ESTIMATION IN DYNAMIC PANEL DATA MODELS: IMPROVING ON THE PERFORMANCE OF THE STANDARD GMM ESTIMATORS

Richard Blundell
Stephen Bond
Frank Windmeijer

THE INSTITUTE FOR FISCAL STUDIES
WP 00/12

Estimation in Dynamic Panel Data Models: Improving on the Performance of the Standard GMM Estimator

Richard Blundell
Institute for Fiscal Studies and University College London

Stephen Bond
Institute for Fiscal Studies and Nuffield College, Oxford

Frank Windmeijer
Institute for Fiscal Studies

Summary

This chapter reviews developments to improve on the poor performance of the standard GMM estimator for highly autoregressive panel series. It considers the use of the "system" GMM estimator that relies on relatively mild restrictions on the initial condition process. This system GMM estimator encompasses the GMM estimator based on the non-linear moment conditions available in the dynamic error components model and has substantial asymptotic efficiency gains. Simulations, that include weakly exogenous covariates, find large finite sample biases and very low precision for the standard first differenced estimator. The use of the system GMM estimator not only greatly improves the precision but also greatly reduces the finite sample bias. An application to panel production function data for the US is provided and confirms these theoretical and experimental findings.

1. Introduction

Much of the recent literature on dynamic panel data estimation has focussed on providing optimal linear Generalised Methods of Moments (GMM) estimators under relatively weak auxiliary assumptions about the exogeneity of the covariate processes and the properties of the heterogeneity and error term processes. A standard approach is to first-difference the equation to remove permanent unobserved heterogeneity, and to use lagged levels of the series as instruments for the predetermined and endogenous variables in first-differences (see Anderson and Hsiao (1981), Holtz-Eakin, Newey and Rosen (1988) and Arellano and Bond (1991)). However, in dynamic panel data models where the series are highly autoregressive and the number of time series observations is moderately small, this standard GMM estimator has been found to have large finite sample bias and poor precision in simulation studies (see the experimental evidence and theoretical discussions in Ahn and Schmidt (1995) and Alonso-Borrego and Arellano (1999), for example).

The poor performance of the standard GMM panel data estimator is also reflected in empirical experience with estimation on relatively short panels with highly persistent data. To quote from the extensive review of production function estimation by Griliches and Mairesse (1998) - one of the original applications for panel data estimation - "In empirical practice, the application of panel methods to micro-data produced rather unsatisfactory results: low and often insignificant capital coefficients and unreasonably low estimates of returns to scale." One simple explanation of these findings in the production function context is that lagged levels of the series provide weak instruments for first-differenced variables in this case (see Blundell and Bond (1999)).

One response to these findings has been to consider the use of further moment conditions that have improved properties for the estimates of the parameters of interest. For example, Ahn and Schmidt (1995) consider the non-linear moment conditions implied by the standard error components formulation and show that asymptotic variance ratios can be considerably improved. Blundell and Bond (1998) consider alternative estimators that require further restrictions on the initial conditions process, designed to improve the properties of the standard first-differenced instrumental variables estimator.

This also provides the motivation for the discussion in this chapter. The idea is to consider the performance of a "system" GMM estimator that relies on relatively mild restrictions on the initial condition process to improve the performance of the GMM estimator in the dynamic panel data context. The material presented draws extensively from the existing literature. For example, Arellano and Bover (1995) and Blundell and Bond (1998) show that mean stationarity in an AR(1) panel data model is sufficient to justify the use of lagged *differences* of the dependent variable as instruments for equations in levels, in addition to lagged *levels* as instruments for equations in first-differences. This result naturally extends to models with weakly exogenous covariates. The Monte Carlo simulations and asymptotic variance calculations reported in this paper show that this extended GMM estimator can offer considerable efficiency gains in the situations where the standard first-differenced GMM estimator performs poorly. Given this restriction on the initial conditions, the system GMM estimator is also shown to encompass the GMM estimator based on the non-linear moment conditions available in the dynamic error components model (see Ahn and Schmidt (1995)). The system GMM estimator has substantial asymptotic efficiency gains relative to this non-linear GMM estimator, and these are reflected in their finite sample properties.

The chapter is organised in the following way. The next section reviews the standard error components structure for a linear dynamic panel data model and lays out the underlying assumptions. Recalling that Within Groups, GLS and

OLS on the levels and first-differenced models all suffer from bias even when the cross-section dimension is large, this section also briefly considers the biases that occur for standard panel data estimators in dynamic models. Section 3 then presents the linear GMM estimator for this model that uses lagged information to instrument current differences in a first-differenced specification. The following section then outlines the problem of weak instruments in this case. Following the discussion in Ahn and Schmidt (1995), Section 5 considers the use of further *non-linear* moment conditions that are implied by the model outlined in section 2. Section 6 derives a linear moment restriction for the levels model using initial condition restrictions and this is then incorporated into the full system GMM estimator. Asymptotic variance comparisons among these various GMM estimators are given in section 8. The detailed discussion in these earlier sections uses an AR(1) model and the extension to a multivariate setting is presented in section 9. Finally, before moving to the Monte Carlo results and empirical application, overidentification tests are reviewed.

The Monte Carlo results presented in section 11 are the first in the literature to consider the properties of these GMM estimators in dynamic models with weakly exogenous regressors. As this is perhaps the most common case in empirical applications, these results have important bearing on applied work. The analysis finds both a large bias and very low precision for the standard first-differenced estimator when the individual series are highly autoregressive. The use of the system GMM estimator not only greatly improves the precision but also greatly reduces the finite sample bias. Exploiting the non-linear moment conditions also provides significant gains compared to the standard first-differenced GMM estimator, but these gains are much less dramatic than those provided by the system GMM estimator when the initial conditions restriction is valid.

The empirical application returns to the Griliches and Mairesse discussion.

The application uses production function data for the US and confirms the Griliches and Mairesse findings for the capital and labor coefficients in a Cobb-Douglas model. Using the standard first-differenced GMM estimator, the estimated coefficient on capital is very low and all coefficient estimates have poor precision. Constant returns to scale is easily rejected. Moreover, an examination of the individual series suggests that they are highly autoregressive thus hinting at a weak instruments problem for standard GMM on this data. These production function results are improved by using the system estimator. The capital coefficient is now more precise and takes a reasonable value and constant returns to scale is not rejected. These Monte Carlo and empirical results indicate that a careful examination of the original series and use of the system GMM estimator can overcome many of the disappointing features of the standard GMM estimator in the context of highly persistent series.

2. Dynamic Models and the Biases from Standard Panel Data Estimators

To analyse the properties of estimators of the parameters in linear dynamic panel data models we consider an autoregressive panel data model of the form

$$y_{it} = \alpha y_{it-1} + \beta' x_{it} + u_{it} \tag{2.1}$$

$$u_{it} = \eta_i + v_{it} \tag{2.2}$$

for $i = 1, ..., N$ and $t = 2, ..., T$, where $\eta_i + v_{it}$ is the usual 'error components' decomposition of the error term; N is large, T is fixed and $|\alpha| < 1$.[1] This model specification is sufficient to cover most of the standard cases encountered in linear dynamic panel applications. Allowing the inclusion of x_{it-1} provides the autoregressive panel data model

$$y_{it} = \alpha y_{i,t-1} + \beta_1' x_{it} + \beta_2' x_{it-1} + \eta_i + v_{it}$$

which has the corresponding 'common factor' restricted ($\beta_2 = -\alpha\beta_1$) form

$$y_{it} = \beta_1' x_{it} + f_i + \zeta_{it},$$

with $\zeta_{it} = \alpha\zeta_{i,t-1} + v_{it}$ and $\eta_i = (1-\alpha)f_i$.

In our Monte Carlo study and application to panel data production function equations presented in Sections 11 and 12 we allow for the inclusion of x_{it} regressors, but for the evaluation of the various estimators we use an AR(1) model with unobserved individual-specific effects

$$\begin{aligned} y_{it} &= \alpha y_{i,t-1} + u_{it} \\ u_{it} &= \eta_i + v_{it} \end{aligned} \tag{2.3}$$

for $i = 1, ..., N$ and $t = 2, ..., T$.[2] At the outset we will assume that η_i and v_{it} have the familiar error components structure in which

$$E(\eta_i) = 0\ ,\ \ E(v_{it}) = 0,\ \ E(v_{it}\eta_i) = 0 \text{ for } i = 1, ..., N \text{ and } t = 2, ..., T \tag{2.4}$$

and

$$E(v_{it}v_{is}) = 0 \text{ for } i = 1, ..., N \text{ and } \forall\ t \neq s. \tag{2.5}$$

In addition there is the standard assumption concerning the initial conditions y_{i1}(see Ahn and Schmidt (1995), for example)

$$E(y_{i1}v_{it}) = 0 \text{ for } i = 1, ..., N \text{ and } t = 2, ..., T. \tag{2.6}$$

These 'standard assumptions' (2.4), (2.5) and (2.6) imply moment restrictions that are sufficient to (identify and) estimate α for $T \geq 3$.[3]

Further restrictions on the initial conditions define a *mean stationary* process as

$$y_{i1} = \frac{\eta_i}{1-\alpha} + \varepsilon_{i1} \quad \text{for } i = 1, ..., N \tag{2.7}$$

and

$$E(\varepsilon_{i1}) = E(\eta_i \varepsilon_{i1}) = 0 \quad \text{for } i = 1, ..., N, \tag{2.8}$$

and a *covariance stationary* process by further specifying

$$\begin{aligned} E\left(v_{it}^2\right) &= \sigma_v^2 \quad \text{for } i = 1, ..., N \text{ and } t = 2..., T \\ E\left(\varepsilon_{i1}^2\right) &= \frac{\sigma_v^2}{1-\alpha^2} \quad \text{for } i = 1, ..., N. \end{aligned}$$

For completeness and to conclude this brief outline of the dynamic error components model, we consider the biases from the standard panel data estimators in this model. We consider here the biases found under covariance stationarity (for more detail see Baltagi (1995) and Hsiao (1986)).

The asymptotic bias of the simple OLS estimator for α in model (2.3), is given by

$$\text{plim}\left(\widehat{\alpha}_{OLS} - \alpha\right) = (1-\alpha)\frac{\sigma_\eta^2/\sigma_v^2}{\sigma_\eta^2/\sigma_v^2 + k}, \text{ with } k = \frac{1-\alpha}{1+\alpha},$$

where $\sigma_\eta^2 = E\left(\eta_i^2\right)$, and therefore the OLS estimator is biased upwards, with $\alpha < \text{plim}\left(\widehat{\alpha}_{OLS}\right) < 1$.

The asymptotic bias of the Within Groups estimator for α has been documented by Nickell (1981) and is given by

$$\text{plim}\left(\widehat{\alpha}_{WG} - \alpha\right) = -\frac{\frac{1+\alpha}{T-1}\left(1 - \frac{1}{T}\frac{1-\alpha^T}{(1-\alpha)}\right)}{1 - \frac{2\alpha}{(1-\alpha)(T-1)}\left(1 - \frac{1}{T}\frac{1-\alpha^T}{(1-\alpha)}\right)},$$

and so, when $\alpha > 0$, $\text{plim}\left(\widehat{\alpha}_{WG}\right) < \alpha$.

