maximum age of individuals in the sample. For a given age a and lag s, we average $cov(\Delta^{\rho} y_{i,a}, \Delta^{\rho} y_{i,a+s})$ across all birth cohorts and thereby integrate out calendar time effects. Let the resulting "average" autocovariance matrix be denoted by $\overline{C}$ and define the stacked vector[2] of its unique elements by

$$\hat{M} = vech\left(\overline{C}\right).$$

Let $\Theta = \{\sigma_a^2, \omega_a^2, var(\alpha_i), \theta\}$ denote the parameters[3] to be estimated for a given ρ. The equally-weighted minimum distance estimator $\hat{\Theta}$ is then the solution to

$$\hat{\Theta} = \arg\min_{\Theta} \left[M(\Theta;\rho) - \hat{M} \right]' \left[M(\Theta;\rho) - \hat{M} \right]$$

where $M(\Theta;\rho)$ are the corresponding stacked vector of theoretical moments given by

$$M(\Theta;\rho) = \begin{bmatrix} (1-\rho)^2 var(\alpha_i) + \sigma_{26}^2 + \omega_{26}^2 + (\theta-\rho)^2 \omega_{25}^2 + \theta^2 \rho^2 \omega_{24}^2 \\ \vdots \\ (1-\rho)^2 var(\alpha_i) + \sigma_{60}^2 + \omega_{60}^2 + (\theta-\rho)^2 \omega_{59}^2 + \theta^2 \rho^2 \omega_{58}^2 \\ (1-\rho)^2 var(\alpha_i) + (\theta-\rho)\left(\omega_{26}^2 - \theta\rho\omega_{25}^2\right) \\ \vdots \\ (1-\rho)^2 var(\alpha_i) + (\theta-\rho)\left(\omega_{59}^2 - \theta\rho\omega_{58}^2\right) \\ (1-\rho)^2 var(\alpha_i) - \theta\rho\omega_{26}^2 \\ \vdots \\ (1-\rho)^2 var(\alpha_i) - \theta\rho\omega_{58}^2 \\ (1-\rho)^2 var(\alpha_i) \\ \vdots \\ (1-\rho)^2 var(\alpha_i) \end{bmatrix}.$$

[2]Note that $\hat{M}$ is a $\frac{(A+1)A}{2} \times 1$ vector.
[3]$\sigma_{i,a}^2$ is a $A \times 1$ vectors and $\omega_{i,a}^2$ is a $(A+2) \times 1$ vector.

Thus the objective function can be written as

$$f(\Theta;\rho) = \frac{1}{2}\left[\ M(\Theta;\rho) - \ \hat{M}\ \right]'\left[\ M(\Theta;\rho) - \ \hat{M}\ \right].$$

To estimate the parameters we employ a Gauss-Newton algorithm. The gradient of the objective function reads

$$\nabla f(\Theta;\rho) = J(\Theta;\rho)'\left[\ M(\Theta;\rho) - \ \hat{M}\ \right],$$

where the Jacobian[4] can be represented as

$$J(\Theta;\rho) = \begin{bmatrix} J_{11} & J_{12} & J_{13} \\ J_{21} & J_{22} & J_{23} \\ J_{31} & J_{32} & J_{33} \\ J_{41} & J_{42} & J_{43} \end{bmatrix}$$

[4]The Jacobian is of dimension $\frac{(A+1)A}{2} \times (A + (A+2) + 2)$

with

$$\underset{A\times A}{J_{11}} = \begin{bmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{bmatrix}$$

$$\underset{A\times(A+2)}{J_{12}} = \begin{bmatrix} \theta^2\rho^2 & (\theta-\rho)^2 & 1 & 0 & 0 \\ 0 & \ddots & \ddots & \ddots & 0 \\ 0 & 0 & \theta^2\rho^2 & (\theta-\rho)^2 & 1 \end{bmatrix}$$

