A Non-Invariance Problem in Panel GMM Estimators When Level Instruments Are Used for Differenced Equations

Seung C. Ahn
Sogang University (Correspondence, Email: smahn@sogang.ac.kr)

Yoshitsugu Kitazawa
Kyushu Sangyo University (Email: kitazawa@ip.kyusan-u.ac.jp)

Presented at 20th International Panel Data Conference held at the Hitotsubashi Hall in Tokyo on 9 and 10 July 2014.

1. Introduction

LIDE (level-instruments-for-differenced-equations) approach

- The most popular way of controlling for the **fixed effects**: to remove them by **first-differencing** or **quasi-differencing** regression equations.
- The differenced equations are estimated by the Generalized Method of Moments (GMM) of Hansen (1982); for example, Arellano and Bond (1991), and Ahn and Schmidt (1995), for **linear dynamic models**, and Chamberlain (1992), Wooldridge (1997), Windmeijer (2000), for **count data models**.
- The typical instruments used for the **differenced equations** are **lagged level regressors**. (LIDE "level-instruments-for-differenced-equations" approach).

Purpose of this paper: to remind the LIDE users of the importance of including a **constant** (typically, one) in instrument sets

- “GMM-C” estimator: using **a constant** as an instrument in addition to lagged level regressors for the differenced equations
- “GMM-WC” estimator: without a constant in the instruments set
- The GMM-C estimator is asymptotically more efficient than the GMM-WC estimator (Crépon, Kramarz and Trognon (1997)).
- The GMM-WC estimation results are not invariant to the means of regressors, but the GMM-C estimation results are invariant to the means of regressors asymptotically and often in finite sample, when using the first-difference and quasi-difference transformations.

Structure of Paper

- Section 2: considering three different LIDE methods: Arellano-Bond (1991) method for dynamic panel data models and the quasi-differencing methods of Chamberlain (1992) and Wooldridge (1997, endnote 2) for count panel data models, and showing that both the asymptotic and finite-sample distributions of the GMM-WC estimators from these methods are not invariant to the means of regressors.
- Section 3: Small-scale Monte Carlo simulation
- Section 4: Concluding remark

2. Asymptotic Distributions of GMM Estimators

We only consider the cases with *T*=2.

- We suppose the case of $N \to \infty$ and T being fixed ($i = 1, \dots, N$ and $\mathrm{t} = 1, \dots, T$). *i* and *t* denote the individual and time respectively.
- β is the regression coefficient.
- $m_{wc,i}(\beta)$: moment function used for GMM-WC estimator.
- $m_i(\beta) = (m_{wc,i}(\beta)', m_{c,i}(\beta)')'$: moment function used for GMM-WC estimator, where $m_{c,i}(\beta)$ is simply the differenced error.
- we use notation β_0 to denote the true value of β.

$$M_{wc} = E\left(\frac{\partial m_{wc,i}(\beta_o)}{\partial \beta'}\right); M_c = E\left(\frac{\partial m_{c,i}(\beta_o)}{\partial \beta'}\right); M = \begin{pmatrix} M_{wc} \\ M_c \end{pmatrix}$$

$$V_{j,j'} = E[m_{j,i}(\beta_o) m_{j',i}(\beta_o)'] \qquad V = \left[V_{j.j'}\right]$$

$$j, j' = wc, c$$

2.1. Simple Dynamic Panel Model

Model

$$y_{it} = \beta_o y_{i,t-1} + \alpha_i + \varepsilon_{it}$$

Dependent variable

Parameter of interest (-1,1)

fixed effect

Usual disturbance

Covariance Restriction Ahn and Schmidt (1995)

$$\begin{pmatrix} y_{i0} \\ \alpha_i \\ \varepsilon_{i1} \\ \varepsilon_{i2} \end{pmatrix} \sim N\left(\begin{pmatrix} \mu_0 \\ \mu_\alpha \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_0^2 & \sigma_{\alpha 0} & 0 & 0 \\ \sigma_{\alpha 0} & \sigma_\alpha^2 & 0 & 0 \\ 0 & 0 & \sigma_\varepsilon^2 & 0 \\ 0 & 0 & 0 & \sigma_\varepsilon^2 \end{pmatrix}\right)$$

i: individual

t: time period (*t=1, 2*)

Generally, the GMM-C estimator is strictly more efficient than the GMM-WC estimator, as Crépon, Kramarz and Trognon (1997) projected.

