Separating uncertainty from heterogeneity in life cycle earnings

Cunha, Heckman and Navarro

October 17 2006

A central element in macro is the income process

Example: Storesletten et al JME 2004

- We can we write a general income process as

$$y_t^i = g(\theta^0, X_t^i) + u_t^i$$
$$u_t^i = f(\theta^i, X_t^i) + z_t^i + \varepsilon_t^i$$
$$z_t^i = \rho z_{t-1}^i + \eta_t^i$$

where in Storesletten et al (2004) $f(\theta^i, X_t^i) = \alpha^i$

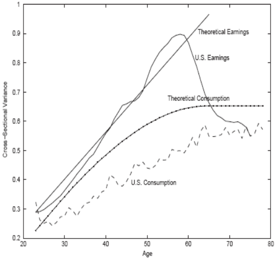
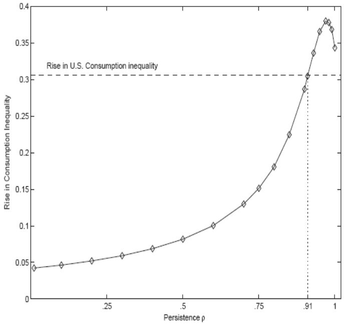


Fig. 4. This graph compares the population moments from our benchmark model with those of the data. The solid lines (without dots) represent the theoretical and empirical cross-sectional variance of log earnings. The dashed line represents the empirical cross-sectional variance of consumption and the solid-dotted line represents the theoretical cross-sectional variance of consumption from the benchmark economy.

The parameters $(\rho, \sigma_\alpha^2, \sigma_\varepsilon^2, \sigma_\eta^2)$ are key to match the model with data

In Guvenen (2006), the high persistence is replaced by heterogeneity and learning about components of u

Storesletten et al (2004):



Guvenen (2006):

$$y_t^i = g(\theta^0, X_t^i) + u_t^i$$
$$u_t^i = f(\theta^i, X_t^i) + z_t^i + \varepsilon_t^i$$
$$z_t^i = \rho z_{t-1}^i + \eta_t^i$$

- $f(\theta^i, X_t^i) = \alpha^i + \beta^i \cdot age_i$ (Heterogenous Income Profiles)
- Storesletten et al (2004): $f(\theta^i, X_t^i) = \alpha^i$ (Restricted Income Profiles)

Separating uncertainty from heterogeneity in life-cycle earnings - the Roy model

Under complete markets the present discounted value of income under either choice for each individual i is

$$Y_{1,i} = \sum_{t=0}^{T} \frac{Y_{1,i,t}}{(1+r)^t}$$

$$Y_{0,i} = \sum_{t=0}^{T} \frac{Y_{0,i,t}}{(1+r)^t}$$

where only one of the two objects will be observed for each individual.

The choice of the agent is governed by

$$S_i = \begin{cases} 1 \text{ if } E\left[Y_{1,i} - Y_{0,i} - C_i | J_{i,0}\right] \geq 0 \\ 0, \text{ otherwise} \end{cases}$$

where S_i is a schooling choice at time $t = 0$, $S_i = 1$ for college and $S_i = 0$ for high school, and C_i a cost for $S_i = 1$

Choice, income and another way of specifying heterogeneity in u

The choice of the agent generates a choice index

$$I_i = E\left[\sum_{t=1}^{5} \frac{Y_{1,i,t} - Y_{0,i,t}}{(1+r)^t} - C_i \middle| J_{i,0}\right]$$

Agent i's income and cost at time t for choice s=0,1

$$\begin{cases} Y_{s,i,t} = X_{i,t}\beta_{s,t} + \underbrace{\theta_{i1}\alpha_{s,1,t} + \theta_{i2}\alpha_{s,2,t} + \varepsilon_{s,i,t}}_{=U_{s,i,t}} & t = 1,2,3 \\ Y_{s,i,t} = X_{i,t}\beta_{s,t} + \underbrace{\theta_{i1}\alpha_{s,1,t} + \theta_{i2}\alpha_{s,2,t} + \theta_{i3}\alpha_{s,3,t} + \varepsilon_{s,i,t}}_{=U_{s,i,t}} & t = 4,5 \\ C_i = Z_i\gamma + \underbrace{\theta_{i1}\alpha_{C,1,t} + \theta_{i2}\alpha_{C,2,t} + \varepsilon_{C,i,t}}_{=U_{C,i}} & \end{cases}$$

where $(\mathbf{X}_{it}, Z_i)$ are vectors of observable characteristics, $\boldsymbol{\theta}_i$ is a vector of latent factors, and $\boldsymbol{\alpha}$ factor loadings (skill prices).

