Economics Bulletin

Volume 45, Issue 4

Default and collapsed instruments in dynamic panel GMM: a Monte Carlo comparison

John Levendis
Tulane University

Abstract

The Arellano-Bond type dynamic panel estimators are known to suffer from over-instrumentation. Using Monte Carlo simulations, we compare Difference and System GMM with both default and collapsed instrument sets across varying numbers of countries and time periods typical in macroeconomic analysis. Collapsing consistently outperforms default instrumentation, providing more accurate estimates, especially as T increases. System GMM with collapsed instruments performs best overall, especially for endogenous variables. These findings strongly support using collapsed instruments for the analysis of macroeconomic panels, regardless of the number of countries or time periods.

Citation: John Levendis, (2025) "Default and collapsed instruments in dynamic panel GMM: a Monte Carlo comparison", *Economics Bulletin*, Volume 45, Issue 4, pages 1915-1925
Contact: John Levendis - jlevendis@tulane.edu.
Submitted: April 18, 2025. **Published:** December 30, 2025.

1. Introduction

Many macroeconomic datasets have a panel structure. The World Bank, for example, publishes annual macroeconomic data for approximately 150 countries. The Federal Reserve publishes quarterly data for the fifty US states.

Economic processes often exhibit strong autocorrelation—this year's GDP is similar to last year's GDP, plus or minus an adjustment—making a lagged dependent variable a necessary component of most econometric models. Heterogeneity across countries also necessitates a fixed-effects term in these models.

For these reasons, it is common to estimate dynamic panel models with both exogenous and endogenous covariates:

$$y_{it} = \rho y_{i,t-1} + \beta x_{it} + \gamma z_{it} + \eta_i + \epsilon_{it} \quad (1)$$

To eliminate the fixed effects, researchers typically estimate the model in first differences:

$$\Delta y_{it} = \rho \Delta y_{i,t-1} + \beta \Delta x_{it} + \gamma \Delta z_{it} + \Delta \epsilon_{it} \quad (2)$$

While differencing eliminates the fixed effects and induces stationarity, it does so at a cost: the lagged and differenced dependent variable, $\Delta y_{i,t-1}$, becomes correlated with the differenced error term, $\Delta \epsilon_{it}$ (Nickell, 1981).

In a series now famous papers, Arellano, Bond, Blundell, and Bover developed Generalized Method of Moments (GMM) estimators that address these challenges by constructing instruments from within the dataset itself. Arellano and Bond (1991) introduced the Difference GMM estimator for dynamic panel data. As the name implies, Difference GMM estimates equation (2), the equation in first-differences, but it uses lagged levels as instruments. The System GMM estimator is due to the combined work of Arellano and Bover (1995) and Blundell and Bond (1998). Arellano and Bover (1995) introduced the idea of combining the levels and differences equations—equations (1) and (2)—into a system to be estimated simultaneously. Whereas the level equation is instrumented with lagged differences, the difference equation is instrumented by lagged levels.

These GMM estimators have become remarkably common in applied economics research. As of April 2025, Arellano and Bond (1991) has been cited over 45,000 times, Blundell and Bond (1998) more than 33,000 times, and Arellano and Bover (1995) over 26,000 times. Even methodological papers about implementation and refinement, such as David Roodman's papers on the Stata command xtabond2 (2009a) and instrument proliferation (2009b) have been cited 14,000 times and 6,800 times, respectively. Such extraordinary numbers suggest that these GMM estimators have become default tools in applied research, often implemented without critically assessing their appropriateness for specific applications. For textbook treatments of these methods, see Levendis (2023).

Despite their widespread adoption, these methods are not without limitations. When the time dimension (T) grows moderately large, the number of instruments can increase dramatically, potentially leading to overfitting. As Roodman (2009b) demonstrated, instrument proliferation can re-introduce the very bias they were designed to eliminate.

