On GMM Estimation and Inference with Bootstrap Bias-Correction in Linear Panel Data Models[*]

Takashi Yamagata[†]
Department of Economics and Related Studies, University of York,
Heslington, York, UK

January 2009

Abstract

This paper proposes a simple bootstrap-bias correction (BBC) of the first-differenced (DIF) and system (SYS) GMM estimators, and investigates their finite sample behaviour, especially with many instruments and under the weak instrument problem. It is found that the BBC DIF and SYS estimators have much less bias, and that among variants of DIF estimators, the BBC DIF estimator can be a reliable alternative in terms of mean square errors. Also, a t-test based on bootstrap standard errors of BBC estimator is proposed. The evidence shows that the proposed t-test of BBC DIF estimator can outperform the standard bootstrap t-test of DIF estimator. Nonetheless, the crude SYS estimator outperforms other BBC estimators considered here, in terms of root mean square errors. Therefore, in the applications where the additional mean stationarity assumption for the validity of SYS estimator are not believed to be satisfied, the BBC DIF estimator together with the proposed bootstrap t-test, can be a reliable alternative.

Key Words: method of moments; dynamic panel data; bootstrap; bias-correction.
JEL Classification: C12, C13, C23.

---

[*] I am grateful to Frank Windmeijer for helpful comments.

[†] Corresponding author. E-mail address: ty509@york.ac.uk

1 Introduction

It is well documented in the literature that, in linear dynamic panel data models, the widely used first-differenced (DIF) generalised method of moments (GMM) estimator proposed by Arellano and Bond (1991) can be largely biased when there are too many moment conditions and/or under the weak instrument problem. Ziliak (1997) has shown that the bias of DIF estimator increases as the number of time observations increases (see also Roodman, 2007). Blundell and Bond (1998) have illustrated the bias of the DIF estimator rises as the weak instrument problem deteriorates, and in view of this, they proposed system (SYS) GMM estimator, imposing further restrictions on the initial conditions process. They have found that SYS estimator can be much less biased and more efficient than the DIF estimator.

There have been two concerns on SYS estimators. Firstly, the mean stationarity assumption which validates the SYS estimation method might not be plausible in some applications. For example, in the growth convergence literature, most developing countries are not in a steady state yet but on the way to it, and in such a situation it is not very clear whether the mean stationarity assumption of SYS estimator – the deviations of the initial conditions from a country steady state are uncorrelated with the level of country steady state itself – is plausible or not.[1] Note that even in such a situation, DIF estimator seems valid, thus, its bias-correction would become important. Secondly, recently some papers have found that SYS estimator is not free from the weak instrument problem. Hayakawa (2007) and Bun and Windmeijer (2007) show that the bias of SYS estimator increases when the variance ratio of individual effects to idiosyncratic errors rises. Here, again, the bias-correction could be an issue.

Some methods of bias-correction of GMM estimator have been proposed in the literature. Alvarez and Arellano (2003) derive the asymptotic bias of DIF estimator, and propose a simple correction which is a linear function of the number of observations and the coefficient, in the context of panel autoregressive (AR) models. Newey and Smith (2004) derive the asymptotic bias of (possibly non-linear) GMM estimator, then Newey, Smith and Ramalho (2001) and Ramalho (2006) propose a bootstrap bias-correction of GMM estimators and compares the finite sample behaviour of these and other moment based estimators. Their results suggest that the bootstrap methods allow the bias of the GMM estimator to be substantially reduced, although at the expense of an increment in its dispersion. As pointed out by MacKinnon and Smith (1998), the question that arises is whether the benefit of bias-reduction exceeds the increase of variation.

In view of these, this paper aims at investigating two important issues which have not been done yet in the literature. Firstly, a simple bootstrap-bias correction of DIF and SYS estimator, based on Hall and Horowitz (1996) case sampling bootstrap, is proposed and finite sample behaviour of these estimators are investigated in dynamic linear panel models, especially with many instruments and under the weak instrument problem. Also, the bootstrap-bias corrected DIF estimator is compared to the bias-corrected estimator of Alvarez and Arellano (2003) in terms of root mean square errors. Note that the bootstrap bias-correction can be applied to a variety of linear dynamic models, meanwhile the latter is applied only for the AR models. Secondly, extending the results of Orme (1995) and Dhaene and Hoorelbeke (2004), Wald test (or t-test) based on the bootstrap standard

---

[1] Bond, Hoeffler and Temple (2001) justifies SYS estimator, while Roodman (2007) claims it may not be justified. Also see Arellano (2003;p.110) for an example of practical implication of the mean stationarity assumption.

errors of the bias-corrected estimator, which are coherent to the bias-corrected estimator, is proposed. The finite sample behaviour of this proposed test, and single and double bootstrap tests is investigated and compared.

The rest of the paper is organised as follows. Section 2 sets out the model and the estimation methods. Section 3 discusses linear bias-corrections and bootstrap bias-correction methods. Section 4 explains the bootstrap Wald test or t-test. Section 5 gives results of Monte Carlo experiments, then Section 6 contains some concluding remarks.

2 Model and Estimation Method

Consider the following linear dynamic model

$$y_{it} = \alpha_i + \lambda y_{i,t-1} + \boldsymbol{\beta}'\mathbf{x}_{it} + u_{it},\ i = 1, 2, ..., N,\ t = 1, 2, ..., T, \tag{1}$$

where α_i is an individual effect with zero-mean and finite variance, $|\lambda| < 1$, $\boldsymbol{\beta}$ is a $(K \times 1)$ parameter vector which is bounded, $\mathbf{x}_{it} = (x_{1it}, x_{2it}, ..., x_{Kit})'$ is a $(K \times 1)$ vector of predetermined regressors such that $E(\mathbf{x}_{is}u_{it}) \neq 0$ for $s > t$, zero otherwise. This can be rewritten as

$$y_{it} = \alpha_i + \boldsymbol{\theta}'\mathbf{w}_{it} + u_{it}, \tag{2}$$

where $\mathbf{w}_{it} = (y_{i,t-1,}\mathbf{x}'_{it})'$, $\boldsymbol{\theta} = (\lambda, \boldsymbol{\beta}')'$. First differencing (2) gives

$$\Delta y_{it} = \boldsymbol{\theta}'\Delta\mathbf{w}_{it} + \Delta u_{it},\ i = 1, 2, ..., N,\ t = 2, 3, ..., T \tag{3}$$

where $\Delta y_{it} = y_{it} - y_{it-1}$, $\Delta\mathbf{w}_{it} = \mathbf{w}_{it} - \mathbf{w}_{it-1}$, $\Delta u_{it} = u_{it} - u_{it-1}$. For further discussion, stacking (1) for each i yields

$$\mathbf{y}_i = \alpha_i\boldsymbol{\iota}_T + \mathbf{W}_i\boldsymbol{\theta} + \mathbf{u}_i,\ i = 1, 2, ..., N, \tag{4}$$

where $\mathbf{y}_i = (y_{i1}, y_{i2}, ..., y_{iT})'$, $\boldsymbol{\iota}_g$ is a $(g \times 1)$ vector of unity with natural number g, $\mathbf{W}_i = (\mathbf{w}_{i1}, \mathbf{w}_{i2}, ..., \mathbf{w}_{iT})'$, $\mathbf{u}_i = (u_{i1}, u_{i2}, ..., u_{iT})'$. The matrix version of the first differenced equation is defined by

$$\Delta\mathbf{y}_i = \Delta\mathbf{W}_i\boldsymbol{\theta} + \Delta\mathbf{u}_i,\ i = 1, 2, ..., N, \tag{5}$$

where $\Delta\mathbf{y}_i = (\Delta y_{i2}, \Delta y_{i3}, ..., \Delta y_{iT})'$, $\Delta\mathbf{W}_i = (\Delta\mathbf{w}_{i2}, \Delta\mathbf{w}_{i3}, ..., \Delta\mathbf{w}_{iT})'$, $\Delta\mathbf{u}_i = (\Delta u_{i2}, \Delta u_{i3}, ..., \Delta u_{iT})'$.