$$\underset{A\times 2}{J_{13}} = \begin{bmatrix} (1-\rho)^2 & 2(\theta-\rho)\,\omega^2_{a_{\min}} + 2\theta\rho^2\omega^2_{a_{\min}-1} \\ \vdots & \vdots \\ (1-\rho)^2 & 2(\theta-\rho)\,\omega^2_{a_{\max}-1} + 2\theta\rho^2\omega^2_{a_{\max}-2} \end{bmatrix}$$

$$\underset{(A-1)\times A}{J_{21}} = \mathbf{0}$$

$$\underset{(A-1)\times(A+2)}{J_{22}} = \begin{bmatrix} 0 & -(\theta-\rho)\,\theta\rho & (\theta-\rho) & 0 & 0 & 0 \\ 0 & \ddots & \ddots & \ddots & 0 & 0 \\ 0 & 0 & 0 & -(\theta-\rho)\,\theta\rho & (\theta-\rho) & 0 \end{bmatrix}$$

$$\underset{(A-1)\times 2}{J_{23}} = \begin{bmatrix} (1-\rho)^2 & \omega^2_{a_{\min}+1} + \left(\rho^2-2\theta\rho\right)\omega^2_{a_{\min}} \\ \vdots & \vdots \\ (1-\rho)^2 & \omega^2_{a_{\max}-1} + \left(\rho^2-2\theta\rho\right)\omega^2_{a_{\max}-2} \end{bmatrix}$$

$$\underset{(A-2)\times A}{J_{31}} = \mathbf{0}$$

$$\underset{(A-2)\times(A+2)}{J_{32}} = \begin{bmatrix} 0 & 0 & -\theta\rho & 0 & 0 & 0 & 0 \\ 0 & \ddots & \ddots & \ddots & 0 & \vdots & \vdots \\ 0 & 0 & 0 & 0 & -\theta\rho & 0 & 0 \end{bmatrix}$$

$$\underset{(A-2)\times 2}{J_{33}} = \begin{bmatrix} (1-\rho)^2 & -\rho\omega^2_{a_{\min}+1} \\ \vdots & \vdots \\ (1-\rho)^2 & -\rho\omega^2_{a_{\max}-2} \end{bmatrix}$$

$$\underset{\frac{(A-3)(A-2)}{2}\times A}{J_{41}} = \mathbf{0}$$

$$\underset{\frac{(A-3)(A-2)}{2}\times(A+2)}{J_{42}} = \mathbf{0}$$

$$\underset{\frac{(A-3)(A-2)}{2}\times 2}{J_{43}} = \begin{bmatrix} (1-\rho)^2 & 0 \\ \vdots & \vdots \\ (1-\rho)^2 & 0 \end{bmatrix}$$

The Hessian is approximated by

$$\nabla^2 f\,(\Theta;\rho) \approx J(\Theta;\rho)' J(\Theta;\rho)$$

Once we have solved for $\hat{\Theta}$ for each value of ρ on our grid, we select the estimator $\hat{\Theta}$ together with ρ that together minimise the objective function.

Restrictions. As initial conditions we impose

$$\omega^2_{a_{\min}-1} = \omega^2_{a_{\min}} = \omega^2_{a_{\min}+1}$$

and we further impose

$$\sigma^2_{a_{\max}} = \sigma^2_{a_{\max}-1}; \quad \omega^2_{a_{\max}} = \omega^2_{a_{\min}-1}$$

at the end of the working life. In addition, variances are non-negative,

$$\begin{aligned} &\sigma^2_a \geq 0; \quad \omega^2_a \geq 0 \quad \forall a. \\ &\sigma^2_\alpha \geq 0 \end{aligned}$$

