Code documentation

César Garro-Marín[*]

July 10, 2023

1 Notation

Our GMM model is based on the moment equations:

$$\mathbb{E}\left(\psi(w_i, \mu)\right) = 0$$

where:

- w_i is the vector of data for observation i.
- ψ is a $P \times 1$ vector of functions.
- μ is an $R \times 1$ vector of parameters.

We estimate the model's parameters μ by solving he problem:

$$\hat{\mu} = \arg\min_c \left(\frac{1}{N}\sum_{i=1}^{N}\psi(w_i, c)\right)' A \left(\frac{1}{N}\sum_{i=1}^{N}\psi(w_i, c)\right)$$

which we can write as:

$$\hat{\mu} = \arg\min_c \left(\frac{1}{N}\varepsilon(c)'Z\right) A \left(\frac{1}{N}Z'\varepsilon(c)\right)$$

where $\varepsilon(c)$ is the $N \times 1$ vector of errors of the model, and Z is the $N \times P$ matrix of instruments. Here $N = N_1 + N_2 + N_3$ is the total number of observations and N_i denotes the number of observations that belong to equation i.

$\varepsilon(c)$ is defined as follows. Define the functions $J(l)$, $E(l)$, $T(l)$, and $I(l)$ which return the job, education level, year, and skill that correspond to observation l. In addition, for observations belonging to the employment equation, define as $EET(l)$ as the function that returns the education pair-year cell of the observation. Then the error for observation l is:

$$\varepsilon_l(\mu) = \begin{cases} \Delta\overline{\ln S_{ijet+1}} - \sum_k \theta_{ke}\overline{S_{kejt}}\pi_{kjt} + \pi_{ijt} & l \leq N_1 \\ 1 - \sum_k \theta_{ke}\overline{S_{kejt}} & N_1 < l \leq N_1 + N_2 \\ \Delta\left[\ln \frac{q_{ejt+1}}{q_{e'jt+1}}\right] - \beta_j \left[\sum_k \left(\theta_{ke}\overline{S_{kjet}} - \theta_{ke'}\overline{S_{kje't}}\right)\pi_{kjt}\right] - \gamma_{e,e't} & N_1 + N_2 < l \leq N \end{cases}$$

[*] Boston University, email: cesarlgm@bu.edu

2 GMM standard errors

The GMM estimates are distributed as:

$$\sqrt{N}(\hat{\mu}-\mu) \to N(0,\tilde{V})$$

where $\bar{V}=(D'AD)^{-1}D'AVAD(D'AD)^{-1}$. Here:

- D is the model gradient.
- V is defined as:

$$V=\lim_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N}\mathbb{E}\left(\psi(w_i,\mu)\psi(w_i,\mu)'\right)$$

- A is the weighting matrix $(Z'Z)^{-1}$

We estimate $\bar{V}$ as:

$$\hat{\bar{V}}=(\hat{D}'A\hat{D})^{-1}\hat{D}'A\hat{V}A\hat{D}(\hat{D}'A\hat{D})^{-1}$$

where:

- $\hat{V}=\frac{1}{N}\sum_{i=1}^{N}\psi(w_i,\hat{\mu})\psi(w_i,\hat{\mu})'$.
- $\hat{D}=\frac{1}{N}\sum_{i=1}^{N}\frac{\partial\psi(w_i,\hat{\mu})}{\partial c'}$

2.1 Computing the gradient:

Let Ξ be the $N\times R$ matrix with general term $\Xi_{lr}=\frac{\partial\varepsilon(w_l,c)}{\partial c_r}$. Then our estimate of the gradient is $P\times R$ matrix:

$$\hat{D}=\frac{1}{N}Z'\Xi$$

Thus, we jut have to compute Ξ_{lr}. We need to compute derivatives with respect to four types of parameters: θ_{ke}, π_{kjt}, and β_j:

$$\frac{\partial\varepsilon_l(\mu)}{\partial\theta_{ie}} = \begin{cases} -\overline{S_{kejt}}\pi_{ijt} & l\le N_1, e=E(l) \\ -\overline{S_{iejt}} & N_1<l\le N_1+N_2, e=E(l) \\ -\beta_j\overline{S_{ijet}}\pi_{ijt} & N_1+N_2<l\le N, (e,\cdot,\cdot)=EET(l) \\ \beta_j\overline{S_{ijet}}\pi_{ijt} & N_1+N_2<l\le N, (\cdot,e,\cdot)=EET(l) \\ 0 & \text{otherwise} \end{cases}$$

$$\frac{\partial\varepsilon_l(\mu)}{\partial\pi_{ijt}} = \begin{cases} -\theta_{ie}\overline{S_{iejt}} & l\le N_1, i\ne I(l), j=J(l), t=T(l) \\ -\theta_{ie}\overline{S_{iejt}}+1 & l\le N_1, i=I(l), j=J(l), t=T(l) \\ -\beta_j\left(\theta_{ie}\overline{S_{ijet}}-\theta_{ie'}\overline{S_{ije't}}\right) & N_1+N_2<l\le N \\ 0 & \text{otherwise} \end{cases}$$

$$\frac{\partial\varepsilon_l(\mu)}{\partial\beta_j} = \begin{cases} -\left[\sum_k\left(\theta_{ke}\overline{S_{kjet}}-\theta_{ke'}\overline{S_{kje't}}\right)\pi_{kjt}\right] & N_1+N_2<l\le N, j=J(l) \\ 0 & \text{otherwise} \end{cases}$$

$$\frac{\partial\varepsilon_l(\mu)}{\partial\gamma_{ee't}} = \begin{cases} -1 & N_1+N_2<l\le N, (e,e',t)=EET(l) \\ 0 & \text{otherwise} \end{cases}$$

We compute Ξ with the function `get_xi_matrix`, and compute $\bar{V}$ with `estimate_v`. Finally, the function `get_standard_errors` computes the standard errors.