Statistica Sinica **33** (2023), 2693-2714
doi:https://doi.org/10.5705/ss.202021.0119

IMPACT ANALYSIS FOR SPATIAL AUTOREGRESSIVE MODELS: WITH APPLICATION TO AIR POLLUTION IN CHINA

Hsuan-Yu Chang and Jihai Yu

Chung-Hua Institution for Economic Research and Peking University

Abstract: We investigate impact analysis and its asymptotic inference for spatial autoregressive models. We propose using the delta method, which enables us to obtain the dispersion in an explicit form. In addition, we provide an element-wise impact analysis. We first study the cross-sectional case, where various impacts are introduced to measure the interaction and feedback effects in a space dimension. We then study the spatial dynamic panel case, with simultaneous spatial and dynamic feedback in the effects. Monte Carlo results show that the proposed impact analysis has satisfactory finite-sample properties. Finally, we apply the impact analysis to investigate how meteorological factors and air pollutants affect $PM_{2.5}$ in Chinese cities.

Key words and phrases: Air pollution, dynamic panels, impact analysis, spatial autoregression.

1. Introduction

Considerable progress has been made in the past decade on the theoretical aspects of spatial econometrics. In cross-sectional spatial econometrics, the spatial autoregressive (SAR) model of Cliff and Ord (1973) has received the most attention. Various estimation methods and their asymptotic analysis have been developed, such as the 2SLS of Kelejian and Prucha (1998), the quasi-maximum likelihood (QML) of Lee (2004), and the generalized method of moments (GMM) of Lee (2007). For spatial panel data, fixed effects and random effects models can be found in Baltagi, Song and Koh (2003); Baltagi et al. (2007), Kapoor, Kelejian and Prucha (2007), and Lee and Yu (2010b, 2012), among others. When there is a dynamic feature in the model, Elhorst (2005) and Su and Yang (2015) investigate dynamic panel data with spatial disturbances, and Yu, de Jong and Lee (2008, 2012) and Yu and Lee (2010) study stable spatial cointegration and unit root models that include individual time lags, spatial time lags, and contemporaneous

Corresponding author: Hsuan-Yu Chang, The Center for Green Economy, Chung-Hua Institution for Economic Research, Taipei 10672. Email: hychang@cier.edu.tw.

spatial lags.

However, few studies in the applied literature examine impact analyses in spatial econometric models (LeSage and Pace (2009)), that is, how a change in a single region associated with any given explanatory variable affects the region itself (direct impact) and potentially indirectly affects all other regions (indirect impact). The two most important theoretical works are LeSage and Pace (2009) and Debarsy, Ertur and LeSage (2012). In a cross-sectional setting, LeSage and Pace (2009) provide a computationally efficient simulation approach for producing empirical estimates of the dispersion for the scalar summary measures of the impacts, and then make a pseudo-inference based on either an MLE or a Bayesian MCMC estimation. Debarsy, Ertur and LeSage (2012) extend the aforementioned approach to spatial dynamic panel data (SDPD) models. The contribution of this study is to establish the asymptotic properties of these impact estimates based on the delta method, which enables researchers to explicitly obtain the variance matrix and is useful for empirical applications. In addition, we provide an element-wise impact analysis to evaluate the unit-to-unit impact and to conduct a corresponding inference. The asymptotic distribution of the effects using the delta method has the advantage of an explicit variance formula for this dispersion of the scalar impact, but it can be computationally burdensome to compute the matrix inverse when the matrix dimension is very large. On the other hand, the MCMC estimation of the dispersion in LeSage and Pace (2009) and Debarsy, Ertur and LeSage (2012) does not require a matrix inversion computation, but might produce posterior distributions that are heavily influenced by the priors.

The aim of this study is to provide a unified asymptotic inference method for the estimated effects in various types of spatial econometric models, in both cross-sectional and panel data settings, which could be helpful for applied researchers and policy makers. Section 2 investigates the impact analysis for cross-sectional SAR models. Section 3 covers spatial panel data models, where the SDPD case is the main focus. Section 4 provides Monte Carlo results to evaluate the finite-sample performance of the impact analysis. Section 5 investigates how various pollutants and meteorological factors affect air pollution in Chinese cities. Section 6 concludes the paper. To conserve space, some algebra and additional simulation and empirical results are provided in the Supplementary Material.

2. Impact Analysis for SAR Models

Let us first consider the following cross-sectional SAR model:

$$Y_n = \alpha_0 l_n + \lambda_0 W_n Y_n + \sum_{k=1}^{H} \beta_{k0} X_{nk} + \sum_{k=1}^{K} \delta_{k0} W_n X_{nk} + \epsilon_n, \tag{2.1}$$

where Y_n is an $n \times 1$ vector of the dependent variable, l_n is an $n \times 1$ vector of ones, W_n is the so-called spatial weights matrix or interaction matrix, $W_n Y_n$ is the spatially lagged dependent variable, and λ_0 is the scalar spatial autoregressive coefficient. Here, X_{nk} is an $n \times 1$ vector of exogenous variables, and β_{k0} is the corresponding scalar coefficient. Compared with the standard SAR model in Cliff and Ord (1973) and LeSage and Pace (2009), we allow exogenous regressors to have additional cross-neighbor effects, the so-called spatial Durbin terms, where the corresponding scalar coefficients are δ_{k0}. In contrast to the classical linear model, it is straightforward to see that β_{k0} and δ_{k0} cannot be interpreted as impact coefficients, because the model is in an implicit form. To express the partial derivatives, we first compute the reduced form of the model. In equilibrium, assuming that $I_n - \lambda_0 W_n$ is invertible, we have

$$Y_n = \alpha_0 (I_n - \lambda_0 W_n)^{-1} l_n + (I_n - \lambda_0 W_n)^{-1} \left(\sum_{k=1}^{H} \beta_{k0} X_{nk} + \sum_{k=1}^{K} \delta_{k0} W_n X_{nk} + \epsilon_n \right).$$

2.1. Definition of impacts

Following LeSage and Pace (2009), we take the partial derivatives of Y_n relative to X_{nk}, assuming that W_n does not depend on X_{nk}, for all k:

$$\frac{\partial Y_n}{\partial X'_{nk}} = (I_n - \lambda_0 W_n)^{-1} C_{nk} = (I_n + \lambda_0 W_n + \lambda^2 W_n^2 + \cdots) C_{nk},$$

where

$$C_{nk} = \begin{cases} \beta_{k0} I_n + \delta_{k0} W_n, & \text{for } 1 \le k \le \min(K, H) \\ \beta_{k0} I_n, & \text{for } K+1 \le k \le H \text{ if } H > K \\ \delta_{k0} W_n, & \text{for } H+1 \le k \le K \text{ if } H < K \end{cases} \tag{2.2}$$

This expression differs from that of a simple cross-sectional linear regression model if there is no interaction (i.e., $\lambda_0 = 0$, $\delta_0 = 0$), which would be $\beta_{k0} I_n$.

The matrix $(I_n - \lambda_0 W_n)^{-1}$ is the so-called *global spatial multiplier* or *global interaction multiplier*. Note that if W_n is row-normalized, then

$$(I_n - \lambda_0 W_n)^{-1} l_n = (I_n + \lambda_0 W_n + \lambda_0^2 W_n^2 + \cdots) l_n = \frac{1}{1 - \lambda_0} l_n.$$

Define $R_{nk} = (I_n - \lambda_0 W_n)^{-1} C_{nk}$ as the *impact matrix* associated with the kth explanatory variable, with the following elements:

$$R_{nk} = \begin{pmatrix} r_{11,k} & r_{12,k} & \cdots & r_{1n,k} \\ r_{21,k} & r_{22,k} & \cdots & r_{2n,k} \\ \vdots & \vdots & \ddots & \vdots \\ r_{n1,k} & r_{n2,k} & \cdots & r_{nn,k} \end{pmatrix} \neq \beta_{k0} I_n.$$

In contrast to the classical linear regression model, the diagonal elements of this matrix are different from each other, the off-diagonal elements are non-null, and the matrix is not symmetric.