2 Heterogenous Profiles

The Income Process. For each education group and birth cohort, we write log-income net of predictable components[5]of individual i at age $a \in [a_{\min}, ..., a_{\max}]$ and as

$$y_{i,a} = \alpha_i + \beta_i a + v_{i,a} + \tau_{i,a},$$

Individual-specific differences in levels and growth rates are captured by the parameters α_i and β_i drawn from a bivariate mean-zero distribution,

$$\begin{pmatrix} \alpha_i \\ \beta_i \end{pmatrix} \sim (\mathbf{0}, \mathbf{\Sigma}),$$

with covariance matrix $\mathbf{\Sigma}$. Transitory income $\tau_{i,a}$ is assumed to follow a mean-reverting MA(1) process,

$$\tau_{i,a} = \varepsilon_{i,a} + \theta\varepsilon_{i,a-1}, \quad \varepsilon_{i,a},$$

while the permanent component $v_{i,a}$ is assumed to follow an AR(1) process,

$$v_{i,a} = \rho v_{i,a-1} + u_{i,a},$$

with persistence parameter ρ. The shocks to income $\varepsilon_{i,a}$ and $u_{i,a}$ are assumed to be independent of each other, serially uncorrelated, independent of α_i and β_i, with zero-mean and variances ω_a^2 and σ_a^2 respectively. The quasi-difference $\Delta^\rho y_{i,a} \equiv y_{i,a} - \rho y_{i,a-1}$ can be written as

$$\Delta^\rho y_{i,a} = \alpha_i(1-\rho) + \beta_i\xi_a + u_{i,a} + \Delta^\rho\varepsilon_{i,a} + \theta\Delta^\rho\varepsilon_{i,a-1}, \quad a = a_{\min}+1, ..., a_{\max}, \tag{3}$$

with $\xi_a \equiv (1-\rho)\, a + \rho$ and its autocovariance is given by

$$\begin{aligned} &\operatorname{cov}\left(\Delta^\rho y_{i,a}, \Delta^\rho y_{i,a+s}\right) = \left[(1-\rho), \xi_a\right] \mathbf{\Sigma} \left[(1-\rho), \xi_{a+s}\right]' \\ &\begin{cases} +\sigma_a^2 + \omega_a^2 + (\theta-\rho)^2\,\omega_{a-1}^2 + \theta^2\rho^2\omega_{a-2}^2 & \text{if } s=0 \\ +\left[\theta-\rho\right]\left[\omega_a^2 - \theta\rho\omega_{a-1}^2\right] & \text{if } s=1 \\ -\theta\rho\omega_a^2 & \text{if } s=2 \\ 0 & \text{if } s>2 \end{cases} \end{aligned}$$

Estimation. For each value of ρ on a grid we calculate the $A \times A$ autocovariance matrix for the male member of the household for each cohort and each education group in our baseline sample, where $A \equiv a_{\max} - a_{\min}$, with $a_{\min}$and $a_{\max}$the minimum and maximum age of individuals in the sample. For a given age a and lag s, we average $cov(\Delta^\rho y_{i,a}, \Delta^\rho y_{i,a+s})$ across all birth cohorts. Let the resulting "average" autocovariance matrix be denoted by $\overline{C}$ and define the stacked vector[6] of its unique elements by

$$\hat{M} = vech\left(\overline{C}\right).$$

[5]Observable income characteristics including a second polynomial in age, dummies for education, region, family size and marital status, the interaction of family size and marital status and the interaction of the polynomial in age and the education dummy. Family income is given by total household income (NOTE: I removed the equivalence scale adjustment, because it did not matter for our results in the previous version and to be consistent with the analysis of Magne's β in case we include it)

[6]Note that $\hat{M}$ is a $\frac{(A+1)A}{2} \times 1$ vector.

