On the Effect of Mean-Nonstationarity in Dynamic Panel Data Models*

Kazuhiko Hayakawa†
Department of Economics, Hiroshima University

This version: February 2009

Abstract

In this paper, we investigate the effect of mean-nonstationarity on the first-difference generalized method of moments (FD-GMM) estimator in dynamic panel data models. We find that when data is mean-nonstationary and the variance of individual effects is significantly larger than that of disturbances, the FD-GMM estimator performs quite well. We demonstrate that this is because the correlation between the lagged dependent variable and instruments gets larger owing to the unremoved individual effects, i.e., instruments become strong. This implies that, under mean-nonstationarity, the FD-GMM estimator does not always suffer from the weak instruments problem even when data is persistent.

Keywords: Dynamic panel data models, strength of instruments, generalized method of moments estimator, mean-nonstationarity.
JEL classification: C13, C23.

*This paper is a revised version of chapter three of my Ph.D dissertation submitted to Hitotsubashi University.

†I am deeply grateful to the editor Cheng Hsiao, the two anonymous referees, Taku Yamamoto, Satoru Kanoh, Katsuto Tanaka, Eiji Kurozumi, Ryo Okui, and the seminar participants at Hitotsubashi University for helpful comments. I also acknowledge the financial support from the JSPS Fellowship and the Grant-in-Aid for Scientific Research (KAKENHI 20830056) provided by the JSPS. All remaining errors are mine. Corresponding author: Kazuhiko Hayakawa, E-mail : kazuhaya@hiroshima-u.ac.jp

1 Introduction

In the literature of dynamic panel data models, since the work of Anderson and Hsiao (1981, 1982), it has been recognized that appropriate formulation of initial conditions is very important because it determines the feature of a variable. For the GMM estimators, it is known that the consistency of the FD-GMM estimator by Arellano and Bond (1991) does not depend on the formulation of initial conditions. This is a substantial distinction between the FD-GMM estimator and the level and system GMM estimators (Arellano and Bover, 1995; Blundell and Bond, 1998) whose consistency is obtained by assuming a specific form for initial conditions that render the variable mean-stationary. However, little is known about the effect of initial conditions on the behavior of the FD-GMM estimator. Therefore, in this paper, we investigate how the finite sample properties of the FD-GMM estimator are affected by initial conditions, especially, initial conditions that render a variable mean-nonstationary[1].

2 Setup

Let us consider an AR(1) panel data model given by

$$y_{it} = \alpha y_{i,t-1} + \eta_i + v_{it} \qquad (i = 1, ..., N;\ t = 1, ..., T),$$

where α is the parameter of interest with $|\alpha| < 1$. We assume that $v_{it} \sim iid(0, \sigma_v^2)$, $\eta_i \sim iid(0, \sigma_\eta^2)$, and both are mutually independent. By letting $x_{it} = y_{i,t-1}$ and removing the individual effects by the forward orthogonal deviation, we have

$$y_{it}^* = \alpha x_{it}^* + v_{it}^* \qquad (i = 1, ..., N;\ t = 1, ..., T-1),$$

where $y_{it}^* = c_t \left[y_{it} - (y_{i,t+1} + \cdots + y_{iT})/(T-t) \right]$ and $c_t^2 = (T-t)/(T-t+1)$. x_{it}^* and v_{it}^* are defined in the same way. The FD-GMM estimator is given by

$$\widehat{\alpha} = \frac{\sum_{t=1}^{T-1} x_t^{*\prime} M_t y_t^*}{\sum_{t=1}^{T-1} x_t^{*\prime} M_t x_t^*} = \alpha + \frac{\sum_{t=1}^{T-1} x_t^{*\prime} M_t v_t^*}{\sum_{t=1}^{T-1} x_t^{*\prime} M_t x_t^*},$$

where $x_t^* = (x_{1t}^*, ..., x_{Nt}^*)'$, $y_t^* = (y_{1t}^*, ..., y_{Nt}^*)'$, $M_t = Z_t (Z_t' Z_t)^{-1} Z_t'$, $Z_t = (z_{1t}, ..., z_{Nt})'$, and $z_{it} = (y_{i0}, ..., y_{i,t-1})'$.

In order to investigate the effect of initial conditions, we need to formulate them. Although there are several ways to formulate initial conditions, in this paper, we assume that

$$y_{i0} = \delta \mu_i + w_{i0} \qquad (i = 1, ..., N), \quad (1)$$

where $\mu_i = \eta_i/(1-\alpha)$ and $w_{i0} = \sum_{j=0}^{\infty} \alpha^j v_{i,-j}$[2]. Note that this type of initial conditions renders y_{it} to be mean-nonstationary in the sense that the conditional mean of y_{it} given η_i depends on t as follows:

$$E(y_{it}|\eta_i) = \left[1 - (1-\delta)\alpha^t \right] \mu_i.$$

From this, we find that when $\delta \neq 1$, y_{it} is mean-nonstationary owning to the dependence on t, and when $\delta = 1$, y_{it} is mean-stationary.

