

ISSN Online: 2456-1452
NAAS Rating (2026): 4.49
IJSAM 2026; 11(10): 49-55
www.mathsjournal.com
Received: 17-08-2026
Accepted: 26-09-2026
Published: 09-10-2026

Sreedevi PN

Department of Statistics,
Ramanujan School of
Mathematical Sciences,
Pondicherry University,
Puducherry, India

Dr. J Prabhakara Naik

Department of Statistics,
Ramanujan School of
Mathematical Sciences,
Pondicherry University,
Puducherry, India

Ponnada Asha

Department of Statistics,
Ramanujan School of
Mathematical Sciences,
Pondicherry University,
Puducherry, India

Corresponding Author:

Dr. J Prabhakara Naik

Department of Statistics,
Ramanujan School of
Mathematical Sciences,
Pondicherry University,
Puducherry, India

Instrument validity in dynamic panel GMM under general measurement-error dependence

Sreedevi PN¹, Dr. J Prabhakara Naik², Ponnada Asha³

DOI: <https://www.doi.org/10.22271/math.2026.v11.i10a.2717>

ABSTRACT

The contamination of both the Difference-GMM and the System-GMM moment conditions depends on the error process only through three consecutive autocovariances. The contamination of the level moment condition at instrument depth s is the exact negative of the contamination of the differenced moment condition at depth $s+1$, whatever the error process, so the System-GMM level block offers no first-order protection against measurement error. In moving-average and autoregressive error processes carry opposite practical implications: under an MA(q) error, instruments at depth $q+3$ and beyond are exactly valid and consistency is recoverable at finite cost, whereas under an autoregressive error no finite truncation restores validity. Monte Carlo evidence confirms the closed-form expressions to within simulation error and establishes a result of direct practical importance: a minimum-depth rule calibrated to autoregressive error is not merely insufficient under moving-average error but actively harmful, producing bias up to eight times larger than the conventional instrument set it was intended to improve upon. We therefore characterize the lag profile of the contamination sequence, which terminates under MA(q) and decays geometrically under AR(1), as a diagnostic that renders the maintained error-process assumption an empirical question rather than an untestable premise.

Keywords: Dynamic panel data, Measurement error, Generalized method of moments, Instrument validity, Moving-average processes

1. Introduction

Dynamic panel data models are routinely estimated by the generalized method of moments, using internal instruments constructed from lags of the dependent variable. The Difference-GMM estimator of Arellano *et al.* (1991) [1] instruments the first-differenced equation with sufficiently deep lagged levels; the System-GMM estimator of Arellano *et al.* (1995) [2] and Blundell *et al.* (1998) [3] augments this with a level equation instrumented by lagged first differences. Both rest entirely on orthogonality between the chosen instruments and the model's composite disturbance.

That orthogonality fails when the dependent variable is measured with error. Griliches *et al.* (1986) [4] established the consequences of errors-in-variables for panel estimation, and Biörn (2000, 2012) traced how measurement error corrupts the instrument sets on which dynamic panel GMM depends. Wansbeek (2001) [5] characterized the moment conditions available under mismeasurement, and Xiao, Shao, and Palta (2010) [6] developed instrumental-variable and GMM procedures for panels with measurement error under more general error structures. Lee, Moon, and Weidner (2012) [7] and Lee, Moon, and Zhou (2017) [8] take up the problem in fixed-effects and large- N , large- T settings respectively. Staudenmayer *et al.* (2005) [9] analyse measurement error in linear autoregressive models directly.

A companion analysis derives, in closed form, the contamination of both estimators' moment conditions when the measurement error follows a stationary AR(1) process, showing that the contamination vanishes at only two points in the persistence parameter, and that the shallowest lag of each instrument set is contaminated whenever any measurement error is present. Those results, like the wider literature, are conditional on the assumed parametric form of the error process. This is an uncomfortable position for applied work. The measurement-error process is by construction unobserved, so the maintained assumption cannot be checked, and there is no particular reason to expect reporting lags, periodic revisions, or aggregation across sources to generate autoregressive rather than moving-average dependence.

This paper shows that the parametric assumption is dispensable in the derivation and partially testable in the data. Section 3 establishes that the contamination of both moment conditions is a function of three consecutive autocovariances of the error process and of nothing else. The AR(1) results follow by substitution, as do the corresponding results for any moving-average process. Three findings emerge that are not visible within a single parametric family.