When the model is transformed into first-differences to eliminate the unobserved individual heterogeneity component η_i,

$$\Delta y_{it} = \alpha \Delta y_{it-1} + \Delta u_{it},$$

the bias of the OLS estimator is given by

$$\text{plim}\left(\widehat{\alpha}_{OLSd} - \alpha\right) = -\frac{1+\alpha}{2},$$

and so $\text{plim}\left(\widehat{\alpha}_{OLSd}\right) = \frac{\alpha-1}{2} < 0$.

3. A First-Differenced GMM Estimator

3.1. The standard moment conditions

In the absence of any further restrictions on the process generating the initial conditions, the autoregressive error components model (2.3) - (2.6) implies the following $m_d = 0.5(T-1)(T-2)$ orthogonality conditions which are linear in the α parameter

$$E(y_{i,t-s}\Delta u_{it}) = 0; \text{ for } t = 3, ..., T \text{ and } 2 \leq s \leq t-1, \tag{3.1}$$

where $\Delta u_{it} = u_{it} - u_{i,t-1}$. These depend only on the assumed absence of serial correlation in the time varying disturbances v_{it}, together with the restriction (2.6).

The moment restrictions in (3.1) can be expressed more compactly as

$$E(\mathbf{Z}'_{di}\mathbf{\Delta u}_i) = \mathbf{0},$$

where $\mathbf{Z}_{di}$ is the $(T-2) \times m_d$ matrix given by

$$\mathbf{Z}_{di} = \begin{bmatrix} y_{i1} & 0 & 0 & ... & 0 & ... & 0 \\ 0 & y_{i1} & y_{i2} & ... & 0 & ... & 0 \\ . & . & . & ... & . & ... & . \\ 0 & 0 & 0 & ... & y_{i1} & ... & y_{iT-2} \end{bmatrix},$$

and $\mathbf{\Delta u}_i$ is the $(T-2)$ vector $(\Delta u_{i3,}\Delta u_{i4}, ..., \Delta u_{iT})'$.

The Generalised Method of Moments (GMM) estimator based on these moment conditions minimises the quadratic distance $\mathbf{\Delta u}'\mathbf{Z}_d\mathbf{W}_N\mathbf{Z}'_d\mathbf{\Delta u}$ for some metric $\mathbf{W}_N$, where $\mathbf{Z}'_d$ is the $m_d \times N(T-2)$ matrix $(\mathbf{Z}'_{d1}, \mathbf{Z}'_{d2}, ..., \mathbf{Z}'_{dN})$ and $\mathbf{\Delta u}'$ is the $N(T-2)$ vector $(\mathbf{\Delta u}'_1, \mathbf{\Delta u}'_2, ..., \mathbf{\Delta u}'_N)$. This gives the GMM estimator for α as

$$\widehat{\alpha}_d = (\mathbf{\Delta y}'_{-1}\mathbf{Z}_d\mathbf{W}_N\mathbf{Z}'_d\mathbf{\Delta y}_{-1})^{-1}\mathbf{\Delta y}'_{-1}\mathbf{Z}_d\mathbf{W}_N\mathbf{Z}'_d\mathbf{\Delta y},$$

where $\mathbf{\Delta y}'_i$ is the $(T-2)$ vector $(\Delta y_{i3,}\Delta y_{i4}, ..., \Delta y_{iT})$, $\mathbf{\Delta y}'_{i,-1}$ is the $(T-2)$ vector $(\Delta y_{i2}, \Delta y_{i3}, ..., \Delta y_{i,T-1})$, and $\mathbf{\Delta y}$ and $\mathbf{\Delta y}_{-1}$ are stacked across individuals in the same way as $\mathbf{\Delta u}$.

Alternative choices for the weights $\mathbf{W}_N$ give rise to a set of GMM estimators based on the moment conditions in (3.1), all of which are consistent for large N and finite T, but which differ in their asymptotic efficiency.[4] In general the optimal weights are given by

$$\mathbf{W}_N = (\frac{1}{N}\sum_{i=1}^{N} \mathbf{Z}'_{di}\widehat{\mathbf{\Delta u}}_i\widehat{\mathbf{\Delta u}}'_i\mathbf{Z}_{di})^{-1} \tag{3.2}$$

where $\widehat{\mathbf{\Delta u}}_i$ are residuals from an initial consistent estimator. We refer to this as the *two-step* GMM estimator.[5] In the absence of any additional knowledge about the process for the initial conditions, this estimator is asymptotically efficient in the class of estimators based on the linear moment conditions (3.1) (see Hansen (1982) and Chamberlain (1987)).

3.2. Homoskedasticity

Ahn and Schmidt (1995) show that additional linear moment conditions are available if the v_{it} disturbances are homoskedastic through time, i.e. if

$$E(v_{it}^2) = \sigma_i^2 \text{ for } t = 2, ..., T. \tag{3.3}$$

This implies $T-3$ orthogonality restrictions of the form

$$E(y_{i,t-2}\Delta u_{i,t-1} - y_{i,t-1}\Delta u_{it}) = 0; \text{ for } t = 4, ..., T \tag{3.4}$$

and allows a further $T-3$ columns to be added to the instrument matrix $\mathbf{Z}_{di}$. The additional columns $\mathbf{Z}_{hi}$ are

$$\mathbf{Z}_{hi} = \begin{pmatrix} y_{i2} & -y_{i3} & 0 & ... & 0 & 0 \\ 0 & y_{i3} & -y_{i4} & ... & 0 & 0 \\ . & . & . & ... & . & . \\ 0 & 0 & 0 & ... & y_{iT-2} & -y_{iT-1} \end{pmatrix}'.$$

Calculation of the one-step and two-step GMM estimators then proceeds exactly as described above.

4. Weak Instruments

The instruments used in the standard first-differenced GMM estimator become less informative in two important cases. First, as the value of the autoregressive parameter α increases towards unity; and second, as the variance of the individual effects η_i increases relative to the variance of v_{it}. To examine this further consider the case with $T = 3$. In this case, the moment conditions corresponding to the standard GMM estimator reduce to a single orthogonality condition. The corresponding method of moments estimator reduces to a simple two stage least squares (2SLS) estimator, with first stage (instrumental variable) regression

$$\Delta y_{i2} = \pi_d y_{i1} + r_i \text{ for } i = 1, ..., N.$$

For sufficiently high autoregressive parameter α or for sufficiently high relative variance of the individual effects, the least squares estimate of the reduced form coefficient π_d can be made arbitrarily close to zero. In this case the instrument y_{i1} is only weakly correlated with Δy_{i2}. To see this notice that the model (2.3) implies that

$$\Delta y_{i2} = (\alpha - 1)y_{i1} + \eta_i + v_{i2} \text{ for } i = 1, ..., N. \tag{4.1}$$

The least squares estimator of $(\alpha - 1)$ in (4.1) is generally biased upwards, towards zero, since we expect $E(y_{i1}\eta_i) > 0$. Assuming covariance stationarity and letting $\sigma_\eta^2 = \text{var}(\eta_i)$ and $\sigma_v^2 = \text{var}(v_{it})$, the plim of $\widehat{\pi}_d$ is given by

$$\text{plim}\,\widehat{\pi}_d = (\alpha - 1)\frac{k}{\frac{\sigma_\eta^2}{\sigma_v^2} + k}; \text{ with } k = \frac{1-\alpha}{1+\alpha}. \tag{4.2}$$

The bias term effectively scales the estimated coefficient on the instrumental variable y_{i1} toward zero. We find that $\text{plim}\,\widehat{\pi}_d \to 0$ as $\alpha \to 1$ or as $(\sigma_\eta^2/\sigma_v^2) \to \infty$, which are the cases in which the first stage F-statistic is $O_p(1)$. A graph showing both $\text{plim}\,\widehat{\pi}_d$ and $\alpha - 1$ against α is given in Figure 1, for $\sigma_\eta^2 = \sigma_v^2$, $T = 3$.

FIGURE 1 ABOUT HERE

We are interested in inferences using this first-differenced instrumental variable (IV) estimator when π_d is local to zero, that is where the instrument y_{i1} is only weakly correlated with Δy_{i2}. Following Nelson and Startz (1990a,b) and Staiger and Stock (1997) we characterise this problem of weak instruments using the concentration parameter. First note that the F-statistic for the first stage instrumental variable regression converges to a noncentral chi-squared with one degree of freedom. The concentration parameter is then the corresponding noncentrality parameter which we label τ in this case. The IV estimator performs poorly when τ approaches zero. Assuming covariance stationarity, τ has the following simple characterisation in terms of the parameters of the AR model

$$\tau = \frac{(\sigma_v^2 k)^2}{\sigma_\eta^2 + \sigma_v^2 k}; \text{ with } k = \frac{1-\alpha}{1+\alpha}.$$

The performance of the standard GMM differenced estimator in this AR(1) specification can therefore be seen to deteriorate as $\alpha \to 1$, as well as for decreasing values of σ_v^2 and for increasing values of σ_η^2. To illustrate this further Figure 2 provides a plot of τ against α for the case $\sigma_\eta^2 = \sigma_v^2 = 1$, $T = 3$.

FIGURE 2 ABOUT HERE

Blundell and Bond (1999) note that the finite sample bias of the first-differenced GMM estimator for the AR(1) model with weak instruments is likely to be in the direction of the Within Groups estimator. This is because the (one-step) first-differenced GMM estimator coincides with a 2SLS estimator based on the 'orthogonal deviations' transformation of Arellano and Bover (1995), and 2SLS estimators are biased in the direction of OLS in the presence of weak instruments (see, for example, Bound, Jaeger and Baker (1995)).[6] We explore the finite sample behaviour of the first-differenced GMM estimator further in Section 11 below.

5. Non-linear Moment Conditions

5.1. Standard assumptions

The standard assumptions (2.4), (2.5) and (2.6) also imply non-linear moment conditions which are not exploited by the standard linear first-differenced GMM estimator described in Section 3.1. Ahn and Schmidt (1995) show that there are a further $T-3$ non-linear moment conditions, which can be written as

$$E(u_{it}\Delta u_{i,t-1}) = 0; \text{ for } t = 4, 5, ..., T \tag{5.1}$$

and which could be expected to improve efficiency. These conditions relate directly to the absence of serial correlation in v_{it} and do not require homoskedasticity. Thus, under the standard assumptions, the complete set of second-order moment conditions available is (3.1) and (5.1). Asymptotic efficiency comparisons reported in Ahn and Schmidt (1995) confirm that these non-linear moments are particularly informative in cases where α is close to unity and/or where σ^2_η/σ^2_v is high.

