Working Paper Series
Department of Economics
University of Verona

Improving GMM efficiency in dynamic models for panel data with mean stationarity

Giorgio Calzolari, Laura Magazzini

WP Number: 12 July 2014

ISSN: 2036-2919 (paper), 2036-4679 (online)

Improving GMM efficiency in dynamic models for panel data with mean stationarity

Giorgio Calzolari* Laura Magazzini†‡

July 1, 2014

Abstract

Within the framework of dynamic panel data models with mean stationarity, one additional moment condition may remarkably increase the efficiency of the system GMM estimator. This additional condition is essentially a condition of "homoskesdasticity" of the individual effects; it is "implicitly satisfied" in all the Monte Carlo simulations on dynamic panel data models available in the literature (including the experiments with heteroskedasticity, which is always confined to the idiosyncratic errors), but not "explicitly" exploited. Monte Carlo experiments show remarkable efficiency improvements when the distribution of individual effects, and thus of y_{i0}, are skewed, thus including the very important cases in economic applications that include variables like individual wages, sizes of the firms, number of employees, etc.

Keywords: panel data, dynamic model, GMM estimation, mean stationarity, skewed individual effects
JEL Codes: C23, C13

1 Introduction

We consider the estimation of dynamic linear panel data models on N independent units observed over T time periods ($i = 1, ..., N;\ t = 1, ..., T$):

$$y_{it} = \beta y_{it-1} + \alpha_i + e_{it} = \beta y_{it-1} + \varepsilon_{it} \quad (1)$$

where α_i captures individual 'unobserved heterogeneity', constant over time and different across units, and e_{it} is an idiosyncratic component changing both over time and across units. We

*Department of Statistics, Informatics, Applications, Università di Firenze, <calzolar@disia.unifi.it>
†Department of Economics, Università di Verona, <laura.magazzini@univr.it>
‡We gratefully acknowledge comments and suggestions from Francesca Mantese, and conference participants at the Fifth Italian Congress of Econometrics and Empirical Economics, Genova, January 16-18, 2013. This work has benefited from financial support of the project MIUR (PRIN): MISURA – Multivariate Models for Risk Assessment.

consider the typical case of micro panels, where N is "large" and T is "small". The initial y_{i0} is observed. We further assume $|\beta| < 1$.

Since the seminal work of Arellano and Bond (1991), estimation of dynamic panel data models relies on the generalized method of moments (GMM; see Hansen, 1982).[1]

In the context of the linear dynamic panel data models, two GMM approaches are customarily considered for estimation: the difference GMM estimator and the system GMM estimator (Arellano and Bond, 1991; Arellano and Bover, 1995; Blundell and Bond, 1998). The latter approach is known to produce efficiency improvements, which are particularly strong when the autoregression coefficient (β) approaches unity. This is obtained at the cost of a more restrictive hypothesis on the initial y_{i0}, known as the "mean stationarity" assumption. Throughout the paper, we will assume that such initial conditions are satisfied (Blundell and Bond, 1998).

What we shall show in this paper is that if the initial conditions of mean stationarity are interpreted as they are in "all" simulation experiments presented in the rich literature on this subject, further remarkable efficiency improvements can be obtained from the simple addition of "one more" moment condition to the system GMM estimator.

In the next Section we briefly describe the two traditional approaches with their assumptions, and the additional moment condition we propose. A detailed set of Monte Carlo experiments in Section 3 shows the efficiency improvements over system GMM estimator. The efficiency gains are sometimes quite remarkable. For example, in the case $N = 3000$ and 3 observation in time on the dependent variable, under the assumption of y_{i0} composed of a log-normally distributed individual effect and a log-normally distributed idiosyncratic component, the variance of the estimator we propose is about one-fifth of the variance of the system GMM estimator.[2] A kernel density estimation of the distributions of the estimated β in the corresponding Monte Carlo experiment is reported in Figure 1.

[1]Other approaches are also available for estimation of linear dynamic panel data models. Hsiao, Pesaran, and Tahmiscioglu (2002) developed a transformed likelihood approach for the estimation of linear dynamic model within a fixed effect framework. Kiviet (1995) proposed a method to correct the small sample bias of the least squares dummy variable estimator. The performance of simulation methods (indirect inference) has also been analysed (Gouriéroux, Phillips, and Yu, 2010). In this paper we focus on GMM methods.

[2]For further details, please refer to Section 3.

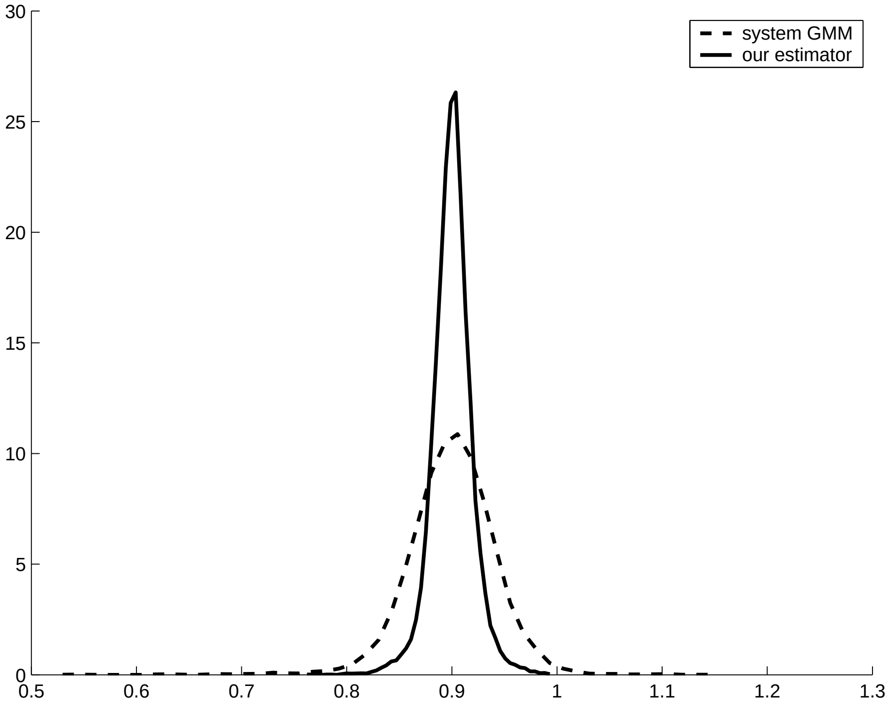

Figure 1: Monte Carlo results of one of the experiments reported in Section 3 ($\beta = 0.9$)