Let $\Theta = \{\sigma_a^2, \omega_a^2, var(\alpha_i), var(\beta_i), \rho_{\alpha\beta}, \theta\}$ denote the parameters[7] to be estimated for a given ρ. The equally-weighted minimum distance estimator $\hat{\Theta}$ is then the solution to

$$\hat{\Theta} = \arg\min_{\Theta} \left[\ M(\Theta;\rho) - \ \hat{M} \ \right]' \left[\ M(\Theta;\rho) - \ \hat{M} \ \right]$$

where $M(\Theta;\rho)$ are the corresponding stacked vector of theoretical moments given by

$$M(\Theta;\rho) = \begin{bmatrix} [(1-\rho),\xi_{26}]\,\mathbf{\Sigma}\,[(1-\rho),\xi_{26}]' + \sigma_{26}^2 + \omega_{26}^2 + (\theta-\rho)^2\omega_{25}^2 + \theta^2\rho^2\omega_{24}^2 \\ \vdots \\ [(1-\rho),\xi_{60}]\,\mathbf{\Sigma}\,[(1-\rho),\xi_{60}]' + \sigma_{60}^2 + \omega_{60}^2 + (\theta-\rho)^2\omega_{59}^2 + \theta^2\rho^2\omega_{58}^2 \\ [(1-\rho),\xi_{26}]\,\mathbf{\Sigma}\,[(1-\rho),\xi_{27}]' + (\theta-\rho)\left(\omega_{26}^2 - \theta\rho\omega_{25}^2\right) \\ \vdots \\ [(1-\rho),\xi_{59}]\,\mathbf{\Sigma}\,[(1-\rho),\xi_{60}]' + (\theta-\rho)\left(\omega_{59}^2 - \theta\rho\omega_{58}^2\right) \\ [(1-\rho),\xi_{26}]\,\mathbf{\Sigma}\,[(1-\rho),\xi_{28}]' - \theta\rho\omega_{26}^2 \\ \vdots \\ [(1-\rho),\xi_{58}]\,\mathbf{\Sigma}\,[(1-\rho),\xi_{60}]' - \theta\rho\omega_{58}^2 \\ [(1-\rho),\xi_{26}]\,\mathbf{\Sigma}\,[(1-\rho),\xi_{29}]' \\ [(1-\rho),\xi_{27}]\,\mathbf{\Sigma}\,[(1-\rho),\xi_{30}]' \\ \vdots \\ [(1-\rho),\xi_{26}]\,\mathbf{\Sigma}\,[(1-\rho),\xi_{60}]' \end{bmatrix}.$$

Thus the objective function can be written as

$$f(\Theta;\rho) = \frac{1}{2}\left[\ M(\Theta;\rho) - \ \hat{M} \ \right]' \left[\ M(\Theta;\rho) - \ \hat{M} \ \right].$$

To estimate the parameters we employ a Gauss-Newton algorithm. The gradient of the objective function reads

$$\nabla f(\Theta;\rho) = J(\Theta;\rho)' \left[\ M(\Theta;\rho) - \ \hat{M} \ \right],$$

where the Jacobian[8] can be represented as

$$J(\Theta;\rho) = \begin{bmatrix} J_{11} & J_{12} & J_{13} & J_{14} \\ J_{21} & J_{22} & J_{23} & J_{24} \\ J_{31} & J_{32} & J_{33} & J_{34} \\ J_{41} & J_{42} & J_{43} & J_{44} \end{bmatrix}$$

[7] $\sigma_{i,a}^2$ is a $A \times 1$ vectors and $\omega_{i,a}^2$ is a $(A+2)\times 1$ vector.

[8] The Jacobian is of dimension $\frac{(A+1)A}{2} \times (A + (A+2) + 2)$

with

$$\underset{A\times A}{J_{11}} = \begin{bmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{bmatrix}$$

$$\underset{A\times(A+2)}{J_{12}} = \begin{bmatrix} \theta^2\rho^2 & (\theta-\rho)^2 & 1 & 0 & 0 \\ 0 & \ddots & \ddots & \ddots & 0 \\ 0 & 0 & \theta^2\rho^2 & (\theta-\rho)^2 & 1 \end{bmatrix}$$

$$\underset{A\times 3}{J_{13}} = \begin{bmatrix} diag_0\left(\mathbf{\Sigma}_1\right) & diag_0\left(\mathbf{\Sigma}_2\right) & diag_0\left(\mathbf{\Sigma}_3\right) \end{bmatrix}$$

