OXFORD BULLETIN OF ECONOMICS AND STATISTICS, 76, 1 (2014) 0305-9049
doi: 10.1111/obes.12011

Estimating and Forecasting with a Dynamic Spatial Panel Data Model*

BADI H. BALTAGI†, BERNARD FINGLETON‡ and ALAIN PIROTTE§,¶

†*Department of Economics and Center for Policy Research, Syracuse University, Syracuse, NY, USA (e-mail: bbaltagi@maxwell.syr.edu)*
‡*Department of Land Economy, University of Cambridge, Cambridge, UK (e-mail: bf100@cam.ac.uk)*
§*ERMES (CNRS), University of Panthéon-Assas Paris II/Sorbonne University, Paris, France (e-mail: alain.pirotte@u-paris2.fr)*
¶*IFSTTAR-DEST, French Institute of Science and Technology for Transport, Development and Networks, Noisy-le-Grand, France (e-mail: alain.pirotte@ifsttar.fr)*

Abstract

This study focuses on the estimation and predictive performance of several estimators for the dynamic and autoregressive spatial lag panel data model with spatially correlated disturbances. In the spirit of Arellano and Bond (1991) and Mutl (2006), a dynamic spatial generalized method of moments (GMM) estimator is proposed based on Kapoor, Kelejian and Prucha (2007) for the spatial autoregressive (SAR) error model. The main idea is to mix non-spatial and spatial instruments to obtain consistent estimates of the parameters. Then, a linear predictor of this spatial dynamic model is derived. Using Monte Carlo simulations, we compare the performance of the GMM spatial estimator to that of spatial and non-spatial estimators and illustrate our approach with an application to new economic geography.

I. Introduction

This study considers spatial panel data models in which there is variation across time and space involving simultaneous spatial (network) dependence together with dynamic interaction. Spatial dependence models are popular in regional science and urban economics, where the cross-sectional units are typically locations (cities, countries, regions) which are affected by common factors or spillover effects from neighbouring locations. Forecasting studies using spatial panel data models are rare, and those involving forecasting with a dynamic component are almost absent from the literature. Recently, Baltagi and Pirotte

*This paper was presented at the Economics Seminar Series, Strathclyde University, Glasgow, Scotland, 30 March 2011. Also, at the Vth Conference of the Spatial Econometrics Association, Toulouse University, France, 6–8 July, 2011. We would like to thank Patrick Sevestre, the participants of this seminar and this conference for their useful comments and suggestions, and also the referees and editor of the journal, Anindya Banerjee, for their help and advice.
JEL Classification number: C33.

(2010) showed that tests of hypotheses based on the usual panel data estimators that ignore spatial dependence can produce misleading inference.

For dynamic panel data models with no spatial autocorrelation, it is well known that the ordinary least squares (OLS) estimator is biased (see Trognon, 1978; Sevestre and Trognon, 1983, 1985). Also, the fixed effects estimator is biased (see Nickell, 1981; Kiviet, 1995). Anderson and Hsiao (1981, 1982) proposed an IV estimator which is consistent. Subsequent developments focused on generalized method of moments (GMM) estimators including Arellano and Bond (1991) and Blundell and Bond (1998) to mention just a few. See Blundell, Bond and Windmeijer (2000), Arellano and Honoré (2001), Hsiao (2003), Harris *et al*. (2008) and Baltagi (2008) for good reviews and a textbook treatment of this subject.

Spatial dependence models deal with spatial interaction and spatial heterogeneity (see Anselin, 1988; LeSage and Pace, 2009). The structure of the spatial dependence can be related to location and distance, both in a geographical space as well as a more general economic or social network space (see Anselin, Le Gallo and Jayet, 2008). Typically, cross-section dependence is modelled as proportional to some observable distance (see Anselin, 1988; LeSage and Pace, 2009) introduced through an endogenous spatial lag variable or via spatially correlated disturbances, or both. Combining cross-section dependence with autoregressive (temporal) dependence leads us to Elhorst (2005), who derives a maximum likelihood estimator (MLE). Another way to estimate autoregressive models with spatial dependence is to extend the GMM approach to the spatial case in order to obtain consistent parameter estimates. Jacobs, Ligthart and Vrijburg (2009) focus on a dynamic autoregressive fixed effects model which includes the spatial lag of the dependent variable together with spatially correlated disturbances. They propose a three-step GMM approach. Elhorst (2010) considers the same model except that the disturbances are not spatially autocorrelated. He develops Bias-corrected least squares dummy variables (BCLSDV), unconditional ML and GMM estimators. Mutl (2006) mixes the Arellano and Bond (1991) and Kapoor, Kelejian and Prucha (2007) approaches to estimate a dynamic model with spatially correlated disturbances under less restrictive assumptions. Yu, de Jong and Lee (2008) propose a quasi maximum likelihood estimator (QMLE) for spatial dynamic panel data with fixed effects when both N and T are large. Lee and Yu (2010 *a*,*b*) extend this approach under different assumptions about N and T. Kukenova and Monteiro (2009) consider a system-GMM to estimate a dynamic spatial panel model (i.e. first order spatial autoregressive panel data model). They compare its properties with those of the usual estimators (MLE, QMLE, LSDV, etc.).

In this study, we propose a spatial GMM estimator in the spirit of Arellano and Bond (1991) and Mutl (2006) under the assumptions that the model includes temporal and spatial lags on the endogenous variable together with SAR-RE disturbances. Only a few articles study the predictive performance of spatial panel models. Baltagi and Li (2006), Fingleton (2009, 2010) and Baltagi, Bresson and Pirotte (2012) focus on the particular case of a static model under spatially correlated disturbances. In contrast, Longhi and Nijkamp (2007) and Kholodilin, Siliverstovs and Kooths (2008) use dynamic spatial models. Longhi and Nijkamp (2007) compare different models designed to compute short-term *ex post* forecasts of regional employment in a panel of 326 West German regions observed over the period from 1987 to 2002. They show that forecasts can be improved by simply taking into account the distances across regions. Nevertheless, this study is specific to the area and variables under investigation. Kholodilin *et al.* (2008) focus on multi-step forecasts of the annual growth rates of real GDP for each of the 16 German Länder simultaneously over the period from 2002 to 2006 (estimation period is 1993–2001). They do not consider any explanatory variables. Moreover, the forecasts are computed using reduced forms and ML estimates of the parameters. In this article, we present a GMM spatial procedure and derive a linear predictor (LP) for the more general case of a spatial dynamic model (i.e. one that includes a spatial lag together with spatially correlated disturbances). Using Monte Carlo simulations and an empirical illustration, we compare the empirical performance of our GMM spatial estimator to that of OLS, within and GMM, each of which takes no account of the spatial structure of the disturbances. We also compare our spatial estimator to other spatial GMM estimators, such as Mutl (2006). Last, we also evaluate the predictive performance of our dynamic spatial model. The plan of the article is as follows: section II presents the model, section III focuses on our spatial GMM estimators. Section IV derives a linear predictor. section V describes the Monte Carlo design. Section VI presents the results, section VII illustrates our approach using an application to new economic geography, and the last section concludes.

II. The spatial dynamic panel model

Consider a first-order spatial autoregressive panel data model

$$y_{it} = \gamma y_{it-1} + \rho_1 \sum_{j=1}^{N} w_{ij} y_{jt} + x_{it}\beta + \varepsilon_{it} \quad i = 1, \ldots, N; \; t = 1, \ldots, T, \tag{1}$$

where y_{it} is the dependent variable for region i at time t, x_{it} is a $(1 \times K)$ vector of explanatory (exogenous) variables, γ and β represent corresponding (1×1) and $(K \times 1)$ parameters to be estimate. w_{ij} is the (i,j) element of the matrix W_N. W_N is an $(N \times N)$ known spatial weights matrix which has zero diagonal elements. If W_N is defined as first-order contiguity, $w_{ij} = 1$ if i and j share a boundary and $w_{ij} = 0$ otherwise. ρ_1 is the spatial lag coefficient. Parameter space must be defined so that $(I_N - \rho_1 W_N)$ is non-singular. This holds when its determinant is non-zero which in turn requires that ρ_1 is not equal to $1/r_i$ for all eigenvalues r_i of the matrix W_N. This is guaranteed by the assumption that ρ_1 is from the interval $]1/r_{\min}, 1/r_{\max}[$. In case of row-normalized matrix $r_{\max} = 1$ and hence the simplification $]1/r_{\min}, 1[$, where $r_{\min}$ equals the most negative purely real characteristic root of W_N (see LeSage and Pace, 2009, pp. 88–89). Moreover, the model (1) is dynamically stable if $[(I_N - \rho_1 W_N)^{-1}\gamma]$ has the largest absolute eigenvalue smaller than one, that is, $|\gamma| < |1 - \rho_1 r_i|$ for all eigenvalues r_i of W_N. This is equivalent to $|\gamma| < 1 - \rho_1 r_{\max} (= 1$ when W_N is row-normalized) for $\rho_1 > 0$, and $|\gamma| < 1 - \rho_1 r_{\min}$ for $\rho_1 < 0$ but the definition of $r_{\min}$ and $r_{\max}$ here does not exclude the complex eigenvalues. The stationarity assumption also requires $|\gamma| < 1$. Moreover, in contrast to the usual panel data framework, we allow ε_{it} to be contemporaneously correlated according to the spatial autoregressive (SAR) error model with row-normalized matrix M_N where it satisfies the same assumptions as W_N but with elements $m_{ij} \neq w_{ij}$:

$$\varepsilon_{it} = \rho_2 \sum_{j=1}^{N} m_{ij} \varepsilon_{jt} + u_{it}, \tag{2}$$

where the assumptions imposed on ρ_2 are like those for ρ_1, in this case with respect to the real eigenvalues of M_N. Moreover, the remainder term u_{it} follows an error component structure

$$u_{it} = \mu_i + v_{it} \tag{3}$$

where μ_i is an individual-specific time-invariant effect which is assumed to be i.i.d. $(0, \sigma^2_\mu)$, and v_{it} is a remainder effect which is assumed to be i.i.d.$(0, \sigma^2_v)$. μ_i and v_{it} are independent of each other and among themselves. Combining (2) and (3), we obtain the SAR-RE specification of the disturbance ε_{it}. For simplicity, in practice, we assume that $M_N = W_N$.

III. A spatial GMM estimator

Following Anderson and Hsiao (1981, 1982) and Arellano and Bond (1991), we eliminate the individual effect μ_i in equation (3), which is correlated with the lagged dependent variable, by differencing the model (1) yielding

$$\Delta y_{it} = \gamma \Delta y_{it-1} + \rho_1 \sum_{j=1}^{N} w_{ij} \Delta y_{jt} + \Delta x_{it}\beta + \Delta \varepsilon_{it}. \tag{4}$$

In contrast to the classical literature on panel data, grouping the data by periods rather than units is more convenient when we consider the spatial correlation due to equation (2). For a cross-section t, we have:

$$\Delta y_t = \gamma \Delta y_{t-1} + \rho_1 W_N \Delta y_t + \Delta x_t \beta + \Delta \varepsilon_t \tag{5}$$

where $y_t = (y_{1t}, \ldots, y_{Nt})'$ is a $(N \times 1)$ vector, and the matrix $x_t = (x_{1t}, \ldots, x_{Nt})'$ is of dimension $(N \times K)$. For the error vector ε_t of dimension $(N \times 1)$, we assume $E(\varepsilon_t) = 0$,

$$E(\varepsilon_t \varepsilon_t') = \sigma^2_\varepsilon (B_N' B_N)^{-1} \tag{6}$$

where $\sigma^2_\varepsilon = \sigma^2_\mu + \sigma^2_v$, $B_N = (I_N - \rho_2 W_N)$. The matrix B_N is assumed to be non-singular, and the row and column sums of the matrix W_N are bounded uniformly in absolute value. The corresponding vector $(N(T-1) \times N(T-1))$ covariance matrix of ε is given by

$$\Omega = (\sigma^2_\mu J_{T-1} + \sigma^2_v I_{T-1}) \otimes (B_N' B_N)^{-1} \tag{7}$$

where J_{T-1} is a $(T-1 \times T-1)$ matrix of ones, I_{T-1}, an identity matrix of order $T-1$.