Proper instrumentation requires a delicate balancing act. On the one hand, if you do not use IVs at all, then you cannot get rid of endogeneity, and the coefficient estimates are guaranteed to be biased. On the other hand, using too many instruments is a problem, too. This is most easily seen in two-stage IV estimators. In the first stage, endogenous X is regressed on the IVs and the remaining control variables (but not Y). A fitted value of X is then constructed. This $\hat{X}$ should be purged of its correlation with Y. But if we have too many IVs then we could have a perfectly fitting first stage regression. This, in turn, implies that $\hat{X}$ is identical to the original X. In this case, the endogeneity problem hasn't been fixed at all. The variable has simply been renamed from X to $\hat{X}$! Thus, the balancing act. We need to have just the right number of instruments, not too few and not too many. How many should we use?

To address instrument proliferation, Roodman (2009b) proposed "collapsing" the instrument matrix: combining moment conditions across time periods and thereby reducing the number of instruments.

Despite the enormous influence of these estimators, there remains considerable uncertainty about which variant performs best under different conditions. In this paper, we conduct Monte Carlo simulations to compare four estimators: Difference GMM and System GMM, each implemented with either default or collapsed instrument sets. Our simulation design focuses on sample dimensions typical in macroeconomic research: between 10 and 200 countries with 20 to 50 time periods. By systematically varying these dimensions, we provide practical guidance for applied researchers on the appropriate estimator choice for different panel structures.

This research addresses a fundamental paradox in applied econometrics: dynamic GMM estimators are most needed precisely in the challenging empirical contexts where they perform poorly with default instruments. AB estimators are theoretically designed for short, wide panels ($T \approx$ 5-10, $N \approx$ 100+), yet they are frequently applied in macroeconomic studies with longer time dimensions (T = 15-30+), smaller cross-sections (regional studies, developing country samples), and multiple endogenous variables that exacerbate instrument proliferation. Researchers face a difficult trade-off: they have limited sample sizes but complex endogeneity problems, long time series with persistent regressors requiring many instrument lags, and multiple potentially endogenous controls. In these non-ideal but common applications, standard default settings often fail, yet researchers lack systematic guidance on which estimator variants perform best under such constraints.

While Roodman (2009b) introduced the concept of instrument proliferation and proposed collapsing as a solution, his Monte Carlo analysis focused exclusively on System GMM variants with different instrument configurations. Our study provides a more comprehensive framework by systematically comparing both Difference GMM and System GMM estimators alongside their collapsed instrument counterparts. Since practitioners typically rely on standard software implementations, our systematic comparison across both major GMM estimator families and realistic sample dimensions offers empirical evidence for establishing better default practices in dynamic panel estimation.

Next, we review instrument proliferation, and explain "collapsing."

2. Proliferation of Moments

To illustrate how the number of instruments can grow during GMM estimation, we examine a simple case with $T = 4$ time periods and $p = 2$ endogenous variables (lagged Y and Z).

2.1. Difference GMM with Default Instruments

With differencing, $t = 1$ is lost in creating the first-difference, leaving $t = 2$ and 3 for estimation. For each time period, we create instruments from all valid lags:

For lagged Y at $t = 2$:

$$E[y_{i0}\Delta\varepsilon_{i2}] = 0$$

Since y_{i0} occurs before the differenced error term $\Delta\epsilon_{i2} = (\epsilon_{i2} - \epsilon_{i1})$, it is uncorrelated with the error term and is thereby a valid instrument.

For lagged Y at $t = 3$:

$$E[y_{i0}\Delta\varepsilon_{i3}] = 0$$
$$E[y_{i1}\Delta\varepsilon_{i3}] = 0$$

Similarly for endogenous Z at $t = 2$:

$$E[z_{i0}\Delta\varepsilon_{i2}] = 0$$

And for endogenous Z at $t = 3$:

$$E[z_{i0}\Delta\varepsilon_{i3}] = 0$$
$$E[z_{i1}\Delta\varepsilon_{i3}] = 0$$

Difference GMM with default instrumentation creates 6 distinct instruments (columns in the instrument matrix).