The contamination of the System-GMM level moment condition at depth 's' equals minus the contamination of the Difference-GMM condition at depth s+1, for every covariance-stationary error process. The additional block that motivates System-GMM therefore carries the same contamination as the block it supplements, shifted by one lag and with the sign reversed, and offers no first-order protection. Under an MA(q) error, the contamination is exactly zero from lag depth q+3 onward, so discarding a finite number of shallow instruments restores exact validity and consistency. Under an AR(1) error the contamination is geometric in the depth and never reaches zero, so no truncation suffices. The two cases are usually conflated under the informal heading of serially correlated measurement error, yet they support opposite recommendations. From the simulation a minimum-depth rule calibrated to autoregressive error discard the shallowest lag is not merely insufficient when the error is moving average. It is ac-

tively harmful at moderate moving-average dependence it produces bias roughly eight times larger than the conventional instrument set it was meant to improve upon, because the lag it retains carries contamination that grows with the dependence parameter while the lag it discards carries contamination of the opposite sign. Since the practitioner cannot observe which process generated the error, this is an argument for the deeper rule, or for the diagnostic of Proposition 3.10, rather than for the shallower one.

The remainder of the paper is organized as follows. Section 2 states the model and assumptions. Section 3 contains the theoretical results. Section 4 reports the Monte Carlo evidence. Section 5 concludes.

2. Model and assumptions

Consider a panel of N cross-sectional units observed over T periods, $i = 1, \dots, N$ and $t = 1, \dots, T$. The latent dynamic panel data model is

$$y_{it}^* = \rho y_{i,t-1}^* + \alpha_i + u_{it}, \quad |\rho| < 1 \quad (1)$$

where y_{it}^* is the true but unobserved dependent variable, α_i a time-invariant individual effect, and u_{it} an idiosyncratic disturbance.

Assumption 2.1 (Structural innovations). $E(u_{it}) = 0$, $\text{Var}(u_{it}) = \sigma_u^2$, $\text{Cov}(u_{it}, u_{is}) = 0$ for $t \neq s$, and $\text{Cov}(u_{it}, \alpha_i) = 0$.

The observed variable is contaminated additively,

$$y_{it} = y_{it}^* + e_{it} \quad (2)$$

Assumption 2.2 (General stationary measurement error). The measurement error $\{e_{it}\}$ is covariance stationary with $E(e_{it}) = 0$ and autocovariance sequence

$$\gamma_k = \text{Cov}(e_{it}, e_{i,t-k}) = \gamma_{-k}, \quad k = 0, 1, 2, \dots,$$

satisfying $\sum_{k \geq 0} |\gamma_k| < \infty$ and $\gamma_0 > 0$. The error process is independent of $\{u_{it}\}$, of $\{y_{it}^*\}$ and of α_i , and is independent across individuals. Assumption 2.2 imposes no parametric structure. Any stationary ARMA process, any moving average of finite or infinite order, and the classical white-noise case are all admitted. Where a specific

process is considered below, we impose the usual identification restrictions: $|\phi| < 1$ for the autoregressive case, and invertibility $|\theta| < 1$ for the moving-average case, since θ and θ^{-1} generate proportional autocovariance sequences and are therefore observationally equivalent up to a rescaling of the innovation variance.

Substituting Equation (2) into Equation (1) gives the observed model

$$y_{it} = \rho y_{i,t-1} + \alpha_i + \varepsilon_{it}, \quad \varepsilon_{it} = u_{it} + e_{it} - \rho e_{i,t-1} \quad (3)$$

and first-differencing removes α_i ,

$$\Delta y_{it} = \rho \Delta y_{i,t-1} + \Delta \varepsilon_{it}, \quad \Delta \varepsilon_{it} = \Delta u_{it} + \Delta e_{it} - \rho \Delta e_{i,t-1} \quad (4)$$

Difference-GMM exploits $E(y_{i,t-s} \Delta \varepsilon_{it}) = 0$ for $s \geq 2$; System-GMM adds $E(\Delta y_{i,t-s} \varepsilon_{it}) = 0$ for $s \geq 1$, the latter requiring the mean-stationarity condition $E(\Delta y_{i2}^* \alpha_i) = 0$ of [Blundell et al. \(1998\)](#) [3]. Both conditions hold in the absence of measurement error. Section 3 evaluates them when $\gamma_0 > 0$.