Using the Arellano and Bond (1991) methodology, we can obtain a GMM estimator based on the following moment conditions:

$$E(y_{il} \Delta v_{it}) = 0 \quad \forall i, l = 1, 2, \ldots, T-2; \quad t = 3, 4, \ldots, T \tag{8}$$

$$E(x_{k,im} \Delta v_{it}) = 0 \quad \forall i, k, \quad m = 1, 2, \ldots, T; \quad t = 3, 4, \ldots, T \tag{9}$$

where equation (9) assumes that the explanatory variables $x_{k,im}$ are strictly exogenous. Moreover, we can use spatially dependent and explanatory variables as instruments. The validity of this strategy requires the following moments conditions:

$$E\left(\sum_{i\neq j} w_{ij} y_{jl} \Delta v_{it}\right) = 0 \quad l = 1, 2, \ldots, T-2; \quad t = 3, 4, \ldots, T \tag{10}$$

$$E\left(\sum_{i\neq j} w_{ij} x_{k,jm} \Delta v_{it}\right) = 0 \quad \forall i, k, \quad m = 1, 2, \ldots, T; \quad t = 3, 4, \ldots, T. \tag{11}$$

Let us define the matrix Z which contains the non-spatial instruments (i.e. related to the conditions (8) and (9)) as

$$Z = \begin{pmatrix} Z_3 & 0 & \cdots & 0 \\ 0 & Z_4 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & Z_T \end{pmatrix} \tag{12}$$

where

$$Z_t = (y_1, \ldots, y_{t-2}, x_1, \ldots, x_T) \tag{13}$$

is an $(N \times (t-2) + KT)$ matrix of instruments at time t, y_l is a vector of dimension $(N \times 1)$ and x_r is a matrix of dimension $(N \times K)$. Moreover, we can define a matrix Z^s which contains the spatial instruments (i.e. related to the conditions (10) and (11)) as

$$Z^s = \begin{pmatrix} Z_3^s & 0 & \cdots & 0 \\ 0 & Z_4^s & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & Z_T^s \end{pmatrix} \tag{14}$$

with

$$Z_t^s = (y_1^s, \ldots, y_{t-2}^s, x_1^s, \ldots, x_T^s) \tag{15}$$

where

$$y_l^s = \begin{pmatrix} \sum_{j=1}^N w_{1j} y_{jl} \\ \sum_{j=1}^N w_{2j} y_{jl} \\ \vdots \\ \sum_{j=1}^N w_{Nj} y_{jl} \end{pmatrix} \tag{16}$$

and

$$x_r^s = \begin{pmatrix} \sum_{j=1}^N w_{1j} x_{1jr} & \sum_{j=1}^N w_{1j} x_{2jr} & \cdots & \sum_{j=1}^N w_{1j} x_{kjr} \\ \sum_{j=1}^N w_{2j} x_{1jr} & \sum_{j=1}^N w_{2j} x_{2jr} & \cdots & \sum_{j=1}^N w_{2j} x_{kjr} \\ \vdots & \vdots & & \vdots \\ \sum_{j=1}^N w_{Nj} x_{1jr} & \sum_{j=1}^N w_{Nj} x_{2jr} & \cdots & \sum_{j=1}^N w_{Nj} x_{kjr} \end{pmatrix}. \tag{17}$$

If we stack the matrices Z and Z^s, we obtain the valid instruments for the model (4), namely Z^*. Moreover, we use the weight matrix of moments

$$A_N = [E[Z'(\Delta\varepsilon)(\Delta\varepsilon)'Z]]^{-1} \tag{18}$$

with

$$E[(\Delta\varepsilon)(\Delta\varepsilon)'] = \sigma_v^2 (I_{T-2} \otimes H_N)(G \otimes I_N)(I_{T-2} \otimes H_N') \tag{19}$$

and

$$G = \begin{pmatrix} 2 & -1 & 0 & \cdots & 0 & 0 \\ -1 & 2 & -1 & \cdots & 0 & 0 \\ 0 & -1 & \ddots & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & \ddots & -1 & 0 \\ 0 & 0 & 0 & -1 & 2 & -1 \\ 0 & 0 & 0 & \cdots & -1 & 2 \end{pmatrix}, \quad H_N = B_N^{-1} = (I_N - \rho_2 W_N)^{-1}. \tag{20}$$

A consistent estimate of the autoregressive parameter ρ_2 can be obtained using the Kapoor, *et al.* (2007) approach, hereafter KKP. In fact, KKP generalized the generalized moments (GM) procedure from cross-section data proposed by Kelejian and Prucha (1999) to panel data and derived its large sample properties when T is fixed and $N \to \infty$. They proposed three GM estimators of ρ_2, σ_v^2 and $\sigma_1^2 (= \sigma_v^2 + T\sigma_\mu^2)$ based on the following six moment conditions:

$$E\begin{bmatrix} \frac{1}{N(T-1)} u_N' Q_{0,N} u_N \\ \frac{1}{N(T-1)} \bar{u}_N' Q_{0,N} \bar{u}_N \\ \frac{1}{N(T-1)} \bar{u}_N' Q_{0,N} u_N \\ \frac{1}{N} u_N' Q_{1,N} u_N \\ \frac{1}{N} \bar{u}_N' Q_{1,N} \bar{u}_N \\ \frac{1}{N} \bar{u}_N' Q_{1,N} u_N \end{bmatrix} = \begin{bmatrix} \sigma_v^2 \\ \sigma_v^2 \frac{1}{N} tr\left(W_N' W_N\right) \\ 0 \\ \sigma_1^2 \\ \sigma_1^2 \frac{1}{N} tr\left(W_N' W_N\right) \\ 0 \end{bmatrix} \tag{21}$$

where

$$u_N = \varepsilon_N - \rho_2 \bar{\varepsilon}_N \tag{22}$$

$$\bar{u}_N = \bar{\varepsilon}_N - \rho_2 \bar{\varepsilon}_N \tag{23}$$

$$\bar{\varepsilon}_N = (I_T \otimes W_N)\varepsilon_N \tag{24}$$

$$\bar{\varepsilon}_N = (I_T \otimes W_N)\bar{\varepsilon}_N. \tag{25}$$

Following equations (21), (22)–(25), and if we consider the sample moments counterparts based on $\widetilde{\varepsilon}_N$, we can write:

$$G_N[\rho_2, \rho_2^2, \sigma_v^2, \sigma_1^2]' - g_N = \xi_N(\rho_2, \sigma_v^2, \sigma_1^2) \tag{26}$$

where

$$G_N = \begin{bmatrix} g_{11}^0 & g_{12}^0 & g_{13}^0 & 0 \\ g_{21}^0 & g_{22}^0 & g_{23}^0 & 0 \\ g_{31}^0 & g_{32}^0 & g_{33}^0 & 0 \\ g_{11}^1 & g_{12}^1 & 0 & g_{13}^1 \\ g_{21}^1 & g_{22}^1 & 0 & g_{23}^1 \\ g_{31}^1 & g_{32}^1 & 0 & g_{33}^1 \end{bmatrix}, \quad g_N = \begin{bmatrix} g_1^0 \\ g_2^0 \\ g_3^0 \\ g_1^1 \\ g_2^1 \\ g_3^1 \end{bmatrix} \tag{27}$$

and

$$\begin{aligned}
g_{11}^i &= \frac{2}{N(T-1)^{1-i}} \widetilde{\varepsilon}'_N Q_{i,N} \widetilde{\bar{\varepsilon}}_N, \quad g_{12}^i = -\frac{1}{N(T-1)^{1-i}} \widetilde{\bar{\varepsilon}}'_N Q_{i,N} \widetilde{\bar{\varepsilon}}_N, \\
g_{21}^i &= \frac{2}{N(T-1)^{1-i}} \widetilde{\bar{\bar{\varepsilon}}}'_N Q_{i,N} \widetilde{\bar{\varepsilon}}_N, \quad g_{22}^i = -\frac{1}{N(T-1)^{1-i}} \widetilde{\bar{\bar{\varepsilon}}}'_N Q_{i,N} \widetilde{\bar{\bar{\varepsilon}}}_N, \\
g_{31}^i &= \frac{1}{N(T-1)^{1-i}} (\widetilde{\varepsilon}'_N Q_{i,N} \widetilde{\bar{\bar{\varepsilon}}}_N + \widetilde{\bar{\varepsilon}}'_N Q_{i,N} \widetilde{\bar{\varepsilon}}_N), \quad g_{32}^i = -\frac{1}{N(T-1)^{1-i}} \widetilde{\bar{\varepsilon}}'_N Q_{i,N} \widetilde{\bar{\bar{\varepsilon}}}_N, \\
g_{13}^i &= 1, \quad g_1^i = \frac{1}{N(T-1)^{1-i}} \widetilde{\varepsilon}'_N Q_{i,N} \widetilde{\varepsilon}_N \\
g_{23}^i &= \frac{1}{N} tr(W'_N W_N), \quad g_2^i = \frac{1}{N(T-1)^{1-i}} \widetilde{\bar{\varepsilon}}'_N Q_{i,N} \widetilde{\bar{\varepsilon}}_N \\
g_{33}^i &= 0, \quad g_3^i = \frac{1}{N(T-1)^{1-i}} \widetilde{\varepsilon}'_N Q_{i,N} \widetilde{\bar{\varepsilon}}_N
\end{aligned}$$

with

$$Q_{0,N} = \left(I_T - \frac{J_T}{T} \right) \otimes I_N \tag{28}$$

$$Q_{1,N} = \frac{J_T}{T} \otimes I_N. \tag{29}$$

KKP specified a static model rather than a dynamic one, that is, $\gamma = 0$, and also their model does not have a spatial lag, that is, $\rho_1 = 0$ in equation (1). In this context, under the random effects specification (3), the OLS estimator of β is consistent. Using $\widetilde{\beta}_{\text{OLS}}$ one gets a consistent estimator of the disturbances $\widetilde{\varepsilon} = y - X\widetilde{\beta}_{\text{OLS}}$. The GM estimators of σ_1^2, σ_v^2 and ρ_2 are the solution of the sample moments (26). KKP suggest three GM estimators. The first involves only the first three moments which do not involve σ_1^2 and yield estimates of ρ_2 and σ_v^2. The fourth moment condition is then used to solve for σ_1^2 given estimates of ρ_2 and σ_v^2. The second GM estimator is based upon weighting the moment equations by the inverse of a properly normalized variance–covariance matrix of the sample moments evaluated at the true parameter values. A simple version of this weighting matrix is derived under normality of the disturbances. The third GM estimator is motivated by computational considerations and replaces a component of the weighting matrix for the second GM estimator by an identity matrix. KKP perform Monte Carlo experiments comparing ML and these three GM estimation methods. They find that on average, the RMSE of ML estimator and their weighted GM estimators are quite similar.[1] $\xi_N(\rho_2, \sigma_v^2, \sigma_1^2)$ can be viewed as a vector of residuals. The GM estimator of ρ_2 is defined as the nonlinear least squares estimator corresponding to equation (26). Here, the OLS estimator of equation (1) is not consistent. Therefore, we modify the KKP approach as follows:

(i) In the first step, we use an IV or GMM estimator to get consistent estimates of γ, ρ_1 and β. For example, we can use the Anderson and Hsiao (1981, 1982) IV estimator by adding $W_N y_{t-2}$ to the list of instruments.

(ii) In the second step, the IV or GMM residuals are used to obtain consistent estimates of the autoregressive parameter ρ_2 and the variance components σ_v^2 and σ_1^2.

- In the third step, we compute the preliminary one-step consistent spatial GMM estimator which is given by

$$\widehat{\delta}_1 = (\Delta\widetilde{X}'Z^*\widehat{A}_N Z^{*'}\Delta\widetilde{X})^{-1}\Delta\widetilde{X}'Z^*\widehat{A}_N Z^{*'}\Delta y \tag{30}$$

where $\Delta\widetilde{X} = (\Delta y_{-1}, (I_{T-2} \otimes W_N)\Delta y, \Delta x)$, $\delta' = (\gamma, \rho_1, \beta')$ and

$$\widehat{A}_N = \left[Z^{*\prime}(I_{T-2} \otimes \widehat{H}_N)(G \otimes I_N)(I_{T-2} \otimes \widehat{H}_N')Z^* \right]^{-1} \tag{31}$$

with $\widehat{H}_N = \widehat{B}_N^{-1} = (I_N - \widehat{\rho}_2 W_N)^{-1}$.