2.2. Difference GMM with Collapsed Instruments

With collapsed instruments, on the other hand, the instrument count is much smaller. We create a single instrument for each lag by summing across time periods:

For lagged Y, the instrument set is:

$$E[y_{i0}(\Delta\varepsilon_{i2} + \Delta\varepsilon_{i3})] = 0$$
$$E[y_{i1}\Delta\varepsilon_{i3}] = 0$$

For endogenous Z:

$$E[z_{i0}(\Delta\varepsilon_{i2} + \Delta\varepsilon_{i3})] = 0$$
$$E[z_{i1}\Delta\varepsilon_{i3}] = 0$$

Collapsing in Difference GMM thereby creates only 4 instruments.

2.3. System GMM with Default Instruments

System GMM adds level equations with differenced instruments to the difference equations. The number of instruments are therefore the original six instruments from the Difference GMM matrix, plus the following four instruments for the level equation:

$$E[\Delta y_{i1}(\eta_i + \varepsilon_{i2})] = 0 \quad \text{(at } t = 2\text{)}$$
$$E[\Delta y_{i2}(\eta_i + \varepsilon_{i3})] = 0 \quad \text{(at } t = 3\text{)}$$
$$E[\Delta z_{i1}(\eta_i + \varepsilon_{i2})] = 0 \quad \text{(at } t = 2\text{)}$$
$$E[\Delta z_{i2}(\eta_i + \varepsilon_{i3})] = 0 \quad \text{(at } t = 3\text{)}$$

2.4. System GMM with Collapsed Instruments

Collapsed instruments in System GMM combine both difference and level equations, each with their own collapsed instrument structure. First, we include all 4 instruments from the collapsed Difference GMM approach described earlier. Then, we add collapsed instruments for the level equations:

$$E[(\Delta y_{i1} + \Delta y_{i2})((\eta_i + \varepsilon_{i2}) + (\eta_i + \varepsilon_{i3}))] = 0$$
$$E[(\Delta z_{i1} + \Delta z_{i2})((\eta_i + \varepsilon_{i2}) + (\eta_i + \varepsilon_{i3}))] = 0$$

These collapsed level equations combine all time periods into a single instrument for each endogenous variable. This yields 6 instruments total (4 from collapsed difference equations plus 2 from collapsed level equations).

2.5. Summary and Comparison

In general, the instrument count (the number of columns in the instrument matrix) for each estimator is:

Method	**Columns in the Instrument Matrix**
Difference GMM (default):	$\frac{1}{2}p(T-1)(T-2)$
Difference GMM (collapsed):	$p(T-2)$
System GMM (default):	$\frac{1}{2}p(T-2)(T+1)$
System GMM (collapsed):	$p(T-1)$

For the default instrument sets, the number of columns in the instrument matrix (i.e. what is commonly referred to as the "instrument count") grows quadratically with T. With moderate number of time periods ($T = 20$ to 50), the default GMM estimators create instrument sets which quickly exceed the effective sample size. For the collapsed instrument set, on the other hand, the number of columns only grows linearly with T.

3. The Data Generating Process

In our simulations, $\rho = 0.7$ represents the autoregressive parameter, $\beta = 1.0$ is the coefficient on the exogenous variable x, and $\gamma = 0.5$ is the coefficient on the endogenous variable z. The model incorporates country-specific fixed effects (η_i) and idiosyncratic errors (ϵ_{it}) drawn from standard normal distributions. The endogenous variable z is constructed to be correlated with the contemporaneous error term, creating the endogeneity problem that GMM methods are designed to address.

The simulation incorporates an endogenous regressor z_{it} that exhibits properties commonly encountered in empirical applications. The endogeneity is established through two distinct channels. First, the variable is correlated with the country-specific effects. Second, and more importantly, it is correlated with the contemporaneous idiosyncratic error term.