[1] In the working paper version (Hayakawa, 2008), we also consider GMM and LIML estimators using instruments in first difference and instruments proposed by Hayakawa (2009), and a within-groups estimator.

[2] Although we can consider other types of initial conditions such as $y_{i0} = \delta \mu_i + \dot{w}_{i0}$ with $\dot{w}_{i0} \neq \sum_{j=0}^{\infty} \alpha^j v_{i,-j}$, we only focus on form (1) to simplify the discussion.

Allowing for mean-nonstationarity has important implications for empirical analyses. For example, when we consider a cross-country panel data set that begins after a war or another large historical event, or panel data of young workers or new firms, mean-stationarity may not hold[3]. In fact, Arellano (2003) provides empirical evidence of mean-nonstationarity from the estimation result of employment and wage equations using Spanish firms panel data.

3 Main result

To investigate the effect of mean-nonstationarity on the FD-GMM estimator, we consider an alternative expression of the FD-GMM estimator as follows:

$$\widehat{\alpha} = \frac{\sum_{t=1}^{T-1} x_t^{*\prime} M_t y_t^*}{\sum_{t=1}^{T-1} x_t^{*\prime} M_t x_t^*} = \frac{\sum_{t=1}^{T-1} x_t^{*\prime} M_t x_t^* \cdot \widehat{\alpha}_t}{\sum_{t=1}^{T-1} x_t^{*\prime} M_t x_t^*},$$

where $\widehat{\alpha}_t = x_t^{*\prime} M_t y_t^* / x_t^{*\prime} M_t x_t^*$ is the cross section two-stage least squares estimator at time t obtained from

$$\begin{aligned} y_{it}^* &= \alpha x_{it}^* + v_{it}^* \qquad (i = 1, ..., N), \\ x_{it}^* &= \pi_t' z_{it} + \varepsilon_{it}. \end{aligned}$$

Since $\widehat{\alpha}$ is a weighted sum of $\widehat{\alpha}_t$, it is expected that the properties of $\widehat{\alpha}_t$ will carry over to $\widehat{\alpha}$. Hence, in the following, we focus on $\widehat{\alpha}_t$.

As is well known, the correlation between x_{it}^* and z_{it} plays a very important role in the performance of $\widehat{\alpha}_t$. Before computing the correlation, note that x_{it}^* can be written as

$$x_{it}^* = \psi_t \left[w_{i,t-1} - (1-\delta)\alpha^{t-1}\mu_i \right] - c_t \tilde{v}_{itT}, \tag{2}$$

where $\psi_t = c_t \left(1 - \frac{\alpha\phi_{T-t}}{T-t}\right)$, $\phi_j = (1-\alpha^j)/(1-\alpha)$, $w_{i,t-1} = \sum_{j=0}^{\infty} \alpha^j v_{i,t-1-j}$, and $\tilde{v}_{itT} = (\phi_{T-t} v_{it} + \cdots + \phi_1 v_{i,T-1})/(T-t)$. It is obvious that the individual effect μ_i is removed from x_{it}^* only when $\delta = 1$. Using (2), $E(x_{it}^* z_{it})$ is given by

$$\begin{aligned} E(x_{it}^* z_{it}) &= \psi_t E[w_{i,t-1} z_{it}] - (1-\delta)\psi_t \alpha^{t-1} E[\mu_i z_{it}] \qquad (3) \\ &= \text{"idiosyncratic part"} + \text{"individual effects part."} \end{aligned}$$

Note that for given $\alpha, t,$ and T, the magnitudes of $E[w_{i,t-1} z_{it}]$ and $E[\mu_i z_{it}]$ are determined by σ_v^2 and σ_η^2, respectively. This implies that the relative ratio σ_η^2/σ_v^2 plays a key role in comparing the magnitudes of the "idiosyncratic part" and "individual effects part."

From (3), we find that when $\delta = 1$, the correlation between x_{it}^* and z_{it} is solely due to the idiosyncratic term; otherwise, it is due to both the "idiosyncratic part" *and* the "individual effects part." This indicates that mean-nonstationarity provides an additional correlation between x_{it}^* and z_{it} through the unremoved individual effects. However, we have to investigate the "individual effects part" carefully since it can be both positive and negative, while the "idiosyncratic part" is always positive. When $\delta > 1$, since the "individual effects part" is always positive, $E[x_{it}^* z_{it}]$ increases with σ_η^2. However, when $\delta < 1$, $E(x_{it}^* z_{it})$ might be close to zero since the "idiosyncratic part" is positive while the "individual effects part" is negative. In this case, the instruments may be weak. However, if σ_η^2/σ_v^2 is large enough, the "individual effects part" becomes much smaller than the "idiosyncratic part" and $E(x_{it}^* z_{it})$ can be large in terms of its absolute value. Therefore, when σ_η^2/σ_v^2 is large, the instruments become strong regardless of whether $\delta > 1$ or $\delta < 1$.

