

Code documentation

César Garro-Marín*

July 21, 2023

1 Notation

Our GMM model is based on the moment equations:

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

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 the 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:

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

*Boston University, email: cesarlgm@bu.edu

2 GMM standard errors

The GMM estimates are distributed as:

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

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

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

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

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

We estimate \tilde{V} as:

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

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

$$\begin{aligned}\frac{\partial \varepsilon_l(\mu)}{\partial \theta_{ie}} &= \begin{cases} -\overline{S_{kejt}}\pi_{ijt} & l \leq N_1, e = E(l), i \neq 1 \\ -\overline{S_{iejt}} & N_1 < l \leq N_1 + N_2, e = E(l), i \neq 1 \\ -\beta_j \overline{S_{ijet}}\pi_{ijt} & N_1 + N_2 < l \leq N, (e, \cdot, \cdot) = EET(l), i \neq 1 \\ \beta_j \overline{S_{ijet}}\pi_{ijt} & N_1 + N_2 < l \leq N, (\cdot, e, \cdot) = EET(l), i \neq 1 \\ 0 & \text{otherwise} \end{cases} \\ \frac{\partial \varepsilon_l(\mu)}{\partial \theta_1} &= \begin{cases} -\overline{S_{kejt}}\pi_{ijt} & l \leq N_1 \\ -\overline{S_{iejt}} & N_1 < l \leq N_1 + N_2 \\ \beta_j \overline{S_{ijet}}\pi_{ijt} - \beta_j \overline{S_{ije't}}\pi_{ijt} & N_1 + N_2 < l \leq N \\ 0 & \text{otherwise} \end{cases} \\ \frac{\partial \varepsilon_l(\mu)}{\partial \pi_{ijt}} &= \begin{cases} -\theta_{ie} \overline{S_{iejt}} & l \leq N_1, i \neq I(l), j = J(l), t = T(l) \\ -\theta_{ie} \overline{S_{iejt}} + 1 & l \leq N_1, i = I(l), j = J(l), t = T(l) \\ -\beta_j (\theta_{ie} \overline{S_{ijet}} - \theta_{ie'} \overline{S_{ije't}}) & N_1 + N_2 < l \leq N \\ 0 & \text{otherwise} \end{cases} \\ \frac{\partial \varepsilon_l(\mu)}{\partial \beta_j} &= \begin{cases} -[\sum_k (\theta_{ke} \overline{S_{kjet}} - \theta_{ke'} \overline{S_{kje't}}) \pi_{kjt}] & N_1 + N_2 < l \leq N, j = J(l) \\ 0 & \text{otherwise} \end{cases} \\ \frac{\partial \varepsilon_l(\mu)}{\partial \gamma_{ee't}} &= \begin{cases} -1 & N_1 + N_2 < l \leq N, (e, e', t) = EET(l) \\ 0 & \text{otherwise} \end{cases}\end{aligned}$$

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.

2.2 Standard errors for σ_j and $d \ln A_{ijt}$

The above derivation gives standard errors for $\beta_j = \frac{\sigma_j - 1}{\sigma_j}$. Using the delta method, it is straightforward to compute standard errors for $\sigma_j = \frac{1}{1 - \beta_j}$.

$$\sqrt{N}(\hat{\sigma}_j - \sigma_j) \rightarrow N \left(0, \text{var}(\sigma_j) \left[\frac{\partial \sigma_j}{\partial \beta_j} \right]^2 \right)$$

where:

$$\left[\frac{\partial \sigma_j}{\partial \beta_j} \right]^2 = \frac{1}{(1 - \beta_j)^2} = \sigma_j^2$$

By an analogous argument:

$$\sqrt{N}(d \ln \hat{A}_{ijt} - d \ln A_{ijt}) \rightarrow N \left(0, G(\gamma)' \tilde{V} G(\gamma) \right)$$

where \tilde{V} is the variance matrix of $\psi = (\pi_{ijt}, \beta_j)'$ and $G(\psi) = (\beta_j, \pi_{ijt})'$