Dynamic Panel Data Models Featuring Endogenous Interaction and Spatially Correlated Errors

Jan P.A.M. Jacobs
University of Groningen
[j.p.a.m.jacobs@rug.nl]

Jenny E. Ligthart*
Georgia State University and Tilburg University
[j.ligthart@uvt.nl]

Hendrik Vrijburg
Erasmus University Rotterdam
[vrijburg@few.eur.nl]

June 1, 2009

Abstract

We extend the three-step Generalized Methods of Moments (GMM) approach of Kapoor, Kelejian, and Prucha (2007), which corrects for spatially correlated errors in static panel data models, by introducing a spatial lag and a one-period lag of the endogenous variable as additional explanatory variables. Combining the extended KKP approach with the dynamic panel data model GMM estimators of Arellano and Bond (1991) and Blundell and Bond (1998) and supplementing the dynamic instruments by lagged and weighted exogenous variables as suggested by Kelejian and Robinson (1993) yields new spatial dynamic panel data estimators. The performance of these spatial dynamic panel data estimators is investigated by means of Monte Carlo simulations. We show that differences in bias as well as root mean squared error between spatial GMM estimates and corresponding GMM estimates in which spatial error correlation is ignored are small.

JEL codes: C15, C21, C22, C23

Keywords: Dynamic panel models, spatial lag, spatial error, GMM estimation

---

*Corresponding author: CentER and Department of Economics, Tilburg University, P.O. Box 90153, 5000 LE Tilburg, The Netherlands. Phone: +31-13-466-8755, Fax: +31-13-466-4032.

1 Introduction

The literatures on dynamic panel data models and spatial econometric models have matured rapidly and have reached (graduate) textbooks during the last decade.[1] Panel data may feature state dependence—that is, the dependent variable is correlated over time—as well as display spatial dependence, that is, the dependent variable is correlated in space. Applied economists' interest in frameworks that integrate spatial considerations into dynamic panel data models is a fairly recent development, however.[2] Elhorst (2003, 2008a,b) and Yu, De Jong, and Lee (2007) have analyzed the properties of Maximum Likelihood (ML) estimators and combinations of ML and Corrected Least Squares Dummy Variables (CLSDV) for this model class. Elhorst (2008b) briefly touches upon spatial Generalized Methods of Moments (GMM) estimators in comparison to ML estimators. The properties of spatial Generalized Methods of Moments (GMM) estimators have not been comprehensively studied in a dynamic panel data context yet.[3] This paper fills this void by comparing the performance of various spatial GMM estimators of dynamic panel data models with fixed effects.

Spatial panel data applications typically employ either a spatial lag model or a spatial error model. Many economic interactions among agents are characterized by a spatially lagged dependent variable, which consists of observations on the dependent variable in other locations than the "home" location. In the public finance literature, for example, local governments take into account the behavior of neighboring governments in setting their tax rates (cf. Wilson, 1999; and Brueckner, 2003) and deciding on the provision of public goods (cf. Case, Rosen, and Hines, 1993). In the trade literature, for example, foreign direct investment (FDI) inflows into the host country depend on FDI inflows into proximate

[1]See Arellano (2003) and Baltagi (2008, Chapter 8) for an analysis of dynamic panel data models and Anselin (1988, 2006) for a treatment of spatial econometrics.

[2]Badinger, Mueller, and Tondl (2004) and Jacobs, Ligthart, and Vrijburg (2009) provide empirical applications of spatial dynamic panel data models.

[3]See Cameron and Trivedi (2005, pp. 744–745) for a general discussion on GMM estimators.

host countries (cf. Blonigen et al., 2007). Spatial error dependence is an alternative way of capturing spatial aspects. It may arise, for example, in house price models in which air quality affects house prices but is not included as an explanatory variable (cf. Kim, Phipps, and Anselin, 2003).[4] Spatially correlated errors can be thought of as analogous to the well-known practice of clustering error terms by groups, which are defined based on some direct observable characteristic of the group. In spatial econometrics, the groups are based on spatial "similarity," which is typically captured by some geographic characteristic (e.g., proximity).

Recently, Kapoor, Kelejian, and Prucha (2007) designed a GMM procedure to deal with spatial error correlation. We extend their three-step spatial procedure to panels with a spatially lagged dependent variable *and* a one-period time lag of the dependent variable. The performance of the spatial GMM estimators—which is measured in terms of bias and root mean squared error (RMSE)—is investigated by means of Monte Carlo simulations. We consider panels with a small number of time periods relative to the number of units. Rather than modeling either a spatial error or a spatial lag model, we allow both processes to be present simultaneously. In economic interaction models, spatial error dependence may exist above and beyond the theoretically motivated spatial lag structure,[5] reflecting the potential presence of omitted spatial variables. Ignoring spatial error correlation in static panels may give rise to a loss in efficiency of the estimates and may thus erroneously suggest that strategic interaction is absent. In contrast, disregarding spatial dependency in the dependent variable comes at a relatively high cost because it gives rise to biased estimates (cf. LeSage and Pace, 2009, p. 158).

The time lag of the (endogenous) dependent variable is correlated with the unit-specific

[4]Spatial error correlation may also result from measurement error in variables or a misspecified functional form of the regression equation.

[5]In their study on commodity tax competition, Egger, Pfaffermayr, and Winner (2005) find a significantly positive coefficient of the spatial lag while controlling for spatially correlated errors; the spatially autoregressive coefficient is shown to be significantly negative and non-negligible in size.

effect. Consequently, the standard fixed effects estimator for (non-spatial) panels with a fixed time span and a large number of units is biased and inconsistent. Dynamic panel data models are usually estimated using the GMM estimator of Arellano and Bond (1991), which differs from static panel GMM estimators in the set of moment conditions and the matrix of instruments. Alternatively, authors have used Blundell and Bond's system (1998) approach, which extends Arellano and Bond's instrument set by allowing differences of variables to be included. In the following, we contribute to the literature by developing spatial and spatial correlation corrected variants of the Arellano-Bond and Blundell-Bond estimators. This involves defining appropriate instruments to control for the endogeneity of the spatial lag and time lag of the dependent variable while controlling for spatial error correlation. For this we use an instrumental variables approach based on the set of instruments suggested by Kelejian and Robinson (1993).

The Monte Carlo evidence indicates the following. Differences in bias as well as RMSE between spatial GMM estimates and corresponding GMM estimates in which spatial error correlation is ignored are small. Choosing appropriate instruments for both the time and spatial lag is sufficient to model the spatial characteristics in our setting. However, it remains to be seen whether our estimators also perform well when exogenous variables are allowed to have spatial attributes.

The paper is organized as follows. Section 2 sets out the dynamic spatial panel data model. Section 3 develops the two estimators for dynamic spatial panel data models, that is, the spatial Arellano-Bond and spatial Blundell-Bond estimators. Section 4 presents Monte Carlo simulation outcomes. Finally, Section 5 concludes.