Define the matrix of instruments

$$\mathbf{Z}_i = \begin{bmatrix} \mathbf{w}'_{i1} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & (\mathbf{w}'_{i1}; \mathbf{w}'_{i2}) & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \left(\mathbf{w}'_{i1}; ...; \mathbf{w}'_{iT-1}\right) \end{bmatrix} (T - 1 \times h), \tag{6}$$

where $h = (K + 1)T(T - 1)/2$. GMM estimation is based on the moment restrictions

$$E[\mathbf{Z}'_i\Delta\mathbf{u}_i] = \mathbf{0}. \tag{7}$$

Arellano and Bond (1991) first-differenced two-step (DIF2) GMM estimator is defined as

$$\ddot{\boldsymbol{\theta}}_{DIF2} = \left(\mathbf{A}'_{DIF}\ddot{\boldsymbol{\Omega}}^{-1}_{DIF}\mathbf{A}_{DIF}\right)^{-1}\mathbf{A}'_{DIF}\ddot{\boldsymbol{\Omega}}^{-1}_{DIF}\mathbf{b}_{DIF},$$

where $\mathbf{A}_{DIF} = N^{-1}\sum_{i=1}^{N}\mathbf{Z}_i'\Delta\mathbf{W}_i$, $\mathbf{b}_{DIF} = N^{-1}\sum_{i=1}^{N}\mathbf{Z}_i'\Delta\mathbf{y}_i$, $\ddot{\boldsymbol{\Omega}}_{DIF} = N^{-1}\sum_{i=1}^{N}\mathbf{Z}_i'\Delta\dot{\mathbf{u}}_i\Delta\dot{\mathbf{u}}_i'\mathbf{Z}_i$ with $\Delta\dot{\mathbf{u}}_i = \Delta\mathbf{y}_i - \Delta\mathbf{W}_i\dot{\boldsymbol{\theta}}_{DIF1}$, where $\dot{\boldsymbol{\theta}}_{DIF1}$ is the one-step (DIF1) GMM estimator

$$\dot{\boldsymbol{\theta}}_{DIF1} = \left(\mathbf{A}_{DIF}'\dot{\boldsymbol{\Omega}}_{DIF}^{-1}\mathbf{A}_{DIF}\right)^{-1}\mathbf{A}_{DIF}'\dot{\boldsymbol{\Omega}}_{DIF}^{-1}\mathbf{b}_{DIF},$$

where $\dot{\boldsymbol{\Omega}}_{DIF} = N^{-1}\sum_{i=1}^{N}\mathbf{Z}_i'\mathbf{H}\mathbf{Z}_i$, $\mathbf{H}$ is a square matrix of order $T-1$ with diagonal elements being 2's, the first off-diagonal elements being -1's, and zeros elsewhere. The variance-covariance estimators of $\ddot{\boldsymbol{\theta}}_{DIF2}$ and $\dot{\boldsymbol{\theta}}_{DIF1}$ are given by

$$\ddot{\mathbf{V}}_{DIF2} = \left(\mathbf{A}_{DIF}'\ddot{\boldsymbol{\Omega}}_{DIF}^{-1}\mathbf{A}_{DIF}\right)^{-1}$$

and

$$\dot{\mathbf{V}}_{DIF1} = \left(\mathbf{A}_{DIF}'\dot{\boldsymbol{\Omega}}_{DIF}^{-1}\mathbf{A}_{DIF}\right)^{-1}\mathbf{A}_{DIF}'\dot{\boldsymbol{\Omega}}_{DIF}^{-1}\ddot{\boldsymbol{\Omega}}_{DIF}\dot{\boldsymbol{\Omega}}_{DIF}^{-1}\mathbf{A}_{DIF}\left(\mathbf{A}_{DIF}'\dot{\boldsymbol{\Omega}}_{DIF}^{-1}\mathbf{A}_{DIF}\right)^{-1},$$

respectively. The Wald test statistic or t-test statistic is based on $\ddot{\boldsymbol{\theta}}_{DIF2}$ and $\ddot{\mathbf{V}}_{DIF2}$ or $\dot{\boldsymbol{\theta}}_{DIF1}$ and $\dot{\mathbf{V}}_{DIF1}$, respectively.

It is well known that the DIF estimators can be heavily biased under the weak instrument problem. In view of this, Blundell and Bond (1998) proposed system (SYS) GMM estimator, imposing further restrictions on the initial conditions process, which is called mean stationarity condition in Arellano (2003;p.110), and has found that it can be much less biased and more efficient than the DIF estimator. Under this extra assumption, one could use

$$E[\mathbf{Z}_i^{+\prime}\mathbf{u}_i^{+}] = \mathbf{0}, \tag{8}$$

where

$$\mathbf{Z}_i^{+} = \begin{bmatrix} \mathbf{Z}_i & \mathbf{0} \\ \mathbf{0} & \mathbf{Z}_i^{L} \end{bmatrix} \ (2(T-1)\times h_s) \tag{9}$$

with

$$\mathbf{Z}_i^{L} = \begin{bmatrix} \Delta\mathbf{w}_{i2}' & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \Delta\mathbf{w}_{i3}' & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \Delta\mathbf{w}_{iT}' \end{bmatrix} \ (T-1\times h_L),$$

$h_s = h + h_L$, $h_L = (K+1)(T-1)$, $\mathbf{u}_i^{+} = (\Delta\mathbf{u}_i', \mathbf{u}_i^{L\prime})'$ with $\mathbf{u}_i^{L} = (u_{i2}, u_{i3}, ..., u_{iT})'$. Defining $\mathbf{y}_i^{+} = (\Delta\mathbf{y}_i', \mathbf{y}_i^{L\prime})'$, $\mathbf{W}_i^{+} = (\Delta\mathbf{W}_i'; \mathbf{W}_i^{L\prime})'$, with $\mathbf{y}_i^{L} = (y_{i2}, y_{i3}, ..., y_{iT})'$ and $\mathbf{W}_i^{L} = (\mathbf{w}_{i2}, \mathbf{w}_{i3}, ..., \mathbf{w}_{iT})'$, Blundell and Bond (1998) system two-step (SYS2) GMM estimator is defined as

$$\ddot{\boldsymbol{\theta}}_{SYS2} = \left(\mathbf{A}_{SYS}'\ddot{\boldsymbol{\Omega}}_{SYS}^{-1}\mathbf{A}_{SYS}\right)^{-1}\mathbf{A}_{SYS}'\ddot{\boldsymbol{\Omega}}_{SYS}^{-1}\mathbf{b}_{SYS},$$

where $\mathbf{A}_{SYS} = N^{-1}\sum_{i=1}^{N}\mathbf{Z}_i^{+\prime}\mathbf{W}_i^{+}$, $\mathbf{b}_{SYS} = N^{-1}\sum_{i=1}^{N}\mathbf{Z}_i^{+\prime}\mathbf{y}_i^{+}$, $\ddot{\boldsymbol{\Omega}}_{SYS} = N^{-1}\sum_{i=1}^{N}\mathbf{Z}_i^{+\prime}\dot{\mathbf{u}}_i^{+}\dot{\mathbf{u}}_i^{+\prime}\mathbf{Z}_i^{+}$ with $\dot{\mathbf{u}}_i^{+} = \mathbf{y}_i^{+} - \mathbf{W}_i^{+}\dot{\boldsymbol{\theta}}_{SYS1}$, where $\dot{\boldsymbol{\theta}}_{SYS1}$ is the one-step (SYS1) GMM estimator

$$\dot{\boldsymbol{\theta}}_{SYS1} = \left(\mathbf{A}_{SYS}'\dot{\boldsymbol{\Omega}}_{SYS}^{-1}\mathbf{A}_{SYS}\right)^{-1}\mathbf{A}_{SYS}'\dot{\boldsymbol{\Omega}}_{SYS}^{-1}\mathbf{b}_{SYS},$$

where $\dot{\boldsymbol{\Omega}}_{SYS} = N^{-1}\sum_{i=1}^{N}\mathbf{Z}_i^{+\prime}\mathbf{H}^{+}\mathbf{Z}_i^{+}$ with

$$\mathbf{H}^{+} = \begin{bmatrix} \mathbf{H} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{T-1} \end{bmatrix}.$$

The variance-covariance estimators of $\ddot{\boldsymbol{\theta}}_{SYS2}$ and $\dot{\boldsymbol{\theta}}_{SYS1}$ are given by

$$\ddot{\mathbf{V}}_{SYS2} = \left(\mathbf{A}'_{SYS}\ddot{\boldsymbol{\Omega}}^{-1}_{SYS}\mathbf{A}_{SYS}\right)^{-1}$$

and

$$\dot{\mathbf{V}}_{SYS1} = \left(\mathbf{A}'_{SYS}\dot{\boldsymbol{\Omega}}^{-1}_{SYS}\mathbf{A}_{SYS}\right)^{-1}\mathbf{A}'_{SYS}\dot{\boldsymbol{\Omega}}^{-1}_{SYS}\ddot{\boldsymbol{\Omega}}_{SYS}\dot{\boldsymbol{\Omega}}^{-1}_{SYS}\mathbf{A}_{SYS}\left(\mathbf{A}'_{SYS}\dot{\boldsymbol{\Omega}}^{-1}_{SYS}\mathbf{A}_{SYS}\right)^{-1},$$

respectively. The Wald test statistic or t-test statistic are based on $\ddot{\boldsymbol{\theta}}_{SYS2}$ and $\ddot{\mathbf{V}}_{SYS2}$ or $\dot{\boldsymbol{\theta}}_{SYS1}$ and $\ddot{\mathbf{V}}_{SYS2}$, respectively.