The partial derivatives of y_i relative to x_{ik} or x_{jk}, for $i, j = 1, \ldots, n$, $j \neq i$, are then

$$\frac{\partial y_i}{\partial x_{ik}} = r_{ii,k} \quad \text{and} \quad \frac{\partial y_i}{\partial x_{jk}} = r_{ij,k}.$$

In general, $r_{ii,k} \neq r_{jj,k}$ and $r_{ij,k} \neq r_{ji,k}$, for $i, j = 1, \ldots, n$, $j \neq i$. The diagonal elements of this matrix, $diag\{R_{nk}\}$, represent the *direct impacts*, including feedback effects, where individuals i and j affect each other, and there are also longer paths that can go from individual i to j to k and back to i. The feedback effects corresponding to $diag\{R_{nk}\} - \beta_{k0} I_n$ are inherently heterogenous in the presence of spatial autocorrelation, owing to differentiated interaction terms in the W_n matrix. Note that the feedback effects are zero if there are no spatial effects. The magnitudes of these direct effects depend on (1) the degree of interaction between individuals, which is governed by the W_n matrix, (2) the parameter λ_0, measuring the strength of the spatial correlation between individuals, and (3) the parameters β_{k0} and δ_{k0}.

Finally, the off-diagonal elements of the impact matrix $R_{nk} - diag\{R_{nk}\}$ represent *indirect impacts*. Note that because $r_{ij,k} \neq r_{ji,k}$, the impact of a unit change in the kth explanatory variable for individual j on the dependent variable for individual i will, in general, be different from that of a unit change in the kth explanatory variable for individual i on the dependent variable for individual j.

Moreover, considering column j of matrix R_{nk}, note that a variation Δx_{jk} in the kth explanatory variable for individual j affects each individual in the sample differently. The sum of the jth column yields the total impact of a change of x_{jk} for individual j on all n individuals. The total impacts, direct and indirect, from each of the individuals $j = 1, \ldots, n$ are then collected in the row vector $l_n' R_{nk}$. The total indirect impacts from each unit $j = 1, \ldots, n$ are usefully collected in the row vector $l_n'(R_{nk} - diag\{R_{nk}\})$.

Considering row i of matrix R_{nk}, note that an identical variation ΔX_k in the

kth explanatory variable across all individuals of the sample affects individual i differently. The sum across the ith row represents the total impact on y_i of an identical change in x_{jk} ($j = 1, \ldots, n$) across all n individuals in the sample. The total impacts, direct and indirect, on each individual $i = 1, \ldots, n$ are then collected in the column vector $R_{nk} l_n$.

For notational convenience, we denote $\theta_0 = (\lambda_0, \beta_0', \delta_0')'$, where $\beta_0 = (\beta_{10}, \ldots, \beta_{H0})'$ and $\delta_0 = (\delta_{10}, \ldots, \delta_{K0})'$, and $R_{nk}(\theta) = (I_n - \lambda W_n)^{-1} C_{nk}(\theta)$, where $C_{nk}(\theta)$ is from (2.2). When we are interested in an element-wise analysis, the subject of interest is

$$r_{ij,k}(\theta) = e_{ni}' R_{nk}(\theta) e_{nj}, \tag{2.3}$$

where $e_{ni} = (0, \ldots, 0, 1, 0, \ldots, 0)'$, with one in its ith position. The (i, j) element of $R_{nk}(\theta)$ measures how a one-unit change in the jth unit of X_{nk} influences the ith unit in Y_n.

Here, summary scalar measures for the direct, indirect, feedback, and total effects are useful, given the complexity and amount of information available in such $n \times n$ impact matrices. The average direct impact is defined as

$$f_{k,direct}(\theta) \equiv n^{-1} tr R_{nk}(\theta), \tag{2.4}$$

and the average total impact is defined as

$$f_{k,total}(\theta) \equiv n^{-1} l_n' R_{nk}(\theta) l_n. \tag{2.5}$$

Finally, the average indirect impact is, by definition, the difference between the average total impact and the average direct impact:

$$f_{k,indirect}(\theta) \equiv n^{-1} l_n' R_{nk}(\theta) l_n - n^{-1} tr R_{nk}(\theta). \tag{2.6}$$

We may also be interested in the average feedback effect:

$$f_{k,feedback}(\theta) \equiv n^{-1} tr R_{nk}(\theta) - \beta_k, \tag{2.7}$$

where $(1/n) tr(C_{nk}) = \beta_{k0}$, because $tr(W_n) = 0$. The feedback effect can be applied to test the significance of feedback loops, where observation i affects observation j via a longer path that might go from observation i to j to k and back to i.

2.2. Estimation and inference of impacts

Assume that we have already obtained the estimate $\hat{\theta}_n$ and its asymptotic distribution such that $\sqrt{n}(\hat{\theta}_n - \theta_0) \xrightarrow{d} N(0, \lim_{n\to\infty} \Sigma_{\theta_0,n})$, where $\lim_{n\to\infty} \Sigma_{\theta_{0,n}}$ is nonsingular. For the cross-sectional SAR model in (2.1), the estimate $\hat{\theta}_n$ can be obtained using a 2SLS, ML, or GMM estimation (see Kelejian and Prucha (1998), Lee (2004), and Lee (2007), respectively). We wish to use the distributions of $r_{ij,k}(\hat{\theta}_n)$, $f_{k,direct}(\hat{\theta}_n)$, $f_{k,total}(\hat{\theta}_n)$, $f_{k,indirect}(\hat{\theta}_n)$, and $f_{k,feedback}(\hat{\theta}_n)$ to make statistical inferences for these impact estimates. In contrast to LeSage and Pace (2009), we derive the asymptotic distribution of these impacts and provide an explicit variance formula for empirical researchers.

If we are interested in how a one-unit change in the jth unit of X_{nk} influences the ith unit in Y_n, we can investigate the estimation and statistical inference of (2.3). For $r_{ij,k}(\theta_0) = e'_{ni}R_{nk}(\theta_0)e_{nj}$, we might obtain a superconsistent estimate, because some elements of $R_{nk}(\theta_0)$ will be of a smaller order. Because $R_{nk}(\theta_0)$ is row and column sum bounded, some elements must have a smaller order magnitude; otherwise, the row or column sums of $R_{nk}(\theta_0)$ would not be bounded. However, we can still perform a statistical inference, regardless of the rate of convergence for $r_{ij,k}(\hat{\theta}_n)$. Similarly to the analysis above, from the Taylor expansion, we have $r_{ij,k}(\hat{\theta}_n) = r_{ij,k}(\theta_0) + \partial r_{ij,k}(\bar{\theta}_n)/\partial\theta'(\hat{\theta}_n - \theta_0)$, where $\bar{\theta}_n$ lies between $\hat{\theta}_n$ and θ_0, and $\partial r_{ij,k}(\theta_n)/\partial\theta'$ has the form shown in (S1.1) in the Supplementary Material. Here, $\partial r_{ij,k}(\bar{\theta}_n)/\partial\theta'$ may be of a smaller order magnitude. Assume that $\Upsilon(\partial r_{ij,k}(\bar{\theta}_n)/\partial\theta - \partial r_{ij,k}(\theta_0)/\partial\theta) = o_p(1)$ and $\Upsilon(\partial r_{ij,k}(\theta_0)/\partial\theta) = O(1)$, where Υ can be of a higher order magnitude so that $\partial r_{ij,k}(\bar{\theta}_n)/\partial\theta - \partial r_{ij,k}(\theta_0)/\partial\theta$ can be superconsistent. As $\Upsilon(\partial r_{ij,k}(\bar{\theta}_n)/\partial\theta - \partial r_{ij,k}(\theta_0)/\partial\theta) \xrightarrow{p} 0$,

$$\Upsilon\sqrt{n}(r_{ij,k}(\hat{\theta}_n) - r_{ij,k}(\theta_0)) \xrightarrow{d} N\left(0, \lim_{n\to\infty}\left(\Upsilon^2\frac{\partial r_{ij,k}(\theta_0)}{\partial\theta'}\Sigma_{\theta_0,n}\frac{\partial r_{ij,k}(\theta_0)}{\partial\theta}\right)\right). \quad (2.8)$$

