

# Bias-corrected estimation in dynamic panel data models

Citation for published version (APA):

Bun, M. J. G., & Carree, M. A. (2002). *Bias-corrected estimation in dynamic panel data models*. METEOR, Maastricht University School of Business and Economics. METEOR Research Memorandum No. 025 <https://doi.org/10.26481/umamet.2002025>

## Document status and date:

Published: 01/01/2002

## DOI:

[10.26481/umamet.2002025](https://doi.org/10.26481/umamet.2002025)

## Document Version:

Publisher's PDF, also known as Version of record

## Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

## General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

[www.umlib.nl/taverne-license](http://www.umlib.nl/taverne-license)

## Take down policy

If you believe that this document breaches copyright please contact us at:

[repository@maastrichtuniversity.nl](mailto:repository@maastrichtuniversity.nl)

providing details and we will investigate your claim.

# Bias-corrected estimation in dynamic panel data models

MAURICE J.G. BUN AND MARTIN A. CARREE\*

*version:* September 2002

*JEL-code:* C13; C23

*Keywords:* bias correction, dynamic panel data model, unemployment dynamics

## Abstract

This study develops a new bias-corrected estimator for the fixed-effects dynamic panel data model and derives its limiting distribution for fixed  $T$  and  $N$  large. The bias-corrected estimator is derived as a bias correction of the least-squares dummy variable (within) estimator. It does not share some of the drawbacks of recently developed IV and GMM estimators and is relatively easy to compute. Monte Carlo experiments provide evidence for the bias-corrected estimator to perform well even in small samples. The paper contains an application to a model of unemployment dynamics at the U.S. state level for the 1991-2000 period.

## 1. Introduction

The estimation of fixed-effects dynamic panel data models has been one of the main challenges in econometrics during the last two decades. Various instrumental variables (IV) estimators or generalized method of moments (GMM) estimators have been proposed and compared (see e.g. Anderson and Hsiao, 1981, 1982; Arellano and Bond, 1991; Arellano and Bover, 1995; Ahn and Schmidt, 1995; Kiviet, 1995; Wansbeek and Bekker, 1996; Ziliak, 1997; Blundell and Bond, 1998, Hahn,

---

\*Maurice Bun is affiliated with University of Amsterdam. His email address is MBUN@FEE.UVA.NL. Martin Carree is affiliated with Erasmus University Rotterdam and University of Maastricht. His email address is M.CARREE@MW.UNIMAAS.NL. Martin Carree is corresponding author. He is grateful to the Royal Academy of Arts and Sciences (KNAW) for financial support.

1999; Judson and Owen, 1999). The development and comparison of such new estimators was necessary because the traditional within estimator (least-squares dummy variable estimator) is inconsistent for fixed  $T$ . Despite the increasing sophistication of the IV and GMM estimators, two important drawbacks remain. First, the complexity of the new estimators is a barrier for applied researchers (see e.g. Baltagi et al., 2000). This should partly be a temporary drawback as the new estimators will be incorporated in the statistical packages. However, the newly developed estimators may require additional decisions on, for example, which and how many instruments to use. For example, by evaluating the expectation of asymptotic expansions of estimation errors Bun and Kiviet (2002b) show that finite sample bias of GMM-estimators increases with the number of moment conditions used. This makes application less straightforward. Second, the new estimators introduce problems of their own. For example, the GMM-estimators may suffer from a substantial bias in case the autoregressive parameter becomes close to one (see Blundell and Bond, 1998; Meghir and Windmeijer, 1999; Kitazawa, 2001). Furthermore, the performance of these estimators depends strongly upon the ratio of variance of the individual specific effects and the variance of the general error term (see e.g. Kitazawa, 2001).

This paper introduces a new and simple estimator for dynamic panel data models with or without additional exogenous variables. It is computed as a bias correction to the least squares dummy variable (LSDV) estimator (also referred to as within estimator) and is, as such, related to the bias-corrected estimator developed by Kiviet (1995).<sup>1</sup> An important advantage of the estimator is that it does not depend upon the ratio of the variance of the fixed effects and the variance of the error term. The bias adjustment of the newly developed estimator is done without resorting to outside initial consistent estimates, as is the case with Kiviet's bias-corrected estimator and, in addition, appears to perform well in comparison. MacKinnon and Smith (1998) already indicate that bias of parameter estimates may be virtually eliminated in some common cases, though at the expense of increased variance of the estimators. This paper shows that this is also the case for dynamic panel data models.

---

<sup>1</sup>Judson and Owen (1999) present Monte Carlo simulation results indicating that Kiviet's bias-corrected estimator outperforms the Anderson-Hsiao IV estimator and the Arellano-Bond GMM estimators. The current paper provides further support for the usefulness of Kiviet's central idea of bias-correcting the least squares dummy variable estimator. Kiviet proposed to consistently estimate the extent of the bias by using a preliminary consistent estimator, e.g. based on generalized method of moments. This allows for a consistent corrected estimator based on *additive* bias-correction. An obvious disadvantage of the Kiviet-estimator is that its performance depends upon the preliminary estimator chosen. Hahn and Kuersteiner (2002) recently introduced a bias-corrected estimator related to that developed by Kiviet. However, their estimator is not designed for samples with small  $T$ .

The rest of the paper is organized as follows. In section 2 we explain the principle of bias-correction in dynamic panel data models. In section 3 we derive the limiting distribution for fixed  $T$  of the bias-corrected estimator. In section 4 we discuss the special case of the AR(1)-model in which no additional exogenous variables are included. In this section the bias-corrected estimator is compared with other possible corrections on the LSDV-estimator. Section 5 contains results from Monte Carlo experiments for the model with an additional exogenous regressor. In section 6 the estimators are applied to a simple model of intertemporal dynamics of the unemployment rate in U.S. states in the 1991-2000 period. In section 7 we discuss extensions and limitations of the proposed estimator in more general models and provide concluding remarks.

## 2. Bias-corrected estimation in dynamic panel data models

In this section we will illustrate the principle of bias-corrected estimation in the first-order dynamic panel data model. For ease of exposition we assume only one additional time varying regressor (next to the lagged dependent variable regressor) and the panel to be balanced. Consider the following first-order dynamic panel data model

$$y_{it} = \gamma y_{i,t-1} + \beta x_{it} + \eta_i + \varepsilon_{it}, \quad i = 1, \dots, N; \quad t = 1, \dots, T. \quad (2.1)$$

In this model the dependent variable  $y_{it}$  is determined by the one-period lagged value of the dependent variable  $y_{i,t-1}$ , the additional regressor  $x_{it}$ , the unobserved individual specific effect  $\eta_i$  and a general disturbance term  $\varepsilon_{it}$ . The regressor  $x_{it}$  may be correlated with the individual specific effect  $\eta_i$ , but we assume that it is strictly exogenous with respect to the general error term  $\varepsilon_{it}$ . Regarding the latter we assume that it has mean zero, constant variance  $\sigma_\varepsilon^2$ , finite fourth moment, not correlated either over time or across individuals and not correlated with  $\eta_i$ . Considering the start-up observations  $y_{i,0}$  we assume that they are uncorrelated with subsequent error terms  $\varepsilon_{it}$ . Finally, the model (2.1) is assumed to be dynamically stable, or  $|\gamma| < 1$ .

The unknown individual effects in (2.1) can be eliminated by expressing each variable in deviation of its individual-specific mean. We introduce  $\tilde{y}_{it} = y_{it} - \bar{y}_i$ ,  $\tilde{y}_{i,t-1} = y_{i,t-1} - \bar{y}_{i,-1}$ ,  $\tilde{x}_{it} = x_{it} - \bar{x}_i$  and  $\tilde{\varepsilon}_{it} = \varepsilon_{it} - \bar{\varepsilon}_i$  and rewrite model (2.1) as

$$\tilde{y}_{it} = \gamma \tilde{y}_{i,t-1} + \beta \tilde{x}_{it} + \tilde{\varepsilon}_{it}, \quad i = 1, \dots, N; \quad t = 1, \dots, T. \quad (2.2)$$

The LSDV-estimators are computed by applying ordinary least squares to this equation to give

$$\hat{\gamma}_{lsdv} = \frac{\sum \sum \tilde{x}_{it}^2 \sum \sum \tilde{y}_{i,t-1} \tilde{y}_{it} - \sum \sum \tilde{x}_{it} \tilde{y}_{i,t-1} \sum \sum \tilde{x}_{it} \tilde{y}_{it}}{\sum \sum \tilde{x}_{it}^2 \sum \sum \tilde{y}_{i,t-1}^2 - (\sum \sum \tilde{x}_{it} \tilde{y}_{i,t-1})^2}, \quad (2.3)$$

$$\hat{\beta}_{lsdv} = \frac{-\sum \sum \tilde{x}_{it} \tilde{y}_{i,t-1} \sum \sum \tilde{y}_{i,t-1} \tilde{y}_{it} + \sum \sum \tilde{y}_{i,t-1}^2 \sum \sum \tilde{x}_{it} \tilde{y}_{it}}{\sum \sum \tilde{x}_{it}^2 \sum \sum \tilde{y}_{i,t-1}^2 - (\sum \sum \tilde{x}_{it} \tilde{y}_{i,t-1})^2}, \quad (2.4)$$

where the summations are for  $i = 1, \dots, N$  and  $t = 1, \dots, T$ .