$$\underset{A\times 1}{J_{14}} = \begin{bmatrix} 2\left(\theta-\rho\right)\omega_{a_{\min}}^2 + 2\theta\rho^2\omega_{a_{\min}-1}^2 \\ \vdots \\ 2\left(\theta-\rho\right)\omega_{a_{\max}-1}^2 + 2\theta\rho^2\omega_{a_{\max}-2}^2 \end{bmatrix}$$

$$\underset{(A-1)\times A}{J_{21}} = \mathbf{0}$$

$$\underset{(A-1)\times(A+2)}{J_{22}} = \begin{bmatrix} 0 & -\left(\theta-\rho\right)\theta\rho & \left(\theta-\rho\right) & 0 & 0 & 0 \\ 0 & \ddots & \ddots & \ddots & 0 & 0 \\ 0 & 0 & 0 & -\left(\theta-\rho\right)\theta\rho & \left(\theta-\rho\right) & 0 \end{bmatrix}$$

$$\underset{(A-1)\times 3}{J_{23}} = \begin{bmatrix} diag_1\left(\mathbf{\Sigma}_1\right) & diag_1\left(\mathbf{\Sigma}_2\right) & diag_1\left(\mathbf{\Sigma}_3\right) \end{bmatrix}$$

$$\underset{(A-1)\times 2}{J_{24}} = \begin{bmatrix} \omega_{a_{\min}+1}^2 + \left(\rho^2 - 2\theta\rho\right)\omega_{a_{\min}}^2 \\ \vdots \\ \omega_{a_{\max}-1}^2 + \left(\rho^2 - 2\theta\rho\right)\omega_{a_{\max}-2}^2 \end{bmatrix}$$

$$\underset{(A-2)\times A}{J_{31}} = \mathbf{0}$$

$$\underset{(A-2)\times(A+2)}{J_{32}} = \begin{bmatrix} 0 & 0 & -\theta\rho & 0 & 0 & 0 & 0 \\ 0 & \ddots & \ddots & \ddots & 0 & \vdots & \vdots \\ 0 & 0 & 0 & 0 & -\theta\rho & 0 & 0 \end{bmatrix}$$

$$\underset{(A-2)\times 3}{J_{33}} = \begin{bmatrix} diag_2\left(\mathbf{\Sigma}_1\right) & diag_2\left(\mathbf{\Sigma}_2\right) & diag_2\left(\mathbf{\Sigma}_3\right) \end{bmatrix}$$

$$\underset{(A-2)\times 2}{J_{34}} = \begin{bmatrix} -\rho\omega_{a_{\min}+1}^2 \\ \vdots \\ -\rho\omega_{a_{\max}-2}^2 \end{bmatrix}$$

$$\underset{\frac{(A-3)(A-2)}{2}\times A}{J_{41}} = \mathbf{0}$$

$$\underset{\frac{(A-3)(A-2)}{2}\times(A+2)}{J_{42}} = \mathbf{0}$$

$$\underset{A\times 3}{J_{43}} = vech\begin{bmatrix} diag_{3+}\left(\mathbf{\Sigma}_1\right) & diag_{3+}\left(\mathbf{\Sigma}_2\right) & diag_{3+}\left(\mathbf{\Sigma}_3\right) \end{bmatrix}$$

$$\underset{\frac{(A-3)(A-2)}{2}\times 2}{J_{44}} = \mathbf{0}$$

with

$$\Sigma_1 = [(1-\rho), \xi_a] \begin{bmatrix} 1 & \frac{1}{2}\rho_{\alpha\beta}\sqrt{\frac{var(\beta_i)}{var(\alpha_i)}} \\ \frac{1}{2}\rho_{\alpha\beta}\sqrt{\frac{var(\beta_i)}{var(\alpha_i)}} & 0 \end{bmatrix} [(1-\rho), \xi_a]'$$

