MONASH

WP 4/96

ISSN 1032-3813
ISBN 0 7326 0784 1

MONASH UNIVERSITY

MUDET

AUSTRALIA

A COMPARATIVE ANALYSIS OF DIFFERENT ESTIMATORS FOR DYNAMIC PANEL DATA MODELS

Mark N. Harris and Laszlo Matyas

Working Paper 4/96
June 1996

DEPARTMENT OF ECONOMETRICS

A Comparative Analysis of Different Estimators for Dynamic Panel Data Models[+]

Mark N. Harris[*] and László Mátyás[**]

[*]Dept. of Econometrics
Monash University
Clayton, 3168
Australia

[**]Monash University and Budapest University of Economics

Abstract

It has become increasingly obvious that the estimation of dynamic panel data models has become one of the major issues in recent econometric research as evidenced by the plethora of papers on the subject. It is well known that the usual techniques for estimating panel data models are inconsistent in the dynamic setting. However, numerous consistent estimators have been proposed in the literature. In this paper, two new estimators are offered (one each for the fixed and random effects specifications), and their small sample performance compared with that of all of the existing estimators. It is hoped that the results of these experiments will provide invaluable guidance to applied researchers as to which is the preferred estimator(s). Finally, the divergences in point estimates of all of these estimators is illustrated with an application to a consumer demand schedule of laundry detergent in the metropolitan district of Melbourne, Australia.

JEL Classification: C13, C15 and C23.

May-96

[+] Research assistance by Ritchard Longmire and László Konya is kindly acknowledged. We are also grateful to Merran Evans for helpful suggestions. Any remaining errors are our own.

1. Introduction

As a consequence of increases in both the availability of panel data sets, and in the stock of tools the applied researcher has to analyse them, the area of panel data econometrics has become very popular over the last decade. Moreover, given the vast array of economic theories espousing some form of partial adjustment of economic variables to an equilibrium level, it has become increasingly obvious that attention must be paid to estimation of dynamic panel models. That is, panel models which include lagged value(s) of the endogenous variable as explanatory variable(s). Indeed, this topic has been the focus of many recent theoretical and simulation papers (see for example, Arellano and Bond [1991], Arellano and Bover [1993], Kiviet [1994], Ahn and Schmidt [1993, 1995], and Crépon *et al.* [1996]).

Panel data sets are likely to be characterised by unobserved individual (and eventually time) heterogeneity. To account for this heterogeneity, the two model specifications most frequently used are the *fixed* and *random* effects models. Whereas the later assumes that the individual effects are drawings from a particular distribution, the former treats them as fixed parameters. There is much debate in the literature as to which is the "preferred" specification (see for example, Mundlak [1978a, b] and Hsiao[1985, 1986]), although in this paper we consider estimators of both specifications.

Estimation of dynamic panel models, irrespective of the specification of the heterogeneity, is unfortunately problematic. For the fixed effects specification, the problem arises as a consequence of the relatively short time series component, typical of most panel data sets. Thus the usual Hurwicz type bias is instigated into ordinary least squares (OLS) estimation of a fixed effects dynamic panel model (Nickell [1981]). In the random effects specification, traditional (feasible) generalised least squares estimators - (F)GLS - are similarly biased due to the correlation between the equation's disturbance terms (via the individual effect) and the lagged dependent variable (Sevestre and Trognon [1985]).

Consistent estimators for both specifications however are available. Such estimators generally take the form of instrumental variables (IV's). IV estimation involves utilising certain orthogonality conditions, primarily that the "instruments" are asymptotically uncorrelated with the equation's disturbance terms. However, using a wider set of such orthogonality conditions leads to the more general area of Generalised Method of Moments (GMM) estimation. Indeed, such GMM estimation has spawned much interest in attempting to identify the maximum (and optimal) number of such conditions (Ahn and Schmidt [1993, 1995] and Crépon *et al.* [1996]).

It is the purpose of this paper to firstly compare all of the existing IV/GMM estimators' small sample performance. Secondly, two new estimators (one each for the fixed and random effects specifications) are offered, and their small sample performance is compared to that of the existing ones. With these results an applied researcher can be confident in using the most appropriate estimator for his/her particular data set.

The plan of this paper is as follows. Sections 2 and 3 deal with model specification, "traditional" estimation and (semi-)consistent estimators for dynamic fixed and random effects models, respectively. Section 4 describes the simulation study utilised and discusses its results. The divergences across the estimators are illustrated by an application of them to consumer purchases of laundry detergent (using the Roy Morgan Research Centre's Consumer Panel of Australia data set on consumer purchases in the Melbourne Metropolitan area) in Sections 5. Finally, some concluding remarks are drawn in Section 6.

2. The Fixed Effects Dynamic Panel Model

2.1 The Model

It is assumed that the variable of interest y_{it}, is a linear function of the individual's previous realisation of this variable, and of their contemporaneous personal characteristics x_{it}, with unknown coefficients, δ and $\underline{\beta}$ respectively. Thus one may write:

(1) $$y_{it} = \alpha_i + \delta y_{it-1} + \underline{x'}_{it}\underline{\beta} + u_{it},$$

where: α_i are the *individual effects* (constant for each *i*)

and u_{it} are the usual white noise disturbance terms,

or in matrix form:

(2) $$\underline{y} = D\underline{\alpha} + \delta\underline{y}_{-1} + X\underline{\beta} + \underline{u}$$

where: $D = I_N \otimes \underline{\iota}_T$ and ι_T is the $T \times 1$ unit vector.

The usual method of estimating equation (1) or (2), *i.e.,* when there is no lagged dependent variable (LDV), consists of estimating the equation directly by OLS (the *Least Squares Dummy Variable Estimator - LSDV*), which also leads to the well known *Within* estimator. It is generally assumed that (see for example, Balestra [1992]):

HF1: the x-variables are non-stochastic and uncorrelated with the disturbances, u_{it}..

HF2: The disturbances have zero mean.

HF3: The disturbances are serially uncorrelated.

HF4: The individual effects are time invariant.

Given the short time series component typical of panel data sets, the OLS and *Within* estimators are well known to be biased and inconsistent as $N \to \infty$ and finite T (see Nickell [1981] and Sevestre and Trognon [1985] for a theoretical approach, and Nerlove [1967, 1971] for a simulation based study).

2.2 Instrumental Variable (IV) Estimators

The Balestra-Nerlove Estimator ($BN^{(L)}$).

Balestra and Nerlove (1966) show that consistent parameter estimates in an autoregressive error components model can be obtained by use of lagged exogenous variables as appropriate instruments. This method can be adapted to the fixed effects specification (see Sevestre and Trognon [1992], for example). IV estimates of δ, $\underline{\alpha}$ and $\underline{\beta}$ would be obtained by utilising the transformed model of:

(3) $$Z'\underline{y} = Z'\widetilde{X}_D \underline{\gamma}_D + Z'\underline{u},$$

where: $\widetilde{X}_D = (\underline{y}_{-1} \vdots X \vdots D)$,

$$\underline{\gamma}_D = \left(\delta \vdots \underline{\beta}' \vdots \underline{\alpha}'\right)'$$

and: $Z = (X_{-1} \vdots X \vdots D)$.

Once these concatenations have been made, the *B-N* estimator is obtained by applying GLS to (3) using $\sigma_u^2 Z\left(Z' Z\right)^{-1} Z'$ as the variance of the transformed disturbance term, $Z' \underline{u}$.

The remaining estimators for the fixed effects specification consider the model in terms of first differences:

(4) $$\Delta \underline{y} = \Delta \widetilde{X} \underline{\gamma} + \Delta \underline{u},$$

where: $\Delta y_{it} = y_{it} - y_{it-1}$ (and so on)

$$\Delta \widetilde{X} = (\Delta \underline{y}_{-1} \vdots \Delta X)$$

and: $\underline{\gamma} = \left(\delta \vdots \underline{\beta}'\right)'$.

This procedure is favoured as more orthogonality conditions implied by the usual assumptions *i.e.*, HF1-HF4, are utilised (see below) and, by assumption HF4, this operation removes the individual effects. However, first differencing does create problems of its own. Firstly, the now transformed model still cannot be consistently estimated by OLS, as the lagged endogenous variable $\Delta \underline{y}_{-1}$ is correlated with the model's disturbance vector $\Delta \underline{u}$. Secondly, if the original disturbances u_{it}, are "well-behaved" (*i.e.*, maintaining HF2 and HF3), the transformed ones Δu_{it} will follow a first order moving average (MA[1]) process (see below).

The Anderson-Hsiao (AH) and Arellano (AR) Estimators

Anderson and Hsiao (1982) suggest both y_{it-2} and Δy_{it-2} as an appropriate instrument for Δy_{it-1} in IV estimation of (4). However, in some instances the second latter yields inefficient estimates and therefore y_{it-2} is a more appropriate instrument (Arellano [1988]). Moreover, as $\Delta y_{it-2} = y_{it-2} - y_{it-3}$ use of it as an instrument necessitates removal of an additional time period for estimation purposes (as opposed to using y_{it-2}). This may mean curtailment of an already "short" time-series. Therefore the instruments are defined as:

(5) $$Z_{it}^{AR} = (y_{it-2} \vdots \underline{x}'_{it} - \underline{x}'_{it-1}) \quad \text{and} \quad Z_{it}^{AH} = (y_{it-2} - y_{it-3} \vdots \underline{x}'_{it} - \underline{x}'_{it-1}).$$

The simple *AH* and *AR* estimators are obtained by estimating the transformed model:

(6) $$Z'\Delta \underline{y} = Z'\Delta\widetilde{X}\underline{\gamma} + Z'\Delta\underline{u},$$

by OLS. If the number of instruments is the same as the number of explanatory variables, the resulting estimator will, in general, have no finite moments (Kinal [1980]). Thus perhaps the small sample behaviour of the *AH* and *AR* estimators can be improved by including additional instruments, ΔX_{-1} for example. The augmented instrument sets become:

(7) $$Z_{it}^{AR+} = (y_{it-2} \vdots \underline{x}'_{it-1} - \underline{x}'_{it-2} \vdots \underline{x}'_{it} - \underline{x}'_{it-1}) \text{ and } Z_{it}^{AH+} = (y_{it-2} - y_{it-3} \vdots \underline{x}'_{it-1} - \underline{x}'_{it-2} \vdots \underline{x}'_{it} - \underline{x}'_{it-1}).$$

Given the excess number of instruments over exogenous regressors, (6) can now be estimated by GLS using $\left(Z'\Omega_\Delta Z\right)$ as the variance of its transformed disturbance terms. Note that due to the first differencing operation, $\mathrm{Var}(\Delta\underline{u}) = \sigma_u^2\Omega_\Delta$:

(8) $$\Omega_\Delta = I_N \otimes \Sigma_\Delta = I_N \otimes \begin{pmatrix} 2 & -1 & 0 & \cdots & 0 \\ -1 & 2 & -1 & \ddots & \vdots \\ 0 & -1 & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & -1 \\ 0 & \cdots & 0 & -1 & 2 \end{pmatrix}.$$

The Sevestre-Trognon ($ST^{(b)}$ and $ST^{(c)}$) Estimators

Sevestre and Trognon (1992) again suggest using (5), but before applying instruments the model is turned into a spherical one by the usual GLS transformation of pre-multiplication of $\Omega_\Delta^{-1/2}$. Although this will result in a spherical model, there will still be correlations between the lagged endogenous variable and the transformed disturbance terms. Moreover, as pre-multiplication of (5) leads to disturbances that are linear combinations of the u_{it}'s, the only valid instruments for $\Omega_\Delta^{-1/2}\Delta\underline{y}_{-1}$ will be current and lagged values of both $\Omega_\Delta^{-1/2}\Delta X$ and ΔX.[1]

Although these estimators will be more efficient than those *using the same instruments* on the untransformed model (Sevestre and Trognon [1992] and White [1984]), a direct comparison with

[1] *i.e.*, methods b) and c) respectively (see p. 13 below).

the Anderson-Hsiao estimator, for example, is not straight forward as different instrument sets are used.

The Arellano-Bond One Step IV (AB) Estimator

If the time series is assumed to start at $t = 0$, the variable Δy_{it-1} will only be defined at $t = 2$. At $t = 2$ the only valid instrument for Δy_{it-1} is y_{i0} (recall that at $t = 2$, $\Delta u_{it} = u_{i2} - u_{i1}$, which is therefore independent of y_{i0}). However, at $t = 3$, the valid set of instruments for Δy_{it-1} is now expanded to include y_{i1}. This triangular expansion continues for successive time periods, defining the complete set of instruments at $t = 4$ as:

$$(9) \qquad Z_i^{AB} = (Z_i^{\bullet} \vdots \Delta X_i),$$

$$= \begin{pmatrix} y_{i0} & & & & & 0 & \underline{x}'_{i2} - \underline{x}'_{i1} \\ & y_{i0} & y_{i1} & & & 0 & \underline{x}'_{i3} - \underline{x}'_{i2} \\ 0 & & & y_{i0} & y_{i1} & y_{i2} & \underline{x}'_{i4} - \underline{x}'_{i3} \end{pmatrix}.$$

Moreover, assuming that the x's are strictly exogenous (HF1), they are all valid instruments for each time equation, and the instrument set becomes augmented:

$$(10) \qquad Z_i^{\bullet +} = \begin{pmatrix} y_{i0} & \underline{x}'_{i0} \cdots \underline{x}'_{iT} & & & & & & & 0 \\ & & y_{i0} & y_{i1} & \underline{x}'_{i0} \cdots \underline{x}'_{iT} & & & & \\ 0 & & & & & y_{i0} & y_{i1} & y_{i2} & \underline{x}'_{i0} \cdots \underline{x}'_{iT} \end{pmatrix},$$

$$Z_i^{AB^+} = (Z_i^{\bullet +} \vdots \Delta X_i).$$

Stacking the instrument matrices for each individual $Z = (Z_i' \vdots, \ldots, \vdots Z_N')'$, the Arellano and Bond (1991) estimators are again obtained by applying a GLS type estimator to the transformed model of (6).

The AB estimator is the most semi-asymptotically efficient of all IV estimators using lagged values of the dependent variable as instruments (Sevestre and Trognon [1992]), although more efficient GMM estimators can be derived (see below). However, computationally both of the AB estimators may prove problematic due to: the size of the instrument matrix (especially as T increases); the loss of two time periods for estimation; and difficulty in coding matrices such as (9) and (10) in standard econometric software packages.

The Balestra-Nerlove First Difference Estimator ($BN^{(\Delta)}$).

The Balestra-Nerlove estimator can also be applied to the first difference model, where the instruments for $\Delta\underline{y}_{-1}$ are simply ΔX_{-1}. Thus the full instrument set is given by:

$$Z = (\Delta X_{-1} \vdots \Delta X), \tag{11}$$

which differs from the *ST* estimators described earlier by the fact that (6) is directly estimated by GLS.

Two Step Estimators IV Estimators of the First Differenced Model

Following White (1984) one can relax assumptions HF2 and HF3 and consistently estimate the matrix $Z'\Omega_{\Delta}Z$ required for many of the above estimators as:

$$\hat{\Omega}_{\Delta Z} = \left(1/N \sum_{i=1}^{N} Z_i' \Delta\underline{\hat{u}}_i \Delta\underline{\hat{u}}'_i Z\right), \tag{12}$$

as, assuming independent individuals, the covariance matrix Ω_{Δ} is block diagonal. However, although $\mathrm{E}\left(\sum_{i=1}^{N} Z_i' \Delta\underline{\hat{u}}_i \Delta\underline{\hat{u}}'_i Z\right) = \mathrm{E}(Z'\Delta\underline{u}\,\Delta\underline{u}'Z)$, these two estimators of $Z'\Omega_{\Delta}Z$ will differ numerically.[2] For most of the estimators the difference in the resulting parameter estimates was very small (in the order of 10^{-3}). It was noted that using the latter gave numerically identical estimates between the one and two-step variants of an estimator (see Section 4), therefore in subsequent random effects experiments the former was used.

[2] Consider an *AB* estimator of the simple AR(1) model with no exogenous variables with two individuals and $t = 0,\ldots,2.$, $Z'\Delta\underline{u}\,\Delta\underline{u}'Z = (y_{10}\Delta u_{12})^2 + (y_{20}\Delta u_{22})^2 + y_{10}\Delta u_{12} y_{20}\Delta u_{22} + y_{20}\Delta u_{22} y_{10}\Delta u_{12} \neq \sum_i Z_i'\Delta\underline{u}_i\Delta\underline{u}'_i Z_i = (y_{10}\Delta u_{12})^2 + (y_{20}\Delta u_{22})^2$.

The residual vector $\Delta\underline{\hat{u}}_i$ is obtained from an initial consistent estimate of $\underline{\gamma}$, produced by setting Ω_Δ *as per* (5). The two-step and one-step variants of the various estimators will be asymptotically equivalent if the u_{it} are independent and homoscedastic (Arellano and Bond [1991]). Note also that for those estimators which required it, Ω_Δ was directly estimated as $\hat{\Omega}_\Delta = \left(1/N\sum_{i=1}^{N}\Delta\underline{\hat{u}}_i\Delta\underline{\hat{u}}'_i\right)$.