3 Bias-Correction Methods

It is well-known that the DIF estimator can be largely biased when there are too many moment conditions and/or under the weak instrument problem (Ziliak, 1997). The SYS estimator is much less biased in general, however, the required mean stationarity assumption for the validity of the SYS estimator may not be plausible in some applications, and, furthermore, recent literature illustrates that the bias of the SYS estimator increases when the variance ratio of individual effects to idiosyncratic errors rises (Hayakawa, 2007 and Bun and Windmeijer, 2007). Therefore, the bias-correction of both DIF and SYS estimators can be important.

In this section, a linear bias-correction method and the bootstrap bias-correction methods are explained.

3.1 Linear Bias-Correction

Consider the panel autoregressive of order one (AR(1)) models, namely, $\boldsymbol{\beta} = \mathbf{0}$ in (1) and $w_{it} = y_{it-1}$. In such a case, Alvarez and Arellano (2003) prove that, provided $T/N \to c$, $0 \leq c < \infty$,

$$\sqrt{NT}\left[\ddot{\lambda}_{DIF2} - \left(\lambda - N^{-1}\left(1+\lambda\right)\right)\right] \xrightarrow{d} N\left(0, 1-\lambda^2\right). \tag{10}$$

The straightforward linear bias-corrected (LBC) DIF2 estimator would be

$$\ddot{\lambda}_{DIF2-LBC} = \ddot{\lambda}_{DIF2} + N^{-1}\left(1 + \ddot{\lambda}_{DIF2}\right),$$

and the variance estimator of $\ddot{\lambda}_{DIF2-LBC}$ is

$$\ddot{V}_{DIF2-LBC} = \left(1 - \ddot{\lambda}^2_{DIF2-LBC}\right)/NT. \tag{11}$$

The counterparts of DIF1 estimator are similarly defined. This bias-correction is appealing because of its simplicity. On the other hand, similar to the case of the bias of ordinary least square (OLS) estimator of the root of AR(1) regression, provided by MacKinnon and Smith (1998), the bias function of $\ddot{\lambda}_{DIF2}$ may be non-linear in λ, especially where λ is close to 1. Also, the variance estimator may have a poor finite sample approximation.[2] Finally, this approach is valid only for pure autoregressive type models. Given these, the bootstrap bias-correction is explained next.[3]

---

[2] The bootstrap variance estimator is proposed later.

[3] Other more general analytical bias- corrections, which can be used for non-linear GMM estimation, are proposed and examined in Newey, Ramalho and Smith (2003), Newey and Smith (2004), Ramalho

3.2 Bootstrap Bias-correction

The bootstrap bias-correction in the GMM models is not new. See, for example, Newey, Smith and Ramalho (2001) and Ramalho (2006). Given the recent advancement in personal computers and the popularity of bootstrap methods gained in empirical literature, the bootstrap bias-correction is relatively simple and easily accessible by practitioners. However, I have not found so far any applications of the bootstrap bias-correction in linear dynamic panel data models after GMM estimation.[4] For the bootstrap, Hall and Horowitz (1996) case resampling bootstrap is considered here.[5] All illustrations are for DIF1 and DIF2 estimators, and the procedures for SYS1 and SYS2 estimators are analogous.

Procedure to Obtain Bootstrap bias-corrected Estimator

1. Obtain $\dot{\boldsymbol{\theta}}_{DIF1}$ and $\ddot{\boldsymbol{\theta}}_{DIF2}$.

2. Generate j^{th} bootstrap samples, $\{\mathbf{y}_i^{*(j)}, \mathbf{W}_i^{*(j)}\}_{i=1}^N$, by resampling from $\{\mathbf{y}_i, \mathbf{W}_i\}_{i=1}^N$, with replacement.

3(i). Using the bootstrap samples, obtain j^{th} bootstrap estimator minimising the objective function of recentred moment conditions:

$$\ddot{\boldsymbol{\theta}}^{*j}_{DIF2} = \left(\mathbf{A}^{*j\prime}_{DIF}\left(\ddot{\boldsymbol{\Omega}}^{*j}_{DIF}\right)^{-1}\mathbf{A}^{*j}_{DIF}\right)^{-1}\mathbf{A}^{*j\prime}_{DIF}\left(\ddot{\boldsymbol{\Omega}}^{*j}_{DIF}\right)^{-1}\ddot{\mathbf{b}}^{*j}_{DIF},$$

where $\mathbf{A}^{*j}_{DIF} = N^{-1}\sum_{i=1}^N \mathbf{Z}_i^{*j\prime}\Delta\mathbf{W}_i^{*j}$, $\mathbf{b}^{*j}_{DIF} = N^{-1}\sum_{i=1}^N \mathbf{Z}_i^{*j\prime}\Delta\mathbf{y}_i^{*j} - N^{-1}\sum_{i=1}^N \mathbf{Z}_i'\Delta\ddot{\mathbf{u}}_i$, $\ddot{\boldsymbol{\Omega}}^{*j}_{DIF} = N^{-1}\sum_{i=1}^N \left(\mathbf{Z}_i^{*j\prime}\Delta\dot{\mathbf{u}}_i^{*j} - N^{-1}\sum_{i=1}^N \mathbf{Z}_i'\Delta\dot{\mathbf{u}}_i\right)\left(\mathbf{Z}_i^{*j\prime}\Delta\dot{\mathbf{u}}_i^{*j} - N^{-1}\sum_{i=1}^N \mathbf{Z}_i'\Delta\dot{\mathbf{u}}_i\right)'$ with $\Delta\dot{\mathbf{u}}_i^{*j} = \Delta\mathbf{y}_i^{*j} - \Delta\mathbf{W}_i^{*j}\dot{\boldsymbol{\theta}}^{*j}_{DIF1}$, where $\dot{\boldsymbol{\theta}}^{*j}_{DIF1}$ is the one-step (DIF1) GMM estimator

$$\dot{\boldsymbol{\theta}}^{*j}_{DIF1} = \left(\mathbf{A}^{*j\prime}_{DIF}\left(\dot{\boldsymbol{\Omega}}^{*j}_{DIF}\right)^{-1}\mathbf{A}^{*j}_{DIF}\right)^{-1}\mathbf{A}^{*j\prime}_{DIF}\left(\dot{\boldsymbol{\Omega}}^{*j}_{DIF}\right)^{-1}\dot{\mathbf{b}}^{*j}_{DIF},$$

where $\dot{\boldsymbol{\Omega}}^{*j}_{DIF} = N^{-1}\sum_{i=1}^N \mathbf{Z}_i^{*j\prime}\mathbf{H}\mathbf{Z}_i^{*j}$, $\dot{\mathbf{b}}^{*j}_{DIF} = N^{-1}\sum_{i=1}^N \mathbf{Z}_i^{*j\prime}\Delta\mathbf{y}_i^{*j} - N^{-1}\sum_{i=1}^N \mathbf{Z}_i'\Delta\dot{\mathbf{u}}_i$.

4. Repeat steps 2-4 B_1 times and obtain the bootstrap bias-corrected estimates

$$\dot{\boldsymbol{\theta}}_{DIF1-BBC} = 2\dot{\boldsymbol{\theta}}_{DIF1} - \overline{\dot{\boldsymbol{\theta}}}^{*}_{DIF1}$$

where $\overline{\dot{\boldsymbol{\theta}}}^{*}_{DIF1} = B_1^{-1}\sum_{j=1}^{B_1}\dot{\boldsymbol{\theta}}^{*j}_{DIF1}$. $\ddot{\boldsymbol{\theta}}_{DIF2-BBC}$ is computed similarly.