5.2. Homoskedasticity

Under the homoskedasticity through time restriction (3.3), there is one further non-linear moment condition available, in addition to (3.1), (3.4) and (5.1) (see Ahn and Schmidt (1995)). This can be written as

$$E(\overline{u}_i\Delta u_{i3}) = 0 \text{ where } \overline{u}_i = \frac{1}{T-1}\sum_{t=2}^{T} u_{it}. \tag{5.2}$$

Thus, under the homoskedasticity assumption in addition to the standard assumptions, the complete set of moment conditions available comprises the linear conditions (3.1) and (3.4), and the non-linear conditions (5.1) and (5.2).

6. Initial Conditions and a Levels GMM Estimator

In addition to the standard assumptions set out in Section 2, we now consider the additional assumption

$$E\left(\eta_i \Delta y_{i2}\right) = 0 \text{ for } i = 1, ..., N. \tag{6.1}$$

Notice that, given (2.3) - (2.6) which specifies y_{i2} given y_{i1}, assumption (6.1) is a restriction on the initial conditions process generating y_{i1}.[7]

If this initial conditions restriction holds in addition to the standard assumptions (2.4), (2.5) and (2.6), the following $T-2$ linear moment conditions are valid

$$E(u_{it}\Delta y_{i,t-1}) = 0; \text{ for } t = 3, 4, ..., T. \tag{6.2}$$

Moreover, given the standard assumptions, these linear moment conditions imply the $T-3$ non-linear moment conditions given in (5.1), and render these non-linear conditions redundant for estimation. Thus the complete set of second order moment restrictions implied by (2.3)-(2.6) and (6.1) can be implemented as a linear GMM estimator.

To consider when the first-differences Δy_{it} are uncorrelated with the individual effects, notice that for the AR(1) model (2.3)

$$\Delta y_{it} = \alpha^{t-2}\Delta y_{i2} + \sum_{s=0}^{t-3} \alpha^s \Delta u_{i,t-s}$$

so that Δy_{it} will be uncorrelated with η_i if and only if Δy_{i2} is uncorrelated with η_i. This is precisely the assumption (6.1). To guarantee this, we require the initial conditions restriction

$$E\left[\left(y_{i1} - \frac{\eta_i}{1-\alpha}\right)\eta_i\right] = 0,$$

which is satisfied under mean stationarity of the y_{it} process, as defined by (2.3)-(2.8).

To show that the moment conditions (6.2) remain informative when α approaches unity or σ_η^2/σ_v^2 becomes large, we again consider the case of $T = 3$. Here we can use one equation in levels

$$y_{i3} = \alpha y_{i2} + \eta_i + v_{i3}$$

for which the instrument available is Δy_{i2}, and the first stage regression is

$$y_{i2} = \pi_l \Delta y_{i2} + r_i.$$

In this case, assuming covariance stationarity, the plim $\widehat{\pi}_l$ is given by

$$\text{plim}\, \widehat{\pi}_l = \frac{1}{2} \tag{6.3}$$

and therefore this moment condition stays informative for high values of α, in contrast to the moment condition available for the first-differenced model.

The $0.5\,(T+1)\,(T-2)$ linear moment conditions (3.1) and (6.2) comprise the full set of second-order moment conditions under mean stationarity in conjunction with the standard assumptions listed in Section 2, and form the basis for a system GMM estimator which will be discussed in the next section. However, as this system GMM estimator combines the moment conditions for the model in first-differences with those for the model in levels, we also consider a simpler GMM levels estimator, that is based on the $m_l = 0.5\,(T-1)\,(T-2)$ moment conditions

$$E\left(u_{it}\Delta y_{i,t-s}\right) = 0; \text{ for } t = 3, ..., T \text{ and } 1 \leq s \leq t-2, \tag{6.4}$$

that relate only to the equations in levels. These can be expressed as

$$E\left(\mathbf{Z}_{li}'\mathbf{u}_i\right) = 0,$$

where $\mathbf{Z}_{li}$ is the $(T-2) \times m_l$ matrix given by

$$\mathbf{Z}_{li} = \begin{bmatrix} \Delta y_{i2} & 0 & 0 & ... & 0 & ... & 0 \\ 0 & \Delta y_{i2} & \Delta y_{i3} & ... & 0 & ... & 0 \\ . & . & . & ... & . & ... & . \\ 0 & 0 & 0 & ... & \Delta y_{i2} & ... & \Delta y_{iT-1} \end{bmatrix},$$

and $\mathbf{u}_i$ is the $(T-2)$ vector $(u_{i3}, u_{i4}, ..., u_{iT})'$. Calculation of the one-step and two-step GMM estimators then proceeds in a similar way to that described above. In this case though, unless $\sigma_\eta^2 = 0$, there is no one-step GMM estimator that is asymptotically equivalent to the two-step estimator, even in the special case of i.i.d. disturbances.[9]

7. A System GMM Estimator

7.1. The Optimal Combination of Differenced and Levels Estimators

Calculation of the GMM estimator using the full set of linear moment conditions (3.1) and (6.2) can be based on a stacked system comprising all $T-2$ equations in first-differences and the $T-2$ equations in levels corresponding to periods $3, ..., T$, for which instruments are observed. The $m_s = 0.5\,(T+1)\,(T-2)$ moment conditions are[10]

$$E(y_{i,t-s}\Delta u_{it}) = 0;\ \text{for } t = 3, ..., T \text{ and } 2 \leq s \leq t-1 \tag{7.1}$$

$$E\left(u_{it}\Delta y_{i,t-1}\right) = 0;\ \text{for } t = 3, ..., T. \tag{7.2}$$

These can be expressed as

$$E\left(\mathbf{Z}'_{si}\mathbf{q}_i\right) = 0,$$

where

$$\mathbf{q}_i = \begin{bmatrix} \Delta\mathbf{u}_i \\ \mathbf{u}_i \end{bmatrix}$$

$$\mathbf{Z}_{si} = \begin{bmatrix} \mathbf{Z}_{di} & 0 \\ 0 & \mathbf{Z}_{li}^p \end{bmatrix} = \begin{bmatrix} \mathbf{Z}_{di} & 0 & 0 & ... & 0 \\ 0 & \Delta y_{i2} & 0 & ... & 0 \\ 0 & 0 & \Delta y_{i3} & ... & 0 \\ . & . & . & ... & 0 \\ 0 & 0 & 0 & ... & \Delta y_{i,T-1} \end{bmatrix};$$

with $\mathbf{Z}_{di}$ as defined in section 3, and $\mathbf{Z}_{li}^p$ is the non-redundant subset of $\mathbf{Z}_{li}$.

The calculation of the two-step GMM estimator is then analogous to that described above. Again in this case, unless $\sigma_\eta^2 = 0$, there is no one-step GMM estimator that is asymptotically equivalent to the two-step estimator, even in the special case of i.i.d. disturbances.[11]

The system GMM estimator is clearly a combination of the GMM differenced estimator and a GMM levels estimator that uses only (7.2). This combination is linear for the system 2SLS estimator which is given by

$$\widehat{\alpha}_s = \left(\mathbf{q}'_{-1}\mathbf{Z}_s\left(\mathbf{Z}'_s\mathbf{Z}_s\right)^{-1}\mathbf{Z}'_s\mathbf{q}_{-1}\right)^{-1}\mathbf{q}'_{-1}\mathbf{Z}_s\left(\mathbf{Z}'_s\mathbf{Z}_s\right)^{-1}\mathbf{Z}'_s\mathbf{q}.$$

Because

$$\mathbf{q}'_{-1}\mathbf{Z}_s\left(\mathbf{Z}'_s\mathbf{Z}_s\right)^{-1}\mathbf{Z}'_s\mathbf{q}_{-1} = \mathbf{\Delta y}'_{-1}\mathbf{Z}_d\left(\mathbf{Z}'_d\mathbf{Z}_d\right)^{-1}\mathbf{Z}'_d\mathbf{\Delta y}_{-1} + \mathbf{y}'_{-1}\mathbf{Z}^p_l\left(\mathbf{Z}^{p\prime}_l\mathbf{Z}^p_l\right)\mathbf{Z}^p_l\mathbf{y}_{-1}$$

the system 2SLS estimator is equivalent to the linear combination

$$\widehat{\alpha}_s = \gamma\widehat{\alpha}_d + (1-\gamma)\,\widehat{\alpha}^p_l,$$

where $\widehat{\alpha}_d$ and $\widehat{\alpha}^p_l$ are the 2SLS first-differenced and levels estimators respectively, with the levels estimator utilising only the $T-2$ moment conditions (7.2), and

$$\begin{aligned}\gamma &= \frac{\mathbf{\Delta y}'_{-1}\mathbf{Z}_d\left(\mathbf{Z}'_d\mathbf{Z}_d\right)^{-1}\mathbf{Z}'_d\mathbf{\Delta y}_{-1}}{\mathbf{\Delta y}'_{-1}\mathbf{Z}_d\left(\mathbf{Z}'_d\mathbf{Z}_d\right)^{-1}\mathbf{Z}'_d\mathbf{\Delta y}_{-1} + \mathbf{y}'_{-1}\mathbf{Z}^p_l\left(\mathbf{Z}^{p\prime}_l\mathbf{Z}^p_l\right)\mathbf{Z}^{p\prime}_l y_{-1}} \\ &= \frac{\widehat{\pi}'_d\mathbf{Z}'_d\mathbf{Z}_d\widehat{\pi}_d}{\widehat{\pi}'_d\mathbf{Z}'_d\mathbf{Z}_d\widehat{\pi}_d + \widehat{\pi}'_l\mathbf{Z}^{p\prime}_l\mathbf{Z}^p_l\widehat{\pi}_l},\end{aligned}$$

where $\widehat{\pi}_d$ and $\widehat{\pi}_l$ are the OLS estimates of the first stage regression coefficients underlying these 2SLS estimators. From (4.2) and (6.3) it follows that $\gamma \to 0$ if $\alpha \to 1$ and/or $(\sigma_\eta^2/\sigma_v^2) \to \infty$, so all the weight for the system estimator will in these cases be given to the informative levels moment conditions (7.2).

7.2. Homoskedasticity Restrictions

In the case where the initial conditions satisfy restriction (6.1) and the v_{it} satisfy restriction (3.3), Ahn and Schmidt (1995, equation (12b)) show that the

$T-2$ homoskedasticity restrictions (3.4) and (5.2) can be replaced by a set of $T-2$ moment conditions

$$E(y_{it}u_{it} - y_{i,t-1}u_{i,t-1}) = 0; \text{ for } t = 3, ..., T,$$

which are all linear in the parameter α. The non-linear condition (5.2) is again redundant for estimation given (6.1), and the complete set of second order moment restrictions implied by (2.3)-(2.6), (3.3) and (6.1) can be implemented in a linear GMM estimator.