To test whether $r_{ij,k}(\hat{\theta}_n)$ is significantly different from zero, we construct a z-test from (2.8). Here, although Υ is unknown, by using $z \equiv (r_{ij,k}(\hat{\theta}_n) - r_{ij,k}(\theta_0))/\sqrt{(1/n)(\partial r_{ij,k}(\theta_0)/\partial\theta')\Sigma_{\theta_0,n}(\partial r_{ij,k}(\theta_0)/\partial\theta)} = \Upsilon\sqrt{n}(r_{ij,k}(\hat{\theta}_n) - r_{ij,k}(\theta_0)) /\sqrt{\Upsilon^2(\partial r_{ij,k}(\theta_0)/\partial\theta')\Sigma_{\theta_0,n}(\partial r_{ij,k}(\theta_0)/\partial\theta)} \xrightarrow{d} N(0,1)$, the construction of z does not depend on Υ. The above analysis can be extended to test the equivalence of two element-wise impacts; see the Supplementary Material.

Using the Taylor expansion $f_{k,direct}(\hat{\theta}_n) = f_{k,direct}(\theta_0) + (\partial f_{k,direct}(\bar{\theta}_n)/\partial\theta')(\hat{\theta}_n - \theta_0)$, where $\bar{\theta}_n$ lies between $\hat{\theta}_n$ and θ_0, we have

$$\begin{aligned}&\sqrt{n}(f_{k,direct}(\hat{\theta}_n) - f_{k,direct}(\theta_0))\\ &\xrightarrow{d} N\left(0, \lim_{n\to\infty}\left(\frac{\partial f_{k,direct}(\theta_0)}{\partial\theta'}\Sigma_{\theta_0,n}\frac{\partial f_{k,direct}(\theta_0)}{\partial\theta}\right)\right),\end{aligned} \quad (2.9)$$

where $\partial f_{k,direct}(\theta_0)/\partial\theta$ can be estimated from (S1.2) in the Supplementary Material and $\Sigma_{\theta_0,n}$ is obtained from the corresponding 2SLS, ML, or GMM estimate. Similarly, using $\partial f_{k,total}(\theta)/\partial\theta$ in (S1.3) in the Supplementary Material and $\partial f_{k,indirect}(\theta)/\partial\theta = \partial f_{k,total}(\theta)/\partial\theta - \partial f_{k,direct}(\theta)/\partial\theta$, we have

$$\sqrt{n}(f_{k,total}(\hat{\theta}_n) - f_{k,total}(\theta_0)) \xrightarrow{d} N\left(0, \lim_{n\to\infty}\left(\frac{\partial f_{k,total}(\theta_0)}{\partial\theta'}\Sigma_{\theta_0,n}\frac{\partial f_{k,total}(\theta_0)}{\partial\theta}\right)\right), \text{ and} \tag{2.10}$$

$$\sqrt{n}(f_{k,indirect}(\hat{\theta}_n) - f_{k,indirect}(\theta_0)) \xrightarrow{d} N\left(0, \lim_{n\to\infty}\left(\frac{\partial f_{k,indirect}(\theta_0)}{\partial\theta'}\Sigma_{\theta_0,n}\frac{\partial f_{k,indirect}(\theta_0)}{\partial\theta}\right)\right). \tag{2.11}$$

For the average feedback effect in (2.7), we have $\partial f_{k,feedback}(\theta)/\partial\theta = (\partial f_{k,direct}(\theta)/\partial\theta) - \big(0, 1, \mathbf{0}_{(H-1)\times 1}, \mathbf{0}_{K\times 1}\big)'$, because the diagonal elements of W_n in typical empirical applications are specified to be zero . Thus,

$$\sqrt{n}(f_{k,feedback}(\hat{\theta}_n) - f_{k,feedback}(\theta_0)) \xrightarrow{d} N\left(0, \lim_{n\to\infty}\left(\frac{\partial f_{k,feedback}(\theta_0)}{\partial\theta'}\Sigma_{\theta_0,n}\frac{\partial f_{k,feedback}(\theta_0)}{\partial\theta}\right)\right). \tag{2.12}$$

Proposition 1. *Under some regularity conditions, the estimates for the impacts in* (2.3)–(2.7) *are consistent and asymptotically normally distributed, as in* (2.8)–(2.12)*). Specifically, the estimates for the element-wise effects, direct effects, total effects, indirect effects, and feedback effect are given in* (2.8)–(2.12)*, respectively..*

Note that the required regularity conditions are the assumptions used to derive the asymptotic properties of the relevant estimators. Therefore, we might have different conditions, depending on the estimators we use. For example, the GMM and ML estimators have stronger conditions than the 2SLS estimators in the boundedness of the higher moments of disturbances. See Kelejian and Prucha (1998), Lee (2004), and Lee (2007) for the conditions for deriving the 2SLSE, MLE, and GMME, respectively.

If W_n is row normalized, we have $(1/n)l_n'[R_{nk}(\theta_0)]l_n = (\beta_{k0} + \delta_{k0})/(1 - \lambda_0)$ for $1 \leq k \leq \min(K, H)$, $(1/n)l_n'[R_{nk}(\theta_0)]l_n = \beta_{k0}/(1 - \lambda_0)$ for $K + 1 \leq k \leq H$ if $H > K$, and $(1/n)l_n'[R_{nk}(\theta_0)]l_n = \delta_{k0}/(1 - \lambda_0)$ for $H + 1 \leq k \leq K$ if $H < K$. For the empirical procedure, we first obtain a consistent and asymptotically normally distributed estimate $\hat{\theta}_n$. Using its variance matrix $\Sigma_{\theta_0,n}$, we can estimate and make statistical inferences for the impact analyses in (2.8)–(2.12).

3. Impact Analysis for Spatial Panel Models

The analysis in Section 2 can be generalized to panel data models:

$$Y_{nt} = \lambda_0 W_n Y_{nt} + \sum_{k=1}^{H} \beta_{k0} X_{nk,t} + \sum_{k=1}^{K} \delta_{k0} W_n X_{nk,t} + \mathbf{c}_{n0} + \alpha_{t0} l_n + V_{nt}, \quad (3.1)$$

where $Y_{nt} = (y_{1t}, y_{2t}, \ldots, y_{nt})'$ and $V_{nt} = (v_{1t}, v_{2t}, \ldots, v_{nt})'$ are $n \times 1$ column vectors, and v_{it} are independent and identically (i.i.d.) across i and t with zero mean. Furthermore, the spatial weights matrix W_n is nonstochastic, $X_{nk,t}$ is an $n \times 1$ vector of nonstochastic regressors, $\mathbf{c}_{n0}$ is an $n \times 1$ column vector of individual effects, and α_{t0} is the time effect. Presuming that $S_n \equiv (I_n - \lambda_0 W_n)$ is invertible, (3.1) can be rewritten as

$$Y_{nt} = S_n^{-1} \left(\sum_{k=1}^{H} \beta_{k0} X_{nk,t} + \sum_{k=1}^{K} \delta_{k0} W_n X_{nk,t} \right) + S_n^{-1} \left(\mathbf{c}_{n0} + \alpha_{t0} l_n + V_{nt} \right).$$

Thus, a change in X_{nt} affects the dependent variable for the current period (Y_{nt}), but not for other periods. Therefore, the analysis in Section 2 for the cross-sectional model can be extended straightforwardly to static spatial panel data models. However, when we have dynamic features, changes in the exogenous variables affects both current and future periods. In the following, we focus on an impact analysis for an SDPD model.

