



Article

# FDML versus GMM for Dynamic Panel Models with Roots Near Unity

Adrian Mehic

Department of Economics, Lund University, SE-223 63 Lund, Sweden; adrian.mehic@nek.lu.se

**Abstract:** This paper evaluates the first-differenced maximum likelihood (FDML) and the continuously updating system generalized method of moments (CU-GMM) estimators of dynamic panel models when the data is close to non-stationary. This case is far from trivial, as a high degree of persistence is the norm rather than the exception in economic panels, particularly in financial management. While the CU-GMM is shown to have lower bias and higher power, it suffers from severe size distortions, which are exacerbated when the data approaches non-stationarity.

**Keywords:** dynamic panel data; persistence; FDML estimation

**JEL Classification:** C13; C23



**Citation:** Mehic, Adrian. 2021. FDML versus GMM for Dynamic Panel Models with Roots Near Unity. *Journal of Risk and Financial Management* 14: 405. <https://doi.org/10.3390/jrfm14090405>

Academic Editor: Shuddhasattwa Rafiq

Received: 12 July 2021

Accepted: 23 August 2021

Published: 26 August 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Dynamic panel data models are of crucial importance in empirical econometrics, since their applications can be found in virtually all subfields of economics. For empirical finance research in particular, several of the most cited papers in recent years have applied the dynamic panel model Beck et al. (2000); Midrigan and Xu (2014); Wintoki et al. (2012). In the financial management literature, the dynamic panel model has been used to answer, for instance, whether cash flow impacts innovation (Brown et al. 2009; Carpenter and Guariglia 2008), and how dividend smoothing differs between private and public firms (Michaely and Roberts 2012). The model is chiefly used when the number of cross-sections,  $N$ , is high, but the number of time periods,  $T$ , is relatively small.

The econometric techniques for parameter estimation in the dynamic panel model have traditionally been based on the generalized method of moments, GMM. Two GMM-based methods for dynamic panels have been particularly successful: the *difference GMM* estimator, which is due to Arellano and Bond (1991), and the *system GMM* of Arellano and Bover (1995) and Blundell and Bond (1998). A likelihood-based estimator, the *first-differenced ML* (FDML), was developed by Hsiao et al. (2002). However, the performance of both these classes of estimators in the situation in which the data is close to non-stationary remains relatively unexplored. For the GMM, the system estimator is usually preferred to the difference estimator (cf. Blundell and Bond 1998; Hayakwa and Pesaran 2015). This is because of the considerable increase in bias that results from using the difference GMM in lieu of the system GMM for highly persistent data. Additionally, Hansen et al. (1996) have made further contributions to GMM theory by introducing the so-called *continuously updating GMM* (CU-GMM), which can be shown to decrease the finite-sample bias of GMM estimators.<sup>1</sup> As a high degree of persistence is the norm rather than an exception in economic panels, this is a situation with potentially serious ramifications. The finite-sample properties of the FDML in the almost non-stationary setting has not been evaluated either. Failing to perform in the nearly non-stationary case would be considered a serious drawback of this relatively new method, potentially limiting its usability in practical situations.

In this paper, I compare the local-to-unit-root performance of the system CU-GMM, which was tailor-made for this setting, with the FDML. The results of the Monte Carlo

simulations show that the FDML has higher absolute bias in the nearly non-stationary case, particularly for low values of  $T$ , compared to the continuously updating system GMM. Moreover, for low values of  $T$ , the power of the FDML is lower than for the system CU-GMM. However, the CU-GMM estimator suffers from severe oversizing, which is exacerbated as  $N$  and  $T$  increase. This size distortion problem was not seen when applying the FDML.

This paper contributes to a rich literature on the CU-GMM and FDML estimators of dynamic panel models. Recent contributions on the CU-GMM include the work by [Ashley and Sun \(2016\)](#), who adjust the standard two-step estimator by applying continuous updating on the autoregressive parameter only, and by [Kleibergen and Zhan \(2021\)](#) on CU-GMM robustness tests for under weak identification and potential misspecification. Further, a second strain in the literature focuses on the sensitivity of the CU-GMM to the choice of instruments. [Newey \(2004\)](#) gives standard errors for the CU-GMM that adjust for the number of overidentifying restrictions, whereas [Donald et al. \(2009\)](#) develop moment selection criteria for various GMM-based estimators, including the CU-GMM. Recent research on the FDML includes, for instance, extensions to the situation with arbitrary initial conditions ([Kruiniger 2018b](#)).

Despite its good theoretical properties, the FDML is more seldom used by practitioners compared to the GMM. What is more, recent research has shown that the FDML outperforms the GMM in terms of size, power, and bias in most finite-sample cases: [Elhorst \(2010\)](#) and [Hsiao and Zhang \(2015\)](#) provide evidence when comparing with the Arellano-Bond difference GMM, and [Hayakwa and Pesaran \(2015\)](#) extend this to include the system GMM, including the CU-adjusted estimator. However, none of these papers consider the situation with local to unit root data<sup>2</sup>, for which the Arellano-Bond estimator is inconsistent due to weak instruments. The present paper aims to fill this gap.