8. Asymptotic Variance Comparisons

To quantify the gains in asymptotic efficiency that result from exploiting the linear moment conditions (6.2), Table 1 reports the ratio of the asymptotic variance of the standard first-differenced GMM estimator described in Section 3.1 to the asymptotic variance of the system GMM estimator described in Section 7.1. These asymptotic variance ratios are calculated assuming both covariance stationarity and homoskedasticity. They are presented for $T = 3$ and $T = 4$, for two fixed values of σ_η^2/σ_v^2, and for a range of values of the autoregressive parameter α. For comparison, we also reproduce from Ahn and Schmidt (1995) the corresponding asymptotic variance ratios comparing first-differenced GMM to the non-linear GMM estimator which uses the quadratic moment conditions (5.1), but *not* the extra linear moment conditions (6.2). In the $T = 3$ case there are no quadratic moment restrictions available. These calculations suggest that exploiting conditions (6.2) can result in dramatic efficiency gains when $T = 3$, particularly at high values of α and high values of σ_η^2/σ_v^2. These are indeed the cases where we find the instruments used to obtain the first-differenced estimator to be weak.

In the $T = 4$ case we still find dramatic efficiency gains at high values of α.

Comparison to the results for the non-linear GMM estimator also shows that the gains from exploiting conditions (6.2) can be much larger than the gains from simply exploiting the non-linear restrictions (5.1).

In the Monte Carlo simulations presented in Section 11 we investigate whether similar improvements are found in finite samples.

TABLE 1 ABOUT HERE

9. Multivariate Dynamic Panel Data Models

In this section the dynamic panel data model with additional regressors is considered.[12] In particular, we focus on the model

$$\begin{aligned} y_{it} &= \alpha y_{it-1} + \beta x_{it} + u_{it} \\ u_{it} &= \eta_i + v_{it} \end{aligned} \tag{9.1}$$

where x_{it} is a scalar. The error components η_i and v_{it} again satisfy the conditions (2.4)-(2.6). The x_{it} process is correlated with the individual effects η_i and we consider three possible correlation structures between the x_{it} process and the v_{it} error process that determine the instruments that can be used to estimate α and β.

First, the x_{it} process is strictly exogenous:

$$E\left(x_{is}v_{it}\right) = 0;\ \text{for } s = 1, ..., T;\ t = 2, ..., T. \tag{9.2}$$

Secondly, the x_{it} process is weakly exogenous, or predetermined

$$\begin{aligned} E\left(x_{is}v_{it}\right) &= 0;\ \text{for } s = 1, ..., t;\ t = 2, ..., T \\ E\left(x_{is}v_{it}\right) &\neq 0;\ \text{for } s = t+1, ..., T;\ t = 2, ..., T \end{aligned} \tag{9.3}$$

and thirdly, the x_{it} process is endogenously determined

$$\begin{aligned} E\left(x_{is}v_{it}\right) &= 0;\ \text{for } s=1,...,t-1;\ t=2,...,T \\ E\left(x_{is}v_{it}\right) &\neq 0;\ \text{for } s=t,...,T;\ t=2,...,T. \end{aligned} \tag{9.4}$$

We are especially interested in the case when the x_{it} process is endogenously determined, which includes simultaneous processes, but also measurement error.

For the GMM first-differenced estimator, the $0.5\left(T-1\right)\left(T-2\right)$ moment conditions (3.1)

$$E(y_{i,t-s}\Delta u_{it})=0;\ \text{for } t=3,...,T \text{ and } 2\leq s\leq t-1$$

remain valid. When the x_{it} process is strictly exogenous, the following additional $T\left(T-2\right)$ moment conditions are valid

$$E(x_{is}\Delta u_{it})=0;\ \text{for } t=3,...,T \text{ and } 1\leq s\leq T. \tag{9.5}$$

When x_{it} is predetermined there are only the $0.5\left(T+1\right)\left(T-2\right)$ additional moment conditions

$$E(x_{i,t-s}\Delta u_{it})=0;\ \text{for } t=3,...,T \text{ and } 1\leq s\leq t-1, \tag{9.6}$$

whereas when x_{it} is endogenously determined only the following $0.5\left(T-1\right)\left(T-2\right)$ additional moment conditions are valid

$$E(x_{i,t-s}\Delta u_{it})=0;\ \text{for } t=3,...,T \text{ and } 2\leq s\leq t-1. \tag{9.7}$$

For the non-linear GMM estimator, moment conditions (5.1) remain valid, and no further moment conditions result from the presence of x_{it} variables.

For the system GMM estimator, we first consider under what conditions both Δy_{it} and Δx_{it} are uncorrelated with η_i. In order to illustrate this, we specify the following process for the regressor

$$x_{it}=\rho x_{i,t-1}+\tau\eta_i+e_{it}.$$

Thus $\tau \neq 0$ allows the level of x_{it} to be correlated with η_i, and the covariance properties between v_{it} and e_{is} determine whether x_{it} is strictly exogenous, predetermined or endogenously determined. First notice that

$$\Delta x_{it} = \rho^{t-2}\Delta x_{i2} + \sum_{s=0}^{t-3} \rho^s \Delta e_{i,t-s},$$

so that Δx_{it} will be correlated with η_i if and only if Δx_{i2} is correlated with η_i. To guarantee $E\left[\Delta x_{i2}\eta_i\right] = 0$ we require the initial conditions restriction

$$E\left[\left(x_{i1} - \frac{\tau\eta_i}{1-\rho}\right)\tau\eta_i\right] = 0 \tag{9.8}$$

which is satisfied under mean stationarity of the x_{it} process.

Given this restriction, writing Δy_{it} as

$$\Delta y_{it} = \alpha^{t-2}\Delta y_{i2} + \sum_{s=0}^{t-3} \alpha^s \left(\beta\Delta x_{i,t-s} + \Delta u_{i,t-s}\right) \tag{9.9}$$

shows that Δy_{it} will be correlated with η_i if and only if Δy_{i2} is correlated with η_i. To guarantee $E\left[\Delta y_{i2}\eta_i\right] = 0$ we then require the similar initial conditions restriction

$$E\left[\left(y_{i1} - \frac{\beta\left(\frac{\tau\eta_i}{1-\rho}\right) + \eta_i}{1-\alpha}\right)\eta_i\right] = 0 \tag{9.10}$$

which would again be satisfied under stationarity. Thus, there are additional moment restrictions available for the equations in levels when the y_{it} and x_{it} processes are both mean stationary.

Whilst jointly stationary means is sufficient to ensure that both Δy_{it} and Δx_{it} are uncorrelated with η_i, this condition is stronger than is necessary. For example, if the conditional model (9.1) has generated the y_{it} series for sufficiently long time prior to our sample period for any influence of the true initial conditions to be negligible, then an expression analogous to (9.9) shows that Δy_{it} will be uncorrelated with η_i provided that Δx_{it} is uncorrelated with η_i, even if the mean of x_{it} (and hence y_{it}) is time-varying. Moreover we can note that it is perfectly possible for Δx_{it} to be uncorrelated with η_i in cases where Δy_{it} is correlated with η_i (for example, when (9.8) holds or $\tau = 0$ but (9.10) is not satisfied). However, given (9.9), it seems very unlikely that Δy_{it} will be uncorrelated with η_i in contexts where Δx_{it} is correlated with η_i.

When both Δy_{it} and Δx_{it} are uncorrelated with η_i, the extra moment conditions for the GMM system estimator are, as before, (7.2),

$$E\left(u_{it}\Delta y_{i,t-1}\right) = 0; \text{ for } t = 3, ..., T$$

and

$$E\left(u_{it}\Delta x_{it}\right) = 0; \text{ for } t = 2, ..., T \tag{9.11}$$

in the case where x_{it} is strictly exogenous or predetermined; or

$$E\left(u_{it}\Delta x_{it-1}\right) = 0; \text{ for } t = 3, ..., T, \tag{9.12}$$

when x_{it} is endogenously determined. Therefore, when for example x_{it} is endogenous, the GMM system estimator is based on the moment conditions (7.1), (9.7), (7.2) and (9.12).

10. Tests of Overidentifiying Restrictions

The standard test for testing the validity of the moment conditions used in the GMM estimation procedure is the Sargan test of overidentifying restrictions (see Sargan (1958) and the development for GMM in Hansen (1982*))*. For the GMM estimator in the first-differenced model this test statistic is given by

$$\text{Sar}_d = \frac{1}{N}\widehat{\mathbf{\Delta u}}'\mathbf{Z}_d\mathbf{W}_N\mathbf{Z}_d'\widehat{\mathbf{\Delta u}}$$

where $\mathbf{W}_N$ is the optimal weight matrix as in (3.2) and $\widehat{\mathbf{\Delta u}}$ are the two-step residuals in the differenced model. In general, under the null that the moment conditions are valid, Sar_d is asymptotically chi-squared distributed with $m_d - k$ degrees of freedom, where m_d is the number of moment conditions and k is the number of estimated parameters.

For the system estimator, the same test is readily defined. Call this test Sar_s. A test for the validity of the level moment conditions that are utilised by the system estimator is then obtained as the difference between Sar_s and Sar_d:

$$\text{Dif-Sar} = \text{Sar}_s - \text{Sar}_d \tag{10.1}$$

and Dif-Sar is asymptotically chi-squared distributed with $m_s - m_d$ degrees of freedom under the null that the level moment conditions are valid.

11. Monte Carlo Results

This section illustrates the performance of the various estimators, as discussed above, for a dynamic multivariate panel data model. In particular, the effect of weak instruments and the potential gains from exploiting initial conditions restrictions are investigated.

The model specification is

$$y_{it} = \alpha y_{it-1} + \beta x_{it} + \eta_i + v_{it} \tag{11.1}$$

$$x_{it} = \rho x_{it-1} + \tau \eta_i + \theta v_{it} + e_{it} \tag{11.2}$$

with

$$\eta_i \sim N\left(0, \sigma_\eta^2\right);\ v_{it} \sim N\left(0, \sigma_v^2\right);\ e_{it} \sim N\left(0, \sigma_e^2\right)$$

and the initial observations are drawn from the covariance stationary distribution. Although these errors are homoskedastic, we do not consider any of the additional moment conditions that require homoskedasticity in the simulated estimators.

We choose the error process parameters in such a way that the x_{it} process is highly persistent for high values of ρ. Further, x_{it} is positively correlated with

η_i and the value of θ is negative to mimic the effects of measurement error. The values of the parameters that are kept fixed in the various Monte Carlo simulations presented below are

$$\begin{aligned} \beta &= 1,\, \tau = 0.25,\, \theta = -0.1, \\ \sigma_\eta^2 &= 1,\, \sigma_v^2 = 1,\, \sigma_e^2 = 0.16. \end{aligned}$$

The parameters that are varied in the simulations are the autoregressive coefficients α and ρ. We consider four designs with α and ρ both taking the values of 0.5 and 0.95. The case when $\alpha = 0.5$ and $\rho = 0.95$ resembles the production function data that will be analysed in the next section. The sample size is $N = 500$, and the simulation results for the various estimators are presented in Tables 2 and 3 for $T = 4$ and in Tables 4 and 5 for $T = 8$.