The SDPD model is

$$\begin{aligned} Y_{nt} = {} & \lambda_0 W_n Y_{nt} + \gamma_0 Y_{n,t-1} + \rho_0 W_n Y_{n,t-1} \\ & + \sum_{k=1}^{H} \beta_{k0} X_{nk,t} + \sum_{k=1}^{K} \delta_{k0} W_n X_{nk,t} + \mathbf{c}_{n0} + \alpha_{t0} l_n + V_{nt}, \end{aligned} \quad (3.2)$$

where γ_0 is the dynamic effect coefficient and ρ_0 is the spatial-dynamic coefficient. For the DGP, we assume that the initial values in Y_{n0} are observable. When T is large, the role of the initial observation Y_{n0} is not important. For the ML estimation, we need $n/T^3 \to 0$ for the bias-corrected estimates to work. However, when T is small, the MLE is inconsistent and the GMM estimation requires that the initial observation has some boundedness feature. The reduced form of (3.2) is

$$Y_{nt} = A_n Y_{n,t-1} + S_n^{-1} \left(\sum_{k=1}^{H} \beta_{k0} X_{nk,t} + \sum_{k=1}^{K} \delta_{k0} W_n X_{nk,t} + \mathbf{c}_{n0} + \alpha_{t0} l_n + V_{nt} \right),$$

where $A_n = S_n^{-1}(\gamma_0 I_n + \rho_0 W_n)$. Assuming that the process is stable, so that infinite sums are well defined, by continuous substitution,

$$Y_{nt} = \sum_{h=0}^{\infty} A_n^h S_n^{-1} \left(\sum_{k=1}^{H} \beta_{k0} X_{nk,t-h} + \sum_{k=1}^{K} \delta_{k0} W_n X_{nk,t-h} + \mathbf{c}_{n0} + \alpha_{t-h,0} l_n + V_{n,t-h} \right). \tag{3.3}$$

Equation (3.3) expresses the model in terms of a space-time multiplier (Anselin, LeGallo and Jayet (2008)) specifying how the joint determination of the dependent variables is a function of both the spatial and the time lags of the explanatory variables and disturbances at all locations of the spatial units. This equation is useful for calculating the effects of changes of exogenous variables on outcomes over time and across spatial units.

3.1. Definition of impacts for SDPD model

For analytical purposes, we assume that the time effects have zero means, so that $E(\alpha_{t0}) = 0$. The regressors X_{nt} and weights matrix W_n are assumed to be given. From (3.3), we have

$$\mathrm{E}(Y_{nt}) = \sum_{h=0}^{\infty} A_n^h S_n^{-1} \left(\sum_{k=1}^{H} \beta_{k0} X_{nk,t-h} + \sum_{k=1}^{K} \delta_{k0} W_n X_{nk,t-h} + \mathbf{c}_{n0} \right).$$

As in LeSage and Pace (2009), we may be interested in the impact of changing the regressor by the same amount across all spatial units at a time. In a more general space-time setting, we consider changing the regressor by the same amount across all spatial units in some consecutive periods, for instance, from period t_1 to t_2, where $t_1 \leq t_2 \leq t$. Hence, we have $\partial \mathrm{E}(Y_{nt})/\partial X_k = \sum_{h=t-t_2}^{t-t_1} A_n^h S_n^{-1} C_{nk}$, where C_{nk} is defined in (2.2).

By denoting $\theta = (\lambda, \gamma, \rho, \beta', \delta')'$, the element-wise impact is

$$r_{ij,kt}(\theta_0) \equiv \frac{\partial [\mathrm{E}(Y_{nt})]_i}{\partial x_{kj,t_1}} = \sum_{h=t-t_2}^{t-t_1} \left[e'_{ni} A_n^h S_n^{-1} C_{nk} e_{nj} \right]. \tag{3.4}$$

In addition, the average direct impact $f_{kt,direct}(\theta_0) \equiv (1/n) tr \left[\partial \mathrm{E}(Y_{nt})/\partial X_k \right]$ is

$$f_{kt,direct}(\theta_0) = \sum_{h=t-t_2}^{t-t_1} \frac{1}{n} tr \left[A_n^h S_n^{-1} C_{nk} \right]. \tag{3.5}$$

Thus, marginal changes in a dynamic model also affect future periods. Here, $A_n^h S_n^{-1} C_{nk}$ provides the space-time multiplier of $X_{n,t-h}$ at time period $t-h$ on

h-periods-ahead outcome Y_{nt}, and (3.5) summarizes the average direct impact of changing x_k for all spatial units in the consecutive periods t_1 to t_2 on the expected outcomes Y_{nt} at time t. Similarly, we define the average total effect, average indirect effect, and average feedback effect as follows:

$$f_{kt,total}(\theta_0) = \sum_{h=t-t_2}^{t-t_1} \frac{1}{n}\left[l_n' A_n^h S_n^{-1} C_{nk} l_n\right], \tag{3.6}$$

$$f_{kt,indirect}(\theta_0) = \sum_{h=t-t_2}^{t-t_1} \left(\frac{1}{n}\left[l_n' A_n^h S_n^{-1} C_{nk} l_n\right] - \frac{1}{n} tr\left[A_n^h S_n^{-1} C_{nk}\right]\right), \tag{3.7}$$

and

$$f_{kt,feedback}(\theta_0) \equiv \sum_{h=t-t_2}^{t-t_1} \frac{1}{n} tr\left[A_n^h S_n^{-1} C_{nk}\right] - \sum_{h=t-t_2}^{t-t_1} \gamma^h \beta_{k0}, \tag{3.8}$$

respectively.

Instead of considering a change in the regressors in consecutive periods t_1 to t_2, where $t_1 \leq t_2 \leq t$, we might only be interested in how a current change at time t influences the future outcome at $(t+\tau)$, where we can have either a one-time change only at t, or a continuous change from t to $t+\tau$. Here, the average direct effect is $(1/n)tr\left[A_n^\tau S_n^{-1} C_{nk}\right]$ and $\sum_{h=0}^{\tau}(1/n)tr\left[A_n^h S_n^{-1} C_{nk}\right]$ for the marginal and accumulative cases, respectively. Debarsy, Ertur and LeSage (2012) study these cases to determine how a permanent change in $X_{nk,t}$ affects the future horizons (accumulatively). We investigate both the marginal and the accumulative impacts in our Monte Carlo analysis in Section 4.

For the special case of $t_1 = t_2 = t$ (so that we have a change in x only at period t), the average total impact on the expected outcome $\mathrm{E}(Y_{nt})$ is simply $(1/n)l_n' S_n^{-1} C_{nk} l_n$, which is the same as the cross-sectional SAR model. When the weights matrix W_n is row normalized, we have some interesting implications for the effects. Under a row normalization of W_n, we have $\sum_{h=t-t_2}^{t-t_1}(1/n)l_n' A_n^h S_n^{-1} C_{nk} l_n = (\beta_{k0}+\delta_{k0})/(1-\lambda_0)\sum_{h=t-t_2}^{t-t_1}((\gamma_0+\rho_0)/(1-\lambda_0))^h$. Note that if we have changes in x_k for all times from the infinite past to t, that is, $t_1 = -\infty$ and $t = t_2$, the total impact is $(\beta_{k0}+\delta_{k0})/(1-(\lambda_0+\gamma_0+\rho_0))$.