The LSDV-estimators of  $\gamma$  and  $\beta$  are biased and inconsistent for fixed  $T$  because of the correlation between  $\tilde{y}_{i,t-1}$  and  $\tilde{\varepsilon}_{it}$ . The extent of the inconsistency can be computed as follows. We rewrite equations (2.3) and (2.4) as

$$\hat{\gamma}_{lsdv} = \gamma + \frac{\sum \sum \tilde{x}_{it}^2 \sum \sum \tilde{y}_{i,t-1} \tilde{\varepsilon}_{it} - \sum \sum \tilde{x}_{it} \tilde{y}_{i,t-1} \sum \sum \tilde{x}_{it} \tilde{\varepsilon}_{it}}{\sum \sum \tilde{x}_{it}^2 \sum \sum \tilde{y}_{i,t-1}^2 - (\sum \sum \tilde{x}_{it} \tilde{y}_{i,t-1})^2}, \quad (2.5)$$

$$\hat{\beta}_{lsdv} = \beta - \frac{\sum \sum \tilde{x}_{it} \tilde{y}_{i,t-1} \sum \sum \tilde{y}_{i,t-1} \tilde{\varepsilon}_{it} - \sum \sum \tilde{y}_{i,t-1}^2 \sum \sum \tilde{x}_{it} \tilde{\varepsilon}_{it}}{\sum \sum \tilde{x}_{it}^2 \sum \sum \tilde{y}_{i,t-1}^2 - (\sum \sum \tilde{x}_{it} \tilde{y}_{i,t-1})^2}. \quad (2.6)$$

From equation (2.1) we use continuous substitution to obtain

$$\begin{aligned} y_{it} &= \gamma^t y_{i0} + \beta (x_{it} + \gamma x_{i,t-1} + \dots + \gamma^{t-1} x_{i1}) + \frac{1 - \gamma^t}{1 - \gamma} \eta_i + \\ &\quad + \varepsilon_{it} + \gamma \varepsilon_{i,t-1} + \dots + \gamma^{t-1} \varepsilon_{i1}. \end{aligned} \quad (2.7)$$

In order to obtain an expression for  $\tilde{y}_{i,t-1}$  we require the mean  $\bar{y}_{i,-1}$ . The sum of  $y_{i0}$  through  $y_{i,T-1}$  equals

$$\begin{aligned} \sum_{t=1}^T y_{i,t-1} &= \frac{1 - \gamma^T}{1 - \gamma} y_{i0} + \beta \left( x_{i,T-1} + \dots + \frac{1 - \gamma^{T-1}}{1 - \gamma} x_{i1} \right) + \\ &\quad + \frac{(T-1) - T\gamma + \gamma^T}{(1 - \gamma)^2} \eta_i + \varepsilon_{i,T-1} + \dots + \frac{1 - \gamma^{T-1}}{1 - \gamma} \varepsilon_{i1}. \end{aligned} \quad (2.8)$$

From this it can be derived that when  $y_{i0}$  is uncorrelated with subsequent error terms  $\varepsilon_{it}$  that

$$\text{plim}_{N \rightarrow \infty} \frac{\sum \sum \tilde{y}_{i,t-1} \tilde{\varepsilon}_{it}}{NT} = -\sigma_\varepsilon^2 \frac{(T-1) - T\gamma + \gamma^T}{T^2(1 - \gamma)^2} = -\sigma_\varepsilon^2 h(\gamma, T). \quad (2.9)$$

This expression is always negative because the function  $h(\gamma, T)$  is positive. Having  $N$  tending to infinity and using  $\text{plim}_{N \rightarrow \infty} \frac{1}{NT} \sum \sum \tilde{x}_{it} \tilde{\varepsilon}_{it} = 0$  (as the error term  $\tilde{\varepsilon}_{it}$  is assumed to be uncorrelated with  $\tilde{x}_{it}$ ), we find that the inconsistency of the LSDV coefficient estimators equals (see also Nickell, 1981, p.1424 and Kiviet, 1995, p.61)

$$\gamma^* = \text{plim}_{N \rightarrow \infty} (\hat{\gamma}_{lsdv} - \gamma) = \text{plim}_{N \rightarrow \infty} \frac{\sum \sum \tilde{y}_{i,t-1} \tilde{\varepsilon}_{it} / \sum \sum \tilde{y}_{i,t-1}^2}{1 - (\sum \sum \tilde{x}_{it} \tilde{y}_{i,t-1})^2 / (\sum \sum \tilde{x}_{it}^2 \sum \sum \tilde{y}_{i,t-1}^2)}, \quad (2.10)$$

$$\beta^* = \text{plim}_{N \rightarrow \infty} (\hat{\beta}_{lsdv} - \beta) = -\text{plim}_{N \rightarrow \infty} \frac{\sum \sum \tilde{x}_{it} \tilde{y}_{i,t-1}}{\sum \sum \tilde{x}_{it}^2} (\hat{\gamma}_{lsdv} - \gamma). \quad (2.11)$$

We introduce the following expressions of the (asymptotic) variances of  $\tilde{y}_{i,t-1}$  and  $\tilde{x}_{it}$  and their (limiting) covariance:  $\sigma_{y-1}^2 = \text{plim}_{N \rightarrow \infty} \frac{1}{NT} \sum \sum \tilde{y}_{i,t-1}^2$ ,  $\sigma_x^2 = \text{plim}_{N \rightarrow \infty} \frac{1}{NT} \sum \sum \tilde{x}_{it}^2$  and  $\sigma_{xy-1} = \text{plim}_{N \rightarrow \infty} \frac{1}{NT} \sum \sum \tilde{x}_{it} \tilde{y}_{i,t-1}$ . The inconsistency of the LSDV coefficient estimators is now conveniently expressed as

$$\gamma^* = \frac{-\sigma_\varepsilon^2 h(\gamma, T)}{(1 - \rho_{xy-1}^2) \sigma_{y-1}^2}; \quad \beta^* = -\zeta \gamma^*, \quad (2.12)$$

where  $\rho_{xy-1} = \sigma_{xy-1} / \sigma_x \sigma_{y-1}$  and  $\zeta = \sigma_{xy-1} / \sigma_x^2$  are the (asymptotic) correlation coefficient between  $\tilde{y}_{i,t-1}$  and  $\tilde{x}_{it}$  and the (asymptotic) regression coefficient of  $\tilde{y}_{i,t-1}$  on  $\tilde{x}_{it}$ . Note that the denominator  $(1 - \rho_{xy-1}^2) \sigma_{y-1}^2$  in the first equation of (2.12) is the conditional variance of  $\tilde{y}_{i,t-1}$  given  $\tilde{x}_{it}$ .

From the first expression in (2.12) it is clear that the LSDV-estimator  $\hat{\gamma}_{lsdv}$  is always downward biased. The extent of the (asymptotic) bias depends upon five parameters, viz.  $\gamma$ ,  $T$ ,  $\sigma_\varepsilon^2$ ,  $\sigma_{y-1}^2$  and  $\rho_{xy-1}^2$ . The bias of the LSDV-estimator will be especially severe in case (i) the value of  $\gamma$  is close to one; (ii) the number of time periods  $T$  is low; (iii) the ratio of variances  $\sigma_\varepsilon^2 / \sigma_{y-1}^2$  is high; (iv) the lagged endogenous variable and the exogenous variable are highly correlated, either positively or negatively. The second expression in (2.12) shows that the inconsistency of  $\hat{\beta}_{lsdv}$  is proportional to that of  $\hat{\gamma}_{lsdv}$ . The bias of the LSDV-estimator  $\hat{\beta}_{lsdv}$  can be either positive or negative dependent upon the sign of the (asymptotic) covariance between  $\tilde{y}_{i,t-1}$  and  $\tilde{x}_{it}$ .

The principle of bias-correction can be explained straightforwardly using equation (2.12). First, assume that we would know the values for  $\sigma_\varepsilon^2$ ,  $\rho_{xy-1}$ ,  $\sigma_{y-1}^2$  and  $\zeta$ . Then we may use as a bias-corrected estimator  $\hat{\gamma}_{bc}$  that value of  $\gamma$  for which

$$\hat{\gamma}_{lsdv} = \gamma - \frac{\sigma_\varepsilon^2 h(\gamma, T)}{(1 - \rho_{xy-1}^2) \sigma_{y-1}^2}. \quad (2.13)$$

This estimator can then be inserted into the second expression in (2.12) to find the bias-corrected estimator  $\hat{\beta}_{bc} = \hat{\beta}_{lsdv} + \zeta (\hat{\gamma}_{lsdv} - \hat{\gamma}_{bc})$ .<sup>2</sup>

The function  $h(\gamma) = ((T-1) - T\gamma + \gamma^T) / T^2 (1-\gamma)^2$  plays an important role in this non-linear bias-correction procedure. This function is always positive and monotonically increasing for  $|\gamma| < 1$ . Because  $h(1) = (T-1)/2T$  using l'Hôpital's rule, we have that, irrespective of the values of  $\gamma$  and  $T$  that the inconsistency  $\gamma^*$  is at most  $-\sigma_\varepsilon^2 / 2(1 - \rho_{xy-1}^2) \sigma_{y-1}^2$ . For  $T = 2$  the function  $h(\gamma)$  is equal to  $1/4$ , for

---

<sup>2</sup>The subscript *bc* means "bias-corrected". The letters *bc* also being the initials of the authors' surnames is purely coincidental.