$$\Sigma_2 = [(1-\rho), \xi_a] \begin{bmatrix} 0 & \frac{1}{2}\rho_{\alpha\beta}\sqrt{\frac{var(\alpha_i)}{var(\beta_i)}} \\ \frac{1}{2}\rho_{\alpha\beta}\sqrt{\frac{var(\alpha_i)}{var(\beta_i)}} & 1 \end{bmatrix} [(1-\rho), \xi_a]'$$

$$\Sigma_3 = [(1-\rho), \xi_a] \begin{bmatrix} 0 & \sqrt{var(\alpha_i)var(\beta_i)} \\ \sqrt{var(\alpha_i)var(\beta_i)} & 0 \end{bmatrix} [(1-\rho), \xi_a]'$$

and $diag_k(x)$ represents the diagonal k of matrix x, $k = 0$ is the main diagonal. The Hessian is approximated by

$$\nabla^2 f(\Theta; \rho) \approx J(\Theta; \rho)' J(\Theta; \rho)$$

Once we have solved for $\hat{\Theta}$ for each value of ρ on our grid, we select the estimator $\hat{\Theta}$ together with ρ that together minimise the objective function.

Restrictions. As initial conditions we impose

$$\omega^2_{a_{\min}-1} = \omega^2_{a_{\min}} = \omega^2_{a_{\min}+1}$$

and we further impose

$$\sigma^2_{a_{\max}} = \sigma^2_{a_{\max}-1}; \quad \omega^2_{a_{\max}} = \omega^2_{a_{\min}-1}$$

at the end of the working life. In addition, variances are non-negative,

$$\begin{aligned} \sigma^2_a \geq 0; \quad \omega^2_a &\geq 0 \quad \forall a. \\ -1 \leq \rho_{\alpha\beta} &\leq 1 \\ var(\alpha_i) &\geq 0 \\ var(\beta_i) &\geq 0 \end{aligned}$$

3 Dual Earner Families

3.1 The Income Process

For each education group and birth cohort, we write log-income net of predictable components[9]of individual i of household j at age $a \in [a_{\min}, ..., a_{\max}]$ as

$$y_{i,j,a} = \alpha_{i,j} + \tau_{i,j,a} + v_{i,j,a}$$

Transitory income $\tau_{i,j,a}$ is assumed to follow a mean-reverting MA(1) process,

$$\tau_{i,j,a} = \varepsilon_{i,j,a} + \theta_i \varepsilon_{i,j,a-1}, \quad \varepsilon_{i,j,a},$$

[9]Observable income characteristics including a second polynomial in age, dummies for education, region, family size and marital status, the interaction of family size and marital status and the interaction of the polynomial in age and the education dummy. Family income is given by total household income (NOTE: I removed the equivalence scale adjustment, because it did not matter for our results in the previous version and to be consistent with the analysis of Magne's β in case we include it)

while the permanent component $v_{i,j,a}$ is assumed to follow an AR(1) process,

$$v_{i,j,a} = \rho_i v_{i,j,a-1} + u_{i,j,a}.$$

with persistence parameter ρ_i. Shocks are assumed to be serially uncorrelated, but potentially correlated within household j,

$$\mathbb{E}\left(u_{i,j,a}u_{i',j',a+s}\right) = \begin{cases} \sigma^2_{i,a} & \text{if } i = i',\, j = j' \text{and } s = 0 \\ \sigma_{i,i',a} & \text{if } i \neq i',\, j = j' \text{and } s = 0 \\ 0 & otherwise \end{cases} \quad \mathbb{E}\left(\varepsilon_{i,j,a}\varepsilon_{i',j,a+s}\right) = \begin{cases} \omega^2_{i,a} & \text{if } i = i',\, j = j' \text{and } s = 0 \\ \omega_{i,i',a} & \text{if } i \neq i',\, j = j' \text{and } s = 0 \\ 0 & otherwise \end{cases}$$

and

$$\mathbb{E}\left(\alpha_{i,j}, \alpha_{i',j'}\right) = \begin{cases} var\left(\alpha_i\right) & \text{if } i = i',\, j = j' \\ cov\left(\alpha_i, \alpha_{i'}\right) & \text{if } i \neq i',\, j = j' \\ 0 & otherwise \end{cases}$$