3. Contamination under general error dependence

3.1 The contamination lemma

Lemma 3.1 (Contamination under general error dependence). Under Assumptions 2.1 and 2.2, the contamination of the Difference-GMM moment condition at instrument depth s is

$$b_s^D \equiv E(y_{i,t-s} \Delta \varepsilon_{it}) = (\gamma_s - \gamma_{s-1}) - \rho (\gamma_{s-1} - \gamma_{s-2}), \quad s \geq 2 \quad (5)$$

and the contamination of the System-GMM level moment condition is

$$b_s^S \equiv E(\Delta y_{i,t-s} \varepsilon_{it}) = (\gamma_s - \gamma_{s+1}) - \rho (\gamma_{s-1} - \gamma_s), \quad s \geq 1 \quad (6)$$

Proof: Writing $y_{i,t-s} = y_{i,t-s}^* + e_{i,t-s}$ and applying Assumption 2.1 to the terms in u and Assumption 2.2 to the independence of y^* and e ,

$$E(y_{i,t-s} \Delta \varepsilon_{it}) = E(e_{i,t-s} \Delta e_{it}) - \rho E(e_{i,t-s} \Delta e_{i,t-1})$$

By stationarity $E(e_{i,t-s} \Delta e_{it}) = \gamma_s - \gamma_{s-1}$ and $E(e_{i,t-s} \Delta e_{i,t-1}) = \gamma_{s-1} - \gamma_{s-2}$, which gives Equation (5). For the

level condition, writing $\Delta y_{i,t-s} = \Delta y_{i,t-s}^* + \Delta e_{i,t-s}$, the same arguments leave

$$E(\Delta y_{i,t-s} \varepsilon_{it}) = E(\Delta e_{i,t-s} e_{it}) - \rho E(\Delta e_{i,t-s} e_{i,t-1})$$

and $E(\Delta e_{i,t-s} e_{it}) = \gamma_s - \gamma_{s+1}$, $E(\Delta e_{i,t-s} e_{i,t-1}) = \gamma_{s-1} - \gamma_s$, giving Equation (6).

Lemma 3.1 says that the entire effect of measurement error on instrument validity is transmitted through three consecutive autocovariances of the error process. Nothing else about that process matters — not its order, not whether it is autoregressive or moving average, not whether it admits a finite parameterization at all.

$$-b_{s+1}^D = -(\gamma_{s+1} - \gamma_s) + \rho(\gamma_s - \gamma_{s-1}) = (\gamma_s - \gamma_{s+1}) - \rho(\gamma_{s-1} - \gamma_s)$$

which is Equation (6).

Remark 3.3. System-GMM is adopted because its level moment conditions add information the differenced conditions do not contain. Under measurement error, however the contamination of the added block is not merely of the same order as that of the differenced block: it is the same quantity with the opposite sign, shifted by one lag. The level equation therefore offers no first-order protection against measurement error relative to the differenced equation at matched depths, and the two estimators should be expected to exhibit near-identical

3.2 A mirror relation between the two blocks

Proposition 3.2 (Mirror equivalence). Under Assumptions 2.1 and 2.2, for every $s \geq 1$,

$$b_s^S = -b_{s+1}^D \quad (7)$$

Proof. Evaluating Equation (5) at $s+1$ and negating,

bias. Any advantage System-GMM retains under contamination is an advantage in sampling variability, not in location. Section 4.3 confirms both halves of this prediction.

3.3 The autoregressive case

Corollary 3.4 (AR(1) measurement error). Suppose $e_{it} = \phi e_{i,t-1} + \eta_{it}$ with $|\phi| < 1$, so that $\gamma_k = \phi^k \sigma_e^2$ with $\sigma_e^2 = \frac{\sigma_\eta^2}{(1-\phi^2)}$. Then

$$b_s^D = (\phi - 1)(\phi - \rho) \phi^{s-2} \sigma_e^2, \quad b_s^S = (1 - \phi)(\phi - \rho) \phi^{s-1} \sigma_e^2 \quad (8)$$

Proof. Substituting $\gamma_k = \phi^k \sigma_e^2$ into Equation (5),

$$b_s^D = \sigma_e^2 \phi^{s-2} [(\phi^2 - \phi) - \rho(\phi - 1)] = \sigma_e^2 \phi^{s-2} (\phi - 1)(\phi - \rho),$$

and similarly for Equation (6).

Corollary 3.4 reproduces exactly the closed-form contamination results derived for the AR(1) case elsewhere in this line of work. Note that $\phi^{s-2} > 0$ for every finite s whenever $\phi \neq 0$: the contamination decays but never reaches zero, so no truncation of the instrument set restores exact validity. The practitioner is left choosing a depth at which the residual contamination is judged tolerable, trading bias against the loss of instrument relevance that deeper lags entail.

3.4 The moving-average case

Corollary 3.5 (MA(1) measurement error). Suppose $e_{it} = \eta_{it} + \theta \eta_{i,t-1}$ with $\text{Var}(\eta_{it}) = \sigma_\eta^2$ and $|\theta| < 1$, so that $\gamma_0 = (1 + \theta^2) \sigma_\eta^2$, $\gamma_1 = \theta \sigma_\eta^2$, and $\gamma_k = 0$ for $k \geq 2$. Then

$$b_2^D = \sigma_\eta^2 [\rho(1 - \theta + \theta^2) - \theta] \quad (9)$$

$$b_s^D = 0 \text{ for } s \geq q+3, \quad b_s^S = 0 \text{ for } s \geq q+2 \quad (12)$$

Instrument sets constructed from lag depth $q+3$ onward for the differenced equation, and $q+2$ onward for the level equation, are therefore exactly valid, and the resulting GMM estimators are consistent.