2 GMM estimation of dynamic panel data models

Considering the estimation of dynamic linear panel data model (1), a consistent estimation of β can be obtained by transforming the model in first differences, thus removing the individual effect α_i. Then, under suitable assumptions on the correlation structure of e_{it}, lagged values of y_{it} can be used as instruments in the differenced equation (Anderson and Hsiao, 1981, 1982; Arellano and Bond, 1991). In particular, as e_{it} is customarily assumed to be uncorrelated over time, lag 2 or more of y_{it} are valid instruments for Δy_{it-1}. Thus, the following $T(T-1)/2$ linear moment conditions can be exploited for estimation (Arellano and Bond, 1991):

$$E(y_{it-j}\Delta\varepsilon_{it}) = 0 \tag{2}$$

with $j, t \geq 2$. The GMM estimator based on the set of moment conditions in (2) is known in the literature as difference GMM estimator.

Simulation studies have shown that the difference GMM estimator may have non-negligible finite sample bias and poor performances when β approaches 1, or when the variance of α_i is large with respect to the variance of e_{it} (Blundell and Bond, 1998).[3]

Blundell and Bond (1998) introduce an additional assumption on the initial observation (see also Arellano and Bover, 1995):

$$E(\varepsilon_{i2}\Delta y_{i1}) = 0 \tag{3}$$

This is known as the "mean stationarity" assumption.

If this is satisfied, the following T moment conditions hold ($t \geq 2$):[4]

$$E(\varepsilon_{it}\Delta y_{it-1}) = 0 \tag{4}$$

These conditions exclude cases in which y_{i0} is exogenous or when its correlation with the individual effect α_i is different from the correlation of α_i with y_{it} at different times.

The GMM estimator based on the moment conditions in (2) and (4) is known as the "system

[3] Ahn and Schmidt (1995) show that the difference GMM estimator is not fully exploiting all the information in the data: under the standard assumptions, they propose additional moment conditions that allow improving efficiency.

[4] Of course, longer lags could be considered, but these conditions are redundant when the moment conditions (2) are included for estimation (Arellano and Bover, 1995; Blundell and Bond, 1998).

GMM estimator". When the moment conditions (4) are added to the moment conditions in (2), substantial efficiency gains are achieved with respect to difference GMM estimator (Hahn, 1997). Besides gains in (asymptotic) efficiency, the system GMM estimator also performs better in terms of small sample bias (Hayakawa, 2007).[5]

Blundell and Bond (1998) state that condition (3) is satisfied in cases in which

$$y_{i0} = \frac{\alpha_i}{1-\beta} + e_{i0} \tag{5}$$

with e_{i0} uncorrelated with α_i (see also Roodman, 2009). Equation (5) has become the hallmark of simulation studies in which the data generating process for y_{i0} obeys the mean stationarity assumption. Thus we propose to assume explicitly eq. (5) and exploit the fact that for all individuals i, α_i enters the first observation y_{i0} multiplied by the same factor for each i (equal to $1/(1-\beta)$). This allows to add one moment condition to the system GMM estimator.

The condition we propose is more restrictive than assumptions needed for consistency of the system GMM estimator. For instance, condition (3) would hold even if

$$E(y_{i0}\varepsilon_{it}) = \frac{\sigma^2_{\alpha,i}}{(1-\beta)}$$

for each $t \geq 1$, thus with a variance $\sigma^2_{\alpha,i}$ changing across individuals. However, (5) is implied by "all" Monte Carlo designs of the literature (at least to the best of our knowledge), where in fact condition (5) is always adopted (Arellano and Bover, 1995; Blundell and Bond, 1998; Windmeijer, 2000; Roodman, 2009).[6] Even though settled in the Monte Carlo experiments, none of the above GMM estimators exploits such a condition. It seems therefore quite reasonable to try to further improve the GMM efficiency with the introduction of one additional moment condition that explicitly exploits this condition.

It is quite unlikely (perhaps impossible) that linear conditions analogous to those used by the methods available in the literature can distinguish between initial values y_{i0} generated as

[5] In addition, moment conditions (4) and (2) are shown to encompass the moment conditions proposed by Ahn and Schmidt (1995) (Blundell and Bond, 1998).

[6] Even when simulation studies deal explicitly with heteroskedasticity (e.g. (Windmeijer, 2000)), it is always the e_{it} effect to be considered heteroskedastic, while the α_i effects are always homoskedstic. This is also true for Monte Carlo experiments related to estimators other than GMM, e.g. Gouriéroux, Phillips, and Yu (2010).

in (5) from values generated by more complicated structures.[7] Nonlinear moment conditions should therefore be considered. In principle, many new moment conditions could be introduced. Being nonlinear, it is not difficult to prevent possible redundancies. But, to avoid excess of moment conditions that could be dangerous for moderately small samples, the additional moment conditions should be very few (Roodman, 2009).

We therefore propose the "one" nonlinear moment condition, to be added to the system GMM moments (in the case in which mean stationarity can be assumed), leading to efficiency gains (sometimes quite remarkable, see our experiments in Section 3):[8]

$$E\left(z_{i0}^2\bar{\alpha}_{i1} - \frac{1}{1-\beta}z_{i0}\bar{\alpha}_{i2}\right) = 0 \tag{6}$$

where

- $z_{i0} = y_{i0} - \bar{y}_0$ with $\bar{y}_0 = \sum_{i=1}^N y_{i0}/N$ (the "demeaned" initial value of y_{i0}),
- $\bar{\alpha}_{i1} = \sum_{t=1}^T \varepsilon_{it}/T$,
- $\bar{\alpha}_{i2} = \sum_{t=2}^T \sum_{s<t} \varepsilon_{it}\varepsilon_{is}/(T(T-1)/2)$.