3.2. Estimation and inference of impacts for SDPD model

The impacts described in (3.4)–(3.8) are nonstochastic and depend on the true parameter value. If we replace those unknown parameters with estimates, we have an estimate of the expectation, and its variance can be obtained using the delta method. Assume that we have initial estimate $\hat{\theta}_{nT}$ available so that

$\sqrt{nT}(\hat{\theta}_{nT} - \theta_0) \xrightarrow{d} N(0, \Sigma_{\theta_0,nT})$.

For the element-wise analysis, assume that $\Upsilon(\partial r_{ij,kt}(\hat{\theta}_n)/\partial\theta - \partial r_{ij,kt}(\theta_0)/\partial\theta) = o_p(1)$ and $\Upsilon(\partial r_{ij,kt}(\theta_0)/\partial\theta) = O(1)$, where Υ can be of higher-order magnitude so that $\partial r_{ij,kt}(\hat{\theta}_{nT})/\partial\theta - \partial r_{ij,kt}(\theta_0)/\partial\theta$ can be superconsistent. We have

$$\begin{aligned} &\Upsilon\sqrt{nT}(r_{ij,kt}(\hat{\theta}_{nT}) - r_{ij,kt}(\theta_0)) \\ &\xrightarrow{d} N\left(0, \lim_{n,T\to\infty}\left(\Upsilon^2 \frac{\partial r_{ij,kt}(\theta_0)}{\partial\theta'}\Sigma_{\theta_0,nT}\frac{\partial r_{ij,kt}(\theta_0)}{\partial\theta}\right)\right), \end{aligned} \tag{3.9}$$

where $\partial r_{ij,kt}(\theta_0)/\partial\theta'$ can be found in the Supplementary Material. Similarly to the cross-sectional case, the statistical inference of $r_{ij,kt}$ does not depend on Υ. The above analysis can be extended to test the equivalence of element-wise impacts in different periods, which is provided in the Supplementary Material.

Similarly,

$$\begin{aligned} &\sqrt{nT}(f_{kt,direct}(\hat{\theta}_{nT}) - f_{kt,direct}(\theta_0)) \\ &\xrightarrow{d} N\left(0, \lim_{n,T\to\infty}\left(\frac{\partial f_{kt,direct}(\theta_0)}{\partial\theta'}\Sigma_{\theta_0,nT}\frac{\partial f_{kt,direct}(\theta_0)}{\partial\theta}\right)\right), \end{aligned} \tag{3.10}$$

$$\begin{aligned} &\sqrt{nT}(f_{kt,total}(\hat{\theta}_{nT}) - f_{kt,total}(\theta_0)) \\ &\xrightarrow{d} N\left(0, \lim_{n,T\to\infty}\left(\frac{\partial f_{kt,total}(\theta_0)}{\partial\theta'}\Sigma_{\theta_0,nT}\frac{\partial f_{kt,total}(\theta_0)}{\partial\theta}\right)\right), \end{aligned} \tag{3.11}$$

$$\begin{aligned} &\sqrt{nT}(f_{kt,indirect}(\hat{\theta}_{nT}) - f_{kt,indirect}(\theta_0)) \\ &\xrightarrow{d} N\left(0, \lim_{n,T\to\infty}\left(\frac{\partial f_{kt,indirect}(\theta_0)}{\partial\theta'}\Sigma_{\theta_0,nT}\frac{\partial f_{kt,indirect}(\theta_0)}{\partial\theta}\right)\right), \end{aligned} \tag{3.12}$$

and

$$\begin{aligned} &\sqrt{nT}(f_{kt,feedback}(\hat{\theta}_{nT}) - f_{kt,feedback}(\theta_0)) \\ &\xrightarrow{d} N\left(0, \lim_{n,T\to\infty}\left(\frac{\partial f_{kt,feedback}(\theta_0)}{\partial\theta'}\Sigma_{\theta_0,nT}\frac{\partial f_{kt,feedback}(\theta_0)}{\partial\theta}\right)\right), \end{aligned} \tag{3.13}$$

where $\partial f_{kt,direct}(\theta_0)/\partial\theta'$, $\partial f_{kt,total}(\theta_0)/\partial\theta'$, $\partial f_{kt,indirect}(\theta_0)/\partial\theta'$, and $\partial f_{kt,feedback}(\theta_0)/\partial\theta'$ can be found in the Supplementary Material.

Proposition 2. *Under some regularity conditions, the estimates for the impacts in* (3.4)–(3.8) *are consistent and asymptotically normally distributed, as in* (3.9)–(3.13)*. Specifically, the estimate for the element-wise effects is in* (3.9)*, the estimate for the direct effects is specified in* (3.10)*, the estimate for the total effects is in* (3.11)*, the estimate for the indirect effects is in* (3.12)*, and the estimate for the feedback effect is in* (3.13)*.*

Similarly to Proposition 1, these regularity conditions correspond to the assumptions used to derive the asymptotic properties of the relevant estimators.

4. Monte Carlo

4.1. Impacts for cross-sectional SAR

We first investigate the cross-sectional case. The DGP is

$$Y_n = \alpha_0 l_n + \lambda_0 W_n Y_n + X_{n,1}\beta_{01} + X_{n,2}\beta_{02} + W_n X_{n,1}\delta_{01} + \epsilon_n, \qquad (4.1)$$

where $X_{n,1}$ and $X_{n,2}$ are two independently generated standard normal (vector) variables and are *i.i.d.* for all i. Moreover, ϵ_{ni} are drawn independently from the standard normal distribution. We choose $n = 49$ and $n = 400$ as our sample size. The spatial weights matrix we use is the rook matrix based on an r board (so that $n = r^2$). The rook matrix represents a square tessellation with a connectivity of four for the inner fields on the chessboard, and two and three for the corner and border fields, respectively. Most empirically observed regional structures in spatial econometrics are made up of regions with connectivity close to the range of the rook tessellation. Given a rook matrix, we normalize it by its maximum row sum, and obtain the corresponding ML estimator of λ_0. The impact estimates are then based on the MLE. We compute the average direct, indirect, total, and feedback effects for a unit change in $X_{n,1}$; we also compute these effects for a change in $X_{n,2}$. In addition, we report the element-wise impact of how a change of the $j = 2$ or $j = 5$ units in $X_{n,1}$ affects the $i = 1$ unit in Y_n. Note that because of the structure of a rook matrix, the (1,2) element of W_n is nonzero, and its (1,5) element is zero. This choice of pair will help us to investigate the magnitude of a pairwise impact from a directly connected neighbor and an indirectly connected neighbor. The number of repetitions is 1,000 for each case. We report the mean (Mean), theoretical standard deviation (T-SD), empirical standard deviation (E-SD), and coverage probability (CP) at the 5% significance level. The true parameter λ_0 takes the value 0.5, and β_{01}, β_{02}, and δ_{01} are all set to 1.

Table 1 shows that various impact estimates are close to the true values. The estimate of the total impact is larger than the direct effect, as expected. We observe that the variance of the total impact is larger than that of the direct impact, because $\partial f_{k,direct}(\theta)/\partial\lambda = (1/n)tr(S_n^{-1}(\lambda)G_n(\lambda)C_{nk}(\theta))$ is smaller than $\partial f_{k,total}(\theta)/\partial\lambda = (1/n)l_n' S_n^{-1}(\lambda)G_n(\lambda)C_{nk}(\theta)l_n$, and so are $\partial f_{k,direct}(\theta)/\partial\beta$ and $\partial f_{k,total}(\theta)/\partial\beta$. The T-SD is similar to the E-SD, which implies that a statistical

Table 1. Impact Analysis for SAR Model.