Test scores and factor properties

Let $\mathbf{M}_i$ denote a vector of test scores on cognitive ability (ASVAB)

$$\begin{cases} M_{ij} = & X_{Mi}\beta_{M,j} + \theta_{i1}\alpha_{M,j} + \varepsilon_{Mi,j} \qquad j = 1, ..., 5 \\ C_i = & Z_i\gamma + \theta_{i1}\alpha_{C,1} + \theta_{i2}\alpha_{C,2} + \varepsilon_{C,i} \end{cases}$$

where $\theta_{i1} \perp \theta_{i2}$ and $\varepsilon_i \perp \boldsymbol{\theta}_i$.Each factor follows a mixture of two normals:

$$\theta_k \sim \textstyle\sum_{j=1}^{2} p_{kj} f\left(\theta_k | \mu_{kj}, \sigma_{kj}^2\right)$$

where, WLOG, we shift means (μ_{kj}) so that

$$E\left[\theta_k\right] = \textstyle\sum_{j=1}^{2} p_{kj}\mu_{kj} = 0$$

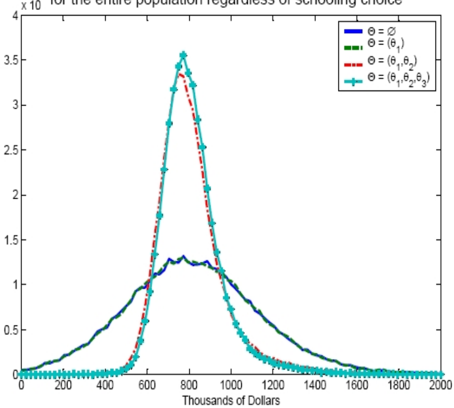
The complete set-up

$$\begin{cases} M_{ij} & = X_{Mi}\beta_{M,j} + \theta_{i1}\alpha_{M,j} + \varepsilon_{Mi,j} & j = 1, ..5 \\ Y_{1,i,t} & = X_{i,t}\beta_{1,t} + U_{1,i,t} & t = 1, \dots T \\ Y_{0,i,t} & = X_{i,t}\beta_{0,t} + U_{0,i,t} & t = 1, \dots T \\ C_i & = Z_i\gamma + U_{C,i} & \\ I_i & = E\left[\sum_{t=1}^{T} \frac{Y_{1,i,t} - Y_{0,i,t}}{(1+r)^t} - C_i \middle| J_{i,0}\right] & \\ & = E\left[\mu_I(X_i, Z_i) - U_{I,i} \middle| J_{i,0}\right] & \end{cases}$$

- M: ASVAB test scores
- $\theta_k \sim \sum_{j=1}^{2} p_{kj} f\left(\theta_k | \mu_{kj}, \sigma_{kj}^2\right)$
  - Shift means (μ_{kj}) so that $E\left[\theta_k\right] = \sum_{j=1}^{2} p_{kj}\mu_{kj} = 0$

CHN are able to derive the density of PDV of income under the information set of the agent, i.e. tell which components of U that are known

Figure 9.1
Densities of present value of high school earnings under different information sets for the agent calculated for the entire population regardless of schooling choice



Identifying the information set of agents

A simplified set-up:

$$\begin{cases} Y_{s,i} & = \mathbf{X}_i \boldsymbol{\beta}_s + \underbrace{\theta_i \alpha_s + \varepsilon_{s,i}}_{=U_{s,i}} \qquad s = 0,1 \\ I_i & = E\left[Y_{1,i} - Y_{0,i} \middle| J_{i,0}\right] \\ & = E\left[\mu_I(X_i) + U_{I,i} \middle| J_{i,0}\right] \end{cases}$$

Identifying the information set of agents (cont.)