2 The Dynamic Spatial Panel Data Model

Consider $i = 1, ..., N$ spatial units and $t = 1, ..., T$ time periods. The focus is on panels with a small number of time periods relative to the number of units. Assume that the data at time t are generated according to the following model:

$$\mathbf{y}(t) = \lambda \mathbf{y}(t-1) + \delta \mathbf{W}_N \mathbf{y}(t) + \mathbf{x}(t)\boldsymbol{\beta} + \mathbf{u}(t), \tag{1}$$

where $\mathbf{y}(t)$ is an $N \times 1$ vector of observations on the dependent variable, $\mathbf{y}(t-1)$ is a one-period time lag of the dependent variable, $\mathbf{W}_N$ is an $N \times N$ matrix of spatial weights, $\mathbf{x}(t)$ is an $N \times K$ matrix of observations on the (exogenous) explanatory variables (where K denotes the number of covariates), and $\mathbf{u}(t)$ is an $N \times 1$ vector of error terms (see below).[6] The scalar parameter λ is the autocorrelation coefficient, δ is the spatial autoregressive coefficient (which measures the endogenous interaction effect among units), and $\boldsymbol{\beta}$ is a $K \times 1$ vector of slope coefficients.

The *spatial lag* is denoted by $\mathbf{W}_N \mathbf{y}(t)$, which captures the contemporaneous correlation between unit i's behavior and a weighted sum of the behavior of units $j \neq i$. The elements of $\mathbf{W}_N$ (denoted by ω_{ij}) are exogenously given, non-negative, and zero on the diagonal of the matrix. In addition, the elements are row normalized so that each row sums to one. Note that there is little formal guidance on choosing the "correct" spatial weights because many definitions of neighbors are possible. The literature usually employs contiguity (having common borders) or physical distance as weighting factors.

In the absence of a time lag of the dependent variable ($\lambda = 0$), the reduced form of Equation (1) amounts to: $\mathbf{y}(t) = (\mathbf{I}_N - \delta \mathbf{W}_N)^{-1}\mathbf{x}(t)\boldsymbol{\beta} + (\mathbf{I}_N - \delta \mathbf{W}_N)^{-1}\mathbf{u}(t)$, where

$$(\mathbf{I}_N - \delta \mathbf{W}_N)^{-1} = \sum_{i=0}^{\infty} (\delta \mathbf{W}_N)^i = \mathbf{I}_N + \delta \mathbf{W}_N + \delta^2 \mathbf{W}_N^2 + \delta^3 \mathbf{W}_N^3 + ..., \tag{2}$$

[6]Note that our exogenous explanatory variables are not spatially weighted. We leave this for further research.

where $\mathbf{I}_N$ is an identity matrix of dimension $N \times N$ and $\mathbf{W}_N^0 = \mathbf{I}_N$. Hence, the dependent variable is affected not only by the characteristics of the own jurisdiction but also those of direct neighbors and of "neighbors of neighbors." Stationarity of the model requires that $1/\zeta_L < \delta < 1/\zeta_U$, where ζ_L and ζ_U denote the smallest and largest characteristic root of $\mathbf{W}_N$, respectively. If $\mathbf{W}_N$ is (row) normalized then $\zeta_U = 1$ and thus $1/\zeta_L < \delta < 1$. If a time lag of the dependent variable is present ($\lambda > 0$), the stationarity condition $|\lambda| < 1$ needs to be imposed (cf. Elhorst, 2008b).

Spatial error correlation may arise when omitted variables follow a spatial pattern, yielding a non-diagonal variance-covariance matrix of the error term $\mathbf{u}(t)$. In the case of spatial correlation, the error structure in Equation (1) is a spatially weighted average of the error components of neighbors, where $\mathbf{M}_N$ is an $N \times N$ matrix of spatial weights (with typical element m_{ij}). More formally, the spatially autoregressive process is given by:

$$\mathbf{u}(t) = \rho \mathbf{M}_N \mathbf{u}(t) + \boldsymbol{\varepsilon}(t), \tag{3}$$

where $\mathbf{M}_N \mathbf{u}(t)$ is the *spatial error* term, ρ is a (second) spatially autoregressive coefficient, and $\boldsymbol{\varepsilon}(t)$ denotes a vector of innovations. The interpretation of the "nuisance" parameter ρ is very different from δ in the spatial lag model, in that there is no particular relation to a substantive theoretical underpinning of the spatial interaction. We follow the common practice in the literature by assuming that $\mathbf{W}_N \neq \mathbf{M}_N$.[7] The spatial error process in reduced form is: $\mathbf{u}(t) = (\mathbf{I}_N - \rho \mathbf{M}_N)^{-1} \boldsymbol{\varepsilon}(t) = \boldsymbol{\varepsilon}(t) + \rho \mathbf{M}_N \boldsymbol{\varepsilon}(t) + \rho^2 \mathbf{M}_N^2 \boldsymbol{\varepsilon}(t) + \rho^3 \mathbf{M}_N^3 \boldsymbol{\varepsilon}(t) + \ldots$, where $\mathbf{I}_N$ is an identity matrix of dimension N. Shocks in the spatial error representation have a global effect. Intuitively, a shock in location k directly affects the error term of location k but also indirectly transmits to other locations (with a non-zero m_{ij}) and eventually works its way back to k. If $|\rho| < 1$, the spatial error process is stable thus yielding feedback effects that are bounded.

[7] If $\mathbf{W}_N = \mathbf{M}_N$, then parameters δ and ρ cannot be disentangled when estimated by ML (cf. Anselin, 2006). When using GMM estimation, however, δ and ρ can be separately identified.

The vector of innovations is defined as:

$$\boldsymbol{\varepsilon}(t) = \mathbf{I}_N\boldsymbol{\eta}(t) + \mathbf{v}(t), \qquad \mathbf{v} \sim \text{iid}(0, \sigma_v^2\mathbf{I}_N), \tag{4}$$

where $\boldsymbol{\eta}$ is an $N \times 1$ vector representing (unobservable) unit-specific effects, and $\mathbf{v}(t)$ is an $N \times 1$ vector of independently and identically distributed (iid) error terms with variance σ_v^2 (which is assumed to be constant across units and time periods). In the following, we focus on a specification in which $\boldsymbol{\eta}$ is correlated with the regressors.

Using conventional notation, Equations (1), (3), and (4) can be written concisely as:

$$\mathbf{y}(t) = \mathbf{z}(t)\boldsymbol{\theta} + \mathbf{u}(t), \tag{5}$$

$$\mathbf{u}(t) = (\mathbf{I}_N - \rho\mathbf{M}_N)^{-1}[\mathbf{I}_N\boldsymbol{\eta}(t) + \mathbf{v}(t)], \tag{6}$$

where $\mathbf{z}(t) \equiv [\mathbf{y}(t-1), \mathbf{W}_N\mathbf{y}(t), \mathbf{x}(t)]$ denotes the matrix of regressors, $\boldsymbol{\theta} \equiv [\lambda, \delta, \boldsymbol{\beta}']'$ is a vector of $K+2$ parameters, and a prime denotes a transpose. Our general dynamic spatial panel data model embeds various special cases discussed in the literature. If $\lambda = \rho = 0$ and $\delta > 0$, our model reduces to the familiar spatial lag model (also known as the mixed regressive spatial autoregressive model; see Anselin, 1988), whereas for $\lambda = \rho = \beta = 0$ we get a pure spatial autoregressive model. If $\lambda = \delta = 0$ and $\rho > 0$, we obtain the spatial error model. If $\lambda > 0$ and $\delta = \rho = 0$, we arrive at Arellano and Bond's dynamic panel data model. Finally, the dynamic general spatial panel data model boils down to a standard static panel data model if $\lambda = \delta = \rho = 0$.