The rest of the paper is structured as follows. Section 2 introduces the CU-GMM and FDML estimators for the dynamic panel model. Section 3 describes the Monte Carlo design. Section 4 presents the results of the simulation study. The paper concludes with Section 5.

## 2. Theory

The AR(1) dynamic panel data model can be described by

$$y_{it} = \alpha_i + \phi y_{i,t-1} + u_{it} \quad (1)$$

for individuals  $i = 1, \dots, N$  and time periods  $t = 1, \dots, T$ , where  $\alpha_i$  are the fixed effects,  $\phi$  is the autoregressive (AR) parameter, and  $u_{it}$  is the idiosyncratic error term. It is assumed that the error terms  $u_{it}$  are independent and identically distributed, and  $\mathbb{E}[u_{it}] = 0$ ,  $\mathbb{V}[u_{it}] = \sigma_u^2$ ,  $\mathbb{E}[|u_{it}|^4] < \infty$  and  $\mathbb{E}[u_{it}u_{is}] = 0$  for each  $t \neq s$ . Further, it is assumed that the initial observations  $y_{i0} = O_p(1)$  are observed, and that  $\mathbb{E}[\alpha_i] = 0$ ,  $\mathbb{V}[\alpha_i] = \sigma_\alpha^2 = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N \alpha_i^2 < \infty$ ,  $\mathbb{E}[\alpha_i u_{it}] = 0$ , and  $\mathbb{E}[y_{i0} u_{it}] = 0$ . Using the fixed effects estimator to estimate (1) gives biased estimates of  $\phi$  ([Nickell 1981](#)). For  $T$  large, we can write  $\text{plim}_{N \rightarrow \infty} (\hat{\phi} - \phi) \approx -(1 + \phi)/(T - 1)$ , meaning that when  $\phi$  is near unity, the bias can be sizable.

The first to propose an unbiased estimator of (1) were [Anderson and Hsiao \(1981\)](#). However, the Anderson-Hsiao estimator is asymptotically inefficient for all values of  $\phi$ , and the absolute bias and variance of the estimator both increase significantly when  $\phi$  approaches unity (cf. [Arellano and Bover 1995](#)). These problems have caused the empirical importance of the Anderson-Hsiao estimator to decline significantly. Instead, this paper considers GMM and FDML, which dominate present-day empirical research.

### 2.1. CU-GMM

The first estimation technique considered in this paper is the system CU-GMM. I shall focus particularly on the performance of the CU version of the GMM in this paper, as it has received relatively little attention in the literature compared to the one-step and two-step GMM estimators.

Let  $\pi_{it} = \alpha_i + u_{it}$ . Then, [Arellano and Bond \(1991\)](#) show that for  $t = 3, \dots, T$ , the moment conditions  $\mathbb{E}[y_{is}\Delta\pi_{is}] = 0$ , where  $\Delta$  is the difference operator, can be utilized. If  $u_{it} \sim MA(0)$ , it holds that  $s \geq 2$ , and  $s \geq 3$  if  $u_{it} \sim MA(1)$ . Now, the instrument matrix  $\mathbf{Z}_i$  is

$$\mathbf{Z}_i = \begin{pmatrix} y_{i1} & 0 & 0 & \dots & 0 & \dots & 0 \\ 0 & y_{i1} & y_{i2} & \dots & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \dots & \vdots & \dots & \vdots \\ 0 & 0 & 0 & \dots & y_{i1} & \dots & y_{i,T-2} \end{pmatrix} \tag{2}$$

and the vector of first-differenced errors is  $\Delta\pi_i = (\Delta\pi_{i3}, \dots, \Delta\pi_{iT})'$ . Using this notation, the moment conditions can be written  $\mathbb{E}[\mathbf{Z}'_i\Delta\pi_i] = \mathbf{0}$  for  $i = 1, \dots, N$ . However, [Blundell and Bond \(1998\)](#) show that the Arellano-Bond estimator significantly underestimates  $\phi$  in the local-to-unity case. In order to remedy this problem, it is possible to introduce additional moment conditions, namely  $\mathbb{E}[\pi_{it}\Delta y_{i,t-s}] = 0$  for  $t = 3, \dots, T$  and  $i = 1, \dots, N$ . For  $u_{it} \sim MA(0)$ , it holds that  $s = 1$ , and if  $u_{it} \sim MA(1)$ , then  $s = 2$ . The joint moment conditions can be written in matrix form as  $\mathbb{E}[\tilde{\mathbf{Z}}'_i\pi_i^*] = \mathbf{0}$  where  $\tilde{\mathbf{Z}}_i = \text{diag}(\mathbf{Z}_i, \Delta y_{i2}, \Delta y_{i3}, \dots, \Delta y_{i,T-1})$ ,  $\pi_i^* = (\Delta\pi_i, \pi_i)'$ , where  $\Delta\pi_i$  is as defined previously and  $\pi_i = (\pi_{i3}, \dots, \pi_{iT})'$ , for  $i = 1, \dots, N$ . Using these additional moment conditions, the Blundell-Bond (otherwise known as the *system GMM*) CU estimator is the solution to the optimization problem