- In the fourth step, following Arellano and Bond (1991), we replace equation (31) in equation (30) by

$$V_N = \left[Z^{*\prime}(I_{T-2} \otimes \widehat{H}_N)(\Delta v)(\Delta v)'(I_{T-2} \otimes \widehat{H}_N')Z^* \right]^{-1}.$$

To operationalize this estimator, Δv is replaced by differenced residuals obtained from the preliminary one-step consistent spatial GMM estimator (30). The resulting estimator is the two-step spatial GMM estimator

$$\widehat{\delta}_2 = (\Delta\widetilde{X}'Z^*\widehat{V}_N Z^{*'}\Delta\widetilde{X})^{-1}\Delta\widetilde{X}'Z^*\widehat{V}_N Z^{*'}\Delta y. \tag{32}$$

IV. Prediction

For the static model with $(\gamma = \rho_1 = 0)$, Goldberger (1962) showed that, for a given Ω, the best linear unbiased predictor (BLUP) for the ith individual at a future period $T+\tau$ is given by:

$$\widehat{y}_{i,T+\tau} = x_{i,T+\tau}\widehat{\beta}_{\text{GLS}} + \omega'\Omega^{-1}\widehat{\varepsilon}_{\text{GLS}} \tag{33}$$

where $\omega = E[\varepsilon_{i,T+\tau}\varepsilon]$ is the covariance between the future disturbance $\varepsilon_{i,T+\tau}$ and the sample disturbances ε. $\widehat{\beta}_{\text{GLS}}$ is the GLS estimator of β based on Ω and $\widehat{\varepsilon}_{\text{GLS}}$ denotes the corresponding GLS residual vector. For the static RE model with $(\gamma = \rho_2 = \rho_1 = 0)$, this predictor in equation (33) was considered by Wansbeek and Kapteyn (1978), Lee and Griffiths (1979) and Taub (1979). The BLUP in equation (33) reduces to:

[1] In our Monte Carlo experiments, we report the results from the GM estimator called the weighted GM estimator by Kapoor *et al*. (2007) in order to save space. The differences relative to the other two GM estimators in our Monte Carlo experiments were minor.

$$\widehat{y}_{i,T+\tau} = x_{i,T+\tau}\widehat{\beta}_{\mathrm{GLS}} + \frac{\sigma_\mu^2}{\sigma_1^2}(\iota_T' \otimes l_i')\widehat{\varepsilon}_{\mathrm{GLS}} \tag{34}$$

where $\sigma_1^2 = T\sigma_\mu^2 + \sigma_v^2$ and l_i is the ith column of I_N. The typical element of the last term of equation (34) is $(T\sigma_\mu^2/\sigma_1^2)\bar{\varepsilon}_{i.,\mathrm{GLS}}$ where $\bar{\varepsilon}_{i.,\,\mathrm{GLS}} = \sum_{t=1}^{T}\widehat{\varepsilon}_{ti,\mathrm{GLS}}/T$. Therefore, the BLUP of $y_{i,T+\tau}$ for the RE model modifies the usual GLS forecasts by adding a fraction of the mean of the GLS residuals corresponding to the ith individual. Baltagi, *et al* (2012) derived the BLUP and demonstrated that the second term of equation (34) is not modified if the structure of the disturbances is SAR-RE with ($\gamma = \rho_1 = 0$). Another interesting case is when the model (1) is a spatial lag model with an error component structure (i.e. $\gamma = \rho_2 = 0$). The BLUP is given by[2]:

$$\widehat{y}_{i,T+\tau} = \sum_{k=1}^{K}\widehat{\beta}_k \sum_{j=1}^{N} h_{ij}x_{k,j,T+\tau} + \frac{T\sigma_\mu^2}{\sigma_1^2}\sum_{j=1}^{N} h_{ij}\widehat{\varepsilon}_{j.} \tag{35}$$

where h_{ij} is the (i,j) element of the matrix $G_N^{-1} = (I_N - \rho_1 W_N)^{-1}$, $\sigma_1^2 = T\sigma_\mu^2 + \sigma_v^2$ and $\widehat{\varepsilon}_{j.} = (1/T)\sum_{t=1}^{T}\widehat{\varepsilon}_{jt}$. In practice, the variance components and the spatial lag parameter ρ_1 of equation (35) are unknown. So, these are replaced by their ML estimates.

When $\gamma \neq 0$ (i.e. a dynamic model), and $\rho_1 \neq 0$ (i.e. including a spatial lag on the dependent variable) and $\rho_2 \neq 0$ (i.e. including a SAR process on the disturbances ε), the derivation of the predictor is more complicated mainly because the lagged endogenous variable is correlated with the individual effects. Following Chamberlain (1984) and Sevestre and Trognon (1996), we derive the linear predictor of y_{it} conditional upon $(y_{10},\ldots,y_{N0},x_{11},\ldots,x_{N1},\ldots,x_{1T},\ldots,x_{NT})$ which is given[3] by

$$\begin{aligned} &E^*[y_{it}\,|\,y_{10},\ldots,y_{N0},x_{11},\ldots,x_{N1},\ldots,x_{1T},\ldots,x_{NT}] \\ &= \gamma^t \sum_{j=1}^{N} h_{ij}^{(t)} y_{j0} + \sum_{l=1}^{t}\gamma^{l-1}\sum_{j=1}^{N} h_{ij}^{(l)} x_{jt-l+1}\beta \\ &+ \sum_{l=1}^{t}\gamma^{l-1}\sum_{j=1}^{N} p_{ij}^{(l)} E^*[\mu_j\,|\,y_{10},\ldots,y_{N0}], \end{aligned} \tag{36}$$

where $h_{ij}^{(l)}$ is the (i,j) element of the matrix $(G_N^{-1})^l$, $p_{ij}^{(l)}$ is the (i,j) element of the matrix $\left(((G_N)^{-1})^l B_N^{-1}\right)$. μ_j and y_{j0} are assumed to be uncorrelated with the sequence $(x_{j1},\ldots,x_{jT})$, $\forall j$. Following Chamberlain (1984), we assume that:

$$E^*[\mu_j|y_{10},\ldots,y_{N0}] = \psi + \lambda_1 y_{10} + \cdots + \lambda_N y_{N0} = \psi + \lambda' y_0 \tag{37}$$

where λ is an $N \times 1$ vector with

$$\lambda = V[y_0]^{-1}\mathrm{cov}[\mu_j, y_0] = V[y_0]^{-1}E[\mu_j y_0] \tag{38}$$

[2]The proof is available at http://www.spatialeconomics.ac.uk/textonly/SERC/publications/download/sercdp0095.pdf.

[3]To facilitate the presentation, we consider only one exogenous variable. The proof is available at http://www.spatialeconomics.ac.uk/textonly/SERC/publications/download/sercdp0095.pdf.

and

$$\psi = E[\mu_j] - \lambda_1 E[y_{10}] - \cdots - \lambda_N E[y_{N0}]. \tag{39}$$

Following Sevestre and Trognon (1983), Chamberlain (1984) and Mutl (2006), we can write:

$$E^*[\mu_j | y_{10}, \ldots, y_{N0}] = \lambda'[y_0 - E[y_0]] \tag{40}$$

with

$$y_0 = \frac{1}{1-\gamma} P_N \mu_N + \frac{1}{\sqrt{1-\gamma^2}} P_N v_0, \tag{41}$$

$$V[y_0] = \left(\frac{\sigma_\mu^2}{(1-\gamma)^2} + \frac{\sigma_v^2}{1-\gamma^2} \right) P_N P_N', \tag{42}$$

$$E[\mu_j y_0] = \frac{\sigma_\mu^2}{1-\gamma} P_N l_j \tag{43}$$

where $P_N = (B_N G_N)^{-1}$ and l_j is the jth column of I_N.

V. Monte Carlo design

In all the experiments the dependent variable y_{it} was generated from a model of the form:

$$y_{it} = a + \gamma y_{it-1} + \rho_1 \sum_{j=1}^{N} w_{ij} y_{jt} + \beta x_{it} + \varepsilon_{it} \quad i = 1, \ldots, N; \quad t = 1, \ldots, T, \tag{44}$$

where the disturbance ε_{it} follows a SAR process:

$$\varepsilon_{it} = \rho_2 \sum_{j=1}^{N} w_{ij} \varepsilon_{jt} + u_{it} \tag{45}$$

where w_{ij} is the (i,j) element of the spatial matrix W_N, and u_{it} has an error component structure

$$u_{it} = \mu_i + v_{it} \tag{46}$$

with $\mu_i \sim \text{i.i.d.}N(0, \sigma_\mu^2)$, $v_{it} \sim \text{i.i.d.}N(0, \sigma_v^2)$ and $(\sigma_\mu^2, \sigma_v^2) = (0.8, 0.2), (0.2, 0.8)$. The explanatory variable x_{it} is generated as:

$$x_{it} = \delta x_{it-1} + \xi_{it} \tag{47}$$

with $\delta = 0.6$, $\xi_{it} \sim \text{i.i.d.}N(0, \sigma_\xi^2)$, $\sigma_\xi^2 = 5$ and $x_{i0} = 0$. One sample size is considered $(N, T) = (100, 7)$. For the coefficients of equations (44) and (45), we assume the values $(a, \gamma, \beta) = (1, 0.2, 1)$, $(1, 0.5, 1)$, $\rho_1 = 0.7$, 0.4, 0.2 and $\rho_2 = 0.4$. The first 10 cross-sections were discarded in order to reduce the dependency on initial values. Moreover, following Kelejian and Prucha (1999), two weight matrices are used which essentially differ in their degree of sparseness. The weight matrices are labelled as 'j ahead and j behind' with the non-zero elements being $1/2j, j = 1$ and 5, hereafter respectively $W(1,1)$ and $W(5,5)$. For all experiments, 1,000 replications are performed. We compute the mean, standard deviation, bias and RMSE of the coefficients $\widehat{\gamma}$, $\widehat{\beta}$, $\widehat{\rho}_1$ and $\widehat{\rho}_2$.[4] Following Kapoor *et al*. (2007), we adopt a measure of dispersion which is closely related to the standard measure of root mean square error (RMSE), but is based on quantiles. It is defined as

$$\mathrm{RMSE} = \left[\mathrm{bias}^2 + \left(\frac{\mathrm{IQ}}{1.35} \right)^2 \right]^{1/2} \tag{48}$$

where bias is the difference between the median and the true value of the parameter, and IQ is the interquantile range defined as $c_1 - c_2$ where c_1 is the 0.75 quantile and c_2 is the 0.25 quantile. Clearly, if the distribution is normal the median is the mean and, aside from a slight rounding error, IQ/1.35 is the standard deviation. In this case, the measure (48) reduces to the standard RMSE.

We compare the performance of eight estimators in our Monte Carlo experiments. These are as follows:

(i) Ordinary least squares which does not deal with the endogeneity of the spatial lag $W_N y$ and the endogeneity of the lagged dependent variable. OLS also ignores the individual effects and the SAR process for the disturbances.
(ii) The Within (W) estimator which wipes out the individual effects, but otherwise does not deal with the endogeneity of the spatial lag $W_N y$ and the endogeneity of the lagged dependent variable nor the SAR process for the disturbances.
(iii) The Arellano and Bond (1991) GMM (1) estimator which differences the individual effects, and handles the presence of the lagged dependent variable by using the orthogonality conditions (8) and (9). However, this estimator ignores the spatial lag $W_N y$ and the SAR process for the disturbances.
(iv) GMM (2) is an application of the Arellano and Bond (1991) estimator as in GMM (1) but including the spatial lag $W_N y$ as an extra regressor. This estimator ignores the SAR process for the disturbances.
(v) GMM-SAR-RE (1) is the estimator suggested by Mutl (2006) which accounts for the lagged dependent variable and the SAR-RE process in the spirit of KKP (2007). However, this estimator ignores the spatial lag $W_N y$.
(vi) GMM-SAR-RE (2) is an application of the Mutl (2006) estimator as in GMM-SAR-RE (1) but including the spatial lag $W_N y$ as an extra regressor but does not accounted for its endogeneity.
(vii) GMM-SL-RE is an estimator that uses the orthogonality conditions (8) and (9) of Arellano and Bond (1991) as well as the spatial orthogonality conditions (10) and (11). However, this estimator ignores the SAR-RE process for the disturbances. This GMM estimator uses similar orthogonality conditions to those of Elhorst (2010).

[4]As one referee suggests, it would have been more interesting to compare the performance of the marginal effects rather than that of the parameter estimates themselves, see LeSage and Pace (2009). Due to space limitations, we have not done so in this article, but we agree that this is important for practitioners and should be the focus of all future Monte Carlo experiments.

(viii) GMM-SL-SAR-RE is an estimator that uses the orthogonality conditions (8) and (9) of Arellano and Bond (1991) as well as the spatial orthogonality conditions (10) and (11) as in GMM-SL-RE, but it also accounts for the SAR-RE structure of the disturbances using a KKP approach. This estimator is described in section III.

Last, we check the prediction-performance of the estimators considered. We use the RMSE criterion and compute the out of sample forecast errors for each predictor associated with the alternative estimators for one to five step ahead forecasts (see Baltagi *et al.* 2012). We also compute the Theil's U statistic to avoid the scaling problem of RMSE, as discussed by Trapani and Urga (2009). Average values of RMSE are calculated across N for all five step forecasts.[5]

VI. Monte Carlo results

Mean, bias and RMSE of the estimators

Table 1 presents the mean, bias and RMSE of the coefficients $\widehat{\gamma}, \widehat{\rho}_1, \widehat{\rho}_2$ and $\widehat{\beta}$ for $(N, T) = (100, 7), (\sigma_\mu^2, \sigma_v^2) = (0.8, 0.2)$, $\beta = 1$, $\rho_2 = 0.4$ considering a W(1,1) matrix, one neighbour ahead and one neighbour behind.