Means, standard deviations and root mean squared errors (RMSE) from 10,000 simulations are tabulated for the OLS levels estimator (OLS), Within Groups estimator (WG), the GMM first-differenced estimator (DIF), the non-linear GMM estimator (AS),[13] the levels GMM estimator (LEV), and the system GMM estimator (SYS). Thus for the case of estimating the AR(1) model for x_{it}, DIF uses the moment conditions (3.1); AS uses the moment conditions (3.1) and (5.1); LEV uses the moment conditions (6.4); and SYS uses the moment conditions (3.1) and (6.2). The reported results are for the two-step GMM estimators.

Tables 2 and 4 present results for $\rho = 0.5$. The row labelled 'ρ' presents the results for the estimates of ρ in model (11.2), where the various GMM estimators only utilise lagged information on x as instruments, and potential information from the lagged values of y is not used. Our results for the DIF and SYS estimators can therefore be compared to those reported in, for example, Blundell and Bond (1998) and Alonso-Borrego and Arellano (1999). As expected, the OLS estimates are biased upward and the WG estimates are biased downwards. In this experiment where x_{it} is not highly persistent and the instruments available for the equations in first-differences are not weak, all four GMM estimators are virtually unbiased. The AS, LEV and SYS estimators all provide an improvement in precision compared to the standard DIF estimator. As we would expect from the asymptotic variance ratios in Table 1, there is a greater gain in precision from using SYS rather than AS at $T = 4$, although in Table 4 we can observe that this difference becomes very small at $T = 8$.

The next two rows in Tables 2 and 4 present the estimation results for α and β in model (11.1) when $\alpha = 0.5$ and $\rho = 0.5$. The OLS estimates for α are biased upwards, whereas those for β are biased downwards. The WG estimates for α and β are both biased downwards. Again, as expected, since both the y and x series have a low degree of persistence, the four GMM estimators perform quite well in this experiment. The SYS estimator has the smallest RMSE for both parameters, but the gains are not dramatic at $T = 8$.

The final two rows in Tables 2 and 4 are for the model with $\alpha = 0.95$ and $\rho = 0.5$. As this makes the y process highly persistent, the DIF estimator suffers from a serious weak instrument bias, as well as being very imprecise. We can notice that the DIF estimates of α and β are both biased downwards, in the direction of the Within Groups estimates. The AS estimator is better behaved, as a result of exploiting the non-linear moment conditions (5.1). However the LEV and SYS estimators which exploit the initial conditions restrictions provide more dramatic gains in precision, particularly for the estimation of α and particularly in the case with $T = 4$. With $T = 8$, the LEV and SYS estimates of α are biased upwards, in the direction of the OLS estimate, but still dominate on the RMSE criterion.

TABLES 2-5 ABOUT HERE

Tables 3 and 5 present the results for the cases where the x_{it} process is highly persistent, with $\rho = 0.95$. The estimates for ρ show the familiar pattern: OLS is upward biased, WG is downward biased, and DIF is downward biased towards WG as a result of weak instruments. The AS estimator provides a substantial improvement in both bias and precision. However the LEV and SYS estimators provide more dramatic gains, particularly when $T = 4$.

When $\alpha = 0.5$, the DIF estimator estimates α quite well, but the DIF estimate of β is very imprecise, biased downwards and on average very similar to the WG estimate of β. The AS, LEV and SYS estimates of α are all close to the true value. The AS estimates of β are much less biased than DIF but still imprecise, particularly at $T = 4$. The LEV and SYS estimates of β show a little finite-sample bias, but again dominate in terms of RMSE. This experiment is intended to capture salient features of the production function data we consider in Section 12, notably a highly persistent explanatory variable that is measured with error, and a significant autoregressive parameter that is not close to one. The simulation results confirm that the system GMM estimator has reasonable properties in this context.

When both α and ρ are equal to 0.95 the estimators display a similar pattern. One surprise is that the LEV and SYS estimators actually estimate both parameters better than in the experiments with $\alpha = 0.5$, and the gain from using either of these estimators compared to AS is rather more striking in this case. Also the DIF estimator now estimates α quite well (though not β); this may be because by increasing α whilst keeping the variance of η_i and v_{it} fixed, we have greatly increased the variance of the y_{it} series.

To investigate the size properties of the Sargan test of overidentifying restrictions, we present in Figures 3-12 p-value plots (see Davidson and MacKinnon, 1996) for the Sargan test statistics for the DIF and SYS GMM estimators. We also present the p-value plots for the Dif-Sar statistic as defined in (10.1), testing the validity of the additional levels moment conditions exploited by the SYS estimator.

The x-axis of the p-value plots represents the nominal size using the asymptotic critical values of the corresponding chi-squared distributions; the y-axis represents the actual size of the test statistics in the experiments.

Figures 3-6 are the p-value plots for the Sargan tests for the GMM estimators in the univariate model for x_{it}, (11.2). When $\rho = 0.5$, the distributions of the test statistics are all very close to the asymptotic distribution, with a slight over-rejection when $T = 8$. When the series are persistent, $\rho = 0.95$, the tests over-reject, especially for larger T, with the Dif-Sar test having the largest size distortion when $T = 4$.

FIGURES 3-6 ABOUT HERE

Figures 7-14 present the p-value plots for the Sargan test statistics for the multivariate dynamic panel data model (11.1). These appear to be well behaved in the case with $\alpha = 0.5$ and $\rho = 0.5$. In general, the Dif-Sar test is oversized when either y or x or both are persistent. An interesting case is when $\alpha = 0.5$, $\rho = 0.95$ and $T = 8$. The Sar_s and Dif-Sar tests are considerably oversized in this case, whereas the Sar_d test has the correct size.

FIGURES 7-14 ABOUT HERE

12. An Application: the Cobb-Douglas Production Function

As Griliches and Mairesse (1998) have argued, the estimation of production functions has highlighted the poor performance of standard GMM estimators for short panels. Here we use the problem of estimating production function parameters to evaluate the practical significance of the alternative estimators reviewed in this chapter. In particular attention is focused on the estimation of the Cobb-Douglas production function

$$\begin{aligned} y_{it} &= \beta_n n_{it} + \beta_k k_{it} + \gamma_t + (\eta_i + v_{it} + m_{it}) \qquad (12.1) \\ v_{it} &= \alpha v_{i,t-1} + e_{it} \qquad |\alpha| < 1 \\ e_{it}, m_{it} &\sim MA(0), \end{aligned}$$

where y_{it} is log sales of firm i in year t, n_{it} is log employment, k_{it} is log capital stock and γ_t is a year-specific intercept reflecting, for example, a common technology shock. Of the error components, η_i is an unobserved time-invariant firm-specific effect, v_{it} is a possibly autoregressive (productivity) shock and m_{it} reflects serially uncorrelated (measurement) errors. Constant returns to scale would imply $\beta_n + \beta_k = 1$, but this is not necessarily imposed.

Interest is in the consistent estimation of the parameters $(\beta_n, \beta_k, \alpha)$ when the number of firms (N) is large and the number of years (T) is fixed. We maintain that both employment (n_{it}) and capital (k_{it}) are potentially correlated with the firm-specific effects (η_i), and with both productivity shocks (e_{it}) and measurement errors (m_{it}).

The model has a dynamic (common factor) representation

$$\begin{aligned} y_{it} = {} & \beta_n n_{it} - \alpha\beta_n n_{i,t-1} + \beta_k k_{it} - \alpha\beta_k k_{i,t-1} + \alpha y_{i,t-1} \qquad (12.2) \\ & +(\gamma_t - \alpha\gamma_{t-1}) + (\eta_i(1-\alpha) + e_{it} + m_{it} - \alpha m_{i,t-1}) \end{aligned}$$

or

$$y_{it} = \pi_1 n_{it} + \pi_2 n_{i,t-1} + \pi_3 k_{it} + \pi_4 k_{i,t-1} + \pi_5 y_{i,t-1} + \gamma_t^* + (\eta_i^* + w_{it}) \qquad (12.3)$$

subject to two non-linear (common factor) restrictions $\pi_2 = -\pi_1\pi_5$ and $\pi_4 = -\pi_3\pi_5$. Given consistent estimates of the unrestricted parameter vector $\pi =$

$(\pi_1, \pi_2, \pi_3, \pi_4, \pi_5)$ and $\text{var}(\pi)$, these restrictions can be (tested and) imposed using minimum distance to obtain the restricted parameter vector $(\beta_n, \beta_k, \alpha)$. Notice that $w_{it} = e_{it} \sim MA(0)$ if there are no measurement errors ($\text{var}\,(m_{it}) = 0$), and $w_{it} \sim MA(1)$ otherwise.

12.1. Data and Results

The data used is a balanced panel of 509 R&D-performing US manufacturing companies observed for 8 years, 1982-89. These data were kindly made available to us by Bronwyn Hall, and is similar to that used in Mairesse and Hall (1996), although the sample of 509 firms used here is larger than the final sample of 442 firms used in Mairesse and Hall (1996). Capital stock and employment are measured at the end of the firm's accounting year, and sales is used as a proxy for output. Further details of the data construction can be found in Mairesse and Hall (1996).

Table 6 reports results for the basic production function, not imposing constant returns to scale, for a range of estimators. We report results for both the unrestricted model (12.3) and the restricted model (12.1), where the common factor restrictions are tested and imposed using minimum distance.[14] We report results here for the one-step GMM estimators, for which inference based on the asymptotic variance matrix has been found to be more reliable than for the (asymptotically) more efficient two-step estimator. Simulations suggest that the loss in precision that results from not using the optimal weight matrix is unlikely to be large (cf. Blundell and Bond, 1998).

As expected in the presence of firm-specific effects, OLS levels appears to give an upwards-biased estimate of the coefficient on the lagged dependent variable, whilst Within Groups appears to give a downwards-biased estimate of this coefficient. Note that even using OLS, we reject the hypothesis that $\alpha = 1$, and even using Within Groups we reject the hypothesis that $\alpha = 0$. Although the pattern of signs on current and lagged regressors in the unrestricted models are consistent with the AR(1) error-component specification, the common factor restrictions are rejected for both these estimators. They also reject constant returns to scale.[15]

The validity of lagged levels dated $t-2$ as instruments in the first-differenced equations is clearly rejected by the Sargan test of overidentifying restrictions. This is consistent with the presence of measurement errors. Instruments dated t-3 (and earlier) are accepted, and the test of common factor restrictions is easily passed in these first-differenced GMM results. However the estimated coefficient on the lagged dependent variable is barely higher than the Within Groups estimate. We expect this coefficient to be biased downwards if the instruments available are weak, as the Monte Carlo results in the previous section indicated. Indeed the differenced GMM parameter estimates are all very close to the Within Groups results. The estimate of β_k is low and statistically weak, and the constant returns to scale restriction is rejected.