$$\arg \min_{\phi \in \Phi} \left( \frac{1}{N} \sum_{i=1}^N \pi_i^*{}' \tilde{\mathbf{Z}}_i \right) \mathbf{W}_N(\phi) \left( \frac{1}{N} \sum_{i=1}^N \tilde{\mathbf{Z}}_i' \pi_i^* \right) \tag{3}$$

where  $\Phi$  is the compact set of all possible parameters and  $\mathbf{W}_N$  is a positive semi-definite weight matrix, for which it holds that  $\mathbf{W}_N \xrightarrow{P} \mathbf{W}$  by the law of large numbers. The CU technique allows for the weight matrix to be a function of  $\phi$ . Thus, instead of fixing  $\mathbf{W}_N$  in each stage of the estimation, it is altered as the value of  $\hat{\phi}$  changes during the minimization process. This reduces the finite-sample bias of the estimator without altering its asymptotic properties ([Hansen et al. 1996](#); [Newey and Smith 2004](#); [Pakes and Pollard 1989](#)).

I will use  $\hat{\phi}_{GMM}$  to denote the continuously updating system GMM. Under stationarity, the following theorem regarding the limiting distribution of  $\hat{\phi}_{GMM}$  holds.

**Theorem 1.** *Given  $|\phi| < 1$ , if  $(\log T)^2/N \rightarrow 0$ , and provided that  $T/N \rightarrow c, 0 \leq c < \infty$ , the limiting distribution of the GMM estimator  $\hat{\phi}_{GMM}$  of  $\phi$  is*

$$\sqrt{NT} \left[ \hat{\phi}_{GMM} - \left( \phi - \frac{1}{N}(1 + \phi) \right) \right] \xrightarrow{L} \mathcal{N}(0, 1 - \phi^2) \tag{4}$$

**Proof.** See [Alvarez and Arellano \(2003\)](#). □

The asymptotic bias is  $O_p(\sqrt{T/N})$ , which implies that as  $N/T \rightarrow \infty$ , the asymptotic bias disappears. Moreover, the limiting variance in (4) is equal to the Cramér-Rao lower bound ([Hahn and Kuersteiner 2002](#)). Hence, as  $N/T \rightarrow \infty$ , the GMM estimator is asymptotically efficient. However, if  $|\phi| = 1$ , the limiting distribution is no longer normal, as manifested by Theorem 2.

**Theorem 2.** *For  $|\phi| = 1$ , the limiting distribution of the GMM estimator  $\hat{\phi}_{GMM}$  of  $\phi$  is*

$$\sqrt{T}(\hat{\phi}_{GMM} - 1) \xrightarrow{L} 2\mathcal{C} \tag{5}$$

as  $N, T \rightarrow \infty$  jointly.

**Proof.** See [Phillips \(2014\)](#). □

Here,  $\mathcal{C}$  denotes a standard Cauchy variate. However, it can be shown that normality can break down even for values of  $\phi$  local to unity, although it has not been empirically

tested how close to unity  $\phi$  must be in order for the limiting distribution to become Cauchy (Phillips 2014).

### 2.2. FDML

An alternative approach to estimating  $\phi$  is by using FDML. In order to eliminate  $\alpha_i$ , take the first difference of (1) to obtain

$$\Delta y_{it} = \phi \Delta y_{i,t-1} + \Delta u_{it} \tag{6}$$

For  $t = 1$ , the above expression is not well defined, since  $\Delta y_{i1} = \phi \Delta y_{i0} + \Delta u_{i1}$  and  $\Delta y_{i0}$  is not observable. However, by continuous substitution,

$$\Delta y_{i1} = \phi^m \Delta y_{i,-m+1} + \sum_{j=0}^{m-1} \phi^j \Delta u_{i,1-j} = \phi^m \Delta y_{i,-m+1} + \eta_{i1} \tag{7}$$