2.3 A Generalised Method of Moments (GMM) Estimator

Since Hansen's (1982) seminal paper, the method of GMM estimation has found much favour with applied econometricians (see Pagan and Vella [1989] for a useful summary). The estimation technique is very broad, nesting many well know other techniques (for example, IV estimation). The essence of GMM estimation involves explicit exploitation of theoretical moment conditions which, for estimation purposes, are replaced by their sample counterparts. Due to the recent work of Ahn and Schmidt (1995) and Crépon *et al.* (1996) for example, attention has turned to GMM estimation of dynamic error component panel models, however the technique can also be applied to fixed effects dynamic models.

Firstly, define the initial values as:

$$y_{i0} = \alpha_i + \underline{x}'_{i0}\underline{\beta} + u_{i0}. \tag{13}$$

Note that the parameter vector corresponding to the exogenous variables β is assumed to be identical across equations (1) and (13). This is a requirement of the need for consistent starting values of the full parameter vector for GMM estimation (see below). Equations (1) and (13), along with assumptions HF1-HF4 allow a number of implicit orthogonality conditions to be expressed as:

2a) $E(y_{i0} - \alpha_i - \underline{x}'_{i0}\underline{\beta}) = 0$.

2b) $E(y_{i0} - \alpha_i - \underline{x}'_{i0}\underline{\beta})^2 = \sigma_0^2$.

2c) $E(y_{i0} - \alpha_i - \underline{x}'_{i0}\underline{\beta})(y_{it} - \alpha_i - \delta y_{i,t-1} - \underline{x}'_{it}\underline{\beta}) = 0,\ \forall t = 1,\ldots,T$.

2d) $E(y_{it} - \alpha_i - \delta y_{i,t-1} - \underline{x}'_{it}\underline{\beta}) = 0,\ t = 1,\ldots,T$.

2e) $E(y_{it} - \alpha_i - \delta y_{i,t-1} - \underline{x}'_{it}\underline{\beta})(y_{is} - \alpha_i - \delta y_{i,s-1} - \underline{x}'_{is}\underline{\beta}) = 0,\ t,\ s = 1,\ldots,T,\ t \neq s$.

2f) $E(y_{it} - \alpha_i - \delta y_{i,t-1} - \underline{x}'_{it}\underline{\beta})(y_{it} - \alpha_i - \delta y_{i,t-1} - \underline{x}'_{it}\underline{\beta}) = \sigma_u^2,\ t = 1,\ldots,T$.

2g) $E(y_{i0} - \alpha_i - \underline{x}'_{i0}\underline{\beta})x_{it}^k = 0,\ \forall k, t = 0,\ldots,T$.

2h) $E(y_{it} - \alpha_i - \delta y_{i,t-1} - \underline{x}'_{it}\underline{\beta})x_{i0}^k = 0,\ \forall k, t = 1,\ldots,T$.

2i) $E(y_{it} - \alpha_i - \delta y_{i,t-1} - \underline{x}'_{it}\underline{\beta})x_{it}^k = 0,\ \forall k, t = 1,\ldots,T$.

2j) $E(y_{it} - \alpha_i - \delta y_{i,t-1} - \underline{x}'_{it}\underline{\beta})x_{is}^k = 0,\ \forall k, t \neq s, t = 1,\ldots,T, s = 1,\ldots,T$.

For (GMM) estimation purposes, these ten orthogonality conditions translate into identifying equations expressed in terms of observed variables and parameters (an example of such for t = 0,...,2, is given in Appendix I). Moreover, generally the IV estimators presented above can be shown to be based upon a subset of such identifying equations. The majority of the estimators for the fixed effects specification were actually proposed in the context of the random effects model (such that by first differencing the individual effects whether fixed or random are removed). Therefore, the relationships between such orthogonality conditions and IV estimators are presented only for the latter (see Appendix II).

Once the total number of identifying equations have been identified, the question arises as to how many of these one should use. Asymptotic efficiency arguments suggest all of them. However, Crépon *et al.* (1996) have shown that there is no efficiency loss in disregarding those equations in which any of the parameters of interest (α_i, $\underline{\beta}$ and δ) do not feature. Moreover, it is also possible to re-arrange some of the equations such that not all of the nuisance parameters need be estimated (see Appendices I and II). Defining the full parameter vector as $\underline{\gamma}_D^+$, which contains the parameter vector of interest $\underline{\gamma}_D = \left(\underline{\alpha}' \vdots \underline{\beta}' \vdots \delta\right)'$, as well as other nuisance parameters, the GMM estimator is given by the value that minimises the criterion function:

(14) $$\hat{\underline{\gamma}}_D^+ = \min_{\gamma_D} m_N(\underline{\gamma}_D^+)' \hat{W}^{-1} m_N(\underline{\gamma}_D^+),$$

where: $$m_N = N^{-1}\sum_i m_i(\underline{\gamma}_D^+)$$

$$W = \lim_{N\to\infty} \mathrm{cov}\left(N^{-1/2}\sum_i m_i\right) = \mathrm{cov}(m_N)$$

and: $\hat{W} = N^{-1}\sum_i m_i(\underline{\gamma}_D^+)\, m_i(\underline{\gamma}_D^+)'$, evaluated at an initial consistent parameter estimate of $\underline{\gamma}_D^+$.

A computational point not often addressed in the literature, is that as one increases the number of orthogonality conditions (especially as T increases) the columns of the matrix $\hat{W}$ are likely to become increasingly collinear, such that the matrix $\hat{W}$ is not invertible. In this case, one must disregard some conditions, although the consequences on efficiency are likely to be small, given the strong correlation between the conditions. That is, the maximum number of orthogonality conditions the GMM estimator can utilise is dependent upon the sample size in both N and T. Table 1a below lists those conditions which could be used in the samples considered in this study, whilst Table 1b summarises the IV-type estimators.

Table 1a: GMM-Type Estimators for the Dynamic Fixed Effects Model.

Sample Size T	Sample Size N	Conditions Used	Weighting Matrix[1]	Estimator Mnemonic
4	25	2a) - 2j)	I	GMM_F1
4	50	2a) - 2j)	I	GMM_F1
10	25	2a) - 2j)	I	GMM_F1
10	50	2a) - 2j)	I	GMM_F1
4	25	2a) - 2g) and 2i)[2]	$\hat{W}$	GMM_F2
4	50	2a) - 2j)	$\hat{W}$	GMM_F3
10	25	2a), 2b), 2d) and 2f)	$\hat{W}$	GMM_F4
10	50	2a), 2b), 2d), 2f), 2i) and 2g)[2]	$\hat{W}$	GMM_F5

Notes: [1] $\hat{W}$ is the estimated covariance matrix of the empirical moments. [2] 2i) for $k = 1$ only.

Table 1b: IV-Type Estimators for the Dynamic Fixed Effects Model, $N \to \infty$, finite T.

Method	Consistency	Model Estimated in: Levels	First Differences	Instrument(s) for $y_{i,t-1}$ or $\Delta y_{i,t-1}$
OLS	×	✓	✓	$y_{i,t-1}$ or $\Delta y_{i,t-1}$
Within	×	✓	×	$y_{i,t-1}$
$BN^{(L)}$	✓	✓	×	X_{-1}
AH	✓	×	✓	$y_{i,t-2}$
AR	✓	×	✓	$\Delta y_{i,t-2}$
AH^{+}	✓	×	✓	$y_{i,t-2}$, ΔX_{-1}
AR^{+}	✓	×	✓	$\Delta y_{i,t-2}$, ΔX_{-1}
$ST^{(b)}$	✓	×	✓	$\Omega_{\Delta}^{-1/2}\Delta X_{-1}$
$ST^{(c)}$	✓	×	✓	ΔX_{-1}
AB	✓	×	✓	y_{i0}; y_{i0}, y_{i1};…
AB^{+}	✓	×	✓	y_{i0}, $\underline{x}'_{i0}\dots\underline{x}'_{iT}$; y_{i0}, y_{i1}, $\underline{x}'_{i0}\dots\underline{x}'_{iT}$;…
$BN^{(\Delta)}$	✓	×	✓	ΔX_{-1}

3 The Random Effects Dynamic Panel Model

3.1 The Model

Under the random effects specification, the α_i terms of (1) are treated as independent random drawings from a particular distribution and the disturbance term becomes "composite", $v_{it} = \alpha_i + u_{it}$. As with the fixed effects specification, the traditional estimators (*Within* and GLS) of the static random effects panel model are semi-inconsistent in the dynamic setting (Sevestre and Trognon [1985]).

The semi-consistent estimators for the dynamic random effects model similarly rely on certain maintained hypotheses, not necessarily the same for all estimators:

HR1: The u_{it}'s are uncorrelated with y_{i0}, $\forall\, t,i$.

HR2: The u_{it}'s are uncorrelated with α_i, $\forall\ t,i$.

HR3: The u_{it}'s are uncorrelated, $\forall\ t,i$.

HR4: The u_{it}'s have zero mean and scalar variance σ_u^2.

HR5: The α_i's have zero mean and variance σ_α^2.

HR6: The x variables are non-stochastic and their individual means uncorrelated with either the u_{it}'s or the α_i's, $\forall\ t,i$.

HR7: The x-variables are non-stochastic and are uncorrelated with either the u_{it}'s or the α_i's, $\forall\ t,i$.

Again, HR6 and HR7 are violated by the inclusion of the lagged dependent variable. Also, the latter could be easily modified along the lines of Hausman and Taylor (1981) - see below - to allow a subset of the x-variables to be correlated with α_i. However, this simply affects the choice of valid instruments and unnecessarily complicates the following arguments. The assumptions concerning the equation's disturbances imply that variance-covariance matrix of the composite disturbance term will be:

(15)
$$\Omega_v = V(\underline{v}) = I_N \otimes E(\underline{v}_i \underline{v}_i') = I_N \otimes \Sigma_v,$$

$$\Sigma_v = \sigma_\alpha^2 J_T + \sigma_u^2 I_T = \sigma_v^2 \begin{pmatrix} 1 & \rho & \cdots & \rho \\ \rho & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \rho \\ \rho & \cdots & \rho & 1 \end{pmatrix},$$

where: ρ is the *intra-class correlation coefficient*, $\rho = \sigma_\alpha^2 / (\sigma_\alpha^2 + \sigma_u^2)$.

3.2 IV Estimators

As noted above, all of the IV estimators for the first differenced model are similarly applicable to both the random and fixed effects specification. Therefore it only remains to consider estimation of the random effects model in levels.

The Balestra-Nerlove (BN), Hausman-Taylor (HT), Amemiya-MaCurdy (AM) and Breusch-Mizon-Schmidt (BMS) Two-Step IV Estimators

Several estimators have been proposed in the context of a static random effects panel model, that first transform the model into a spherical one by the usual GLS pre-multiplication of $\Omega_v^{-1/2}$. However, generally they can be suitably adapted to dynamic models as well.[3]

Whatever the estimator, the usual *Feasible Generalised Least Squares* (FGLS) problem arises, as Ω_v is unknown. Indeed, Ω_v cannot be estimated using the residuals from an OLS regression of (1), as the parameter estimates will be semi-inconsistent, as would any estimate of Ω_v based upon the estimated residuals $\hat{\underline{v}}$. Asymptotically, *any* semi-consistent estimate of $\hat{\underline{v}}$ (and hence Ω_v) can be used, although differing methods are likely to cause small sample divergences. Once a consistent estimator of Ω_v has been obtained and an appropriate instrument set (Z) defined, one has a choice of three estimation procedures:

a) Pre-multiply (1) by Z' and estimate by GLS, using $\mathrm{Var}(Z'\underline{v}) = Z'\Omega_v Z$.
b) Transform (1) into a scalar model by pre-multiplying by $\Omega_v^{-1/2}$. Transform again by pre-multiplying the (transformed) model by $Z'\Omega_v^{-1/2}$. Finally, estimate this twice transformed model by GLS using $\mathrm{Var}(Z'\Omega_v^{-1}\underline{v}) = Z'\Omega_v^{-1}Z$.
c) As b) above, except in the second stage the untransformed instrument set Z is used (as opposed to $\Omega_v^{-1/2}Z$), and using $\mathrm{Var}(Z'\Omega_v^{-1/2}\underline{v}) = Z'Z$.

The Generalised Balestra-Nerlove (G-BNran) Two-Step IV Estimator

The Balestra and Nerlove (1966) estimator again uses current and (one period) lagged exogenous variables as an instrument set.

Hausman-Taylor (HT) IV Estimator

Hausman and Taylor (1981) partition the X-matrix such that $X = (X_1 \vdots X_2)$, where X_1 are uncorrelated with the individual effects, but X_2 is not. In a dynamic panel data setting, the lagged

[3] The exception is the Breusch-Mizon-Schmidt estimator (see below).

dependent variable is analogous to X_2 and under assumption HR7, the remaining explanatory variables (X in the notation of this paper) are analogous to Hausman-Taylor's X_1.

Following similar logic to that of the *G-BNran* estimator, the *HT* estimator also considers the means and deviations from the means of the original exogenous variables as valid instruments in addition to lagged values of X. As Breusch, Mizon and Schmidt (1989, p.696) show, this amounts to using the following instrument set:

(17) $$Z = (W_n X_{-1} \vdots\ \overline{B}_n X_{-1} \vdots X),$$

where: $\overline{B}_n = I_N \otimes J_T/T$, $W_n = I_N \otimes (I_T - J_T/T)$ and J_i is a matrix of ones of order i.

The Amemiya-MaCurdy (AM) IV Estimator

If the x's are strictly exogenous all past, present and future values become valid instruments. Thus the Amemiya and MaCurdy (1986) estimator further extends the instrument set to include X^*, defined as:

(18) $$X^* = \begin{pmatrix} x_{11}^{(1)} & x_{11}^{(2)} & \cdots & x_{11}^{(k)} & \cdots & x_{1T}^{(1)} & \cdots & x_{1T}^{(k)} \\ x_{21}^{(1)} & x_{21}^{(2)} & \cdots & x_{21}^{(k)} & \cdots & x_{2T}^{(1)} & \cdots & x_{2T}^{(k)} \\ \vdots & \vdots & & \vdots & & \vdots & & \vdots \\ x_{N1}^{(1)} & x_{N1}^{(2)} & \cdots & x_{N1}^{(k)} & \cdots & x_{NT}^{(1)} & \cdots & x_{NT}^{(k)} \end{pmatrix} \otimes \iota_T.$$

Note that each column of X^* contains values of x_{it} *for only one t*, as opposed to X which contains values of x_{it} for $t = 1,...T$. In effect X is being used $(T + 1)$ times, T times as X^* and once as $W_n X$. In the case of the dynamic model however, X^*_{-1} is required which is defined as:

(19) $$X^*_{-1} = \begin{pmatrix} x_{10}^{(1)} & x_{10}^{(2)} & \cdots & x_{10}^{(k)} & \cdots & x_{1T}^{(1)} & \cdots & x_{1T}^{(k)} \\ x_{20}^{(1)} & x_{20}^{(2)} & \cdots & x_{20}^{(k)} & \cdots & x_{2T}^{(1)} & \cdots & x_{2T}^{(k)} \\ \vdots & \vdots & & \vdots & & \vdots & & \vdots \\ x_{N0}^{(1)} & x_{N0}^{(2)} & \cdots & x_{N0}^{(k)} & \cdots & x_{NT}^{(1)} & \cdots & x_{NT}^{(k)} \end{pmatrix} \otimes \iota_T.$$

The full *AM* instrument set for the dynamic model is therefore:

(20) $$Z = (W_n X_{-1} \vdots X^*_{-1} \vdots X).$$

The *AM* estimator, if consistent, is at least as efficient as the *HT* estimator (Amemiya and MaCurdy [1986], pp.871-872).

Breusch-Mizon-Schmidt (BMS) IV Estimator

The *BMS* estimator (Breusch, Mizon and Schmidt [1989]) again extends the instrument set. In terms of a dynamic setting, this amounts to including $(W_n \underline{y}_{-1})^*$, similarly defined as X^* in (19). However, these additional instruments are not valid in the case of a dynamic model, as this is the source of inconsistency of the *Within* estimator. Therefore, in this instance the *BMS* and *AM* estimators are identical.

The Wansbeek-Bekker (WB) IV Estimator

Although consistent for finite T and $N \rightarrow \infty$, the proposed IV estimators will still be biased for finite N. In addition to small sample bias, estimators may also be preferred in terms of semi-asymptotic efficiency, which is the approach adopted by Wansbeek and Bekker (1993).

The *WB* approach extends that of Anderson and Hsiao (1982), such that now both lags and leads (and linear combinations of these) of the dependent variable are included in the instrument set. That is, by defining the variable y from period $t = 1$ to $t = T$, the *WB* estimator considers linear functions of $\underline{y}_+$ as instruments, where $\underline{y}_+$ is the stacked vector of observations defined from $t = 0$ to $t = T$ for each individual. The linear functions are defined by the $(T+1) \times T$ matrix A_i, which yields $A'\underline{y}_+$ as the full instrument set (where $A = I_N \otimes A_i$). Restrictions are imposed on A such that:

(22) $$A\underline{\iota}_T = 0, \qquad \text{and} \qquad \mathrm{E}(\underline{y}'_+ A\underline{u}) = \mathrm{tr}A\mathrm{E}(\underline{u}\underline{y}'_+) = 0,$$

which respectively ensure elimination of the individual effects and consistency of the estimator.

