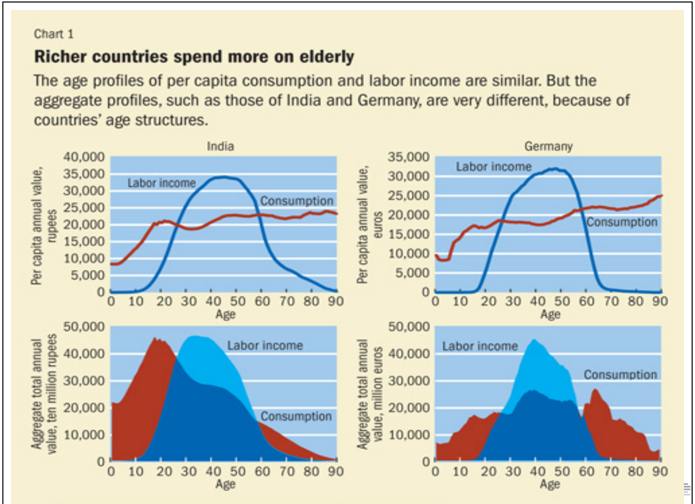
Calibrating OG-India: Lifetime Income Profiles

August 21, 2019

What is an earning profile?

- An empirical fact is that earnings vary over one's lifetime:



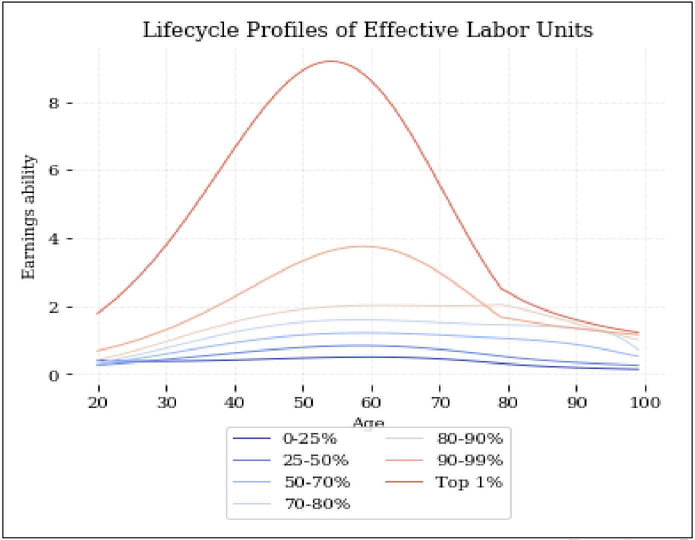
Implications

- The fact that earnings are not constant over the lifecycle has important implications for:
  - Savings
  - Consumption
  - Labor supply
- It is therefore important to capture the shape of earnings profiles in the model

Heterogeneity

- In addition to variance across the lifecycle, we are going to want to capture variance across households in earnings
- This is important for questions of:
  - inequality
  - the impact of taxes across low and high-skilled households
- Therefore, we will want to estimate earnings profiles separately across households

Heterogeneity



Heterogeneity

- We will considering profiles that vary by “ability type”
- We define “ability type” by ones *lifetime income group*
    - Lifetime income group means the present discounted value of *potential* lifetime earnings
        - That is, what is the present value of earnings if you worked full-time in every year of life?
    - In the model and estimation, we will divide household into a finite number of groups based on lifetime earnings
    - Each group will have a lifecycle profile of earnings

Earnings in theory

- Total labor earnings in the model is given by: $w_t e_{j,s} n$
  - $n_{j,s,t}$ is the number of units of labor supplied (e.g., hours)
  - $e_{j,s}$ are the number of effective labor units per unit of labor supplied (i.e., labor productivity)
  - w_t is wage rate per effective labor unit

Earnings in theory

- The $e_{j,s}$ come into the budget constraint:

$$c_{j,s,t} + b_{j,s+1,t+1} = (1 + r_t)b_{j,s,t} + w_t e_{j,s} n_{j,s,t} + \zeta_{j,s}\frac{BQ_t}{\lambda_j \omega_{s,t}} + \eta_{j,s,t}\frac{TR_t}{\lambda_j \omega_{s,t}} - T_{s,t}$$

- and the first order condition for the household's choice of labor supply:

$$w_t e_{j,s}\left(1 - \tau^{mtrx}_{s,t}\right)(c_{j,s,t})^{-\sigma} = e^{g_y(1-\sigma)}\chi^n_s\left(\frac{b}{\tilde{l}}\right)\left(\frac{n_{j,s,t}}{\tilde{l}}\right)^{\upsilon-1}\left[1 - \left(\frac{n_{j,s,t}}{\tilde{l}}\right)^{\upsilon}\right]^{\frac{1-\upsilon}{\upsilon}}$$

$$\forall j, t, \quad \text{and} \quad E + 1 \leq s \leq E + S$$

What do we want to estimate?