The validity of lagged levels dated $t-3$ (and earlier) as instruments in the first-differenced equations, combined with lagged first-differences dated $t-2$ as instruments in the levels equations, appears to be marginal in the system GMM estimator. However we have seen that these tests do have some tendency to

TABLE 6 ABOUT HERE

overreject in samples of this size. Moreover the Dif-Sar statistic that specifically tests the additional moment conditions used in the levels equations accepts their validity at the 10% level. The system GMM parameter estimates appear to be reasonable. The estimated coefficient on the lagged dependent variable is higher than the Within Groups estimate, but well below the OLS levels estimate. The common factor restrictions are easily accepted, and the estimate of β_k is both higher and better determined than the differenced GMM estimate. The constant returns to scale restriction is easily accepted in the system GMM results.[16]

Blundell and Bond (1999) explore this data in more detail and conclude that the system GMM estimates in the final column of Table 6 are their preferred results. In particular they find that the individual series used here are highly persistent, and that the instruments available for the first-differenced equations are only weakly correlated with the explanatory variables in first-differences. This is consistent with the similarity between the first-differenced GMM and Within Groups results. Blundell and Bond (1999) also find that when constant returns to scale is imposed on the production function - it is not rejected in the preferred system GMM results - then the results obtained using the first-differenced GMM estimator become more similar to the system GMM estimates.

13. Summary and Conclusions

The aim of this chapter has been to review developments in the recent literature which have tried to improve on the poor performance of the standard first-differenced GMM estimator for highly autoregressive panel series by using additional moment conditions. In particular, we discuss the use of the "system" GMM estimator that relies on relatively mild restrictions on the initial conditions process. This system GMM estimator encompasses the GMM estimator based on the non-linear moment conditions available in the dynamic error components model and has substantial asymptotic efficiency gains relative to this non-linear GMM estimator. The chapter systematically sets out the assumptions required and moment conditions used by each estimator and provides a Monte Carlo simulation comparison as well as an application to production function estimation.

The simulation results are the first in the literature to consider the properties of these GMM estimators in dynamic models with endogenous regressors. Our analysis suggests that similar issues arise in this case to those that have been found in previous Monte Carlo studies for the AR(1) model. In particular, we find both a large bias and very low precision for the standard first-differenced estimator when the individual series are highly persistent. By exploiting instruments available for the equations in levels, the system GMM estimator can both greatly improve the precision and greatly reduce the finite sample bias when these additional moment conditions are valid. Intermediate results are found for the non-linear GMM estimator considered, which suggests that this estimator could also be useful in applications with persistent series where the validity of the initial conditions restrictions required for the system GMM estimator are rejected.

The empirical application uses company accounts data for the US to estimate a simple Cobb-Douglas production function. For the standard GMM estimator that uses moment conditions only for the first-differenced equations, we confirm the problems noted by Griliches and Mairesse: the estimated coefficient on capital is very low, all coefficient estimates are imprecise, and constant returns to scale is easily rejected. We notice that the first-differenced GMM results are similar to the Within Groups results, which suggests there may be a problem of weak instruments. This suggestion is consistent with the persistence of the underlying sales, employment and capital stock series. The additional moment conditions used by the system GMM estimator are not rejected in this context, and lead to a marked improvement in the empirical results.

Taken together, these Monte Carlo and empirical results suggest that careful consideration of the underlying series and comparisons between different panel data estimators can be useful in detecting situations where the standard first-differenced GMM estimator is likely to be subject to serious weak instruments biases. Where appropriate, the use of the system GMM estimator offers a simple and powerful alternative, that can overcome many of the disappointing features of the standard first-differenced GMM estimator in the context of highly persistent series.

Acknowledgements: This research is part of the programme of research at the ESRC Centre for the Micro-Economic Analysis of Fiscal Policy at IFS. Financial support from the ESRC is gratefully acknowledged.

Endnotes

1. All of the estimators discussed and their properties extend in an obvious fashion to higher order autoregressive models.

2. Extensions to dynamic models with additional regressors are considered in Section 9.

3. With $T = 3$, the absence of serial correlation in v_{it} (2.5) and predetermined initial conditions (2.6) are required to identify α (in the absence of any strictly exogenous instruments). With $T > 3$, α can be identified in the presence of suitably low order moving average autocorrelation in v_{it}.

4. These estimators are all based on the normalisation (2.3). Alonso-Borrego and Arellano (1999) consider a symmetrically normalised instrumental variable estimator based on the normalisation invariance of the standard LIML estimator.

5. As a choice of $\mathbf{W}_N$ to yield the initial consistent estimator, Arellano and Bond (1991) suggest

$$\mathbf{W}_N = (\frac{1}{N}\sum_{i=1}^{N}\mathbf{Z}'_{di}\mathbf{H}_d\mathbf{Z}_{di})^{-1}$$

where $\mathbf{H}_d$ is the $(T-2)\times(T-2)$ matrix given by

$$\mathbf{H}_d = \begin{pmatrix} 2 & -1 & 0 & ... & 0 \\ -1 & 2 & -1 & ... & 0 \\ 0 & -1 & 2 & ... & 0 \\ ... & ... & ... & ... & . \\ 0 & 0 & 0 & ... & 2 \end{pmatrix}$$

which can be calculated in one step. The use of this $\mathbf{H}_d$ matrix accounts for the first-order moving average structure in Δu_{it} induced by the first-differencing transformation. Note that when the v_{it} are i.i.d., the one-step and two-step estimators are asymptotically equivalent in this model. We follow this suggestion in the Monte Carlo simulations in Section 11.

6. As shown by Arellano and Bover (1995), OLS on the model transformed to orthogonal deviations coincides with the Within Groups estimator.

7. In this section we focus only on moment conditions that are valid under heteroskedasticity. The case with homoskedasticity and assumption (6.1) is considered in Section 7.2.

8. This corrects the expression for plim $\widehat{\pi}_l$ as given in Blundell and Bond (1998, p.125).

9. As a choice of $\mathbf{W}_N$ to yield the initial consistent estimator, we use

$$\mathbf{W}_N = (\frac{1}{N}\sum_{i=1}^{N}\mathbf{Z}_{li}'\mathbf{Z}_{li})^{-1}$$

in the Monte Carlo simulations reported below.

10. The use of moment conditions $E\left(u_{it}\Delta y_{i,t-s}\right) = 0$ for $s > 1$ can be shown to be redundant, given (7.1) and (7.2). For balanced panels, the $T-2$ equations in levels may be replaced by a single levels equation for period T, with (7.2) replaced by the equivalent moment conditions $E\left(u_{iT}\Delta y_{i,T-s}\right) = 0$ for $s = 1, ..., T-1$. However this approach does not extend easily to the case of unbalanced panels.

11. For an analysis of the potential loss in efficiency due to specific choices of the initial weight matrix for these system estimators, see Windmeijer (2000). As a choice of $\mathbf{W}_N$ to yield the initial consistent estimator, we use

$$\mathbf{W}_N = (\frac{1}{N}\sum_{i=1}^{N}\mathbf{Z}_{si}'\mathbf{H}_s\mathbf{Z}_{si})^{-1}$$

in our Monte Carlo simulations, where $\mathbf{H}_s$ is the matrix

$$\begin{pmatrix} \mathbf{H}_d & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{T-2} \end{pmatrix},$$

$\mathbf{I}_{T-2}$ is the $(T-2)$ identity matrix and $\mathbf{H}_d$ is defined in Section 3.

12. Here we only consider moment conditions that do not require any homoskedasticity assumptions.

13. Define $\mathbf{s}_i = [u_{i3} - u_{i2}, ..., u_{iT} - u_{iT-1}, u_{i4}(u_{i3} - u_{i2}), ..., u_{iT}(u_{iT-1} - u_{iT-2})]'$ and $\mathbf{Z}_{nli} = \begin{bmatrix} \mathbf{Z}_{di} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{T-3} \end{bmatrix}$, then the non-linear moment conditions can be written as $E[\mathbf{Z}'_{nli}\mathbf{s}_i] = 0$. As an initial weight matrix we use $\mathbf{W}_N = \left(\frac{1}{N}\sum_{i=1}^{N}\mathbf{Z}'_{nli}\mathbf{Z}_{nli}\right)^{-1}$, see Meghir and Windmeijer (1999).

14. The unrestricted results are computed using DPD98 for GAUSS (see Arellano and Bond, 1998).

15. The table reports p-values from minimum distance tests of the common factor restrictions and Wald tests of the constant returns to scale restrictions.

16. One puzzle is that we find little evidence of second-order serial correlation in the first-differenced residuals (i.e. an $MA(1)$ component in the error term in levels), although the use of instruments dated $t-2$ is strongly rejected. It may be that the e_{it} productivity shocks are also $MA(1)$, in a way that happens to offset the appearance of serial correlation that would otherwise result from measurement errors.

References

[1] Ahn, S.C. and P. Schmidt (1995), Efficient Estimation of Models for Dynamic Panel Data, *Journal of Econometrics*, 68, 5-28.

[2] Alonso-Borrego, C. and M. Arellano (1999), Symmetrically Normalised Instrumental-Variable Estimation using Panel Data, *Journal of Business and Economic Statistics*, 17, 36-49.

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

[4] Arellano, M. and S.R. Bond (1991), Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations, *Review of Economic Studies*, 58, 277-297.

[5] Arellano, M. and S.R. Bond (1998), Dynamic Panel Data Estimation using DPD98 for GAUSS, http://www.ifs.org.uk/staff/steve_b.shtml.

[6] Arellano, M. and O. Bover (1995), Another Look at the Instrumental-Variable Estimation of Error-Components Models, *Journal of Econometrics*, 68, 29-52.

[7] Baltagi, B.H. (1995), *Econometric Analysis of Panel Data*, Wiley.

[8] Bhagarva, A. and J.D. Sargan (1983), Estimating Dynamic Random Effects Models from Panel Data Covering Short Time Periods, *Econometrica*, 51, 1635-1659.

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

[10] Blundell, R., and S. Bond (1999), GMM Estimation with Persistent Panel Data: An Application to Production Function Estimation, Working Paper Series No. W99/4, The Institute for Fiscal Studies, London.

[11] Bound, J., D.A. Jaeger and R.M. Baker (1995), Problems with Instrumental Variables Estimation when the Correlation between the Instruments and the Endogenous Explanatory Variable is Weak, *Journal of the American Statistical Association* 90, 443-450.

[12] Chamberlain, G. (1987), Asymptotic Efficiency in Estimation with Conditional Moment Restrictions, *Journal of Econometrics*, 34, 305-334.

[13] Davidson, R. and J.G. MacKinnon (1996), Graphical Methods for Investigating the Size and Power of Hypothesis Tests, *Manchester School*, 66, 1-26.

[14] Griliches, Z. and J. Mairesse (1998), Production Functions: the Search for Identification, in S. Strom (ed.), *Essays in Honour of Ragnar Frisch*, Econometric Society Monograph Series, Cambridge University Press, Cambridge.

[15] Hansen, L.P. (1982), Large Sample Properties of Generalised Method of Moment Estimators, *Econometrica*, 50, 1029-1054.

[16] Holtz-Eakin, D., W. Newey, and H.S. Rosen (1988), Estimating Vector Autoregressions with Panel Data, *Econometrica*, 56, 1371-1396.