Wansbeek and Bekker (1993) show that these conditions for A define its structure such that its rows sum to zero, as do each of its lowest T quasi-diagonal elements (in particular, the lower left element is zero). Transformation of a variable by the matrix A will, in some instances have a "usual" interpretation, whilst in others it will not (see Wansbeek and Bekker [1993] for examples of A). The full WB instrument set is therefore defined as:

$$(23) \qquad Z = (A'\underline{y}_{+} \vdots X),$$

and using $\sigma_u^2(Z'Z)$ as the variance of $Z'\underline{u}$, the WB estimator is obtained by applying GLS to the transformed model:[4]

$$(24) \qquad Z'\underline{y} = Z'\tilde{X}\underline{\gamma} + Z'\underline{u}.$$

The estimator's semi-asymptotic variance will be given by:

$$(25) \qquad \sigma_u^2\left(\operatorname*{plim}_{N\to\infty}\frac{1}{N}(\tilde{X}'P_Z\tilde{X})^{-1}\right),$$

$$\text{where: } P_Z = Z(Z'Z)^{-1}Z',$$

which from (23) is a function of A. The optimal choice of A is that which minimises (25), such that A conforms to its appropriate restrictions . However, A is unspecified apart from these restrictions. The form of A for the WB estimator of the simple AR(1) model can be found by constrained optimisation (Wansbeek and Bekker [1993]). However, when the model additionally contains exogenous variables, numerical methods must be used as the variance of the estimator is a matrix not a scalar.

If one is only interested in the variance of the parameter vector (and not covariances of particular elements of it), the optimal WB estimator can be obtained by constrained optimisation, where A is that which minimises the trace of (25), treating σ_u^2 as a constant, subject to the restrictions of (22). Note that the list of valid instruments can be expanded to include not only $A'\underline{y}_{+}$, but also $A'X_{+}$ for example (WB and WB^{+}, respectively), such that:

[4] Note that this expression for the variance of $Z'\underline{u}$ is only an approximation, differing from the true variance to the extent that $E(\underline{y}'_{+}\underline{u}) \neq 0$, and this cross correlation is not taken into account.

(26) $$Z^{+} = (A'\underline{y}_{+} \vdots A'X_{+} \vdots X).$$

These estimators can also be adapted to the model where the assumption of a scalar covariance matrix of the disturbance terms u_{it} is relaxed. The corresponding estimators are still obtained by applying GLS to (24), but where the variance of $Z'\underline{u}$ is now $(Z'\Omega_u Z)$, where Ω_u is unspecified. The variance of the unrestricted *WB* estimator is:

(27) $$\operatorname*{plim}_{N\to\infty} \frac{1}{N}(\widetilde{X}' P_{Z\Omega} \widetilde{X})^{-1},$$

$$\text{where: } P_{Z\Omega} = Z(Z'\Omega_u Z)^{-1} Z',$$

and where once more the inverse term of $P_{Z\Omega}$ is estimated from initial preliminary estimates of $\underline{u}_i$ along the lines of equation (12). The unrestricted *WB* estimator is again obtained by minimising the trace of (27) with respect to *A* subject to the latter conforming to its necessary restrictions. The unrestricted WB^{+} is as above but where $P_{Z\Omega}$ is replaced by $P^{+}_{Z\Omega}$ whose definition is obvious.

The Arellano and Bover (ABov) IV Estimator

As with the *WB* estimator, the Arellano and Bover (1993) estimator first involves transforming the system of *T* equations. The nonsingular transformation is now given by:

(28) $$H_i = \begin{bmatrix} K \\ \underline{\iota}'_T / T \end{bmatrix},$$

where *K* is similar to Wansbeek and Bekker's *A*, in that $K\iota_T = 0$, where *K* is any $(T-1)\times T$ matrix of rank $(T-1)$. For example, *K* could be the first $(T-1)$ rows of the *Within* group operator, or the first difference operator. As the first $(T-1)$ transformed errors,

(29) $$\underline{v}_i^{+} = H_i \underline{v}_i = \begin{bmatrix} K\underline{v}_i \\ \bar{v}_i \end{bmatrix},$$

are free of α_i, all exogenous variables are valid instruments for these first $(T-1)$ equations. Moreover, assuming serial independence of the disturbance terms v_{it} along the lines of the Arellano-Bond estimator, the series $(y_{i0}, y_{i1},\ldots,y_{i,t-1})$ is also a valid instrument. However, this assumption requires more structure for *K*, which now additionally has to be upper triangular (Arellano and Bover [1993] p.16). This defines the matrix of valid instruments to be:

$$Z_i = \begin{bmatrix} (\underline{x}'_i, y_{i0}) & & & & 0 \\ & (\underline{x}'_i, y_{i0}, y_{i1}) & & & \\ & & \ddots & & \\ & & & (\underline{x}'_i, y_{i0},\ldots, y_{i,T-2}) & \\ 0 & & & & \underline{x}'_i \end{bmatrix}, \tag{30}$$

where $\underline{x}'_i = (\underline{x}'_{i0}\vdots,\ldots,\vdots\underline{x}'_{iT})$. Letting $Z = (Z'_1\vdots,\ldots,\vdots Z'_N)'$ and $H = I_N \otimes H_i$, the *A-Bov* estimator is obtained by estimating the transformed model:

$$Z'H\underline{y} = Z'H\widetilde{X}\underline{\gamma} + Z'H\underline{v}, \tag{31}$$

where: $\mathrm{Var}(Z'H\underline{v}) = Z'H\Omega_v H'Z$.

by GLS. Operationally, Arellano and Bover (*ibid* p.18), state that *provided K satisfies the above restrictions*, the *ABov* estimator is invariant to the choice of *K*.[5]

As with previous estimators, the covariance of the transformed system, $\Omega^+ = H\Omega_v H'$ must be estimated from residuals obtained from preliminary semi-consistent estimates. Following White (1984), Arellano and Bover (*ibid* p.6), suggest:

$$\hat{\Omega}^+_{Hv} = \frac{1}{N}\sum_{i=1}^{N} \underline{\hat{v}}^+_i \underline{\hat{v}}^{+\prime}_i, \tag{32}$$

where: $\underline{\hat{v}}^+_i$ are semi-consistent preliminary estimates of $H_i\underline{v}_i$.

[5] Using this result, the first difference operator was used in subsequent simulation experiments (see Section 4 below).

which is an *unrestricted* estimator of Ω^+. The *restricted* estimator under the usual assumptions of the error components model is:

(33) $$\hat{\Omega}_v^+ = H\hat{\Omega}_v H',$$

where: $\hat{\Omega}_v = I_N \otimes (\hat{\sigma}_\alpha^2 J_T + \hat{\sigma}_u^2 I_T)$

and: $\hat{\sigma}_\alpha^2$ and $\hat{\sigma}_u^2$ are consistent estimates of σ_α^2 and σ_u^2.

GMM Estimators

Important advances have been made recently in GMM estimation of dynamic panel data models, most notably by Ahn and Schmidt (1995) and Crépon *et al.* (1996). Again the assumptions HR1 - HR7 along with specification of initial values (*as per* equation [20]) allows one to write orthogonality conditions, expressed in terms of parameters and data:

3a) $\mathrm{E}(y_{i0} - \underline{x}'_{i0}\underline{\beta}) = 0$.

3b) $\mathrm{E}(y_{i0} - \underline{x}'_{i0}\underline{\beta})^2 = \sigma_0^2$.

3c) $\mathrm{E}(y_{i0} - \underline{x}'_{i0}\underline{\beta})(y_{it} - \delta y_{i,t-1} - \underline{x}'_{it}\underline{\beta}) = \sigma_\alpha^2,\ \forall t = 1,\ldots,T$.

3d) $\mathrm{E}(y_{it} - \delta y_{i,t-1} - \underline{x}'_{it}\underline{\beta}) = 0,\ t = 1,\ldots,T$.

3e) $\mathrm{E}(y_{it} - \delta y_{i,t-1} - \underline{x}'_{it}\underline{\beta})(y_{is} - \delta y_{i,s-1} - \underline{x}'_{is}\underline{\beta}) = \sigma_\alpha^2,\ \forall t \neq s$.

3f) $\mathrm{E}(y_{it} - \delta y_{i,t-1} - \underline{x}'_{it}\underline{\beta})^2 = \sigma_\alpha^2 + \sigma_u^2,\ t = 1,\ldots,T$.

3g) $\mathrm{E}y_{i0}(y_{it} - \delta y_{i,t-1} - \underline{x}'_{it}\underline{\beta}) = c,\ t = 1,\ldots,T$, where c is a constant.

3h) $\mathrm{E}(y_{i0} - \underline{x}'_{i0}\underline{\beta})x_{it}^k = 0,\ \forall k, t = 0,\ldots,T$.

3i) $\mathrm{E}(y_{it} - \delta y_{i,t-1} - \underline{x}'_{it}\underline{\beta})x_{i0}^k = 0,\ \forall k, t = 0,\ldots,T$.

3j) $\mathrm{E}(y_{it} - \delta y_{i,t-1} - \underline{x}'_{it}\underline{\beta})x_{it}^k = 0,\ \forall k, t = 1,\ldots,T$.

3k) $\mathrm{E}(y_{it} - \delta y_{i,t-1} - \underline{x}'_{it}\underline{\beta})x_{is}^k = 0,\ \forall k, t \neq s, t = 1,\ldots,T, s = 1,\ldots,T$.

IV estimators are based upon only those conditions which are linear in the parameters of interest [*c.f.* the *AB* estimator and condition 3g) above], whereas Ahn and Schmidt (1995) focus on the non-linear conditions [such as 3f)]. However, as Crépon *et al.* (1996) point out, such estimators ignore the first order moments [3a) and 3d)] at the cost of reduced efficiency. Again, an example of these identifying equations for $t = 0,...,2$ is given in Appendix II, along with how IV estimators are related to these. One could conceivably consider numerous GMM estimators based upon different subsets of conditions 3a) to 3k) above (Ahn and Schmidt [1993]). However, once more numerical considerations will to a large extent determine the number of such conditions one can use (see Table 2b).

The IV-type estimators for the random effects specification in levels, are summarised in Table 2a below (refer to Table 1a for the first differenced estimators). Again the number of orthogonality conditions that can be used for GMM-type estimators was dictated by both N and T. These GMM-type estimators are summarised in Table 2b below.

Table 2a: IV-Type Estimators for the Dynamic Random Effects Model, $N \to \infty$, finite T

Method	Consistency	Transformation into Spherical Model	Instrument(s) for $y_{i,t-1}$
OLS	✗	-	$y_{i,t-1}$
Within	✗	W_n	$y_{i,t-1}$
FGLS	✗	$\Omega_v^{-1/2}$	$y_{i,t-1}$
G-BNran$^{(a)}$	✓	-	X_{-1}
G-BNran$^{(b)}$	✓	$\Omega_v^{-1/2}$	$\Omega_v^{-1/2} X_{-1}$
G-BNran$^{(c)}$	✓	$\Omega_v^{-1/2}$	X_{-1}
HT$^{(a)}$	✓	-	$W_n X_{-1}$, $\overline{B}_n X_{-1}$
HT$^{(b)}$	✓	$\Omega_v^{-1/2}$	$\Omega_v^{-1/2} W_n X_{-1}$, $\Omega_v^{-1/2} \overline{B}_n X_{-1}$
HT$^{(c)}$	✓	$\Omega_v^{-1/2}$	$W_n X_{-1}$, $\overline{B}_n X_{-1}$
AM$^{(a)}$	✓	-	$W_n X_{-1}$, X_{-1}^{*}
AM$^{(b)}$	✓	$\Omega_v^{-1/2}$	$\Omega_v^{-1/2} W_n X_{-1}$, $\Omega_v^{-1/2} X_{-1}^{*}$
AM$^{(c)}$	✓	$\Omega_v^{-1/2}$	$W_n X_{-1}$, X_{-1}^{*}
WB	✓	A	$\underline{y}_{+}$
WB^{+}	✓	A	$\underline{y}_{+}$, X_{-1}
ABov	✓	H	y_{i0}, $\underline{x}'_{i0} \cdots \underline{x}'_{iT}$; y_{i0}, y_{i1}, $\underline{x}'_{i0} \cdots \underline{x}'_{iT}$; ...

Table 2b: GMM-Type Estimators for the Dynamic Random Effects Model.

Sample Size		Conditions Used	Weighting	Estimator
N	T		Matrix[1]	Mnemonic
4	25	3a) - 3k)	I	GMM_R1
4	50	3a) - 3k)	I	GMM_R1
10	25	3a) - 3k)	I	GMM_R1
10	50	3a) - 3k)	I	GMM_R1
4	25	3a) - 3g), 3j)	$\hat{W}$	GMM_R2
4	50	3a) - 3k)[2]	$\hat{W}$	GMM_R3
10	25	3a), 3b), 3d) and 3e)	$\hat{W}$	GMM_R4

Notes:[1] $\hat{W}$ is the estimated covariance matrix of the empirical moments. [2] for $k = 1$ only.

4. The Simulation Experiments

The data for the simulation experiments was generated in the following manner:

(34)
$$y_{it} = \alpha_i + \delta y_{it-1} + x_{it}^{(1)}\beta_1 + x_{it}^{(2)}\beta_2 + u_{it},$$
$$y_{i0} = \alpha_i + \underline{x}'_{i0}\underline{\beta} + u_{i0}$$
$$\text{where: } u_{it} \sim \mathrm{N}(0,1),$$
$$x_{it}^{(k)} = x_{i,t-1}^{(k)} + \mathrm{uniform}(-0.5, 0.5),\ k = 1,2$$
$$\text{and: } \delta = \beta_1 = \beta_2 = 0.5.$$

The individual effects were generated as $\alpha_i = 1,\ldots,N$ and $\alpha_i \sim \mathrm{N}(0,1)$ for the fixed and random effects specifications respectively. Sample sizes of $T = 4$, 10 and $N = 25$, 50 were chosen. Finally, due to computation time, the number of Monte Carlo repetitions was limited to 100. In each case, analysis is focussed upon the estimation of δ. Computing was undertaken in GAUSS, using the Constrained Optimisor.[6] The results of the experiments for the fixed effects specification can be

[6] Code is available from the authors on request. $T = 4$ is the smallest sample size that can accommodate all of the estimators. The results for β_1 and β_2 are available on request from the authors. Due to time constraints, the largest sample size for the random effects simulations ($N = 50$ and $T = 10$) was not run. It was estimated that such an experiment would take in excess of two months on a Pentium 120 personal computer.

found in Tables A2 to A5 and those for the random effects in Tables A6 to A8, in Appendices III and IV respectively. Figures 1 to 7 in Appendix V, present the same results in graphics.

4.1 The Fixed Effects Results

In the smallest sample size ($N = 25$, $T = 4$) one can immediately disregard the simple OLS estimator in terms of excessive bias and range of bias (Figure 1). The *AB* estimator that uses the expanded instrument set AB^+ appears to suffer heavily from the resulting small sample bias, especially due to the range of bias. In terms of the remaining semi-consistent estimators, there appears to be very little between the variants of the *AR* and simple *AB* estimators. However, given their relatively small bias and stability, one may be tempted to use either the *Within* estimator or the simple OLS estimator of the first differenced model. Finally, the GMM-type estimators, which respectively use all of the conditions and the identity matrix as a weighting matrix (GMM_F1) and a subset of these conditions with their empirical covariance matrix as a weighting matrix (GMM_F2), have a very similar performance. The latter may be preferred in terms of a slightly smaller bias, and range of bias, although both are similarly dominated by the *AR* and *AB* estimators.

When T is held constant and N is increased to fifty, a similar pattern emerges. That is, the *AR* and simple *AB* estimators dominate, followed closely by the GMM-type estimators. However, for the same reasons as before, one still may be tempted to use either of the (inconsistent) *Within* or OLS (Δ) estimators. Other points of interest are firstly that the AB^+ estimators can be disregarded due to excessive instability. Secondly, although the increase in the number of individuals allows the GMM estimator using the empirical covariance matrix (GMM_F3) to use all of the conditions (see Table 1a), this does not appear to improve its performance relative to the one that uses the identity matrix (GMM_F1). Both of these estimators are still marginally inferior to the *AR* and *AB* ones.

Increasing the number of time periods to ten (with $N = 25$) has several noticeable consequences. Firstly, it quite severely worsens the performance of the AB^+ estimators and also, somewhat surprisingly, that of the OLS (Δ) estimator (Figure 3). Those estimators that performed well in the previous samples continue to perform the best. The rankings do however change, with the *AB* estimators now having marginally smaller bias than the *AR* estimators. The *Within* estimator continues to perform well, on a par with the *AB* estimators. The GMM-type estimators again closely follow these estimators, and although increasing the number of time periods severely restricts the number of orthogonality conditions that the GMM_F4 estimator can use (Table 1a), this estimator performs well relative to its identity matrix counterpart.

