05_GMM

QuantFit Estimator Standard Operating Procedure

SOP: Generalized Method of Moments (GMM)

Moment-based estimator including Arellano-Bond and Blundell-Bond panel GMM

=> Use GMM for dynamic panels where the lagged dependent variable is a regressor.

1. Purpose

GMM minimises a quadratic form in sample moment conditions E[Z'epsilon] = 0. Difference GMM (Arellano-Bond) first-differences the model and uses lagged levels as instruments. System GMM (Blundell-Bond) augments with the equation in levels, using lagged differences as additional instruments - more efficient when persistence is high.

2. When to use this estimator

Dynamic panels: Y_it = rho Y_{i,t-1} + X_it beta + alpha_i + epsilon_it.

Endogenous regressors with no exogenous instruments outside the panel.

Short T, large N panels (typical in microeconomic applications).

3. Required data structure

Panel with N entities x T periods (T >= 3 for AB, >= 4 ideal for BB).

At least one lagged dependent variable as regressor.

No requirement for external instruments - uses internal lags.

4. Mathematical formulation

Sample moment vector and GMM criterion:

g_N(theta) = (1/N) Sum Z_i' epsilon_i(theta)

theta?_GMM = argmin_theta g_N(theta)' W g_N(theta)

Two-step optimal W = (Avar(g_N))^-1

5. Pre-estimation diagnostics

Stationarity tests on the dependent variable.

Hausman or Sargan/Hansen for instrument validity.

Sargan J test of over-identifying restrictions.

Arellano-Bond AR(1) and AR(2) tests on differenced residuals - AR(2) should not reject.

6. Estimation procedure

First-difference the equation to remove alpha_i (AB) or stack difference +

level (BB).

Build instrument matrix Z from lagged levels (AB) and lagged differences (BB).

Run one-step GMM with weighting matrix W1.

Use one-step residuals to update W to optimal W2; run two-step GMM.

Compute Sargan-Hansen J and AR(1)/AR(2) tests; report Windmeijer-corrected SE.

7. Output produced

8. Output interpretation

rho? on lagged Y measures persistence; |rho?| < 1 implies stationary autoregression.

AR(2) p >= 0.10 is required for instrument validity; rejection invalidates GMM.

Sargan-Hansen J p >= 0.05 supports the instrument set.

Number of instruments should not exceed the number of entities (avoid weak-instrument problem).

9. Post-estimation diagnostics

Always report AR(1) and AR(2) test results.

Sargan-Hansen J in over-identified setting.

Windmeijer correction in two-step SE.

Discuss instrument count vs N.

10. Common pitfalls

Too many instruments overfit and weaken Hansen J.

AR(2) rejection is fatal - return to OLS-based dynamic models or reduce the lag set.

BB (system GMM) requires stationarity of initial conditions.

Two-step SE without Windmeijer correction are severely downward biased.

11. Reporting checklist

Estimator: AB or BB; one-step or two-step.

Number of entities N, periods T, instruments.

Hansen J statistic and p-value.

AR(1) and AR(2) z-statistics and p-values.

Windmeijer-corrected SE if two-step.

12. References

Arellano, M., Bond, S. (1991). Some tests of specification for panel data. Review of Economic Studies.

Blundell, R., Bond, S. (1998). Initial conditions and moment restrictions in dynamic panel data.

Roodman, D. (2009). How to do xtabond2: An introduction to difference and system GMM. Stata Journal.

Field | Meaning

coefficients | GMM slopes including lagged Y

metadata['jStat'] / jStatP / jStatDf | Sargan-Hansen over-id test

metadata['ar1Z'] / ar1P | AR(1) test (expected to reject - first differences)

metadata['ar2Z'] / ar2P | AR(2) test (must NOT reject)

metadata['nInstr'] / N | Instrument and entity counts

metadata['abStep'] | AB / BB and one- / two-step indicator