The first panel in Table 1 shows the results for $\gamma = 0.2$ (low value of the lagged dependent variable coefficient) and $\rho_1 = 0.2$ and 0.7, (low and high values of the spatial lag coefficient). The second panel in Table 1 shows the results for $\gamma = 0.5$ (high value of the lagged dependent variable coefficient). In particular, for $\rho_1 = 0.2$ in Table 1, the GMM estimator associated with model (1) (i.e. including both the spatial lag variable and SAR-RE process, namely GMM-SL-SAR-RE) has lower RMSEs for the coefficients ρ_1 and ρ_2 than those obtained for GMM-SAR-RE (2) and GMM spatial lag including RE disturbances, namely GMM-SL-RE. The RMSEs of γ and β are similar. The GMM-SAR-RE (2) estimator is an extension of the Mutl (2006) estimator (GMM-SAR-RE (1)) with the additional regressor $W_N y$. GMM-SL-RE is an estimator that uses the orthogonality conditions (8) and (9) of Arellano and Bond (1991) as well as the spatial orthogonality conditions (10) and (11). However, this estimator ignores the SAR-RE process for the disturbances. Note the huge bias and RMSE for OLS, Within and GMM (1) for the coefficients γ, ρ_1, β. These results are not surprising. OLS does not deal with the endogeneity of the spatial $W_N y$ and the endogeneity of the lagged dependent variable. OLS also ignores the individual effects and the SAR process for the disturbances. The Within estimator does not deal with the endogeneity of the spatial lag $W_N y$ and the endogeneity of the temporally lagged dependent variable nor the SAR process for the disturbances. Moreover GMM (1) ignores the spatial lag $W_N y$ and the SAR process for the disturbances.

Consider next the outcomes in the top panel of Table 1 for $\gamma = 0.2$ (i.e. low value of the lagged dependent variable coefficient) but for the high value of the spatial lag coefficient ($\rho_1 = 0.7$). In this case, GMM-SL-SAR-RE is the *best* in terms of RMSE for the coefficients ρ_1, ρ_2, γ and β. Note that the bias and RMSE for OLS, Within and GMM (1) for the coefficients γ, ρ_1, β increases. It is apparent that the larger value of spatial lag coefficient improves the results produced by the GMM-SL-SAR-RE estimator.

[5]Tables reporting Theil's U statistic are available upon request from the authors.

TABLE 1

Mean, bias and RMSE of the coefficients $\hat{\gamma}, \hat{\rho}_1, \hat{\beta}$ and $\hat{\rho}_2$ for W(1,1), $\sigma_\mu^2 = 0.8$, $\sigma_v^2 = 0.2$, (N,T) = (100,7), 1,000 replications

		Non-spatial estimators						Spatial estimators									
		OLS		Within		GMM (1) (without spatial lag variable)		GMM (2) (with spatial lag variable)		GMM-SAR-RE (1) (without spatial lag variable)		GMM-SAR-RE (2) (with spatial lag lag variable)		GMM-SL-RE		GMM-SL-SAR-RE	
ρ_1		0.7	0.2	0.7	0.2	0.7	0.2	0.7	0.2	0.7	0.2	0.7	0.2	0.7	0.2	0.7	0.2
$\gamma = 0.2$, $\beta = 1$ and $\rho_2 = 0.4$ (SAR-RE)																	
$\hat{\gamma}$	Mean	0.2349	0.3275	0.1806	0.1868	0.3914	0.2148	0.1825	0.1930	0.3914	0.2112	0.1878	0.2005	0.1937	0.1967	0.1957	0.1981
	Bias	0.0349	0.1275	−0.0193	−0.0131	0.1914	0.0148	−0.0174	−0.0070	0.1914	0.0112	−0.0121	0.0005	−0.0062	−0.0032	−0.0042	−0.0018
	RMSE	0.0364	0.1289	0.0202	0.0160	0.1914	0.0192	0.0187	0.0119	0.1914	0.0150	0.0134	0.0085	0.0090	0.0099	0.0076	0.0087
$\hat{\rho}_1$	Mean	0.7347	0.2718	0.7328	0.2404	—	—	0.7314	0.2393	—	—	0.7273	0.2318	0.7120	0.2115	0.7080	0.2051
	Bias	0.0347	0.0718	0.0328	0.0404	—	—	0.0314	0.0393	—	—	0.0273	0.0318	0.0120	0.0115	0.0080	0.0051
	RMSE	0.0373	0.0762	0.0340	0.0428	—	—	0.0330	0.0416	—	—	0.0287	0.0346	0.0153	0.0181	0.0122	0.0145
$\hat{\beta}$	Mean	0.9489	0.9111	0.9896	1.0021	1.2715	1.0052	0.9897	1.0003	1.2004	0.9984	0.9912	0.9988	0.9959	1.0004	1.0012	1.0018
	Bias	−0.0510	−0.0888	−0.0103	0.0021	0.2715	0.0052	−0.0102	0.0003	0.2004	−0.0015	−0.0087	−0.0011	−0.0040	0.0004	0.0012	0.0018
	RMSE	0.0547	0.0911	0.0136	0.0100	0.2719	0.0138	0.0139	0.0099	0.2024	0.0107	0.0124	0.0098	0.0097	0.0099	0.0084	0.0098
$\hat{\rho}_2$	Mean	—	—	—	—	—	—	—	—	0.5426	0.4826	0.3387	0.3434	—	—	0.3911	0.3921
	Bias	—	—	—	—	—	—	—	—	0.1426	0.0826	−0.0612	−0.0565	—	—	−0.0088	−0.0078
	RMSE	—	—	—	—	—	—		—	0.1605	0.1202	0.1107	0.1055	—	—	0.1072	0.0992
		0.4	0.2	0.4	0.2	0.4	0.2	0.4	0.2	0.4	0.2	0.4	0.2	0.4	0.2	0.4	0.2
$\gamma = 0.5$, $\beta = 1$ and $\rho_2 = 0.4$ (SAR-RE)																	
$\hat{\gamma}$	Mean	0.5680	0.5985	0.4789	0.4852	0.6163	0.5245	0.4836	0.4913	0.6121	0.5216	0.4924	0.4990	0.4942	0.4964	0.4973	0.4983
	Bias	0.0680	0.0985	−0.0210	−0.0147	0.1163	0.0245	−0.0164	−0.0086	0.1121	0.0216	−0.0076	−0.0009	−0.0057	−0.0035	−0.0026	−0.0016
	RMSE	0.0691	0.0995	0.02227	0.0168	0.1165	0.0270	0.0181	0.0122	0.1122	0.0234	0.0107	0.0074	0.0097	0.0088	0.0076	0.0076
$\hat{\rho}_1$	Mean	0.4177	0.2418	0.4314	0.2292	—	—	0.4277	0.2269	—	—	0.4197	0.2202	0.4098	0.2085	0.4045	0.2032
	Bias	0.0177	0.0419	0.0314	0.0292	—	—	0.0277	0.0269	—	—	0.0197	0.0202	0.0098	0.0085	0.0046	0.0032
	RMSE	0.0229	0.0451	0.0328	0.0313	—	—	0.0293	0.0293	—	—	0.0224	0.0233	0.0144	0.0143	0.0110	0.0120
$\hat{\beta}$	Mean	0.9396	0.9200	0.9988	1.0014	1.0298	1.0034	0.9989	1.0006	1.0046	0.9938	0.9949	0.9967	0.9995	1.0002	1.0013	1.0010
	Bias	−0.0603	−0.0799	−0.0011	0.0014	0.0298	0.0034	−0.0010	0.0006	0.0046	−0.0061	−0.0050	−0.0032	−0.0004	0.0002	0.0013	0.0010
	RMSE	0.0631	0.0819	0.0092	0.0098	0.0335	0.0128	0.0097	0.0098	0.0128	0.0121	0.0105	0.0097	0.0095	0.0097	0.0088	0.0090
$\hat{\rho}_2$	Mean	—	—	—	—	—	—	—	—	0.5117	0.4959	0.3367	0.3397	—	—	0.3910	0.3919
	Bias	—	—	—	—	—	—	—	—	0.1117	0.0959	−0.0632	−0.0602	—	—	−0.0089	−0.0080
	RMSE	—	—	—	—	—	—	—	—	0.1349	0.1289	0.1139	0.1111	—	—	0.1084	0.1056

If we consider the second panel in Table 1, that is $\gamma = 0.5$ (high value of the lagged dependent variable coefficient) and $\rho_1 = 0.2$ (low value of the spatial lag coefficient), GMM-SL-SAR-RE is still the best in terms of RMSE for the coefficients ρ_1, γ, β and ρ_2. OLS, Within and GMM (1) estimation of the coefficients γ, ρ_1, β produces more bias and a larger RMSE compared with the outcomes with $\gamma = 0.2$. If we maintain the value of γ at 0.5 but increase ρ_1 so that it takes the value 0.4 (i.e. a higher value for the spatial lag coefficient), the GMM estimator for the model (1) (i.e. including spatial lag variable and SAR-RE process) remains the best in terms of RMSE for the coefficients ρ_1, γ, β and ρ_2. In other words the results are the same, but the magnitudes are different.

Table 2 considers a lower level of individual heterogeneity, $(\sigma_\mu^2, \sigma_v^2) = (0.2, 0.8)$ but otherwise leaves all the other parameters as in Table 1. In terms of RMSE, the GMM-SL-SAR-RE estimator remains the *best* whatever the values of ρ_1 and γ. Overall, the RMSEs are greater than those of Table 1, where we assumed that $(\sigma_\mu^2, \sigma_v^2) = (0.8, 0.2)$.[6]

Forecast accuracy

Table 3 gives the RMSEs for the 1- and 5-year ahead forecasts along with the average RMSE for all 5 years. These are out-of-sample forecasts when the true model is (1) (i.e. including both temporally and spatially lagged dependent variables) with SAR-RE disturbances. The sample size is $N = 100$ and $T = 7$, the weights matrix is $W(1, 1)$, that is, one neighbour behind and one neighbour ahead.

If we consider the case $(\sigma_\mu^2, \sigma_v^2) = (0.8, 0.2)$ and $(\gamma, \rho_1) = (0.2, 0.2)$, the lowest RMSE is that of GMM-SL-SAR-RE, followed closely by the RSMEs produced by the GMM-SAR-RE and GMM-SL-RE estimators. It appears that misspecifying the disturbances affects the forecast performance, as the RMSEs of the OLS and Within estimators are approximately double those produced by GMM-SL-SAR-RE. If the spatial lag parameter ρ_1 increases from 0.2 to 0.7, the RMSEs increase sharply, but the ranking remains intact. If $(\gamma, \rho_1) = (0.5, 0.2)$ or $(\gamma, \rho_1) = (0.5, 0.4)$, the results remain essentially the same, although the magnitudes are different. The GMM estimators that take endogeneity and/or heterogeneity into account perfom better than the OLS and Within estimators that do not. This forecast comparison is robust to whether we are predicting one period ahead or five periods ahead and is also reflected in the average over the five years. Thus the gain in forecast performance is substantial once we account for endogeneity and heterogeneity.

If we consider the case of a low level of individual heterogeneity, that is $(\sigma_\mu^2, \sigma_v^2) = (0.2, 0.8)$, the lowest RMSE remains that of GMM-SL-SAR-RE. However, the difference between the GMM-SL-SAR-RE and GMM-SL-RE outcomes is reduced.

Sensitivity analysis

For the various estimators considered, Tables 4 and 5 report the mean, bias and RMSE results, as was done for Tables 1 and 2 except that the weight matrix is changed from a $W(1, 1)$ to $W(5, 5)$, that is, five neighbours behind and five neighbours ahead. Table 6 reports

[6]Additional Monte Carlo experiments confirm the robustness of GMM-SL-SAR-RE estimator to misspecification of W_N and/or M_N. Supplementary tables showing this are available upon request.