A similar pattern is found in the largest sample size, with the simple *AB* and *AR* estimators performing the best. However, the *Within* estimator remains on a par with the former. The GMM type estimators continue to have reasonable performance, although they are now surpassed by the *AH* estimators.

In summary, the findings illustrated the known biases of the OLS and *Within* estimators (although the OLS estimator for the first differenced model performs relatively well in the smallest sample), and the relatively small variance of the *Within* estimator (Kiviet [1994]). The increased efficiency of the *AR* over the *AH* estimators was also evident (Arellano [1988]). However, somewhat surprising was the good performance of the *Within* estimator, especially in light of previous results (see Nickell [1981] and Maddala [1991]).[7] It was also found that there is effectively no difference between those variants of an estimator that use an estimated covariance matrix and those which use the theoretical one. Similarly, somewhat surprisingly, adding further instruments to the original *AR* and *AH* estimators has little effect on their performance.

Thus, for preferred semi-consistent estimators, those which only use current lagged values of the exogenous variables (and transformations of these) *i.e.,* the *ST* and *BN* estimators, are dominated by those which, in some form, use previous values of the dependent variable. In the smaller samples, the simple *AR* estimator appears a good candidate, closely followed by the simple *AB* and GMM-type estimators. In the larger samples, the *AB* estimator starts to dominate and the performance of the GMM-type estimators also relatively improves (that is, although additional parameters require estimation as *N* increase, better estimation of the theoretical moments is afforded by a larger cross-sectional component of the sample).

For ease of computation, however, one would have to choose the simple *AR* estimator, as coding of the *AB* instrument matrix may prove difficult in standard software. Even easier to estimate would

[7] This is not strictly true as the simulated (absolute) biases were estimated as: 0.022, 0.008, 0.014 and 0.0038 (Tables A4 to A7). The exact asymptotic biases can be calculated using Nickell's (1981) equation (25). These were respectively calculated as: 0.023, 0.007, 0.010 and 0.0025.

be the simple *Within* and OLS (Δ) estimators, which do tend to perform rather well, especially in small samples. Moreover, following Nickell (1981) one can calculate the likely biases of these estimators quite accurately. Finally, if an optimising package is available, the GMM-type estimators may well prove attractive, especially in larger samples, and especially if one has doubts about any of the underlying assumptions concerning the true data generating process.

4.2 The Random Effects Results

Estimation of the random effects specification appears to be much more troublesome. For example in the smallest sample size, one would be reticent to use any of the semi-consistent estimators apart from the GMM-type, WB^{+} and *AM* estimators, and probably in that order (Figure 5). There is however, effectively no difference between any of the *AM* variants. Once more one may be tempted here to use either of the inconsistent simple OLS or FGLS estimators. Although these estimators are dominated in terms of absolute bias by the GMM-type estimators, they do tend to be more stable (Figure 5). Note that there is effectively no difference between the GMM-type estimators, even though several conditions have to be dropped in the GMM_R2 estimator (Table 2b).

Increasing N to fifty (Figure 6) allows the GMM-estimators to more accurately estimate the empirical moments, and their performance increases accordingly. Indeed, in this sample size (T = 4, N = 50), the GMM estimators clearly dominate. Again, even though the increase in N allows the estimated GMM-type estimator to use virtually all of the conditions (Table 2b), there is very little difference between the two GMM-type variants. Of the other estimators, the WB^{+} estimator again performs well, only being dominated by the GMM-type estimators. Also, although the biases of the *HT*, *AR* and *BN* estimators, for example, are slightly smaller than those of the *AM* ones, the latter may be preferred in terms of a smaller range.

Increasing the number of time periods to ten significantly improves the performance of most of the estimators (Figure 7). Indeed, in line with previous studies (see for example, Arellano and Bond [1991] and Kiviet [1994]) variants of the *AR* and *AB* estimators for example, could now be considered possible candidates. However, the best performing estimators are quite easily the *WB* and WB^{+} estimators (the latter marginally more so), with the smallest biases, low MSE's of biases and small ranges of such. The GMM-type estimators again continue to perform well. However, the rise in T severely reduces the number of orthogonality conditions that GMM_R4 can utilise (Table

2b), which appears to adversely affect its performance, especially relative to GMM_R1. Other candidates in this sample are the *BN* (especially $BN^{(c)}$) estimators and the simple *AB* estimator, as most of the other estimators can be disregarded in terms of either excessive bias or range of such. Finally, although inconsistent, the *Within* estimator still performs well, being out performed only by the *WB* and WB^{+} estimators.

In summary, unlike the fixed effects estimators, the semi-consistent random effects estimators do not appear to be dominated by inconsistent traditional estimators (for example, only with $T = 10$, $N = 25$ does the *Within* estimator outperform most of the consistent estimators). With small *T*, for ease of computation the choice would appear to be between the *AM* and (inconsistent) OLS and FGLS estimators. However, if an optimising package is available, either the GMM-type or the WB^{+} estimators are good candidates. In samples with more time periods, the clear winner is the WB^{+} estimator, closely followed by its simpler counterpart, *WB*. GMM-type estimation could again be considered, especially if one has any strong feelings concerning the true data generating process. However, if no such optimising package is available, again the simple *AB* estimator appears appropriate.

Interesting anomalies arise between different variants of a particular estimator. Firstly, now there is a significant divergence in the simple *AR* and *AH* estimators and their expanded instrument set counterparts (the latter performing better in all samples for both *AR* and *AH*). Although using $1/N \cdot \sum_{i=1}^{N} Z_i' \, \underline{\hat{u}}_i^* \underline{\hat{u}}_i^{*\prime} \, Z_i$ (where $\underline{\hat{u}}_i^*$ represents estimates of the generalised transformed disturbance term) - see footnote 2 - no longer yields identical estimates, they are still very close for most of the estimators which have "estimated" counterparts. Finally, concerning the GMM-type estimators, unlike the fixed effects specification, the "estimated" version does not appear to uniformly dominate its identity matrix counterpart. Indeed, in "large" *T* samples the latter performs at least as well as the former, and often outperforms it both in terms of lower bias and efficiency.

Four estimators (*BN*, *HT*, *AM* and *ST*) use the three GLS-type IV variants *a*) to *c*).[8] In all samples, the *AM* appears invariant to the choice and, as noted above, clearly exhibits increased efficiency

[8] Noting that $BN^{(\Delta)}$ would be equivalent to $ST^{(a)}$.

over the other three estimators. However, apart from this estimator, no clear pattern emerges either for any particular estimator or indeed for any particular sample size.

An interesting question is how these results position themselves with respect to previous ones. Three studies have considered the *AB*, *AR* and *AH* estimators (Arellano and Bond [1991], Arellano and Bover [1993], who only consider the *AH* and *AR* estimators, and Kiviet [1994]). Arellano and Bover (1993) consider a purely AR(1) process with $T = 3$. The disturbance terms were (pseudo) random drawings from a standard normal distribution. The individual effects were similarly (but independently) normally distributed, although different variances were considered. Initial values were defined as:

$$(35) \quad y_{i0} = \alpha_i/(1-\delta) + (1-\delta^2)^{-1/2} u_{i0}.$$

For $\delta = 0.5$, $\mathrm{Var}(\alpha_i) = 1$ and $N = 100$, they find the value of the simulated *AH* estimator to be 0.8, with a (relatively unstable) standard deviation of just under three. The one and two step *AB* estimators had very similar performances which were much more stable (with standard deviations of 0.24) and accurate (simulated values 0.4762 and 0.4748, respectively) than that of *AH*. Although the biases of these estimators decreased significantly when N was increased to 500, were an increasing function of the true parameter δ and also varied with the variance of the individual effect, their rankings invariably remained unchanged.

Arellano and Bond (1991) consider a sample of $N = 100$ and $T = 7$, and extend the model to additionally include one strictly exogenous variable (generated with AR parameter equal to 0.8) with a corresponding coefficient of $\beta = 1$. All of the disturbance terms were independent random drawings from the standard normal distribution. Although initial values were set to zero, the first ten observations were then discarded. With $\delta = 0.5$, the simulated estimates of the one and two step *AB* estimators were found to be 0.4884 and 0.4920, respectively. The simple *AH* estimator had very poor performance, with a mean simulated value of -2.4753 and standard deviation of over 45. The simple *AR* estimator fared much better with a mean value of 0.5075 (and a standard deviation of 0.0821). Again biases generally increased with δ (values of 0.2, 0.5 and 0.8 were considered) and, although in the experiments with $\delta = 0.2$ and 0.8 biases were very close across estimators, the *AB* ones may be preferred in terms of smaller variances.

Kiviet (1994) presents the results of a much more extensive set of simulation experiments, where the model also includes one exogenous variable. Given the range of results, the reader is referred directly to the paper. However, the broad conclusions were that the *AH* estimators were very volatile, often having large variances and poor performances (in particular, they were susceptible to changes in the true value of δ). Of the estimators considered in this study, the simple one and two-step *AB* and *AR* estimators again appeared the most appropriate.

Comparing these results with those of the present experiment, we found much larger biases. These increased biases can be primarily attributed to the smaller cross-sectional sample sizes considered (*i.e.*, $N = 50$ compared to $N = 100$). The results of our experiments thus confirm those from these previous studies, highlighting the unreliability of the *AH* estimator (its performance varying markedly with the sample size and the true parameter value). Moreover, our results confirm the relative superiority of the *AB* and *AR* over the *AH* ones, although in most instances there appears to be very little between these two, (one may prefer the *AB* estimator in terms of smaller variances, but the *AR* in terms of ease of coding). However, our study considered many more estimators, and although the *AR* and *AB* estimators did preform relatively well, they were clearly surpassed by the GMM and *WB* estimators.

5. An Application to a Consumer Demand Schedule

Panel data is especially useful in estimating demand schedules, as by using individual unit data, one avoids the identification problem encountered at an aggregate level. In its simplest form, demand theory postulates that price and income will be important determinants in consumer purchases. Although in the short run consumer purchases of necessities will be invariant to income, price may be influential, possibly affecting the timing of purchases. However, a lagged dependent variable in such instances will probably capture the effects of any remaining, and possibly unobserved, variables. In this application the variable of interest is such a necessity, that of laundry detergent.

The data comes from The Roy Morgan Consumer Panel of Australia (CPA), which consists of an Australia wide sample of 2831 households, recording information on consumer purchases and personal demographics. Although many product fields are available, attention was restricted to laundry detergent as, by being a necessity, it was expected to be the most appropriately modelled using a lagged dependent variable. To avoid any problems caused by regional differences, the analysis was based solely upon the purchases made by the panel members in the Melbourne Metropolitan area. As laundry detergent is in general an infrequently purchased product, the monthly data for the financial year in question (1992/93) was aggregated into quarters so as to be comparable with the simulation experiment (*i.e.*, $T = 4$) but primarily to reduce the (troublesome) occurrence of zeros in the data set. Individuals who made no purchases in any one of the four quarters were removed from the sample (this reduced the sample to 113 individuals). Units are the quarterly number of kilograms of laundry detergent purchased and the average price per kilogram of such. Finally, although we have made no attempt in this paper to justify use of either fixed or random effects models, we only consider the latter in the empirical application primarily due to the large number of individuals in the data set.

The results for the estimators that explicitly consider the model in terms of first differences are presented in Table 3a below, and the remainder in Table 3b (the latter differs by virtue of the fact that a constant was included). *A priori* one would expect the constant (when estimated) and the lagged dependent variable to exert positive and price negative influences on current demand.

Tables 3a and 3b clearly illustrate, different methods can provide extremely different point estimates for the parameters of interest. On the basis of asymptotic standard errors, the lagged dependent variable appears to be insignificant for most of the estimators. The *ABov* estimator did yield a significant coefficient, however the estimator's poor performance in the simulations casts doubts as to the validity of this estimate. The AR^{+} estimates gave significantly *negative* results, which we would be reticent to accept, but the AB^{+} estimator did yield significant and sensible loyalty coefficient of just over 0.3. Of the other estimators, many can be disregarded in terms of perverse signs (most notably the GMM variants), however there appears to be some agreement between the *AB* and *WB* estimators, suggesting a loyalty coefficient in the range of between 0.2 - 0.3.

Table 3a Parameter Estimates of a Consumer Demand Schedule for Laundry Detergent (in First Differences). [1]

Method	Lagged Demand	Price	Method	Lagged Demand	Price
Δ OLS	**-0.491**	**-0.196**	$ST^{(c)}$ ($\Omega_{\Delta}^{-1/2}$)	-0.078	-0.274
	(0.064)	(0.093)		(0.575)	(0.152)
AH	**0.836**	**-0.458**	$ST^{(b)}$ ($\hat{\Omega}_{\Delta}$)	-0.116	-0.224
	(0.323)	(0.167)		(1.185)	(0.352)
AH^{+} (Ω_{Δ})	**-0.478**	**-0.215**	*AB* (Ω_{Δ})	0.298	**-0.342**
	(0.072)	(0.098)		(0.184)	(0.131)
AH^{+} ($\hat{\Omega}_{\Delta Z}$)	**-0.488**	-0.187	AB^{+} (Ω_{Δ})	0.308	-0.343
	(0.110)	(0.166)		(0.581)	(0.416)
AR	-0.052	**-0.324**	*AB*($\hat{\Omega}_{\Delta Z}$)	**0.322**	**-0.414**
	(0.139)	(0.095)		(0.158)	(0.117)
AR^{+} (Ω_{Δ})	-0.053	**-0.324**	AB^{+} ($\hat{\Omega}_{\Delta Z}$)	0.249	-0.195
	(0.198)	(0.135)		(0.480)	(0.361)
AR^{+} ($\hat{\Omega}_{\Delta Z}$)	-0.053	-0.324	$BN^{(\Delta)}$	0.040	-0.301
	(0.560)	(0.384)		(0.635)	(0.158)
$ST^{(b)}$ (Ω_{Δ})	-0.290	-0.169	$BN^{(\Delta)}$ ($\hat{\Omega}_{\Delta Z}$)	0.040	-0.301
	(0.516)	(0.145)		(1.703)	(0.441)

Notes:

[1] Quarterly data concerning laundry detergent in the Melbourne metropolitan district, taken from the Roy Morgan Research Consumer Panel of Australia (N = 113) in the financial year 1992/3. Dependent variable is the quarterly change in the number of kilograms purchased, explanatory variables are a lagged dependent variable and the quarterly change in the average price of purchases. Asymptotic standard errors are given in parenthesis, for inconsistent estimators standard errors are similarly inconsistent. Parameters in bold indicate statistical significance at 5% size two-sided test based upon estimated standard errors.

There is more consensus with the coefficient on price, with all the estimators agreeing on a negative influence on demand. Moreover, the range of estimates is tighter, ranging from about 0.2 to 0.6 (the GMM ($\hat{W}$) estimator yielding the largest parameter estimate). Again though several estimates appear to be statistically insignificant, most notably in the first difference model. The appropriate estimate would again appear to be around 0.2, which includes the (significant) *HT*, *AM*, *WB* and *ABov* estimators, although the significant first difference estimators generally suggest a slightly higher 0.3 - 0.5. Finally, of the estimators that consider the model in levels, all suggest a statistically significant habitual purchase of approximately 3.5 to 6 kilos of laundry detergent per quarter. Again the GMM ($\hat{W}$) estimator yields an estimate (of over 10) that appears "too large".

Table 3b Parameter Estimates of a Consumer Demand Schedule for Laundry Detergent (in Levels).[1]

Method	Lagged Demand	Price	Constant
OLS	**0.647** (0.045)	-0.158 (0.084)	**1.929** (0.421)
Within	**-0.349** (0.051)	**-0.155** (0.068)	-
FGLS	**0.733** (0.036)	-0.108 (0.072)	**1.353** (0.355)
$BN^{(a)}$	-0.165 (0.559)	-0.207 (0.130)	**5.710** (2.396)
$BN^{(b)}$	-0.226 (0.413)	-0.190 (0.107)	**5.916** (1.647)
$BN^{(c)}$	-0.010 (0.439)	-0.266 (0.128)	**5.772** (1.851)
$HT^{(a)}$	-0.223 (0.413)	-0.204 (0.128)	**5.956** (1.801)
$HT^{(b)}$	-0.209 (0.404)	-0.193 (0.106)	**5.850** (1.616)
$HT^{(c)}$	-0.064 (0.416)	**-0.252** (0.123)	**5.958** (1.784)
$AM^{(a)}$	-0.102 (0.352)	**-0.208** (0.095)	**5.439** (1.466)
$AM^{(b)}$	-0.102 (0.352)	**-0.208** (0.095)	**5.439** (1.466)
$AM^{(c)}$	0.005 (0.352)	**-0.230** (0.098)	**5.345** (1.459)
WB	0.208 (0.233)	**-0.185** (0.083)	**3.973** (0.405)
WB^{+}	0.281 (0.202)	**-0.180** (0.079)	**3.634** (0.373)
$WB\ (\hat{P}_{Z\Omega})$	0.226 (0.177)	-0.184 (0.098)	**3.893** (0.902)
$WB^{+}\ (\hat{P}_{Z\Omega})$	0.286 (0.165)	**-0.195** (0.089)	**3.586** (0.827)
$ABov\ (\hat{\Omega}^{+}_{Hv})$	**0.413** (0.132)	**-0.188** (0.047)	**3.172** (0.667)
$ABov\ (\hat{\Omega}^{+}_{v})$	-0.136 (0.316)	**-0.222** (0.097)	**5.571** (1.345)
$GMM\ (I)$	-0.008 (0.027)	-0.333 (0.199)	**5.579** (0.850)
$GMM\ (\hat{W})^{2}$	-0.007 (0.125)	**-0.634** (0.221)	**10.131** (1.951)

Notes:

[1] See Table 3a, except variables are now expressed as levels. [2] Conditions used: all except 3k).