Quasi-differencing yields

$$\begin{aligned} \Delta^{\rho_i} y_{i,j,a} &= \alpha_{i,j}(1-\rho_i) + u_{i,j,a} + \Delta^{\rho_i}\varepsilon_{i,j,a} + \theta_i \Delta^{\rho_i}\varepsilon_{i,j,a-1} \\ &= \alpha_{i,j}(1-\rho_i) + u_{i,j,a} + \varepsilon_{i,j,a} + \left[\theta_i - \rho_i\right]\varepsilon_{i,j,a-1} - \theta_i\rho_i\varepsilon_{i,j,a-2}. \end{aligned}$$

The auto-covariance $cov\left(\Delta^{\rho_i} y_{i,j,a}, \Delta^{\rho_{i'}} y_{i,j,a+s}\right)$ is given by

$$cov(\Delta^{\rho_i} y_{i,j,a}, \Delta^{\rho_i} y_{i,j,a+s}) =$$
$$(1-\rho_i)^2\, var\left(\alpha_i\right) \begin{cases} +\sigma^2_{i,a} + \omega^2_{i,a} + \left(\theta_i - \rho_i\right)^2 \omega^2_{i,a-1} + \theta_i^2\rho_i^2\omega^2_{i,a-2} & \text{if } s = 0 \\ +\left[\theta_i - \rho_i\right]\left[\omega^2_{i,a} - \theta_i\rho_i\omega^2_{i,a-1}\right] & \text{if } s = 1 \\ -\theta_i\rho_i\omega^2_{i,a} & \text{if } s = 2 \\ +0 & \text{if } s > 2 \end{cases}$$

The cross-covariance $cov\left(\Delta^{\rho_i} y_{i,j,a}, \Delta^{\rho_{i'}} y_{i',j',a+s}\right)$ for $j = j'$ and $i \neq i'$ is given by

$$cov(\Delta^{\rho_i} y_{i,j,a}, \Delta^{\rho_{i'}} y_{i',j,a+s}) =$$
$$\left(1-\rho_i\right)\left(1-\rho_{i'}\right) cov\left(\alpha_i, \alpha_{i'}\right) \begin{cases} +\sigma_{i,i',a} + \omega_{i,i',a} + \left(\theta_i - \rho_i\right)\left(\theta_{i'} - \rho_{i'}\right)\omega_{i,i',a-1} + \rho_i\theta_i\rho_{i'}\theta_{i'}\omega_{i,i',a-2} & \text{if } s = 0 \\ +\left[\theta_{i'} - \rho_{i'}\right]\omega_{i,i',a} - \theta_{i'}\rho_{i'}\left[\theta_i - \rho_i\right]\omega_{i,i',a-1} & \text{if } s = 1 \\ -\theta_{i'}\rho_{i'}\omega_{i,i',a} & \text{if } s = 2 \\ +0 & \text{if } s > 2 \end{cases}$$

3.2 Estimation

On a 2-dimensional grid $\rho_i \times \rho_{i'}$ we calculate the

- the $A \times A$ autocovariance matrix for the male member of the household $cov(\Delta^{\rho_i} y_{i,j,a}, \Delta^{\rho_i} y_{i,j,a+s})$
- the $A \times A$ autocovariance matrix of the spouse $cov(\Delta^{\rho_{i'}} y_{i',j,a}, \Delta^{\rho_{i'}} y_{i',j,a+s})$
- the $A \times A$ cross-covariance matrix $cov(\Delta^{\rho_i} y_{i,j,a}, \Delta^{\rho_{i'}} y_{i',j,a+s})$

for each cohort and each education group in our sample of dual earner families. For a given age a and lag s, we average these matrices across all birth cohorts. Let the resulting "average" matrices be denoted by $\overline{C}_i, \overline{C}_{i'}$ and $\overline{C}_{i,i'}$ and define the stacked vectors of their unique elements by

$$\hat{M}_i = vech\left(\overline{C}_i\right) \quad \hat{M}_{i'} = vech\left(\overline{C}_{i'}\right) \quad \hat{M}_{i,i'} = vech\left(\overline{C}_{i,i'}\right)$$

where the first two vectors are of length[10] $\frac{(A-1)A}{2}$, while the latter vector is of length A^2 because the cross-covariance matrix is not symetric.