[17] Hsiao, C. (1986), *Analysis of Panel Data*, Cambridge: Cambridge University Press.

[18] Mairesse, J. and B.H. Hall (1996), Estimating the Productivity of Research and Development in French and US Manufacturing Firms: an Exploration of Simultaneity Issues with GMM Methods, in Wagner, K. and B. Van Ark (eds.), *International Productivity Differences and Their Explanations*, Elsevier Science, 285-315.

[19] Meghir, C. and F. Windmeijer (1999), Moment Conditions for Dynamic Panel Data Models with Multiplicative Individual Effects in the Conditional Variance, *Annales d'Économie et de Statistique*, 55/56, 317-330.

[20] Nelson, C.R. and R. Startz (1990a), Some Further Results on the Exact Small Sample Properties of the Instrumental Variable Estimator, *Econometrica*, 58, 967-976.

[21] Nelson, C.R. and R. Startz (1990b), The Distribution of the Instrumental Variable Estimator and its t-ratio when the Instrument is a Poor One, *Journal of Business and Economic Statistics*, 63, 5125-5140.

[22] Nickell, S.J. (1981), Biases in Dynamic Models with Fixed Effects, *Econometrica*, 49, 1417-1426.

[23] Sargan, J.D. (1958), The Estimation of Economic Relationships using Instrumental Variables, *Econometrica*, 26, 329-338.

[24] Staiger, D. and J.H. Stock (1997), Instrumental Variables Regression with Weak Instruments, *Econometrica, 65, 557-586.*

[25] Windmeijer, F. (2000), Efficiency Comparisons for a System GMM Estimator in Dynamic Panel Data Models, in Heijmans, R.D.H., D.S.G. Pollock and A. Satorra (eds.), *Innovations in Multivariate Statistical Analysis. A Festschrift for Heinz Neudecker*, Kluwer Academic Publishers, 175-184.

Table 1: Asymptotic Variance Ratios

		$\sigma^2_\eta/\sigma^2_v = 1.00$		$\sigma^2_\eta/\sigma^2_v = 0.25$	
	α	SYS	NON-LINEAR	SYS	NON-LINEAR
$T = 3$	0.0	1.33	n/a	1.33	n/a
	0.3	2.15		1.89	
	0.5	4.00		2.91	
	0.8	28.00		13.10	
	0.9	121.33		47.91	
$T = 4$	0.0	1.75	1.67	1.40	1.29
	0.3	2.31	1.91	1.77	1.33
	0.5	3.26	2.10	2.42	1.35
	0.8	13.97	2.42	8.88	1.41
	0.9	55.40	2.54	30.90	1.45

Source: Blundell and Bond (1998)

Table 2. Monte-Carlo results, $T = 4$, $\rho = 0.5$, $\beta = 1$, $N = 500$

		OLS		WG		DIF		AS		LEV		
		Mean	St D rmse	Mean	St D rmse	Mean	St D rmse	Mean	St D rmse	Mean	St D rmse	M
	ρ	0.762	0.017	-0.036	0.030	0.496	0.090	0.501	0.075	0.502	0.059	0.
			0.263		0.538		0.091		0.075		0.059	
$\alpha = 0.5$	α	0.820	0.011	0.010	0.031	0.469	0.131	0.516	0.095	0.512	0.070	0.
			0.320		0.491		0.135		0.096		0.070	
	β	0.775	0.053	0.318	0.080	0.915	0.420	1.006	0.351	1.029	0.336	1.
			0.231		0.687		0.428		0.351		0.337	
$\alpha = 0.95$	α	0.990	0.001	0.300	0.032	0.350	0.487	0.840	0.242	0.980	0.029	0.
			0.040		0.651		0.773		0.266		0.042	
	β	0.583	0.053	0.194	0.075	-0.195	0.994	0.790	0.524	1.004	0.289	1.
			0.420		0.809		1.554		0.565		0.289	

Means and standard devations of 10,000 replications. DIF, AS, LEV and SYS are two-step estimator

Table 3. Monte-Carlo results, $T = 4$, $\rho = 0.95$, $\beta = 1$, $N = 500$

		OLS		WG		DIF		AS		LEV		
		Mean	St D rmse	Mean	St D rmse	Mean	St D rmse	Mean	St D rmse	Mean	St D rmse	Me
	ρ	0.997	0.002	0.221	0.032	0.472	0.825	0.868	0.221	0.961	0.144	0.9
			0.047		0.729		0.954		0.235		0.145	
$\alpha = 0.5$	α	0.650	0.014	0.089	0.031	0.466	0.103	0.500	0.065	0.518	0.053	0.5
			0.151		0.412		0.109		0.065		0.056	
	β	0.830	0.034	0.551	0.090	0.517	1.438	1.021	0.461	1.078	0.160	1.0
			0.174		0.458		1.522		0.461		0.178	
$\alpha = 0.95$	α	0.962	0.001	0.661	0.026	0.907	0.104	0.936	0.072	0.957	0.008	0.9
			0.012		0.290		0.112		0.074		0.010	
	β	0.904	0.026	0.465	0.089	0.233	1.769	0.863	0.853	1.020	0.091	1.0
			0.100		0.543		1.928		0.864		0.093	

Means and standard devations of 10,000 replications. DIF, AS, LEV and SYS are two-step estimators

Table 4. Monte Carlo results, $T = 8$, $\rho = 0.5$, $\beta = 1$, $N = 500$

		OLS		WG		DIF		AS		LEV		
		Mean	St D rmse	Mean	St D rmse	Mean	St D rmse	Mean	St D rmse	Mean	St D rmse	Me
	ρ	0.762	0.012 0.262	0.265	0.018 0.236	0.494	0.034 0.035	0.495	0.025 0.026	0.503	0.029 0.029	0.5
$\alpha = 0.5$	α	0.820	0.007 0.320	0.311	0.017 0.190	0.480	0.040 0.045	0.497	0.029 0.029	0.523	0.034 0.041	0.5
	β	0.775	0.034 0.228	0.490	0.045 0.512	0.930	0.136 0.153	0.944	0.134 0.145	1.041	0.157 0.162	0.9
$\alpha = 0.95$	α	0.990	0.001 0.040	0.662	0.016 0.289	0.548	0.177 0.440	0.969	0.030 0.036	0.982	0.007 0.032	0.9
	β	0.581	0.035 0.421	0.388	0.044 0.613	0.226	0.356 0.852	0.972	0.134 0.137	0.979	0.108 0.110	0.9

Means and standard devations of 10,000 replications. DIF, AS, LEV and SYS are two-step estimator

Table 5. Monte Carlo results, $T = 8$, $\rho = 0.95$, $\beta = 1$, $N = 500$

		OLS		WG		DIF		AS		LEV		
		Mean	St D	Mean	St D	Mean	St D	Mean	St D	Mean	St D	Me
			rmse		rmse		rmse		rmse		rmse	
	ρ	0.997	0.001	0.591	0.017	0.676	0.222	0.903	0.061	0.973	0.022	0.9
			0.047		0.359		0.350		0.077		0.032	
$\alpha = 0.5$	α	0.650	0.009	0.396	0.015	0.480	0.033	0.508	0.024	0.523	0.022	0.5
			0.150		0.106		0.039		0.025		0.032	
	β	0.830	0.022	0.796	0.040	0.800	0.290	1.099	0.125	1.084	0.058	1.0
			0.171		0.208		0.352		0.159		0.101	
$\alpha = 0.95$	α	0.962	0.001	0.882	0.009	0.927	0.025	0.956	0.007	0.957	0.002	0.9
			0.012		0.068		0.034		0.009		0.007	
	β	0.902	0.017	0.745	0.040	0.615	0.400	1.016	0.118	1.017	0.028	1.0
			0.100		0.258		0.555		0.119		0.033	

Means and standard devations of 10,000 replications. DIF, AS, LEV and SYS are two-step estimator

Table 6. Production Function Estimates

	OLS Levels	Within Groups	DIF $t-2$	DIF $t-3$	SYS $t-2$	SYS $t-3$
n_t	0.479	0.488	0.513	0.499	0.629	0.472
	(0.029)	(0.030)	(0.089)	(0.101)	(0.106)	(0.112)
n_{t-1}	-0.423	-0.023	0.073	-0.147	-0.092	-0.278
	(0.031)	(0.034)	(0.093)	(0.113)	(0.108)	(0.120)
k_t	0.235	0.177	0.132	0.194	0.361	0.398
	(0.035)	(0.034)	(0.118)	(0.154)	(0.129)	(0.152)
k_{t-1}	-0.212	-0.131	-0.207	-0.105	-0.326	-0.209
	(0.035)	(0.025)	(0.095)	(0.110)	(0.104)	(0.119)
y_{t-1}	0.922	0.404	0.326	0.426	0.462	0.602
	(0.011)	(0.029)	(0.052)	(0.079)	(0.051)	(0.098)
m1	-2.60	-8.89	-6.21	-4.84	-8.14	-6.53
m2	-2.06	-1.09	-1.36	-0.69	-0.59	-0.35
Sar	-	-	.001	.073	.000	.032
Dif-Sar	-	-	-	-	.001	.102
β_n	0.538	0.488	0.583	0.515	0.773	0.479
	(0.025)	(0.030)	(0.085)	(0.099)	(0.093)	(0.098)
β_k	0.266	0.199	0.062	0.225	0.231	0.492
	(0.032)	(0.033)	(0.079)	(0.126)	(0.075)	(0.074)
α	0.964	0.512	0.377	0.448	0.509	0.565
	(0.006)	(0.022)	(0.049)	(0.073)	(0.048)	(0.078)
Comfac	.000	.000	.014	.711	.012	.772
CRS	.000	.000	.000	.006	.922	.641

Asymptotic standard errors in parentheses. Year dummies included in all models. m1 and m2 are tests for first- and second-order serial correlation, asymptotically N(0,1). We test the levels residuals for OLS levels, and the first-differenced residuals in all other columns.

Comfac is a minimum distance test of the non-linear common factor restrictions imposed in the restricted models. P-values are reported (also for Sar and Dif-Sar).

CRS is a Wald test of the constant resturns to scale hypothesis $\beta_n + \beta_k = 1$ in the restricted models. P-values are reported.

Source: Blundell and Bond (1999).