6. Conclusions

In this paper we propose two new estimators for dynamic panel data models, and evaluated their small sample performance along with that of the numerous existing ones. In terms of the fixed effects specification, the new (GMM) estimator for the fixed effects specification performed well and it appears that it will be useful in many empirical applications. However, its small sample performance is marginally surpassed by that of the simple *AR* and *AB* estimators, which indeed may be favoured due to ease of computation. Somewhat surprisingly, those estimators which performed well in the fixed effects specification, did not do so in the random effects specification (especially in samples with few time periods). The GMM-type estimators appear appropriate for this latter specification, generally irrespective of the sample size. However, the other new estimator proposed, the *WB* estimator, also appears to have very desirable small sample properties. Indeed, in samples with more time periods, this estimator has the best small sample performance. Therefore we would also expect this new estimator to be appropriate for many empirical applications.

Finally, the importance of the choice of estimator was highlighted by an application to a consumer demand schedule for laundry detergent, we found that there was a vast difference not only in the sign and magnitude of parameter estimates, but also in the statistical significance of such as determined by estimated asymptotic standard errors. This emphases the importance of the proper choice of estimator for particular problems and data sets.

References

Ahn, S.C. and Schmidt, P. (1993); Efficient Estimation of Dynamic Panel Data Models under Alternative Sets of Assumptions, mimeo Michigan State University.

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

Anderson, T.W. and Hsiao, C. (1982); Formulation and Estimation of Dynamic Models Using Panel Data, *Journal of Econometrics*, Vol. 18, pp. 578-606.

Amemiya, T. and MaCurdy, T.E. (1986); Instrumental Estimation of an Error Components Model, *Econometrica*, Vol. 54, pp. 869-881.

Arellano, M. (1988); A Note on the Anderson-Hsiao Estimator for Panel Data, *mimeo* Institute of Economics, Oxford University.

Arellano, M. and Bond, S. (1991); Some Tests of Specification for Panel Data: Monte-Carlo Evidence and an Application to Employment Equations, *Review of Economic Studies*, Vol. 58, pp. 127-134.

Arellano, M. and Bover, O. (1993); Another Look at the Instrumental Variables Estimation of Error-Components Models *Journal of Econometrics*, Vol. 68, pp. 29-52.

Balestra, P. (1992); Fixed Effects Models and Fixed Coefficient Models, in (chpt 3) *The Econometrics of Panel Data*, Mátyás and Sevestre (eds.), 1992, Kluwer Academic Publishers, the Netherlands.

Balestra, P. and Nerlove, M. (1966); Pooling Cross-Section and Time-Series Data in the Estimation of a Dynamic Model, *Econometrica*, Vol. 34, pp. 585-612.

Breusch, T.S., Mizon, G.E. and Schmidt, P. (1989); Efficient Estimation Using Panel Data, *Econometrica*, Vol. 57, pp. 695-700.

Chamberlain, G. (1980); Analysis of Covariance with Qualitative Data, *Review of Economic Studies*, Vol. 47, pp. 225-238.

Crépon, B., Kramarz, F. and Trogon, A. (1996); Parameters of Interest, Nuisance Parameter and Orthogonality Conditions: An Application to Error Component Models, Discussion Paper no. 9335, Institut National de la Statistique et des Etudes Economiques.

Hansen, L. (1982); Large Sample Properties of Generalised Method of Moments Estimators, *Econometrica*, Vol. 50, pp. 1029-1054.

Hausman, J.A. (1978); Specification Tests in Econometrics, *Econometrica*, Vol. 46, pp. 1251-1273.

Hausman, J.A. and Taylor, W.E. (1981); Panel Data and Unobservable Individual Effects, *Econometrica*, Vol. 49, pp. 1377-1398.

Hsiao, C. (1985); Benefits and Limitations of Using Panel Data, *Econometric Reviews*, Vol. 4, pp. 121-174.

Hsiao, C. (1986); *Analysis of Panel Data*, Cambridge University Press, Cambridge.

Kinal, T.W. (1980); The Existence of *k*-class Estimators, *Econometrica*, Vol. 48, pp. 241-249.

Kiviet, J.F (1994); On Bias, Inconsistency and Efficiency of Various Estimators in Dynamic Panel Data Models, *Journal of Econometrics*, Vol. 68, pp. 53-78.

Maddala, G.,S. (1971); The Use of Variance Components Models in Pooling Cross-Section and Time Series Data, *Econometrica*, Vol. 39, pp. 341-358.

Mundlak, Y. (1978a); On the Pooling of Time Series and Cross Section Data, *Econometrica*, Vol. 46, pp. 69-85.

Mundlak, Y. (1978b); Models with Varying Coefficients: Integration and Extension, *Annales de l'INSEE*, 30-1, pp. 483-509.

Nerlove, M. (1967); Experimental Evidence on the Estimation of Dynamic Economic Relationships from a Time Series of Cross-Sections, *Economic Studies Quarterly*, Vol. 18, pp. 42-74.

Nerlove, M. (1971); Further Evidence on the Estimation of Dynamic Economic Relationships from a Time Series of Cross-Sections, *Econometrica*, Vol. 39, pp. 359-387.

Nickell, S. (1981); Biases in Models With Fixed Effects, *Econometrica*, Vol. 49, pp. 1417-1426.

Pagan, A. and Vella, F. (1989); Diagnostic Tests for Models Based Upon Individual Unit Data: A Survey, *Journal of Applied Econometrics*, Vol. 4, Supplement, pp. S29-S59.

Sevestre, P. and Trognon, A. (1985); A Note on Autoregressive Error Component Models, *Journal of Econometrics*, Vol. 29, pp. 231-245.

Sevestre, P., and Trognon, A. (1992); Linear Dynamic Models, in *The Econometrics of Panel Data*, Mátyás and Sevestre (eds.), 1992, Kluwer Academic Publishers, the Netherlands.

Wansbeek, T. and Bekker, P. (1993); Efficient Instrumental Variables Estimators in a Dynamic Panel Data Model, mimeo.

White, H. (1984); *Asymptotic Theory for Econometricians*, Academic Press, Orlando.

Appendix I: Identifying Equations and IV/GMM Estimators for the Fixed Effects Model

Assuming that the time series runs from t = 0,...,2 the six orthogonality conditions 2a) to 2f), translate into nine identifying equations which can be used for estimation purposes:

AI.1) $\frac{1}{N}\sum_{i=1}^{N} y_{i0} - \underline{x}'_{i0}\underline{\beta} - \alpha_i = 0$.

AI.2) $\frac{1}{N}\sum_{i=1}^{N} (y_{i0} - \underline{x}'_{i0}\underline{\beta} - \alpha_i)^2 = \sigma_0^2$.

AI.3) $\frac{1}{N}\sum_{i=1}^{N} (y_{i0} - \underline{x}'_{i0}\underline{\beta} - \alpha_i)(y_{i1} - \delta y_{i0} - \underline{x}'_{i1}\underline{\beta} - \alpha_i) = 0$.

AI.4) $\frac{1}{N}\sum_{i=1}^{N} (y_{i0} - \underline{x}'_{i0}\underline{\beta} - \alpha_i)(y_{i2} - \delta y_{i1} - \underline{x}'_{i2}\underline{\beta} - \alpha_i) = 0$.

AI.5) $\frac{1}{N}\sum_{i=1}^{N} y_{i1} - \delta y_{i0} - \underline{x}'_{i1}\underline{\beta} - \alpha_i = 0$.

AI.6) $\frac{1}{N}\sum_{i=1}^{N} y_{i2} - \delta y_{i1} - \underline{x}'_{i2}\underline{\beta} - \alpha_i = 0$.

AI.7) $\frac{1}{N}\sum_{i=1}^{N} (y_{i2} - \delta y_{i1} - \underline{x}'_{i2}\underline{\beta} - \alpha_i)(y_{i1} - \delta y_{i0} - \underline{x}'_{i1}\underline{\beta} - \alpha_i) = 0$.

AI.8) $\frac{1}{N}\sum_{i=1}^{N} (y_{i1} - \delta y_{i0} - \underline{x}'_{i1}\underline{\beta} - \alpha_i)^2 = \sigma_u^2$.

AI.9) $\frac{1}{N}\frac{1}{N}\sum_{i=1}^{N} (y_{i2} - \delta y_{i1} - \underline{x}'_{i2}\underline{\beta} - \alpha_i)^2 = \sigma_u^2$

Including the equations implied by the orthogonality conditions 2g) to 2k), yield the further nine equations:

AI.10) $\frac{1}{N}\sum_{i=1}^{N} (y_{i0} - \underline{x}'_{i0}\underline{\beta} - \alpha_i)\underline{x}'_{i0} = 0$, $k = 1,...K$.

AI.11) $\frac{1}{N}\sum_{i=1}^{N} (y_{i0} - \underline{x}'_{i0}\underline{\beta} - \alpha_i)\underline{x}'_{i1} = 0$, $k = 1,...K$.

AI.12) $\frac{1}{N}\sum_{i=1}^{N} (y_{i0} - \underline{x}'_{i0}\underline{\beta} - \alpha_i)\underline{x}'_{i2} = 0$, $k = 1,...K$.

$$AI.13)\ \frac{1}{N}\sum_{i=1}^{N} (y_{i1} - \delta y_{i0} - \underline{x}'_{i1}\underline{\beta} - \alpha_i)\underline{x}'_{i0} = 0,\ k = 1,...K.$$

$$AI.14)\ \frac{1}{N}\sum_{i=1}^{N} (y_{i2} - \delta y_{i1} - \underline{x}'_{i2}\underline{\beta} - \alpha_i)\underline{x}'_{i0} = 0,\ k = 1,...K.$$

$$AI.15)\ \frac{1}{N}\sum_{i=1}^{N} (y_{i1} - \delta y_{i0} - \underline{x}'_{i1}\underline{\beta} - \alpha_i)\underline{x}'_{i1} = 0,\ k = 1,...K.$$

$$AI.16)\ \frac{1}{N}\sum_{i=1}^{N} (y_{i2} - \delta y_{i1} - \underline{x}'_{i2}\underline{\beta} - \alpha_i)\underline{x}'_{i2} = 0,\ k = 1,...K.$$

$$AI.17)\ \frac{1}{N}\sum_{i=1}^{N} (y_{i2} - \delta y_{i1} - \underline{x}'_{i2}\underline{\beta} - \alpha_i)\underline{x}'_{i1} = 0,\ k = 1,...K.$$

$$AI.18)\ \frac{1}{N}\sum_{i=1}^{N} (y_{i1} - \delta y_{i0} - \underline{x}'_{i1}\underline{\beta} - \alpha_i)\underline{x}'_{i2} = 0,\ k = 1,...K.$$

Appendix II: Identifying Equations and IV/GMM Estimators for the Random Effects Model

Assuming that the time series runs from t = 0,...,2 the six orthogonality conditions 3a) to 3g), translate into the following eleven identifying equations which can be used for estimation purposes:

$$AII.1)\ \frac{1}{N}\sum_{i=1}^{N} y_{i0} - \underline{x}'_{i0}\underline{\beta} = 0.$$

$$AII.2)\ \frac{1}{N}\sum_{i=1}^{N} (y_{i0} - \underline{x}'_{i0}\underline{\beta})^2 = \sigma_0^2.$$

$$AII.3)\ \frac{1}{N}\sum_{i=1}^{N} (y_{i0} - \underline{x}'_{i0}\underline{\beta}_i)(y_{i1} - \delta y_{i0} - \underline{x}'_{i1}\underline{\beta}) = \sigma_\alpha^2.$$

$$AII.4)\ \frac{1}{N}\sum_{i=1}^{N} (y_{i0} - \underline{x}'_{i0}\underline{\beta}_i)(y_{i2} - \delta y_{i1} - \underline{x}'_{i2}\underline{\beta}) = \sigma_\alpha^2.$$

$$AII.5)\ \frac{1}{N}\sum_{i=1}^{N} y_{i1} - \delta y_{i0} - \underline{x}'_{i1}\underline{\beta} = 0.$$

$$AII.6)\ \frac{1}{N}\sum_{i=1}^{N} y_{i2} - \delta y_{i1} - \underline{x}'_{i2}\underline{\beta} = 0.$$

$$AII.7)\ \frac{1}{N}\sum_{i=1}^{N} (y_{i2} - \delta y_{i1} - \underline{x}'_{i2}\underline{\beta})(y_{i1} - \delta y_{i0} - \underline{x}'_{i1}\underline{\beta}) = \sigma_\alpha^2.$$

AII.8) $\frac{1}{N}\sum_{i=1}^{N}(y_{i1}-\delta y_{i0}-\underline{x}'_{i1}\underline{\beta})^2=\sigma_\alpha^2+\sigma_u^2$.

AII.9) $\frac{1}{N}\sum_{i=1}^{N}(y_{i2}-\delta y_{i1}-\underline{x}'_{i2}\underline{\beta})^2=\sigma_\alpha^2+\sigma_u^2$.

AII.10) $\frac{1}{N}\sum_{i=1}^{N}y_{i0}(y_{i1}-\delta y_{i0}-\underline{x}'_{i1}\underline{\beta})=c$.

AII.11) $\frac{1}{N}\sum_{i=1}^{N}y_{i0}(y_{i2}-\delta y_{i1}-\underline{x}'_{i2}\underline{\beta})=c$.

Again, including additional conditions 3h) to 3k) augments the list by nine further equations:

AII.12) $\frac{1}{N}\sum_{i=1}^{N}(y_{i0}-\underline{x}'_{i0}\underline{\beta})\underline{x}'_{i0}=0$, $k=1,\ldots K$.

AII.13) $\frac{1}{N}\sum_{i=1}^{N}(y_{i0}-\underline{x}'_{i0}\underline{\beta})\underline{x}'_{i1}=0$, $k=1,\ldots K$.

AII.14) $\frac{1}{N}\sum_{i=1}^{N}(y_{i0}-\underline{x}'_{i0}\underline{\beta})\underline{x}'_{i2}=0$, $k=1,\ldots K$.

AII.15) $\frac{1}{N}\sum_{i=1}^{N}(y_{i1}-\delta y_{i0}-\underline{x}'_{i1}\underline{\beta})\underline{x}'_{i0}=0$, $k=1,\ldots K$.

AII.16) $\frac{1}{N}\sum_{i=1}^{N}(y_{i2}-\delta y_{i1}-\underline{x}'_{i2}\underline{\beta})\underline{x}'_{i0}=0$, $k=1,\ldots K$.

AII.17) $\frac{1}{N}\sum_{i=1}^{N}(y_{i1}-\delta y_{i0}-\underline{x}'_{i1}\underline{\beta})\underline{x}'_{i1}=0$, $k=1,\ldots K$.

AII.18) $\frac{1}{N}\sum_{i=1}^{N}(y_{i2}-\delta y_{i1}-\underline{x}'_{i2}\underline{\beta})\underline{x}'_{i2}=0$, $k=1,\ldots K$.

AII.19) $\frac{1}{N}\sum_{i=1}^{N}(y_{i2}-\delta y_{i1}-\underline{x}'_{i2}\underline{\beta})\underline{x}'_{i1}=0$, $k=1,\ldots K$.

AII.20) $\frac{1}{N}\sum_{i=1}^{N}(y_{i1}-\delta y_{i0}-\underline{x}'_{i1}\underline{\beta})\underline{x}'_{i2}=0$, $k=1,\ldots K$.

The relationship between these total number of orthogonality conditions and the various IV estimators is summarised in Tables A1a and Tables A1b below.

Table A1a: IV/GMM Estimators: Identifying Equations Used (First Difference Model)[1]

Method	Disturbance Terms Equations	Exogeneity of X Equations
AR	AII.2) - A.II4), AII.7) - AII.11)	AII.17) - AII.20)
AR^{+}	AII.2) - A.II4), AII.7) - AII.11)	AII.15) - AII.20)
$ST^{(a)}$	AII2) - AII4), AII.7) - AII.9)	AII.15) - AII.20)
$ST^{(b)}$	AII2) - AII4), AII.7) - AII.9)	AII.15) - AII.20)
AB	AII.2) - A.II4), AII.7) - AII.11)	AII.17) - AII.20)
AB^{+}	AII.2) - A.II4), AII.7) - AII.11)	AII.12) - AII.20)
$BN^{(\Delta)}$	AII2) - AII4), AII.7) - AII.9)	AII.15) - AII.20)

Notes: [1]*A-H* estimators are not appropriate as the time series is not long enough.