Now, the analysis will differ slightly depending on whether the process is stationary or not. Assume first that  $|\phi| < 1$  and  $m \rightarrow \infty$ . Then, it holds for  $t = 3, \dots, T$  and  $i = 1, \dots, N$  that  $\mathbb{E}[\Delta y_{i1}] = 0$ ,  $\mathbb{V}[\Delta y_{i1}] = 2\sigma_u^2 / (1 + \phi)$ ,  $\mathbb{C}[\eta_{i1}, \Delta u_{i2}] = -\sigma_u^2$ , and  $\mathbb{C}[\eta_{i1}, \Delta u_{it}] = 0$ . Alternatively, if  $|\phi| \geq 1$ , the process has started from a finite point  $m$  that is behind the 0:th time point, so that  $\mathbb{E}[\Delta y_{i1}] = b$ ,  $\mathbb{V}[\Delta y_{i1}] = c\sigma_u^2$ ,  $\mathbb{C}[\eta_{i1}, \Delta u_{i2}] = -\sigma_u^2$ , and  $\mathbb{C}[\eta_{i1}, \Delta u_{it}] = 0$ , where  $b, c \in \mathbb{R}_+$ . Assuming stationary increments, let  $\Delta \mathbf{y}_i$  be as defined previously, and  $\Delta \mathbf{u}_i^* = (\Delta y_{i1} - b^*, \Delta u_{i2}, \dots, \Delta u_{iT})'$ . Here,  $b^* = 0$  if  $|\phi| < 1$  and  $b^* = b$  if  $|\phi| \geq 1$ . The covariance matrix of  $\Delta \mathbf{u}_i^*$  is  $\mathbb{C}(\Delta \mathbf{u}_i^*) = \mathbf{\Omega} = \sigma_u^2 \mathbf{\Omega}^*$ , where  $\mathbf{\Omega}^*$  is

$$\mathbf{\Omega}^* = \begin{pmatrix} \omega & -1 & 0 & \dots & 0 \\ -1 & 2 & -1 & \dots & 0 \\ 0 & -1 & 2 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & -1 & 2 \end{pmatrix} \tag{8}$$

where  $\omega = (1/\sigma_u^2)\mathbb{V}(\Delta y_{i1})$ . This is equal to  $2/(1 + \phi)$  if  $|\phi| < 1$ , and  $c$  else. To find the likelihood function of  $\Delta \mathbf{y}_i$ , use that  $\Delta \mathbf{u}_i^*$  is a linear combination of  $\Delta \mathbf{y}_i$ , and that the Jacobian of this transformation is equal to unity. Thus, the joint probability density functions (p.d.fs) of  $\Delta \mathbf{u}_i^*$  and  $\Delta \mathbf{y}_i$  are equal. Then, assuming that the  $u_{it}$ 's are independent normal, the joint p.d.f. of  $\Delta \mathbf{y}_i$  is equal to the likelihood function of (1), and is given by

$$\mathcal{L} = \prod_{i=1}^N (2\pi)^{-T/2} |\mathbf{\Omega}|^{-1/2} \exp \left\{ -\frac{1}{2} \mathbf{u}_i^{*'} \mathbf{\Omega}^{-1} \mathbf{u}_i^* \right\} \tag{9}$$

The corresponding log-likelihood is

$$\log \mathcal{L} = \frac{-NT}{2} \log(2\pi) - \frac{N}{2} \log |\mathbf{\Omega}| - \frac{1}{2} \sum_{i=1}^N \mathbf{u}_i^{*'} \mathbf{\Omega}^{-1} \mathbf{u}_i^* \tag{10}$$

The two unknown elements of  $\mathbf{\Omega}$  are  $\sigma_u^2$  and  $\omega$ . Proceeding from here, the FDML technique involves utilizing the Anderson-Hsiao estimator  $\hat{\phi}_{AH}$  to find an initial estimate of  $\phi$ . Then, the variance  $\sigma_u^2$  is estimated by  $\hat{\sigma}_u^2 = \sum_{i=1}^N \sum_{t=2}^T (\Delta y_{it} - \hat{\phi}_{AH} \Delta y_{i,t-1})^2 / 2N(T - 2)$ . In the stationary case,  $\omega$  can be estimated by  $2(1 + \hat{\phi}_{AH})^{-1}$ . In the non-stationary case,  $\omega$  is estimated by  $\hat{\omega} = \frac{1}{(N-1)\hat{\sigma}_u^2} \sum_{i=1}^N (\Delta y_{i1} - \hat{b})^2$ , where a consistent estimator of  $b$  is  $\hat{b} = N^{-1} \sum_{i=1}^N \Delta y_{i1}$ . Then, (10) is maximized numerically until convergence.

Theorems 3 and 4 provide the asymptotic results for the FDML.

**Theorem 3.** Given  $|\phi| < 1$ , for  $T$  large and  $N$  arbitrary, the limiting distribution of the FDML estimator  $\hat{\phi}_{FDML}$  of  $\phi$  is

$$\sqrt{NT}(\hat{\phi}_{FDML} - \phi) \xrightarrow{L} \mathcal{N}(0, 1 - \phi^2) \tag{11}$$

**Proof.** See [Kruiniger \(2008a\)](#). □

Hence, the FDML estimator is asymptotically unbiased, asymptotically normal and, for  $T$  large, asymptotically efficient. For the unit root case, Theorem 4 holds.