- Assume agents know $\boldsymbol{\beta}_s, \alpha_s$
- Assume agents do not know $\varepsilon_{1,i}, \varepsilon_{0,i}$
- From the point of view of the econometrician, treat $\boldsymbol{\beta}_s, \alpha_s$ as known for now
- True choice index:

$$\begin{aligned} I_i &= E\left[Y_{1,i} - Y_{0,i} | J_{i,0}\right] = \\ &= E[X_i(\beta_1 - \beta_0) + \theta_i(\alpha_1 - \alpha_0) + \varepsilon_{1,i} - \varepsilon_{0,i} | J_{i,0}] = \\ &= \underbrace{E[X_i | J_{i,0}](\beta_1 - \beta_0)}_{= E[\mu_I(X_i) | J_{i,0}]} \\ &+ \underbrace{E[\theta_{i1} | J_{i,0}](\alpha_1 - \alpha_0) + E[\varepsilon_{1,i} - \varepsilon_{0,i} | J_{i,0}]}_{= E[U_{I,i} | J_{i,0}]} \end{aligned}$$

Identifying the information set of agents (cont.)

- We can also think of any $\tilde{J_{i,0}}$ and decompose components into knowns and unknowns
  - $E[X_i|\tilde{J_{i,0}}], [X_i - E[X_i|\tilde{J_{i,0}}]]$
  - $E[\theta_i|\tilde{J_{i,0}}], [\theta_i - E[\theta_i|\tilde{J_{i,0}}]]$
  - $E[\varepsilon_{1,i} - \varepsilon_{0,i}|\tilde{J_{i,0}}], [\varepsilon_{1,i} - \varepsilon_{0,i} - E[\varepsilon_{1,i} - \varepsilon_{0,i}|\tilde{J_{i,0}}]]$
- If we only decompose θ_i the choice index $\tilde{I_i}$ would be

$$
\begin{aligned}
\tilde{I_i} &= E\left[Y_{1,i} - Y_{0,i}|\tilde{J_{i,0}}\right] \\
&= E[X_i|\tilde{J_{i,0}}](\beta_1 - \beta_0) \\
&+ E[\theta_i|\tilde{J_{i,0}}](\alpha_1 - \alpha_0) + E[\varepsilon_{1,i} - \varepsilon_{0,i}|\tilde{J_{i,0}}] \\
&+ [\theta_i - E[\theta_i|\tilde{J_{i,0}}]](\alpha_1 - \alpha_0)\Delta^*_\theta
\end{aligned}
$$

Identifying the information set of agents (cont.)

- In practice, the proposed $J_{i,0}^{\tilde{}}$ and the associated choice index $I_i^{\tilde{}}$ is of the following kind
  - assume without testing $X_i \in J_{i,0}^{\tilde{}}, \varepsilon_{1,i} - \varepsilon_{0,i} \notin J_{i,0}^{\tilde{}}$
  - test $\theta_i \notin J_{i,0}^{\tilde{}} \Leftrightarrow H_0 : E[\theta_i | J_{i,0}^{\tilde{}}] = E[\theta_i] = \sum_{j=1}^{2} p_j \mu_j = 0$
  - innovation under $J_{i,0}^{\tilde{}}$: $\theta_i - E[\theta_i | J_{i,0}^{\tilde{}}]$
- The choice index under $J_{i,0}^{\tilde{}}$

$$
\begin{aligned}
I_i^{\tilde{}} \quad &= E[X_i | J_{i,0}^{\tilde{}}](\beta_1 - \beta_0) + E[\theta_i | J_{i,0}^{\tilde{}}](\alpha_1 - \alpha_0) + E[\varepsilon_{1,i} - \varepsilon_{0,i} | J_{i,0}^{\tilde{}}] \\
&+ [\theta_i - E[\theta_i | J_{i,0}^{\tilde{}}]](\alpha_1 - \alpha_0)\Delta_\theta^* = \\
&= X_i(\beta_1 - \beta_0) + \theta_i(\alpha_1 - \alpha_0)\Delta_\theta^* = \\
&= X_i(\beta_1 - \beta_0) + \theta_i \Delta_\theta
\end{aligned}
$$

Identifying the information set of agents (cont.)