Proof. b_s^D depends only on $\gamma_s, \gamma_{s-1}, \gamma_{s-2}$, all of which vanish when $s-2 > q$; b_s^S depends only on $\gamma_{s+1}, \gamma_s, \gamma_{s-1}$, all of which vanish when $s-1 > q$.

Remark 3.7 (The two error classes support opposite recommendations). Corollaries 3.4 and 3.6 describe fundamentally different situations. Under a moving-average error of known order, consistency is recoverable exactly and at finite cost: discard the first $q+2$ lags and the remaining moment conditions hold without approximation. Under an autoregressive error no such remedy exists at any finite depth; the contamination is merely made small. Since the two cases are usually indistinguishable at the level of informal reasoning about serially correlated measurement error, and since they carry different recommendations on instrument depth, the distinction matters.

and $b_1^S = -b_2^D$, $b_2^S = -b_3^D$, and $b_s^S = 0$ for $s \geq 3$.

Proof. Direct substitution into Lemma 3.1. For $s \geq 4$ each of γ_s, γ_{s-1} and γ_{s-2} has index at least two and is therefore zero. The level results follow from Proposition 3.2.

Corollary 3.6 (MA(q) measurement error). If $\{e_{it}\}$ is a moving average of order q , so that $\gamma_k = 0$ for $k > q$, then

Remark 3.8 (Consistency with the classical benchmark). Setting $\theta = 0$ in Equation (9) gives $b_2^D = \rho \sigma_\eta^2$ and $b_3^D = 0$, matching the classical white-noise case obtained by setting $\phi = 0$ in Corollary 3.4. Both parameterizations agree that the shallowest instrument lag is contaminated whenever any measurement error is present, and that the contamination there equals ρ times the error variance. The shallow-lag exception established elsewhere in this line of work is therefore a consequence of Lemma 3.1, not of the AR(1) specification.

Remark 3.9 (A vanishing point in the moving-average parameter). Equation (9) vanishes where $\rho \theta^2 - (1 + \rho) \theta + \rho = 0$. For $\rho = 0.70$ the invertible root is $\theta^* \approx 0.526$. At that value the shallowest differenced instrument is exactly valid despite the presence of measurement error, which is the moving-average analogue of the point $\phi = \rho$ at which the autoregressive contamination vanishes. Unlike that point, however, θ^* leaves $b_3^D = \rho \theta^* \sigma_\eta^2 \neq 0$, so validity is restored at one lag only rather than at all of them.

3.5 A testable implication

Proposition 3.10 (The lag profile identifies the error class). Under Assumptions 2.1 and 2.2:

- (i) if the error process is MA(q), then $b_s^D = 0$ for all $s \geq q + 3$: the contamination sequence terminates;
- (ii) if the error process is AR(1) with $\phi \notin \{0, \rho\}$, then $b_s^D \neq 0$ for every finite s , and successive terms satisfy $b_{s+1}^D / b_s^D = \phi$, a constant independent of s , ρ and σ_e^2 .

The two classes are therefore distinguished by whether the contamination sequence terminates at a finite depth or decays geometrically without reaching zero.

Proof. Claim (i) is Corollary 3.6. Claim (ii) follows from Equation (8), where the ratio of successive terms is ϕ because ρ , σ_e^2 and the factor $(\phi - 1)(\phi - \rho)$ are common to both.

Remark 3.11 (Implementation). The quantities b_s^D are population moments of observed data evaluated at the true parameter, so their sample analogues $\widehat{b}_s = N^{-1} \sum_i \sum_t y_{i,t-s} \Delta \varepsilon_{it}$ are computable from any preliminary consistent estimate of ρ — for instance one based on instruments at a depth greater than the largest q under consideration, which Corollary 3.6 guarantees to be valid. Of the two claims in Proposition 3.10, the termination criterion is by far the more useful in

$$\sigma_\eta = \sigma_{me} \sqrt{1 - \phi^2} \text{ for AR(1), } \sigma_\eta = \frac{\sigma_{me}}{\sqrt{1 + \theta^2}} \text{ for MA(1)} \quad (13)$$

and without it a comparison across the dependence grid would confound persistence with magnitude. The dependence parameter takes the values 0, 0.3, 0.5 and 0.8 in each case.