$T = 3$  it is equal to  $(2 + \gamma)/9$ . Hence, for  $T = 2$  the bias-corrected estimator can be explicitly expressed as

$$\hat{\gamma}_{bc} = \hat{\gamma}_{lsdv} + \frac{\sigma_\varepsilon^2}{4(1 - \rho_{Xy_{-1}}^2)\sigma_{y_{-1}}^2}, \quad \text{for } T = 2. \quad (2.14)$$

For  $T = 3$  it can be explicitly expressed as

$$\hat{\gamma}_{bc} = \frac{9\hat{\gamma}_{lsdv} + 2\sigma_\varepsilon^2/(1 - \rho_{Xy_{-1}}^2)\sigma_{y_{-1}}^2}{9 - \sigma_\varepsilon^2/(1 - \rho_{Xy_{-1}}^2)\sigma_{y_{-1}}^2}, \quad \text{for } T = 3. \quad (2.15)$$

For  $T > 3$  equation (2.13) has to be solved numerically.<sup>3</sup>

In practice we do not know the values for  $\sigma_\varepsilon^2$ ,  $\rho_{xy_{-1}}$ ,  $\sigma_{y_{-1}}^2$  and  $\zeta$ . The values of the latter three variables can be estimated consistently by using their sample analogues  $\hat{\rho}_{xy_{-1}} = \hat{\sigma}_{xy_{-1}}/\hat{\sigma}_x\hat{\sigma}_{y_{-1}}$ ,  $\hat{\sigma}_{y_{-1}}^2$  and  $\hat{\zeta} = \hat{\sigma}_{xy_{-1}}/\hat{\sigma}_x^2$ . However, the LSDV-estimator of  $\sigma_\varepsilon^2$  is inconsistent and the variance of the error term can only be consistently estimated when the LSDV-estimators for  $\gamma$  and  $\beta$  have been bias-corrected. The procedure of finding bias-corrected estimates can therefore be achieved by using an iterative procedure for equation (2.13). We then substitute the LSDV-estimate for  $\sigma_\varepsilon^2$  in (2.13) to achieve 1-step estimates for  $\gamma$  and  $\beta$ . These estimators are used to compute the 1-step estimate for  $\sigma_\varepsilon^2$ . This 1-step estimate is again substituted in (2.13) to achieve 2-step estimates for  $\gamma$  and  $\beta$  and so on until convergence. An alternative procedure is to use the expression for the inconsistency of the LSDV-estimate<sup>4</sup> for  $\sigma_\varepsilon^2$

$$\begin{aligned} \text{plim}_{N \rightarrow \infty} \hat{\sigma}_{lsdv}^2 &= \text{plim}_{N \rightarrow \infty} \frac{\left( \tilde{y}_{it} - \hat{\gamma}_{lsdv} \tilde{y}_{i,t-1} - \hat{\beta}_{lsdv} \tilde{x}_{it} \right)^2}{N(T-1)} \\ &= \text{plim}_{N \rightarrow \infty} \frac{\left( (\gamma - \hat{\gamma}_{lsdv}) \tilde{y}_{i,t-1} + (\beta - \hat{\beta}_{lsdv}) \tilde{x}_{it} + \tilde{\varepsilon}_{it} \right)^2}{N(T-1)} \\ &= \sigma_\varepsilon^2 - (1 - \rho_{xy_{-1}}^2) \sigma_{y_{-1}}^2 \gamma^{*2}. \end{aligned} \quad (2.16)$$

The expression for  $\sigma_\varepsilon^2 = \hat{\sigma}_{lsdv}^2 + (1 - \rho_{xy_{-1}}^2) \sigma_{y_{-1}}^2 (\hat{\gamma}_{lsdv} - \gamma)^2$  is then substituted into equation (2.13) to arrive at an expression from which  $\hat{\gamma}_{bc}$  can be derived in one

<sup>3</sup>Equation (2.13) can be numerically solved as follows. Define  $C = \sigma_\varepsilon^2/(1 - \rho_{xy_{-1}}^2)\sigma_{y_{-1}}^2$  and take  $\hat{\gamma}_{(0)} = \hat{\gamma}_{lsdv}$ . The iterative procedure to converge towards the bias-corrected estimate (from below) is  $\hat{\gamma}_{(j+1)} = \hat{\gamma}_{lsdv} + Ch(\hat{\gamma}_{(j)}, T)$ . In case the estimate becomes one or larger during this procedure, the procedure has to be cancelled as there is no valid estimate.

<sup>4</sup>Alternatively, the denominator could be taken to be  $NT - N - 2$  to take into account that two parameters are estimated next to the  $N$  fixed effects.

step, i.e.

$$\hat{\gamma}_{lsdv} = \gamma - \frac{\hat{\sigma}_{lsdv}^2 h(\gamma, T)}{(1 - \rho_{xy_{-1}}^2) \sigma_{y_{-1}}^2} - (\hat{\gamma}_{lsdv} - \gamma)^2 h(\gamma, T). \quad (2.17)$$

Using either the iterative procedure (2.13) or (2.17) results in the same bias-corrected estimates  $\hat{\gamma}_{bc}$ ,  $\hat{\beta}_{bc}$  and  $\hat{\sigma}_{bc}^2$ .

### 3. Asymptotic properties of bias-corrected estimators

In this section we discuss the asymptotic properties of the proposed bias-corrected LSDV-estimators. We will derive consistency and asymptotic normality for the corrected estimators. We generalize the discussion to the case with  $K$  additional exogenous variables  $x_{1it}$  through  $x_{Kit}$  and use matrix notation. Stacking the observations over time, i.e.  $y_i = (y_{i1}, \dots, y_{iT})'$ ,  $y_{i,-1} = (y_{i0}, \dots, y_{i,T-1})'$ ,  $\beta = (\beta_1, \dots, \beta_K)'$ ,  $\varepsilon_i = (\varepsilon_{i1}, \dots, \varepsilon_{iT})'$  and  $X_i$  a matrix with the  $(t, k)$ -element equal to  $x_{itk}$ , we extend (2.1) to

$$y_i = \gamma y_{i,-1} + X_i \beta + \iota_T \eta_i + \varepsilon_i, \quad i = 1, \dots, N, \quad (3.1)$$

where  $\iota_T = (1, \dots, 1)'$  is a  $T \times 1$  vector of ones. Stacking once more over individuals, i.e.  $y = (y_1, \dots, y_N)'$ ,  $y_{-1} = (y_{1,-1}, \dots, y_{N,-1})'$ ,  $\eta = (\eta_1, \dots, \eta_N)'$ ,  $\varepsilon = (\varepsilon_1, \dots, \varepsilon_N)'$  and  $X = (X_1, \dots, X_N)'$ , we have the following model

$$\begin{aligned} y &= \gamma y_{-1} + X \beta + (I_N \otimes \iota_T) \eta + \varepsilon \\ &= W \delta + (I_N \otimes \iota_T) \eta + \varepsilon, \end{aligned} \quad (3.2)$$

where we have defined the  $NT \times (K + 1)$ -matrix  $W = [y_{-1}; X]$  and the  $(K + 1)$ -parameter-vector  $\delta = (\gamma, \beta)'$ . The LSDV-estimator for model (3.2) is equal to

$$\hat{\delta}_{lsdv} = (W'AW)^{-1}W'Ay = \delta + (W'AW)^{-1}W'A\varepsilon, \quad (3.3)$$

where the  $NT \times NT$  idempotent matrix  $A = I_N \otimes (I_T - \frac{1}{T} \iota_T \iota_T')$  is the within transformation matrix which eliminates the individual effects.

Define the inconsistency of the LSDV-estimator as  $\delta^* = \text{plim}_{N \rightarrow \infty} (\hat{\delta}_{lsdv} - \delta)$ . We now introduce  $\rho_{Xy_{-1}}^2$  as the multiple correlation coefficient of the regression of  $\tilde{y}_{i,t-1}$  on  $\tilde{x}_{1it}$  through  $\tilde{x}_{Kit}$  and  $\zeta = (\zeta_1, \dots, \zeta_K)$  as the corresponding vector of regression coefficients. This allows us to generalize equation (2.12) and express the inconsistency  $\delta^* = (\gamma^*, \beta^{*'})'$  as

$$\gamma^* = \frac{-\sigma_\varepsilon^2 h(\gamma, T)}{(1 - \rho_{Xy_{-1}}^2) \sigma_{y_{-1}}^2}; \quad \beta_k^* = -\zeta_k \gamma^*, \quad k = 1, \dots, K. \quad (3.4)$$



Although inconsistent, the LSDV-estimator has a limiting distribution for  $N \rightarrow \infty$  and fixed  $T$ . Bun and Kiviet (2001) derive the limiting distribution to be

$$\sqrt{N} \left( \hat{\delta}_{lsdv} - \delta^* - \delta \right) \xrightarrow[N \rightarrow \infty]{d} \mathcal{N}[0, V_X], \quad (3.5)$$

where

$$V_X = \sigma_\varepsilon^2 \Sigma_{WAW}^{-1} + \sigma_\varepsilon^4 z(\gamma, T) \Sigma_{WAW}^{-1} e_{K+1} e'_{K+1} \Sigma_{WAW}^{-1}, \quad (3.6)$$

with  $\Sigma_{WAW} = \text{plim}_{N \rightarrow \infty} \frac{1}{N} W'AW$ ,  $e_{K+1}$  the  $(K+1)$ -vector with the first element equal to one and the other elements equal to zero and  $z(\gamma, T)$  equal to<sup>5</sup>

$$z(\gamma, T) = -\frac{1 + 2\gamma^{T-1}}{(1 - \gamma)^2} + \frac{2(1 - \gamma^T)}{T(1 - \gamma)^3} + \frac{(1 - \gamma^T)^2}{T^2(1 - \gamma)^4}. \quad (3.7)$$

Using notation introduced above the variance-covariance matrix  $V_X$  of the limiting distribution (3.5) of the LSDV-estimator can be expressed as

$$V_X = \begin{pmatrix} V_X^{11} & V_X^{12} \\ V_X^{21} & V_X^{22} \end{pmatrix} = \sigma_\varepsilon^2 \Sigma_{WAW}^{-1} + \frac{\sigma_\varepsilon^4 z(\gamma, T)}{(1 - \rho_{Xy_{-1}}^2)^2 \sigma_{y_{-1}}^4} \begin{pmatrix} 1 & -\zeta' \\ -\zeta & \zeta \zeta' \end{pmatrix} \quad (3.8)$$

The result (3.5) of Bun and Kiviet (2001) shows that the LSDV-estimator has a limiting normal distribution for finite  $T$  and  $N \rightarrow \infty$ , but it is not centered at  $\delta$  and it has a non-standard asymptotic variance-covariance matrix.