Table A1b: IV/GMM Estimators: Identifying Equations Used (Levels Model)

Method	Disturbance Terms Equations	Exogeneity of X Equations
$G\text{-}BNran^{(a)}$	AII2) - AII4), AII.7) - AII.9)	AII.16) - AII.19)
$G\text{-}BNran^{(b)}$	AII2) - AII4), AII.7) - AII.9)	AII.16) - AII.19)
$G\text{-}BNran^{(c)}$	AII2) - AII4), AII.7) - AII.9)	AII.16) - AII.19)
$HT^{(a)}$	AII2) - AII4), AII.7) - AII.9)	AII.13), AII.14), AII.17) - AII.20)
$HT^{(b)}$	AII2) - AII4), AII.7) - AII.9)	AII.13), AII.14), AII.17) - AII.20)
$HT^{(c)}$	AII2) - AII4), AII.7) - AII.9)	AII.13), AII.14), AII.17) - AII.20)
$AM^{(a)}$	AII2) - AII4), AII.7) - AII.9)	AII.12) - AII.20)
$AM^{(b)}$	AII2) - AII4), AII.7) - AII.9)	AII.12) - AII.20)
$AM^{(c)}$	AII2) - AII4), AII.7) - AII.9)	AII.12) - AII.20)
WB	AII2) - AII4), AII.7) - AII.9), AII.10), AII.11)	AII.17), AII.18)
WB^{+}	AII2) - AII4), AII.7) - AII.9), AII.10), AII.11)	AII.15), AII.17) - AII.19)
ABov	AII2) - AII4), AII.7) - AII.9), AII.10), AII.11)	AII.12) - AII.20)

Appendix III: Simulation Results for the Fixed Effects Estimators

Table A2: Estimators of δ in the Dynamic Fixed Effects Model, $T = 4$, $N = 25$.[1]

Method	Mean Bias[2]	Mean Squared Error[3]	Range of Bias[4]	
OLS	0.626961	0.393384	0.576235	0.657811
Within	0.02223	0.000766	6.77E-05	0.068714
$BN^{(L)}$	0.092606	0.012683	0.003159	0.254004
Δ OLS	0.027923	0.001264	0.000417	0.083383
AH	0.060932	0.006231	7.35E-05	0.195061
AH^+ $(\hat{\Omega}_{\Delta Z})$	0.060932	0.006231	7.35E-05	0.195061
AH^+ (Ω_Δ)	0.060837	0.006158	0.000944	0.187598
AR	0.025142	0.000954	0.00066	0.087974
AR^+ $(\hat{\Omega}_{\Delta Z})$	0.025142	0.000954	0.00066	0.087974
AR^+ (Ω_Δ)	0.025012	0.000958	0.000228	0.089501
$ST^{(b)}$ (Ω_Δ)	0.092606	0.012683	0.003159	0.254004
$ST^{(c)}$ $(\Omega_\Delta^{-1/2})$	0.123255	0.022971	0.005993	0.555902
$ST^{(b)}$ $(\hat{\Omega}_\Delta)$	0.092606	0.012683	0.003159	0.254004
AB (Ω_Δ)	0.022702	0.000756	2.95E-05	0.067831
AB^+ (Ω_Δ)	0.105704	0.133651	1.81E-04	3.29119
AB $(\hat{\Omega}_{\Delta Z})$	0.022702	0.000756	2.95E-05	0.067831
AB^+ $(\hat{\Omega}_{\Delta Z})$	0.105704	0.133651	1.81E-04	3.29119
$BN^{(\Delta)}$	0.145435	0.033843	0.00247	0.710408
GMM_F1	0.04596	0.004483	0.000189	0.173703
GMM_F2	0.037867	0.002988	0.000129	0.162064

<u>Notes:</u>

[1] Q = number of Monte Carlo Repetitions = 100. [2] $\frac{1}{Q}\sum_i \left|\hat{\delta}_i - \delta\right|$. [3] $\frac{1}{Q}\sum_i (\hat{\delta}_i - \delta)^2$.

[4] $\mathrm{Min}(\hat{\delta}_i)$ to $\mathrm{Max}(\hat{\delta}_i)$.

Table A3: Estimators of δ in the Dynamic Fixed Effects Model, $T = 4$, $N = 50$.[1]

Method	Mean Bias[2]	Mean Squared Error[3]	Range of Bias[4]	
OLS	0.634094	0.402151	0.600266	0.652773
Within	0.008109	9.72E-05	0.000239	0.02626
$BN^{(L)}$	0.050695	0.003905	0.001824	0.158318
Δ OLS	0.010092	0.00016	0.000237	0.02943
AH	0.022658	0.000715	0.000222	0.056046
$AH^{+}\ (\hat{\Omega}_{\Delta Z})$	0.022658	0.000715	0.000222	0.056046
$AH^{+}\ (\Omega_{\Delta})$	0.022568	0.000714	6.98E-05	0.056903
AR	0.008136	0.000102	0.000105	0.024882
$AR^{+}\ (\hat{\Omega}_{\Delta Z})$	0.008136	0.000102	0.000105	0.024882
$AR^{+}\ (\Omega_{\Delta})$	0.008149	0.000103	0.000592	0.025736
$ST^{(b)}\ (\Omega_{\Delta})$	0.050695	0.003905	0.001824	0.158318
$ST^{(c)}\ (\Omega_{\Delta}^{-1/2})$	0.068122	0.008047	0.001224	0.262963
$ST^{(b)}\ (\hat{\Omega}_{\Delta})$	0.050695	0.003905	0.001824	0.158318
$AB\ (\Omega_{\Delta})$	0.007594	8.56E-05	0.000535	0.02383
$AB^{+}\ (\Omega_{\Delta})$	0.03149	2.28E-02	6.13E-05	1.478822
$AB\ (\hat{\Omega}_{\Delta Z})$	0.007594	8.56E-05	0.000535	0.02383
$AB^{+}\ (\hat{\Omega}_{\Delta Z})$	0.03149	2.28E-02	6.13E-05	1.478822
$BN^{(\Delta)}$	0.113507	0.068733	0.001609	2.321198
GMM_F1	0.02663	0.001524	0.000227	0.119583
GMM_F3	0.018544	0.000684	0.000377	0.093935

Notes:

[1-4] See Table A2.

Table A4: Estimators of δ in the Dynamic Fixed Effects Model, T = 10, N = 25.[1]

Method	Mean Bias[2]	Mean Squared Error[3]	Range of Bias[4]	
OLS	0.528715	0.279573	0.515311	0.540887
Within	0.01447	0.000318	2.58E-05	0.047925
$BN^{(L)}$	0.070544	0.007827	0.000325	0.209585
Δ OLS	0.094141	0.009384	0.052919	0.155376
AH	0.043144	0.00318	0.001558	0.1854
$AH^{+}\ (\hat{\Omega}_{\Delta Z})$	0.043144	0.00318	0.001558	0.1854
$AH^{+}\ (\Omega_{\Delta})$	0.041868	0.002992	0.00039	0.1859
AR	0.019785	0.000603	0.000581	0.068485
$AR^{+}\ (\hat{\Omega}_{\Delta Z})$	0.019785	0.000603	0.000581	0.068485
$AR^{+}\ (\Omega_{\Delta})$	0.019916	0.000601	0.000583	0.063629
$ST^{(b)}\ (\Omega_{\Delta})$	0.070544	0.007827	0.000325	0.209585
$ST^{(c)}\ (\Omega_{\Delta}^{-1/2})$	0.195972	0.074054	0.001611	1.008407
$ST^{(b)}\ (\hat{\Omega}_{\Delta})$	0.070544	0.007827	0.000325	0.209585
$AB\ (\Omega_{\Delta})$	0.012862	0.00026	0.000147	0.042427
$AB^{+}\ (\Omega_{\Delta})$	0.374869	0.914585	0.000835	6.314226
$AB\ (\hat{\Omega}_{\Delta Z})$	0.012862	0.00026	0.000147	0.042427
$AB^{+}\ (\hat{\Omega}_{\Delta Z})$	0.374869	0.914585	0.000835	6.314226
$BN^{(\Delta)}$	0.229236	0.173915	0.002428	3.097474
GMM_F1	0.051264	0.004914	0.00147	0.191849
GMM_F3	0.021225	0.000913	0.000143	0.131132

Notes:

1-4 See Table A2.

Table A5: Estimators of δ in the Dynamic Fixed Effects Model, T = 10, N = 50.[1]

Method	Mean Bias[2]	Mean Squared Error[3]	Range of Bias[4]	
OLS	0.532876	0.283967	0.522233	0.539038
Within	0.003836	2.57E-05	7.83E-08	0.015
$BN^{(L)}$	0.045384	0.003057	0.001725	0.135072
Δ OLS	0.024326	0.000671	0.001507	0.053246
AH	0.014663	0.000335	9.43E-05	0.055326
$AH^{+}(\hat{\Omega}_{\Delta Z})$	0.014663	0.000335	9.43E-05	0.055326
$AH^{+}(\Omega_{\Delta})$	0.014543	0.000333	7.57E-05	0.055273
AR	0.007418	8.09E-05	3.12E-05	0.022434
$AR^{+}(\hat{\Omega}_{\Delta Z})$	0.007418	8.09E-05	3.12E-05	0.022434
$AR^{+}(\Omega_{\Delta})$	0.007343	8.01E-05	2.78E-05	0.023025
$ST^{(b)}(\Omega_{\Delta})$	0.045384	0.003057	0.001725	0.135072
$ST^{(c)}(\Omega_{\Delta}^{-1/2})$	0.110731	0.021933	0.001822	0.422576
$ST^{(b)}(\hat{\Omega}_{\Delta})$	0.045383	0.003057	0.001725	0.135072
$AB(\Omega_{\Delta})$	0.004269	2.99E-05	5.14E-05	0.01642
$AB^{+}(\Omega_{\Delta})$	0.097396	0.127879	0.000378	2.662699
$AB(\hat{\Omega}_{\Delta Z})$	0.004269	2.99E-05	5.14E-05	0.01642
$AB^{+}(\hat{\Omega}_{\Delta Z})$	0.097396	0.127879	0.000378	2.662699
$BN^{(\Delta)}$	0.146942	0.049532	0.006799	1.054854
GMM_F1	0.033789	0.002097	0.000157	0.121882
GMM_F4	0.028738	0.001576	3.77E-05	0.103539

Notes:

1-4 See Table A2.

Appendix IV: Simulation Results for the Random Effects Estimators

Table A6: Estimators of δ in the Dynamic Random Effects Model, $T = 4$, $N = 25$.[1]

Method	Mean Bias[2]	Mean Squared Error[3]	Range of Bias[4]	
OLS	0.393918	0.161595	0.151584	0.571731
Within	0.433624	0.203912	0.034334	0.71087
FGLS	0.415627	0.161595	0.163908	0.611003
$BN^{(a)}$	0.524715	0.745830	0.003487	4.685131
$BN^{(b)}$	0.490185	0.424652	0.000721	3.491835
$BN^{(c)}$	0.496658	0.466430	4.39E-05	2.912157
$HT^{(a)}$	0.495023	0.494498	0.00357	3.088249
$HT^{(b)}$	0.599695	2.306882	0.015717	13.77857
$HT^{(c)}$	0.508347	0.722328	0.006374	6.088942
$AM^{(a)}$	0.584756	0.362804	0.103801	0.920026
$AM^{(b)}$	0.586823	0.364954	0.100193	0.91939
$AM^{(c)}$	0.586154	0.364290	0.102883	0.919934
WB	0.481925	0.374013	0.007922	2.079929
WB^{+}	0.400603	0.210255	0.016837	1.140163
$WB\ (\hat{P}_{Z\Omega})$	0.49612	0.384152	0.005073	1.875037
$WB^{+}\ (\hat{P}_{Z\Omega})$	0.456712	0.312176	0.012567	1.794976
$ABov\ (\hat{\Omega}^{+}_{Hv})$	2.94097	91.80679	0.005289	82.80889
$ABov\ (\hat{\Omega}^{+}_{v})$	5.229269	183.0042	0.065776	96.12697

Notes:

[1-4] See Table A2.

Table A6 (cont): Estimators of δ in the Dynamic Random Effects Model, $T = 4$, $N = 25$.[1]

Method	Mean Bias[2]	Mean Squared Error[3]	Range of Bias[4]	
Δ OLS	0.637704	0.422294	0.316655	0.924967
AH	2.462849	22.16758	0.002307	27.48942
$AH^{+}\ (\Omega_{\Delta})$	0.776385	1.060268	0.015223	5.054092
$AH^{+}\ (\hat{\Omega}_{\Delta Z})$	0.761807	0.991632	0.016161	3.883861
AR	2.04976	16.05741	0.00814	26.88921
$AR^{+}\ (\Omega_{\Delta})$	0.601406	0.665818	0.023453	3.386397
$AR^{+}\ (\hat{\Omega}_{\Delta Z})$	0.581013	0.622498	0.00872	3.531342
$ST^{(b)}\ (\Omega_{\Delta})$	0.582355	0.681493	0.010648	4.520265
$ST^{(c)}\ (\Omega_{\Delta}^{-1/2})$	0.741805	1.190795	0.001692	5.615191
$ST^{(b)}\ (\hat{\Omega}_{\Delta})$	0.581547	0.548405	0.001725	2.210813
$AB\ (\Omega_{\Delta})$	0.721149	0.877135	0.013853	2.840147
$AB^{+}\ (\Omega_{\Delta})$	1.179342	36.20289	0.000233	59.54117
$AB(\hat{\Omega}_{\Delta Z})$	0.819952	1.087605	0.020533	3.001129
$AB^{+}\ (\hat{\Omega}_{\Delta Z})$	0.496043	0.350279	0.022255	1.60222
$BN^{(\Delta)}$	0.744876	1.160283	0.006272	5.755997
$BN^{(\Delta)}\ (\hat{\Omega}_{\Delta Z})$	0.756441	1.151141	0.023782	5.893527
GMM_R1	0.232456	0.095405	0.000807	0.854169
GMM_R2	0.231362	0.097492	0.002344	0.837758

Notes:

[1-4] See Table A2.

Table A7: Estimators of δ in the Dynamic Random Effects Model, T = 4, N = 50.[1]

Method	Mean Bias[2]	Mean Squared Error[3]	Range of Bias[4]	
OLS	0.407284	0.169355	0.195414	0.545194
Within	0.432606	0.197192	0.153578	0.629748
FGLS	0.436067	0.169355	0.24182	0.578482
$BN^{(a)}$	0.326858	0.212239	0.000448	2.294027
$BN^{(b)}$	0.363174	0.242707	0.004332	1.788564
$BN^{(c)}$	0.365902	0.334716	0.005475	3.795044
$HT^{(a)}$	0.34798	0.208258	0.011685	1.566645
$HT^{(b)}$	0.342544	0.205303	0.000622	1.793141
$HT^{(c)}$	0.332238	0.200151	0.001668	1.749771
$AM^{(a)}$	0.555434	0.332069	0.060325	0.988295
$AM^{(b)}$	0.56034	0.337082	0.016344	0.988512
$AM^{(c)}$	0.558211	0.334956	0.040682	0.988404
WB	0.500255	0.362852	0.001635	1.802567
WB^{+}	0.345414	0.175011	0.003925	0.889079
$WB\ (\hat{P}_{Z\Omega})$	0.480236	0.395676	0.003944	2.504127
$WB^{+}\ (\hat{P}_{Z\Omega})$	0.380087	0.212942	0.005724	1.21549
$ABov\ (\hat{\Omega}^{+}_{Hv})$	3.41846	129.8076	0.014757	86.58404
$ABov\ (\hat{\Omega}^{+}_{v})$	4.919443	139.109	0.076941	98.1967

Notes:

[1-4] See Table A2.

Table A7 (cont): Estimators of δ in the Dynamic Random Effects Model, T = 4, N = 50.[1]

Method	Mean Bias[2]	Mean Squared Error[3]	Range of Bias[4]	
Δ OLS	0.64885	0.430132	0.420547	0.894157
AH	2.811272	53.00273	0.001269	51.28071
$AH^{+}\ (\Omega_{\Delta})$	0.68784	1.642277	0.029691	10.6288
$AH^{+}\ (\hat{\Omega}_{\Delta Z})$	0.723536	1.791566	0.036575	11.00441
AR	2.23888	.27.18908	0.008949	35.24582
$AR^{+}\ (\Omega_{\Delta})$	0.467652	0.325748	0.000139	1.529921
$AR^{+}\ (\hat{\Omega}_{\Delta Z})$	0.491642	0.344175	0.007286	1.468837
$ST^{(b)}\ (\Omega_{\Delta})$	0.52218	0.553057	0.002877	2.741581
$ST^{(c)}\ (\Omega_{\Delta}^{-1/2})$	0.553803	0.611454	0.003062	2.794898
$ST^{(b)}\ (\hat{\Omega}_{\Delta})$	0.48151	0.452121	0.002628	2.244147
$AB\ (\Omega_{\Delta})$	0.619902	0.700934	0.004647	2.833015
$AB^{+}\ (\Omega_{\Delta})$	0.689637	2.760235	0.01456	14.33668
$AB(\hat{\Omega}_{\Delta Z})$	0.64421	0.776305	0.003817	3.036889
$AB^{+}\ (\hat{\Omega}_{\Delta Z})$	0.553798	0.869881	0.00709	6.773862
$BN^{(\Delta)}$	0.633138	1.229593	0.025214	6.188318
$BN^{(\Delta)}\ (\hat{\Omega}_{\Delta Z})$	0.640798	1.215717	0.007483	6.016841
GMM_R1	0.174712	0.061927	0.001009	0.952286
GMM_R3	0.169053	0.059289	0.002049	0.960462

Notes:

1-4 See Table A2.