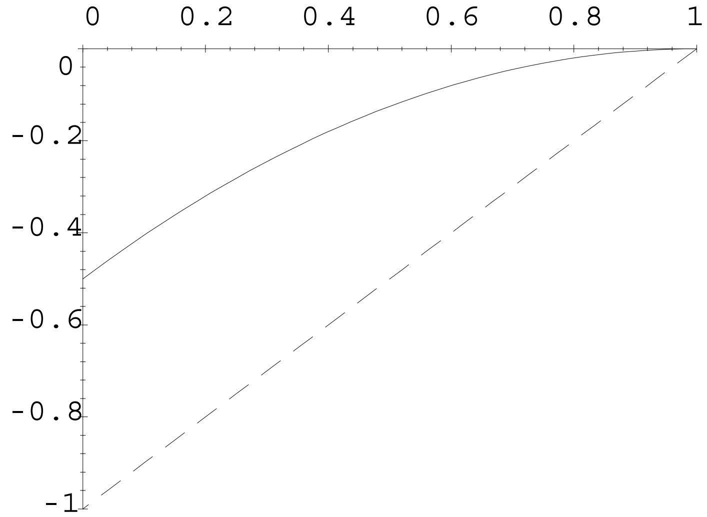

Figure 1: plim $\widehat{\pi}_d$ and $\alpha - 1$, $\sigma_\eta^2 = \sigma_v^2$, $T = 3$.
Source: Blundell and Bond (1998)

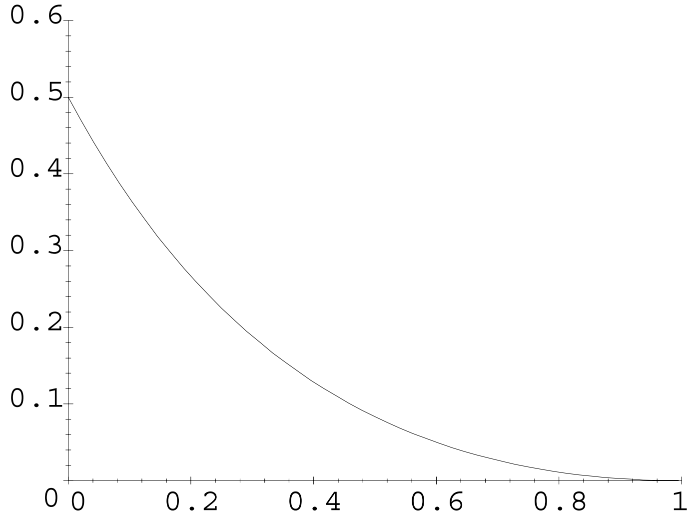

Figure 2: Concentration Parameter τ, $\sigma_\eta^2 = \sigma_v^2 = 1$, $T = 3$.
Source: Blundell and Bond (1998)

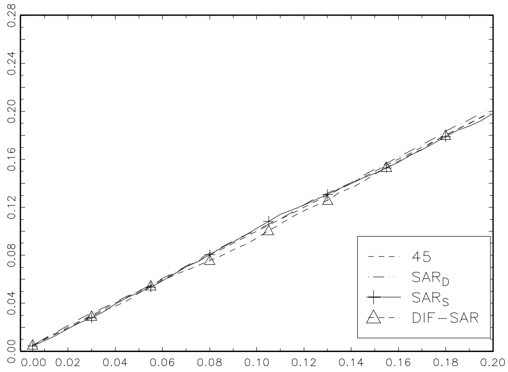

Figure 3. p-value plot, $\rho = 0.5$, $T = 4$

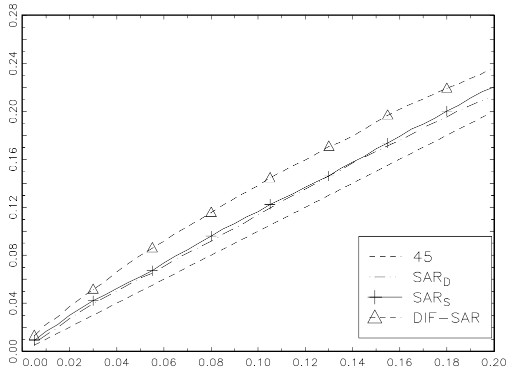

Figure 4. p-value plot, $\rho = 0.95$, $T = 4$

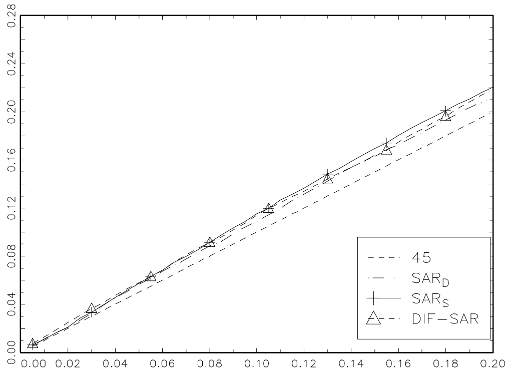

Figure 5. p-value plot, $\rho = 0.5$, $T = 8$

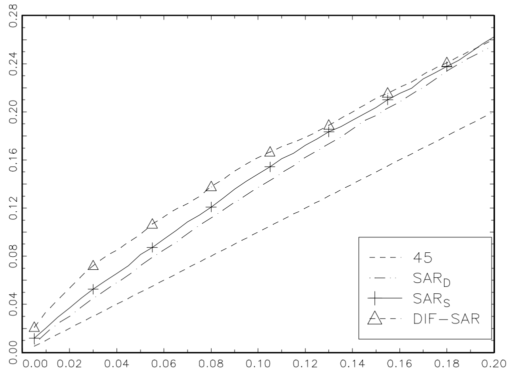

Figure 6. p-value plot, $\rho = 0.95$, $T = 8$

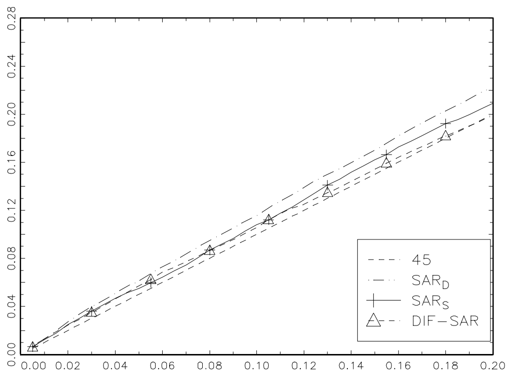

Figure 7. $\rho = 0.5$, $\alpha = 0.5$, $T = 4$

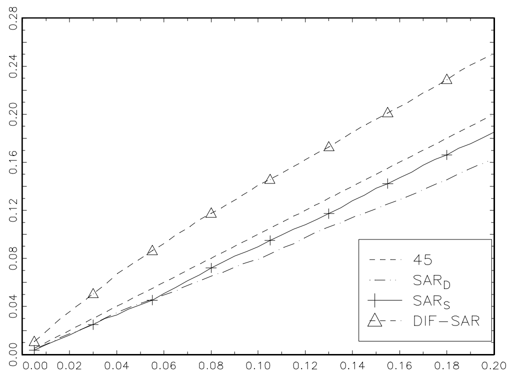

Figure 8. $\rho = 0.5$, $\alpha = 0.95$, $T = 4$

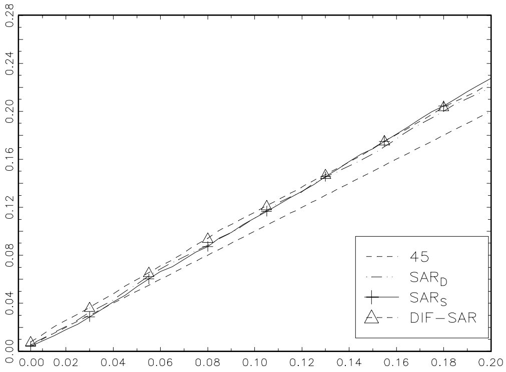

Figure 9. $\rho = 0.5$, $\alpha = 0.5$, $T = 8$

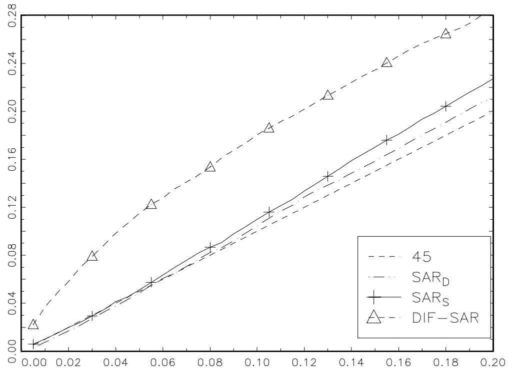

Figure 10. $\rho = 0.5$, $\alpha = 0.95$, $T = 8$

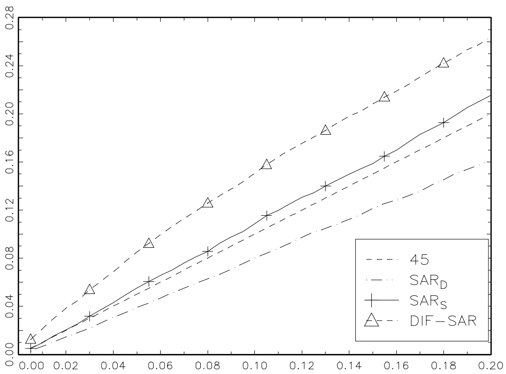

Figure 11. $\rho = 0.95$, $\alpha = 0.5$, $T = 4$

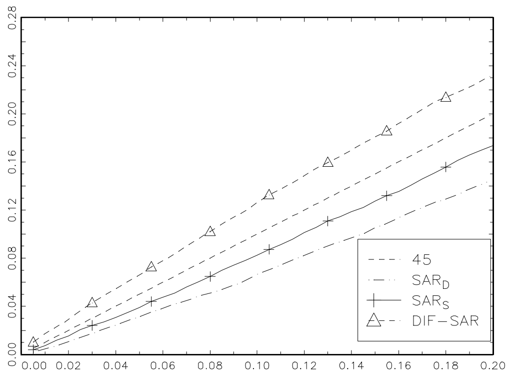

Figure 12. $\rho = 0.95$, $\alpha = 0.95$, $T = 4$

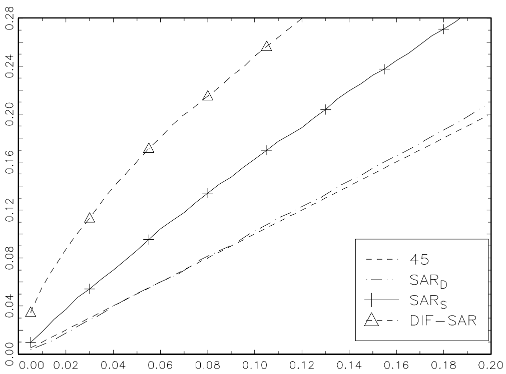

Figure 13. $\rho = 0.95$, $\alpha = 0.5$, $T = 8$

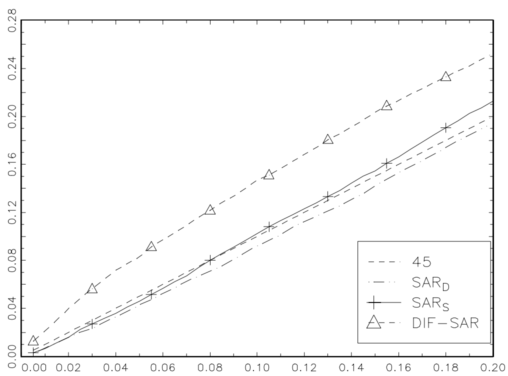

Figure 14. $\rho = 0.95$, $\alpha = 0.95$, $T = 8$