3 Dynamic Spatial Panel Data Estimators

This section develops the dynamic spatial estimators to be used in the Monte Carlo simulations of Section 4. We extend Kapoor, Kelejian, and Prucha's (2007) approach—which explicitly corrects for spatial error correlation—to include a spatial lag and a time lag of the dependent variable. Because the latter two regressors are endogenous, we also propose sets of instruments. This procedure yields "spatial" Arellano-Bond and Blundell-Bond estimators in Sections 3.1 and 3.2, respectively.

3.1 Spatial Arellano-Bond Estimator

This section extends Arellano and Bond's (1991) dynamic panel data approach to a spatial setting in which both a spatial lag and a spatial error term are present. The Spatial Arellano-Bond estimator (denoted by SAB) will be derived in three stages.

3.1.1 The First Stage

In the first stage, we derive a consistent estimate of $\boldsymbol{\theta}$, which we use to calculate $\hat{\mathbf{u}}(t) = \mathbf{y}(t) - \mathbf{z}(t)\hat{\boldsymbol{\theta}}$.[8] To eliminate $\boldsymbol{\eta}$ from $\boldsymbol{\varepsilon}(t)$, we take first differences of (5) and (6):

$$\begin{aligned} \Delta\mathbf{y}(t) &= \Delta\mathbf{z}(t)\boldsymbol{\theta} + \Delta\mathbf{u}(t), & (7) \\ \Delta\mathbf{u}(t) &= (\mathbf{I}_N - \rho\mathbf{M}_N)^{-1}\Delta\mathbf{v}(t), \qquad \text{for} \quad t = 3, ..., T, & (8) \end{aligned}$$

where $\Delta\mathbf{q}(t) \equiv \mathbf{q}(t) - \mathbf{q}(t-1)$ for $\mathbf{q}(t) = \{\mathbf{y}(t), \mathbf{z}(t), \mathbf{u}(t), \mathbf{v}(t)\}$. Because of the endogeneity of the spatial lag and the time lag of the dependent variable, we apply a panel GMM procedure. Both endogenous regressors are correlated. Therefore, the challenge is to find instruments that are more strongly correlated with the spatial lag than with the time lag of the dependent variable. The panel GMM procedure relies on the existence of an $N(T-2) \times F$ instrument matrix (i.e., $\mathbf{H}_{SAB}$, see below) that satisfies the following F moment conditions: $\mathrm{E}[\mathbf{H}'_{SAB}\Delta\mathbf{u}] = 0$, where we use stacked notation (sorting the data first by time and then by unit) and E denotes the expectations operator. If the model is

[8] Strictly speaking, consistency still has to be proved. Both the Arellano-Bond estimator and the spatial GMM estimator of Kelejian and Robinson (1993) are consistent. What remains to be proved is that the combination of the two is also consistent. The asymptotics of our spatial GMM estimators are postponed to a next version of this paper.

just identified, the panel GMM estimator simplifies to an instrumental variables estimator. The resulting first-stage SAB estimator becomes:

$$\hat{\boldsymbol{\theta}}_{SAB} = [\Delta\mathbf{z}'\mathbf{H}_{SAB}\mathbf{A}_{SAB}\mathbf{H}'_{SAB}\Delta\mathbf{z}]^{-1}\,\Delta\mathbf{z}'\mathbf{H}_{SAB}\mathbf{A}_{SAB}\mathbf{H}'_{SAB}\Delta\mathbf{y}, \tag{9}$$

where $\mathbf{A}_{SAB} = [\mathbf{H}'_{SAB}\mathbf{G}\mathbf{H}_{SAB}]^{-1}$ is an $F \times F$ matrix of instruments and $\mathbf{G} = \mathbf{I}_N \otimes G_{ij}$ is an $N(T-2) \times N(T-2)$ weighting matrix with elements:[9]

$$G_{ij} \equiv \begin{cases} 2 & \text{if } i = j \\ -1 & \text{if } i = j+1 \\ -1 & \text{if } j = i+1 \\ 0 & \text{otherwise} \end{cases}, \tag{10}$$

where $\otimes$ denotes the Kronecker product. The instrument matrix $\mathbf{H}_{SAB}(t) = [\mathbf{y}(t-2), ..., \mathbf{y}(1), \mathbf{H}(t)]$ consists of two sets: (i) the dynamic instruments suggested by Arellano and Bond (which we call the AB component); and (ii) the instruments for the spatial lag (denoted by $\mathbf{H}(t)$).

Arellano and Bond (1991) propose to use the *levels* of the dependent variable (i.e., $\mathbf{y}(t-2), ..., \mathbf{y}(1)$) as instruments for the time lag of the dependent variable in *first differences* (i.e., $\Delta\mathbf{y}(t-1)$). Because of time dependency in the model, the instruments are correlated with the time lag of the dependent variable in first differences $\Delta\mathbf{y}(t-1)$, but uncorrelated with the error term in first differences (i.e., $\Delta\mathbf{v}(t)$) as the unit-specific effect is eliminated from the first differenced variable. Recall that $\mathbf{y}(t-2)$ is correlated with $\mathbf{v}(t-2), ..., \mathbf{v}(1)$, but not with $\mathbf{v}(t)$ and $\mathbf{v}(t-1)$. The Arellano-Bond procedure implies the moment conditions

$$\mathrm{E}[\mathbf{y}(t-s)'\Delta\mathbf{v}(t)] = 0, \tag{11}$$

for $t = 3, ..., T$ and $s = 2, ..., T-1$. Equation (11) yields $(T-2)(T-1)/2$ potential instruments.

[9] Arellano and Bond (1991) use (10) to yield a one-step estimator that is asymptotically equivalent to the two-step estimator (if the v_{ij}'s are independent and homoskedastic both across units and over time).