Table A8: Estimators of δ in the Dynamic Random Effects Model, T = 10, N = 25.[1]

Method	Mean Bias[2]	Mean Squared Error[3]	Range of Bias[4]	
OLS	0.357973	0.129908	0.21698	0.437301
Within	0.168219	0.032122	0.00142	0.287052
FGLS	0.340495	0.129908	0.149222	0.459341
$BN^{(a)}$	0.19501	0.072658	0.000633	0.99399
$BN^{(b)}$	0.180563	0.055859	0.003271	0.959967
$BN^{(c)}$	0.175239	0.05101	0.003376	0.680279
$HT^{(a)}$	0.278751	0.119341	0.002866	1.315486
$HT^{(b)}$	0.272477	0.122454	0.000456	1.408965
$HT^{(c)}$	0.282452	0.120456	0.001961	0.976591
$AM^{(a)}$	0.541212	0.293677	0.469585	0.608461
$AM^{(b)}$	0.544877	0.297639	0.468689	0.608542
$AM^{(c)}$	0.543157	0.295791	0.470934	0.611055
WB	0.139995	0.025189	0.020817	0.356473
WB^{+}	0.116944	0.018814	0.001734	0.328883
$WB\ (\hat{P}_{Z\Omega})$	0.128627	0.022608	0.004928	0.391376
$WB^{+}\ (\hat{P}_{Z\Omega})$	0.146853	0.031379	0.000759	0.53842
$ABov\ (\hat{\Omega}^{+}_{Hv})$	3.536692	55.29187	0.065163	44.98672
$ABov\ (\hat{\Omega}^{+}_{v})$	4.314346	168.9669	0.058389	98.38756

Notes:

[1-4] See Table A2.

Table A8 (cont): Estimators of δ in the Dynamic Random Effects Model, T = 10, N = 25.[1]

Method	Mean Bias[2]	Mean Squared Error[3]	Range of Bias[4]	
Δ OLS	0.709456	0.50621	0.553588	0.847541
AH	0.410727	0.366219	0.001319	2.677441
AH^+ (Ω_Δ)	0.286218	0.125063	0.003623	1.05947
AH^+ $(\hat{\Omega}_{\Delta Z})$	0.285825	0.131871	0.000208	1.106822
AR	0.240821	0.091636	0.003948	0.836004
AR^+ (Ω_Δ)	0.197702	0.06385	0.001307	0.730897
AR^+ $(\hat{\Omega}_{\Delta Z})$	0.219418	0.075447	0.001853	0.748879
$ST^{(b)}$ (Ω_Δ)	0.205432	0.080156	0.001335	1.058902
$ST^{(c)}$ $(\Omega_\Delta^{-1/2})$	0.411042	0.317658	0.002711	2.462767
$ST^{(b)}$ $(\hat{\Omega}_\Delta)$	0.232086	0.100317	0.001964	1.304938
AB (Ω_Δ)	0.183523	0.043741	0.004904	0.490483
AB^+ (Ω_Δ)	0.962885	5.443846	0.000774	17.95726
$AB(\hat{\Omega}_{\Delta Z})$	0.183523	0.043741	0.004904	0.490483
AB^+ $(\hat{\Omega}_{\Delta Z})$	0.962885	5.443846	0.000774	17.95726
$BN^{(\Delta)}$	0.419188	0.311296	0.000274	2.478788
$BN^{(\Delta)}$ $(\hat{\Omega}_{\Delta Z})$	0.429478	0.339284	0.003255	2.489361
GMM_R1	0.187588	0.047743	0.000301	0.481215
GMM_R4	0.215933	0.060855	0.00171	0.4962

Notes:

[1-4] See Table A2.

Figure 1

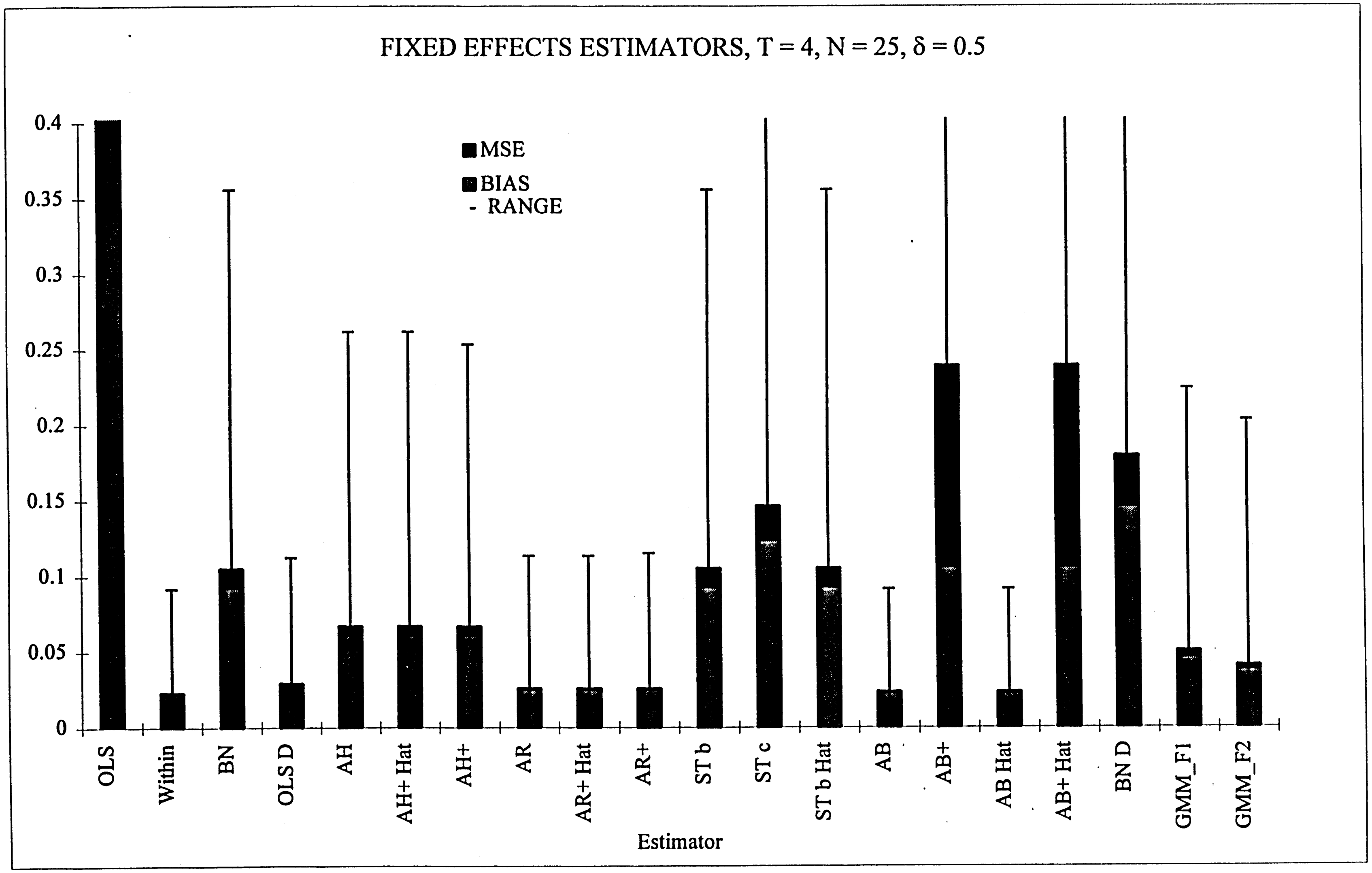

FIXED EFFECTS ESTIMATORS, T = 4, N = 25, δ = 0.5
MSE
BIAS
RANGE
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
OLS
Within
BN
OLS D
AH
AH+ Hat
AH+
AR
AR+ Hat
AR+
ST b
ST c
ST b Hat
AB
AB+
AB Hat
AB+ Hat
BN D
GMM_F1
GMM_F2
Estimator

Figure2

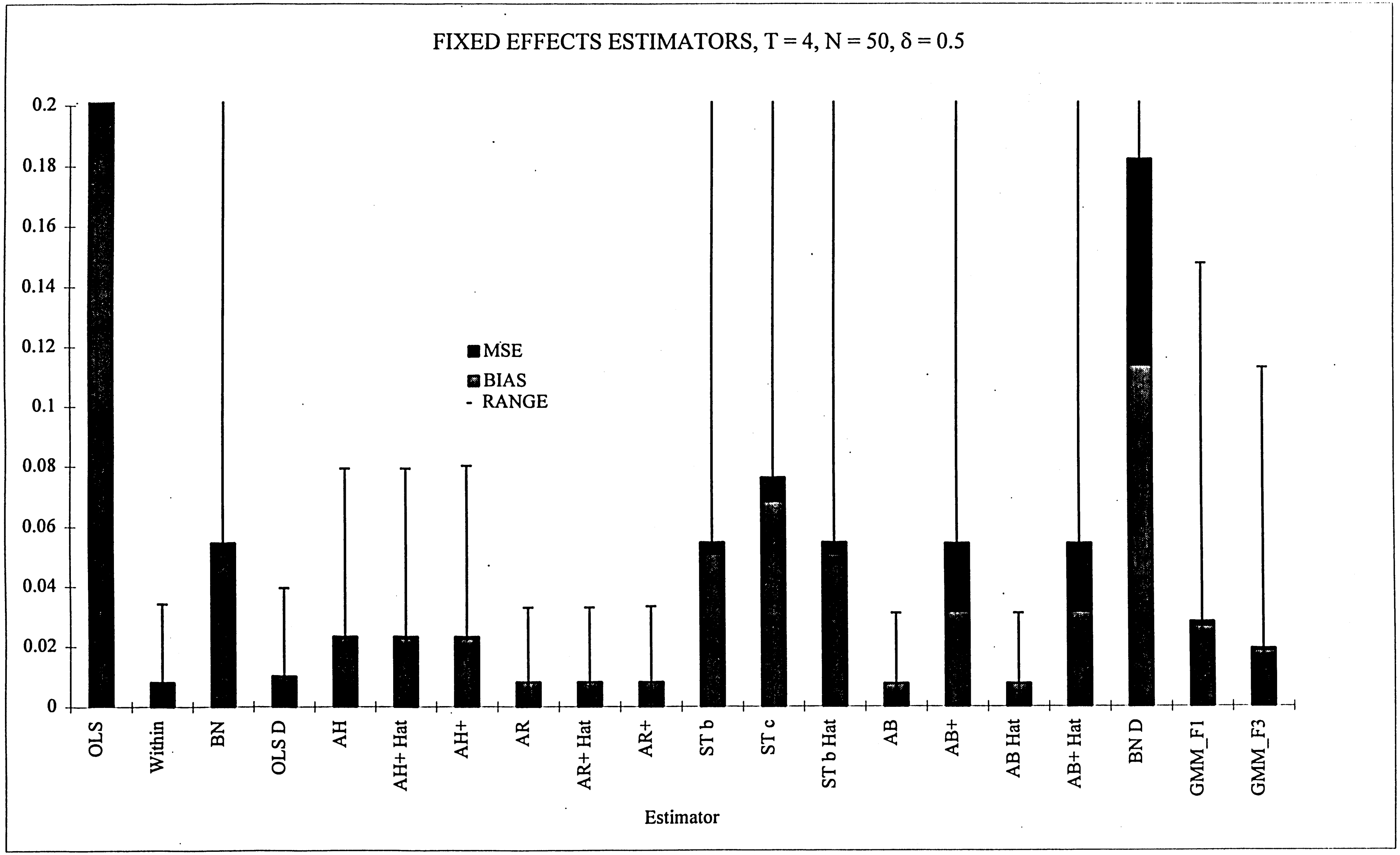
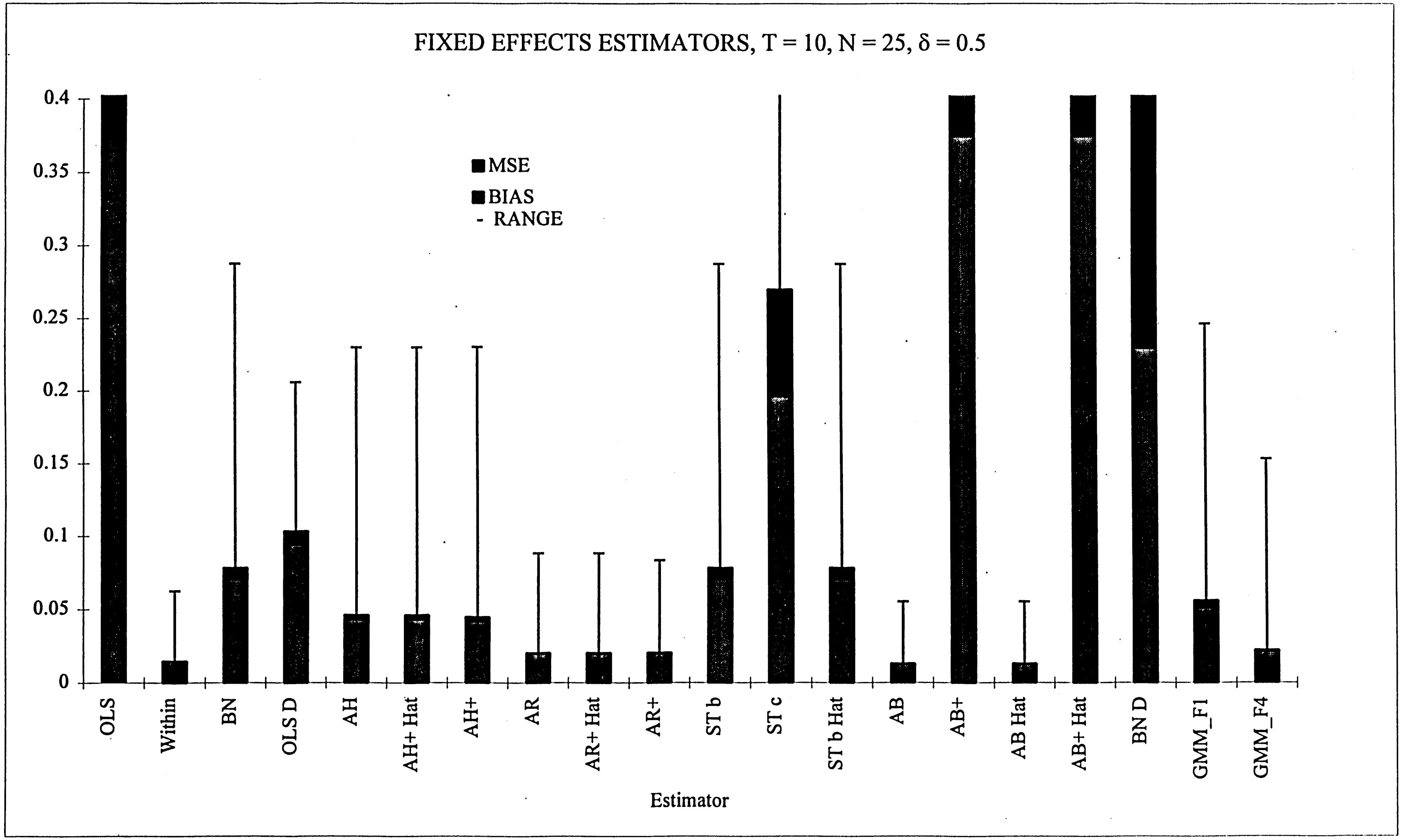

FIXED EFFECTS ESTIMATORS, T = 4, N = 50, δ = 0.5
MSE
BIAS
- RANGE
0.2
0.18
0.16
0.14
0.12
0.1
0.08
0.06
0.04
0.02
0
OLS
Within
BN
OLS D
AH
AH+ Hat
AH+
AR
AR+ Hat
AR+
ST b
ST c
ST b Hat
AB
AB+
AB Hat
AB+ Hat
BN D
GMM_F1
GMM_F3
Estimator

Figure 3



FIXED EFFECTS ESTIMATORS, T = 10, N = 25, δ = 0.5
MSE
BIAS
- RANGE
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
OLS
Within
BN
OLS D
AH
AH+ Hat
AH+
AR
AR+ Hat
AR+
ST b
ST c
ST b Hat
AB
AB+
AB Hat
AB+ Hat
BN D
GMM_F1
GMM_F4
Estimator