**Theorem 4.** For  $|\phi| = 1$ , the limiting distribution of the FDML estimator is

$$T\sqrt{N}(\hat{\phi}_{FDML} - 1) \xrightarrow{L} \mathcal{N}(0, 8) \tag{12}$$

as  $N, T \rightarrow \infty$  jointly.

**Proof.** See [Kruiniger \(2008a\)](#). □

The rate of convergence is  $O_p(T\sqrt{N})$ , which is faster than the rate of convergence in the stationary case, which is  $O_p(\sqrt{NT})$  according to Theorem 3. Note that normality of the FDML does not break down in the unit root case, which is a clear contrast to the GMM.

### 3. Monte Carlo Setup

I now proceed to evaluate the performance of the CU-GMM and FDML in the close-to-non-stationary setting. The model of interest is

$$y_{it} = (1 - \phi_0)\alpha_i + \phi_0 y_{i,t-1} + u_{it} \tag{13}$$

where  $u_{it} \sim \mathcal{N}(0, 1)$  or  $u_{it} \sim \mathcal{N}(0, 2)$ . Individual effects are generated according to  $\alpha_i = (\lambda_i - 1)/\sqrt{2}$ , where  $\lambda_i \sim \chi^2(1)$ , so that  $\mathbb{E}[\alpha_i] = 0$  and  $\mathbb{V}[\alpha_i] = 1$ .<sup>3</sup> For the initial observations  $y_{i0}$ , I set  $y_{i0} = \alpha_i / (1 - \phi_0) + u_{i0}$ , where  $u_{i0} \sim \mathcal{N}[0, 1 / (1 - \phi_0^2)]$ . The AR parameter is varied according to  $\phi_0 \in \{0.90, 0.95, 0.99\}$ . The number of Monte Carlo replications is set to 1000.

### 4. Results

Tables 1–3 present the mean and median biases, size, and power for the FDML and CU system GMM, using standard standard errors in computing size and power. Table 1 corresponds to  $\phi = 0.90$ , Table 2 to  $\phi = 0.95$ , and Table 3 is for the case when  $\phi = 0.99$ .

**Table 1.** Mean bias, median bias, size and power for  $\phi = 0.90$ .

N/T	FDML					GMM				
	5	10	20	30	50	5	10	20	30	50
	<b>Mean bias (<math>\times 100</math>)</b>									
50	−7.333	−1.501	−0.037	0.025	0.090	0.739	−1.496	−	−	−
150	−4.692	0.120	0.225	−0.006	−0.056	−2.088	−1.087	−0.709	−0.659	−0.050
500	−2.345	0.537	0.052	0.048	−0.034	−0.942	−0.439	−0.114	−0.173	−0.123
	<b>Median bias (<math>\times 100</math>)</b>									
50	−4.397	−0.826	−0.288	−0.135	−0.0090	3.323	0.889	−	−	−
150	−2.993	0.077	−0.017	−0.010	0.423	−2.933	−0.071	−0.407	−0.369	1.342
500	−1.512	0.184	−0.041	0.023	−0.035	0.196	−0.053	−0.240	−0.056	−0.058
	<b>Size</b>									
50	15.0	16.6	17.7	9.6	4.8	54.5	80.0	−	−	−
150	14.5	20.1	10.2	6.1	5.8	32.6	40.7	57.0	71.2	97.7
500	13.5	18.8	6.3	5.3	5.0	16.5	18.4	22.8	22.8	33.3
	<b>Power (<math>H_1 : \phi = 0.80</math>)</b>									
50	26.6	43.1	75.6	98.8	100.0	60.4	89.3	−	−	−
150	32.1	51.2	98.5	99.9	100.0	32.3	86.0	97.2	97.5	98.7
500	38.7	79.2	99.1	100.0	100.0	83.3	98.1	100.0	100.0	100.0

**Table 2.** Mean bias, median bias, size and power for  $\phi = 0.95$ .

N/T	FDML					GMM				
	5	10	20	30	50	5	10	20	30	50
<b>Mean bias (<math>\times 100</math>)</b>										
50	-9.397	-3.618	-0.593	0.018	0.042	2.526	-0.551	-	-	-
150	-6.623	-2.045	-0.075	0.129	0.034	-0.148	-1.419	-1.165	-0.842	0.355
500	-4.297	-0.785	0.172	0.074	0.053	-1.053	-0.876	-0.314	0.192	-0.141
<b>Median bias (<math>\times 100</math>)</b>										
50	-5.272	-1.964	-0.184	-0.055	-0.071	3.286	3.543	-	-	-
150	-3.379	-1.419	-0.261	-0.003	0.079	-3.679	-0.010	-0.301	-0.369	1.268
500	-2.458	0.424	-0.033	-0.047	0.056	-0.310	-0.111	-0.113	-0.076	-0.023
<b>Size</b>										
50	14.9	15.9	17.9	19.3	12.5	68.6	87.1	-	-	-
150	14.4	14.2	19.5	16.8	7.5	51.3	63.4	76.4	80.2	98.6
500	18.5	13.9	17.7	9.4	3.6	29.4	36.8	38.4	41.0	50.1
<b>Power (<math>H_1 : \phi = 0.85</math>)</b>										
50	23.5	37.0	68.0	95.3	99.6	89.0	94.8	-	-	-
150	31.2	46.8	96.3	99.4	99.9	86.1	91.5	96.3	97.5	99.6
500	35.7	64.2	98.9	99.8	100.0	87.0	95.2	100.0	100.0	100.0