		Normalized rook matrix							Normalized distance matrix with ϕ_d= 10						
$n = 49$			direct	indirect	total	feedback	(1,2)	(1,5)		direct	indirect	total	feedback	(1,2)	(1,5)
$X_{n,1}$		Impact_0	1.1991	2.4086	3.6076	0.1991	0.4083	0.0020	Impact_0	1.1796	2.1393	3.3189	0.1796	0.4024	0.0010
		Mean	1.2033	2.3415	3.5448	0.1802	0.4070	0.0021	Mean	1.1854	2.0941	3.2795	0.1636	0.4020	0.0011
		T-SD	0.1610	0.6755	0.7572	0.0734	0.0887	0.0018	T-SD	0.1596	0.5987	0.6801	0.0685	0.0917	0.0010
		E-SD	0.1739	0.6953	0.7832	0.0747	0.0944	0.0021	E-SD	0.1725	0.6200	0.7092	0.0706	0.0970	0.0011
		CP	0.9350	0.9030	0.9140	0.9030	0.9230	0.7890	CP	0.9310	0.9140	0.9080	0.8880	0.9220	0.7860
		Std1	0.0303	Std2	0.1138				Std1	0.0350	Std2	0.1091			
$n = 400$			direct	indirect	total	feedback	(1,2)	(1,5)		direct	indirect	total	feedback	(1,2)	(1,5)
$X_{n,1}$		Impact_0	1.2129	2.6582	3.8712	0.2129	0.4083	0.0010	Impact_0	1.2022	2.5506	3.7529	0.2022	0.4024	0.0010
		Mean	1.2139	2.6480	3.8619	0.2103	0.4089	0.0010	Mean	1.2033	2.5415	3.7448	0.1996	0.4034	0.0010
		T-SD	0.0557	0.2450	0.2718	0.0251	0.0287	0.0003	T-SD	0.0554	0.2354	0.2621	0.0243	0.0290	0.0003
		E-SD	0.0554	0.2461	0.2727	0.0254	0.0286	0.0003	E-SD	0.0552	0.2366	0.2633	0.0246	0.0287	0.0003
		CP	0.9550	0.9440	0.9480	0.9460	0.9460	0.9130	CP	0.9530	0.9450	0.9490	0.9380	0.9510	0.9030
		Std1	0.0195	Std2	0.0447				Std1	0.0252	Std2	0.0436			

Note: 1. Std1 is the standard deviation of the diagonal elements, and Std2 is that of the off-diagonal elements. 2. The weights matrix is normalized by its maximum row sum.

inference based on the T-SD would be reliable. This is confirmed from the CP values, which are close to the 95% theoretical value. When n is larger, the T-SD and E-SD are smaller, which is consistent with the theoretical prediction for a larger sample size; also, CPs would improve under a larger n. Furthermore, the variation of the off-diagonal elements in the impact matrix is larger than that of the diagonal elements in the impact matrix.

In addition to using a sparse weights matrix, such as the rook matrix, we investigate the impact analysis under a less sparse weights matrix. We construct Euclidean distances for the units on a regular lattice ($d_{ij} = \sqrt{(y_i - y_j)^2 + (x_i - x_j)^2}$, where (x_i, y_i) are the coordinates of the ith unit), and then use an exponential decay function to construct the weights ($w_{ij} = e^{-\phi_d d_{ij}}$ with $\phi_d = 10$). The results are reported in Table 1, and are similar to those for the rook matrix. However, the variation of the diagonal elements and off-diagonal elements in the impact matrix are larger in the distance weights matrix setting.

To investigate how the variance of the disturbances and the sparseness of W_n affect the performance of the impact coefficients, we increase the variance of the disturbances to four, and change ϕ_d from 10 to 5 and 1. With a larger variance of disturbances, the means of estimates are basically the same, but the E-SD becomes larger. As a result, the CPs become slightly smaller. Additionally, under a less sparse spatial weights matrix (ϕ_d is smaller), the CPs are slightly smaller, on average, whereas the biases and SDs are similar. Note that under a less sparse weights matrix, the variances of the elements in the direct and

Table 2. Direct and Indirect Effects for Spatial Dynamic Panel

Direct Effect		Marginal impacts					Accumulative impacts				
		Impact_0	Mean	T-SD	E-SD	CP	Impact_0	Mean	T-SD	E-SD	CP
$X_{nt,1}$	τ=0	1.0590	1.0571	0.0241	0.0243	0.955	1.0590	1.0571	0.0241	0.0243	0.955
	1	0.2976	0.2969	0.0163	0.0172	0.944	1.3566	1.3540	0.0344	0.0354	0.945
	2	0.0974	0.0979	0.0098	0.0103	0.938	1.4540	1.4519	0.0412	0.0426	0.941
	3	0.0352	0.0358	0.0054	0.0057	0.934	1.4892	1.4877	0.0451	0.0467	0.941
	4	0.0136	0.0140	0.0029	0.0031	0.932	1.5028	1.5017	0.0472	0.0488	0.944
	5	0.0055	0.0058	0.0015	0.0016	0.931	1.5083	1.5075	0.0482	0.0500	0.945
	6	0.0023	0.0025	0.0008	0.0009	0.930	1.5106	1.5100	0.0488	0.0506	0.944
	7	0.0010	0.0011	0.0004	0.0005	0.924	1.5116	1.5111	0.0491	0.0509	0.944
	8	0.0004	0.0005	0.0002	0.0002	0.919	1.5120	1.5116	0.0492	0.0510	0.944
	9	0.0002	0.0002	0.0001	0.0001	0.912	1.5122	1.5118	0.0493	0.0511	0.945
	10	0.0001	0.0001	0.0001	0.0001	0.908	1.5123	1.5119	0.0493	0.0511	0.945
Indirect Effect		Marginal impacts					Accumulative impacts				
		Impact_0	Mean	T-SD	E-SD	CP	Impact_0	Mean	T-SD	E-SD	CP
$X_{nt,1}$	τ=0	1.3737	1.3774	0.0644	0.0679	0.932	1.3737	1.3774	0.0644	0.0679	0.932
	1	0.8910	0.8992	0.0724	0.0804	0.926	2.2647	2.2766	0.1197	0.1323	0.924
	2	0.4849	0.4938	0.0653	0.0719	0.931	2.7497	2.7704	0.1755	0.1956	0.926
	3	0.2507	0.2587	0.0483	0.0532	0.938	3.0003	3.0290	0.2187	0.2440	0.929
	4	0.1269	0.1333	0.0323	0.0357	0.933	3.1273	3.1623	0.2482	0.2770	0.932
	5	0.0636	0.0683	0.0203	0.0227	0.934	3.1909	3.2306	0.2669	0.2980	0.930
	6	0.0318	0.0349	0.0124	0.0140	0.923	3.2227	3.2655	0.2783	0.3109	0.932
	7	0.0158	0.0179	0.0073	0.0084	0.920	3.2385	3.2834	0.2851	0.3186	0.932
	8	0.0079	0.0092	0.0043	0.0050	0.912	3.2463	3.2926	0.2890	0.3232	0.932
	9	0.0039	0.0047	0.0025	0.0029	0.910	3.2502	3.2973	0.2913	0.3258	0.932
	10	0.0019	0.0024	0.0014	0.0017	0.903	3.2521	3.2998	0.2926	0.3273	0.932

indirect impact matrices are much smaller. Detailed results are provided in the Supplementary Material.