[3]See Barro and Sala-i-Martin (1995) and Hause (1980).

Further, we find that α determines the magnitude of the "individual effects part". It is easy to see that when α is small, the magnitude of the "individual effects part" also becomes small and when α approaches unity, the effect of the "individual effects part" becomes strong, i.e., the instruments become strong even for the near unit root case if σ_η^2/σ_v^2 is large. This may suggest that the weak instruments problem of the FD-GMM estimator pointed out by Blundell and Bond (1998) may not occur even when α is close to unity. Note that this difference comes from the assumption of the initial conditions[4].

We now provide some simulation results. Figures 1 and 2 depict simulation results for the mean of $\widehat{\alpha}$ with $(T, N) = (9, 100)$, $\alpha = 0.3, 0.9$, and $\sigma_\eta^2/\sigma_v^2 = 0.2, 0.5, 1, 10$. With regards to δ, we use $\delta = (1-\alpha)/(1-\bar{\alpha})$ with $\bar{\alpha}$ ranging from $\alpha - 0.1000$ to $\alpha + 0.0975$ in steps of 0.0025.[5] Note that $\bar{\alpha} = 0.3$, 0.9 correspond to the case of $\delta = 1$. Also note that $0.875 \le \delta \le 1.162$ when $\alpha = 0.3$, and $0.5 \le \delta \le 40$ when $\alpha = 0.9$.[6] The number of replications is 5,000 for each design[7].

From the figures, it is confirmed that (i) when α is small, the effect of mean-nonstationarity is small, (ii) when $\alpha = 0.9$, the behavior of $\widehat{\alpha}$ heavily depends on δ and σ_η^2/σ_v^2, and (iii) when σ_η^2/σ_v^2 is large and δ is moderately deviated from 1, the bias of $\widehat{\alpha}$ becomes quite small.

In the above discussion, we showed that the behavior of the FD-GMM estimator is heavily affected by the initial conditions. However, in time series analysis, it is widely known that the effects of initial conditions vanish as T gets larger. We show that the same applies to the dynamic panel data models considered here. To see this, note that using an asymptotic expansion, we have $E\left(\widehat{\alpha} - \alpha\right) = E\left(\sum_{t=1}^{T-1} x_t^{*\prime} M_t v_t^*\right) / E\left(\sum_{t=1}^{T-1} x_t^{*\prime} M_t x_t^*\right) + O(1/NT)$ (see Bun and Kiviet 2006). With regard to the numerator, Alvarez and Arellano (2003) shows that $E\left((NT)^{-1} \sum_{t=1}^{T-1} x_t^{*\prime} M_t v_t^*\right) = -\sigma_v^2/N(1-\alpha^2)\left[1 - 1/(T(1-\alpha)) \sum_{t=1}^{T} (1-\alpha^t)/t\right]$. This implies that the numerator is not affected by the initial conditions since it does not depend on δ. However, this is not the case for the denominator. Specifically, we have $(NT)^{-1} \sum_{t=1}^{T-1} E(x_t^{*\prime} M_t x_t^*) = (NT)^{-1} \sum_{t=1}^{T-1} E(\psi_t^2 w_{t-1}' M_t w_{t-1} + c_t^2 \tilde{v}_{tT}' M_t \tilde{v}_{tT} + (1-\delta)^2 \psi_t^2 \alpha^{2(t-1)} \mu' M_t \mu - 2(1-\delta) \psi_t^2 \alpha^{t-1} \mu' M_t w_{t-1})$. After some algebra, it is shown that the first and second terms are $O(1)$ and $O(\log T/N)$, respectively, while the last two terms that are associated with the initial conditions are both $O(1/T)$. Thus, we find that the effect of initial conditions characterized by δ gets weaker as T increases.

4 Conclusion

In this paper, we investigated how the FD-GMM estimator is affected by mean-nonstationarity. We showed that under mean-nonstationarity, unremoved individual effects provide an additional correlation between the lagged dependent variable and instruments, which substantially affects the finite sample behavior of the FD-GMM estimator. An important result found in this paper is that the weak instruments problem does not always occur in the FD-GMM estimator even if the persistency of data is strong. On the contrary, we found that if data is mean-nonstationary and the variance of individual effect is much larger than that of the idiosyncratic term, instruments become strong.

Finally, we note possible extensions. First, since we usually do not know the strength of instruments in practice, it may be interesting to investigate the performance of a robust

[4]Blundell and Bond (1998) pointed out the weak instruments problem under the assumption of mean-stationarity.