The set of instruments for the spatial lag, that is, $\mathbf{H}(t)$, is based on a modified version of the approach of Kelejian and Robinson (1993). We expand the expected value of the spatial lag to arrive at $\mathbf{H}(t) = [\mathbf{W}_N\mathbf{x}(t), \mathbf{x}(t)]$. This yields $2K$ additional instruments, which satisfy the following moment condition:

$$\mathrm{E}[\mathbf{H}(t)'\Delta\mathbf{v}(t)] = 0. \tag{12}$$

The total instrument matrix $\mathbf{H}_{SAB}$ consists of the two instrument sets discussed above. For example, for $T = 5$, we find the following:

$$\mathbf{H}_{SAB} = \begin{bmatrix} \mathbf{y}(1) & 0 & 0 & 0 & 0 & 0 & \mathbf{H}(1) \\ 0 & \mathbf{y}(1) & \mathbf{y}(2) & 0 & 0 & 0 & \mathbf{H}(2) \\ 0 & 0 & 0 & \mathbf{y}(1) & \mathbf{y}(2) & \mathbf{y}(3) & \mathbf{H}(3) \end{bmatrix}. \tag{13}$$

The AB component has $T - 2$ instrument blocks. Note that if T increases, the number of instruments rises exponentially, potentially giving rise to multicollinearity. The matrix contains $F = (T-2)(T-1)/2 + 2K$ instruments in total.

3.1.2 The Second Stage

In the second step, consistent GMM estimates of ρ and σ_v^2 are obtained using $\hat{\mathbf{u}}(t)$ and the three moment conditions of Kapoor, Kelejian, and Prucha (2007), which we modify to correct for the time lag in the model. These moment conditions (in stacked format) are:

$$\begin{bmatrix} \frac{1}{N(T-2)}\boldsymbol{\varepsilon}'\mathbf{Q}_N\boldsymbol{\varepsilon} \\ \frac{1}{N(T-2)}\bar{\boldsymbol{\varepsilon}}'\mathbf{Q}_N\bar{\boldsymbol{\varepsilon}} \\ \frac{1}{N(T-2)}\bar{\boldsymbol{\varepsilon}}'\mathbf{Q}_N\boldsymbol{\varepsilon} \end{bmatrix} = \begin{bmatrix} \sigma_v^2 \\ \sigma_v^2\frac{1}{N}\mathrm{tr}(\mathbf{M}_N'\mathbf{M}_N) \\ 0 \end{bmatrix}, \tag{14}$$

where tr(.) is the trace of the matrix $\mathbf{M}_N'\mathbf{M}_N$, $\mathbf{Q}_N \equiv (\mathbf{I}_{T-1} - \mathbf{J}_{T-1}/(T-1)) \otimes \mathbf{I}_N$ is the "within" transformation matrix, $\mathbf{I}_{T-1}$ is an identity matrix of dimension $T-1$, and $\mathbf{J}_{T-1} \equiv$

$\mathbf{e}_{T-1}\mathbf{e}'_{T-1}$ is a $(T-1)\times(T-1)$ matrix of unit elements. The $N(T-1)\times 1$ vectors $\boldsymbol{\varepsilon}$ and $\bar{\boldsymbol{\varepsilon}}$ are:

$$\boldsymbol{\varepsilon} \equiv \mathbf{u} - \rho\bar{\mathbf{u}}, \qquad \bar{\boldsymbol{\varepsilon}} \equiv \bar{\mathbf{u}} - \rho\bar{\bar{\mathbf{u}}}, \tag{15}$$

where $\bar{\mathbf{u}} = (\mathbf{I}_{T-1}\otimes\mathbf{M}_N)\mathbf{u}$ and $\bar{\bar{\mathbf{u}}} = (\mathbf{I}_{T-1}\otimes\mathbf{M}_N)\bar{\mathbf{u}}$. The estimated values of $\mathbf{u}$ in Equation (15) are substituted into (14) to obtain three equations in two unknowns:[10]

$$\frac{1}{N}\begin{bmatrix} \frac{2}{T-2}\hat{\mathbf{u}}'\mathbf{Q}_N\hat{\bar{\mathbf{u}}} & -\frac{1}{T-2}\hat{\bar{\mathbf{u}}}'\mathbf{Q}_N\hat{\bar{\mathbf{u}}} & 1 \\ \frac{2}{T-2}\hat{\bar{\bar{\mathbf{u}}}}'\mathbf{Q}_N\hat{\bar{\mathbf{u}}} & -\frac{1}{T-2}\hat{\bar{\bar{\mathbf{u}}}}'\mathbf{Q}_N\hat{\bar{\bar{\mathbf{u}}}} & \mathrm{tr}(\mathbf{M}_N'\mathbf{M}_N) \\ \frac{2}{T-2}\left[\hat{\mathbf{u}}'\mathbf{Q}_N\hat{\bar{\bar{\mathbf{u}}}} + \hat{\bar{\mathbf{u}}}'\mathbf{Q}_N\hat{\bar{\mathbf{u}}}\right] & -\frac{1}{T-2}\hat{\bar{\mathbf{u}}}'\mathbf{Q}_N\hat{\bar{\bar{\mathbf{u}}}} & 0 \end{bmatrix}\begin{bmatrix} \rho \\ \rho^2 \\ \sigma_v^2 \end{bmatrix} = \frac{1}{N}\begin{bmatrix} \frac{1}{T-2}\hat{\mathbf{u}}'\mathbf{Q}_N\hat{\mathbf{u}} \\ \frac{1}{T-2}\hat{\bar{\mathbf{u}}}'\mathbf{Q}_N\hat{\bar{\mathbf{u}}} \\ \frac{1}{T-2}\mathbf{u}'\mathbf{Q}_N\hat{\bar{\mathbf{u}}} \end{bmatrix}. \tag{16}$$

This nonlinear system of equations can be solved to obtain estimates of ρ and σ_v^2.

3.1.3 The Third Stage

In the final stage, the estimate of ρ is used to spatially transform the variables in (5) to yield:

$$\Delta\tilde{\mathbf{y}} = \Delta\tilde{\mathbf{z}}\boldsymbol{\theta}_{SAB} + \Delta\boldsymbol{\varepsilon}, \tag{17}$$

where $\tilde{\mathbf{p}} = [\mathbf{I}_N - \hat{\rho}\mathbf{M}_N]\mathbf{p}$ for $\mathbf{p} = \{\mathbf{y}, \mathbf{z}\}$. The transformed model is used to derive the final-stage SAB estimator:

$$\hat{\boldsymbol{\theta}}_{SAB} = \left[\Delta\tilde{\mathbf{z}}'\tilde{\mathbf{H}}_{SAB}\tilde{\mathbf{A}}_{SAB}\tilde{\mathbf{H}}'_{SAB}\Delta\tilde{\mathbf{z}}\right]^{-1}\Delta\tilde{\mathbf{z}}'\tilde{\mathbf{H}}_{SAB}\tilde{\mathbf{A}}_{SAB}\tilde{\mathbf{H}}'_{SAB}\Delta\tilde{\mathbf{y}}, \tag{18}$$

where $\tilde{\mathbf{A}}_{SAB} \equiv \left[\tilde{\mathbf{H}}'_{SAB}\mathbf{G}\tilde{\mathbf{H}}_{SAB}\right]^{-1}$ and $\tilde{\mathbf{H}}_{SAB} \equiv [\mathbf{I}_N - \hat{\rho}\mathbf{M}_N]\mathbf{H}_{SAB}$.

3.2 Spatial Blundell-Bond Estimator

The standard Arellano-Bond estimator is known to be rather inefficient when instruments are weak because it makes use of information contained in first differences of variables only.