We now turn to bias-corrected estimation of  $\delta = (\gamma, \beta)'$ . Generalizing the results of section 2 (see equation (2.13)), using (3.4) the bias-corrected estimator for  $\gamma$  is that  $\gamma$  which solves

$$\hat{\gamma}_{lsdv} = \gamma - \frac{\sigma_\varepsilon^2 h(\gamma, T)}{(1 - \rho_{Xy_{-1}}^2) \sigma_{y_{-1}}^2}. \quad (3.9)$$

---

<sup>5</sup>The function  $z(\gamma, T)$  is equal to  $\text{tr}(\Pi_T^2)$  where  $\Pi_T = A_T L_T \Gamma_T$  with  $A_T = I_T - \frac{1}{T} \iota_T \iota_T'$  the within transformation matrix,  $L_T = \begin{bmatrix} 0 & 0 & \cdot & \cdot & 0 & 0 \\ 1 & 0 & 0 & \cdot & \cdot & \cdot \\ 0 & 1 & 0 & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & 0 & 0 \\ 0 & 0 & \cdot & 0 & 1 & 0 \end{bmatrix}$  and  $\Gamma_T =$

$$\begin{bmatrix} 1 & 0 & \cdot & \cdot & 0 & 0 \\ \gamma & 1 & 0 & \cdot & \cdot & \cdot \\ \gamma^2 & \gamma & 1 & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & 1 & 0 \\ \gamma^{T-1} & \gamma^{T-2} & \cdot & \gamma^2 & \gamma & 1 \end{bmatrix}, \text{ see Bun and Kiviet (2001).}$$

The resulting estimator can then be inserted into the second expression in (3.4) to find the bias-corrected estimator for  $\beta$ . In short, we solve  $\hat{\delta}_{lsdv} = g(\delta)$  for  $\delta$  with

$$g(\delta) = \delta + \delta^* = \begin{pmatrix} \gamma - \sigma_\varepsilon^2 h(\gamma) / (1 - \rho_{Xy_{-1}}^2) \sigma_{y_{-1}}^2 \\ \beta + \sigma_\varepsilon^2 \zeta h(\gamma) / (1 - \rho_{Xy_{-1}}^2) \sigma_{y_{-1}}^2 \end{pmatrix}. \quad (3.10)$$

Defining  $f(\delta) = g^{-1}(\delta)$  the expression for the bias-corrected estimator is

$$\hat{\delta}_{bc} = f(\hat{\delta}_{lsdv}). \quad (3.11)$$

The function  $f$  is unknown but it can be evaluated numerically using a few lines of computer code only, see for details section 2.

From (3.11) it is seen that

$$\begin{aligned} \text{plim}_{N \rightarrow \infty} \hat{\delta}_{bc} &= \text{plim}_{N \rightarrow \infty} f(\hat{\delta}_{lsdv}) \\ &= g^{-1} \left( \text{plim}_{N \rightarrow \infty} \hat{\delta}_{lsdv} \right) \\ &= \delta, \end{aligned}$$

hence the bias-corrected estimator is a consistent estimator of  $\delta$  for finite  $T$  and  $N \rightarrow \infty$ . Furthermore, using the delta-method we have

$$\sqrt{N} \left( \hat{\delta}_{bc} - \delta \right) \xrightarrow[N \rightarrow \infty]{d} \mathcal{N} [0, FV_X F'], \quad (3.12)$$

where  $F$  is the  $(K + 1) \times (K + 1)$ -matrix of first partial derivatives of the vector function  $f$ . Hence, the bias-corrected estimator (3.11) has a limiting normal distribution, which is centered at  $\delta$ . Its asymptotic variance is depending on  $V_X$  and  $F$ . The latter matrix is simply  $F = G^{-1}$  with

$$G = \begin{pmatrix} 1 - \sigma_\varepsilon^2 h'(\gamma) / (1 - \rho_{Xy_{-1}}^2) \sigma_{y_{-1}}^2 & 0 \\ \sigma_\varepsilon^2 \zeta h'(\gamma) / (1 - \rho_{Xy_{-1}}^2) \sigma_{y_{-1}}^2 & I \end{pmatrix},$$

as the Jacobian matrix of  $g(\delta)$  and

$$h'(\gamma) = \frac{(T - 2)(1 - \gamma^T) - T\gamma(1 - \gamma^{T-2})}{T^2(1 - \gamma)^3}.$$

Using results on partitioned matrix inversion the matrix  $F = G^{-1}$  can be written as

$$F = \frac{1}{1 - \sigma_\varepsilon^2 h'(\gamma) / \sigma_{y_{-1}|X}^2} \begin{pmatrix} 1 & 0 \\ -\sigma_\varepsilon^2 \zeta h'(\gamma) / \sigma_{y_{-1}|X}^2 & (1 - \sigma_\varepsilon^2 h'(\gamma) / \sigma_{y_{-1}|X}^2) I \end{pmatrix} \quad (3.13)$$

where  $\sigma_{y_{-1}|X}^2$  is the conditional variance of  $\tilde{y}_{-1}$ , i.e.  $(1 - \rho_{Xy_{-1}}^2)\sigma_{y_{-1}}^2$ . This implies, for example, that the first diagonal element of the asymptotic variance-covariance matrix  $FV_XF'$ , or  $N * var(\hat{\gamma}_{bc})$ , is simply equal to  $V_X^{11}/(1 - \sigma_\varepsilon^2 h'(\gamma)/\sigma_{y_{-1}|X}^2)$ . The asymptotic variance-covariance matrix of  $\hat{\delta}_{bc}$  is a known function of  $\delta$ ,  $\sigma_\varepsilon^2$ , and  $\Sigma_{WAW}$ . Hence, it can be estimated consistently by  $\frac{1}{N}\hat{F}\hat{V}_X\hat{F}'$  using the bias-corrected estimators.

#### 4. Bias-correction in the panel AR(1) model

In this section we apply the limiting distribution theory of the previous section to a special case, i.e. the first-order dynamic panel data model without additional exogenous variables. We analyze the following model

$$y_{it} = \gamma y_{i,t-1} + \eta_i + \varepsilon_{it}, \quad i = 1, \dots, N; \quad t = 1, \dots, T. \quad (4.1)$$

This model is a special case of equation (3.1) where  $\beta = 0$ . An important difference with the previous sections is that we make an explicit assumption about the distribution of the initial observations  $y_{i0}$ . For the initial observations we assume that the process (4.1) has been going on for a long time, i.e.

$$y_{i0} = \frac{1}{1 - \gamma} \eta_i + \frac{\varepsilon_{i0}}{\sqrt{1 - \gamma^2}}, \quad i = 1, \dots, N, \quad (4.2)$$

with the same assumptions about  $\varepsilon_{i0}$  as for the other disturbance terms  $\varepsilon_{it}$ ,  $t = 1, \dots, T$  (see section 2). Note that this specific assumption about  $y_{i0}$  matches our earlier assumption about the initial observations made in section 2, i.e. that all  $N$  start-up observations  $y_{i0}$  are uncorrelated with all  $\varepsilon_{it}$  for  $t > 0$ . However, the additional assumptions about  $y_{i0}$  enable us to derive explicit expressions for the inconsistency of the LSDV-estimator and its asymptotic variance as a function of  $\gamma$  and  $T$ , as we will see below. This makes it possible to analytically compute and compare asymptotic efficiency of original and bias-corrected LSDV-estimators.

Stacking the observations over time and across individuals one gets

$$y = \gamma y_{-1} + (I_N \otimes \iota_T) \eta + \varepsilon. \quad (4.3)$$

Focusing on the autoregressive parameter  $\gamma$  estimation of model (4.3) by ordinary least squares yields

$$\hat{\gamma}_{lsdv} = (y'_{-1} A y_{-1})^{-1} y'_{-1} A y = \gamma + (y'_{-1} A y_{-1})^{-1} y'_{-1} A \varepsilon. \quad (4.4)$$

The inconsistency of the LSDV-estimator for  $\gamma$  when  $N$  tends to infinity has been named the "Nickell-bias" (see Nickell, 1981, Hsiao, 1986, p.74). Under the

assumptions made it can be expressed as

$$\begin{aligned}
\gamma^* &= \text{plim}_{N \rightarrow \infty} (\hat{\gamma}_{lsdv} - \gamma) \\
&= \left( \text{plim}_{N \rightarrow \infty} \frac{1}{N} (y'_{-1} A y_{-1}) \right)^{-1} \text{plim}_{N \rightarrow \infty} \frac{1}{N} (y'_{-1} A \varepsilon) \\
&= -\frac{1 + \gamma}{T - 1} \left( 1 - \frac{1 - \gamma^T}{T(1 - \gamma)} \right) \left( 1 - \frac{2\gamma}{(1 - \gamma)(T - 1)} \left( 1 - \frac{1 - \gamma^T}{T(1 - \gamma)} \right) \right)^{-1} \quad (4.5)
\end{aligned}$$

Note that the inconsistency of the LSDV-estimator is a function of  $\gamma$  for fixed  $T$ , i.e. we have  $\text{plim}_{N \rightarrow \infty} (\hat{\gamma}_{lsdv}) = \gamma^* + \gamma = g(\gamma)$  for given  $T$ . In the interval  $[-1, 1]$  the function  $g$  is a monotonically increasing function of  $\gamma$  with minimum value  $g(-1) = -1$  and maximum value  $g(1) = 1 - 3/(T + 1)$  of which the latter is computed using l'Hôpital's rule. Hence, it is possible to invert the function  $g$  and to express  $\gamma$  as a function of  $\text{plim}_{N \rightarrow \infty} (\hat{\gamma}_{lsdv})$ , i.e.  $\gamma = f(\text{plim}_{N \rightarrow \infty} (\hat{\gamma}_{lsdv}))$  with  $f = g^{-1}$ . Analogue to previous sections a consistent bias-corrected estimator can thus be constructed as

$$\hat{\gamma}_{bc} = f(\hat{\gamma}_{lsdv}). \quad (4.6)$$

For example, when  $T = 2$  we find from equation (4.5) that  $\text{plim}_{N \rightarrow \infty} (\hat{\gamma}_{lsdv}) = (\gamma - 1)/2$ . Hence, we use  $2\hat{\gamma}_{lsdv} + 1$  as bias-corrected estimator for  $\gamma$ . However, for higher values of  $T$  the function  $f$  is unknown but it can be evaluated numerically or approximated by a known function. In the latter case consistency is lost, however, due to the approximation<sup>6</sup> and, hence, we do not pursue this approach in this study.