[5]Note that y_{i0} can be written as $y_{i0} = \delta\mu_i + w_{i0} = 1/(1-\bar{\alpha})\eta_i + w_{i0}$, where $\delta = (1-\alpha)/(1-\bar{\alpha})$.

[6]δ becomes quite large since $\bar{\alpha} = 0.9975$ is very close to unity. If we set $\bar{\alpha} = 0.99(0.95)$, $\delta = 10(2)$.

[7]Further results are reported in Hayakawa (2008).

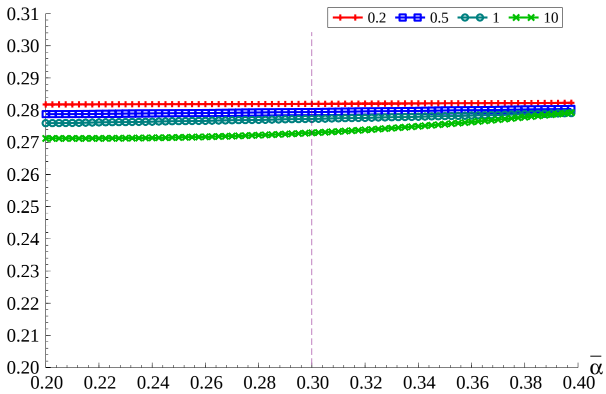

Figure 1: Mean of $\widehat{\alpha}$ ($\alpha = 0.3$)

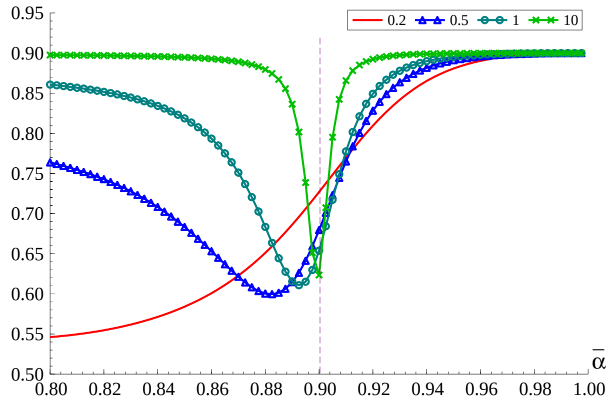

Figure 2: Mean of $\widehat{\alpha}$ ($\alpha = 0.9$)

statistic to weak instruments such as Kleibergen (2005). Second, although the model considered in this paper is limited to a stable AR(1) panel model, it is important to extend the analysis to models with additional regressors. However, the finding that unremoved individual effects provide an additional correlation between the lagged dependent variable and instruments may be useful when considering more general models.

References

[1] Alvarez, J. and M. Arellano (2003) "The Time Series and Cross-Section Asymptotics of Dynamic Panel Data Estimators," *Econometrica*, 71, 4, 1121–1159.

[2] Anderson, T. W. and C. Hsiao (1981) "Estimation of Dynamic Models with Error Components," *Journal of the American Statistical Association*, 76, 375, 598–606.

[3] Anderson, T. W. and C. Hsiao (1982) "Formulation and Estimation of Dynamic Models Using Panel Data," *Journal of Econometrics*, 18, 1, 47–82.

[4] Arellano, M. (2003) "Modelling Optimal Instrumental Variables For Dynamic Panel Data Models," working papers.

[5] Arellano, M. and S. Bond (1991) "Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations," *Review of Economic Studies*, 58, 2, 277–97.

[6] Arellano, M. and O. Bover (1995) "Another Look at the Instrumental Variable Estimation of Error-Components Models," *Journal of Econometrics*, 68, 1, 29–51.

[7] Barro, R. J. and X. S. i Martin (1995) *Economic Growth*, New York: McGraw-Hill.

[8] Blundell, R. and S. Bond (1998) "Initial Conditions and Moment Restrictions in Dynamic Panel Data Models," *Journal of Econometrics*, 87, 1, 115–143.

[9] Bun, M. J. G. and J. F. Kiviet (2006) "The Effects of Dynamic Feedbacks on LS and MM Estimator Accuracy in Panel Data Models," *Journal of Econometrics*, 127, 2, 409–444.

[10] Hause, J. (1980) "The Fine Structure of Earnings and the On-the-Job Training Hypothesis," *Econometrica*, 48, 4, 1013–1029.

[11] Hayakawa, K. (2008) "On the Effect of Nonstationary Initial Conditions in Dynamic Panel Data Models." Hi-Stat DP series, No. 245, Hitotsubashi University.

[12] Hayakawa, K. (2009) "A Simple Efficient Instrumental Variable Estimator in Panel AR(p) Models," forthcoming in *Econometric Theory*.

[13] Kleibergen, F. (2005) "Testing Parameters in GMM Without Assuming that They Are Identified," *Econometrica*, 73, 4, 1103–1123.