- That is,

The asymptotic variance of the GMM-WC estimator

$$\Xi_{wc} = \left[M'_{wc} (V_{wc,wc})^{-1} M_{wc} \right]^{-1} = \frac{2\sigma_\varepsilon^2 (\mu_0^2 + \sigma_0^2)}{\left[(\beta_o - 1)\sigma_0^2 + \sigma_{\alpha 0} + \mu_0 \{ (\beta_o - 1)\mu_0 + \mu_\alpha \} \right]^2}$$

The asymptotic variance of the GMM-C estimator

$$\Xi_c = \left[M' V^{-1} M \right]^{-1} = \frac{2\sigma_\varepsilon^2 \sigma_0^2}{[(\beta_o - 1)\sigma_0^2 + \sigma_{\alpha 0}]^2 + \sigma_0^2 [(\beta_o - 1)\mu_0 + \mu_\alpha]^2}$$

$$\Xi_{wc} - \Xi_c = \frac{2\sigma_\varepsilon^2 \left(\mu_0 \sigma_{\alpha 0} - \sigma_0^2 \mu_\alpha \right)^2}{\left[(\beta_o - 1)\sigma_0^2 + \sigma_{\alpha 0} + \mu_0 \{ (\beta_o - 1)\mu_0 + \mu_\alpha \} \right]^2 \left[\{ (\beta_o - 1)\sigma_0^2 + \sigma_{\alpha 0} \}^2 + \sigma_0^2 \{ (\beta_o - 1)\mu_0 + \mu_\alpha \}^2 \right]} \geq 0$$

The asymptotic variance of the GMM-WC estimator depends on the choice of *b*, but that of the GMM-C does not.

The asymptotic variance of the GMM-WC estimator when $y_{it}^b = y_{it} - b$ is used

The asymptotic variance of the GMM-WC estimator when y_{it} is used

- That is,

$$\Xi_{wc}^{b} = \frac{2\sigma_{\varepsilon}^{2}[(\mu_0 - b)^2 + \sigma_0^2]}{\left[(\beta_o - 1)\{(\mu_0 - b)^2 + \sigma_0^2\} + \sigma_{\alpha 0} + (\mu_0 - b)\{\mu_\alpha + b(\beta_o - 1)\}\right]^2} \neq \Xi_{wc}$$

$$\Xi_{c}^{b} = \frac{2\sigma_{\varepsilon}^{2}\sigma_0^2}{[(\beta_o - 1)\sigma_0^2 + \sigma_{\alpha 0}]^2 + \sigma_0^2[(\beta_o - 1)\mu_0 + \mu_\alpha]^2} = \Xi_{c}$$

The asymptotic variance of the GMM-C estimator when $y_{it}^b = y_{it} - b$ is used

The asymptotic variance of the GMM-C estimator when y_{it} is used

General relationship between $\Xi_{\mathrm{wc}}^{\mathrm{b}}$ and b

- Suppose the stationarity: $y_{it} = \frac{\alpha_i}{1-\beta_0} + v_{i0}$, where $E[v_{i0}] = 0$, $var(v_{i0}) = \sigma_v^2$, and v_{i0} is uncorrelated with ε_{i1}, ε_{i2} and α_i.
- Holding other things equal (given σ_ε^2, σ_α^2, σ_v^2, and β_0), the efficiency of the GMM-WC estimator (i.e. the smallness of $\Xi_{\mathrm{wc}}^{\mathrm{b}}$) has a negative relationship with the absolute value of the mean of y_{it}.

In finite sample, the minimands of the one-step and two-step GMM-C estimators do not depend on the choices of b.

Implying that both the one-step and two-step GMM-C estimators are invariant to b in finite samples, for the case of using $\sum_{i=1}^{N} Z_i^b H Z_i^{b\prime}$ as the one-step weighting matrix.