TABLE 2

Mean, Bias and RMSE of the coefficients $\hat{\gamma}, \hat{\rho}_1, \hat{\beta}\ \hat{\rho}_2$ and for W(1,1), $\sigma_\mu^2 = 0.2$, $\sigma_v^2 = 0.8$, (N,T) = (100,7), 1,000 replications

		Non-spatial estimators						*Spatial estimators*									
		OLS		*Within*		*GMM (1) (without spatial lag variable)*		*GMM (2) (with spatial lag variable)*		*GMM-SAR-RE (1) (without spatial lag variable)*		*GMM-SAR-RE (2) (with spatial lag variable)*		*GMM-SL-RE*		*GMM-SL-SAR-RE*	
	ρ_1	0.7	0.2	0.7	0.2	0.7	0.2	0.7	0.2	0.7	0.2	0.7	0.2	0.7	0.2	0.7	0.2
$\gamma = 0.2, \beta = 1$ and $\rho_2 = 0.4$ (SAR-RE)																	
$\hat{\gamma}$	Mean	0.1671	0.2210	0.1395	0.1555	0.3974	0.2153	0.1440	0.1728	0.3985	0.2117	0.1508	0.1835	0.1773	0.1872	0.1842	0.1931
	Bias	−0.0330	0.0210	−0.0604	−0.0444	0.1974	0.0153	−0.0559	−0.0271	0.1985	0.0116	−0.0491	−0.0164	−0.0226	−0.0127	−0.0157	−0.0068
	RMSE	0.0352	0.0266	0.0614	0.0476	0.2015	0.0279	0.0573	0.0325	0.2000	0.0234	0.0507	0.0252	0.0258	0.0226	0.0198	0.0178
$\hat{\rho}_1$	Mean	0.7794	0.2976	0.8033	0.3374	—	—	0.8008	0.3350	—	—	0.7945	0.3233	0.7429	0.2434	0.7290	0.2187
	Bias	0.0794	0.0976	0.1033	0.1374	—	—	0.1008	0.1349	—	—	0.0945	0.1233	0.0429	0.0434	0.0290	0.0187
	RMSE	0.0808	0.1006	0.1046	0.1405	—	—	0.1021	0.1387	—	—	0.0963	0.1273	0.0462	0.0519	0.0339	0.0329
$\hat{\beta}$	Mean	0.9834	0.9855	0.9676	1.0072	1.2712	1.0045	0.9668	1.0017	1.2181	1.0006	0.9732	1.0021	0.9858	1.0016	1.0048	1.0069
	Bias	−0.0165	−0.0144	−0.0323	0.0072	0.2712	0.0045	−0.0331	0.0016	0.2181	0.0006	−0.0268	0.0021	−0.0142	0.0016	0.0048	0.0069
	RMSE	0.0236	0.0234	0.0373	0.0202	0.2746	0.0246	0.0383	0.0195	0.2203	0.0216	0.0330	0.0191	0.0232	0.0199	0.0177	0.0205
$\hat{\rho}_2$	Mean	—	—	—	—	—	—	—	—	0.4132	0.3711	0.1710	0.1546	—	—	0.3880	0.3880
	Bias	—	—	—	—	—	—	—	—	0.0132	−0.0288	−0.2290	−0.2453	—	—	−0.0120	−0.0120
	RMSE	—	—	—	—	—	—	—	—	0.0783	0.1032	0.2512	0.2693	—	—	0.0981	0.0973
		0.4	0.2	0.4	0.2	0.4	0.2	0.4	0.2	0.4	0.2	0.4	0.2	0.4	0.2	0.4	0.2
$\gamma = 0.5, \beta = 1$ and $\rho_2 = 0.4$ (SAR-RE)																	
$\hat{\gamma}$	Mean	0.4901	0.5155	0.4269	0.4468	0.6118	0.5242	0.4401	0.4650	0.6122	0.5221	0.4505	0.4753	0.4783	0.4858	0.4894	0.4931
	Bias	−0.0098	0.0155	−0.0730	−0.0531	0.1118	0.0242	−0.0599	−0.0349	0.1122	0.0221	−0.0494	−0.0246	−0.0216	−0.0141	−0.0105	−0.0068
	RMSE	0.0164	0.0201	0.0743	0.0550	0.1148	0.0315	0.0620	0.0386	0.1139	0.0284	0.0532	0.0311	0.0266	0.0219	0.0176	0.0164
$\hat{\rho}$	Mean	0.4566	0.2585	0.5082	0.3028	—	—	0.4984	0.2964	—	—	0.4883	0.2867	0.4364	0.2314	0.4168	0.2115
	Bias	0.0566	0.0585	0.1082	0.1028	—	—	0.0984	0.0964	—	—	0.0883	0.0867	0.0365	0.0314	0.0168	0.0115
	RMSE	0.0590	0.0606	0.1098	0.1051	—	—	0.1003	0.0991	—	—	0.0924	0.0904	0.0423	0.0392	0.0266	0.0256
$\hat{\beta}$	Mean	0.9950	0.9867	0.9962	1.0052	1.0260	1.0025	0.9960	1.0026	1.0110	0.9971	0.9940	0.9994	0.9985	1.0011	1.0050	1.0042
	Bias	−0.0050	−0.0132	−0.0038	0.0052	0.0258	0.0025	−0.0040	0.0026	0.0108	−0.0028	−0.0060	−0.0005	−0.0014	0.0011	0.0050	0.0042
	RMSE	0.0192	0.0231	0.0180	0.0195	0.0377	0.0240	0.0190	0.0186	0.0263	0.0209	0.0186	0.0187	0.0188	0.0191	0.0187	0.0186
$\hat{\rho}$	Mean	—	—	—	—	—	—	—	—	0.3724	0.3469	0.1261	0.13121	—	—	0.3880	0.3881
	Bias	—	—	—	—	—	—	—	—	−0.0275	−0.0530	−0.2739	−0.2688	—	—	−0.0120	−0.0118
	RMSE	—	—	—	—	—	—	—	—	0.0826	0.1163	0.2986	0.2960	—	—	0.0978	0.0972

TABLE 3

Forecasting RMSE for W(1,1), (N.T) = (100,7), 1,000 replications

	Non-spatial estimators						*Spatial estimators*									
	OLS		*Within*		*GMM (1) (without spatial lag variable)*		*GMM (2) (with spatial lag variable)*		*GMM-SAR-RE (1) (without spatial lag variable)*		*GMM-SAR-RE (2) (with spatial lag variable)*		*GMM-SL-RE*		*GMM-SL-SAR-RE*	
ρ_1	0.7	0.2	0.7	0.2	0.7	0.2	0.7	0.2	0.7	0.2	0.7	0.2	0.7	0.2	0.7	0.2
$\sigma_\mu^2 = 0.8$, $\sigma_v^2 = 0.2$																
$\gamma = 0.2$, $\beta = 1$ and $\rho_2 = 0.4$ (SAR-RE)																
First year	4.6032	1.3293	4.0919	1.2549	3.0533	0.7389	4.0456	0.7702	4.1043	0.7730	2.2110	0.6723	2.9988	0.7209	2.1353	0.6670
Second year	4.7351	1.3890	4.1787	1.3048	3.1466	0.7952	4.2545	0.8334	4.3176	0.8373	2.3158	0.7285	3.0893	0.7772	2.2379	0.7230
Third year	4.8328	1.4217	4.2222	1.3239	3.1919	0.8175	4.3692	0.8617	4.4329	0.8655	2.3697	0.7512	3.1320	0.7993	2.2885	0.7456
Fourth year	4. 9097	1.4382	4.2534	1.3349	3.2225	0.8298	4.4161	0.8792	4.4752	0.8825	2.4044	0.7637	3.1613	0.8116	2.3216	0.7582
Fifth year	4.9777	1.4470	4.2756	1.3417	3.2458	0.8380	4.5084	0.8934	4.5661	0.8967	2.4336	0.7718	3.1823	0.8195	2.3468	0.7661
5-year average	4.8117	1.4050	4.2044	1.3120	3.1720	0.8039	4.3188	0.8476	4.3792	0.8510	2.3469	0.7375	3.1128	0.7857	2.2660	0.7320
	0.4	0.2	0.4	0.2	0.4	0.2	0.4	0.2	0.4	0.2	0.4	0.2	0.4	0.2	0.4	0.2
$\gamma = 0.5$, $\beta = 1$ and $\rho_2 = 0.4$ (SAR-RE)																
First year	3.8144	2.2716	3.9185	2.1698	3.6442	1.2856	2.0922	1.0203	3.6737	1.2921	1.7741	0.9068	2.0834	1.0025	1.6326	0.8913
Second year	3.9838	2.3365	4.0194	2.2097	3.7328	1.3204	2.1673	1.0686	3.7639	1.3286	1.6935	0.9561	2.1585	1.0511	1.7090	0.9409
Third year	4.1364	2.3800	4.0954	2.2302	3.7875	1.3388	2.2175	1.0922	3.8187	1.3477	1.8306	0.9808	2.2086	1.0747	1.7605	0.9656
Fourth year	4.2790	2.4104	4.1604	2.2445	3.8318	1.3521	2.2579	1.1075	3.8617	1.3607	1.8765	0.9967	2.2488	1.0902	1.8018	0.9816
Fifth year	4.4097	2.4302	4.2156	2.2548	3.9029	1.3700	2.2929	1.1191	3.9329	1.3785	1.9173	1.0088	2.2827	1.1016	1.8373	0.9936
5-year average	4.1247	2.3657	4.0819	2.2218	3.7799	1.3334	2.2056	1.0816	3.8102	1.3415	1.8184	0.9698	2.1964	1.0640	1.7482	0.9546
	0.7	0.2	0.7	0.2	0.7	0.2	0.7	0.2	0.7	0.2	0.7	0.2	0.7	0.2	0.7	0.2
$\sigma_\mu^2 = 0.2$, $\sigma_v^2 = 0.8$																
$\gamma = 0.2$, $\beta = 1$ and $\rho_2 = 0.4$ (SAR-RE)																
First year	3.2639	1.1302	3.2929	1.1459	4.1161	1.1294	3.2314	1.1085	4.1112	1.1273	3.2106	1.1015	2.8654	1.0657	2.7851	1.0611
Second year	3.5015	1.2420	3.5361	1.2569	4.3990	1.2435	3.4723	1.2200	4.3983	1.2417	3.4517	1.2130	3.0780	1.1742	2.9972	1.1695
Third year	3.6148	1.2877	3.6663	1.3005	4.5385	1.2904	3.5983	1.2641	4.5389	1.2886	3.5752	1.2575	3.1752	1.2178	3.0939	1.2130
Fourth year	3.6796	1.3121	3.7331	1.3244	4.5961	1.3169	3.6646	1.2874	4.5926	1.3149	3.6407	1.2808	3.2360	1.2417	3.1541	1.2370
Fifth year	3.7487	1.3272	3.8083	1.3414	4.6983	1.3362	3.7422	1.3043	4.6942	1.3342	3.7189	1.2973	3.2824	1.2569	3.1992	1.2519
5-year average	3.5617	1.2598	3.6073	1.2738	4.4696	1.2633	3.5418	1.2369	4.4670	1.2613	3.5194	1.2301	3.1274	1.1913	3.0459	1.1865

TABLE 3

Continued

	Non-spatial estimators						*Spatial estimators*									
	OLS		*Within*		*GMM (1) (without spatial lag variable)*		*GMM (2) (with spatial lag variable)*		*GMM-SAR-RE (1) (without spatial lag variable)*		*GMM-SAR-RE (2) (with spatial lag variable)*		*GMM-SL-RE*		*GMM-SL-SAR-RE*	
	0.4	0.2	0.4	0.2	0.4	0.2	0.4	0.2	0.4	0.2	0.4	0.2	0.4	0.2	0.4	0.2
$\gamma=0.5$, $\beta=1$ and $\rho_2=0.4$ (SAR-RE)																
First year	2.6342	1.5787	2.6681	1.5857	3.4751	1.6395	2.4435	1.4514	3.4721	1.6353	2.4817	1.4523	2.2790	1.3861	2.2328	1.3766
Second year	2.7856	1.6802	2.8118	1.6863	3.5964	1.7278	2.5857	1.5521	3.5939	1.7240	2.6264	1.5523	2.4128	1.4871	2.3661	1.4774
Third year	2.8788	1.7290	2.8958	1.7343	3.6601	1.7705	2.6688	1.5995	3.6571	1.7667	2.7150	1.5999	2.4901	1.5353	2.4433	1.5256
Fourth year	2.9530	1.7598	2.9590	1.7648	3.7033	1.7977	2.7301	1.6284	3.6988	1.7937	2.7819	1.6289	2.5480	1.5654	2.5006	1.5558
Fifth year	3.0208	1.7817	3.0177	1.7877	3.7734	1.8238	2.7909	1.6516	3.7689	1.8197	2.8475	1.6518	2.5963	1.5872	2.5476	1.5773
5-year average	2.8545	1.7060	2.8705	1.7118	3.6417	1.7519	2.6438	1.5766	3.6382	1.7479	2.6905	1.5771	2.4653	1.5122	2.4181	1.5025