Let $\Theta = \{\sigma^2_{i,a}\sigma^2_{i',a}, \sigma_{i,i',a}, \omega^2_{i,a}, \omega^2_{i',a}, \omega_{i,i',a}, \theta_i, \theta_{i'}, var\left(\alpha_i\right), var\left(\alpha_{i'}\right), cov\left(\alpha_i, \alpha_{i'}\right)\}$ denote the parameters[11] to be estimated for a given ρ_i and $\rho_{i'}$. The equally-weighted minimum distance estimator $\hat{\Theta}$ is then the solution to

$$\hat{\Theta} = \arg\min_{\Theta} \left[\ \mathcal{M}\left(\Theta; \rho_i, \rho_{i'}\right) - \ \ \hat{\mathcal{M}} \ \ \right]' \left[\ \mathcal{M}\left(\Theta; \rho_i, \rho_{i'}\right) - \ \ \hat{\mathcal{M}} \ \ \right]$$

where where $\hat{\mathcal{M}} \equiv [\hat{M}_i; \hat{M}; \hat{M}_{i,i'}]$ is the $\left((A-1)A + A^2\right) \times 1$ vector of empirical moments and $\mathcal{M}\left(\Theta; \rho_i, \rho_{i'}\right)$ denotes the corresponding stacked vectors of theoretical moments. To estimate the parameters we rely on finite difference approximations for the gradient and a dense quasi-Newton BFGS Hessian. Once we have solved for $\hat{\Theta}$ for each point on our grid we select the estimates and ρ_i and $\rho_{i'}$ that yield the minimum of the objective function.

Restrictions. As initial conditions we impose

$$\begin{aligned} \omega^2_{i,a_{\min}-1} &= \omega^2_{i,a_{\min}} = \omega^2_{i,a_{\min}+1} \\ \omega^2_{i',a_{\min}-1} &= \omega^2_{i',a_{\min}} = \omega^2_{i',a_{\min}+1} \\ \omega_{i,i',a_{\min}-1} &= \omega_{i,i',a_{\min}} = \omega_{i,i',a_{\min}+1} \end{aligned}$$

and we further impose

$$\begin{aligned} \sigma^2_{i,a_{\max}} &= \sigma^2_{i,a_{\max}-1}; \quad \omega^2_{i,a_{\max}} = \omega^2_{i,a_{\min}-1} \\ \sigma^2_{i',a_{\max}} &= \sigma^2_{i',a_{\max}-1}; \quad \omega^2_{i',a_{\max}} = \omega^2_{i',a_{\min}-1} \\ \sigma_{i,i',a_{\max}} &= \sigma_{i,i',a_{\max}-1}; \quad \omega_{i,i',a_{\max}} = \omega_{i,i',a_{\min}-1} \end{aligned}$$

at the end of the working life. In addition, variances are non-negative,

$$\begin{aligned} \sigma^2_{j,a} \geq 0; \quad \omega^2_{j,a} \geq 0 \quad \forall a, j = i, i'. \\ var(\alpha_i) \geq 0 \quad var(\alpha_{i'}) \geq 0 \end{aligned}$$

[10] $A \equiv a_{\max} - a_{\min}$, with $a_{\min}$and $a_{\max}$the minimum and maximum age of individuals in our sample
[11] $\sigma^2_{i,a}, \sigma^2_{i',a}$ and $\sigma_{i,i',a}$ are $A \times 1$ vectors and $\omega^2_{i,a}, \omega^2_{i',a}$ and $\omega_{i,i',a}$ are $(A+2) \times 1$ vectors.