All estimators are one-step with an identity weighting matrix and a collapsed instrument set in the sense of Roodman (2009) [10], of fixed depth independent of T . Instruments unavailable in a given period are set to zero, following Holtz-Eakin, Newey, and Rosen (1988) [11]. Five specifications are compared:

- **DIF2**: differenced equation, lagged levels $s = 2, \dots, 5$. The conventional Difference-GMM instrument set.
- **DIF3**: differenced equation, $s = 3, \dots, 6$. The minimum depth implied by the AR(1) analysis, which discards only the shallowest lag.
- **DIF4**: differenced equation, $s = 4, \dots, 7$. The minimum depth implied by Corollary 3.6 for $q = 1$.

practice. The ratio in claim (ii) is exact in population but poorly behaved in finite samples, because b_s^D decays geometrically toward zero while the sampling variability of $\widehat{b}_s$ does not, so beyond the first two or three lags the ratio is essentially noise divided by noise. Section 4.5 documents both behaviours. Testing whether $\widehat{b}_s$ is distinguishable from zero beyond a given depth is correspondingly the version we recommend.

4. Monte Carlo evidence

4.1 Design

We simulate from Equation (1) with $\rho=0.70$, $\sigma_u = \sigma_\alpha = 1$, $N = 500$, $T = 10$ and $R = 1000$ replications. Initial conditions are drawn from the stationary distribution and each panel is run through a burn-in period, so that both the latent process and the error process are stationary from the first retained observation.

The measurement error is generated in two forms. The autoregressive form is $e_{it} = \phi e_{i,t-1} + \eta_{it}$; the moving-average form is $e_{it} = \eta_{it} + \theta \eta_{i,t-1}$. In both cases the stationary standard deviation of the error is held fixed at $\sigma_{me} = 0.50$, so that varying the dependence parameter does not also vary the magnitude of the contamination. This requires

- **SYS1**: DIF2 plus the level equation instrumented with $\Delta y_{i,t-s}$, $s = 1, \dots, 4$. The conventional System-GMM instrument set, retaining the standard Blundell *et al.* (1998) [3] instrument $\Delta y_{i,t-1}$.
- **SYS3**: DIF4 plus the level equation with $s = 3, \dots, 6$, so that both blocks satisfy Corollary 3.6.

Bias, Monte Carlo standard deviation (MCSD) and root mean squared error use divisor R throughout, so that $\text{RMSE}^2 = \text{Bias}^2 + \text{MCSD}^2$ holds exactly; this identity is verified in every cell reported. We also report the empirical rejection frequency of the Hansen (1982) [12] overidentification test at the 5% level and the empirical coverage of nominal 95% intervals. With $R = 1000$ the Monte Carlo standard error of a reported bias is at most 0.006.

4.2 Verification of the closed-form contamination

Table 1 compares the contamination terms predicted by Corollaries 3.4 and 3.5 with their simulated counterparts, computed directly as sample averages of $y_{i,t-s} \Delta \varepsilon_{it}$ evaluated at the true ρ .

Table 1. Predicted and simulated contamination b_s^D ($N = 500$, $T = 10$, $R = 1000$, $\rho = 0.70$, $\sigma_{me} = 0.50$)

Process	Parameter	Predicted $s = 2$	Simulated $s = 2$	Predicted $s = 3$	Simulated $s = 3$	Predicted $s = 4$	Simulated $s = 4$
MA(1)	$\theta = 0.0$	0.1750	0.1748	0.0000	0.0004	0.0000	-0.0007
	$\theta = 0.3$	0.0580	0.0575	0.0482	0.0499	0.0000	-0.0010
	$\theta = 0.5$	0.0050	0.0043	0.0700	0.0697	0.0000	-0.0021
	$\theta = 0.8$	-0.0323	-0.0327	0.0854	0.0851	0.0000	-0.0022
AR(1)	$\phi = 0.0$	0.1750	0.1726	0.0000	-0.0032	0.0000	-0.0024
	$\phi = 0.3$	0.0700	0.0711	0.0210	0.0228	0.0063	0.0066
	$\phi = 0.5$	0.0250	0.0261	0.0125	0.0114	0.0063	0.0067
	$\phi = 0.8$	-0.0050	-0.0043	-0.0040	-0.0015	-0.0032	-0.0017

Agreement is close throughout, and the two qualitative predictions are clearly visible. Under MA(1) the simulated b_4^D is at most 0.0022 in absolute value at every θ , against a b_3^D of up to 0.085; the same holds for b_5^D through b_7^D , none of which exceeds 0.003. The sequence terminates, as Corollary 3.5 requires. Under AR(1) it does not: at $\phi = 0.3$ the simulated values decline as 0.0711, 0.0228, 0.0066, 0.0040 without reaching zero.