The finite sample bias of the estimators bring potential misinterpretation of the estimated model. For example, a downward bias of estimators could result in underestimation of the magnitude of the effect of variables. The bias-correction methods discussed above would be a solution. However, bias-correction can bring another problem, which is discussed next.

(2006). We do not consider these in this paper, since their results show that the performance of the best method among them provide rather similar to that of the bootstrap bias adjustment in terms of bias reduction, RMSEs, and mean absolute error.

[4] See Ziliak (1997) for relating discussions.

[5] Another popular bootstrap method in the panel GMM context is proposed by Brown and Newey (2002). We only focus on the Hall and Horowitz bootstrap, since, fairly large amount of evidence have shown that these two bootstrap methods produce similar results; see Bond and Windmeijer (2005), for example.

4 Wald test based on Bootstrap Estimator of Standard Errors

Another potential problem which may be caused by the finite sample bias is the bias in inference. For instance, because of the bias of the estimator, a t-test with biased estimators could lead to its poor finite sample behaviour. In general, this problem cannot be sorted out solely by its bias-correction. The reason is because the bias-corrected estimator will have a different distribution from the non-bias-corrected one with finite sample, and this difference might not be negligible.[6] Two solutions to this problem can be thought of. First, and probably most popular one, is to adopt the bootstrap t-test. As a bootstrap procedure is aiming at mimicking the finite sample distribution of a test statistic under the null, it will take account of the finite sample bias of an estimator. Second, given that the bias is corrected already, another asymptotically valid test would be the t-test based on the bias-corrected estimator and its bootstrap variance estimator, which is obtained by simulating the variance of the estimator by bootstrap replications.[7] This bootstrap approach is different from the former in that the latter tries to approximate the distribution of the bias-corrected estimator up to the second moment, but the former does not. This proposed bootstrap test procedure requires a second level bootstrap for the bootstrap bias-corrected estimator.

Procedure to Obtain the t-test based on Bootstrap Variance Estimator

Steps 1, 2, 3(i), and 4 are the same as above, but the following Step 3(ii) and Step 5 are inserted right after Step 3(i) and Step 4, respectively:

3(ii) For each first-level bootstrap j, obtain the linear bias-corrected estimates, $\dot{\boldsymbol{\theta}}^{*j}_{DIF1-LBC}$ and $\ddot{\boldsymbol{\theta}}^{*j}_{DIF2-LBC}$. If the bootstrap bias-correction is used, the second level bootstrap is required as below:

  (a) Generate ℓ^{th} second level bootstrap samples, $\{\mathbf{y}_i^{**(\ell)}, \mathbf{W}_i^{**(\ell)}\}_{i=1}^N$, by resampling from $\{\mathbf{y}_i^{*(j)}, \mathbf{W}_i^{*(j)}\}_{i=1}^N$, with replacement.

  (b) Using the bootstrap data, obtain the second level bootstrap estimators $\dot{\boldsymbol{\theta}}^{**\ell}_{DIF1}$ and $\ddot{\boldsymbol{\theta}}^{**\ell}_{DIF2}$ in a similar manner to Step 3(i).

  (c) Repeat steps 3(ii)(a)-(b) B_2 times, and obtain the j^{th} bootstrap bias-corrected estimates

$$\dot{\boldsymbol{\theta}}^{*j}_{DIF1-BBC} = 2\dot{\boldsymbol{\theta}}^{*j}_{DIF1} - \overline{\dot{\boldsymbol{\theta}}}^{**}_{DIF1}$$

  where $\overline{\dot{\boldsymbol{\theta}}}^{**}_{DIF1} = B_2^{-1}\sum_{\ell=1}^{B_2}\dot{\boldsymbol{\theta}}^{**\ell}_{DIF1}$. $\ddot{\boldsymbol{\theta}}^{*j}_{DIF2-BBC}$ is computed similarly.

5. Compute the bootstrap variance covariance estimator of $\dot{\boldsymbol{\theta}}_{DIF1-BBC}$ by

$$\begin{aligned}\dot{\mathbf{V}}^{*}_{DIF1-BBC} &= (B_1-1)^{-1}\sum_{j=1}^{B_1}\left(\dot{\boldsymbol{\theta}}^{*j}_{DIF1-BBC} - \overline{\dot{\boldsymbol{\theta}}}^{*}_{DIF1-BBC}\right) \\ &\qquad \times\left(\dot{\boldsymbol{\theta}}^{*j}_{DIF1-BBC} - \overline{\dot{\boldsymbol{\theta}}}^{*}_{DIF1-BBC}\right)' \end{aligned} \qquad (12)$$

---

[6] This difference in distributions, of course, may be negligible asymptotically.

[7] For the application of this type of idea to non bias-adjusted estimator, see Bond and Windmeijer (2003), and also Orme (1995) and Dhaene and Hoorelbeke (2004).

with $\overline{\dot{\boldsymbol{\theta}}}^{*}_{DIF1-BBC} = B_1^{-1}\sum_{j=1}^{B_1}\dot{\boldsymbol{\theta}}^{*j}_{DIF1-BBC}$. $\ddot{\mathbf{V}}^{*}_{DIF2-BBC}$ is obtained similarly. The Wald test statistic or t-statistics are obtained using $\dot{\boldsymbol{\theta}}_{DIF1-BBC}$ and $\dot{\mathbf{V}}^{*}_{DIF1-BBC}$ or $\ddot{\boldsymbol{\theta}}_{DIF2-BBC}$ and $\ddot{\mathbf{V}}^{*}_{DIF2-BBC}$.

In the next section, finite sample behaviour of the bias-corrected estimators and the proposed test is examined.

5 Monte Carlo Experiments

In this section, we investigate finite sample behaviour of the bias-corrected DIF and SYS GMM estimators by means of Monte Carlo experiments. The proposed t-test is compared with the standard bootstrapped t-test, which is proposed and examined in Bond and Windmeijer (2005), and the double bootstrapped t-test, expecting a potential further refinement. Particularly we consider the environment where the number of moment conditions is large and where the weak instrument problem exists.

5.1 Design

The first data generating process (DGP) considered is a panel AR(1) model, which is

$$y_{it} = \alpha_i + \lambda y_{i,t-1} + u_{it},\ i = 1, 2, ..., N; t = -48, -47, ..., T, \tag{13}$$

where $\alpha_i \sim iidN(0, \sigma_\alpha^2)$, $u_{it} \sim iidN(0, \sigma_u^2)$, $y_{i,-49} = 0$ and first 49 observations are discarded. We consider $\lambda = 0.2, 0.5, 0.8$. We set $\sigma_u^2 = 1$ and

$$\sigma_\alpha^2 = \eta_1^2 \frac{1-\lambda}{1+\lambda}$$

so that the long-run relative impact of α_i to u_{it} is constant for different value of λ, which is controlled by η_1^2. η_1^2 is set to four, which gives the weak instrument problem.[8]

Another DGP considered is a linear model with a predetermined regressor, which is

$$y_{it} = \alpha_i + \beta x_{it} + u_{it},\ i = 1, 2, ..., N; t = 1, 2, ..., T, \tag{14}$$

where $\alpha_i \sim iidN(0, \sigma_\alpha^2)$, $u_{it} \sim iidN(0, \sigma_u^2)$,

$$x_{it} = \alpha_i + \rho x_{i,t-1} + \pi u_{i,t-1} + v_{it}, i = 1, 2, ..., N; t = -48, -47, ..., T,$$

where $v_{it} \sim iidN(0, \sigma_v^2)$, $x_{i,-49} = 0$ and first 50 observations are discarded. We set $\rho = 0.5$, $\pi = 0.5$, $\sigma_u^2 = 1$, $\sigma_v^2 = 0.5$, and

$$\sigma_\alpha^2 = \eta_1^2 \frac{(1-\rho)\left(\beta^2\pi^2 + 1 - \rho^2\right)}{(1+\rho)(1-\pi+\beta)^2}.$$

η_1^2 is set to four again.

We consider all combinations of $N = 100$, $T = 6, 8, 10$, which give 15, 28, 45 moment restrictions, respectively. The weak instrument problem, which is more serious than considered in Blundell and Bond (1998), is generated by a high long-run relative impact of α_i to u_{it} ($\eta_1^2 = 4$). All experiments are based on 2000 replications, except for $T = 10$ whicha are based on 1000 replications, and 99 first and second level bootstrap samples.