If y_{i0} is generated as in (5), then condition (6) holds. Being the additional moment condition only one, it would be quite simple to introduce it into the available algorithms or software packages. The variance of the GMM estimator based on the moment condition we propose involves third moments of y_{i0}. This point will be discussed in Appendix A. Monte Carlo experiments in Section 3 will show that the main efficiency gains are indeed obtained when the distribution of y_{i0} (either due to the distribution of α_i or e_{i0} or both) is asymmetric. However, notice that no explicit assumption on the distribution is ever required. But the additional condition (6) can be particulraly helpful in applications that involve individual effects like "wage of individuals", or "size of the firms", which are variables typically with strongly skewed distributions (in the Monte Carlo experiments we shall use "demeaned" log-normal or χ^2 distributions).

[7] As an example, condition (3) would be also satisfied if $y_{i0} = \gamma_i\alpha_i + e_{i0}$ with a random γ_i independent of α_i and $E(\gamma_i) = 1/(1-\beta)$.

[8] We have also explored moment conditions that involve the fourth order moment of α_i, but without obtaining any significant benefit.

3 Monte Carlo experiments

In this Section a set of Monte Carlo experiments is conducted in order to assess the finite sample performance of the GMM estimator based on moment conditions (2), (4), and (6) as compared to the system GMM estimator.

We generate a dynamic panel data model as:

$$y_{it} = \beta y_{it-1} + \varepsilon_{it} = \beta y_{it-1} + \alpha_i + e_{it} \tag{7}$$

with $|\beta| < 1$, and e_{it} normally distributed with mean zero and unconditional variance equal to 1, independent of α_i. As for the distribution of α_i we considered the normal distribution, and two asymmetric distributions, i.e. the log-normal distribution and the χ^2 distribution with 1 degrees of freedom. In all cases, when needed, appropriate transformation of α_i are considered in order to have mean zero and the chosen value for its unconditional variance.

The initial observation, y_{i0} is generated as:

$$y_{i0} = \frac{\alpha_i}{1-\beta} + e_{i0} \tag{8}$$

and different distributions are also considered for e_{i0}.

Results of Monte Carlo experiments are reported in Table 1, where we set $N = 100$, and both the unconditional variance of α_i and e_{it} equal to 1, so that the two error components give equally weighted contribution to the unconditional variance of y_{it} (both equal to 1). In Table 2, the unconditional variance of α_i is set 4 times larger than the unconditional variance of e_{it} (still set equal to 1). Tables 3 and 4 report the results for the same set of experiments with $N = 3000$.[9]

As expected, when the distribution of α_i and e_{i0} is symmetric (as it is the case with the normal distribution) no gain in performance is detected, whereas the results of the Monte Carlo experiments show remarkable gains in performance when the distribution of α_i and e_{i0} (thus, of y_{i0}) is asymmetric (log-normal or chi-squared distribution). Larger gains in performance are observed when T is small, and they reduce when T increases. Results with large N ($N = 3000$)

[9] In all cases we considered the continosly updated version of the GMM estimator (Hansen *et al.*, 1996).

$T+1$	β	Distribution of α_i, e_{i0}	System GMM			C-M		
			mean	var.	RMSE	mean	var.	RMSE
3	0.5	normal	.4985	.4231E-01	.4231E-01	.5090	.4475E-01	.4483E-01
8	0.5	normal	.5030	.5211E-02	.5220E-02	.5044	.5296E-02	.5315E-02
3	0.7	normal	.6904	.5498E-01	.5507E-01	.7029	.5457E-01	.5458E-01
8	0.7	normal	.7023	.5628E-02	.5633E-02	.7024	.5580E-02	.5586E-02
3	0.9	normal	.8932	.7769E-01	.7774E-01	.9099	.5661E-01	.5671E-01
8	0.9	normal	.8969	.6140E-02	.6150E-02	.8955	.5843E-02	.5863E-02
3	0.5	log-N	.5146	.3599E-01	.3620E-01	.5012	.2003E-01	.2003E-01
8	0.5	log-N	.4983	.4645E-02	.4648E-02	.4902	.3957E-02	.4053E-02
3	0.7	log-N	.7111	.4731E-01	.4743E-01	.7048	.1996E-01	.1998E-01
8	0.7	log-N	.6982	.4467E-02	.4470E-02	.6929	.3479E-02	.3529E-02
3	0.9	log-N	.9173	.6848E-01	.6878E-01	.9091	.2387E-01	.2395E-01
8	0.9	log-N	.8981	.4274E-02	.4278E-02	.8978	.2827E-02	.2832E-02
3	0.5	χ^2_1	.5091	.4133E-01	.4141E-01	.4979	.2178E-01	.2178E-01
8	0.5	χ^2_1	.5004	.4835E-02	.4835E-02	.4884	.4048E-02	.4183E-02
3	0.7	χ^2_1	.7022	.5388E-01	.5388E-01	.7025	.2306E-01	.2307E-01
8	0.7	χ^2_1	.6996	.5000E-02	.5000E-02	.6903	.3799E-02	.3893E-02
3	0.9	χ^2_1	.9070	.7996E-01	.8001E-01	.9062	.3218E-01	.3222E-01
8	0.9	χ^2_1	.8977	.5255E-02	.5260E-02	.8956	.3336E-02	.3355E-02