We now turn to limiting distributions of the LSDV-estimator and the proposed bias-corrected estimator. Exploiting (3.5) the limiting distribution for  $\hat{\gamma}_{lsdv}$  for finite  $T$  and  $N$  large is

$$\sqrt{N} (\hat{\gamma}_{lsdv} - \gamma^* - \gamma) \xrightarrow[N \rightarrow \infty]{d} \mathcal{N}[0, V], \quad (4.7)$$

where  $V = \sigma_\varepsilon^2 \Sigma_{yAy}^{-1} + \sigma_\varepsilon^4 z(\gamma, T) \Sigma_{yAy}^{-2}$  and  $\Sigma_{yAy} = \text{plim}_{N \rightarrow \infty} \frac{1}{N} (y'_{-1} A y_{-1})$  and  $z(\gamma, T)$  as in (3.7). For given  $T$  this limiting distribution depends only on  $\gamma$  as the factor  $\sigma_\varepsilon^2$

---

<sup>6</sup>Carree (2002) proposed to approximate the function  $f$  by a linear specification. His estimate is easy to calculate, but requires the use of a table to obtain values for the intercept and slope. Furthermore, the estimator is inconsistent for  $N \rightarrow \infty$  due to the approximation. These properties make the estimator less appealing to use.

in  $V$  cancels out. We find

$$V = \left( \frac{T}{1-\gamma^2} - \frac{1}{(1-\gamma)^2} + \frac{2\gamma(1-\gamma^T)}{T(1-\gamma^2)(1-\gamma)^2} \right)^{-1} \quad (4.8)$$

$$\times \left( 1 + z(\gamma, T) \left( \frac{T}{1-\gamma^2} - \frac{1}{(1-\gamma)^2} + \frac{2\gamma(1-\gamma^T)}{T(1-\gamma^2)(1-\gamma)^2} \right)^{-1} \right).$$

Regarding the bias-corrected estimator (4.6) we find using (3.12)

$$\sqrt{N}(\hat{\gamma}_{bc} - \gamma) \xrightarrow[N \rightarrow \infty]{d} \mathcal{N} \left( 0, \frac{V}{(g'(\gamma))^2} \right). \quad (4.9)$$

The asymptotic variance is depending on  $V$  and the first derivative of the function  $g$ . Evaluating the latter factor analytically is cumbersome, but it can be approximated numerically. In fact, to compute the variance of  $\hat{\gamma}_{bc}$  we insert this estimate into equation (4.8) to find  $\hat{V}$ . We then approximate the first derivative of  $g$  using the expression for  $\gamma^*(\gamma)$  as given in equation (4.5) by  $g'(\hat{\gamma}_{bc}) = 1 + [\gamma^*(\hat{\gamma}_{bc}) - \gamma^*(\hat{\gamma}_{bc} - \mu)]/\mu$  with  $\mu$  a small number, say 0.001. One may also actually derive the analytic first derivative from (4.5), but this fails to be an elegant expression.

For the dynamic panel data model without additional exogenous regressors (4.1) other estimators can be used that are not consistent for fixed  $T$ , but are simple to compute being *linear* functions of the LSDV-estimator. It is interesting to compare their asymptotic efficiency with the  $\hat{\gamma}_{bc}$ -estimator. A first estimator emerges from taking a linear approximation to (4.5). When we insert in (4.5) values for  $\gamma$  equal to zero and one (using l'Hôpital's rule) we find for  $\gamma = 0$  that  $\text{plim}_{N \rightarrow \infty}(\hat{\gamma}_{lsdv}) = -1/T$  and for  $\gamma \rightarrow 1$  that  $\text{plim}_{N \rightarrow \infty}(\hat{\gamma}_{lsdv}) = 1 - 3/(T + 1)$ . A linear approximation for the function  $f$  in the  $\gamma \in [0, 1)$ -interval is found by connecting the points  $(-1/T, 0)$  and  $(1 - 3/(T + 1), 1)$ . Hence, the proposed estimator is

$$\hat{\gamma}_c = \frac{T^2 + T}{T^2 - T + 1} \hat{\gamma}_{lsdv} + \frac{T + 1}{T^2 - T + 1}. \quad (4.10)$$

The estimator in (4.10) resembles strongly an estimator proposed by Hahn and Kuersteiner (2002, p.1645), i.e.

$$\hat{\gamma}_{hk} = \frac{T + 1}{T} \hat{\gamma}_{lsdv} + \frac{1}{T}. \quad (4.11)$$

Although the estimator (4.10) is also inconsistent for finite  $T$  the leading bias term of order  $O(T^{-1})$  has been accounted for, hence this estimator may perform reasonably well for moderate  $T$ .

Each of the three estimators (4.6), (4.10) and (4.11) are functions of  $\hat{\gamma}_{lsdv}$  of which we know the limiting distribution (4.7), which is dependent on  $\gamma$  and  $T$ . This makes it possible to analytically compute asymptotic bias and variance of the estimators. These are presented in Table 1 for values of  $T$  equal to 3, 6 and 10 and values of  $\gamma$  equal to 0, 0.4 and 0.8. The bias-corrected estimator  $\hat{\gamma}_{bc}$  has (by definition) the lowest bias while the Hahn and Kuersteiner estimator has considerable bias for small  $T$ .<sup>7</sup> The latter estimator has the lowest asymptotic variance of the three estimators, though. In terms of mean squared error  $\hat{\gamma}_{bc}$  will be preferable if we have small  $T$  and  $N$  large because the extent of bias will dominate this measure for such dimensions.

## 5. Monte Carlo experiments

In this section we compare the performance of the bias-corrected estimator (3.11) (labelled *bc*) with some alternative estimators in a first-order dynamic panel model with an additional exogenous regressor. We compare it with (i) the original LSDV-estimator (*lsdv*), (ii) an additive bias-corrected estimator (*ac*)<sup>8</sup> and (iii) the so-called system GMM-estimator (*gmm*) by Blundell and Bond (1998). The GMM-estimator is consistent for finite  $T$  and  $N$  large and its favorable finite sample properties over other GMM implementations have been established by simulation (Blundell and Bond, 1998; Kitazawa, 2001), hence it is a reasonable benchmark for evaluating the various corrected LSDV variants.

The simulation design for the model with an additional explanatory variable is basically the same as in Kiviet (1995), although different values for some of the parameters have been chosen. Data for  $y$  have been generated according to equation (2.1) with  $\eta_i \sim \mathcal{N}[0, \sigma_\eta^2]$  and  $\varepsilon_{it} \sim \mathcal{N}[0, \sigma_\varepsilon^2]$ . The generating equation for the explanatory variable  $x$  is

$$x_{it} = \rho x_{i,t-1} + \xi_{it}, \quad i = 1, \dots, N; \quad t = 1, \dots, T, \quad (5.1)$$

where  $\xi_{it} \sim \mathcal{IIN}[0, \sigma_\xi^2]$ . We use three different research designs. In the first design we choose  $\beta = 1$  and  $\sigma_\eta = \sigma_\varepsilon = \sigma_\xi = 1$ . We use two different values for  $\gamma$ , viz. 0.4 and 0.8. We assume that the panel data set has 600 observations and conduct experiments for several combinations of  $T$  and  $N$  for which  $NT = 600$ . The second design is equal to the first design except that we allow for cross-sectional

---

<sup>7</sup>It can easily be shown that Monte Carlo experiments (with sizeable  $N$ ) confirm the asymptotic results, see e.g. Hahn and Kuersteiner (2002, p.1648).

<sup>8</sup>We use a slightly different version of Kiviet's (1995) estimator, i.e. there is bias correction of the first-order term only. Bun and Kiviet (2002a) show that this first-order term is responsible for the majority of the finite sample bias in the LSDV estimator. The GMM-estimate is used as the first step consistent estimate.

heteroscedasticity in the general error term  $\varepsilon_{it}$ . In this design  $\sigma_\varepsilon^2$  is varying across individuals, i.e. we specify  $\sigma_{\varepsilon,i}^2 \sim \chi_1^2$ . This implies that the mean of the variances is one. The third research design is identical to parameter settings used in Kiviet (1995, Table 1). In all his experiments the long-run effect  $\beta/(1 - \gamma)$  of  $x$  on  $y$  has been set equal to unity, hence the impact multiplier  $\beta$  varies with the chosen values for  $\gamma$ . Cross-sectional homoscedasticity is assumed and the value of  $\sigma_\varepsilon^2$  is set equal to one, but the values of the variances  $\sigma_\eta^2$  and  $\sigma_\xi^2$  differ across experiments. By varying  $\sigma_\eta^2$  the relative impact on  $y$  of the two error components  $\eta$  and  $\varepsilon$  is changed, while the parameter  $\sigma_\xi^2$  determines the signal-to-noise ratio of the model (see for details Kiviet, 1995). For each experiment we performed 1000 Monte Carlo replications.