---

[8] Bun and Windmeijer (2007) consider a panel AR(1) model $\sigma_\alpha^2 = \{1/4, 1, 4\}$ for $\lambda = 0.8$, $\sigma_u^2 = 1$. Corresponding values of η_1^2 to their σ_α^2 are $\{2.25, 9, 36\}$, last two of which might be seen rather extreme.

5.2 Results

5.2.1 Panel AR(1) Models

Panel A of Table 1 reports the properties of various DIF estimators. In terms of the bias, bootstrap bias-correction (BBC) works best. Linear bias-correction (LBC) reduces the bias moderately, though, the bias-correction is not very effective when $\lambda = 0.8$, as expected. DIF1 is less dispersed than DIF2, uniformly, probably because the errors are homoskedastic. Also the bias of DIF2 is not smaller than that of DIF1, in most cases. Reflecting this, DIF1 is always better than DIF2 in terms of root mean square errors (RMSEs). When $\lambda = 0.2$, DIF1-LBC is better than DIF1-BBC in terms of RMSEs. When $\lambda = 0.5$, the RMSEs of DIF1-BBC are almost the same as those of DIF1-LBC, and the former is the best when $\lambda = 0.8$. Among the two-step DIF estimators, DIF2-BBC is always the best, only exception occurs when $\lambda = 0.2$ with $T = 6$.

Panel B of Table 1 provides the properties of various SYS estimators. The SYS2 is always less dispersed than SYS1, and most of the former has less bias than the latter. Thus, in terms of RMSEs, SYS2 is better than SYS1. In terms of bias, when $\lambda = 0.2$ and 0.5, SYS1 and SYS2 are moderately upward biased, but they are not when $\lambda = 0.8$. This is consistent with the results of Hayakawa (2006). On the other hand, SYS1-BBC and SYS2-BBC have much less bias for all values of λ considered.

Nevertheless, when Panel A and B of Table 1 are compared, it is clear that SYS2 is always the best, in terms of RMSEs.

Table 2 reports the estimated size and the power of the t-tests based on DIF estimators. The finite sample behaviour of the t-test based on DIF2 is poor, and that based on DIF1 is poor but slightly less than the former, as shown in the literature. The t-test of the DIF1-LBA and DIF2-LBA, based on the variance estimator given by (11), always reject the null. The behaviour of the test based on DIF1 becomes worse as λ increases. The t-test based on variance estimator of DIF1 using single bootstrap behaves similarly. The size of the t-test based on variance estimator of DIF2 using single bootstrap approaches zero as T rises. Standard bootstrap t-test behaves satisfactorily for $\lambda = 0.2$ and 0.5, but when $\lambda = 0.8$ it largely overrejects. On the other hand, the proposed t-test based on the bootstrap variance estimator of DIF1-BBC, defined by (12), controls the size very well, maintaining reasonable power. The double bootstrapping t-test fails to improve over the standard single bootstrapping for all cases considered.

Table 3 reports the estimated size and power of the t-tests based on SYS estimators. The finite sample behaviour of the t-test based on SYS2 is dreadful. The t-test based on SYS1 moderately overrejects, but its behaviour becomes better as λ increases, probably because the bias goes down. The best performed test is clearly the standard bootstrap t-test of SYS1. The followers are the proposed t-test based on the bootstrap variance estimator of SYS1-BBC, usual t-test of SYS1, and the double bootstrapped t-test of SYS1.

5.2.2 Panel Models with a Predetermined Regressor

Table 4 reports the properties of various DIF and SYS estimators. Interestingly the bias of SYS1 and SYS2 is mostly larger than that of DIF1 and DIF2. Again, the bootstrap-bias correction reduces the bias remarkably well. In this model, the root mean square errors of the non-bias-corrected and the bias-corrected one-step (or two-step) estimators are very similar, though, the former tends to be smaller than the latter. Table 5 provides the estimated size and power of the various t-tests, and similar results to the panel AR(1) models are found.

6 Concluding Remarks

This paper has proposed a simple bootstrap-bias correction of the first-differenced (DIF) and system (SYS) GMM estimators, based on the Hall and Horowitz (1996) case sampling bootstrap, and has investigated finite sample behaviour of these estimators, especially with many instruments and under the weak instrument problem, in dynamic linear panel models. Also, the bootstrap-bias corrected DIF estimator is compared to the bias-corrected estimator of Alvarez and Arellano (2004), in terms of root means square errors. The Monte Carlo evidence shows that the bootstrap bias-correction (BBC) reduces the bias of the DIF and SYS estimators remarkably well, and the bootstrap bias-corrected DIF estimators can be reliable alternative to the linear bias-corrected (LBC) DIF estimators of Arellano and Alvarez (2003), in terms mean square errors.

Also, extending the results of Orme (1995) and Dhaene and Hoorelbeke (2004), the Wald test or the t-test based on the bootstrap standard errors of the bias-corrected estimator, which is coherent to the bias-corrected estimator, is proposed. The finite sample behaviour of this proposed test, and the single and the double bootstrap tests is investigated and compared. The results show that the standard bootstrap t-test of the DIF estimator behaves unsatisfactorily when the data is persistent. On the other hand, the proposed t-test based on the bootstrap variance estimator of the bias-corrected one-step DIF estimator controls the size very well, maintaining reasonable power. The double bootstrapped t-test fails to improve over the standard single bootstrapping for all cases considered.

Nonetheless, the crude system GMM estimator of Blundell and Bond (1998) outperforms other bias-corrected estimators considered here, in terms of root mean square errors. Therefore, when the additional initial conditions for the validity of the system GMM estimator – the deviations of the initial conditions from the individual steady state are uncorrelated with the level of individual steady state itself – are not believed to be satisfied, the bootstrap bias-corrected first-differenced estimator together with the proposed bootstrap Wald test or t-test, can be a reliable alternative.

References

Alvarez, J., Arellano, M., (2003). The time-series and cross-section asymptotics of dynamic panel data estimators, Econometrica 71, 1121-1159.

Arellano, M., (2003). Panel Data Econometrics, Oxford University Press, Oxford.

Arellano, M., Bond, S., (1991). Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. Review of Economic Studies 58, 277-297.

Blundell, R.W., Bond, S.R., (1998). Initial conditions and moment restrictions in dynamic panel data models. Journal of Econometrics 87, 115-143.

Bond, S., Hoeffler, A., Temple, J., (2001). GMM estimation of empirical growth Models. Nuffield College Working Papers 21.

Bond, S., Windmeijer, F., (2003). Bootstrapping for GMM: A comparison of alternative inference procedures. mimeo.

Bond, S., Windmeijer, F., (2005). Reliable inference for GMM estimators? Finite sample procedures in linear panel data models. Econometric Reviews 24, 1-37.

Brown, B.W., Newey, W.K., (2002). Generalized method of moments, efficient bootstrapping, and improved inference. Journal of Business and Economic Statistics 20, 507-517.

Bun, M.J.G., Kiviet, J.F., (2006). The effects of dynamic feedbacks on LS and MM estimator accuracy in panel data models. Journal of Econometrics 132, 409-444.

Bun, M.J.G., Windmeijer, F., (2007). The weak instrument problem of the system GMM estimator in dynamic panel data models. Cemmap Working Paper CWP08/07.

Dhaene G., Hoorelbeke, D., (2004). The information matrix test with bootstrap-based covariance matrix estimation. Economics Letters 82, 341–347.

Hall, P., Horowitz, J., (1996). Bootstrap critical values for tests based on generalized method of moments estimators. Econometrica 64, 891-916.

Hahn, J., Hausman, J., Kuersteiner, G., (2007). Long difference instrumental variables estimation for dynamic panel models with fixed effects. Journal of Econometrics 140, 574-617.

Hansen, L., (1982). Large sample properties of generalized method of moments estimators. Econometrica 50, 1029-1054.

Hayakawa, K., (2007). Small sample bias properties of the system GMM estimator in dynamic panel data models. Economics Letters 95, 32–38.

MacKinnon, J.G., Smith, Jr, A.A., (1998). Approximate bias correction in econometrics. Journal of Econometrics 85, 205-230.

Newey K.N., Ramalho, J.J.S., Smith, R.J (2001). Asymptotic bias for GMM and GEL estimators with estimated nuisance parameters. paper presented at 2001 NSF-Berkeley Econometrics Symposium on “Identification and Inference for Econometric Models.”

Newey K.N., Smith, R.J., (2004). Higher order properties of GMM and generalized empirical likelihood estimators. Econometrica 72, 219-255.