The mirror relation of Proposition 3.2 is confirmed independently. Across all forty combinations of process, dependence parameter and lag depth, the quantity $\widehat{b}_s + \widehat{b}_{s+1}$ never exceeds 0.0035 in absolute value, against contamination terms up to 0.175. Since nothing in the estimation code enforces this identity, its appearance is a joint check on the derivation and on the implementation.

4.3 Estimator performance

Table 2 reports the five specifications under MA(1) error.

Table 2. Performance under MA(1) measurement error ($N = 500$, $T = 10$, $R = 1000$, $\rho = 0.70$, $\sigma_{me} = 0.50$)

θ	Spec.	Bias	MCSD	RMSE	Hansen rej.	Coverage
0	DIF2	-0.1660	0.0533	0.1743	0.443	0.104
	DIF3	-0.0070	0.1115	0.1118	0.062	0.952
	DIF4	-0.0164	0.1810	0.1817	0.057	0.967
	SYS1	-0.1626	0.0361	0.1666	0.612	0.004
	SYS3	-0.0090	0.1082	0.1086	0.084	0.958
0.3	DIF2	-0.0766	0.0514	0.0923	0.084	0.668
	DIF3	-0.0797	0.0932	0.1226	0.085	0.874
	DIF4	-0.0111	0.1719	0.1722	0.066	0.979
	SYS1	-0.0755	0.0337	0.0827	0.095	0.380
	SYS3	-0.0068	0.1056	0.1058	0.073	0.954
0.5	DIF2	-0.0387	0.0510	0.0641	0.277	0.883
	DIF3	-0.1048	0.0913	0.1390	0.128	0.771
	DIF4	-0.0033	0.1886	0.1887	0.069	0.961
	SYS1	-0.0382	0.0327	0.0503	0.382	0.800
	SYS3	-0.0012	0.1086	0.1086	0.098	0.943
0.8	DIF2	-0.0153	0.0510	0.0533	0.628	0.944
	DIF3	-0.1245	0.0878	0.1523	0.173	0.688
	DIF4	-0.0065	0.1801	0.1802	0.053	0.966
	SYS1	-0.0142	0.0314	0.0344	0.818	0.939
	SYS3	-0.0007	0.1049	0.1049	0.093	0.952

Four findings deserve comment.

Corollary 3.6 delivers what it promises. DIF4 has bias between -0.0164 and -0.0033 across the whole θ grid, with Hansen rejection frequencies between 0.053 and 0.069 against a nominal 0.05 and coverage between 0.961 and 0.979 against a nominal 0.95. SYS3 behaves similarly, with bias between -0.0090 and -0.0007 and coverage between 0.943 and 0.958. Both are, to within simulation error, exactly valid at every value of the moving-average parameter, which is the central prediction of the paper.

The AR(1) depth rule fails, and fails badly. DIF3 discards the shallowest lag, which is the correct remedy under autoregressive error. Under moving-average error it is not merely insufficient but counterproductive: its bias grows monotonically from -0.0070 at $\theta = 0$ to -0.1245 at $\theta = 0.8$, where it is eight times the bias of the conventional DIF2 it was meant to improve upon, and its coverage falls to 0.688. The mechanism is transparent from Corollary 3.5: DIF3 retains lag 3, whose contamination $b^D_3 = \rho\theta\sigma_1^2$ grows with θ , while discarding lag 2, whose contamination is of the opposite sign at large θ and therefore partially offsetting. Because the practitioner cannot observe which process generated the error, this is a strong argument for the deeper rule.

The mirror relation is visible in the estimator biases. At every θ , DIF2 and SYS1 differ in bias by less than 0.006, -0.1660 against -0.1626 , -0.0766 against -0.0755 , -0.0387 against -0.0382 , -0.0153 against -0.0142 — while SYS1's MCSD is roughly one third smaller throughout. This is exactly the pattern Remark 3.3 predicts: the level block buys precision, not location. The consequence for inference is severe, and follows the pattern documented in the companion analy-

sis: at $\theta = 0$ SYS1's coverage is 0.004 against DIF2's 0.104, because the same displacement is measured against a narrower interval.

Exact validity is not free. DIF4's MCSD is roughly 0.18 against DIF2's 0.051, a factor of three and a half, and SYS3's is 0.105 against SYS1's 0.033. At $\theta = 0.8$ the conventional DIF2 has both lower RMSE (0.0533 against 0.1802) and adequate coverage (0.944), so the deeper instrument set is dominated there on every criterion except the Hansen test. The recommendation of Corollary 3.6 is therefore about guaranteed validity across the parameter space rather than about minimal RMSE at any particular point. Comparing the two System-GMM specifications is instructive in this regard: SYS3 attains the same near-zero bias as DIF4 at roughly half its sampling variability, so where an exactly valid estimator is wanted, the level block remains worth including even though Proposition 3.2 shows it contributes nothing to bias reduction.