Elements in *H* are constants/

- That is,

$$Q_{1,N}^{b}(\beta) \equiv \left(\Sigma_{i=1}^{N} Z_i^b r_i^b(\beta)\right)' \left[\Sigma_{i=1}^{N} Z_i^b H Z_i^{b\prime}\right]^{-1} \left(\Sigma_{i=1}^{N} Z_i^b r_i^b(\beta)\right)$$
$$= \left(\Sigma_{i=1}^{N} Z_i^0 r_i^0(\beta)\right)' \left[\Sigma_{i=1}^{N} Z_i^0 H Z_i^0\right]^{-1} \left(\Sigma_{i=1}^{N} Z_i^0 r_i^0(\beta)\right)$$
$$\equiv Q_{1,N}^{0}(\beta)$$

$$Q_{2,N}^{b}(\beta) \equiv \left(\Sigma_{i=1}^{N} Z_i^b r_i^b(\beta)\right)' \left[\Sigma_{i=1}^{N} Z_i^b r_i^b(\tilde{\beta}_c^b) r_i^b(\tilde{\beta}_c^b)' Z_i^{b\prime}\right]^{-1} \left(\Sigma_{i=1}^{N} Z_i^b r_i^b(\beta)\right)$$
$$= \left(\Sigma_{i=1}^{N} Z_i^0 r_i^0(\beta)\right)' \left[\Sigma_{i=1}^{N} Z_i^0 r_i^0(\tilde{\beta}_c^0) r_i^0(\tilde{\beta}_c^0)' Z_i^{0\prime}\right]^{-1} \left(\Sigma_{i=1}^{N} Z_i^0 r_i^0(\beta)\right)$$
$$\equiv Q_{2,N}^{0}(\beta).$$

Minimand of **one-step** GMM-C estimator when $y_{it}^b = y_{it} - b$ is used

Minimand of **one-step** GMM-C estimator when y_{it} is used

Minimand of **two-step** GMM-C estimator when $y_{it}^b = y_{it} - b$ is used

Minimand of **two-step** GMM-C estimator when y_{it} is used

2.2. Simple Count Data Models

Model

$\eta_{\mathrm{i}} = \exp(\alpha_i)$

$\mathrm{v}_{it} = \exp(\varepsilon_{it})$

- Model

$$y_{it} = \exp(x_{it}\beta_o + \alpha_i + \varepsilon_{it}) = \exp(x_{it}\beta_o)\eta_i v_{it}$$

Count dependent variable

Continuous explanatory variable

Parameter of interest

Fixed effects

disturbance

$$E(v_{i1} \mid \eta_i, x_{i1}) = 1 \qquad E(v_{i2} \mid \eta_i, x_{i1}, x_{i2}) = 1$$

Assumptions on disturbances: predeterminedness is assumed here.

Both the asymptotic variance of the GMM-C estimator and the GMM minimands of the one-step and two-step GMM-C estimator are invariant to the choice of b, when using Chamberlain transformation.

- Chamberlain Transformation

$$p_i^0(\beta) = y_{i1} - \exp(-\Delta x_{i2}\beta)y_{i2}$$

$$\Xi_c^b = \Xi_c$$

The asymptotic variance of the GMM-C estimator when $x_{it}^b = x_{it} - b$ is used

The asymptotic variance of the GMM-C estimator when x_{it} is used

$$Q_{1,N}^b(\beta) = Q_{1,N}^0(\beta)$$

Minimand of **one-step** GMM-C estimator when $x_{it}^b = x_{it} - b$ is used

Minimand of **one-step** GMM-C estimator when x_{it} is used

$$Q_{2,N}^b(\beta) = Q_{2,N}^0(\beta)$$

Minimand of **two-step** GMM-C estimator when $x_{it}^b = x_{it} - b$ is used

Minimand of **two-step** GMM-C estimator when x_{it} is used

The asymptotic variance of the GMM-C estimator is invariant to the choice of b, but the GMM minimands of the one-step and two-step GMM-C estimators are not so, when using Wooldridge transformation.