Selected simulation results for the first, second and third design are presented in Tables 2, 3 and 4 respectively. Regarding coefficient estimators we present in these tables the bias in estimating  $\gamma$  and  $\beta$  together with the root mean squared error (RMSE). In calculating the RMSE of coefficient estimators we use the variance as estimated from the Monte Carlo as a measure of true variance. This measure has been used also when calculating the bias of variance estimators. Regarding these variance estimators<sup>9</sup> we present the relative bias in estimating the standard deviation of the various coefficient estimators, i.e. as percentage of the true standard deviation as estimated from the Monte Carlo. Finally, in the tables we indicate for each estimator for each experiment the percentage of estimates in which  $\gamma$  was estimated to be larger than one. Although we choose in all experiments  $\gamma < 1$  it can happen that estimates exceed one. Replications, in which *any* of the estimators has an estimate of  $\gamma$  larger than one, have been skipped completely. This will happen quite frequently for very small  $T$ , as we shall see.

Regarding the first design in Table 2 the results for  $\gamma = 0.8$  are presented. The results for  $\gamma = 0.4$  are similar and, hence, are deleted to save space. We observe the following patterns in the simulation results for the coefficient estimators. First, bias in estimating the autoregressive parameter  $\gamma$  is always negative for *lsdv*, while positive bias has been found for *ac* and *gmm*. Second, regarding these estimators bias in estimating both  $\gamma$  and  $\beta$  decreases for larger  $T$  (and smaller  $N$ ). This is a somewhat surprising finite sample result for *gmm*, which one should expect to perform well especially for  $T$  small and  $N$  large. Third, especially for  $\gamma$  bias in *gmm* carries over to bias in *ac* showing the detrimental effect of the dependence of additive bias-correction on preliminary consistent estimators. Fourth, in estimating both  $\gamma$  and  $\beta$  *bc* is virtually unbiased. Finally, based on a mean squared error criterion *bc* is never beaten by the other coefficient estimators.

---

<sup>9</sup>Regarding the *lsdv* and *ac* estimators we use the standard variance expression, i.e.  $\hat{\sigma}_\varepsilon^2(W'AW)^{-1}$ . For *bc* the expression in (3.12) has been used, while for *gmm* so-called one-step estimates have been exploited.

Regarding the variance estimators we observe that for larger  $T$  relative bias in estimating standard deviations is less than around 10%. However,  $gmm$  is found to have substantial relative bias in estimating standard errors when  $T = 2$ .

For the second design with cross-sectional heteroscedasticity the simulation results are presented in Table 3. Again we show the results for  $\gamma = 0.8$  only. In general, results for coefficient estimators are similar to the case of homoscedasticity. This is a somewhat surprising result because bias-corrected estimators are not consistent in case of heteroscedasticity. Nevertheless, violation of this assumption does not seem to have substantial impact on the finite sample properties of coefficient estimators. Regarding variance estimators, however, the performance of most estimators deteriorate, especially those in estimating standard errors of coefficient estimators for  $\gamma$ . This result is not surprising as the asymptotic results of the previous sections are not valid in case of cross-sectional heteroscedasticity.

Finally, we turn to simulation results using the third design, i.e. the parametrizations used by Kiviet (1995). The simulation results for a selection of parametrizations are presented in Table 4. The first panel of this table gives the parametrizations used. Several points can be made before discussing the simulation results. First, the relative impact on  $y$  of the two error components  $\eta$  and  $\varepsilon$  is 1 in experiments I-VIII, but raised to 5 in IX and X. Hence, the individual specific effect is relatively strong in IX and X. Second, the signal-to-noise ratio corresponds with an expected  $R^2$  of 2/3 in all experiments except VIII and X where it is raised to 8/9.

Regarding the third design we observe the following patterns for the coefficient estimators. First, except for  $ac$  bias in estimating the autoregressive parameter  $\gamma$  increases with  $\gamma$  for all estimators. Second,  $lsdv$  and  $bc$  are invariant with respect to the relative strength of the error components, but  $gmm$  and  $ac$  are not. Increasing the relative strength of the individual-specific effect leads to large biases in both  $gmm$  and  $ac$ . Third, increasing the signal-to-noise ratio leads in most cases to more accurate estimates. Fourth, there is no estimation method with lowest RMSE across all parametrizations. The bias-corrected estimator performs very well in all of the designs *except* for designs III and VI. In the latter two designs there is a relatively high value for  $\gamma$  combined with a relatively low signal-to-noise ratio. The bias-corrected estimator fails to converge to an estimate less than one in about one quarter of the simulations in these two designs. Regarding the variance estimators in the third design we observe that relative bias in estimating standard deviations is in many cases less than around 10%.

Summarizing, regarding coefficient estimators we find large bias for  $lsdv$ , moderate bias for  $ac$  and  $gmm$  and little bias for  $bc$ . Regarding variance estimators we find small to moderate biases for all estimators. Finally, the invariance of  $bc$  to the individual specific effects seems to be an important advantage over  $ac$  and  $gmm$ .



The Monte Carlo results do not suggest that there is one estimation technique superior for all parameter combinations. Hence, in empirical applications it may be advisable to compare results using different (consistent) estimation techniques.

## 6. Empirical application: unemployment dynamics at the U.S. state level

In this section we apply the bias-corrected estimation procedure (labelled *bc*) to a model of unemployment dynamics at the U.S. state level. We compare the coefficient estimates and estimated standard errors with those of the LSDV-, additive bias-corrected LSDV-, and GMM(SYS)-estimators (labelled *lsdv*, *ac* and *gmm* respectively). The model relates the current unemployment rate ( $U_{it}$ ) to the unemployment rate and economic growth rate ( $G_{it}$ ) in the previous year. The model has fixed effects ( $\eta_i$ ) included and is as follows

$$U_{it} = \eta_i + \gamma U_{i,t-1} + \beta G_{i,t-1} + \varepsilon_{it}. \quad (6.1)$$

Equation (6.1) can be rewritten in a form which is more easy to interpret, i.e.

$$\Delta U_{it} = (\gamma - 1)(U_{i,t-1} - \alpha_i) + \beta(G_{i,t-1} - \delta) + \varepsilon_{it}, \quad (6.2)$$

where  $(1 - \gamma)\alpha_i - \beta\delta = \eta_i$ . Equation (6.2) indicates that the change in the unemployment rate is determined by an adjustment of the unemployment rate towards a "natural" or "equilibrium" rate of unemployment  $\alpha_i$  which may be different across the states and by the previous economic growth rate. The speed of adjustment of the unemployment rate towards the "natural" or "equilibrium" rate is equal to  $1 - \gamma$ . It is to be expected that there is partial adjustment:  $1 > \gamma > 0$ . A state which shows relatively high economic growth is more likely to show reduced unemployment rates when compared to states in which the economy is growing more slowly. This would imply  $\beta < 0$ .

The data for the unemployment rate for the 1991-2000 period are from the U.S. Bureau of Labor Statistics and data for the (current dollar) gross state product are from the U.S. Bureau of Economic Analysis. The economic growth rate is taken to be the relative growth of the gross state product. Data are available for all U.S. states and Washington D.C. ( $N=51$ ). The number of time periods in estimation is  $T = 9$  because the year 1991 is taken as starting observation.

The various coefficient estimates and their estimated standard deviations are presented in Table 5. The value of the LSDV-estimate of  $\gamma$  is 0.805 which would imply an adjustment rate of almost 20% per year. In contrast, the bias-corrected estimate (*bc*) is equal to 0.943 which implies an adjustment rate of less than 6%. Hence, the speed of adjustment towards a "natural rate of unemployment"

is not as large as the original LSDV-estimator would suggest. The value of the LSDV-estimate of  $\beta$  equals -0.083 while the value of the bias-corrected estimate is -0.071. It implies a somewhat smaller effect of economic growth on the change in unemployment than indicated by the traditional within estimate. The results for additive bias-corrected estimator (*ac*) are not very different from those of the bias-corrected estimator introduced in this paper. However, the results for the GMM-estimator (*gmm*) are remarkably different. The GMM-estimates are estimated to have much lower standard deviations than the other estimators.

A restrictive assumption of bias-corrected LSDV-estimators is that consistency depends on strict exogeneity of the lagged growth rate ( $G_{i,t-1}$ ). As we have assumed strict exogeneity of  $G_{i,t-1}$  in GMM-estimation we can test against exogeneity using the Sargan test. It has a value of 46.90 (p-value is 0.23), hence the validity of the moment conditions is not rejected. Hence, we conclude that a problem of  $G_{i,t-1}$  being only *weakly* exogenous is not an issue in this particular application.

## 7. Extensions and concluding remarks

This paper develops a new bias-corrected estimator for dynamic panel data models. The proposed estimator has desirable asymptotic properties for finite  $T$  and  $N$  large but they have been derived under certain restrictive assumptions: (1) strict exogeneity of regressors in  $x_{it}$ ; (2) homoscedasticity of the disturbances; (3) balanced panels. In this final section we will discuss the limitations and possible extensions of our approach with respect to each of these assumptions.

Regarding the exogeneity assumption, some regressors in  $x_{it}$  could be predetermined as well. Inconsistencies originating from this source is not accounted for in the current bias corrections. It can be shown that the order of magnitude of such inconsistency terms is equal to that of lagged dependent variable regressors, i.e. of order  $O(T^{-1})$ . Addressing the importance of this source of bias, however, requires a full specification of the marginal process of the regressors  $x_{it}$ , which is a major complication in practice. Simulation evidence for the dynamic panel data model with predetermined or endogenous regressors  $x_{it}$  can be found in Bun and Kiviet (2002b) and Blundell et al. (2000) respectively. In general, these simulation results show that lack of strict exogeneity of  $x_{it}$  does influence the finite sample properties of estimators and, hence, it is expected that in practice estimators will be affected also. Note, however, that in the current application on unemployment dynamics strict exogeneity of the additional regressor (lagged growth rate) is not rejected.

Regarding homoscedasticity of the disturbances we provided some simulation results allowing for cross-sectional heteroscedasticity. From the simulation results

it is seen that coefficient estimators behave reasonably well, while variance estimators are biased. The bias-corrected estimator may be extended to incorporate some types of heteroskedasticity, e.g. heteroskedasticity over time or heteroskedasticity as a function of the exogenous variables.