To generate data for z_{it} we set the initial value of z:

$$z_{i0} = 0.5\eta_i + \nu_{i0} \tag{3}$$

where $\nu_{i0} \sim \mathcal{N}(0, 4)$. For subsequent periods ($t = 1, 2, ..., T-1$), z_{it} follows a first-order autoregressive process with contemporaneous endogeneity:

$$z_{it} = 0.4 z_{i,t-1} + 0.3\varepsilon_{it} + \nu_{it} \tag{4}$$

The moderate autoregressive coefficient of 0.4 introduces persistence in the series while maintaining stationarity. ε_{it} is the contemporaneous idiosyncratic error term from the main equation and $\nu_{it} \sim \mathcal{N}(0, 4)$ represents additional random variation specific to the z process. The coefficient of 0.3 on ε_{it} controls the degree of contemporaneous endogeneity.

We simulate the GMM estimators across sample sizes representative of common macroeconomic panels: N = 10, 25, 50, 100, and 200. These values capture the range from small regional groupings (like the 10 ASEAN countries) to medium-sized panels (such as the 46 Sub-Saharan African nations) and large global samples (approaching the World Bank's comprehensive datasets with approximately 150-200 countries).

The time dimension varies from $T = 20$ to 50, reflecting common macroeconomic time series lengths. While this exceeds the theoretical optimum for AB estimators ($T \approx$ 5-10), it captures the reality that these methods are often applied in contexts where they are most needed but where default instruments perform poorly. The lower bound corresponds to post-2000 datasets frequently used in recent research, while the upper bound captures longer panels comparable to Federal Reserve state-level data that often extends back to the 1970s.

For each combination of N and T, we perform 1,000 Monte Carlo replications using Stata's `xtabond2` command to compare four estimators: Difference GMM and System GMM, each implemented with either default or collapsed instrument sets. This comprehensive design evaluates estimator performance across sample dimensions ($N = 10 - 200$, $T = 20 - 50$) typical of applied macroeconomic research

4. Results

Tables I-III present the Monte Carlo simulation results for our four estimators across different levels of N and T. We compare their ability in estimating three key parameters: the lagged dependent variable's coefficient (ρ=0.7), the exogenous variable's coefficient (β=1.0), and the endogenous variable's coefficient (γ=0.5).

Table I shows the estimates of the coefficient on the lagged dependent variable. The true value is 0.70. Difference GMM with default instrumentation displays an increasingly downward bias as T increases, especially with small N. For example, with $N = 10$ and $T = 50$, the estimate is 0.572 compared to the true value of 0.7. On the other hand, Difference GMM with a collapsed instrument set performs remarkably well across all sample sizes, with estimates consistently around 0.69-0.695 (slight downward bias) and smaller standard errors compared to the default version. System GMM with default instruments exhibits upward bias, particularly for moderate N (25-50). With $N = 50$ and $T = 30$, the estimate is 0.754, substantially above the true 0.70 value. System GMM with collapsed instruments produces estimates of 0.71 across almost all configurations, with small standard errors.

Table II shows the coefficient estimates on the exogenous variable. All estimators perform relatively well for this parameter, especially as N increases. The two estimators with default

Table I: Estimated Coefficients on Lagged Dependent Variable
(True Value = $\rho = 0.7$)

	T				T			
	20	30	40	50	20	30	40	50
	Difference GMM (default)				Difference GMM (collapsed)			
N								
10	0.650	0.619	0.603	0.572	0.691	0.692	0.696	0.696
	(0.119)	(0.149)	(0.177)	(0.192)	(0.035)	(0.027)	(0.027)	(0.024)
25	0.685	0.678	0.672	0.663	0.691	0.693	0.695	0.695
	(0.027)	(0.038)	(0.047)	(0.052)	(0.018)	(0.012)	(0.010)	(0.009)
50	0.690	0.690	0.687	0.685	0.692	0.694	0.694	0.695
	(0.011)	(0.014)	(0.018)	(0.022)	(0.013)	(0.008)	(0.007)	(0.006)
100	0.692	0.693	0.694	0.692	0.693	0.694	0.695	0.695
	(0.007)	(0.006)	(0.006)	(0.008)	(0.010)	(0.007)	(0.005)	(0.004)
200	0.694	0.695	0.695	0.695	0.693	0.695	0.695	0.695
	(0.005)	(0.003)	(0.003)	(0.003)	(0.007)	(0.005)	(0.004)	(0.003)
	System GMM (default)				System GMM (collapsed)			
N								
10	0.676	0.657	0.668	0.670	0.709	0.708	0.707	0.707
	(0.112)	(0.128)	(0.126)	(0.133)	(0.034)	(0.028)	(0.025)	(0.025)
25	0.745	0.741	0.739	0.738	0.710	0.709	0.709	0.710
	(0.023)	(0.021)	(0.021)	(0.018)	(0.019)	(0.015)	(0.013)	(0.012)
50	0.734	0.754	0.751	0.750	0.707	0.706	0.707	0.707
	(0.015)	(0.014)	(0.013)	(0.012)	(0.014)	(0.012)	(0.011)	(0.009)
100	0.720	0.727	0.735	0.749	0.704	0.705	0.704	0.705
	(0.010)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.008)	(0.007)
200	0.712	0.716	0.720	0.723	0.701	0.701	0.703	0.704
	(0.007)	(0.007)	(0.006)	(0.006)	(0.006)	(0.005)	(0.005)	(0.006)

