Estimation of Production Function 1

Introduction & Panel Data Approach

Yuta Toyama

Last updated: October 14, 2018

Crash course on Generalized Method of Moments (GMM)

- I start with a quick review on generalized method of moments (GMM).
  - GMM is widely used in empirical IO and structural estimation.
  - OLS and IV are within the class of GMM estimator
- The slide is based on the lecture note by Chris Conlon: `https://chrisconlon.github.io/gradio.html`
- For a more detailed treatment of GMM, see
  1. Hayashi "Econometrics" Chapter 3 and 4.
  2. Newey and McFadden (1994) "Large Sample Estimation and Hypothesis Testing," Handbook of Econometrics Chapter 36

Setup

- Data $\{w_i\}_{i=1}^N$, which may include dependent var y_i, regressors x_i, and excluded instruments z_i
- Suppose that economic model implies the following restriction on our data
$$E[g(w_i, \theta_0)] = 0$$
  - At the true parameter θ_0, our moment condition $g()$ are 0 on average.
  - g might be a vector.
- In empirical analysis, we need to consider the sample analogue:
$$\frac{1}{N}\sum_{i=1}^{N} g(w_i, \theta) \equiv g_N(\theta)$$
- (Roughly-speaking) The estimator $\hat{\theta}$ for θ_0 should satisfy $g_N(\hat{\theta}) = 0$.
  - Later I define estimators more formally.

Examples

- OLS: $y_i = \beta x_i + \epsilon_i, E[\epsilon_i x_i] = 0$.We can write down the moment condition
$$E[x_i(y_i - \beta x_i)] = 0$$
- IV: $y_i = \beta x_i + \epsilon_i, E[\epsilon_i z_i] = 0$ where z_i is instrumental variable. Then,
$$E[z_i(y_i - \beta x_i)] = 0$$
- MLE: maximizing the likelihood function
$$\max_{\theta} \sum_{i=1}^{N} \log f(w_i, \theta).$$
The FOC is
$$\sum_{i=1}^{N} \frac{\partial \log f(w_i, \theta)}{\partial \theta} = 0,$$
which can be seen as moment.

- Euler Equation (Hansen 1982)
- Assume a CRRA utility function

$$u(c) = \frac{c^{1-\gamma} - 1}{1 - \gamma}$$

  and an agent maximized the expected disscounted value of their stream of consumption.
- An Euler equation

$$E\left[\beta \left(\frac{c_{t+1}}{c_t}\right)^{-\gamma} R_{t+1} - 1 | \Omega_t\right] = 0$$

  where R_t is interest rate and Ω_t is the information set of everything known until time t.
- For any $z_t \in \Omega_t$, we have

$$E\left[z_t \left(\beta \left(\frac{c_{t+1}}{c_t}\right)^{-\gamma} R_{t+1} - 1\right)\right] = 0,$$

  which can be used as a moment to estimate (β, γ).

Definition of GMM estimator

- Moment conditions

$$E[\underbrace{g(w_i, \theta_0)}_{(q\times 1)}] = 0$$

  where $\theta \in \mathbb{R}^k$. Assume $k \leq q$ (more on later)
- Sample moments $g_N(\theta) = N^{-1} \sum_i g(w_i, \theta)$
- The GMM estimator

$$\hat{\theta} = \arg\min_{\theta} Q_N(\theta), \;\; Q_N(\theta) = \underbrace{g_N(\theta)'}_{(1\times q)} \underbrace{W_N}_{(q\times q)} \underbrace{g_N(\theta)}_{(q\times 1)}$$

  - W_N: weight matrix, $W_N \xrightarrow{p} W$ (coverge in probability)

How "generalized"?

- If $k = q$ (**just-identified** case), can solve an equation

$$\frac{1}{N}\sum_i g(w_i, \theta) = 0$$

  because (# of equations) = (# of parameters).
  - This is original method of moments.

- If $q \geq k$ (**over-identified** case), this is no longer the case.
  - Try to find $\hat{\theta}$which makes the sample moment as close to as 0.
  - Weighting matrix W_N does matter in estimation.