Finally, the proposed method in this study can be extended to unbalanced panels. In this case not all time observations are available for each individual  $i$ . That is, the data may be observed for certain individuals  $i$  only from a certain date on or the data may be observed for other individuals only up till a certain date. It implies that the starting date and ending date of the data are individual-specific. We denote the begin of the data period by  $B_i$  and the final time period of observation by  $T_i$ . That is, we have  $1 \leq B_i \leq T_i \leq T$ . The individuals are then ordered in terms of the length of the time period  $T_i - B_i + 1$ . The largest value for this length of time period is  $T$  while the smallest value is 2. Denote by  $\varphi_t$  the fraction of individuals with period of time length  $t = 2, \dots, T$ . That is,  $\sum_{t=2}^T \varphi_t = 1$ . Then we should replace the function  $h(\gamma)$  in sections 2 and 3 with

$$h_u(\gamma) = \sum_{t=2}^T \varphi_t \frac{(t-1) - t\gamma + \gamma^t}{t^2 (1-\gamma)^2} \quad (7.1)$$

and likewise derive expressions for the limiting distribution of the estimator. Note that we do not take possible sample selection issues into account in this way.<sup>10</sup>

The bias-corrected estimator has been shown to perform well when  $T$  is small and  $N$  is large. Simulation results on various designs show that based on a root mean squared criterion bias-corrected LSDV-estimators perform well against system GMM-estimators. An important advantage of the bias-corrected estimator developed in this study is its invariance with respect to the individual specific effects. Varying the relative strength of the error components in the simulations clearly shows the potential detrimental effects on the accuracy of estimators that do not share this property. In case both  $T$  and  $N$  are small the limiting distributions for the estimators may have little to say about the actual distribution (especially when  $\gamma$  is close to unity). However, given the strong (relative) performance of the bias-corrected estimator in the Monte Carlo exercises in case  $T$  is as small as two or three, it appears suitable for research efforts with samples with a large number of individuals/firms and a (very) small number of time pe-

---

<sup>10</sup>Research into problems of sample selection in *dynamic* panel data models has started only recently. Kyriazidou (2001) has provided a first contribution in this area. However, the selection equation in that paper does not contain the lagged continuous endogenous variable or other predetermined variables, limiting its applicability.

riods. Many data sets, especially when data is collected yearly, have these panel dimensions.

Recently, new estimators for the dynamic panel data model have been introduced. Each of these estimators have their advantages and disadvantages and it is not clear that any of these would uniformly outperform the bias-corrected estimator. Hahn and Kuersteiner (2002) introduced an estimator that requires the number of time periods to be at least moderate. Their paper also pays most of its attention to the case of no exogenous variables. Pesaran et al. (2002) introduced a (computationally complicated) maximum likelihood estimator which is based upon first differencing the dynamic panel data model to get rid of the unobserved individual effects. Methods based upon first differencing are conceptually different from methods based upon removing unobserved effects by subtracting the individual-specific means.<sup>11</sup> Alvarez and Arellano (1998) introduced a random effects maximum likelihood estimator. However, they do not consider the case of exogenous variables included and they assume in deriving the asymptotic distribution that both  $N$  and  $T$  tend to infinity. Finally, Lancaster (2001) takes a Bayesian approach to dynamic panel data models. He finds a relatively simple set of  $K + 2$  first-order conditions for the maximum of the posterior. However, the paper still has some unresolved issues concerning priors and its inference may not be completely comparable to the classic inference used in the current paper. Nevertheless, research into these and other newly developed estimators for dynamic panel data models remains a very vivid and important area for theoreticians and practitioners.

## References

- Ahn, S.C. and P. Schmidt (1995), Efficient estimation of models for dynamic panel data, *Journal of Econometrics* 68, 29-52.
- Alvarez, J. and M. Arellano (1998), The time series and cross-section asymptotics of dynamic panel data estimators, CEMFI working paper No. 9808, December 1998.
- Anderson, T.W. and C. Hsiao (1981), Estimation of dynamic models with error components, *Journal of the American Statistical Association* 76, 598-606.

---

<sup>11</sup>One potential source of distinction between methods based on either first difference or within transformations are the influence of measurement errors. Mairesse et al. (1999), for example, argue that biases from random measurement errors are more severe in case of first differenced estimates compared to within estimates. In addition, using first differences implies that the first time observation is lost for each cross-sectional unit. This is an important disadvantage when  $T$  is (very) small.

- Anderson, T.W. and C. Hsiao (1982), Formulation and estimation of dynamic models using panel data, *Journal of Econometrics* 18, 47-82.
- Arellano, M. and S.R. Bond (1991), Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations, *Review of Economic Studies* 58, 277-297.
- Arellano, M. and O. Bover (1995), Another look at the instrumental variable estimator of error-components models, *Journal of Econometrics* 68, 29-52.
- Baltagi, B.H., J.M. Griffin and W. Xiong (2000), To pool or not to pool: Homogeneous versus heterogeneous estimators applied to cigarette demand, *Review of Economics and Statistics* 82, 117-126.
- Blundell, R. and S. Bond (1998), Initial conditions and moment restrictions in dynamic panel data models, *Journal of Econometrics* 87, 115-143.
- Blundell, R., S. Bond and F. Windmeijer (2000), Estimation in dynamic panel data models: improving on the performance of the standard GMM estimators, in B. Baltagi (ed.), *Nonstationary Panels, Panel Cointegration, and Dynamic Panels*, Advances in Econometrics 15, Amsterdam: JAI Press, Elsevier Science.
- Bun, M.J.G. and J.F. Kiviet (2001), The accuracy of inference in small samples of dynamic panel data models, Tinbergen Institute Discussion Paper 2001-006/4, Tinbergen Institute Amsterdam.
- Bun, M.J.G. and J.F. Kiviet (2002a), On the diminishing returns of higher-order terms in asymptotic expansions of bias, *Economics Letters*, forthcoming.
- Bun, M.J.G. and J.F. Kiviet (2002b), The effects of dynamic feedbacks on LS and MM estimator accuracy in panel data models, mimeo, University of Amsterdam.
- Carree, M.A. (2002), Nearly Unbiased Estimation in Dynamic Panel Data Models, Tinbergen Institute Discussion Paper 2002-008/2, Tinbergen Institute Rotterdam.
- Hahn, J. (1999), How informative is the initial condition in the dynamic panel model with fixed effects?, *Journal of Econometrics* 93, 309-326.
- Hahn, J. and G. Kuersteiner (2002), Asymptotically unbiased inference for a dynamic panel model with fixed effects when both  $n$  and  $T$  are large, *Econometrica* 70, 1639-1657.
- Hsiao, C. (1986), *Analysis of Panel Data*, Cambridge: Cambridge University Press.
- Hsiao, C., M.H. Pesaran and A.K. Tahmiscioglu (2002), Maximum likelihood estimation of fixed effects dynamic panel data models covering short time periods, *Journal of Econometrics* 109, 107-150.
- Judson, R.A. and A.L. Owen (1999), Estimating dynamic panel data models: a guide for macroeconomists, *Economics Letters* 65, 9-15.
- Kitazawa, Y. (2001), Exponential regression of dynamic panel data models, *Economics Letters* 73, 7-13.

- Kyriazidou, E. (2001), Estimation of dynamic panel data sample selection models, *Review of Economic Studies* 68, 543-572.
- Kiviet, J.F. (1995), On bias, inconsistency, and efficiency of various estimators in dynamic panel data models, *Journal of Econometrics* 68, 53-78.
- Lancaster, T. (2001), Orthogonal parameters and panel data, mimeo, Brown University, Providence, RI.
- MacKinnon, J.G. and A.A. Smith, Jr. (1998), Approximate bias correction in econometrics, *Journal of Econometrics* 85, 205-230.
- Mairesse, J., B.H. Hall and B. Mulkay (1999), Firm-level investment in France and the United States: an exploration of what we have learned in twenty years, *Annales d'Économie et de Statistique* 55-56, 27-67.
- Meghir, C. and F. Windmeijer (1999), Moment conditions for dynamic panel data models with multiplicative individual effects in the conditional variance, *Annales d'Économie et de Statistique* 55-56, 317-330.
- Nickell, S. (1981), Biases in dynamic models with fixed effects, *Econometrica* 49, 1417-1426.
- Wansbeek, T. and P. Bekker (1996), On IV, GMM and ML in a dynamic panel data model, *Economics Letters* 51, 145-152.
- Ziliak, J.P. (1997), Efficient estimation with panel data when instruments are pre-determined: An empirical comparison of moment-condition estimators, *Journal of Business and Economic Statistics* 15, 419-431.

Table 1: Asymptotic bias and variance for the panel AR(1) model

T	$\gamma$	$\gamma^*$	$g'$	V	$N^*\text{var}(\hat{\gamma}_{bc})$	$\text{bias}(\hat{\gamma}_c)$	$N^*\text{var}(\hat{\gamma}_c)$	$\text{bias}(\hat{\gamma}_{hk})$	$N^*\text{var}(\hat{\gamma}_{hk})$
3	0	-0.333	0.611	0.611	1.637	0	1.796	-0.111	1.086
3	0.4	-0.494	0.587	1.427	4.141	0.010	4.194	-0.192	2.537
3	0.8	-0.663	0.569	14.61	45.13	0.006	42.94	-0.284	25.97
6	0	-0.167	0.811	0.188	0.286	0	0.345	-0.028	0.256
6	0.4	-0.251	0.762	0.202	0.348	0.028	0.371	-0.059	0.275
6	0.8	-0.361	0.684	0.535	1.144	0.020	0.982	-0.121	0.728
10	0	-0.100	0.891	0.104	0.131	0	0.152	-0.010	0.126
10	0.4	-0.148	0.864	0.093	0.125	0.026	0.136	-0.023	0.113
10	0.8	-0.218	0.768	0.087	0.148	0.024	0.127	-0.060	0.105

Note: The asymptotic bias( $\hat{\gamma}_{bc}$ ) is always equal to zero. The value for V is  $N^*\text{var}(\hat{\gamma}_{lsdv})$ .