- Wooldridge Transformation

$$q_i^0(\beta) = \exp(-x_{i1}\beta)y_{i1} - \exp(-x_{i2}\beta)y_{i2}$$

The asymptotic variance of the GMM-C estimator when $x_{it}^b = x_{it} - b$ is used

$$\Xi_c^b = \Xi_c$$

The asymptotic variance of the GMM-C estimator when x_{it} is used

Minimand of **two-step** GMM-C estimator when $x_{it}^b = x_{it} - b$ is used

$$Q_{2,N}^b(\beta) \neq Q_{2,N}^0(\beta)$$

Minimand of **one-step** GMM-C estimator when $x_{it}^b = x_{it} - b$ is used

$$Q_{1,N}^b(\beta) \neq Q_{1,N}^0(\beta)$$

Minimand of **one-step** GMM-C estimator when x_{it} is used

Minimand of **two-step** GMM-C estimator when x_{it} is used

3. Simulations

- This section investigates the finite-sample properties of the two-step GMM-C and GMM-WC estimators.
- We also illustrate the invariance property of the GMM-C estimator.
- Our experiments are carried out with a simple dynamic panel model and two count panel models.

3.1. Dynamic Panel Data Model

- Data Generating Process

$$y_{it} = \beta_o y_{i,t-1} + \alpha_i + \varepsilon_{it} \qquad y_{i0} = \frac{\alpha_i}{1-\beta_o} + v_{i0}$$

$$\beta_o = 0.5 \qquad \sigma_\alpha^2 = \sigma_\varepsilon^2 = 1 \qquad \sigma_v^2 = \sigma_\varepsilon^2 / (1-\beta_o^2)$$

$$\alpha_i \sim N(0,\sigma_\alpha^2) \qquad \varepsilon_{it} \sim N(0,\sigma_\varepsilon^2) \qquad v_{i0} \sim N(0,\sigma_v^2)$$

- Number of replications = 1000
- $y_{it}^b = y_{it} - b$ is used in the estimation, by changing *b* variously.

Table 1 presents the simulation results from the two-step GMM estimation of the simple dynamic model.

- *Reported statistics: bias*, *rmse* (root mean squared error), *mcsd* (Monte Carlo standard deviation) and *mcmse* (Monte Carlo mean of standard error)
- The distribution of the GMM-WC estimator changes depending on b (for each N, both the *mcsd* and *rmse* of the GMM-WC estimator increases as $|b|$ increases).
- For given N, the GMM-C estimation results are the same for any choice of b.

Table 1: Monte Carlo Results for a Dynamic Panel Data Model
(T=5, $\beta_0 = 0.5, \sigma_\varepsilon^2 = \sigma_\alpha^2 = 1$)

		N = 100		N = 500		N = 1000	
		bias	rmse	bias	rmse	bias	rmse
		(mcsd)	(mcmse)	(mcsd)	(mcmse)	(mcsd)	(mcsd)
b = 20	GMM-WC	−0.08	0.18	−0.01	0.08	−0.01	0.05
		(0.17)	(0.14)	(0.08)	(0.07)	(0.05)	(0.05)
	GMM-C	−0.06	0.14	−0.01	0.06	−0.01	0.04
		(0.13)	(0.10)	(0.06)	(0.05)	(0.04)	(0.04)
b = 2	GMM-WC	−0.06	0.16	−0.01	0.06	−0.01	0.05
		(0.15)	(0.12)	(0.06)	(0.06)	(0.04)	(0.04)
	GMM-C	−0.06	0.14	−0.01	0.06	−0.01	0.04
		(0.13)	(0.10)	(0.06)	(0.05)	(0.04)	(0.04)
b = 0	GMM-WC	−0.04	0.14	−0.01	0.06	−0.01	0.04
		(0.13)	(0.11)	(0.06)	(0.05)	(0.04)	(0.04)
	GMM-C	−0.06	0.14	−0.01	0.06	−0.01	0.04
		(0.13)	(0.10)	(0.06)	(0.05)	(0.04)	(0.04)
b = −2	GMM-WC	−0.06	0.16	−0.01	0.07	−0.01	0.05
		(0.15)	(0.12)	(0.06)	(0.06)	(0.05)	(0.04)
	GMM-C	−0.06	0.14	−0.01	0.06	−0.01	0.04
		(0.13)	(0.10)	(0.06)	(0.05)	(0.04)	(0.04)
b = −20	GMM-WC	−0.08	0.18	−0.01	0.08	−0.01	0.05
		(0.17)	(0.14)	(0.08)	(0.07)	(0.05)	(0.05)
	GMM-C	−0.06	0.14	−0.01	0.06	−0.01	0.04
		(0.13)	(0.10)	(0.06)	(0.05)	(0.04)	(0.04)