[10]Kapoor, Kelejian, and Prucha (2007) derive six moment conditions, but only the first three conditions are relevant for a model characterized by fixed effects.

To address this shortcoming, the GMM approach of Blundell and Bond (1998) uses both variables in levels and in first differences in one model. The spatial variant of Blundell and Bond's estimator (denoted by SBB) can be derived using the following model:

$$\begin{bmatrix} \mathbf{y}(t) \\ \Delta\mathbf{y}(t) \end{bmatrix} = \begin{bmatrix} \mathbf{z}(t) \\ \Delta\mathbf{z}(t) \end{bmatrix} \boldsymbol{\theta}_{SBB} + \begin{bmatrix} \mathbf{u}(t) \\ \Delta\mathbf{u}(t) \end{bmatrix}, \tag{19}$$

which can be expressed more compactly as:

$$\mathbf{y}_{SBB}(t) = \mathbf{z}_{SBB}(t)\boldsymbol{\theta}_{SBB} + \mathbf{u}_{SBB}(t), \tag{20}$$

where $\mathbf{y}_{SBB}(t)$ is a $2N \times 1$ vector. The Blundell-Bond structure doubles the number of observations (i.e., $2N(T-2)$), which increases estimation efficiency. Again, we can use the three-step spatial estimation procedure as set out in Section 3.1. In the first step, the panel GMM estimator is derived:

$$\hat{\boldsymbol{\theta}}_{SBB} = \left[\mathbf{z}'_{SBB}\mathbf{H}_{SBB}\mathbf{A}_{SBB}\mathbf{H}'_{SBB}\mathbf{z}_{SBB}\right]^{-1} \mathbf{z}'_{SBB}\mathbf{H}_{SBB}\mathbf{A}_{SBB}\mathbf{H}'_{SBB}\mathbf{y}_{SBB}, \tag{21}$$

where $\mathbf{A}_{SBB} = \left[\mathbf{H}'_{SBB}\mathbf{H}_{SBB}\right]^{-1}$ and $\mathbf{H}_{SBB}$ is defined as:

$$\mathbf{H}_{SBB} = \begin{bmatrix} \mathbf{H}_D & 0 \\ 0 & \mathbf{H}_L \end{bmatrix}. \tag{22}$$

The upper left block of $\mathbf{H}_{SBB}$, denoted by $\mathbf{H}_{\mathrm{D}}$, contains the instruments for the model in first differences and the lower right block, $\mathbf{H}_{\mathrm{L}}$, includes the instruments for the model in levels. Note that diagonal structure of $\mathbf{H}_{SBB}$ ensures that the instruments intended for the variables defined in levels do not interact with the variables in first differences and vice versa. The instrument matrix $\mathbf{H}_D$ is based on the following moment conditions:

$$\mathrm{E}[\mathbf{y}(t-s)'\Delta\mathbf{v}(t)] = 0, \quad \mathrm{E}[(\mathbf{W}_N\mathbf{y}(t-s))'\Delta\mathbf{v}(t)] = 0, \quad \mathrm{E}[\mathbf{H}(t)'\Delta\mathbf{v}(t)] = 0, \tag{23}$$

while $\mathbf{H}_L$ is based on:

$$\mathrm{E}[\mathbf{v}(t-s)'\Delta\mathbf{y}(t)] = 0, \quad \mathrm{E}[\mathbf{v}(t-s)'\mathbf{W}_N\Delta\mathbf{y}(t)] = 0, \quad \mathrm{E}[\mathbf{v}(t-s)'\Delta\mathbf{H}(t)] = 0, \tag{24}$$

for $t = 3, ..., T$ and $s = 2, ..., T-1$. Compared to the SAB estimator, we consider an additional set of dynamic instruments for the spatial lag, which are based on various time lags of the spatially lagged dependent variable. Note that $\mathbf{H}(t)$ continues to include the "static" instruments for the spatial lag, as defined above.

Taking $T = 5$, for example, gives the following upper left block of $\mathbf{H}_{SBB}$,

$$\mathbf{H}_D = \begin{bmatrix} \mathbf{y}(1) & 0 & 0 & \mathbf{W}_N\mathbf{y}(1) & 0 & 0 & \mathbf{H}(1) \\ \mathbf{y}(2) & \mathbf{y}(1) & 0 & \mathbf{W}_N\mathbf{y}(2) & \mathbf{W}_N\mathbf{y}(1) & 0 & \mathbf{H}(2) \\ \mathbf{y}(3) & \mathbf{y}(2) & \mathbf{y}(1) & \mathbf{W}_N\mathbf{y}(3) & \mathbf{W}_N\mathbf{y}(2) & \mathbf{W}_N\mathbf{y}(1) & \mathbf{H}(3) \end{bmatrix}, \tag{25}$$

and the lower right block

$$\mathbf{H}_L = \begin{bmatrix} \Delta\mathbf{y}(1) & 0 & 0 & \mathbf{W}_N\Delta\mathbf{y}(1) & 0 & 0 & \Delta\mathbf{H}(1) \\ \Delta\mathbf{y}(2) & \Delta\mathbf{y}(1) & 0 & \mathbf{W}_N\Delta\mathbf{y}(2) & \mathbf{W}_N\Delta\mathbf{y}(1) & 0 & \Delta\mathbf{H}(2) \\ \Delta\mathbf{y}(3) & \Delta\mathbf{y}(2) & \Delta\mathbf{y}(1) & \mathbf{W}_N\Delta\mathbf{y}(3) & \mathbf{W}_N\Delta\mathbf{y}(2) & \mathbf{W}_N\Delta\mathbf{y}(1) & \Delta\mathbf{H}(3) \end{bmatrix}, \tag{26}$$

The two instrument blocks of $\mathbf{H}_{SBB}$ differ from the instruments in $\mathbf{H}_{SAB}$ as follows. First of all, the blocks of $\mathbf{H}_{SBB}$ include dynamic instruments for the spatial lag [columns 4–6 of (25) and (26)] in addition to the standard dynamic instruments (columns 1–3) and the static instruments for the spatial lag (column 7). Second, we operationalize the moment conditions for both the time lag and the spatial lag of the endogenous variable differently,[11] that is, the moment conditions are instrumented by three instead of six dynamic instruments each of which is more strongly correlated with the variable it intends to instrument. Relative

[11] Recall from Section 3.1 that the number of instruments suggested by Arellano and Bond (1991) tends to explode for larger values of T. When used in short panels (i.e., $T < 10$) this need not be a problem, but raises multicollinearity problems for larger time periods. One solution is to cut down on the number of time lags used to instrument the time lag of the dependent variable. Alternatively, as pursued here, one can opt to operationalize the moment conditions differently.

to the operationalization of moment conditions used for $\mathbf{H}_{SAB}$, the new operationalization reduces the total number of instruments (including the dynamic instruments for the spatial lag) from $2\left[(T-2)(T-1)/2+(T-2)(T-1)/2+2K\right]$ to only $2(2T-4+2K)$.[12]

Analogously to Section 3.1, the second stage of the estimation procedure uses $\hat{\mathbf{u}}_{SBB} = \mathbf{y}_{SBB} - \mathbf{z}_{SBB}\hat{\boldsymbol{\theta}}_{SBB}$ from the GMM panel procedure in the three moment conditions (14) to arrive at $\hat{\rho}$ and $\hat{\sigma}_v^2$. The value of $\hat{\rho}$ is employed to transform the variables in the third step.