Table 2: Cross-sectional homoscedasticity,  $\gamma = 0.8$  and  $\beta = 1$

$(N, T)$	(300, 2)	(200, 3)	(150, 4)	(100, 6)	(60, 10)	(40, 15)
	% $\gamma$ estimates outside unit interval					
lsdv	0	0	0	0	0	0
ac	44	8	0	0	0	0
bc	1	0	0	0	0	0
gmm	9	3	0	0	0	0
	bias $\gamma$					
lsdv	-0.377	-0.215	-0.143	-0.080	-0.037	-0.021
ac	0.100	0.088	0.060	0.032	0.015	0.008
bc	-0.014	-0.001	0.001	-0.000	0.000	-0.000
gmm	0.064	0.028	0.024	0.019	0.016	0.012
	RMSE $\gamma$					
lsdv	0.379	0.218	0.146	0.083	0.041	0.025
ac	0.113	0.101	0.069	0.038	0.022	0.015
bc	0.060	0.042	0.032	0.024	0.017	0.014
gmm	0.089	0.081	0.069	0.049	0.035	0.026
	bias $\beta$					
lsdv	-0.096	-0.031	-0.003	0.014	0.022	0.017
ac	0.032	0.014	0.001	-0.008	-0.007	-0.009
bc	0.002	0.001	0.000	-0.002	0.001	-0.001
gmm	-0.116	-0.048	-0.038	-0.031	-0.023	-0.021
	RMSE $\beta$					
lsdv	0.119	0.067	0.051	0.047	0.042	0.037
ac	0.088	0.063	0.053	0.046	0.036	0.034
bc	0.078	0.059	0.052	0.045	0.036	0.033
gmm	0.165	0.139	0.119	0.088	0.067	0.056
	% bias std $\gamma$					
lsdv	17.74	-2.25	-5.89	-9.97	-11.63	-7.89
ac	-1.06	-25.97	-16.60	-0.73	-0.56	-2.19
bc	-8.94	-10.55	-7.14	-7.35	-5.73	-4.83
gmm	122.64	20.59	2.73	0.07	-3.56	-1.27
	% bias std $\beta$					
lsdv	1.05	-3.94	-2.32	-4.41	3.20	3.80
ac	2.47	1.81	-0.05	-1.48	5.19	5.48
bc	1.30	0.90	-0.14	-2.27	3.53	3.56
gmm	111.30	21.59	2.38	-1.37	-2.66	1.35

Note: For the variances we assume  $\sigma_\varepsilon^2 = \sigma_\eta^2 = \sigma_\xi^2 = 1$

Table 3: Cross-sectional heteroscedasticity,  $\gamma = 0.8$  and  $\beta = 1$

$(N, T)$	(300, 2)	(200, 3)	(150, 4)	(100, 6)	(60, 10)	(40, 15)
	% $\gamma$ estimates outside unit interval					
lsdv	0	0	0	0	0	0
ac	54	13	0	0	0	0
bc	2	0	0	0	0	0
gmm	13	6	2	0	0	0
	bias $\gamma$					
lsdv	-0.383	-0.215	-0.144	-0.080	-0.037	-0.021
ac	0.093	0.092	0.064	0.034	0.015	0.008
bc	-0.015	-0.001	0.000	0.000	-0.000	-0.000
gmm	0.070	0.026	0.017	0.015	0.006	0.006
	RMSE $\gamma$					
lsdv	0.386	0.219	0.148	0.084	0.041	0.025
ac	0.112	0.106	0.076	0.042	0.023	0.016
bc	0.077	0.049	0.038	0.026	0.018	0.014
gmm	0.097	0.086	0.081	0.056	0.037	0.028
	bias $\beta$					
lsdv	-0.103	-0.030	-0.005	0.015	0.020	0.020
ac	0.027	0.014	0.001	-0.007	-0.009	-0.006
bc	-0.002	0.001	-0.001	-0.001	-0.000	0.002
gmm	-0.127	-0.043	-0.027	-0.022	-0.009	-0.008
	RMSE $\beta$					
lsdv	0.126	0.063	0.052	0.047	0.044	0.041
ac	0.089	0.062	0.055	0.046	0.040	0.036
bc	0.080	0.057	0.053	0.045	0.039	0.035
gmm	0.177	0.148	0.141	0.099	0.074	0.063
	% bias std $\gamma$					
lsdv	-9.50	-23.73	-26.84	-24.19	-18.81	-11.61
ac	-15.90	-32.85	-32.00	-16.90	-12.80	-7.98
bc	-28.53	-23.88	-20.63	-15.37	-11.30	-5.67
gmm	166.69	35.11	4.69	0.58	-6.67	-8.96
	% bias std $\beta$					
lsdv	-1.51	1.13	-3.03	-4.76	-6.22	-4.54
ac	-0.51	4.26	-2.53	-2.01	-4.06	-2.75
bc	-0.80	3.86	-2.05	-2.90	-5.53	-3.76
gmm	156.58	33.98	3.03	-0.99	-6.98	-8.17

Note: For the variances we assume  $\sigma_{\varepsilon,i}^2 \sim \chi_1^2$ ,  $\sigma_\eta^2 = \sigma_\xi^2 = 1$



Table 4: Simulation results for selected designs from Kiviet (1995)

design nr.	I	II	III	IV	V	VI	VII	VIII	IX	X
$T$	6	6	6	6	6	6	3	3	3	3
$\gamma$	0	0.4	0.8	0	0.4	0.8	0.4	0.4	0.4	0.4
$\rho$	0.8	0.8	0.8	0.99	0.99	0.99	0.8	0.8	0.8	0.8
$\sigma_\eta$	1	0.6	0.2	1	0.6	0.2	0.6	0.6	3	3
$\sigma_\xi$	0.85	0.88	0.4	0.2	0.19	0.07	0.88	1.84	0.88	1.84
% $\gamma$ estimates outside unit interval										
lsdv	0	0	0	0	0	0	0	0	0	0
ac	0	0	0	0	0	0	0	0	49	9
bc	0	0	24	0	0	25	1	0	1	0
gmm	0	0	0	0	0	0	0	0	3	4
bias $\gamma$										
lsdv	-0.104	-0.177	-0.367	-0.164	-0.249	-0.376	-0.381	-0.215	-0.374	-0.214
ac	0.042	0.034	0.003	0.066	0.047	0.002	0.081	0.052	0.415	0.331
bc	-0.001	0.001	-0.033	-0.002	-0.001	-0.037	0.007	0.002	0.008	0.004
gmm	0.001	-0.007	-0.029	0.002	-0.006	-0.029	-0.008	-0.012	0.329	0.240
RMSE $\gamma$										
lsdv	0.110	0.181	0.370	0.169	0.253	0.378	0.386	0.221	0.379	0.219
ac	0.057	0.054	0.040	0.080	0.067	0.040	0.106	0.074	0.430	0.350
bc	0.039	0.047	0.081	0.049	0.060	0.084	0.111	0.063	0.111	0.063
gmm	0.062	0.077	0.081	0.064	0.074	0.081	0.135	0.154	0.345	0.295
bias $\beta$										
lsdv	0.043	0.045	0.011	0.081	0.090	0.025	0.014	0.008	0.014	0.008
ac	-0.021	-0.011	-0.003	-0.045	-0.027	-0.027	-0.003	-0.002	-0.025	-0.012
bc	-0.002	-0.002	-0.001	-0.008	-0.009	-0.025	0.000	0.000	-0.005	-0.001
gmm	-0.002	0.003	0.008	-0.004	0.004	0.021	0.004	0.006	-0.180	-0.131
RMSE $\beta$										
lsdv	0.070	0.069	0.118	0.238	0.254	0.658	0.090	0.044	0.090	0.044
ac	0.059	0.052	0.113	0.217	0.224	0.621	0.092	0.044	0.111	0.053
bc	0.055	0.052	0.112	0.215	0.224	0.619	0.090	0.043	0.090	0.044
gmm	0.071	0.062	0.088	0.112	0.101	0.172	0.094	0.089	0.213	0.173
% bias std $\gamma$										
lsdv	-0.26	-4.11	-1.67	8.68	4.70	-1.04	0.78	0.06	-1.04	-0.52
ac	-6.16	-7.84	6.43	-4.08	-3.77	7.68	5.09	0.66	-25.71	-45.31
bc	-6.27	-9.37	-10.60	-3.94	-9.47	-10.15	-25.58	-9.67	-23.59	-7.72
gmm	-4.35	-5.03	-1.90	-3.37	-2.86	-2.23	-0.73	0.59	43.10	33.93
% bias std $\beta$										
lsdv	-3.71	-6.08	-14.10	-7.33	-10.37	-14.15	-0.54	2.09	-0.54	2.54
ac	-2.30	-1.51	-3.05	0.77	0.12	-1.10	7.99	7.24	10.36	8.81
bc	-4.39	-4.33	-4.30	-2.15	-2.15	-2.38	7.00	6.70	6.87	6.00
gmm	-3.75	-4.62	-2.40	-2.35	-3.05	-2.61	0.38	1.23	12.03	22.21

Note: We assume  $\sigma_\varepsilon^2 = 1$ ,  $N = 100$  and  $\beta = 1 - \gamma$  in all experiments.

*Table 5:* Empirical results for the unemployment-growth model

	lsdv	ac	bc	gmm
$\hat{\gamma}$	0.805	0.920	0.943	0.892
sd( $\hat{\gamma}$ )	0.027	0.028	0.037	0.017
$\hat{\beta}$	-0.083	-0.073	-0.071	-0.070
sd( $\hat{\beta}$ )	0.014	0.015	0.015	0.010

Note: The number of time periods is 9 and the number of U.S. states is 51.