**Table 3.** Mean bias, median bias, size and power for  $\phi = 0.99$ .

N/T	FDML					GMM				
	5	10	20	30	50	5	10	20	30	50
<b>Mean bias (<math>\times 100</math>)</b>										
50	-12.133	-5.181	-2.394	-1.382	-0.711	1.429	0.066	-	-	-
150	-8.169	-3.529	-1.550	-0.850	-0.361	1.411	0.082	0.090	0.064	-0.262
500	-5.734	-2.599	-0.923	-0.465	-0.183	0.479	0.059	0.027	-0.315	-0.224
<b>Median bias (<math>\times 100</math>)</b>										
50	-8.922	-3.323	-1.660	-0.798	-0.384	0.915	0.925	-	-	-
150	-4.920	-2.098	-1.037	-0.265	-0.207	0.973	0.790	0.809	0.785	0.908
500	-3.219	-1.659	-0.547	-0.270	-0.165	0.861	0.636	0.349	0.149	0.082
<b>Size</b>										
50	18.1	15.2	12.5	13.1	14.5	77.6	92.0	-	-	-
150	15.5	13.2	13.3	13.0	12.5	71.4	87.9	93.0	96.2	99.1
500	13.5	15.1	13.5	13.6	14.7	67.6	80.5	82.9	86.6	89.4
<b>Power (<math>H_1 : \phi = 0.89</math>)</b>										
50	15.0	33.7	57.8	90.0	97.2	97.8	99.0	-	-	-
150	26.9	41.7	89.2	95.2	98.1	97.8	98.7	99.6	99.6	99.3
500	36.1	48.2	95.1	97.5	99.8	98.5	98.1	100.0	100.0	100.0

The absolute biases are generally larger when  $\phi$  is closer to unity; this applies for both estimators. For  $\phi = 0.90$  and  $\phi = 0.95$ , the FDML generally outperforms the GMM in terms of bias except when  $T$  is low. When  $\phi = 0.99$ , the GMM estimator is better for virtually every combination of  $N$  and  $T$ ; the exceptions being the mean bias for the combination of  $T = 50$  and  $N = 500$ , as well as the median biases for the  $T = 30$  when  $N = 150$  and  $500$ . The absolute mean and median biases are close to zero for the GMM when  $\phi = 0.99$ . The FDML performs relatively poorly when  $\phi$  is this close to unity, especially when  $T < 20$ .

Considering that the Blundell-Bond version of the GMM is more or less tailor-made for the situation with close-to-non-stationary data, the bias results do not come as a major surprise. Also, due to the close relationship between  $\hat{\phi}_{FDML}$  and the poor initial estimate obtained from the Anderson-Hsiao estimator, one would expect higher absolute FDML bias when  $\phi$  is closer to unity, given the discussion in Section 2. This is confirmed by the Monte Carlo results. The mean and median biases are generally negative for the FDML, while for GMM, they tend to be negative for  $\phi = 0.90$  and  $\phi = 0.95$ , and positive for  $\phi = 0.99$ .

While the performance of the GMM is superior to the FDML in terms of bias, the size is considerably higher than 5% for the GMM, irrespective of the value of  $\phi$ . Additionally, the size of the GMM estimator is increasing both with the AR parameter  $\phi$  and with the number of time periods,  $T$ . This effect is not observed in the likelihood estimator. The results further show that the size problem for the GMM is exacerbated when  $\phi$  is very close to one. For example, when  $\phi = 0.99$ , the size is above 96% even for  $N = 150$  and  $T = 30$ .



Hence, if imposed such that  $|\phi| < 1$ , the value of  $\phi$  is an important factor in explaining the performance of size.

Regarding the power, it is considerably lower for the FDML, especially for small values of  $T$ . Additionally, the power of the FDML is deteriorating as  $\phi$  approaches unity. However, when  $T = 30$  and  $50$ , the power of both estimators is close to 100%.

## 5. Concluding Remarks

This paper compares the bias, size and power of the system CU-GMM and FDML estimators of the AR coefficient in the dynamic panel model when the process is highly persistent, that is, when AR coefficient is local to unit root. This setting is particularly important in empirical finance, as most financial data is persistent.