Table II: Estimated Coefficients on Exogenous Variable
(True Value = $\beta = 1.0$)

	T				T			
	20	30	40	50	20	30	40	50
	Difference GMM (default)				Difference GMM (collapsed)			
N								
10	0.981	0.966	0.958	0.945	0.997	0.998	0.997	0.999
	(0.066)	(0.076)	(0.084)	(0.088)	(0.035)	(0.032)	(0.033)	(0.033)
25	0.997	0.993	0.990	0.985	0.998	0.998	0.999	0.999
	(0.021)	(0.023)	(0.024)	(0.025)	(0.020)	(0.016)	(0.013)	(0.013)
50	0.999	0.999	0.997	0.996	0.998	0.999	1.000	1.000
	(0.013)	(0.012)	(0.012)	(0.012)	(0.014)	(0.011)	(0.009)	(0.008)
100	0.999	1.000	1.000	1.000	0.998	0.999	1.000	1.000
	(0.009)	(0.007)	(0.006)	(0.006)	(0.010)	(0.008)	(0.007)	(0.006)
200	1.000	1.000	1.000	1.001	0.999	0.999	1.000	1.000
	(0.006)	(0.005)	(0.004)	(0.004)	(0.007)	(0.006)	(0.005)	(0.004)
	System GMM (default)				System GMM (collapsed)			
N								
10	0.970	0.962	0.968	0.972	0.988	0.993	0.991	0.993
	(0.059)	(0.063)	(0.060)	(0.061)	(0.035)	(0.036)	(0.031)	(0.028)
25	0.981	0.982	0.983	0.984	0.995	0.995	0.994	0.994
	(0.021)	(0.018)	(0.016)	(0.014)	(0.020)	(0.016)	(0.014)	(0.013)
50	0.986	0.979	0.979	0.980	0.997	0.996	0.997	0.997
	(0.015)	(0.012)	(0.011)	(0.010)	(0.015)	(0.012)	(0.011)	(0.009)
100	0.992	0.989	0.987	0.981	0.998	0.998	0.998	0.998
	(0.010)	(0.008)	(0.008)	(0.007)	(0.010)	(0.009)	(0.007)	(0.007)
200	0.995	0.994	0.992	0.990	0.998	0.999	0.999	0.999
	(0.007)	(0.006)	(0.005)	(0.005)	(0.007)	(0.006)	(0.005)	(0.005)

Table III: Estimated Coefficients on Endogenous Variable
(True Value = $\gamma = 0.50$)

