Econometrics **August 18, 2026**

Dynamic Panel: GMM

Instructor: Kalyan Kolukuluri *Content: Lecture Notes*

Contents

1 What is GMM? (The Big Picture)

Imagine you are a detective trying to estimate an unknown quantity using clues from data.

1.1 Method of Moments (MM)

In statistics, a **"moment"** is simply a summary statistic, such as the mean, variance, or covariance.

- **Population Moment Condition:** Theory tells us that in the population, a specific expectation equals zero:
$$\mathbb{E}\left[Y_i - \mu\right] = 0 \tag{1}$$
- **Sample Moment Condition:** We replace the theoretical expectation $\mathbb{E}[\cdot]$ with the sample average $\frac{1}{N}\sum_{i=1}^{N}(\cdot)$:
$$\frac{1}{N}\sum_{i=1}^{N}(Y_i - \hat{\mu}) = 0 \implies \hat{\mu} = \overline{Y} \tag{2}$$
- **Rule of Thumb:** If you have 1 unknown parameter, you need 1 moment condition to solve for it.

1.2 Generalized Method of Moments (GMM)

What if you have **more clues (moment conditions) than unknown parameters**? For example, you have 10 moment conditions to estimate just 1 unknown parameter.

Because real-world data contains random noise, you cannot make all 10 equations equal zero simultaneously. Hansen (1982) developed **Generalized Method of Moments (GMM)** as a unified framework that:

1. Takes all available moment conditions (clues).
2. Weighs the most reliable/informative clues more heavily.
3. Finds the single parameter value that gets all equations as close to zero as possible.

2 A Simple Concrete Example: Tracking Firm Sales

To see why GMM is necessary, let's follow a very simple real-world problem: **predicting firm sales over time**.

Suppose we observe N firms (e.g., $N = 1,000$) over a short period of T years (e.g., $T = 5$). We want to measure sales persistence—how much last year's sales influence current sales:

$$\text{Sales}_{it} = \alpha\,\text{Sales}_{i,t-1} + \eta_i + \varepsilon_{it} \tag{3}$$

Color Legend for Model Parameters:

- α (Persistence Parameter): The unknown value we want to estimate (e.g., $\alpha = 0.8$ means 80% of last year's sales carry over).
- η_i (Firm-Level Fixed Effect): Unobserved, permanent characteristics of firm i (e.g., management quality, brand value).
- ε_{it} (Idiosyncratic Shock): Unpredictable year-to-year noise (e.g., unexpected weather or sudden economic recessions).

3 Why Standard Estimators Fail

Before learning GMM, we must understand why standard regression methods fail for dynamic panel data.

3.1 Pooled OLS Fails (Upward Bias)

If we run a standard linear regression ignoring η_i:

- Well-managed firms (η_i high) always have higher sales both last year ($\text{Sales}_{i,t-1}$) and this year (Sales_{it}).
- OLS confuses good management quality with sales momentum.
- **Result:** Pooled OLS overestimates α (**Upward Bias**).

3.2 Fixed Effects (Within Estimator) Fails (Downward Bias)

To eliminate η_i, we might subtract firm averages across time. However, Nickell (1981) proved that in short time periods (small T), subtracting averages creates an artificial negative correlation between the transformed lagged variable and the transformed error term.

- **Result:** Fixed Effects underestimates α (**Downward Bias**).

The Golden Rule (Bond, 2002): A valid estimate of α must lie strictly between the downward-biased Fixed Effects estimate and the upward-biased OLS estimate:

$$\hat{\alpha}_{\text{Fixed Effects}} < \alpha_{\text{True}} < \hat{\alpha}_{\text{Pooled OLS}} \tag{4}$$

4 Difference GMM (Arellano and Bond, 1991)

4.1 How Difference GMM Works

To eliminate η_i without creating Nickell bias, Arellano and Bond (1991) proposed taking **first differences** (year-to-year changes):

$$\Delta\text{Sales}_{it} = \alpha\,\Delta\text{Sales}_{i,t-1} + \Delta\varepsilon_{it} \tag{5}$$

Where:

$$\Delta\text{Sales}_{it} = \text{Sales}_{it} - \text{Sales}_{i,t-1} \tag{6}$$

$$\Delta\text{Sales}_{i,t-1} = \text{Sales}_{i,t-1} - \text{Sales}_{i,t-2} \tag{7}$$

$$\Delta\varepsilon_{it} = \varepsilon_{it} - \varepsilon_{i,t-1} \tag{8}$$

Notice that $\eta_i - \eta_i = 0$, so the firm fixed effect completely drops out!

4.2 The Instrument Solution

Because $\Delta\text{Sales}_{i,t-1}$ contains $\text{Sales}_{i,t-1}$, it is correlated with $\varepsilon_{i,t-1}$ in the error term $\Delta\varepsilon_{it}$.