A note on the Hansen test. At $\theta = 0.8$ the test rejects DIF2 in 62.8% of samples although that estimator's bias is only -0.0153 and its coverage is 0.944. The test is detecting invalid moment conditions, which is what it is designed to do, rather than biased estimates; at that value of θ the individual contamination terms are large but largely offsetting in their effect on $\hat{\rho}$. The converse failure also occurs: at $\theta = 0.3$ the rejection frequency is 0.084, close to nominal, while DIF2's coverage has already fallen to 0.668. Neither direction of error should be surprising, but together they caution against reading the overidentification test as a direct indicator of estimator reliability.

4.4 The autoregressive comparison

Table 3 repeats the exercise with AR(1) error on the identical design, so that the two error classes can be compared directly.

Table 3. Performance under AR(1) measurement error ($N = 500$, $T = 10$, $R = 1000$, $\rho = 0.70$, $\sigma_{me} = 0.50$)

ϕ	Spec.	Bias	MCSD	RMSE	Hansen rej.	Coverage
0	DIF2	-0.1620	0.0533	0.1705	0.476	0.119
	DIF3	0.0004	0.1122	0.1122	0.077	0.953

(continued on next page)

0.3	DIF4	-0.0112	0.1871	0.1874	0.064	0.956
	SYS1	-0.1613	0.0356	0.1651	0.616	0.001
	SYS3	-0.0081	0.1118	0.1121	0.101	0.956
	DIF2	-0.0805	0.0515	0.0956	0.090	0.648
	DIF3	-0.0456	0.0974	0.1076	0.066	0.926
	DIF4	-0.0368	0.1778	0.1816	0.064	0.944
0.5	SYS1	-0.0763	0.0325	0.0829	0.114	0.367
	SYS3	-0.0255	0.1063	0.1094	0.107	0.939
	DIF2	-0.0331	0.0531	0.0626	0.069	0.918
	DIF3	-0.0251	0.0922	0.0956	0.069	0.946
	DIF4	-0.0267	0.1575	0.1598	0.070	0.951
	SYS1	-0.0317	0.0340	0.0465	0.064	0.852
0.8	SYS3	-0.0209	0.0973	0.0995	0.080	0.953
	DIF2	0.0039	0.0556	0.0557	0.068	0.949
	DIF3	-0.0005	0.0903	0.0903	0.062	0.954
	DIF4	-0.0069	0.1492	0.1493	0.061	0.966
	SYS1	0.0067	0.0352	0.0359	0.072	0.936
	SYS3	0.0009	0.0925	0.0925	0.082	0.950

The contrast with Table 2 is the substantive point. Under AR(1), increasing the instrument depth reduces the bias but does not eliminate it: at $\phi = 0.3$, DIF4's bias is -0.0368 , smaller than DIF2's -0.0805 but far from zero, exactly as Corollary 3.4 requires, since $b_4^D = 0.0063 \neq 0$ there. Under MA(1) the corresponding cell has DIF4 bias -0.0111 with a predicted contamination of exactly zero. Depth buys exact validity in one case and only approximate improvement in the other.

The small biases at $\phi = 0.8$ reflect proximity to the vanishing point $\phi = \rho = 0.70$, where all contamination terms are zero irrespective of depth; they are not evidence that deep instruments are exactly valid under autoregressive error.

4.5 The lag-profile diagnostic

Proposition 3.10 suggests two ways of discriminating between the error classes. The termination criterion performs well. Under MA(1) the estimated $\hat{b}_s$ for $s \geq 4$ lie within 0.003 of zero at every θ , while $\hat{b}_3$ reaches 0.085; under AR(1) at $\phi = 0.3$ the same terms are 0.0066, 0.0040, 0.0006 and 0.0033, small but not systematically nearer zero than $\hat{b}_3 = 0.0228$. A formal test of $H_0: b_s^D = 0$ for $s \geq q + 3$ is therefore feasible, and we regard this as the usable form of the diagnostic.

The ratio criterion does not perform well, and we report this explicitly. At $\phi = 0.3$ the estimated ratios $\hat{b}_{s+1}/\hat{b}_s$ across $s = 2, \dots, 6$ are 0.320, 0.289, 0.607, 0.151 and 5.454; only the first two are close to $\phi = 0.3$. The reason is structural rather than a matter of sample size: b_s^D decays geometrically toward zero while the sampling variability of $\hat{b}_s$ does not, so beyond the second or third lag the ratio approaches a ratio of two noise terms and is not informative. Claim (ii) of Proposition 3.10 is an exact population identity and remains useful for interpreting the shallowest two ratios, but it should not be relied upon as a practical test.