Orme, C.D., (1995). Simulated conditional moment tests. Economics Letters 49, 239-245.

Ramalho, J.J.S., (2006). Bootstrap bias-corrected GMM estimators. Economics Letters 92, 149-155.

Roodman, D., (2007). A short note on the theme of too many instruments. The Center for Global Development Working Paper Number 125.

Windmeijer, F., (2005). A finite sample correction for the variance of linear efficient two-step GMM estimators. Journal of Econometrics 126, 25-51.

Ziliak, J. P., (1997). Efficient estimation with panel data when instruments are predetermined: an empirical comparison of moment-condition estimators. Journal of Business & Economic Statistics 15, 419–431.

Table 1: Finite Sample Performance of bias-corrected Estimators in Panel AR(1) Models

Panel (A): DIF Estimators												
	$\lambda = 0.2$				$\lambda = 0.5$				$\lambda = 0.8$			
$T = 6$	*mean*	*rmsq*	*stdv*	*se*	*mean*	*rmsq*	*stdv*	*se*	*mean*	*rmsq*	*stdv*	*se*
DIF1	0.174	0.081	0.077	0.075	0.447	0.112	0.099	0.095	0.653	0.211	0.152	0.136
DIF2	0.174	0.089	0.085	0.089	0.448	0.122	0.110	0.116	0.641	0.239	0.178	0.187
DIF1-LBC	0.185	0.079	0.078	0.076	0.462	0.107	0.100	0.096	0.670	0.202	0.154	0.138
DIF2-LBC	0.186	0.087	0.086	0.090	0.462	0.117	0.111	0.117	0.657	0.229	0.180	0.189
DIF1-BBC	0.194	0.082	0.082	0.082	0.487	0.110	0.109	0.109	0.735	0.196	0.185	0.174
DIF2-BBC	0.193	0.089	0.089	0.100	0.487	0.121	0.120	0.134	0.740	0.223	0.215	0.233
$T = 8$	*mean*	*rmsq*	*stdv*	*se*	*mean*	*rmsq*	*stdv*	*se*	*mean*	*rmsq*	*stdv*	*se*
DIF1	0.179	0.063	0.059	0.056	0.457	0.084	0.072	0.068	0.686	0.154	0.104	0.093
DIF2	0.179	0.072	0.069	0.074	0.456	0.095	0.084	0.091	0.675	0.177	0.125	0.139
DIF1-LBC	0.190	0.060	0.060	0.057	0.472	0.078	0.073	0.068	0.703	0.143	0.105	0.094
DIF2-LBC	0.190	0.070	0.069	0.075	0.471	0.090	0.085	0.092	0.692	0.166	0.126	0.140
DIF1-BBC	0.196	0.063	0.063	0.061	0.489	0.079	0.078	0.077	0.749	0.133	0.123	0.116
DIF2-BBC	0.195	0.069	0.069	0.096	0.488	0.087	0.087	0.120	0.752	0.150	0.142	0.189
$T = 10$	*mean*	*rmsq*	*stdv*	*se*	*mean*	*rmsq*	*stdv*	*se*	*mean*	*rmsq*	*stdv*	*se*
DIF1	0.178	0.053	0.049	0.046	0.465	0.066	0.056	0.053	0.708	0.119	0.076	0.070
DIF2	0.178	0.059	0.055	0.080	0.464	0.074	0.065	0.093	0.705	0.131	0.089	0.130
DIF1-LBC	0.190	0.050	0.049	0.047	0.480	0.060	0.057	0.054	0.725	0.107	0.077	0.071
DIF2-LBC	0.190	0.056	0.055	0.080	0.478	0.069	0.065	0.094	0.722	0.119	0.090	0.132
DIF1-BBC	0.194	0.051	0.051	0.051	0.492	0.061	0.060	0.061	0.761	0.097	0.089	0.087
DIF2-BBC	0.194	0.055	0.055	0.614	0.492	0.066	0.065	0.746	0.769	0.102	0.097	0.932

Panel (B): SYS Estimators												
	$\lambda = 0.2$				$\lambda = 0.5$				$\lambda = 0.8$			
$T = 6$	*mean*	*rmsq*	*stdv*	*se*	*mean*	*rmsq*	*stdv*	*se*	*mean*	*rmsq*	*stdv*	*se*
SYS1	0.229	0.081	0.076	0.072	0.522	0.084	0.081	0.074	0.807	0.080	0.080	0.068
SYS2	0.215	0.068	0.066	0.075	0.516	0.077	0.075	0.082	0.803	0.077	0.077	0.079
SYS1-BBC	0.204	0.082	0.082	0.077	0.503	0.091	0.091	0.083	0.799	0.100	0.100	0.085
SYS2-BBC	0.198	0.070	0.070	0.092	0.500	0.083	0.083	0.103	0.797	0.096	0.096	0.107
$T = 8$	*mean*	*rmsq*	*stdv*	*se*	*mean*	*rmsq*	*stdv*	*se*	*mean*	*rmsq*	*stdv*	*se*
SYS1	0.227	0.064	0.059	0.057	0.520	0.066	0.063	0.058	0.807	0.064	0.063	0.054
SYS2	0.217	0.057	0.055	0.068	0.514	0.061	0.059	0.072	0.806	0.061	0.061	0.068
SYS1-BBC	0.204	0.063	0.063	0.061	0.504	0.070	0.070	0.065	0.802	0.078	0.078	0.067
SYS2-BBC	0.200	0.058	0.059	0.098	0.500	0.065	0.065	0.105	0.802	0.075	0.075	0.105
$T = 10$	*mean*	*rmsq*	*stdv*	*se*	*mean*	*rmsq*	*stdv*	*se*	*mean*	*rmsq*	*stdv*	*se*
SYS1	0.223	0.055	0.050	0.048	0.519	0.055	0.052	0.048	0.801	0.053	0.053	0.045
SYS2	0.218	0.051	0.048	0.088	0.516	0.052	0.050	0.090	0.803	0.051	0.051	0.081
SYS1-BBC	0.201	0.054	0.054	0.052	0.505	0.058	0.057	0.054	0.798	0.064	0.064	0.055
SYS2-BBC	0.201	0.055	0.055	0.158	0.503	0.058	0.058	0.162	0.799	0.065	0.065	0.148

Notes: The DGP considered is $y_{it} = \alpha_i + \lambda y_{i,t-1} + u_{it}$, $i = 1, 2, ..., N = 100; t = -48, -47, ..., T$, where $\alpha_i \sim iidN(0, \sigma_\alpha^2)$, $u_{it} \sim iidN(0, 1)$, $y_{i,-49} = 0$ and first 49 observations are discarded. We set $\sigma_\alpha^2 = \eta_1^2(1-\lambda)(1+\lambda)$ so that the long-run relative impact of α_i to u_{it} is constant, $\eta_1^2 = 4$, for different value of λ, which gives weak instruments problem. DIF1(2) is one(two)-step first-differenced GMM estimator, and SYS1(2) is one(two)-step system GMM estimator. LBC stands for linear bias-correction, namely $\hat{\lambda}_{LBC} = \hat{\lambda} + N^{-1}(1 + \hat{\lambda})$. BBC stands for bootstrap bias-correction, explained in Section 3. *mean* is average, *rmsq* is root mean square errors, *stdv* is standard deviations, and *se* is the asymptotic standard errors, which are computed based on 2000 replications except $T = 10$ which are based on 1000 replications, and 99 bootstrap samples.