The SBB estimator of $\boldsymbol{\theta}$ in the third stage is:

$$\boldsymbol{\theta}_{SBB} = \left[\tilde{\mathbf{z}}'_{SBB}\tilde{\mathbf{H}}_{SBB}\tilde{\mathbf{A}}_{SBB}\tilde{\mathbf{H}}'_{SBB}\tilde{\mathbf{z}}_{SBB}\right]^{-1}\tilde{\mathbf{z}}'_{SBB}\tilde{\mathbf{H}}_{SBB}\tilde{\mathbf{A}}_{SBB}\tilde{\mathbf{H}}'_{SBB}\tilde{\mathbf{y}}_{SBB}, \tag{27}$$

where $\tilde{\mathbf{p}} = [\mathbf{I}_N - \hat{\rho}\mathbf{M}_N]\mathbf{p}$ for $\mathbf{p} = \{\mathbf{y}_{SBB}, \mathbf{z}_{SBB}\}$. The instruments of the matrix $\tilde{\mathbf{H}}_{SBB}$ are defined as: $\tilde{\mathbf{H}}_{SBB} = [\mathbf{I}_N - \hat{\rho}\mathbf{M}_N]\mathbf{H}_{BB}$ and $\tilde{\mathbf{A}}_{SBB} = \left[\tilde{\mathbf{H}}'_{SBB}\tilde{\mathbf{H}}_{SBB}\right]^{-1}$ for $t = 1, ..., T-1$.

4 Monte Carlo Simulations

To compare the performance of the estimators presented in Section 3, this section reports a Monte Carlo experiment. The design of the Monte Carlo experiment is discussed first before turning to the results.

4.1 Simulation Design

We report the small sample properties of the estimators using data sets generated based on the dynamic model of Section 2. To this end, we set $T = 5$ and $N = 60$ in the benchmark design. In generating the data, we follow a three-step procedure. First, we generate the matrix of covariates, which includes only one exogenous variable, that is, $\mathbf{x}(t)$, which is an $N \times 1$ vector. Following Baltagi et al. (2007), the exogenous variable is defined as:

$$\mathbf{x}(t) = \boldsymbol{\varsigma} + \boldsymbol{\chi}(t), \qquad \boldsymbol{\varsigma} \sim \text{iid } U[-7.5, 7.5], \qquad \boldsymbol{\chi} \sim \text{iid } U[-5, 5], \tag{28}$$

[12] For $T = 5$ and $K = 1$, the number of instruments for both $\mathbf{H}_D$ and $\mathbf{H}_L$ are cut down by six compared with the SAB operationalization.

where $\boldsymbol{\varsigma}$ represents the unit-specific component and $\boldsymbol{\chi}(t)$ denotes a random component drawn from a uniform distribution, U, defined on a pre-specified interval.

The second step generates the error component $\mathbf{u}(t)$ using:

$$\boldsymbol{\eta} \sim \text{iid } U[-1,1], \qquad \mathbf{v} \sim \text{iid } N(0, \mathbf{I}_N), \tag{29}$$

where N denotes the normal distribution. The third step generates data on the dependent variable $\mathbf{y}(t)$ and the spatial lag, $\mathbf{W}_N\mathbf{y}(t)$. The data generation process is given by:

$$\mathbf{y}(t) = (\mathbf{I}_N - \delta\mathbf{W}_N)^{-1}\left[\lambda\mathbf{y}(t-1) + \beta x(t) + (\mathbf{I}_N - \rho\mathbf{M}_N)^{-1}(\boldsymbol{\eta} + \mathbf{v}(t))\right], \tag{30}$$

for $t = 2, ..., T$ and $\mathbf{y}(1) = \boldsymbol{\eta}$. The first $100 - T$ observations of the Monte Carlo runs are discarded to ensure that the results are not unduly affected by the initial values (cf. Hsiao, Pesaran, and Tahmiscioglu (2002).[13] Following standard practice in the literature, we use different weight matrices for the spatial lag and spatial error component, that is, $\mathbf{W}_N \neq \mathbf{M}_N$. We use two randomly generated spatial weight matrices, which meet the criteria set out in Section 2.

The parameters in (30) take on the following values in the data generation process. The coefficient of the exogenous explanatory variable (β) is set to unity.[14] The spatial autocorrelation coefficient (ρ) is set to -0.3 in the simulations. A negative value of ρ implies that an unobserved positive shock in the equation for spatial unit i reduces the dependent variable in other spatial units $i \neq j$. We also consider alternative values of ρ, that is, -0.4, -0.2, 0, 0.2, and 0.4; the zero value covers a data generation process based on a pure spatial lag model. We set $\lambda = 0.3$ and $\delta = 0.5$, so that the stationarity conditions of Section 2 are satisfied. Finally, we perform some robustness tests with various values of T ranging from 10 to 50.

[13] We have checked the robustness of the results with respect to changes in the initial values.

[14] Setting $\beta = 1$ is standard practice in the literature. Note that the model does not feature a common intercept across all cross-sectional units.

For each experiment, the performance of the estimators is computed based on $1,000$ replications. Following Kapoor, Kelejian, and Prucha (2007) and others, we measure performance by the RMSE $= \sqrt{bias^2 + \left(\frac{q_1 - q_2}{1.35}\right)^2}$, where *bias* denotes the difference between the median and the "true" value of the parameter of interest (i.e., the value imposed in the data-generating process) and $q_1 - q_2$ is the interquartile range (where q_1 is the 0.75 quantile and q_2 is the 0.25 quantile). If the distribution is normal, $(q_1 - q_2)/1.35$ comes close (aside from a rounding error) to the standard deviation of the estimate.[15] The RMSE thus consists of the bias of the estimator (the first term) and a measure of the distribution of the estimate (the second term).

In the simulations, we compare various estimators with each other. We use four different types of GMM estimators all of which instrument the time lag of the dependent variable. We use the Spatial Arellano-Bond estimator (labeled SAB) and Spatial Blundell-Bond estimator (denoted SBB), which correct for spatial error correlation and apply appropriate instruments for the spatial lag of the dependent variables (as discussed in Sections 3.1 and 3.2). In addition, we consider a Modified Arellano-Bond estimator (labeled MAB) and a Modified Blundell-Bond estimator (labeled MBB), which instrument the spatial lag of the dependent variable—using the standard static instruments suggested by Kelejian and Robinson (1993)—but do not correct for spatial error correlation. We compare the GMM estimators with the Modified Least Squares Dummy Variables (MLSDV) estimator, which applies fixed effects and instruments the spatial lag with the static spatial instruments (as defined in $\mathbf{H}(t)$). Nickell (1981) shows that the regular LSDV estimator yields biased estimates in the case of dynamic panels because it does not instrument the time lag of the dependent variable. Although this bias approaches zero as the number of time periods tends to infinity, it cannot be ignored in small samples.