4.2. Impacts for SDPD models

Here, we investigate the SDPD case, where the data are generated from

$$Y_{nt} = \lambda_0 W_n Y_{nt} + \gamma_0 Y_{n,t-1} + \rho_0 W_n Y_{n,t-1} + X_{nt,1}\beta_{01} + X_{nt,2}\beta_{02} + W_n X_{nt,1}\delta_{01} + \mathbf{c}_{n0} + \alpha_{t0} l_n + V_{nt},$$

using $\theta_0 = (0.2, 0.2, 1, 1, 1, 0.2)'$, where $\theta_0 = (\gamma_0, \rho_0, \beta_0', \delta_0', \lambda_0)'$, and X_{nt}, $\mathbf{c}_{n0}$, $\alpha_{T0} = (\alpha_1, \alpha_2, \ldots, \alpha_T)$, and V_{nt} are generated from independent standard normal distributions. We use a rook matrix as the spatial weights matrix. We generate the spatial panel data with $20 + T$ periods, where the starting value is from $N(0, I_n)$, and then take the last T periods as our sample. We use $n = 196$, and the number of periods in the sample is $T = 10$. For each generated sample observation, we calculate the bias-corrected ML estimator $\hat{\theta}_{nT}$ of Lee and Yu (2010a), and construct the average and element-wise impact analysis. We investigate how a one-time change in X_{nt} affects future values of $Y_{n,t+\tau}$, for $\tau \geq 0$, and how a permanent change in X_{nt} affects $Y_{n,t+\tau}$. We conduct the simulation under different horizons so that τ ranges from 0 to 10, and repeat the simulation 1,000 times. We construct the T-SD, E-SD, and CP of these impact estimates. Table 2 contains the results for the direct and indirect effects. Results for the total, feedback, and element-wise effects are provided in the Supplementary Material in order to conserve space.

We find that the average impacts and element-wise impacts have good finite-sample properties. The estimates are close to the true value, and the T-SD is close to the E-SD. The CPs for $n = 49$ are lower than 95%, partially because of the limited sample size, while the CPs perform much better when $n = 400$. The CP of the impact for the (1,5) pair is small, mainly owing to scaling imprecision, because the estimated value is very small.

The element-wise impacts have smaller values than the average impacts; however, a statistical inference based on element-wise impacts is still valid and has satisfactory finite-sample performance. The downward trend of the CPs for the marginal impacts in Table 2 is also the result of scaling imprecision, because the marginal estimates become negligible over time. We see that the CPs for the accumulative effects are still satisfactory. For all average and element-wise impacts, the estimates are close to zero if the time horizon is 10. This means that the influence of a current change in exogenous variables will have diminishing impacts on the future of dependent variables, over time.

5. Impact Analysis of Air Pollution in China

China has been experiencing severe air pollution in recent years, with the public becoming increasingly aware of this issue and paying more attention to pollution indices, such as the air quality index (AQI). The overall AQI is based on individual AQIs constructed from the pollutant concentrations of SO_2, NO_2, PM_{10}, $PM_{2.5}$, O_3, and CO, where the dominant AQI is reported for the location.

$PM_{2.5}$ is defined as fine particulate matter with a diameter of 2.5 micrometers or less, and can affect human health and cause many diseases, such as lung morbidity and respiratory and cardiovascular diseases (see Schwartz and Neas (2000); Pope III et.al. (2002)). Because the concentration of PM_{10} includes that of $PM_{2.5}$ ($PM_{2.5}$ accounts for 55% of PM_{10}, on average, in our sample) and the public is more concerned with $PM_{2.5}$, we focus on $PM_{2.5}$. Thus, it is important to understand the key factors that contribute to $PM_{2.5}$ in order for central and local governments to adopt effective tools to reduce air pollution. For example, China's State Council targeted a 25% reduction of $PM_{2.5}$ by 2017 from the 2012 level for Beijing, which resulted in considerably improved air quality for Beijing in 2018.

Our aim is to identify the key factors that contribute to $PM_{2.5}$ in Chinese cities. Then, we conduct an intercity impact analysis to determine how these factors affect air pollution in neighboring cities. Because the central government of China is considering environmental factors when evaluating the performance of local officials, intercity impact analyses would improve the design of promotion schemes for local officials.

5.1. Estimation equation and data

We link the $PM_{2.5}$ pollution level of a city to its meteorological factors and neighboring cities. To consider secondary atmospheric chemical reactions, we also include several air pollutants in the regression. Thus, the full regression equation is

$$\begin{aligned} Y_{nt} = {} & \lambda_0 W_n Y_{nt} + \gamma_0 Y_{n,t-1} + \rho_0 W_n Y_{n,t-1} + \mathbf{c}_{n0} + \alpha_{t0} l_n + V_{nt} \\ & + SO_{2,nt}\beta_{01} + NO_{2,nt}\beta_{02} + CO_{nt}\beta_{03} + X_{nt}\beta_{04}, \end{aligned}$$

where Y_{nt} is the $PM_{2.5}$ level, $\mathbf{c}_{n0}$ represents city fixed effects, α_{t0} represents time fixed effects, and $SO_{2,nt}$, $NO_{2,nt}$, and CO_{nt} are air pollutants included to allow for atmospheric chemical reactions. In addition, X_{nt} is a vector of meteorological factors, including temperature (°C), relative humidity (%), wind speed in different directions (meters per second), precipitation (millimeters), and atmospheric pressure (hPa, i.e., 100 pascals). Because wind moving in different directions can have heterogenous effects on air pollution, depending on the region (i.e., north or south) and season (i.e., we define the winter heating season as the period when the northern cities provide collective heating), we have four variables for each wind direction. For example, the regressors associated with northeast winds are NE-NorWint, NE-NorSumm, NE-SouWint, and NE-SouSumm, which correspond to combinations of regions and seasons.

For the spatial weights matrix in our empirical analysis, the baseline is constructed from the geographical distance d_{ij}, so that

$$w_{ij}^{g} = \frac{\exp(-\phi_d d_{ij})}{\max_i(\sum_{j=1}^{n} \exp(-\phi_d d_{ij}))}, \tag{5.1}$$

where $100d_{ij}$ is the distance between prefectural cities in kilometers and $\max_i(\sum_{j=1}^{n} \exp(-\phi_d d_{ij}))$ is for row normalization. By searching over different ϕ_d, we choose $\phi_d = 1.8$, which maximizes the log likelihood of the estimation equation. Because the spillover effect of air pollution can be hindered by mountains, we also construct a more realistic spatial weights matrix to consider mountain barriers. To do so, we first obtain the difference in the average elevation of city j and that of the mountains between city i and city j. Denote this difference by h_{ij}. We then construct an index matrix W_n^m, with element w_{ij}^m equal to zero if $h_{ij} > h^*$, where h^* is the critical value for the mountain barrier to hinder air pollution spillover. The (i, j) element of the matrix $W_n^{g \circ m} = W_n^g \circ W_n^m$ is

$$w_{ij}^{g \circ m} = \begin{cases} w_{ij}^{g}, & \text{if } h_{ij} \le h^* \\ 0, & \text{if } h_{ij} > h^* \end{cases}, \tag{5.2}$$

where $\circ$ is the Hadamard product with an element-wise product. By searching over ϕ_d and h^* simultaneously to maximize the log-likelihood function, we specify $\phi_d = 1.6$ and $h^* = 1{,}400$ meters.

Note that it is possible to combine time-varying wind features with time-invariant geographical information to construct a time-varying spatial weights matrix W_{nt} in the estimation. However, this combination might cause W_{nt} to be endogenous. For example, if wind blows from city j to city i, the corresponding spatial weight could be positive or negative, depending on whether the air pollution level in city j is higher or lower than that in city i; this causes the corresponding spatial weight $w_{ij,t}$ to be endogenous. We leave this issue for future research.