Table 1: Results of Monte Carlo experiments, $N = 100$, 10000 replications

$T+1$	β	Distribution of α_i, e_{i0}	System GMM			C-M		
			mean	var.	RMSE	mean	var.	RMSE
3	0.5	normal	.5212	.5519E-01	.5564E-01	.5618	.5615E-01	.5997E-01
8	0.5	normal	.5047	.5880E-02	.5902E-02	.5084	.6279E-02	.6350E-02
3	0.7	normal	.7013	.7109E-01	.7109E-01	.7502	.6036E-01	.6288E-01
8	0.7	normal	.7050	.7034E-02	.7059E-02	.7095	.7204E-02	.7294E-02
3	0.9	normal	.9297	.1010	.1019	.9449	.4570E-01	.4772E-01
8	0.9	normal	.8979	.7672E-02	.7676E-02	.9043	.6879E-02	.6897E-02
3	0.5	log-N	.5303	.4959E-01	.5051E-01	.5119	.3178E-01	.3192E-01
8	0.5	log-N	.4994	.5449E-02	.5449E-02	.4943	.5485E-02	.5517E-02
3	0.7	log-N	.7206	.6392E-01	.6434E-01	.7125	.3366E-01	.3382E-01
8	0.7	log-N	.6970	.5764E-02	.5773E-02	.6910	.4812E-02	.4893E-02
3	0.9	log-N	.9346	.9587E-01	.9707E-01	.9233	.4442E-01	.4496E-01
8	0.9	log-N	.8950	.6075E-02	.6100E-02	.8930	.4314E-02	.4363E-02
3	0.5	χ^2_1	.5279	.5620E-01	.5698E-01	.5199	.4024E-01	.4064E-01
8	0.5	χ^2_1	.5021	.5534E-02	.5538E-02	.4959	.5553E-02	.5570E-02
3	0.7	χ^2_1	.7145	.7309E-01	.7330E-01	.7198	.4347E-01	.4386E-01
8	0.7	χ^2_1	.7005	.6267E-02	.6267E-02	.6917	.5143E-02	.5212E-02
3	0.9	χ^2_1	.9343	.1085	.1097	.9256	.5040E-01	.5106E-01
8	0.9	χ^2_1	.8972	.7020E-02	.7028E-02	.8941	.5130E-02	.5165E-02

Table 2: Results of Monte Carlo experiments, $N = 100$, unconditional variance of α_i equal to 4 and unconditional variance of e_{it} equal to 1, 10000 replications

$T+1$	β	Distribution of α_i, e_{i0}, e_{it}	System GMM			C-M		
			mean	var.	RMSE	mean	var.	RMSE
3	0.5	normal	.4996	.1339E-02	.1339E-02	.4995	.1346E-02	.1346E-02
8	0.5	normal	.5000	.8954E-04	.8954E-04	.5001	.8959E-04	.8960E-04
3	0.7	normal	.7000	.1370E-02	.1370E-02	.6998	.1384E-02	.1384E-02
8	0.7	normal	.7000	.9669E-04	.9669E-04	.7000	.9661E-04	.9661E-04
3	0.9	normal	.9004	.1379E-02	.1379E-02	.9002	.1413E-02	.1413E-02
8	0.9	normal	.9002	.8480E-04	.8484E-04	.8999	.8503E-04	.8504E-04
3	0.5	log-N	.5001	.1367E-02	.1367E-02	.4988	.5377E-03	.5391E-03
8	0.5	log-N	.5001	.8946E-04	.8947E-04	.4989	.7220E-04	.7341E-04
3	0.7	log-N	.7002	.1474E-02	.1474E-02	.6996	.4631E-03	.4633E-03
8	0.7	log-N	.7001	.9328E-04	.9329E-04	.6991	.6133E-04	.6214E-04
3	0.9	log-N	.9002	.1696E-02	.1696E-02	.9003	.3663E-03	.3664E-03
8	0.9	log-N	.9001	.8116E-04	.8117E-04	.8998	.3393E-04	.3397E-04
3	0.5	χ^2_1	.5002	.1314E-02	.1314E-02	.4991	.5172E-03	.5180E-03
8	0.5	χ^2_1	.5001	.9005E-04	.9006E-04	.4990	.7248E-04	.7348E-04
3	0.7	χ^2_1	.7004	.1358E-02	.1358E-02	.6998	.4405E-03	.4405E-03
8	0.7	χ^2_1	.7000	.9569E-04	.9569E-04	.6991	.5986E-04	.6067E-04
3	0.9	χ^2_1	.9006	.1390E-02	.1390E-02	.9000	.3506E-03	.3506E-03
8	0.9	χ^2_1	.9001	.8283E-04	.8284E-04	.8998	.3323E-04	.3327E-04

Table 3: Results of Monte Carlo experiments, $N = 3000$, 10000 replications

$T+1$	β	Distribution of α_i, e_{i0}	System GMM			C-M		
			mean	var.	RMSE	mean	var.	RMSE
3	0.5	normal	.4976	.2543E-02	.2549E-02	.4990	.2521E-02	.2522E-02
8	0.5	normal	.5000	.9712E-04	.9712E-04	.5002	.9723E-04	.9727E-04
3	0.7	normal	.6980	.2151E-02	.2155E-02	.6989	.2123E-02	.2124E-02
8	0.7	normal	.7001	.1234E-03	.1234E-03	.7003	.1236E-03	.1237E-03
3	0.9	normal	.8985	.1928E-02	.1930E-02	.8991	.1843E-02	.1844E-02
8	0.9	normal	.8998	.1251E-03	.1251E-03	.8998	.1244E-03	.1244E-03
3	0.5	log-N	.4984	.2570E-02	.2573E-02	.4976	.1091E-02	.1097E-02
8	0.5	log-N	.5002	.9859E-04	.9863E-04	.4993	.1052E-03	.1057E-03
3	0.7	log-N	.6978	.2444E-02	.2449E-02	.6987	.8700E-03	.8717E-03
8	0.7	log-N	.7002	.1211E-03	.1211E-03	.6992	.1080E-03	.1086E-03
3	0.9	log-N	.8980	.2725E-02	.2729E-02	.8996	.6554E-03	.6556E-03
8	0.9	log-N	.8998	.1194E-03	.1194E-03	.8996	.5807E-04	.5823E-04
3	0.5	χ^2_3	.4981	.2486E-02	.2490E-02	.4981	.1141E-02	.1145E-02
8	0.5	χ^2_3	.5001	.9814E-04	.9815E-04	.4994	.9018E-04	.9054E-04
3	0.7	χ^2_3	.6982	.2132E-02	.2135E-02	.6991	.9430E-03	.9438E-03
8	0.7	χ^2_3	.7001	.1227E-03	.1227E-03	.6992	.9334E-04	.9398E-04
3	0.9	χ^2_3	.8986	.1912E-02	.1914E-02	.8997	.7407E-03	.7408E-03
8	0.9	χ^2_3	.8998	.1211E-03	.1211E-03	.8996	.6192E-04	.6208E-04

Table 4: Results of Monte Carlo experiments, $N = 3000$, unconditional variance of α_i equal to 4 and unconditional variance of e_{it} equal to 1, 10000 replications

largely confirm what has been observed with $N = 100$.