[15] We have used the median instead of the mean in summarizing the distribution because the former is less sensitive to outliers.

4.2 Results

Table 1 reports the bias and RMSE in the parameters λ, δ, β, and ρ based on 1,000 replications for five estimators, that is, MLSDV, MAB, SAB, MBB, and SBB. The parameter ρ is only estimated in case of the spatial estimators SAB and SBB, but features in the data generation process in all cases. The left panel of the table shows the performance of the estimators for various values of T, whereas the right panel considers performance for various values of the spatially autoregressive coefficient in the data generation process. The middle column of the right panel presents the case of zero spatial error correlation. The bias in the coefficient of the lagged dependent variable (λ) is large and negative when using the MLSDV estimator. For larger values of T the bias gets smaller rapidly. The bias in λ is reduced if we use the MAB estimator or its full spatial variant, but it is still quite large. The MBB/SBB estimators yield a substantial reduction in the bias. Surprisingly, the MBB estimator yields a smaller bias than SBB. If T takes on larger values, this pattern is reversed. In the benchmark simulation, the MAB estimator shows a larger bias than the SAB estimator, but for larger T the SAB estimator performs better. The SBB estimator performs consistently better than the MLSDV estimator for all considered values of T. In estimating λ, the MBB/SBB estimators give rise to a smaller bias than the MAB/SAB and MLSDV estimators across all values of ρ.

The bias in the spatial interaction coefficient (δ) when using the MLSDV estimator is negative and in absolute terms much smaller than that for λ. Elhorst (2008b) finds a similar result in his Monte Carlo simulations when using the CLSDV estimator. In contrast to the negative bias obtained in δ when using the MLSDV estimator, the bias when using the MAB/MBB and SAB/SBB estimators is positive in all cases. The bias in the latter four estimators is smaller than that the absolute value of the bias when using the MLSDV estimator. The bias in δ when using MAB, SAB, MBB, and SBB in the benchmark scenario is on average 0.56 percent of the true δ value. This bias is much smaller than that found by

Elhorst (2008b), who also uses a spatial GMM approach, but does not allow spatial error correlation in the data generation process. Surprisingly, the MAB/MBB estimators yield a smaller bias than their spatial variants. This pattern does not uphold, however, for larger values of T. Again, the estimators based on Blundell and Bond's approach (MBB/SBB) perform better. Note that the MAB/MBB estimators yield a small bias for negative values of ρ, which increases for positive values of ρ.

The MLSDV estimator again gives rise to the largest absolute bias in estimating β, but it quickly outperforms the MAB/SAB estimators for larger values of T. The MAB/SAB estimators yield a negative bias, whereas the bias of the MBB/SBB estimators is positive in the benchmark specification. Just like for ρ, the bias in absolute terms of the MAB/MBB estimators is smaller than that of their spatial counterparts. The differences in bias are small, however.

In the benchmark scenario, the SBB estimator produces a smaller bias in ρ than the SAB estimator. When using the SBB estimator for $T = 5$, the bias in ρ amounts to 5.33 percent of its true value.

Table 1 also reports the RMSE of the parameters λ, δ, β, and ρ. As expected, the RMSE of each parameter decreases for larger values of T. Extending the time period from 5 to 60 reduces the RMSE by 76 percent on average. Not surprisingly, the fall in RMSE is large when time periods are added starting from a small value of T; a fall of 62 percent on average materializes if $T = 5$ is extended to $T = 20$.

5 Conclusion

This paper has dealt with Generalized Methods of Moments (GMM) estimation of spatial dynamic panel data models with spatially correlated errors. We extended the three-step GMM approach of Kapoor, Kelejian, and Prucha (2007), which corrects for spatially correlated errors in static panel data models, by introducing a spatial lag and a one-period lag of the endogenous variable as additional explanatory variables. Combining the extended Kapoor, Kelejian, and Prucha (2007) approach with the dynamic panel data model GMM estimators of Arellano and Bond (1991) and Blundell and Bond (1998) and supplementing the dynamic instruments by lagged and weighted exogenous variables as suggested by Kelejian and Robinson (1993) yielded new spatial dynamic panel data estimators.

Monte Carlo simulations indicated that differences in bias as well as root mean squared error between spatial GMM estimates and corresponding GMM estimates in which spatial error correlation is ignored are small. Choosing appropriate instruments for both the time and spatial lag is sufficient to model the spatial characteristics in our setting.

In future research, we will add spatially weighted exogenous variables to the model. Some preliminary analyses with the consumption tax data of Jacobs, Ligthart, and Vrijburg (2009) suggest that including spatially weighted exogenous variables does indeed have an impact.

Table 1: Bias and RMSE of the MLSDV, MAB, SAB, MBB, and SBB estimators for different values of T and ρ