- The choice index under $J_{i,0}^{\tilde{}}$

$$\begin{aligned} I_i^{\tilde{}} &= \mathbf{X}_i(\boldsymbol{\beta}_1 - \boldsymbol{\beta}_0) + \boldsymbol{\theta}_i \boldsymbol{\Delta}_\theta \\ \Leftrightarrow \quad I_i^{\tilde{}} - \mathbf{X}_i(\boldsymbol{\beta}_1 - \boldsymbol{\beta}_0) &= \boldsymbol{\theta}_i \boldsymbol{\Delta}_\theta \end{aligned}$$

- To be correct, add cost $C_i = Z_i\gamma + \varepsilon_{c,i} \in J_{i,0}^{\tilde{}}$ to get the error term:

$$I_i^{\tilde{}} - \mathbf{X}_i(\boldsymbol{\beta}_1 - \boldsymbol{\beta}_0) + Z_i\gamma = \boldsymbol{\theta}_i \boldsymbol{\Delta}_\theta + \varepsilon_{C,i}$$

- Estimate the coefficient on the innovation: Δ_θ
  - $H_0 : E[\theta_i | J_{i,0}^{\tilde{}}] = 0 \Leftrightarrow H_0 : \Delta_\theta = 0$
  - If agents' know $\boldsymbol{\theta}$ and if the model is correctly specified then $\boldsymbol{\Delta}_{\theta k} = (\boldsymbol{\alpha}_{1k} - \boldsymbol{\alpha}_{0k} - \boldsymbol{\alpha}_{ck})$
  - Claim: if $J_{i,0} = J_{i,0}^{\tilde{}}$ then $\Delta_\theta = 0$ and $I_i^{\tilde{}} = I_i$
  - Equivalent claim: if $\Delta_\theta = 0$ then $[\theta_i - E[\theta_i | J_{i,0}^{\tilde{}}]]$ represents a component which is unknown to the agent or a component which is known by the agent but not acted upon

Main finding: two out of three factors are known, so cross-sectional variance is not equal to uncertainty

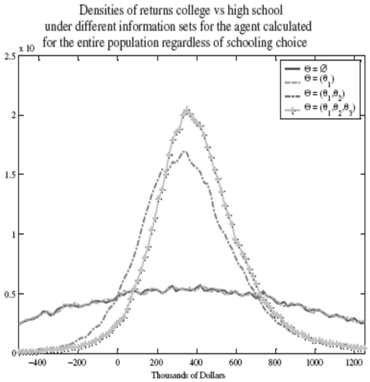
- At t=0 (θ_1, θ_2) are acted upon (known)
- Half of the variability in returns to schooling is forecastable

Table 6.1 Agent's forecast variance of present value of earnings* under different information sets (fraction of the variance explained by Θ).† The calculation is for the entire population regardless of schooling choice

For lifetime:‡	$Var(Y_c)$	$Var(Y_h)$	$Var(Y_c - Y_h)$
Variance when $\Theta = \emptyset$	156402.14	73827.89	267796.38
$\Theta = \{\theta_1\}$	0.95%	0.27%	0.44%
$\Theta = \{\theta_1, \theta_2\}$	29.10%	29.43%	47.42%
$\Theta = \{\theta_1, \theta_2, \theta_3\}$	68.03%	32.27%	62.65%

*We use an interest rate of 3% to calculate the present value of earnings.
†The variance of the unpredictable component of period 1 college earnings, where $\Theta = \{\theta_1\}$ is $(1 - 0.0095) * 156402.14$.
‡Variance of the unpredictable component of earnings between age 19 and 65 as predicted at age 19.

Main finding: returns to college are more forecastable than what we might think



Conclusion: *This has important implications for using measured variability to price risk and predict college attendance*