We also experimented a data generating process that comprise an exogenous independent variable. Results are broadly coherent with the previous simulations and are available from the authors upon request (reported in Appendix B for referees only).

A Why the efficiency gains?

In this Appendix we derive the variance of the estimator we propose, based on moment conditions (2), (4), and (6) and we compare it with the variance of the system GMM estimator.

We denote the set of R moment conditions in the population as $E(\mathbf{f}(\beta)) = 0$. In order to simplify computations we take into account the case of model (1), so that β (and the variance of its estimator) is a scalar.

From general GMM theory (Hansen, 1982), the asymptotic variance of a GMM estimator is given by:

$$V(\hat{\beta}) = (B'\Lambda^{-1}B)^{-1}$$

with

$$B = \frac{\partial \mathbf{f}(\beta)}{\partial \beta}$$

and

$$\Lambda = E(\mathbf{f}(\beta)\mathbf{f}(\beta)')$$

the efficient weighting matrix of the GMM procedure.

In this case we can partition B and Λ into two components: (i) the elements corresponding to the moment conditions that characterize the system GMM estimator, and (ii) the one moment condition we propose in this paper. Assuming that the moment condition we propose is the last moment condition in the full set composed of (2), (4), and (6), we can write:

$$B = \begin{pmatrix} B_1 \\ B_2 \end{pmatrix}$$

with B_1 related to the moment conditions in (2) and (4), and B_1 related to the moment condition in (6). Correspondingly,

$$\Lambda = \begin{pmatrix} \Lambda_1 & \Lambda_2 \\ \Lambda_3 & \Lambda_4 \end{pmatrix}$$

with $\Lambda_3 = \Lambda_2'$. Therefore we have

$$\begin{aligned} V(\hat{\beta}) &= (B'\Lambda^{-1}B)^{-1} \\ &= \left\{ \begin{pmatrix} B_1' & B_2' \end{pmatrix} \begin{pmatrix} \Lambda_1 & \Lambda_2 \\ \Lambda_3 & \Lambda_4 \end{pmatrix}^{-1} \begin{pmatrix} B_1 \\ B_2 \end{pmatrix} \right\}^{-1} \end{aligned}$$

By using the formula of the inverse of a partitioned matrix, we can write:

$$V(\hat{\beta}) = \left\{ \begin{pmatrix} B_1' & B_2' \end{pmatrix} \begin{pmatrix} \Lambda_1^{-1}(I + \Lambda_2 A \Lambda_3 \Lambda_1^{-1}) & -\Lambda_1^{-1}\Lambda_2 A \\ -A\Lambda_3\Lambda_1^{-1} & A \end{pmatrix} \begin{pmatrix} B_1 \\ B_2 \end{pmatrix} \right\}^{-1}$$

with $A = (\Lambda_4 - \Lambda_3\Lambda_1^{-1}\Lambda_2)^{-1} = (\Lambda_4 - \Lambda_2'\Lambda_1^{-1}\Lambda_2)^{-1}$ (a scalar).

By performing all the computations, we get:

$$V(\hat{\beta}) = (B_1'\Lambda_1^{-1}B_1 + (\Lambda_3\Lambda_1^{-1}B_1 - B_2)'A(\Lambda_3\Lambda_1^{-1}B_1 - B_2))^{-1} \quad (9)$$

Note that the first addendum of (9), correspond to the inverse of the variance of the system GMM estimator of β in (1). As a result, the variance of the GMM estimator based on (2), (4), and the moment condition we propose (6) will be smaller than the variance of the system GMM estimator if

$$(\Lambda_3\Lambda_1^{-1}B_1 - B_2)'A(\Lambda_3\Lambda_1^{-1}B_1 - B_2) > 0$$

As the scalar $A = (\det(\Lambda)/\det(\Lambda_1))^{-1} > 0$,[10] this condition will be always satisfied unless

$$\Lambda_3\Lambda_1^{-1}B_1 - B_2 = 0 \quad (10)$$

By relying on simulations and studying more in detail the case in which $T+1 = 3$, it is possible

[10]The equality can be obtained by relying on the fact that $\Lambda_1^{-1} = \Lambda_1^*/\det(\Lambda_1)$ where Λ_1^* is adjugate of Λ_1. Furthermore, for the way in which they are computed, both Λ and Λ_1 are positive definite matrices, so their determinant will be larger than zero.

to show that condition (10) is satisfied only if both α_i and e_{i0} have a symmetric distribution (the third moment of these variables is equal to zero). Coherently, the Monte Carlo experiments reported in Section 3 show no gains in performance in the cases in which the error components are assumed to be normally distributed.

Figure 2 shows the ratio of the asymptotic variance of the estimator we propose with respect to the asymptotic variance of the system GMM estimator. The baseline experiment consider $\beta = 0.5$, $T + 1 = 3$, a χ^2_3 distribution for e_{i0} and α_i, and equal unconditional variances of α_i and e_{it} (both equal to 1). From the baseline experiment, we changed: (i) the level of asymmetry in the distribution of α_i, e_{i0}, accomplished by changing the number of degrees of freedom of their χ^2_3 distributions (not reported in the Figure is the case of perfect symmetry, in which the ratio of the two variances is equal to 1); (ii) the number of time periods $T + 1$ from 3 to 10; (iii) the ratio of unconditional variance of α with respect to e (from 1 to 9; we changed the value of the unconditional variance of α_i, setting the unconditional variance of e equal to 1); (iv) the values of β from 0.1 to 0.9. Coherently with the results in the Monte Carlo experiments, the gains in performance led by the additional moment condition we propose are stronger for larger asymmetry in the distribution of α_i and e_{i0} (thus, of y_{i0}), shorter observation in time and higher value of β. Also, the gains in performance decreases when the unconditional variance of α_i increases with respect to the unconditional variance of e_{it}.