Figure 4

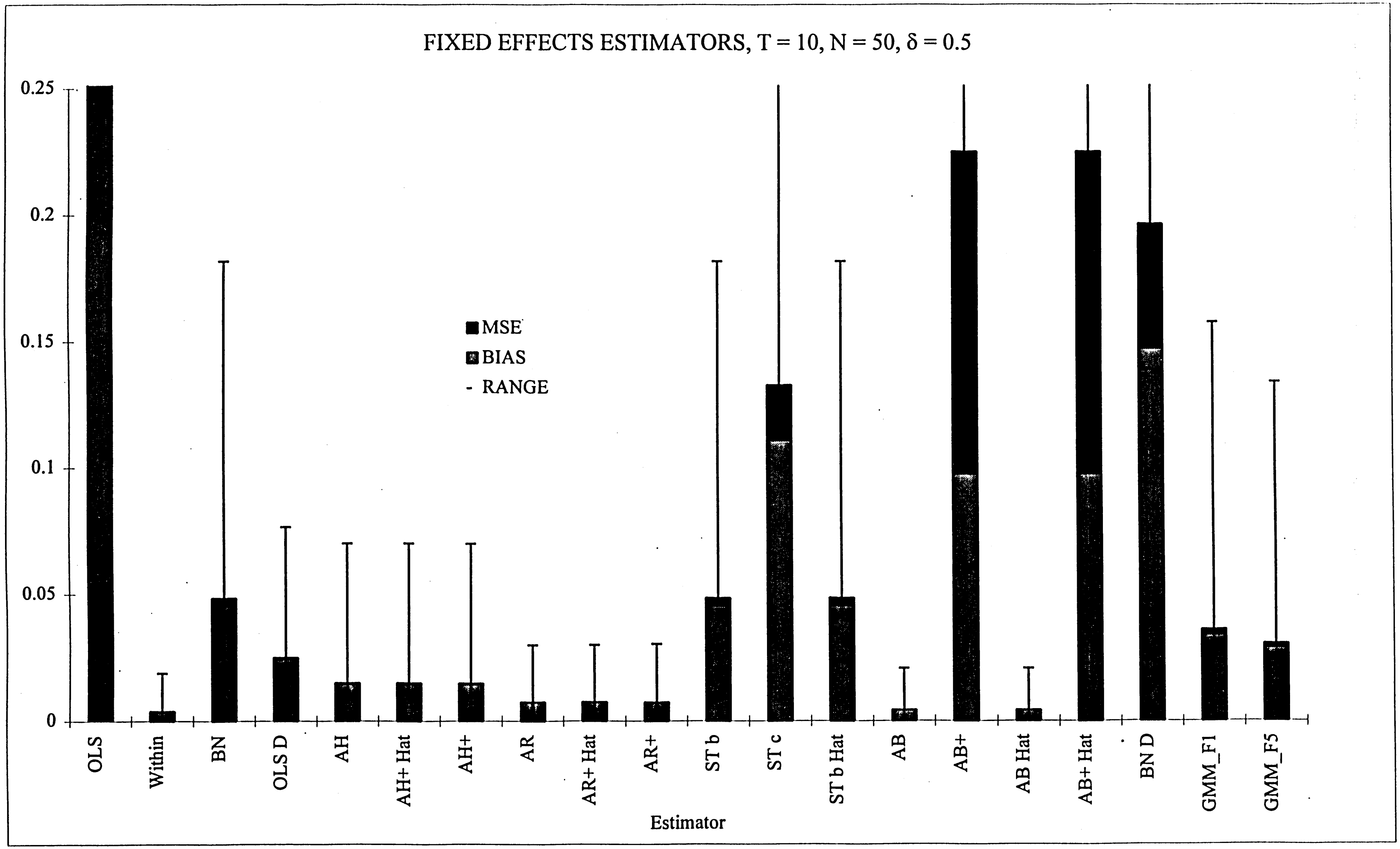

FIXED EFFECTS ESTIMATORS, T = 10, N = 50, δ = 0.5
0.25
0.2
0.15
0.1
0.05
0
MSE
BIAS
RANGE
OLS
Within
BN
OLS D
AH
AH+ Hat
AH+
AR
AR+ Hat
AR+
ST b
ST c
ST b Hat
AB
AB+
AB Hat
AB+ Hat
BN D
GMM_F1
GMM_F5
Estimator

Figure 5

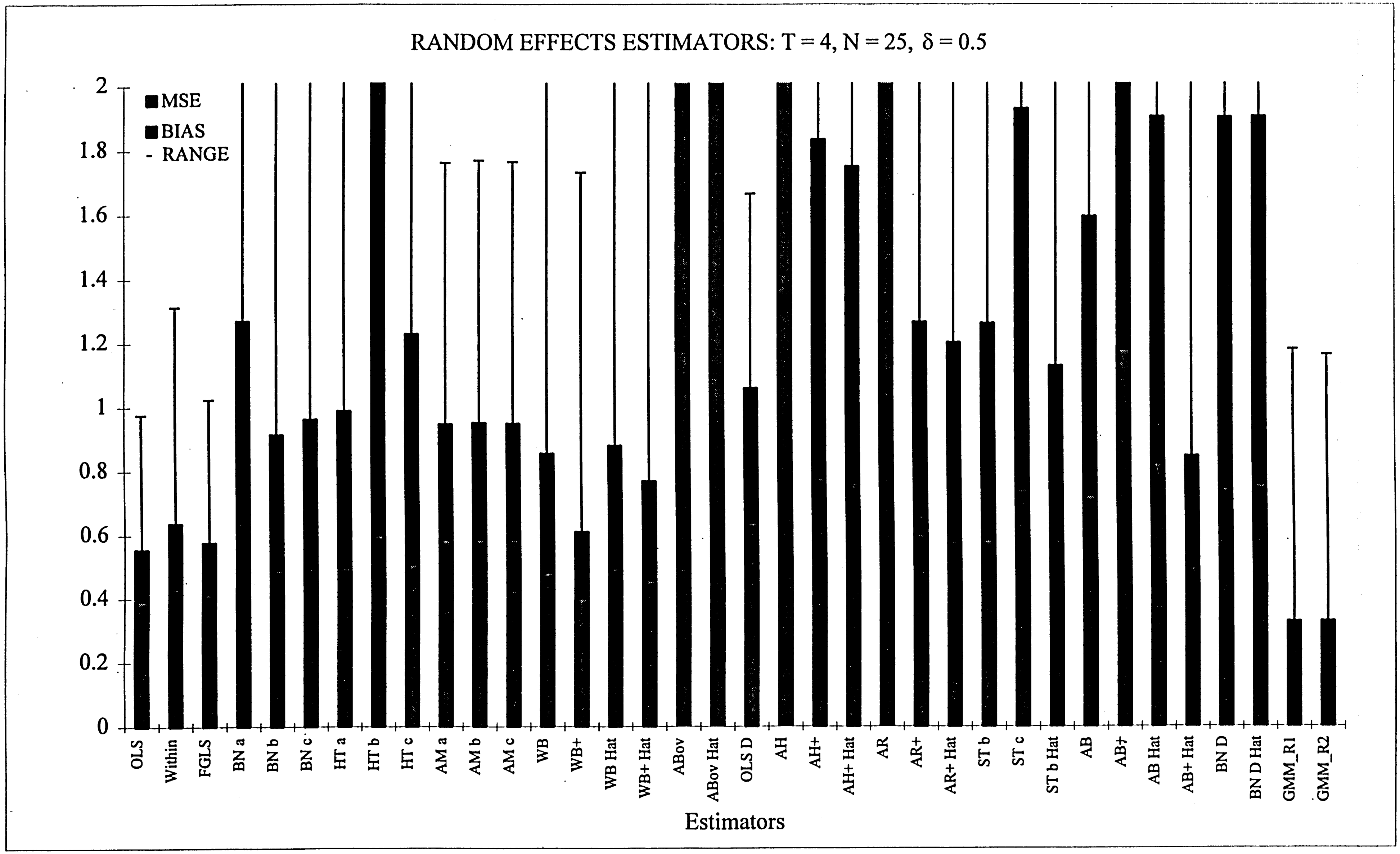

RANDOM EFFECTS ESTIMATORS: T = 4, N = 25, δ = 0.5
MSE
BIAS
- RANGE
2
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
OLS
Within
FGLS
BN a
BN b
BN c
HT a
HT b
HT c
AM a
AM b
AM c
WB
WB+
WB Hat
WB+ Hat
ABov
ABov Hat
OLS D
AH
AH+
AH+ Hat
AR
AR+
AR+ Hat
ST b
ST c
ST b Hat
AB
AB+
AB Hat
AB+ Hat
BN D
BN D Hat
GMM_R1
GMM_R2
Estimators

Figure 6

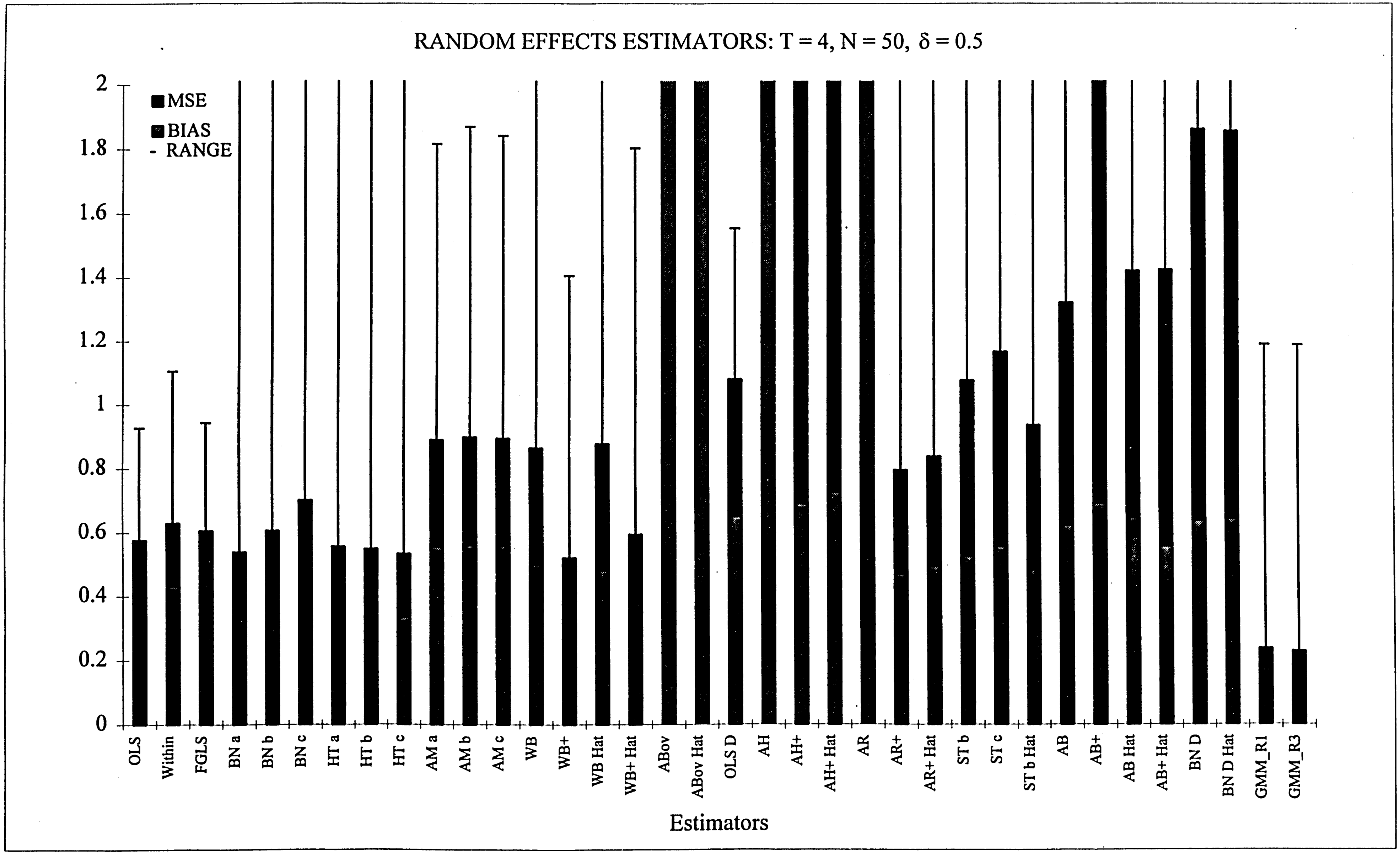
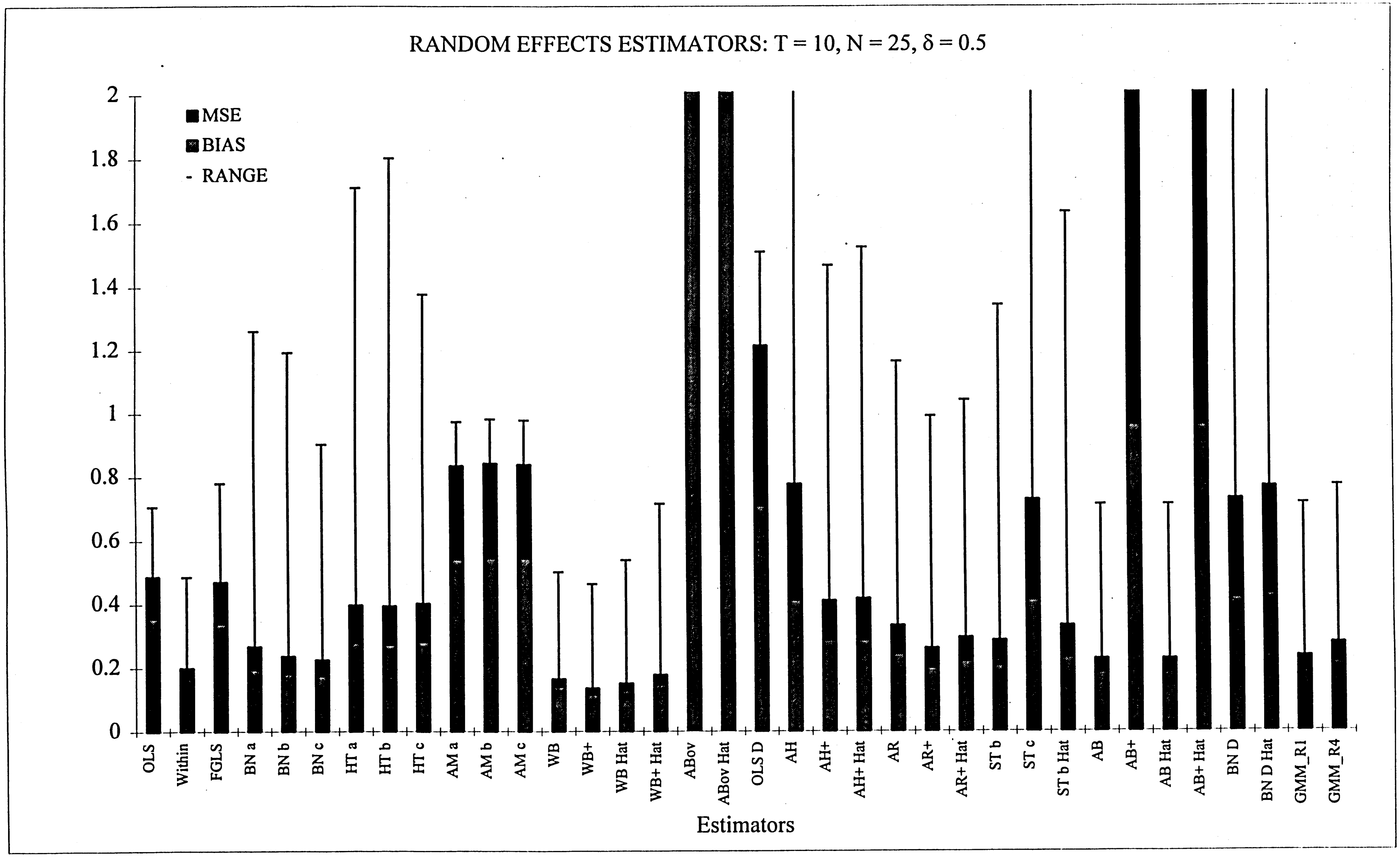

RANDOM EFFECTS ESTIMATORS: T = 4, N = 50, δ = 0.5
MSE
BIAS
RANGE
2
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
OLS
Within
FGLS
BN a
BN b
BN c
HT a
HT b
HT c
AM a
AM b
AM c
WB
WB+
WB Hat
WB+ Hat
ABov
ABov Hat
OLS D
AH
AH+
AH+ Hat
AR
AR+
AR+ Hat
ST b
ST c
ST b Hat
AB
AB+
AB Hat
AB+ Hat
BN D
BN D Hat
GMM_R1
GMM_R3
Estimators

Figure 7


RANDOM EFFECTS ESTIMATORS: T = 10, N = 25, δ = 0.5
MSE
BIAS
RANGE
2
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
OLS
Within
FGLS
BN a
BN b
BN c
HT a
HT b
HT c
AM a
AM b
AM c
WB
WB+
WB Hat
WB+ Hat
ABov
ABov Hat
OLS D
AH
AH+
AH+ Hat
AR
AR+
AR+ Hat
ST b
ST c
ST b Hat
AB
AB+
AB Hat
AB+ Hat
BN D
BN D Hat
GMM_R1
GMM_R4
Estimators