TABLE 4

Mean, Bias and RMSE of the coefficients $\hat{\gamma}, \hat{\rho}_1, \hat{\beta}$ and for W(5,5), $\sigma_\mu^2 = 0.8$, $\sigma_v^2 = 0.2$, (N,T) = (100,7), 1,000 replications

		Non-spatial estimators						*Spatial estimators*									
		OLS		*Within*		*GMM (1) (without spatial lag variable)*		*GMM (2) (with spatial lag variable)*		*GMM-SAR-RE (1) (without spatial lag variable)*		*GMM-SAR-RE (2) (with spatial lag variable)*		*GMM-SL-RE*		*GMM-SL-SAR-RE*	
	ρ_1	0.7	0.2	0.7	0.2	0.7	0.2	0.7	0.2	0.7	0.2	0.7	0.2	0.7	0.2	0.7	0.2
$\gamma = 0.2$, $\beta = 1$ and $\rho_2 = 0.4$ (SAR-RE)																	
$\hat{\gamma}$	Mean	0.2658	0.3212	0.1876	0.1923	0.3390	0.1985	0.1904	0.1976	0.2329	0.1967	0.1932	0.1987	0.1918	0.1974	0.1942	0.1986
	Bias	0.0658	0.1212	−0.0124	−0.0077	0.1390	−0.0014	−0.0095	−0.0024	0.0329	−0.0032	−0.0067	−0.0012	−0.0082	−0.0026	−0.0058	−0.0013
	RMSE	0.0673	0.1230	0.0148	0.0118	0.1445	0.0103	0.0124	0.0088	0.0371	0.0088	0.0102	0.0083	0.0114	0.0089	0.0095	0.0082
$\hat{\rho}_1$	Mean	0.6987	0.3033	0.7230	0.2366	—	—	0.7205	0.2364	—	—	0.7177	0.2312	0.7181	0.2338	0.7155	0.2286
	Bias	−0.0010	0.1033	0.0230	0.0366	—	—	0.0205	0.0364	—	—	0.0177	0.0312	0.0181	0.0338	0.0155	0.0286
	RMSE	0.0242	0.1197	0.0274	0.0505	—	—	0.0257	0.0498	—	—	0.0232	0.0464	0.0237	0.0464	0.0221	0.0441
$\hat{\beta}$	Mean	0.9561	0.9180	1.0005	1.0021	1.0519	1.0002	1.0002	1.0005	0.9971	0.9966	1.0006	1.0013	0.9999	1.0004	1.0010	1.0015
	Bias	−0.0438	−0.0820	0.0004	0.0021	0.0520	0.0003	0.0001	0.0005	−0.0028	−0.0034	0.0006	0.0013	0.00002	0.0004	0.0009	0.0015
	RMSE	0.0481	0.0840	0.0093	0.0098	0.0565	0.0109	0.0099	0.0099	0.0164	0.0106	0.0090	0.0094	0.0093	0.0097	0.0095	0.0096
$\hat{\rho}_2$	Mean	—	—	—	—	—	—	—	—	0.8940	0.5013	0.2551	0.2586	—	—	0.2830	0.2879
	Bias	—	—	—	—	—	—	—	—	0.4940	0.1013	−0.1448	−0.1413	—	—	−0.1170	−0.1120
	RMSE	—	—	—	—	—	—	—	—	0.5080	0.1924	0.2298	0.2170	—	—	0.2564	0.2504
		0.4	0.2	0.4	0.2	0.4	0.2	0.4	0.2	0.4	0.2	0.4	0.2	0.4	0.2	0.4	0.2
$\gamma = 0.5$, $\beta = 1$ and $\rho_2 = 0.4$ (SAR-RE)																	
$\hat{\gamma}$	Mean	0.4931	0.5965	0.3894	0.4914	0.4251	0.5010	0.3941	0.4966	0.4025	0.4968	0.3960	0.4981	0.3946	0.4966	0.3967	0.4981
	Bias	0.0931	0.0965	−0.0105	−0.0086	0.0251	0.0010	−0.0059	−0.0033	0.0025	−0.0031	−0.0035	−0.0018	−0.0054	−0.0033	−0.0033	−0.0018
	RMSE	0.0942	0.0978	0.0129	0.0114	0.0292	0.0094	0.0097	0.0083	0.0098	0.0080	0.0083	0.0074	0.0096	0.0082	0.0083	0.0074
$\hat{\rho}_1$	Mean	0.4236	0.2573	0.4243	0.2236	—	—	0.4214	0.2220	—	—	0.4183	0.2185	0.4198	0.2213	0.4167	0.2175
	Bias	0.0230	0.0573	0.0243	0.0236	—	—	0.0214	0.0220	—	—	0.0183	0.0184	0.0198	0.0213	0.0167	0.0175
	RMSE	0.0395	0.0703	0.0325	0.0350	—	—	0.0308	0.0352	—	—	0.0289	0.0330	0.0296	0.0338	0.0276	0.0324
$\hat{\beta}$	Mean	0.9304	0.9245	1.0018	1.0017	1.0003	0.9962	1.0009	1.0009	0.9854	0.9919	1.0010	1.0011	1.0008	1.0008	1.0012	1.0012
	Bias	−0.0695	−0.0754	0.0018	0.0017	0.0003	−0.0037	0.0009	0.0008	−0.0145	−0.0080	0.0009	0.0011	0.0008	0.0008	0.0012	0.0012
	RMSE	0.0718	0.0774	0.0096	0.0096	0.0125	0.0115	0.0097	0.0097	0.0178	0.0125	0.0094	0.0095	0.0095	0.0092	0.0096	0.0095
$\hat{\rho}_2$	Mean	—	—	—	—	—	—	—	—	0.7992	0.6128	0.2540	0.2553	—	—	0.2829	0.2853
	Bias	—	—	—	—	—	—	—	—	0.3992	0.2128	−0.1459	−0.1447	—	—	−0.1170	−0.1146
	RMSE	—	—	—	—	—	—	—	—	0.4227	0.2663	0.2382	0.2355	—	—	0.2599	0.2590

TABLE 5

Mean, Bias and RMSE of the coefficients $\hat{\gamma}, \hat{\rho}_1, \hat{\beta}$ and $\hat{\rho}_2$ for W(5,5), $\sigma_\mu^2 = 0.2$, $\sigma_v^2 = 0.8$, $(N, T) = (100, 7)$, 1,000 replications

		Non-spatial estimators						*Spatial estimators*									
		OLS		*Within*		*GMM (1) (without spatial lag variable)*		*GMM (2) (with spatial lag variable)*		*GMM-SAR-RE (1) (without spatial lag variable)*		*GMM-SAR-RE (2) (with spatial lag variable)*		*GMM-SL-RE*		*GMM-SL-SAR-RE*	
	ρ_1	0.7	0.2	0.7	0.2	0.7	0.2	0.7	0.2	0.7	0.2	0.7	0.2	0.7	0.2	0.7	0.2
$\gamma = 0.2$, $\beta = 1$ and $\rho_2 = 0.4$ (SAR-RE)																	
ρ_1	0.7	0.2	0.7	0.2	0.7	0.2	0.7	0.2	0.7	0.2	0.7	0.2	0.7	0.2	0.7	0.2	
$\hat{\gamma}$	Mean	0.1945	0.2301	0.1585	0.1722	0.3322	0.1972	0.1678	0.1909	0.2357	0.1949	0.1736	0.1933	0.1715	0.1902	0.1797	0.1946
	Bias	−0.0055	0.0301	−0.0415	−0.0277	0.1322	−0.0027	−0.0322	−0.0090	0.0357	−0.0050	−0.0264	−0.0067	−0.0285	−0.0098	−0.0203	−0.0054
	RMSE	0.0156	0.0332	0.0442	0.0328	0.1382	0.0202	0.0362	0.0191	0.0437	0.0172	0.0311	0.0179	0.0320	0.0194	0.0253	0.0166
$\hat{\rho}_1$	Mean	0.7541	0.3254	0.7774	0.3247	—	—	0.7700	0.3241	—	—	0.7646	0.3133	0.7634	0.3215	0.7550	0.3043
	Bias	0.0541	0.1254	0.0775	0.1247	—	—	0.0701	0.1241	—	—	0.0646	0.1133	0.0634	0.1215	0.0550	0.1043
	RMSE	0.0585	0.1358	0.0812	0.1384	—	—	0.0757	0.1382	—	—	0.0703	0.1292	0.0688	0.1353	0.0621	0.1225
$\hat{\beta}$	Mean	0.9995	0.9817	1.0018	1.0078	1.0485	1.0014	1.0003	1.0020	0.9996	0.9972	1.0015	1.0036	1.0001	1.0020	1.0042	1.0062
	Bias	−0.0004	−0.0183	0.0018	0.0078	0.0485	0.0014	0.0003	0.0019	−0.0004	−0.0027	0.0014	0.0036	0.00006	0.0019	0.0042	0.0062
	RMSE	0.0167	0.0248	0.0181	0.0201	0.0622	0.0219	0.0192	0.0194	0.0250	0.0200	0.0182	0.0191	0.0182	0.0187	0.0183	0.0194
$\hat{\rho}_2$	Mean	—	—	—	—	—	—	—	—	0.8640	0.5055	0.1621	0.1606	—	—	0.2924	0.2910
	Bias	—	—	—	—	—	—	—	—	0.4640	0.1055	−0.2379	−0.2393	—	—	−0.1075	−0.1089
	RMSE	—	—	—	—	—	—	—	—	0.4794	0.1909	0.3029	0.3033	—	—	0.2317	0.2294
	ρ_1	0.4	0.2	0.4	0.2	0.4	0.2	0.4	0.2	0.4	0.2	0.4	0.2	0.4	0.2	0.4	0.2
$\gamma = 0.5$, $\beta = 1$ and $\rho_2 = 0.4$ (SAR-RE)																	
$\hat{\gamma}$	Mean	0.5132	0.5238	0.4597	0.4681	0.5635	0.4984	0.4762	0.4872	0.5060	0.4947	0.4825	0.4904	0.4777	0.4872	0.4857	0.4927
	Bias	0.0132	0.0238	−0.0402	−0.0318	0.0635	−0.0016	−0.0237	−0.0128	0.0060	−0.0053	−0.0175	−0.0095	−0.0222	−0.0128	−0.0143	−0.0073
	RMSE	0.01798	0.0268	0.0427	0.0347	0.0689	0.0179	0.0284	0.0198	0.01971	0.0158	0.02338	0.0179	0.0272	0.0195	0.0208	0.0162
$\hat{\rho}$	Mean	0.4340	0.2714	0.4664	0.2842	—	—	0.4523	0.2790	—	—	0.4447	0.2707	0.4510	0.2790	0.4406	0.2653
	Bias	0.0340	0.0714	0.0664	0.0842	—	—	0.0523	0.0790	—	—	0.0447	0.0707	0.0510	0.0790	0.0406	0.0653
	RMSE	0.0426	0.0820	0.0731	0.0988	—	—	0.0616	0.0969	—	—	0.0564	0.0892	0.0599	0.0935	0.0536	0.0840
$\hat{\beta}$	Mean	0.9907	0.9835	1.0063	1.0064	0.9840	0.9969	1.0042	1.0033	0.9804	0.9926	1.0037	1.0038	1.0040	1.0032	1.0050	1.0052
	Bias	−0.0092	−0.0164	0.0063	0.0064	−0.0159	−0.0031	0.0042	0.0033	−0.0196	−0.0074	0.0037	0.0038	0.0040	0.0032	0.0050	0.0052
	RMSE	0.0204	0.0240	0.0195	0.0196	0.0288	0.0219	0.0195	0.0195	0.0281	0.0202	0.0188	0.0194	0.0188	0.0187	0.0189	0.0189
$\hat{\rho}_2$	Mean	—	—	—	—	—	—	—	—	0.8153	0.6140	0.1870	0.1629	—	—	0.2922	0.2893
	Bias	—	—	—	—	—	—	—	—	0.4153	0.2140	−0.2129	−0.2370	—	—	−0.1077	−0.1106
	RMSE	—	—	—	—	—	—	—	—	0.4363	0.2615	0.2824	0.3029	—	—	0.2441	0.2341