B For referees only: results with an exogenous x

We generate a dynamic panel data model as:

$$y_{it} = \beta_1 y_{it-1} + \beta_2 x_{it} + \varepsilon_{it} = \beta_1 y_{it-1} + \beta_2 x_{it} + \alpha_i + e_{it} \tag{11}$$

with $|\beta_1| < 1$, and e_{it} normally distributed with mean zero and unconditional variance equal to 1, independent of α_i. As for the distribution of α_i we now focus on the log-normal distribution. Still, appropriate transformation of α_i are considered in order to have mean zero and the chosen value for the unconditional variance. A log-normal distribution is also considered for the generation of the exogenous variable x_{it}, as the log-normal distribution is suited for the description of many

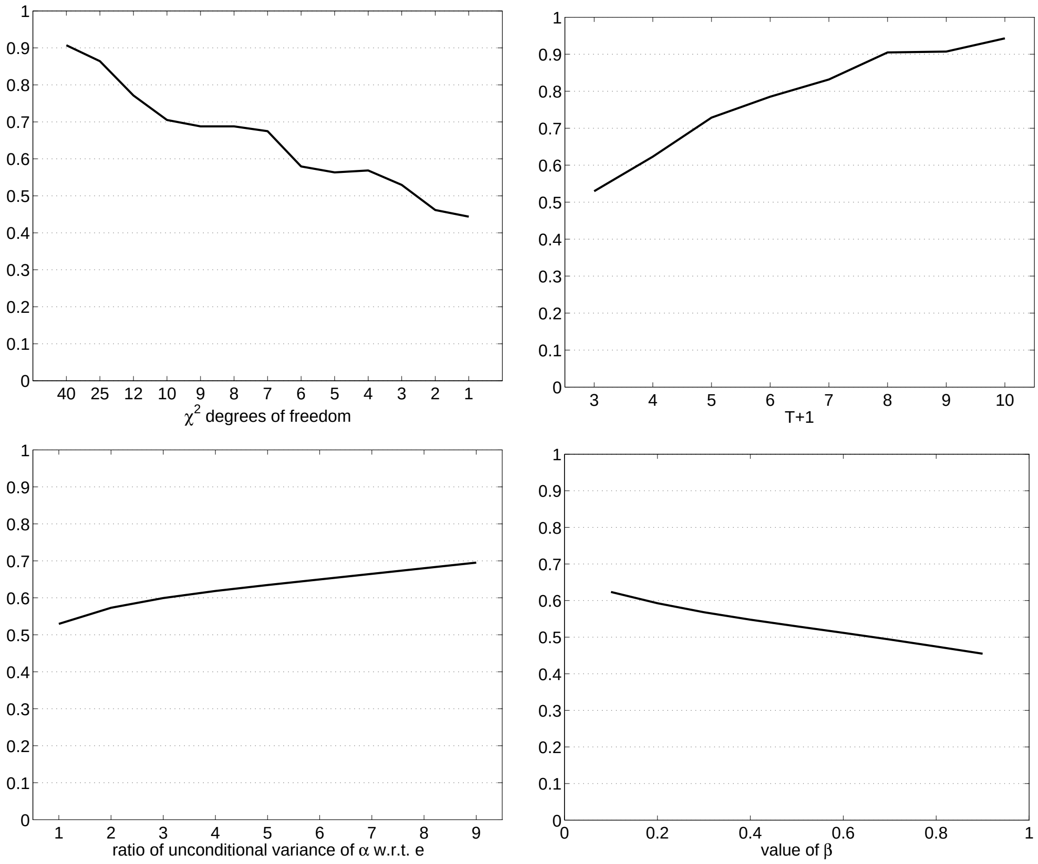

Figure 2: Ratio of the asymptotic variance of the estimator we propose with respect to the asymptotic variance of the system GMM estimator, for different (i) levels of asymmetry in the distribution of α_i, e_{i0}; (ii) time periods; (iii) ratios of unconditional variance of α with respect to e; (iv) values of β

variables in economic applications.

The initial observation, y_{i0} is generated as:

$$y_{i0} = \frac{\alpha_i}{1-\beta} + e_{i0} \tag{12}$$

and different distributions are also considered for e_{i0}.

Besides conditions (2), (4), and (6), the additional moment conditions that exploit the exogeneity of x_{it} are employed in the estimation.

Results of Monte Carlo experiments are reported in Table 5-8. As for the estimates of β_1, gains in performance arise especially in the case in which $T + 1 = 3$. On the contrary, with few exceptions, the estimation of β_2 is not affected (sometimes, larger variances are observed with C-M as compared to the system GMM estimator).

			Estimates of β_1				Estimates of β_2			
			System GMM		C-M		System GMM		C-M	
$T+1$	$\beta_1; \beta_2$	var(X)	mean	var.	mean	var.	mean	var.	mean	var.
3	0.5; 0.1	1.0	.5139	.3780E-01	.5029	.2293E-01	.9544E-01	.1838E-01	.9291E-01	.2107E-01
8	0.5; 0.1	1.0	.5055	.6054E-02	.4968	.5788E-02	.8892E-01	.1323E-01	.8503E-01	.3021E-01
3	0.9; 0.1	1.0	.9052	.5561E-01	.9119	.2380E-01	.9670E-01	.4425E-02	.9495E-01	.4413E-02
8	0.9; 0.1	1.0	.8994	.5759E-02	.9004	.4052E-02	.9482E-01	.3013E-01	.9250E-01	.6776E-01
3	0.5; 0.5	1.0	.5041	.2595E-01	.5011	.1807E-01	.4895	.1460E-01	.4850	.1511E-01
8	0.5; 0.5	1.0	.4953	.4610E-02	.4904	.4436E-02	.4737	.7337E-02	.4702	.7930E-02
3	0.9; 0.5	1.0	.9019	.1511E-01	.9086	.1071E-01	.4968	.3050E-02	.4943	.2994E-02
8	0.9; 0.5	1.0	.9006	.7382E-03	.9002	.7403E-03	.4924	.1448E-02	.4920	.1462E-02
3	0.5; 1.0	1.0	.5010	.1266E-01	.5007	.1093E-01	.9901	.1209E-01	.9843	.1262E-01
8	0.5; 1.0	1.0	.4941	.2299E-02	.4912	.2270E-02	.9709	.6627E-02	.9691	.7078E-02
3	0.9; 1.0	1.0	.9004	.4878E-02	.9032	.4134E-02	.9985	.3039E-02	.9960	.2805E-02
8	0.9; 1.0	1.0	.8999	.1862E-03	.8997	.1873E-03	.9932	.1295E-02	.9930	.1297E-02
3	0.5; 1.0	0.1	.5202	.5116E-01	.5313	.4110E-01	.9619	.9537E-01	.9423	.8379E-01
8	0.5; 1.0	0.1	.5030	.5491E-02	.5008	.5421E-02	.9522	.2211E-01	.9480	.2257E-01
3	0.9; 1.0	0.1	.9062	.3793E-01	.9216	.1867E-01	.9906	.3838E-01	.9767	.1897E-01
8	0.9; 1.0	0.1	.9007	.3701E-03	.9008	.3707E-03	.9915	.3498E-02	.9901	.3480E-02