5. Conclusion

The analysis of measurement error in dynamic panel GMM has generally proceeded by fixing a parametric form for the error process and deriving the consequences. This paper shows that the parametric step is unnecessary. The contamination of both the Difference-GMM and the System-GMM moment conditions depends on the error process only through three consecutive autocovariances, so a single lemma covers every covariance-stationary error process, and the autoregressive results familiar from the literature follow as one instance among many.

Three results follow that are not visible within a single parametric family. The contamination of the System-GMM level moment condition at depth s is exactly minus that of the differenced condition at

depth $s+1$, whatever the error process, so the level block contributes no first-order protection against measurement error; the simulation evidence confirms that the two estimators' biases differ by less than 0.006 at every design point while their sampling variabilities differ by a third. Under MA(q) error the contamination is exactly zero from depth $q+3$ onward, so consistency is recoverable at finite cost, whereas under AR(1) error no finite truncation suffices. And a minimum-depth rule calibrated to the autoregressive case is actively harmful under moving-average error, producing bias up to eight times that of the conventional instrument set it was intended to improve upon.

For practice, three implications follow. Where measurement error is suspected and its form is unknown, the deeper instrument set is the safer choice, and it should include the level block: SYS3 attains the same near-zero bias as the differenced-only DIF4 at roughly half the sampling variability, even though Proposition 3.2 shows the level block contributes nothing to bias reduction. Where the error form can be established, the depth requirement follows directly from Corollary 3.6. And in either case the Hansen test should not be read as an indicator of estimator reliability: our design contains points where it rejects in two-thirds of samples while coverage is at its nominal level, and points where it is near nominal size while coverage has fallen below seventy per cent.

Finally, the maintained assumption about the error process need not be maintained blindly. Proposition 3.10 shows that the two error classes leave different signatures in the lag profile of the contamination sequence: termination under MA(q), geometric decay under AR(1). Developing the limiting distribution of a formal test based on the termination criterion, and extending the analysis to mixed ARMA error processes where both features are present, are natural continuations of this work.

References

1. Arellano M, Bond S. Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *The Review of Economic Studies*. 1991;58(2):277–297.
2. Arellano M, Bover O. Another look at the instrumental variable estimation of error-components models. *Journal of Econometrics*. 1995;68(1):29–51.
3. Blundell R, Bond S. Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics*. 1998;87(1):115–143.
4. Griliches Z, Hausman JA. Errors in variables in panel data. *Journal of Econometrics*. 1986;31(1):93–118.
5. Wansbeek TJ. GMM estimation in panel data models with measurement error. *Journal of Econometrics*. 2001;104(2):259–268.

6. Xiao Z, Shao J, Palta M. Instrumental variable and GMM estimation for panel data with measurement error. *Statistica Sinica*. 2010;20:1725–1747.
7. Lee Y, Moon HR, Weidner M. Dynamic panel models with interactive fixed effects and measurement error. *Economics Letters*. 2012;117(1):239–242.
8. Lee Y, Moon HR, Zhou X. Large panel data models with measurement error. *Journal of Econometrics*. 2017;198(1):1–17.
9. Staudenmayer J, Buonaccorsi JP. Measurement error in linear autoregressive models. *Journal of the American Statistical Association*. 2005;100:841–852.
10. Roodman D. A note on the theme of too many instruments. *Oxford Bulletin of Economics and Statistics*. 2009;71(1):135–158.
11. Holtz-Eakin D, Newey W, Rosen HS. Estimating vector autoregressions with panel data. *Econometrica*. 1988;56(6):1371–1395.
12. Hansen LP. Large sample properties of generalized method of moments estimators. *Econometrica*. 1982;50(4):1029–1054.
13. Biørn E. Panel data with measurement errors: Instrumental variables and GMM estimators combining levels and differences. *Econometric Reviews*. 2000;19(4):391–424.
14. Biørn E. The measurement error problem in dynamic panel data analysis: Modeling and GMM estimation. Memorandum No. 02/2012, Department of Economics, University of Oslo. 2012.
15. Nickell S. Biases in dynamic models with fixed effects. *Econometrica*. 1981;49(6):1417–1426.

How to cite this article: PN S, Naik DJP, Asha P. Instrument validity in dynamic panel GMM under general measurement-error dependence. *International Journal of Statistics and Applied Mathematics*. 2026;11(10):49-55. doi: <https://www.doi.org/10.22271/maths.2026.v11.i10a.2717>