TABLE 6

Forecasting RMSE for W(5,5), $(N,T)=(100,7)$ 1,000 replications

	Non-spatial estimators						*Spatial estimators*									
	OLS		*Within*		*GMM (1) (without spatial lag variable)*		*GMM (2) (with spatial lag variable)*		*GMM-SAR-RE (1) (without spatial lag variable)*		*GMM-SAR-RE (2) (with spatial lag variable)*		*GMM-SL-RE*		*GMM-SL-SAR-RE*	
ρ_1	0.7	0.2	0.7	0.2	0.7	0.2	0.7	0.2	0.7	0.2	0.7	0.2	0.7	0.2	0.7	0.2
$\sigma_\mu^2=0.8$, $\sigma_v^2=0.2$																
$\gamma=0.2$, $\beta=1$ and $\rho_2=0.4$ (SAR-RE)																
1st year	2.4004	1.0835	2.1626	1.0359	2.0637	0.5888	1.8059	0.6509	2.0754	0.5892	1.5627	0.5711	1.8027	0.6490	1.5471	0.5699
2sd year	2.4755	1.1353	2.2232	1.0782	2.1080	0.6364	1.8710	0.6981	2.1155	0.6366	1.6345	0.6183	1.8678	0.6961	1.6193	0.6172
3rd year	2.5264	1.1606	2.2512	1.0949	2.1635	0.6583	1.8993	0.7169	2.1572	0.6587	1.6701	0.6376	1.8958	0.7149	1.6545	0.6364
4th year	2.5581	1.1743	2.2684	1.1034	2.1804	0.6688	1.9165	0.7265	2.1641	0.6691	1.6920	0.6475	1.9129	0.7246	1.6762	0.6463
5th year	2.5841	1.1827	2.2822	1.1093	2.2885	0.6784	1.9313	0.7335	2.2735	0.6787	1.7128	0.6547	1.9277	0.7316	1.6966	0.6534
5-year average	2.5089	1.1473	2.2375	1.0844	2.1608	0.6461	1.8848	0.7052	2.1571	0.6465	1.6544	0.6258	1.8814	0.7032	1.6387	0.6246
	0.4	0.2	0.4	0.2	0.4	0.2	0.4	0.2	0.4	0.2	0.4	0.2	0.4	0.2	0.4	0.2
$\gamma=0.5$, $\beta=1$ and $\rho_2=0.4$ (SAR-RE)																
1st year	1.7686	1.7087	1.7116	1.6644	1.1071	0.7859	1.1551	0.8876	1.1053	0.7863	0.9265	0.7558	1.1534	0.8865	0.9218	0.7544
2sd year	1.8236	1.7571	1.7518	1.6978	1.1615	0.8306	1.2005	0.9288	1.1612	0.8312	0.9765	0.7991	1.1988	0.9277	0.9719	0.7978
3rd year	1.8611	1.7892	1.7712	1.7144	1.1907	0.8535	1.2216	0.9487	1.1904	0.8543	1.0013	0.8201	1.2198	0.9475	0.9966	0.8188
4th year	1.8859	1.8103	1.7834	1.7248	1.2057	0.8667	1.2344	0.9606	1.2051	0.8676	1.0165	0.8329	1.2326	0.9594	1.0118	0.8316
5th year	1.9060	1.8270	1.7927	1.7326	1.2366	0.8804	1.2446	0.9701	1.2356	0.8813	1.0291	0.8431	1.2428	0.9689	1.0242	0.8418
5-year average	1.8490	1.7784	1.7621	1.7068	1.1803	0.8434	1.2113	0.9392	1.1795	0.8441	0.9899	0.8102	1.2095	0.9380	0.9853	0.8088
ρ_1	0.7	0.2	0.7	0.2	0.7	0.2	0.7	0.2	0.7	0.2	0.7	0.2	0.7	0.2	0.7	0.2
$\sigma_\mu^2=0.2$, $\sigma_v^2=0.8$																
$\gamma=0.2$, $\beta=1$ and $\rho_2=0.4$ (SAR-RE)																
1st year	1.7880	0.9395	1.7503	0.9410	2.1673	0.9093	1.7332	0.9129	2.1498	0.9076	1.7385	0.9056	1.7204	0.9118	1.7160	0.9041
2sd year	1.9360	1.0326	1.8906	1.0308	2.2804	1.0007	1.8748	1.0029	2.2613	0.9988	1.8836	0.9958	1.8628	1. 0017	1.8624	0.9942
3rd year	2.0015	1.0727	1.9750	1.0696	2.3511	1.0390	1.9492	1.0418	2.3187	1.0372	1.9575	1.0347	1.9338	1.0406	1.9314	1.0327
4th year	2.0408	1.0923	2.0187	1.0881	2.3784	1.0579	1.9897	1.0603	2.3383	1.0560	1.9966	1.0532	1.9728	1.0591	1.9687	1.0512
5th year	2.0853	1.1059	2.0681	1.1026	2.4885	1.0725	2.0377	1.0748	2.4511	1.0706	2.0443	1.0674	2.0178	1.0735	2.0129	1.0653
5-year average	1.9703	1.0486	1.9405	1.0465	2.3331	1.0158	1.9169	1.0186	2.3038	1.0140	1.9243	1.0113	1.9015	1.0173	1.8983	1.0095

TABLE 6

Continued

	Non-spatial estimators						*Spatial estimators*									
	OLS		*Within*		*GMM (1) (without spatial lag variable)*		*GMM (2) (with spatial lag variable)*		*GMM-SAR-RE (1) (without spatial lag variable)*		*GMM-SAR-RE (2) (with spatial lag variable)*		*GMM-SL-RE*		*GMM-SL-SAR-RE*	
ρ_1	0.4	0.2	0.4	0.2	0.4	0.2	0.4	0.2	0.4	0.2	0.4	0.2	0.4	0.2	0.4	0.2
$\gamma = 0.5$, $\beta = 1$ and $\rho_2 = 0.4$ (SAR-RE)																
1^{st} year	1.6114	1.2273	1.6396	1.2357	1.9238	1.1441	1.5266	1.1371	1.9528	1.1403	1.5224	1.1248	1.5301	1.1371	1.5132	1.1219
2^{sd} year	1.7283	1.3170	1.7452	1.3221	2.0567	1.2331	1.6363	1.2250	2.0997	1.2294	1.6360	1.2133	1.6395	1.2250	1.6262	1.2103
3^{rd} year	1.7975	1.3597	1.8018	1.3631	2.1174	1.2756	1.6959	1.2664	2.1656	1.2716	1.6992	1.2549	1.6991	1.2662	1.6887	1.2516
4^{th} year	1.8526	1.3850	1.8449	1.3870	2.1658	1.2992	1.7407	1.2901	2.2168	1.2950	1.7467	1.2786	1.7443	1.2899	1.7356	1.2753
5^{th} year	1.9076	1.4042	1.8874	1.4061	2.2453	1.3196	1.7866	1.3096	2.2984	1.3154	1.7952	1.2981	1.7907	1.3093	1.7835	1.2946
5-year average	1.7795	1.3386	1.7838	1.3428	2.1018	1.2543	1.6772	1.2457	2.1467	1.2504	1.6799	1.2339	1.6807	1.2455	1.6694	1.2307

the forecast RMSE results in the same way as those of Table 3 except that again the weight matrix is $W(5,5)$ rather than $W(1,1)$. To summarize, we find that the results are essentially the same, although the magnitudes are different.

VII. Empirical illustration

The model

Our empirical illustration is motivated by recent work seeking to apply, and test the viability of, contemporary economic geography theory, as presented in the seminal work of Fujita, Krugman and Venables (1999). This is commonly referred to as the new economic geography (NEG). We draw on one of the model specifications of Fingleton and Fischer (2010), although we do not make a formal link to the NEG theory in this paper. More formally we estimate the following dynamic spatial panel model:

$$\ln y_{it} = a + \gamma \ln y_{it-1} + \rho_1 \sum_{j=1}^{N} w_{ij} \ln y_{jt} + \beta_1 \ln P_{it} + \beta_2 \ln S_{it} + \beta_3 \ln n_{it} + \beta_4 \ln s_{it} + \varepsilon_{it} \quad i=1,\ldots,N; \quad t=1,\ldots,T \tag{49}$$

where y_{it} is gross value added (GVA) per worker for the ith region ($i=1,\ldots,N$) and tth time period ($t=1,\ldots,T$). P_{it} is market potential,[7] S_{it} is schooling, n_{it} is the (adjusted) population growth rate and s_{it} is the investment rate. Moreover, following Koch (2008), who extended the neoclassical growth model to pick up spatial spillover effects, we also include the spatial lag of the dependent variable, given by the matrix product of the W_N matrix and y_{it}. W_N is a first-order contiguity matrix, which considers that two geographical regions i and j are neighbours if they directly share a border. More precisely, the weights matrix is binary, with $w_{ij}=1$ when i and j are neighbours and $w_{ij}=0$ when they are not. By convention, diagonal elements are null: $w_{ii}=0$ and the weights are normalized such that the elements of each row sum to 1 (i.e. row-normalized). As in Fingleton and Fischer (2010), our panel specification also includes a spatial SAR error process with

$$\varepsilon_{it} = \rho_2 \sum_{j=1}^{N} w_{ij}\varepsilon_{jt} + u_{it} \tag{51}$$

and

$$u_{it} = \mu_i + v_{it} \tag{52}$$

in which the remainder term u_{it} is composed of a region-specific individual effect μ_i and an idiosyncratic random shock v_{it}.

[7]The measurement of market potential is difficult and complex. For simplicity, we use an earlier definition of market potential initially attributed to Harris (1954), which is regularly used in the applied literature:

$$P_{it} = \sum_{i \neq j}^{N} G_{jt} d_{ji}^{-\tau} \tag{50}$$

in which G_{jt} is the ‘size’ of the economy in region j, and d_{ij} is the ‘distance’ between region i and region j.

Data description

The data, which originate from Eurostat's Region database, Statistics Norway and the Swiss Office Féderal de la Statistique, comprise 255 NUTS-2 regions, observed over the period 1995–2003, and covering 25 European countries, hence $N = 255$ and $T = 9$. The short time dimension is a consequence of the lack of reliable data for Central and Eastern Europe regions, and also because of the fundamental reorganization of some formerly centrally planned command economies. The data cover Austria (nine regions), Belgium (11 regions), Czech Republic (eight regions), Denmark (one region), Estonia (one region), Finland (five regions), France (22 regions), Germany (40 regions), Greece (13 regions), Hungary (seven regions), Ireland (two regions), Italy (20 regions), Latvia (one region), Lithuania (one region), Luxembourg (one region), the Netherlands (12 regions), Norway (seven regions), Poland (16 regions), Portugal (five regions), Slovakia (four regions), Slovenia (one region), Spain (16 regions), Sweden (eight regions), Switzerland (seven regions) and UK (37 regions).[8]

The dependent variable y_{it} is measured by GVA per worker. GVA is the net result of output at basic prices minus intermediate consumption at purchasers' prices, following the approach used in the European System of Accounts (ESA) 1995. The adjusted population growth rate n_{it} is based on the rate of growth of the working-age population in region i at time t, with working age defined as 15–64 years for details, see Fingleton and Fischer, 2010. The investment rate s_{it} is the share of gross investment in gross regional product. We proxy labour efficiency and workforce skill by the level of educational attainment variable S_{it}, which is defined as the share of population (15 years and older) with higher education as given by the ISCED 1997 classes 5 and 6.

Market potential is a product of the total regional GVA, G_{jt} and $d_{ij}^{-\tau}$ in which d_{ij} is the great-circle distance from (the economic centre of) region i to (the economic centre of) region j. For region i, we sum across all other regions, excluding region i, so as to maximize exogeneity of P_{it} with respect to y_{it}. We assume that the rate of discount with distance ($\tau = 0.5$). We experimented with alternative values of τ and they gave similar outcomes to those described below.[9]

Results

We estimate our model (49) using the period 1995–2001, leaving out the last two years for the purpose of out-of-sample forecasting. Given the dynamic specification, one's first instinct could be to apply the well-known Arellano and Bond (1991) estimator. However, as we have already shown in our Monte Carlo study, the GMM (2) estimator does not produce satisfactory estimates. This is confirmed again in the empirical example in Note that out of the four estimators summarized in Table 7, two ignore the spatial error process (i.e. GMM (2) and GMM-SL-RE), and two take the error process into account (i.e. GMM-SAR-RE (2) and GMM-SL-SAR-RE). The estimated ρ_2 is equal to 0.877 and 0.714 respectively for GMM-SAR-RE (2) and GMM-SL-SAR-RE. Also GMM (2) relies only on

[8] We are grateful to Professor Manfred Fischer for his help in producing these data.

[9] Note that our model differs from that of Fingleton and Fischer (2010) in two respects. They estimate a static model whereas our model is dynamic. They use a different NEG market potential than our Harris measure.