Table 5: Results of Monte Carlo experiments, $N = 100$, 10000 replications

			Estimates of β_1				Estimates of β_2			
			System GMM		C-M		System GMM		C-M	
$T+1$	$\beta_1; \beta_2$	var(X)	mean	var.	mean	var.	mean	var.	mean	var.
3	0.5; 0.1	1.0	.5305	.5398E-01	.5108	.3868E-01	.9251E-01	.1934E-01	.8947E-01	.2224E-01
8	0.5; 0.1	1.0	.5042	.7134E-02	.4939	.7186E-02	.8747E-01	.1914E-01	.8693E-01	.5403
3	0.9; 0.1	1.0	.9006	.8508E-01	.9106	.4177E-01	.9545E-01	.4281E-02	.9227E-01	.4423E-02
8	0.9; 0.1	1.0	.8959	.7681E-02	.8929	.7157E-02	.8884E-01	.8310E-01	.8565E-01	.1046
3	0.5; 0.5	1.0	.5008	.3971E-01	.4914	.3198E-01	.4758	.1689E-01	.4674	.1832E-01
8	0.5; 0.5	1.0	.4857	.6085E-02	.4770	.6009E-02	.4576	.1399E-01	.4582	.2261
3	0.9; 0.5	1.0	.8965	.2329E-01	.9097	.1770E-01	.4919	.3005E-02	.4875	.2972E-02
8	0.9; 0.5	1.0	.8960	.1042E-02	.8951	.1097E-02	.4865	.1750E-02	.4856	.1826E-02
3	0.5; 1.0	1.0	.4993	.2346E-01	.5006	.2257E-01	.9688	.1689E-01	.9568	.1733E-01
8	0.5; 1.0	1.0	.4799	.3849E-02	.4737	.3663E-02	.9503	.1229E-01	.9464	.1214E-01
3	0.9; 1.0	1.0	.9002	.8573E-02	.9064	.7421E-02	.9940	.3991E-02	.9888	.3718E-02
8	0.9; 1.0	1.0	.8978	.2312E-03	.8973	.2341E-03	.9866	.1514E-02	.9857	.1488E-02
3	0.5; 1.0	0.1	.5355	.6575E-01	.5360	.5472E-01	.9158	.1126	.9029	.1010
8	0.5; 1.0	0.1	.4973	.5678E-02	.4930	.5514E-02	.8928	.2801E-01	.8840	.3215E-01
3	0.9; 1.0	0.1	.9110	.4366E-01	.9251	.2109E-01	.9799	.3918E-01	.9673	.1827E-01
8	0.9; 1.0	0.1	.9003	.3875E-03	.9002	.3934E-03	.9792	.3633E-02	.9771	.3648E-02

Table 6: Results of Monte Carlo experiments, $N = 100$, unconditional variance of α_i equal to 4 and unconditional variance of e_{it} equal to 1, 10000 replications

			Estimates of β_1				Estimates of β_2			
			System GMM		C-M		System GMM		C-M	
$T+1$	$\beta_1; \beta_2$	var(X)	mean	var.	mean	var.	mean	var.	mean	var.
3	0.5; 0.1	1.0	.4984	.1267E-02	.4982	.5154E-03	.9963E-01	.1878E-03	.9963E-01	.1866E-03
8	0.5; 0.1	1.0	.5002	.8799E-04	.4986	.7135E-04	.9982E-01	.5453E-04	.9956E-01	.5392E-04
3	0.9; 0.1	1.0	.8994	.1052E-02	.9001	.3056E-03	.9984E-01	.4436E-04	.9979E-01	.4424E-04
8	0.9; 0.1	1.0	.8999	.5438E-04	.8996	.2815E-04	.9987E-01	.1274E-04	.9987E-01	.1263E-04
3	0.5; 0.5	1.0	.4993	.4016E-03	.4985	.2648E-03	.4995	.1443E-03	.4994	.1435E-03
8	0.5; 0.5	1.0	.4991	.5409E-04	.4981	.4557E-04	.4990	.4529E-04	.4987	.4403E-04
3	0.9; 0.5	1.0	.9000	.1095E-03	.9001	.8236E-04	.4999	.2834E-04	.4998	.2826E-04
8	0.9; 0.5	1.0	.8999	.5739E-05	.8998	.5261E-05	.4997	.1015E-04	.4997	.1008E-04
3	0.5; 1.0	1.0	.4997	.1238E-03	.4991	.1075E-03	.9995	.1186E-03	.9993	.1188E-03
8	0.5; 1.0	1.0	.4990	.2242E-04	.4985	.2059E-04	.9989	.4139E-04	.9987	.4104E-04
3	0.9; 1.0	1.0	.9000	.2890E-04	.9000	.2667E-04	.9999	.2575E-04	.9998	.2575E-04
8	0.9; 1.0	1.0	.8999	.1491E-05	.8999	.1458E-05	.9997	.9633E-05	.9997	.9611E-05
3	0.5; 1.0	0.1	.4999	.5408E-03	.4989	.3906E-03	.9990	.8869E-03	.9993	.8438E-03
8	0.5; 1.0	0.1	.4996	.5850E-04	.4990	.5632E-04	.9972	.2485E-03	.9969	.2478E-03
3	0.9; 1.0	0.1	.9002	.1828E-03	.9004	.1253E-03	.9997	.2031E-03	.9994	.1638E-03
8	0.9; 1.0	0.1	.9000	.4497E-05	.9000	.4323E-05	.9993	.4785E-04	.9993	.4689E-04