The main finding of the paper is the relatively large increase in absolute bias of the FDML as the value of the AR parameter approaches unity. Moreover, the absolute bias of the CU-GMM is lower for most combinations of  $N$  and  $T$ . However, the CU-GMM is shown to suffer from severe size distortions, implying the existence of a trade-off between precision and size. The size is increasing both with the value of the AR parameter and with the number of time periods  $T$ . This result is a further contribution of this paper. The power of both the FDML is considerably lower than that of the GMM when  $T$  is small. However, the power of both estimators is shown to be close to 100% with  $N$  and  $T$  sufficiently large.

Overall, although the FDML has slightly higher bias than the CU-GMM when  $\phi$  is close to unity, its high power and correct size makes it a viable option to the hitherto dominating GMM-based methods in most empirical settings. Thus, the findings in this paper have broader implications for applications in the financial management field. In empirical research, FDML estimates could be reported in lieu, or alongside, GMM estimates when using the AR(1) dynamic panel model. Additionally, the FDML allows practitioners to estimate the dynamic panel model when the number of time periods is close to the number cross-sectional units ( $N$ ).<sup>4</sup> This is a fairly common situation when using financial management data.

Finally, since dynamic panel models are highly useful in the financial sector, the results of this paper have implications for decision-makers within the sector. This is because precision in econometric estimates is crucial for making correct investment decisions. Specifically, finance practitioners specialized in global quantitative strategy may use the FDML as an alternative to the CU-GMM when, for example, monitoring for bubbles in equity prices, or for comparing capital structure and payout policy between firms.

A limitation of the study is that there are several other GMM-based estimators widely used by practitioners, for instance the one and two step estimators. However, the continuously updating weight matrix is tailor-made for the situation with local to unit root data, making it the obvious competitor to the FDML in the present setting.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Acknowledgments:** I am grateful to Shudda Rafiq, Sisi Huang, as well as four anonymous referees. This work has benefited from discussions with David Edgerton, Joakim Westerlund, and seminar participants at Lund University.

**Conflicts of Interest:** I declare no conflict of interest.

## Notes

- <sup>1</sup> For additional examples of bias-reducing methods for GMM estimators of dynamic panel models, see [Choi et al. \(2010\)](#) or [Mehic \(2020\)](#).
- <sup>2</sup> I consider processes for which the autoregressive parameter is greater than 0.9 to be highly persistent. The maximum values for the autoregressive parameter were 0.8 in [Elhorst \(2010\)](#) and [Hsiao and Zhang \(2015\)](#), and 0.9 in [Hayakwa and Pesaran \(2015\)](#).
- <sup>3</sup> The parameterization in (15), where the individual effects are multiplied by  $(1 - \phi_0)$ , is a standard approach in the literature when dealing with almost non-stationary data ([Han and Phillips 2013](#); [Bun et al. 2017](#)). Without this correction, the individual effects would have too much of an impact on the results when the true value  $\phi_0$  is close to unity.
- <sup>4</sup> In Tables 1–3, this situation occurs, for instance when  $N = 50$  and  $T = 20$ . Here, the CU-GMM cannot be estimated, as there are too few instruments. However, the FDML can be estimated with low bias and high power in this setting, even with  $\phi = 0.99$ .