- In the end, we want matrix: $\mathbf{e}$, with elements $e_{j,s}$
- Each vector, $\mathbf{e}_j$ represents a lifetime profile of effective labor units for a given “ability type”

Data requirements:

- We need a measure of hourly earnings, these are the $w_t e_{j,s}$ from the model
- We need to be able to identify the "ability group", j
    - We will define "ability group" by one's lifetime earnings potential
    - Specifically, estimate ones potential earnings over the lifetime if working full time for ages 20-80.
    - We put households into groups based on the net present value of these potential earnings

Identifying lifetime income groups

- To do this, we'll want to observe a given househld over many years
  - Without this, it's hard to know what their lifetime earning potential is.
- But even in the longest panel, we still don't observe households over all working years
- We thus fit a regression model to impute wages in year they are not observed"

$$ln(w_{i,t}) = \alpha_i + \beta_1 age_{i,t} + \beta_2 age^2_{i,t} + \beta_3 * age^3_{i,t} + \varepsilon_{i,t} \quad (1)$$

Identifying lifetime income groups

- With fitted values for wages, we then compute lifetime income, assuming full-time work:

$$LI_i = \sum_{t=21}^{80} \left(\frac{1}{1+r}\right)^{t-21} (w_{i,t} * 4000) \tag{2}$$

- We then group households based on their percentile in the distribution of lifetime income
- NOTE: these percentiles need to be consistent with the λ parameter used in the model

Estimating Lifetime Income Profiles

- To find the lifecycle profiles of income, we estimate, separately for each percentile group, the regression:

$$ln(w_{i,t}) = \alpha + \beta_1 age_{i,t} + \beta_2 age^2_{i,t} + \beta_3 * age^3_{i,t} + \varepsilon_{i,t} \quad (3)$$

Income Processes in `OG-India`

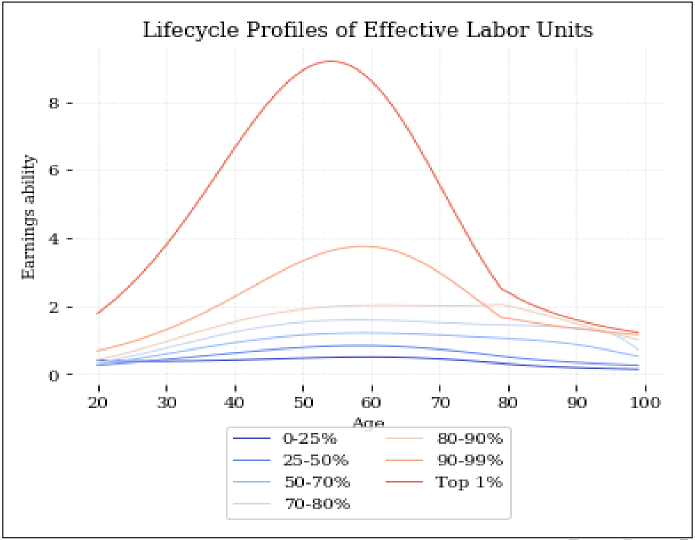
- The place where the estimated processes go into `OG-India` is in the `income.py` script module.
- This script takes the coefficients from the regressions estimated lifetime-income-group-specicic earnings profiles
- These coefficients are hardcoded into the `income.get_e_orig()` function.
- The coefficients come from esstimating regressions for seven lifetime income groups: 0-25%, 25-50%, 50-70%, 70-80%, 80-90%, 90-99%, Top 1%.

Income Processes in `OG-India`

In addition to containing these regression coefficients, `income.get_e_orig()` does the following:

- Scales the $e_{j,s}$ so that weighed average across ability types is 1.
- Uses an inverse tangent function to extrapolate profiles from ages 81-100.
- Interpolates the $e_{j,s}$ for different values of S (the number of economically active periods a model agent lives)
- Allows one to plot the processes

Earnings profiles



How should we calibrate $e_{j,s}$ with Indian data?

Challenges:

- No panel data – hard to identify lifetime-income groups
- Few data with hourly wages by age
- Administrative tax data only represent 3-5% of the population

Some ideas

- Pramanik shared data with wages for rural laborers and "regular workers"
- We could construct lifetime income profiles separately for these (and other) groups
- How:
  - Use cross-sectional data we do have on earnings
  - Make some assumptions about earnings profiles (e.g., rural laborers have flat earnings profiles)