3.2. Count Panel Data Model with Predetermined Regressor

- Data Generating Process

$$y_{it} \sim \text{Poisson}[\exp(x_{it}\beta + \alpha_i + \varepsilon_{it} - (1/2)\sigma_\varepsilon^2)]$$

$$x_{i1} = \frac{\delta}{1-\rho}\eta_i + \frac{1}{\sqrt{1-\rho^2}}(\theta v_i + w_{i1}) \qquad x_{it} = \rho x_{i,t-1} + \delta\eta_i + \theta\varepsilon_{i,t-1} + w_{it}$$

$$\alpha_i \sim N(0,\sigma_\alpha^2) \quad \varepsilon_{it} \sim N(0,\sigma_\varepsilon^2) \quad v_i \sim N(0,\sigma_v^2) \quad w_{it} \sim N(0,\sigma_w^2)$$

- Number of replications = 1000
- $x_{it}^b = x_{it} - b$ is used in the estimation, by changing b variously.

19

Table 2 presents the Monte Carlo results from the two-step Chamberlain GMM estimation.

- The two-step GMM-C estimates are identical to the change of *b*, as discussed in Section 2.2.
- As predicted, the distribution of the two-step GMM-WC estimator changes as *b* changes (The performance of the GMM-WC estimator appears to be quite sensitive to *b*, especially when the sample size *N* is small.).

Table 2: Monte Carlo Results from Count Panel Data Models with Predetermined Regressors

(T=5, $\beta_0 = 0.5, \ \ \delta = 0.1, \ \rho = 0.8, \ \theta = 0.3, \ \sigma_\varepsilon^2 = \sigma_\alpha^2 = 0.3, \ \sigma_w^2 = 0.25$)

		N = 100		N = 500		N = 1000	
		bias (mcsd)	rmse (mcmse)	bias (mcsd)	rmse (mcmse)	bias (mcsd)	rmse (mcsd)
b = 1	GMM-WC	−0.09	0.33	−0.03	0.17	−0.01	0.11
		(0.32)	(0.25)	(0.17)	(0.15)	(0.11)	(0.11)
	GMM-C	−0.14	0.30	−0.03	0.13	−0.02	0.09
		(0.27)	(0.17)	(0.13)	(0.11)	(0.09)	(0.08)
b = 0.5	GMM-WC	−0.09	0.29	−0.03	0.13	−0.01	0.09
		(0.28)	(0.20)	(0.13)	(0.11)	(0.09)	(0.09)
	GMM-C	−0.14	0.30	−0.03	0.13	−0.02	0.09
		(0.27)	(0.17)	(0.13)	(0.11)	(0.09)	(0.08)
b = 0	GMM-WC	−0.14	0.36	−0.03	0.15	−0.02	0.10
		(0.33)	(0.21)	(0.14)	(0.12)	(0.10)	(0.08)
	GMM-C	−0.14	0.30	−0.03	0.13	−0.02	0.09
		(0.27)	(0.17)	(0.13)	(0.11)	(0.09)	(0.08)
b = −0.5	GMM-WC	−0.24	0.51	−0.09	0.30	−0.03	0.15
		(0.45)	(0.24)	(0.29)	(0.13)	(0.14)	(0.09)
	GMM-C	−0.14	0.30	−0.03	0.13	−0.02	0.09
		(0.27)	(0.17)	(0.13)	(0.11)	(0.09)	(0.08)
b = −1	GMM-WC	−0.27	0.54	−0.13	0.34	−0.05	0.21
		(0.46)	(0.25)	(0.32)	(0.15)	(0.20)	(0.10)
	GMM-C	−0.14	0.30	−0.03	0.13	−0.02	0.09
		(0.27)	(0.17)	(0.13)	(0.11)	(0.09)	(0.08)