TABLE 7

Estimates of the NEG-related model and forecasting RMSE

Parameters	GMM (2)	GMM-SAR-RE (2)	GMM-SL-RE	GMM-SL-SAR-RE
γ	0.1120	0.5853	0.1608	0.5705
	(0.0098)	(0.0408)	(0.0063)	(0.0152)
	(11.3901)	(14.3222)	(25.3179)	(37.3081)
ρ_1	0.8783	0.4205	0.8779	0.2488
	(0.0165)	(0.0584)	(0.0106)	(0.0217)
	(53.2201)	(7.1947)	(79.0110)	(11.4647)
β_1	0.0089	0.0099	0.0063	0.1632
	(0.0174)	(0.0801)	(0.0118)	(0.0252)
	(0.5140)	(0.1235)	(0.5386)	(6.4549)
β_2	0.0057	0.0390	0.0079	0.0052
	(0.0060)	(0.0140)	(0.0041)	(0.0048)
	(0.9679)	(2.7828)	(1.9383)	(1.0838)
β_3	0.0001	−0.0161	−0.0001	0.0008
	(0.0026)	(0.0123)	(0.0015)	(0.0024)
	(0.0566)	(−1.3091)	(−0.1209)	(0.3228)
β_4	0.0193	0.0215	0.0156	0.0386
	(0.0047)	(0.0124)	(0.0027)	(0.0035)
	(4.0760)	(1.7399)	(5.6328)	(10.7583)
Forecasting RMSE				
2002	26.3776	9.3194	19.8714	2.7546
2003	27.7578	10.0621	21.3049	2.8229
2-year average	27.0672	9.9607	20.5882	2.7888

Estimated standard errors and t-values in parentheses.

the orthogonality conditions (8) and (9), whereas GMM-SL-RE and GMM-SL-SAR-RE employ the additional spatial orthogonality conditions (10) and (11). Table 7 clearly shows considerable differences among the estimators. In particular, the temporal lag coefficient estimates of γ are smaller for the estimators that exclude the SAR disturbance process. In contrast, the spatial lag (ρ_1) estimates are larger for estimators that exclude the SAR disturbance process. The larger estimated ρ_1 produced under GMM (2) and GMM-SL-RE suggests that the spatial lag is capturing spatial effects that would otherwise be partially captured by an error process model. Note that while the point estimates differ, both temporal and spatial lags are statistically significant for all four estimators. Note also that the Harris-market potential variable is only significant for GMM-SL-SAR-RE. Moreover, the stationarity conditions are only satisfied for this estimator. This is in line with what the NEG theory suggests, and ties in with the interpretation provided by the static panel estimates of Fingleton and Fischer (2010). The educational attainment variable S_{it} is significant for GMM-SAR-RE (2) and GMM-SL-RE. As the schooling variable only changes slowly through time, it appears that the anticipated positive impact of schooling may not be particularly well identified. Of the other variables, population growth n_{it} and the investment rate s_{it}, these represent the rival non-nested neoclassical growth model, so we do not anticipate significant parameter estimates, but expect the neoclassical model to be encompassed by the dominant rival, as found in the formal NEG context by Fingleton and Fischer (2010). In the current dynamic setting, it turns out that while adjusted population growth is insignificant, so that we cannot reject the null hypothesis that $\beta_3 = 0$, on the whole and contrary to expectation the investment rate is a significant variable, with $\beta_4 > 0$. On reflection this is perhaps not surprising given the absence of capital investment in basic NEG theory.

The superiority of our GMM-SL-SAR-RE estimator, which has already been highlighted in our Monte Carlo analysis, is also apparent for our empirical example. This is illustrated using the RMSE for the out-of-sample forecast period. This RMSE is calculated for each of the 255 NUTS-2 regions for each of the two years. The last three lines of Table 7 give the across-region RMSE means for each year, and the 2-year average. From this it is clear that, when one compares model predictions and the actual GVA per worker outcomes, the GMM-SL-SAR-RE produces by far the best RMSE forecasts.

VIII. Conclusion

Our Monte Carlo study finds that when the true model is a dynamic first-order spatial autoregressive specification with SAR-RE disturbances, estimators that ignore the endogeneity of the spatial lag $W_N y$ and the endogeneity of the temporally lagged dependent variable perform badly in terms of bias and RMSE. For our experiments, accounting for heterogeneity and endogeneity improve the forecast performance by a big margin; accounting only for spatial correlation in the disturbances also improves the forecast, but by a smaller margin. Recognizing the presence of a spatially lagged dependent variable among the regressors has an important effect. Ignoring both sources of spatial dependence significantly worsens the forecasting performance and leads to a huge bias in the estimated coefficients. So, a misspecified estimator, especially in terms of spatial effects, has severe consequences in terms of estimation and forecasting for the applied economist. We come to the conclusion, supported by our empirical example, that our dynamic spatial GMM estimator performs well and is recommended in practice.

Final Manuscript Received: xxxx

References

Anderson, T. W. and Hsiao, C. (1981). 'Estimation of dynamic models with error components', *Journal of the American Statistical Association*, Vol. 76, pp. 598–606.

Anderson, T. W. and Hsiao, C. (1982). 'Formulation and estimation of dynamic models using panel data', *Journal of Econometrics*, Vol. 18, pp. 47–82.

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

Anselin, L., Le Gallo, J. and Jayet, H. (2008). 'Spatial panel econometrics' Chapter 19, in Mátyás L. and Sevestre P. (eds), *The Econometrics of Panel Data: Fundamentals and Recent Developments in Theory and Practice*, Springer-Verlag, Berlin, pp. 625–660.

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

Arellano, M. and Honoré, B. (2001). 'Panel data models: some recent developments', in Heckman J. J. and Leamer E. E. (eds), *Handbook of Econometrics*, Vol. 5, North-Holland, Amsterdam, pp. 3229–3296.

Baltagi, B. H. (2008). *Econometric Analysis of Panel Data*, Wiley and Sons, Chichester.

Baltagi, B. H. and Pirotte, A. (2010). 'Panel data inference under spatial dependence', *Economic Modelling*, Vol. 27, pp. 1368–1381.

Baltagi, B. H., Bresson, G. and Pirotte, A. (2012). 'Forecasting with spatial panel data', *Computational Statistics and Data Analysis*, Vol. 56, pp. 3381–3397.
Baltagi, B. H. and Li, D. (2006). 'Prediction in the panel data model with spatial correlation: the case of liquor', *Spatial Economic Analysis*, Vol. 1, pp. 175–185.
Blundell, R. and Bond, S.(1998). 'Initial conditions and moment restrictions in dynamic panel data models', *Journal of Econometrics*, Vol. 87, pp. 115–143.
Blundell, R., Bond, S. and Windmeijer, F. (2000). 'Estimation in dynamic panel data models: improving on the performance of the standard GMM estimator', in Baltagi B. H. (ed.), *Advances in Econometrics: Nonstationary Panels, Panel Cointegration and Dynamic Panels*, Vol. 15, JAI Elsevier, Amsterdam, pp. 8–92.
Chamberlain G. (1984). 'Panel data', in Griliches Z. and Intriligator , M. (eds), *The Handbook of Econometrics*, Chapter 22, North-Holland, Amsterdam, pp. 1247–1318.
Elhorst, J. P. (2005). 'Unconditional maximum likelihood estimation of linear and log-linear dynamic models for spatial models', *Geographical Analysis*, Vol. 37, pp. 85–106.
Elhorst, J. P. (2010). 'Dynamic panels with endogenous interaction effects when T is small', *Regional Science and Urban Economics*, Vol. 40, pp. 272–282.
Fingleton, B. (2009). 'Prediction using panel data regression with spatial random effects', *International Regional Science Review*, Vol. 32, pp. 195–220.
Fingleton, B. (2010). *Predicting the Geography of House Prices*, SERC Discussion paper no. 45, The London School of Economics and Political Science.
Fingleton, B. and Fischer, M. (2010). 'Neoclassical theory versus new economic geography. Competing explanations of cross-regional variation in economic development', *Annals of Regional Science*, Vol. 44, pp. 467–491.
Fujita, M., Krugman, P. R. and Venables, A. (1999). *In The Spatial Economy: Cities, Regions, and International Trade*, MIT press, Cambridge, Massachusetts.
Goldberger, A. S. (1962). 'Best linear unbiased prediction in the generalized linear regression model', *Journal of the American Statistical Association*, Vol. 57, pp. 369–375.
Harris, C. (1954). 'The market as a factor in the localization of industry in the United States', *Annals of the Association of American Geographers*, Vol. 64, pp. 315–348.
Harris, M. N., Sevestre, P. and Mátyás, L. (2008). 'Dynamic models for short panels', Chapter 7, in Mátyás L. and Sevestre, P. (eds), *The Econometrics of Panel Data: Fundamentals and Recent Developments in Theory and Practice*, Springer-Verlag, Berlin, pp. 249–278.
Hsiao, C. (2003). *Analysis of Panel Data*, Cambridge University Press, Cambridge.
Jacobs, J. P. A. M., Ligthart, J. E. and Vrijburg, H. (2009). *Dynamic Panel Data Models Featuring Endogenous Interaction and Spatially Correlated Errors*, Working Paper no. 09-15, International Studies Program, Georgia State University.
Kapoor, M., Kelejian, H. H. and Prucha, I. R. (2007). 'Panel data models with spatially correlated error components', *Journal of Econometrics*, Vol. 140, pp. 97–130.
Kelejian, H. H. and Prucha, I. R. (1999). 'A generalized moments estimator for the autoregressive parameter in a spatial model', *International Economic Review*, 40, pp. 509–533.
Kholodilin, K. A., Siliverstovs B. and Kooths, S. (2008). 'A dynamic panel data approach to the forecasting of the DGP of German Länder', *Spatial Economic Analysis*, Vol. 3, pp. 195–207.
Kiviet, J. (1995). 'On Bias, Inconsistency and Efficiency of Various Estimators in Dynamic Panel Data Models', *Journal of Econometrics*, Vol. 68, pp. 53–78.
Koch, W. (2008). 'Development accounting with spatial effects', *Spatial Economic Analysis*, Vol. 3, pp. 321–341.
Kukenova, M. and Monteiro, J.-A. (2009). *Spatial Dynamic Panel Model and System GMM: A Monte Carlo Investigation*, Munich Personal RePEc Archive paper no.14319. http://mpra.ub.uni-muenchen.de/14319/.
Lee, L.-F. and Griffiths, W. E. (1979). *The Prior Likelihood and Best Linear Unbiased Prediction in Stochastic Coefficient Linear Models*, working paper, Department of Economics, University of Minnesota.
Lee, L.-F. and Yu, J. (2010a). 'Estimation of spatial autoregressive panel data models with fixed effects', *Journal of Econometrics*, Vol. 154, pp. 165–185.

Lee, L.-F. and Yu, J. (2010b). 'Some recent developments in spatial panel data models', *Regional Science and Urban Economics*, Vol. 5, pp. 255–271.

LeSage, J. P. and Pace, R. K. (2009). *Introduction to Spatial Econometrics*, Chapman & Hall/CRC Press, Boca Raton.

Longhi, S. and Nijkamp, P. (2007). 'Forecasting regional labor market developments under spatial autocorrelation', *International Regional Science Review*, Vol. 30, pp. 100–119.

Mutl, J. (2006). *Dynamic Panel Data Models with Spatially Correlated Disturbances*, PhD, University of Maryland.

Nickell, S. (1981). 'Biases in dynamic model with fixed effects', *Econometrica*, Vol. 49, pp. 1417–1426.

Sevestre, P. and Trognon, A. (1983). 'Propriétés de grands échantillons d'une classe d'estimateurs des modèles autorégressifs à erreurs composées', *Annales de l'INSEE*, Vol. 50, pp. 25–48.

Sevestre, P. and Trognon, A. (1985). 'A note on autoregressive error component models', *Journal of Econometrics*, Vol. 28, pp. 231–245.

Sevestre, P. and Trognon, A. (1996). 'Dynamic linear models', in Màtyàs L. and Sevestre P. (eds)., *The Econometrics of Panel Data: A Handbook of Theory with Applications*, Chapter 7, Kluwer Academic Publishers, Dordrecht, pp. 121–144.

Taub, A. J. (1979). 'Prediction in the context of the variance-components model', *Journal of Econometrics*, Vol. 10, pp. 103–108.

Trapani, L. and Urga, G. (2009). 'Optimal forecasting with heterogeneous panels: a Monte Carlo study', *International Journal of Forecasting*, Vol. 25, pp. 567–586.

Trognon, A. (1978). 'Miscellaneous asymptotic properties of ordinary least squares and maximum likelihood estimators in dynamic error components models', *Annales de l'INSEE*, Vol. 30–31, pp. 631–657.

Wansbeek, T. J. and Kapteyn, A. (1978). 'The seperation of individual variation and systematic change in the analysis of panel data', *Annales de l'INSEE*, Vol. 30–31, pp. 659–680.

Yu, J., de Jong, R. and Lee, L. (2008). 'Quasi-maximum likelihood estimators for spatial dynamic panel data with fixed effects when both N and T large', *Journal of Econometrics*, Vol. 146, pp. 118–134.