Table 7: Results of Monte Carlo experiments, $N = 3000$, 10000 replications

			Estimates of β_1				Estimates of β_2			
			System GMM		C-M		System GMM		C-M	
$T+1$	$\beta_1; \beta_2$	var(X)	mean	var.	mean	var.	mean	var.	mean	var.
3	0.5; 0.1	1.0	.4956	.2342E-02	.4955	.1079E-02	.9942E-01	.1919E-03	.9959E-01	.1903E-03
8	0.5; 0.1	1.0	.5002	.9734E-04	.4986	.1086E-03	.9981E-01	.5650E-04	.1033	.1102
3	0.9; 0.1	1.0	.8979	.1365E-02	.8988	.5915E-03	.9973E-01	.4507E-04	.9976E-01	.4440E-04
8	0.9; 0.1	1.0	.8988	.7997E-04	.8985	.4498E-04	.9978E-01	.1290E-04	.9980E-01	.1280E-04
3	0.5; 0.5	1.0	.4975	.6326E-03	.4959	.4516E-03	.4989	.1637E-03	.4986	.1578E-03
8	0.5; 0.5	1.0	.4982	.7516E-04	.4968	.6622E-04	.4985	.5140E-04	.4981	.5042E-04
3	0.9; 0.5	1.0	.8997	.1233E-03	.8995	.1058E-03	.4998	.2916E-04	.4996	.2907E-04
8	0.9; 0.5	1.0	.8994	.7530E-05	.8992	.7056E-05	.4995	.1040E-04	.4995	.1039E-04
3	0.5; 1.0	1.0	.4987	.1861E-03	.4976	.1843E-03	.9988	.1416E-03	.9983	.1484E-03
8	0.5; 1.0	1.0	.4974	.3890E-04	.4967	.3579E-04	.9979	.4801E-04	.9976	.4744E-04
3	0.9; 1.0	1.0	.8999	.3143E-04	.8997	.3032E-04	.9998	.2655E-04	.9996	.2655E-04
8	0.9; 1.0	1.0	.8996	.1920E-05	.8996	.1891E-05	.9995	.9927E-05	.9995	.9923E-05
3	0.5; 1.0	0.1	.4989	.6706E-03	.4967	.5148E-03	.9980	.9693E-03	.9984	.9518E-03
8	0.5; 1.0	0.1	.4984	.6915E-04	.4976	.6572E-04	.9938	.3162E-03	.9934	.3167E-03
3	0.9; 1.0	0.1	.9000	.1889E-03	.8999	.1544E-03	.9995	.2057E-03	.9993	.1815E-03
8	0.9; 1.0	0.1	.8999	.4671E-05	.8998	.4580E-05	.9986	.5115E-04	.9984	.5080E-04

Table 8: Results of Monte Carlo experiments, $N = 3000$, unconditional variance of α_i equal to 4 and unconditional variance of e_{it} equal to 1, 10000 replications

References

Ahn, S.C. and Schmidt, P.: 1995, Efficient Estimation of Models for Dynamic Panel Data, *Journal of Econometrics* **68**(1), 5–27.

Anderson, T. and Hsiao, C.: 1981, “Estimation of Dynamic Models with Error Components”, *Journal of the American Statistical Association* **76**, 598–606.

Anderson, T. and Hsiao, C.: 1982, “Formulation and Estimation of Dynamic Models Using Panel Data”, *Journal of Econometrics* **18**, 47–82.

Arellano, M. and Bond, S.: 1991, “Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations”, *The Review of Economic Studies* **58**(2), 277–297.

Arellano, M. and Bover, O.: 1995, “Another Look at the Instrumental Variables Estimation of Error Components Models’, *Journal of Econometrics* **68**, 29-51.

Bhargava, A. and Sargan, J.D.: 1983, “Estimating Dynamic Random Effects Models from Panel Data Covering Short Time Periods”, *Econometrica* **51**(6), 1635–1659.

Blundell, S. and Bond, S.: “Initial Conditions and Moment Restrictions in Dynamic Panel Data Models”, *Journal of Econometrics* **87**(1), 115–143.

Calzolari, G. and Magazzini, L. (2013): “A powerful test of mean stationarity in dynamic models for panel data: Monte Carlo evidence”, Working Paper Series, Department of Economics, University of Verona n. 14/2013.

Gouriéroux, C., Phillips, P.C.B. and Yu, J.: 2010, “Inderect Inference for Dynamic Panel Models”, *Journal of Econometrics* **157**(1), 68–77.

Hayakawa, K.: 2007, “Small Sample Bias Properties of the System GMM Estimator in Dynamic Panel Data Models”, *Economics Letters* **95**, 32–38.

Han, C. and Phillips, P.C.B.: 2010, “GMM Estimation for Dynamic Panel with Fixed Effects and Strong Instruments at Unity”, *Econometric Theory* **26**(1), 119–151.

Hahn, J.: 1999, "How Informative is the Initial Condition in the Dynamic Panel Model with Fixed Effects?", *Journal of Econometrics* **93**, 309–326.

Hansen, L.: 1982, "Large Sample Properties of Generalized Method of Moments Estimators", *Econometrica* **50**, 1029-1054.

Hansen, L.P., Heaton, J., and Yaron, A.: 1996, Finite-Sample Properties of Some Alternative GMM Estimators, *Journal of Business & Economic Statistics* **14**(3), 262–280.

Hsiao, C., Pesaran, M.H. and Tahmiscioglu A.K.: 2002, "Maximum Likelihood Estimation of Fixed Effects Dynamic Panel Data Models Covering Short Time Periods", *Journal of Econometrics* **109**(1), 107–150.

Kiviet, J.F.: 1995, "On Bias, Inconsistency, and Efficiency of Various Estimators in Dynamic Panel Data Models", *Journal of Econometrics* **68**(1), 53–78.

Roodman, D.: 2009, "A Note on the Theme of Too Many Instruments", *Oxford Bulletin of Economics and Statistics* **71**(1), 135–158.

Windmeijer, F.: 2000, "A Finite Sample Correction for the Variance of Linear Efficient Two-Step GMM Estimators", *Journal of Econometrics* **126**(1), 25–51.