	T				T			
	20	30	40	50	20	30	40	50
	Difference GMM (default)				Difference GMM (collapsed)			
N								
10	0.574	0.561	0.559	0.556	0.563	0.562	0.560	0.558
	(0.213)	(0.258)	(0.282)	(0.304)	(0.124)	(0.139)	(0.146)	(0.150)
25	0.566	0.568	0.567	0.563	0.570	0.567	0.566	0.564
	(0.070)	(0.083)	(0.094)	(0.100)	(0.032)	(0.034)	(0.036)	(0.040)
50	0.569	0.565	0.567	0.567	0.568	0.567	0.567	0.567
	(0.034)	(0.039)	(0.045)	(0.047)	(0.021)	(0.016)	(0.014)	(0.014)
100	0.569	0.566	0.568	0.567	0.569	0.569	0.568	0.567
	(0.018)	(0.019)	(0.020)	(0.022)	(0.015)	(0.011)	(0.010)	(0.008)
200	0.568	0.566	0.565	0.565	0.569	0.569	0.568	0.567
	(0.010)	(0.010)	(0.010)	(0.011)	(0.011)	(0.009)	(0.007)	(0.006)
	System GMM (default)				System GMM (collapsed)			
N								
10	0.579	0.563	0.587	0.610	0.555	0.543	0.546	0.555
	(0.251)	(0.282)	(0.308)	(0.329)	(0.149)	(0.179)	(0.170)	(0.171)
25	0.550	0.557	0.562	0.567	0.546	0.550	0.551	0.548
	(0.080)	(0.088)	(0.101)	(0.102)	(0.038)	(0.041)	(0.043)	(0.044)
50	0.548	0.548	0.551	0.552	0.542	0.547	0.548	0.551
	(0.039)	(0.044)	(0.048)	(0.050)	(0.025)	(0.020)	(0.017)	(0.016)
100	0.547	0.549	0.550	0.548	0.544	0.544	0.547	0.548
	(0.022)	(0.022)	(0.022)	(0.026)	(0.018)	(0.015)	(0.012)	(0.011)
200	0.548	0.550	0.551	0.552	0.554	0.550	0.549	0.546
	(0.012)	(0.013)	(0.012)	(0.012)	(0.011)	(0.009)	(0.009)	(0.008)

instrumentations exhibit slight downward biases for small N and large T. Both of the collapsed-set estimators, on the other hand, recover the true parameter value across virtually all values of N and T.

Table III reports the coefficient estimates of the endogenous variable. Difference GMM consistently overestimates the true parameter by around 10% (with estimates around 0.565 vs the actual value of 0.5). System GMM produces estimates slightly closer to the true value for larger N, generally around 0.545. System GMM with collapsed instruments shows the best overall performance for this parameter, with the most accurate point estimates and the smallest standard errors.

For all three types of parameters, estimators employing collapsed instrument sets consistently outperform their default counterparts. System GMM with collapsed instruments demonstrates the best overall performance, particularly for the autoregressive and endogenous variable parameters. As the number of time periods increases, the advantages of collapsed instruments become more evident. The benefits of collapsed instruments are especially pronounced with larger T values, as expected.

5. Conclusion

Our Monte Carlo results provide clear guidance for applied macroeconomists working with dynamic panel data. As the cross-sectional dimension (N) increases, all estimators show improved accuracy, but the choice of instrument approach remains consequential even in large samples. Collapsed instruments consistently outperform default instruments across virtually all configurations, providing more accurate estimates and smaller standard errors. The benefits of collapsing the instrument set increase the larger the number of time periods. System GMM with collapsed instruments appears to be the best overall performer, especially for endogenous variables. For practitioners, these results provide clear guidance for default estimation practices across the full spectrum of GMM estimator choices. Since researchers typically use standard software settings without extensive sensitivity testing, establishing empirically-supported defaults is crucial. Our comprehensive comparison of both Difference GMM and System GMM variants demonstrates that System GMM with collapsed instruments should serve as the standard default for dynamic panel estimation, regardless of N and T.

References

Arellano, M. and 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.

Arellano, M. and Bover, O. (1995). Another look at the instrumental variable estimation of error-components models. *Journal of Econometrics*, 68(1):29–51.

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

Levendis, J. (2023). *Time Series Econometrics: Learning through replication*. Springer, 2nd edition.

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

Roodman, D. (2009a). How to do xtabond2: An introduction to difference and system GMM in stata. *The Stata Journal*, 9(1):86–136.

Roodman, D. (2009b). A note on the theme of too many instruments. *Oxford Bulletin of Economics and statistics*, 71(1):135–158.