## References

- Alvarez, Javier, and Manuel Arellano. 2003. The Time Series and Cross-Section Asymptotics of Dynamic Panel Data Estimators. *Econometrica* 71: 1121–59. [\[CrossRef\]](#)
- Anderson, Theodore Wilbur, and Cheng Hsiao. 1981. Estimation of Dynamic Models with Error Components. *Journal of the American Statistical Association* 76: 598–606. [\[CrossRef\]](#)
- Arellano, Manuel, and Olympia Bover. 1995. Another Look at the Instrumental Variables Estimation of Error Components Models. *Journal of Econometrics* 68: 29–52. [\[CrossRef\]](#)
- Arellano, Manuel, and Stephen Bond. 1991. Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations. *Review of Economic Studies* 58: 277–97. [\[CrossRef\]](#)
- Ashley, Richard A., and Xiaojin Sun. 2016. Subset-Continuous-Updating GMM Estimators for Dynamic Panel Data Models. *Econometrics* 4: 47. [\[CrossRef\]](#)
- Beck, Thorsten, Ross Levine, and Norman Loayza. 2000. Finance and the Sources of Growth. *Journal of Financial Economics* 58: 261–300. [\[CrossRef\]](#)
- Blundell, Richard, and Stephen Bond. 1998. Initial Conditions and Moment Restrictions in Dynamic Panel Data Models. *Journal of Econometrics* 87: 115–43. [\[CrossRef\]](#)
- Brown, James R., Steven M. Fazzari, and Bruce C. Petersen. 2009. Financing Innovation and Growth: Cash Flow, External Equity, and the 1990s R&D Boom. *Journal of Finance* 64: 151–85.
- Bun, Maurice J.G., Martin A. Carree, and Artūras Juodis. 2017. On Maximum Likelihood Estimation of Dynamic Panel Data Models. *Oxford Bulletin of Economics and Statistics* 79: 463–94. [\[CrossRef\]](#)
- Carpenter, Robert E., and Alessandra Guariglia. 2008. Cash Flow, Investment, and Investment Opportunities: New Tests Using UK Panel Data. *Journal of Banking & Finance* 32: 1894–906.
- Choi, Chi-Young, Nelson C. Mark, and Donggyu Sul. 2010. Bias Reduction in Dynamic Panel Data Models by Common Recursive Mean Adjustment. *Oxford Bulletin of Economics and Statistics* 72: 567–99. [\[CrossRef\]](#)
- Donald, Stephen G., Guido W. Imbens, and Whitney K. Newey. 2009. Choosing Instrumental Variables in Conditional Moment Restriction Models. *Journal of Econometrics* 152: 28–36. [\[CrossRef\]](#)
- Elhorst, J. Paul. 2010. Dynamic panels with endogenous interaction effects when T is small. *Regional Science and Urban Economics* 40: 282. [\[CrossRef\]](#)
- Hahn, Jinyong, and Guido Kuersteiner. 2002. Asymptotically Unbiased Inference for a Dynamic Panel Model with Fixed Effects when Both  $n$  and  $T$  Are Large. *Econometrica* 70: 1639–57. [\[CrossRef\]](#)
- Han, Chirok, and Peter C.B. Phillips. 2013. First difference maximum likelihood and dynamic panel estimation. *Journal of Econometrics* 175: 35–45. [\[CrossRef\]](#)
- Hansen, Lars Peter, John Heaton, and Amir Yaron. 1996. Finite-Sample Properties of Some Alternative GMM Estimators. *Journal of Business and Economic Statistics* 14: 262–80.
- Hayakawa, Kazuhiko, and M. Hashem Pesaran. 2015. Robust Standard Errors in Transformed Likelihood Estimation of Dynamic Panel Data Models with Cross-Sectional Heteroskedasticity. *Journal of Econometrics* 188: 111–34. [\[CrossRef\]](#)
- Hsiao, Cheng, and Junwei Zhang. 2015. IV, GMM or Likelihood Approach to Estimate Dynamic Panel Models When Either N or T or Both Are Large. *Journal of Econometrics* 187: 312–22. [\[CrossRef\]](#)
- Hsiao, Cheng, M. Hashem Pesaran, and A. Kamil Tahmiscioglu. 2002. Maximum Likelihood Estimation of Fixed Effects Dynamic Panel Data Models Covering Short Time Periods. *Journal of Econometrics* 109: 107–50. [\[CrossRef\]](#)
- Kleibergen, Frank, and Zhaoguo Zhan. 2021. Double Robust Inference for Continuous Updating GMM. Unpublished Manuscript.
- Kruiniger, Hugo. 2008a. Maximum Likelihood Estimation and Inference Methods for the Covariance Stationary Panel AR(1)/Unit Root Model. *Journal of Econometrics* 144: 447–64. [\[CrossRef\]](#)
- Kruiniger, Hugo. 2018b. Uniform Quasi ML Based Inference for the Panel AR(1) Model. Unpublished Manuscript.
- Mehic, Adrian. 2020. Half-Panel Jackknife Estimation for Dynamic Panel Models. *Economics Letters* 190: 109082. [\[CrossRef\]](#)
- Michaely, Roni, and Michael R. Roberts. 2012. Corporate Dividend Policies: Lessons from Private Firms. *Review of Financial Studies* 25: 711–46. [\[CrossRef\]](#)



- 
- Midrigan, Virgiliu, and Daniel Yi Xu. 2014. Finance and Misallocation: Evidence from Plant-Level Data. *American Economic Review* 104: 422–58. [\[CrossRef\]](#)
- Newey, Whitney K. 2004. Many Weak Moment Asymptotics for the Continuously Updated GMM Estimator. Unpublished Manuscript.
- Newey, Whitney K., and Richard J. Smith. 2004. Higher Order Properties of GMM and Generalized Empirical Likelihood Estimators. *Econometrica* 72: 219–55. [\[CrossRef\]](#)
- Nickell, Stephen. 1981. Biases in Dynamic Models with Fixed Effects. *Econometrica* 49: 1417–26. [\[CrossRef\]](#)
- Pakes, Ariel, and David Pollard. 1989. Simulation and the Asymptotics of Optimization Estimators. *Econometrica* 57: 1027–57. [\[CrossRef\]](#)
- Phillips, Peter C. B. 2014. *Dynamic Panel GMM with Roots Near Unity*. Working Paper. New Haven: Yale University.
- Wintoki, M. Babajide, James S. Linck, and Jeffrey M. Netter. 2012. Endogeneity and the Dynamics of Internal Corporate Governance. *Journal of Financial Economics* 105: 581–606. [\[CrossRef\]](#)