Asymptotic Variance

- Under suitable assumptions (see textbook), the GMM estimator is asymptotic normal

$$\begin{aligned} \sqrt{N}(\hat{\theta} - \theta_0) &\xrightarrow{d} N(0, V_\theta), \\ V_\theta &= (DWD')^{-1}(DWSW'D')(DWD')^{-1} \end{aligned}$$

where

$$D = E\left[\frac{\partial g(w_i, \theta_0)}{\partial \theta}\right]$$

$$W = \mathrm{plim}_{N\to\infty} W_N$$

and S is given by

$$\frac{1}{\sqrt{N}} \sum_i g(w_i, \theta_0) \xrightarrow{d} N(0, S)$$

Identification

- Definition by words: Given the knowledge of the data generating process, you can get the true parameter θ_0.
- In this setting,
  - If you know the DGP of $\{w_i\}$ in population, then you can calculate $E[g(w_i, \theta)]$.
  - The parameter θ is identified if there exists a unique $\theta_0$that satisfies

$$E[\underbrace{g(w_i, \theta_0)}_{(q\times 1)}] = 0.$$

- Difficult to directly check this identification condition in many economic models.
- Let's focus on a simple model.

Example: Identification of OLS

- Moment condition for OLS:

$$E[x_i(y_i - \beta' x_i)] = 0,$$

  where x_i is $(k \times 1)$ vector.
- The moment condition can be written as

$$\underbrace{E[x_i x_i']}_{(k\times k)} \underbrace{\beta}_{(k\times 1)} = \underbrace{E[x_i y_i]}_{(k\times 1)}$$

- This equation has a unique solution of β if $E[x_i x_i']$ is invertible.
  - What does this mean in a more simple word?

Example: Identification of IV models

- $y_i = \beta' x_i + \epsilon_i$,
  - x_i: explanatory variables $(k \times 1)$ vector
  - z_i : instrumental variables. Assume $(l \times 1)$ vector.
  - $l > k$
- Moment conditions
  $$E[z_i(y_i - \beta' x_i)] = 0,$$
  which amounts to
  $$\underbrace{E[z_i x_i']}_{(l \times k)} \underbrace{\beta}_{(k \times 1)} = \underbrace{E[z_i y_i]}_{(l \times 1)}$$
- Identification conditions
  - Independence: $E[\epsilon_i z_i] = 0$
  - Relevance: The $(l \times k)$ matrix if of full column rank, i.e., $rank(E[z_i x_i']) = k$

- In a simple setting with $x_i = (1, p_i)$ and $z_i = (1, w_i)$, the matrix is

$$\begin{pmatrix} 1 & E[w_i] \\ E[p_i] & E[p_i w_i] \end{pmatrix}$$

and

- The determinant of the matrix is not 0 iff

$$Cov(p_i, w_i) = E[p_i w_i] - E[p_i]E[w_i] \neq 0$$

  - There should be correlation between p_i (endogenous variable) and w_i (insrumental variable).

Motivation: Why production function?

- Production function is a key element in many economic models.
  - Production function gives cost function (duality)
- Interested in distribution of productivity in economy
  - Is allocation of inputs (e.g., inputs, labor) efficient?
  - How do policies (e.g., trade liberalization) affects productivity distribution?
- What is the determinants of productivity?
  - Learning-by-doing, management, R&D, etc...

Data

- Panel data of firms over time
  - indexed by firm j and time t
  - Typically unbalanced-panel (due to entry/exit).
- Output:
  - Ideally phisical unit of outputs.
  - Revenue (in most cases),
    - contain price information.
  - Value-added (= revenue - material costs)
- Inputs: Labor, capital, materials
  - We typically observe these in expenditures, which we need to correct.
  - Quality of inputs (such as labor) should be controlled, though difficult.

Typical Data Sources

- United States:
  - Census of manufacturing
    - Plant-level, every 5 years.
  - COMPUSTAT
    - Balance sheet & financial information for traded firms.
  - Regulated industry (e.g., electricity)
- Japan:
  - Census of Manufacturing
  - Basic Survey of Japanese Business Structure and Activity
  - Nikkei-NEEDS
    - Japanese version of COMPUSTAT
  - Annual Survey of Local Public Firms
    - Public water service, public hospitals, etc.
- Other countries:
  - Census in Chili, Columbia, Slovakia
  - Encuesta Sobre Estrategias Empresariales (ESEE) survey in Spain

Econometric Issues

- Suppose we estimate Cobb-Douglas production function

$$
\begin{aligned}
Q_{it} &= L_{it}^{\beta_1} K_{it}^{\beta_2} \exp(\epsilon_{it}) \\
\Longleftrightarrow \log Q_{it} &= \beta_1 \log L_{it} + \beta_2 \log K_{it} + \epsilon_{it}
\end{aligned}
$$

  - Q_{it}: output, L_{it} : labor, K_{it}: capital
  - ϵ_{it}: unobserved shock (including productivity)

1. Endogneity of inputs
   - Firms choose input based on "unobserved" productivity ϵ_{it}
2. Sample selection due to entry/exit
   - Firms that are relatively efficient can survive.
3. Measurement of output (and input)
   - Revenue data contains the demand & price information.