3.3. Count Panel Data Model with Endogenous Regressor

- Data Generating Process

$$y_{it} \sim \text{Poisson}(\exp(x_{it}\beta + \alpha_i + \varepsilon_{it} - (1/2)\sigma_\varepsilon^2))$$

$$x_{it} = \rho x_{i,t-1} + \delta\eta_i + \theta\varepsilon_{it} + w_{it}$$

$$x_{i0} = \frac{\delta}{1-\rho}\alpha_i + \frac{1}{\sqrt{1-\rho^2}}(\theta\varepsilon_{i0} + w_{i0})$$

$$\alpha_i \sim N(0,\sigma_\alpha^2) \qquad \varepsilon_{it} \sim N(0,\sigma_\varepsilon^2) \qquad w_{it} \sim N(0,\sigma_w^2)$$

- Number of replications = 1000
- $x_{it}^b = x_{it} - b$ is used in the estimation, by changing b variously.

Table 3 presents the Monte Carlo results for the two-step Wooldridge GMM estimators.

- The finite sample performances of both the GMM-WC and GMM-C estimators deteriorate as |b| deviates from zero, especially when N = 100.
- As the distribution of x_{it} becomes more skewed to the positive or negative sides, the mcsd and rmse of the GMM-WC (and GMM-C) estimator get larger.
- However, the finite-sample performance of the GMM-C estimator is less sensitive to *b* than that of the GMM-WC estimator.

Corroborating that the asymptotic variance of the GMM-C estimator is invariant according to the change of b, but that of the GMM-WC is not, when Wooldridge transformation is used.

Table 3: Monte Carlo Results from Count Panel Data Models with Predetermined Regressors
(T=5, $\beta_0 = 0.5,\ \delta = 0.1,\ \rho = 0.8,\ \theta = 0.3,\ \sigma_\alpha^2 = 0.3,\ \sigma_\varepsilon^2 = 0.25,\ \sigma_w^2 = 0.25$)

		N = 100		N = 500		N = 1000	
		bias (mcsd)	rmse (mcmse)	bias (mcsd)	rmse (mcmse)	bias (mcsd)	rmse (mcsd)
b = 1	GMM-WC	−0.48	0.52	−0.22	0.25	−0.13	0.16
		(0.22)	(0.22)	(0.12)	(0.13)	(0.10)	(0.10)
	GMM-C	−0.39	0.44	−0.17	0.21	−0.10	0.13
		(0.19)	(0.18)	(0.11)	(0.11)	(0.08)	(0.08)
b = 0.5	GMM-WC	−0.21	0.27	−0.09	0.14	−0.05	0.10
		(0.17)	(0.19)	(0.10)	(0.12)	(0.08)	(0.09)
	GMM-C	−0.18	0.24	−0.08	0.13	−0.05	0.09
		(0.17)	(0.16)	(0.10)	(0.11)	(0.08)	(0.08)
b = 0	GMM-WC	0.01	0.16	0.00	0.10	0.00	0.08
		(0.16)	(0.19)	(0.10)	(0.12)	(0.08)	(0.09)
	GMM-C	0.01	0.16	0.01	0.10	0.00	0.08
		(0.16)	(0.17)	(0.10)	(0.11)	(0.08)	(0.08)
b = −0.5	GMM-WC	0.26	0.33	0.10	0.15	0.06	0.11
		(0.20)	(0.22)	(0.11)	(0.13)	(0.09)	(0.09)
	GMM-C	0.22	0.29	0.10	0.15	0.06	0.10
		(0.19)	(0.18)	(0.11)	(0.12)	(0.09)	(0.09)
b = −1	GMM-WC	0.67	0.77	0.28	0.33	0.16	0.20
		(0.38)	(0.30)	(0.17)	(0.16)	(0.12)	(0.11)
	GMM-C	0.56	0.64	0.23	0.28	0.13	0.17
		(0.32)	(0.24)	(0.16)	(0.13)	(0.11)	(0.09)

4. Concluding Remark

- Non-invariance problem in the panel GMM estimation based on the level-instruments-for-differenced-equations (LIDE) approach:
- When a constant is not used as instrument, the asymptotic and finite-sample distributions of the GMM estimators depend on overall means of the regressors used.
- This problem can be solved simply by including a constant into the instrument set.