Table 2: Size and Power of t-tests, DIF Estimators in Panel AR(1) Models

Panel (A): $\lambda = 0.2$								
	Size				Power			
$T = 6$	*asy*	*se(b)*	*bts1*	*bts2*	*asy*	*se(b)*	*bts1*	*bts2*
DIF1	0.069	0.081	0.055	0.055	0.823	0.839	0.755	0.713
DIF2	0.145	0.059	0.021	0.042	0.859	0.715	0.521	0.605
DIF1-LBC	0.920	0.067	–	–	0.316	0.793	–	–
DIF2-LBC	0.898	0.049	–	–	0.351	0.662	–	–
DIF1-BBC	–	0.059	–	–	–	0.706	–	–
DIF2-BBC	–	0.035	–	–	–	0.557	–	–
$T = 8$	*asy*	*se(b)*	*bts1*	*bts2*	*asy*	*se(b)*	*bts1*	*bts2*
DIF1	0.078	0.086	0.063	0.056	0.968	0.966	0.934	0.887
DIF2	0.214	0.057	0.006	0.074	0.969	0.853	0.557	0.840
DIF1-LBC	0.979	0.076	–	–	0.254	0.945	–	–
DIF2-LBC	0.961	0.048	–	–	0.334	0.806	–	–
DIF1-BBC	–	0.056	–	–	–	0.898	–	–
DIF2-BBC	–	0.014	–	–	–	0.599	–	–
$T = 10$	*asy*	*se(b)*	*bts1*	*bts2*	*asy*	*se(b)*	*bts1*	*bts2*
DIF1	0.083	0.094	0.061	0.061	0.996	0.996	0.987	0.948
DIF2	0.301	0.011	0.000	0.319	0.999	0.863	0.146	0.998
DIF1-LBC	0.997	0.074	–	–	0.225	0.990	–	–
DIF2-LBC	0.991	0.009	–	–	0.281	0.803	–	–
DIF1-BBC	–	0.057	–	–	–	0.974	–	–
DIF2-BBC	–	0.000	–	–	–	0.034	–	–

Panel (B): $\lambda = 0.5$								
	Size				Power			
$T = 6$	*asy*	*se(b)*	*bts1*	*bts2*	*asy*	*se(b)*	*bts1*	*bts2*
DIF1	0.088	0.109	0.065	0.075	0.731	0.754	0.627	0.586
DIF2	0.174	0.072	0.019	0.048	0.774	0.599	0.374	0.462
DIF1-LBC	1.000	0.086	–	–	0.976	0.702	–	–
DIF2-LBC	1.000	0.056	–	–	0.961	0.545	–	–
DIF1-BBC	–	0.055	–	–	–	0.516	–	–
DIF2-BBC	–	0.031	–	–	–	0.381	–	–
$T = 8$	*asy*	*se(b)*	*bts1*	*bts2*	*asy*	*se(b)*	*bts1*	*bts2*
DIF1	0.101	0.113	0.073	0.077	0.931	0.934	0.860	0.800
DIF2	0.252	0.057	0.008	0.069	0.951	0.771	0.408	0.727
DIF1-LBC	1.000	0.088	–	–	0.998	0.898	–	–
DIF2-LBC	1.000	0.045	–	–	0.992	0.710	–	–
DIF1-BBC	–	0.059	–	–	–	0.777	–	–
DIF2-BBC	–	0.012	–	–	–	0.410	–	–
$T = 10$	*asy*	*se(b)*	*bts1*	*bts2*	*asy*	*se(b)*	*bts1*	*bts2*
DIF1	0.106	0.118	0.072	0.063	0.988	0.987	0.970	0.914
DIF2	0.329	0.012	0.000	0.312	0.997	0.801	0.077	0.993
DIF1-LBC	1.000	0.087	–	–	1.000	0.979	–	–
DIF2-LBC	1.000	0.005	–	–	1.000	0.718	–	–
DIF1-BBC	–	0.053	–	–	–	0.929	–	–
DIF2-BBC	–	0.000	–	–	–	0.017	–	–

Table 2 (continued)

Panel (C): $\lambda = 0.8$								
	Size				Power			
$T = 6$	asy	$se(b)$	$bts1$	$bts2$	asy	$se(b)$	$bts1$	$bts2$
DIF1	0.167	0.211	0.131	0.178	0.669	0.711	0.546	0.542
DIF2	0.291	0.126	0.029	0.088	0.733	0.499	0.228	0.348
DIF1-LBC	1.000	0.181	–	–	0.994	0.669	–	–
DIF2-LBC	0.998	0.107	–	–	0.987	0.455	–	–
DIF1-BBC	–	0.066	–	–	–	0.371	–	–
DIF2-BBC	–	0.032	–	–	–	0.198	–	–
$T = 8$	asy	$se(b)$	$bts1$	$bts2$	asy	$se(b)$	$bts1$	$bts2$
DIF1	0.195	0.237	0.146	0.176	0.883	0.913	0.747	0.683
DIF2	0.395	0.107	0.013	0.120	0.926	0.684	0.214	0.552
DIF1-LBC	1.000	0.191	–	–	1.000	0.875	–	–
DIF2-LBC	1.000	0.082	–	–	0.998	0.613	–	–
DIF1-BBC	–	0.082	–	–	–	0.593	–	–
DIF2-BBC	–	0.015	–	–	–	0.207	–	–
$T = 10$	asy	$se(b)$	$bts1$	$bts2$	asy	$se(b)$	$bts1$	$bts2$
DIF1	0.225	0.274	0.148	0.174	0.975	0.977	0.888	0.792
DIF2	0.507	0.037	0.000	0.401	0.986	0.681	0.023	0.951
DIF1-LBC	1.000	0.197	–	–	1.000	0.964	–	–
DIF2-LBC	1.000	0.022	–	–	1.000	0.596	–	–
DIF1-BBC	–	0.081	–	–	–	0.789	–	–
DIF2-BBC	–	0.000	–	–	–	0.006	–	–

Notes: See notes to Table 1 for DGP and the definitions of estimators. The test is $H_0 : \lambda = \lambda_0$ against $H_0 : \lambda \neq \lambda_0$ at the 0.05 significance level. The true value of λ is as shown in the table. For size, $\lambda_0 = 0.2, 0.5, 0.8$, and for power $\lambda_0 = 0.0, 0.3, 0.6$. asy is the standard t-test referring to standard normal distribution, $se(b)$ is based on t-statistic with bootstrap variance estimator, referring to standard normal distribution, $bts1$ is the t-test referring to single bootstrap critical values, and $bts2$ is the t-test referring to double bootstrap critical values.

Table 3: Size and Power of t-test, SYS Estimators in Panel AR(1) Models

Panel (A): $\lambda = 0.2$								
	Size				Power			
$T = 6$	asy	$se(b)$	$bts1$	$bts2$	asy	$se(b)$	$bts1$	$bts2$
SYS1	0.090	0.100	0.060	0.086	0.664	0.671	0.533	0.549
SYS2	0.225	0.037	0.004	0.052	0.934	0.725	0.375	0.719
SYS1-BBC	–	0.076	–	–	–	0.716	–	–
SYS2-BBC	–	0.013	–	–	–	0.645	–	–
$T = 8$	asy	$se(b)$	$bts1$	$bts2$	asy	$se(b)$	$bts1$	$bts2$
SYS1	0.090	0.095	0.056	0.082	0.859	0.865	0.760	0.761
SYS2	0.339	0.021	0.000	0.092	0.990	0.806	0.251	0.903
SYS1-BBC	–	0.063	–	–	–	0.896	–	–
SYS2-BBC	–	0.001	–	–	–	0.557	–	–
$T = 10$	asy	$se(b)$	$bts1$	$bts2$	asy	$se(b)$	$bts1$	$bts2$
SYS1	0.107	0.109	0.065	0.085	0.954	0.955	0.914	0.889
SYS2	0.456	0.003	0.000	0.000	0.998	0.584	0.000	0.124
SYS1-BBC	–	0.073	–	–	–	0.967	–	–
SYS2-BBC	–	0.000	–	–	–	0.057	–	–

Panel (B): $\lambda = 0.5$								
	Size				Power			
$T = 6$	asy	$se(b)$	$bts1$	$bts2$	asy	$se(b)$	$bts1$	$bts2$
SYS1	0.079	0.104	0.058	0.080	0.626	0.671	0.512	0.522
SYS2	0.241	0.049	0.008	0.053	0.896	0.627	0.263	0.551
SYS1-BBC	–	0.083	–	–	–	0.650	–	–
SYS2-BBC	–	0.025	–	–	–	0.510	–	–
$T = 8$	asy	$se(b)$	$bts1$	$bts2$	asy	$se(b)$	$bts1$	$bts2$
SYS1	0.078	0.092	0.055	0.080	0.847	0.869	0.771	0.767
SYS2	0.339	0.023	0.000	0.092	0.985	0.776	0.202	0.872
SYS1-BBC	–	0.075	–	–	–	0.846	–	–
SYS2-BBC	–	0.001	–	–	–	0.478	–	–
$T = 10$	asy	$se(b)$	$bts1$	$bts2$	asy	$se(b)$	$bts1$	$bts2$
SYS1	0.087	0.097	0.061	0.077	0.952	0.961	0.914	0.900
SYS2	0.451	0.004	0.000	0.001	0.999	0.591	0.000	0.107
SYS1-BBC	–	0.077	–	–	–	0.945	–	–
SYS2-BBC	–	0.000	–	–	–	0.052	–	–