Plan of Lectures

- I mainly focus on the first two issues.

1. Panel data approach
    1.1 Arellano and Bond (1991), Blundell and Bond (1998)
2. Proxy approach (& correction of sample selection)
    2.1 Olley and Pakes (1996), Levinsohn and Petrin (2003), Ackerberg et al (2015), Gandhi et al (2017)
3. Applications
    3.1 Measurement of markup (Olley and Pakes 1996)
    3.2 Reallocation analysis (DeLoecker and Warzynski 2013)

- Survey: Ackerberg et. al. Handbook of Econometrics

Model

- Gross Production function $Y_{jt} = F(L_{jt}, K_{jt}, M_{jt})e^{\omega_{jt}}e^{\epsilon_{jt}}$

$$y_{jt} \quad = \quad \ln F(L_{jt}, K_{jt}, M_{jt}) + \omega_{jt} + \epsilon_{jt} \tag{1}$$

  - Y_{jt}: Gross output, L_{jt}: labor, K_{jt}: capital, M_{jt}: intermediate input
  - lower letters in logarithmic form.

- ω_{jt} : unobserved productivity. Source of endogeneity.

- ϵ_{jt} : ex-post shock (or measurement error). This is realized after firm's decision.

- Assume Cobb-Douglas function $F(L_{jt}, K_{jt}, M_{jt}) = L_{jt}^{\alpha}K_{jt}^{\beta}M_{jt}^{\gamma}$.
  - Note: Can extend to other functional forms (CES, Trans-log, etc).

Instrumental variables for inputs?

- Regression equation under Cobb-Douglas

$$y_{jt} = \alpha l_{jt} + \beta k_{jt} + \gamma m_{jt} + \underbrace{\omega_{jt} + \epsilon_{jt}}_{u_{jt}} \tag{2}$$

- Consider instrumental variables z_{jt} which satisfies

$$E[u_{jt}|z_{jt}] = 0, Cov(z_{jt}, x_{jt}) \neq 0, x_{jt} \in \{l_{jt}, k_{jt}, m_{jt}\}$$

- Potential candidate: Input prices. Does this work?
- Issue 1: No cross-sectional variation
  - Typically the variation is at the market level.
- Issue 2: Even if we observe variation at firm-level, this variation may come from the difference of unobserved productivity (not price taker).

Fixed Effects Estimator

- Assume that unobserved productivity is constant over time: $\omega_{jt} = \omega_j$

$$y_{jt} = \alpha l_{jt} + \beta k_{jt} + \gamma m_{jt} + \omega_j + \epsilon_{jt} \tag{3}$$

- If we eliminate ω_j, we can consistently estimate coefficients (α, β, γ).
- We can use the first-difference estimator / within-transformation.

Review on Fixed Effects Estimator

- Regression equation with fixed effects

$$y_{jt} = \theta X_{jt} + \omega_j + \epsilon_{jt}$$

  where X_{jt} is correlated with ω_j but uncorrelated with ϵ_{jt}.

- Approach 1: Taking the first-difference $\Delta y_{jt} = y_{jt} - y_{j,t-1}$

$$\Delta y_{jt} = \theta \Delta X_{jt} + \Delta \epsilon_{jt}$$

- Approach 2: Within transformation: $\tilde{y}_{jt} = y_{jt} - \frac{1}{T}\sum_{t=1}^{T} y_{jt}$

$$\tilde{y}_{jt} = \theta \tilde{X}_{jt} + \tilde{\epsilon}_{jt}$$

Issues with Fixed Effects Estimator

1. Is productivity constant over time?

2. Magnify measurement error of inputs (especially capital k_{jt}).
   - See, e.g., Griliches and Hausman (1985, JOE), Griliches and Mairesse (1995, NBER)

Dynamic Panel Data Approach

- Rough idea: Use panel structure of data to create instruments.
- Arellano and Bond (1995), Blundell and Bond (1998, 1999)
- Productivity process

$$\omega_j + \omega_{jt}$$

  - ω_j: fixed effect
  - Both ω_j and ω_{jt} affects input choices.
  - Assume ω_{jt} is not auto-correlated: $Cov(\omega_{jt}, \omega_{jt'}) = 0$ for $t \neq t'$
    - We relax this later.