Estimator	Parameter	bias/RMSE	T					ρ				
			5	10	15	20	50	-0.4	-0.2	0	0.2	0.4
MLSDV	λ	bias	-0.068	-0.026	-0.017	-0.013	-0.005	-0.069	-0.067	-0.067	-0.066	-0.074
		RMSE	0.076	0.032	0.023	0.018	0.009	0.077	0.075	0.075	0.075	0.081
	δ	bias	-0.017	-0.002	-0.002	-0.001	-0.001	-0.017	-0.015	-0.015	-0.013	-0.015
		RMSE	0.071	0.037	0.029	0.024	0.016	0.071	0.067	0.067	0.067	0.069
	β	bias	-0.020	-0.005	-0.002	-0.001	0.000	-0.021	-0.021	-0.020	-0.018	-0.024
		RMSE	0.043	0.022	0.017	0.015	0.009	0.045	0.044	0.044	0.044	0.047
MAB	λ	bias	-0.017	-0.009	-0.009	-0.007	-0.008	-0.018	-0.011	-0.009	-0.016	-0.018
		RMSE	0.079	0.031	0.024	0.017	0.011	0.080	0.082	0.080	0.079	0.079
	δ	bias	0.002	0.017	0.021	0.023	0.029	0.003	0.002	0.006	0.012	0.016
		RMSE	0.079	0.050	0.040	0.037	0.033	0.081	0.084	0.074	0.081	0.085
	β	bias	-0.006	-0.005	-0.002	-0.002	-0.002	-0.006	-0.009	-0.001	-0.004	-0.008
		RMSE	0.050	0.028	0.022	0.019	0.011	0.051	0.053	0.056	0.052	0.053
SAB	λ	bias	-0.015	-0.010	-0.010	-0.008	-0.008	-0.014	-0.012	-0.008	-0.011	-0.016
		RMSE	0.075	0.032	0.023	0.017	0.011	0.074	0.081	0.080	0.076	0.075
	δ	bias	0.003	0.017	0.021	0.024	0.028	0.003	0.005	0.007	0.008	0.004
		RMSE	0.073	0.045	0.038	0.036	0.032	0.071	0.081	0.076	0.082	0.081
	β	bias	-0.007	-0.005	-0.004	-0.003	-0.003	-0.007	-0.009	-0.001	-0.003	-0.006
		RMSE	0.049	0.026	0.022	0.017	0.010	0.049	0.052	0.055	0.051	0.050
	ρ	bias	0.025	0.004	0.000	-0.002	-0.006	0.033	0.013	0.012	-0.010	-0.015
		RMSE	0.128	0.072	0.059	0.048	0.031	0.132	0.123	0.116	0.120	0.115
MBB	λ	bias	-0.004	-0.002	-0.003	-0.002	-0.002	-0.005	-0.001	-0.004	-0.004	-0.005
		RMSE	0.025	0.016	0.012	0.011	0.006	0.026	0.024	0.025	0.025	0.026
	δ	bias	0.002	0.002	0.002	0.002	0.002	0.003	0.001	0.003	0.004	0.006
		RMSE	0.032	0.020	0.015	0.014	0.008	0.031	0.031	0.030	0.032	0.035
	β	bias	0.003	0.000	0.002	0.001	0.001	0.003	0.001	0.004	0.002	0.001
		RMSE	0.031	0.020	0.016	0.014	0.008	0.032	0.029	0.032	0.031	0.032
SBB	λ	bias	-0.005	-0.002	-0.002	-0.002	-0.002	-0.005	0.000	-0.004	-0.003	-0.004
		RMSE	0.025	0.016	0.012	0.011	0.006	0.024	0.025	0.025	0.026	0.025
	δ	bias	0.004	0.002	0.002	0.002	0.002	0.003	0.002	0.003	0.003	0.001
		RMSE	0.027	0.017	0.013	0.012	0.007	0.026	0.029	0.029	0.029	0.030
	β	bias	0.005	0.001	0.002	0.001	0.001	0.005	0.001	0.004	0.002	0.000
		RMSE	0.029	0.019	0.015	0.013	0.007	0.028	0.029	0.031	0.032	0.031
	ρ	bias	0.016	0.003	0.000	-0.002	-0.004	0.018	0.001	0.008	-0.007	-0.006
		RMSE	0.126	0.071	0.058	0.048	0.032	0.128	0.119	0.117	0.121	0.117

Notes: Based on Monte Carlo simulations with 1,000 replications. The parameters in the benchmark scenario are: $N = 60$, $\beta = 1$, $\lambda = 0.3$, $\delta = 0.5$, and $\rho = -0.3$. The labels MLSDV, MAB, SAB, MSB, and SBB denote Modified Least Squares Dummy Variables, Modified Arellano-Bond, Spatial Arellano-Bond, Modified Blundell-Bond, and Spatial Blundell-Bond, respectively.

Acknowledgements

The authors would like to thank Rob Alessie, Peter Egger and J. Paul Elhorst for helpful discussions and participants at the Research Day of the Netherlands Network of Economics (NAKE), Amsterdam, October 2007.

References

Anselin, L. (1988): *Spatial Econometrics: Methods and Models.* Kluwer Academic Publishers, Dordrecht.

——— (2006): "Spatial Econometrics," in *Palgrave Handbook of Econometrics, Vol. 1*, ed. by T. C. Mills, and K. Patterson, pp. 901–969. Palgrave Macmillan, Basingstoke.

Arellano, M. (2003): *Panel Data Econometrics.* Oxford University Press, Oxford, 1st edn.

Arellano, M., and S. 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.

Badinger, H., W. G. Mueller, and G. Tondl (2004): "Regional Convergence in the European Union, 1985–1999: A Spatial Dynamic Panel Analysis," *Regional Studies*, 38, 241–253.

Baltagi, B. (2008): *Econometric Analysis of Panel Data.* John Wiley and Sons, West Sussex, 4th edn.

Baltagi, B., S. Song, B. Jung, and W. Koh (2007): "Testing for Serial Correlation, Spatial Autocorrelation and Random Effects Using Panel Data," *Journal of Econometrics*, 40, 5–51.

Blonigen, B. A., R. B. Davies, G. R. Waddell, and H. T. Naugthon (2007): "FDI in Space: Spatial Autoregressive Relationships in Foreign Direct Investment," *European Economic Review*, 51, 1303–1325.

Blundell, R., and S. Bond (1998): "Initial Conditions and Moment Restrictions in Dynamic Panel Data Models," *Journal of Econometrics*, 87, 115–143.

Brueckner, J. (2003): "Strategic Interaction Among Governments: An Overview of Theoretical Studies," *International Regional Science Review*, 26, 175–188.

Cameron, A. C., and P. K. Trivedi (2005): *Microeconometrics: Methods and Applications*. Cambridge University Press, New York.

Case, A. C., H. S. Rosen, and J. R. Hines (1993): "Budget Spillovers and Fiscal Policy Interdependence: Evidence from the States," *Journal of Public Economics*, 52, 285–307.

Egger, P., M. Pfaffermayr, and H. Winner (2005): "An Unbalanced Spatial Panel Data Approach to US State Tax Competition," *Economics Letters*, 88, 329–335.

Elhorst, J. P. (2003): "Unconditional Maximum Likelihood Estimation of Dynamic Models for Spatial Panels," mimeo, University of Groningen, Groningen.

——— (2008a): "Serial and Spatial Error Correlation," *Economics Letters*, 100, 422–424.

——— (2008b): "Estimation of Dynamic Panels with Endogenous Interaction Effects," mimeo, University of Groningen, Groningen.

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

Jacobs, J. P. A. M., J. E. Ligthart, and H. Vrijburg (2009): "Consumption Tax Competition Among Governments: Evidence from the United States," *International Tax and Public Finance*, forthcoming.

Kapoor, M., H. Kelejian, and I. R. Prucha (2007): "Panel Data Models with Spatially Correlated Error Components," *Journal of Econometrics*, 140, 97–130.

Kelejian, H. H., and D. P. Robinson (1993): "A Suggested Method of Estimation for Spatial Interdependent Models with Autocorrelated Errors, and an Application to a County Expenditure Model," *Papers in Regional Science*, 72, 297–312.

Kim, C. W., T. T. Phipps, and L. Anselin (2003): "Measuring the Benefits of Air Quality Improvement: A Spatial Hedonic Approach," *Journal of Environmental Economics and Management*, 45, 2439.

LeSage, J., and R. K. Pace (2009): *Introduction to Spatial Econometrics*. Taylor and Francis, London.

Nickell, S. (1981): "Biases in Dynamic Models with Fixed Effects.," *Econometrica*, 49, 1417–1426.

Wilson, J. D. (1999): "Theories of Tax Competition," *National Tax Journal*, 52, 269–304.

Yu, J., R. De Jong, and L.-F. Lee (2007): "Quasi-Maximum Likelihood Estimators for Spatial Dynamic Panel Data with Fixed Effects when both N and T are Large," *Journal of Econometrics*, 146, 118–134.