To fix this endogeneity, Arellano and Bond used **lagged levels** from two or more periods ago as Instrumental Variables (IVs):

$$\text{Moment Condition:}\quad \mathbb{E}\left[\text{Sales}_{i,t-2}\cdot(\Delta\varepsilon_{it})\right] = 0 \tag{9}$$

Why this instrument works:

1. **Relevance:** Past sales level $\text{Sales}_{i,t-2}$ predicts last year's sales change $\Delta\text{Sales}_{i,t-1}$.
2. **Exogeneity:** Past sales level $\text{Sales}_{i,t-2}$ is completely unaffected by today's shock ε_{it} and yesterday's shock $\varepsilon_{i,t-1}$.

4.3 When Does Difference GMM Apply?

Difference GMM is designed for panel data settings where:

- Large number of individuals/firms (N is large).
- Small number of time periods (T is small, e.g., $T = 3$ to 6).
- Low to moderate persistence in the dependent variable ($\alpha < 0.7$).

4.4 The Weak Instrument Failure

When sales are **highly persistent** ($\alpha \to 1$), today's sales level is nearly identical to yesterday's sales level. Consequently, past levels contain almost no information about future *changes*. Past levels become **weak instruments**, causing Difference GMM to break down and yield downward-biased estimates (Blundell and Bond, 1998).

5 System GMM (Blundell and Bond, 1998)

5.1 How System GMM Works

To solve the weak instrument problem, Blundell and Bond (1998) introduced **System GMM**. Instead of estimating a single equation, System GMM builds a combined **system of two equations**:

$$\begin{cases} \textbf{1. Difference Eq:} & \Delta \text{Sales}_{it} = \alpha\, \Delta \text{Sales}_{i,t-1} + \Delta \varepsilon_{it} \quad \text{instrumented by } \text{Sales}_{i,t-2} \\ \textbf{2. Level Eq:} & \text{Sales}_{it} = \alpha\, \text{Sales}_{i,t-1} + \eta_i + \varepsilon_{it} \quad \text{instrumented by } \Delta \text{Sales}_{i,t-1} \end{cases} \tag{10}$$

5.2 Why the System Trick Works

Even if sales levels are highly persistent, **changes in sales are not persistent**! Therefore, past sales changes ($\Delta\text{Sales}_{i,t-1}$) serve as highly informative, strong instruments for current sales levels ($\text{Sales}_{i,t-1}$).

5.3 When Does System GMM Apply?

System GMM should be applied when:

- The sample has large N and small T.
- The dependent variable is **highly persistent** ($\alpha \geq 0.7$, e.g., GDP per capita, firm capital, long-run sales).
- Fixed effect variance (η_i) is large relative to shock variance (ε_{it}).

5.4 Why System GMM MUST Be Applied

1. **Eliminates Weak Instrument Bias:** When α is close to 1, Difference GMM fails. System GMM restores strong identification.

2. **Dramatically Increases Efficiency:** By combining moment conditions from both levels and differences, System GMM uses more information, yielding smaller standard errors.
3. **Satisfies Bounds Test:** It produces estimates that properly lie within the valid range ($\hat{\alpha}_{\text{FE}} < \hat{\alpha}_{\text{Sys-GMM}} < \hat{\alpha}_{\text{OLS}}$).

6 Summary Comparison Table

Table 1: Comparison of Dynamic Panel Estimators

Estimator	Fixed Effects (η_i)	Performance when $\alpha \geq 0.7$	When to Use?
Pooled OLS	Ignores η_i	Severe Upward Bias	Never for dynamic panel data
Fixed Effects	Subtracts time mean	Severe Downward Bias (Nickell Bias)	Only when T is very large ($T > 30$)
Difference GMM	Takes first difference	Downward Bias (Weak IVs)	Large N, small T, low persistence ($\alpha < 0.7$)
System GMM	Combines differences & levels	**Unbiased & Highly Efficient**	Large N, small T, high persistence ($\alpha \geq 0.7$)

7 Diagnostic Tests in GMM

When estimating GMM models in practice (e.g., using ‘xtabond2‘ in Stata or ‘pgmm‘ in R), two diagnostic tests are essential:

1. **Arellano-Bond Test for Autocorrelation:**
   - AR(1) in differences is expected ($p < 0.05$).
   - AR(2) in differences **must be absent** ($p > 0.05$) to ensure instruments are valid.
2. **Hansen / Sargan Test of Overidentifying Restrictions:**
   - Tests whether the instruments as a group are exogenous ($p > 0.05$ indicates valid instruments).

References

Arellano, M., & Bond, S. (1991). Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *The Review of Economic Studies*, 58(2), 277–297.

Blundell, R., & Bond, S. (1998). Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics*, 87(1), 115–143.

Bond, S. R. (2002). Dynamic panel data models: a guide to micro data methods and practice. *Portuguese Economic Journal*, 1(2), 141–162.

Hansen, L. P. (1982). Large sample properties of Generalized Method of Moments estimators. *Econometrica*, 50(4), 1029–1054.

Nickell, S. (1981). Biases in dynamic models with fixed effects. *Econometrica*, 49(6), 1417–1426.