Panel (C): $\lambda = 0.8$								
	Size				Power			
$T = 6$	asy	$se(b)$	$bts1$	$bts2$	asy	$se(b)$	$bts1$	$bts2$
SYS1	0.070	0.104	0.050	0.080	0.721	0.800	0.601	0.589
SYS2	0.283	0.062	0.023	0.074	0.960	0.730	0.167	0.502
SYS1-BBC	0.000	0.093	0.000	0.000	0.000	0.648	0.000	0.000
SYS2-BBC	0.000	0.040	0.000	0.000	0.000	0.443	0.000	0.000
$T = 8$	asy	$se(b)$	$bts1$	$bts2$	asy	$se(b)$	$bts1$	$bts2$
SYS1	0.074	0.105	0.040	0.062	0.904	0.938	0.868	0.852
SYS2	0.372	0.042	0.005	0.110	1.000	0.868	0.132	0.937
SYS1-BBC	0.000	0.098	0.000	0.000	0.000	0.833	0.000	0.000
SYS2-BBC	0.000	0.011	0.000	0.000	0.000	0.443	0.000	0.000
$T = 10$	asy	$se(b)$	$bts1$	$bts2$	asy	$se(b)$	$bts1$	$bts2$
SYS1	0.068	0.105	0.041	0.054	0.973	0.987	0.949	0.947
SYS2	0.483	0.008	0.000	0.001	1.000	0.753	0.000	0.131
SYS1-BBC	0.000	0.098	0.000	0.000	0.000	0.932	0.000	0.000
SYS2-BBC	0.000	0.000	0.000	0.000	0.000	0.099	0.000	0.000

Notes: See notes to Table 1 for DGP and the definitions of estimators, and notes to Table 2 for the definition of the tests.

Table 4: Finite Sample Performance of bias-corrected DIF and SYS Estimators of $\beta = 0.5$ in a Model with a Predetermined Regressor

Panel (A): DIF Estimators					**Panel (B): SYS Estimators**				
$T = 6$	*mean*	*rmsq*	*stdv*	*se*	$T = 6$	*mean*	*rmsq*	*stdv*	*se*
DIF1	0.470	0.121	0.117	0.112	SYS1	0.536	0.099	0.093	0.084
DIF2	0.470	0.131	0.128	0.133	SYS2	0.541	0.101	0.092	0.097
DIF1-BBC	0.497	0.128	0.128	0.127	SYS1-BBC	0.512	0.105	0.105	0.096
DIF2-BBC	0.495	0.138	0.138	0.152	SYS2-BBC	0.515	0.104	0.103	0.120
$T = 8$	*mean*	*rmsq*	*stdv*	*se*	$T = 8$	*mean*	*rmsq*	*stdv*	*se*
DIF1	0.470	0.089	0.084	0.079	SYS1	0.527	0.078	0.073	0.067
DIF2	0.471	0.099	0.094	0.104	SYS2	0.534	0.082	0.075	0.088
DIF1-BBC	0.491	0.091	0.091	0.089	SYS1-BBC	0.505	0.082	0.082	0.075
DIF2-BBC	0.491	0.099	0.099	0.134	SYS2-BBC	0.507	0.083	0.083	0.127
$T = 10$	*mean*	*rmsq*	*stdv*	*se*	$T = 10$	*mean*	*rmsq*	*stdv*	*se*
DIF1	0.478	0.070	0.066	0.062	SYS1	0.526	0.066	0.061	0.056
DIF2	0.479	0.077	0.074	0.104	SYS2	0.529	0.069	0.063	0.118
DIF1-BBC	0.496	0.072	0.072	0.070	SYS1-BBC	0.506	0.068	0.068	0.063
DIF2-BBC	0.494	0.076	0.075	0.514	SYS2-BBC	0.503	0.071	0.071	0.216

Notes: The DGP considered is $y_{it} = \alpha_i + \beta x_{it} + u_{it}$, $i = 1, 2, ..., N = 100; t = 1, 2, ..., T$, where $\alpha_i \sim iidN(0, \sigma_\alpha^2)$, $\beta = 0.5$, $u_{it} \sim iidN(0, 1)$, $x_{it} = \alpha_i + 0.5x_{i,t-1} + 0.5u_{i,t-1} + v_{it}$, $t = -48, -47, ..., T$, where $v_{it} \sim iidN(0, 0.5)$, $x_{i,-49} = 0$ and first 50 observations are discarded. σ_α^2 is controlled so that the long-run relative impact of α_i to u_{it} is constant, $\eta_1^2 = 4$, which gives weak instruments problem. See also notes to Table 1 for the definitions of estimators, and notes to Table 2 for the definition of the tests.

Table 5: Size and Power, DIF and SYS Estimators in a Model with a Predetermined Regressor with $\beta = 0.5$

Panel (A): DIF Estimator								
	Size				Power			
$T = 6$	*asy*	*se*(*b*)	*bts*1	*bts*2	*asy*	*se*(*b*)	*bts*1	*bts*2
DIF1	0.057	0.068	0.043	0.056	0.512	0.547	0.464	0.483
DIF2	0.126	0.051	0.018	0.045	0.591	0.417	0.264	0.364
DIF1-BBC	–	0.051	–	–	–	0.372	–	–
DIF2-BBC	–	0.032	–	–	–	0.279	–	–
$T = 8$	*asy*	*se*(*b*)	*bts*1	*bts*2	*asy*	*se*(*b*)	*bts*1	*bts*2
DIF1	0.076	0.091	0.061	0.074	0.796	0.813	0.735	0.722
DIF2	0.187	0.046	0.008	0.076	0.856	0.610	0.299	0.638
DIF1-BBC	–	0.061	–	–	–	0.662	–	–
DIF2-BBC	–	0.014	–	–	–	0.317	–	–
$T = 10$	*asy*	*se*(*b*)	*bts*1	*bts*2	*asy*	*se*(*b*)	*bts*1	*bts*2
DIF1	0.077	0.092	0.058	0.067	0.939	0.943	0.894	0.845
DIF2	0.273	0.012	0.000	0.318	0.968	0.588	0.053	0.974
DIF1-BBC	–	0.057	–	–	–	0.832	–	–
DIF2-BBC	–	0.000	–	–	–	0.008	–	–

Panel (B): SYS Estimator								
	Size				Power			
$T = 6$	*asy*	*se*(*b*)	*bts*1	*bts*2	*asy*	*se*(*b*)	*bts*1	*bts*2
SYS1	0.099	0.116	0.078	0.092	0.445	0.491	0.318	0.321
SYS2	0.225	0.093	0.036	0.077	0.648	0.356	0.077	0.210
SYS1-BBC	–	0.085	–	–	–	0.494	–	–
SYS2-BBC	–	0.048	–	–	–	0.299	–	–
$T = 8$	*asy*	*se*(*b*)	*bts*1	*bts*2	*asy*	*se*(*b*)	*bts*1	*bts*2
SYS1	0.090	0.107	0.069	0.090	0.688	0.720	0.586	0.578
SYS2	0.307	0.060	0.005	0.115	0.872	0.450	0.059	0.526
SYS1-BBC	–	0.086	–	–	–	0.735	–	–
SYS2-BBC	–	0.009	–	–	–	0.260	–	–
$T = 10$	*asy*	*se*(*b*)	*bts*1	*bts*2	*asy*	*se*(*b*)	*bts*1	*bts*2
SYS1	0.086	0.105	0.058	0.075	0.846	0.882	0.757	0.731
SYS2	0.406	0.005	0.000	0.002	0.971	0.189	0.000	0.027
SYS1-BBC	–	0.071	–	–	–	0.874	–	–
SYS2-BBC	–	0.000	–	–	–	0.007	–	–

Notes: See notes to Table 4 for DGP, notes to Table 1 for the definitions of estimators. The test is $H_0 : \beta = \beta_0$ against $H_0 : \beta \neq \beta_0$ at the 0.05 significance level. The true value of β is 0.5. For size, $\beta_0 = 0.5$, and for power $\beta_0 = 0.3$. *asy* is the standard t-test referring to standard normal distribution, *se*(*b*) is based on t-statistic with bootstrap variance estimator, referring to standard normal distribution, *bts*1 is the t-test referring to single bootstrap critical values, and *bts*2 is the t-test referring to double bootstrap critical values.