Arellano and Bond (1991)

- 1st-difference equation

$$\Delta y_{jt} = \alpha \Delta l_{jt} + \beta \Delta k_{jt} + \underbrace{\Delta \omega_{jt} + \Delta \epsilon_{jt}}_{\equiv \Delta u_{jt}} \tag{4}$$

- $\Delta \omega_{jt}$ is correlated with Δl_{jt} and Δk_{jt} so we need IV.
- Arellano and Bond (1995) suggests that using lagged-values for IV

$$\{l_{j,t-\tau}, k_{j,t-\tau}\}_{\tau \geq 2}$$

  - Independence? :
  - Relevance? :
- GMM estimation by moment conditions

$$\mathbb{E}\left[\Delta u_{jt} | \{l_{j,t-\tau}, k_{j,t-\tau}\}_{\tau \geq 2}\right] = 0 \tag{5}$$

- Issue 1: Concerns for weak IV.
- Issue 2: No auto-correlation on ω_{jt}

System GMM by Blundell and Bond (1998, 1999)

- Consider both *level* and *difference* equations

$$y_{jt} = \alpha l_{jt} + \beta k_{jt} + \underbrace{\omega_j + \omega_{jt} + \epsilon_{jt}}_{\equiv u_{jt}} \quad (6)$$

$$\Delta y_{jt} = \alpha \Delta l_{jt} + \beta \Delta k_{jt} + \Delta \omega_{jt} + \Delta \epsilon_{jt} \quad (7)$$

- Blundell and Bond propose using lagged differences as instrument for *level* equation

$$\mathbb{E}\left[u_{jt} | \{\Delta l_{j,t-\tau}, \Delta k_{j,t-\tau}\}_{\tau \geq l}\right] = 0$$

- Idea: past *changes* in inputs are due to past ω_{jt} and ϵ_{jt} (not fixed effect!).
- Given adjustment costs in inputs, the current level of inputs is likely to be correlated with past levels.
- The available lags depend on the serial correlation of ω_{jt}
  - $l = 1$ if ω_{jt} is not serially correlated.
  - $l = 2$ if ω_{jt} is AR(1).

- Combine two sets of moment conditions in a system GMM framework.

$$\mathbb{E}\left[\Delta u_{jt} | \{l_{j,t-\tau}, k_{j,t-\tau}\}_{\tau\geq 2}\right] = 0 \tag{8}$$

$$\mathbb{E}\left[u_{jt} | \{\Delta l_{j,t-\tau}, \Delta k_{j,t-\tau}\}_{\tau\geq 2}\right] = 0 \tag{9}$$

Serial correlation in ω_{jt} and no FE ω_j

- Productivity process:

$$\omega_{jt} = \rho\omega_{j,t-1} + \xi_{jt}, \xi_{jt} \textit{ is IID}$$

- Quasi-difference equation

$$\begin{aligned} y_{jt} &= \alpha l_{jt} + \beta k_{jt} + \omega_{jt} + \epsilon_{jt} \\ \rho y_{jt-1} &= \alpha\rho l_{jt-1} + \beta\rho k_{kt-1} + \rho\omega_{jt-1} + \rho\epsilon_{jt-1} \end{aligned}$$

  so that

$$(y_{jt} - \rho y_{jt-1}) = \alpha(l_{jt} - \rho l_{jt-1}) + \beta(k_{jt} - \rho k_{jt-1}) + (\xi_{jt} + \epsilon_{jt} - \rho\epsilon_{jt-1})$$

  where $\tilde{x}_{jt} = x_{jt} - \rho x_{jt-1}$.
- Use $(l_{jt-1}, k_{jt-1}, y_{j,t-2})$ as IVs to form moment conditions.

Both Serial correlation in ω_{jt} and FE ω_j

- Blundell and Bond (2000, ER)
- Productivity process:

$$\omega_j + \omega_{jt}$$

where $\omega_{jt} = \rho\omega_{j,t-1} + \xi_{jt}$ and ξ_{jt} is IID.

- Quasi-difference equation

$$y_{jt} = \alpha l_{jt} + \beta k_{jt} + \omega_j + \omega_{jt} + \epsilon_{jt}$$
$$\rho y_{jt} = \alpha\rho l_{jt-1} + \beta\rho k_{kt-1} + \rho\omega_j + \rho\omega_{jt} + \rho\epsilon_{jt}$$

so that

$$\tilde{y}_{jt} = \alpha\tilde{l}_{jt} + \beta\tilde{k}_{jt} + (1-\rho)\omega_j + (\xi_{jt} + \epsilon_{jt} - \rho\epsilon_{jt})$$

where $\tilde{x}_{jt} = x_{jt} - \rho x_{jt-1}$.

- We apply system GMM to this equation.