We have daily air pollution data with meteorological facts for 363 prefectural-level cities in China for 2016–2018. These data are from 1,630 weather stations and 2,253 monitor stations for air pollution. The pollution data are from the air-quality real-time release system of the Environmental Monitoring Station by MEP of China. The meteorological data are from the China Meteorological Administration. Although the $PM_{2.5}$ data for prefectural cities are available from 2013 for 74 cities, they are not available for the majority of prefectural cities until December 2014 (289 cities). From December 2015, nearly all prefectural cities provide air pollution data (328 cities). Summary statistics for the air pollution and meteorological factors in 2016, 2017, and 2018 are provided in the Supplementary Material. These data sources provide panel data with $n = 336$ and $T = 1{,}096$. We use the bias corrected MLE of Lee and Yu (2010a) to obtain the estimates for the SDPD model, and use (3.9)–(3.13) to compute the variance matrices of the various impacts. Given the the sample size $n = 336$, the computational burden is not an issue.

5.2. Empirical results

The results with W_n^g are presented in Table 3. We see that the spatial effect is significant, so that $PM_{2.5}$ has a significant spillover effect on neighboring cities. The dynamic coefficient of 0.498 indicates moderate state dependence, meaning that $PM_{2.5}$ today will partially affect air pollution tomorrow. We find that all meteorological factors are significant, where temperature and humidity contribute to air pollution, whereas precipitation and atmospheric pressure can alleviate it. For wind speed from different directions, we see that winds from most directions do not alleviate air pollution, except for the northeast wind for northern cities in the winter, southeast wind for southern cities in the winter, and southwest wind for cities in the south during the summer. We also see that the direct and indirect effects are similar in our empirical results, indicating that a one-unit change in x_{it} has a similar impact on its final change in y_{it} and the equilibrium change in y_{jt}.

In the element-wise impact analysis, we assess the impact of different cities on Beijing, where wind plays an important role. The wind in Beijing comes mainly from the northwest (NW) and southeast (SE), where the winter season mainly has NW winds and the summer season has SE winds. Moreover, Beijing has the Yan Mountains to its north and Taihang Mountains to its west. Because the areas to the north and west of Beijing are less industrialized, and thus less polluted, the NW wind usually helps to reduce air pollution in Beijing. However, because the areas to the south and east of Beijing are more populated and have more industries, SW winds usually accelerate air pollution, especially in winter. Additionally, the mountains to the north and west block pollution diffusion in the presence of a SE wind. We can investigate how northern cities, such as Zhangjiakou and Chengde, and southern cities, such as Baoding, Langfang, Tianjin, and Tangshan, affect the air pollution in Beijing (see Figure 1, where darker color indicates higher altitude). Shenyang and Shanghai are far from Beijing and can be used as a robustness check. From the right columns in Table 3, we see that changes in the variables in neighboring cities affect the air pollution in Chinese cities. Furthermore, wind speed plays an important role in meteorological factors.

Table 3. Air Pollution Spillover and Impact Analysis with W_n^g.

	Coefficient	Direct	Indirect	Total	Feedback	Zhangjiakou	Chengde	Baoding	Shenyang
$Y_{n,t-1}$	0.498								
	(0.001)								
W_nY_{nt}	0.618								
	(0.006)								
$W_nY_{n,t-1}$	−0.325								
	(0.004)								
Temp	0.111	0.122	0.169	0.292	0.011	0.005	0.006	0.012	0.000
	(0.009)	(0.009)	(0.013)	(0.022)	(0.001)	(0.000)	(0.000	(0.001)	(0.000)
Humi	0.087	0.096	0.133	0.229	0.008	0.004	0.005	0.009	0.000
	(0.002)	(0.003)	(0.004)	(0.006)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Prec	−0.015	−0.017	−0.023	−0.040	−0.001	−0.001	−0.001	−0.002	0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Pres	−0.010	−0.011	−0.015	−0.025	−0.001	0.000	−0.001	−0.001	0.000
	(0.005)	(0.006)	(0.008)	(0.013)	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)
NE-NorWint	0.101	0.110	0.153	0.263	0.010	0.004	0.006	0.010	0.000
	(0.078)	(0.085)	(0.118)	(0.203)	(0.007)	(0.003)	(0.004)	(0.008)	(0.000)
NW-NorWint	0.640	0.701	0.973	1.674	0.061	0.027	0.037	0.066	0.001
	(0.072)	(0.079)	(0.110)	(0.189)	(0.007)	(0.003)	(0.004)	(0.007)	(0.000)
SE-NorWint	−0.054	−0.059	−0.082	−0.141	−0.005	−0.002	−0.003	−0.006	0.000
	(0.068)	(0.074)	(0.103)	(0.178)	(0.007)	(0.003)	(0.004)	(0.007)	(0.000)
SW-NorWint	0.199	0.218	0.303	0.521	0.019	0.008	0.011	0.021	0.000
	(0.059)	(0.065)	(0.090)	(0.154)	(0.006)	(0.003)	(0.003)	(0.006)	(0.000)
NE-NorSumm	0.745	0.817	1.134	1.951	0.072	0.032	0.043	0.077	0.002
	(0.079)	(0.086)	(0.120)	(0.206)	(0.008)	(0.003)	(0.005)	(0.008)	(0.000)
NW-NorSumm	0.243	0.266	0.369	0.635	0.023	0.010	0.014	0.025	0.001
	(0.071)	(0.078)	(0.108)	(0.186)	(0.007)	(0.003)	(0.004)	(0.007)	(0.000)
SE-NorSumm	0.311	0.341	0.473	0.815	0.030	0.013	0.018	0.032	0.001
	(0.061)	(0.065)	(0.091)	(0.156)	(0.006)	(0.003)	(0.003)	(0.006)	(0.000)
SW-NorSumm	0.647	0.709	0.984	1.693	0.062	0.028	0.037	0.067	0.001
	(0.050)	(0.055)	(0.076)	(0.131)	(0.005)	(0.002)	(0.003)	(0.005)	(0.000)
NE-SouWint	−0.638	−0.699	−0.971	−1.670	−0.061	−0.027	−0.037	−0.066	−0.001
	(0.082)	(0.090)	(0.125)	(0.214)	(0.008)	(0.003)	(0.005)	(0.009)	(0.000)
NW-SouWint	0.290	0.317	0.441	0.758	0.028	0.012	0.017	0.030	0.001
	(0.073)	(0.080)	(0.111)	(0.190)	(0.007)	(0.003)	(0.004)	(0.008)	(0.000)
SE-SouWint	0.092	0.100	0.139	0.240	0.009	0.004	0.005	0.010	0.000
	(0.078)	(0.085)	(0.118)	(0.204)	(0.007)	(0.003)	(0.004)	(0.008)	(0.000)
SW-SouWint	0.524	0.574	0.797	1.371	0.050	0.022	0.030	0.054	0.001
	(0.066)	(0.072)	(0.100)	(0.173)	(0.006)	(0.003)	(0.004)	(0.007)	(0.000)
NE-SouSumm	0.046	0.050	0.070	0.120	0.004	0.002	0.003	0.005	0.000
	(0.078)	(0.085)	(0.119)	(0.204)	(0.007)	(0.003)	(0.004)	(0.008)	(0.000)
NW-SouSumm	0.052	0.057	0.079	0.135	0.005	0.002	0.003	0.005	0.000
	(0.066)	(0.073)	(0.101)	(0.174)	(0.006)	(0.003)	(0.004)	(0.007)	(0.000)
SE-SouSumm	0.153	0.167	0.232	0.399	0.015	0.007	0.009	0.016	0.000
	(0.069)	(0.075)	(0.104)	(0.179)	(0.007)	(0.003)	(0.004)	(0.007)	(0.000)
SW-SouSumm	0.571	0.626	0.870	1.496	0.055	0.024	0.033	0.059	0.001
	(0.057)	(0.063)	(0.087)	(0.150)	(0.006)	(0.002)	(0.003)	(0.006)	(0.000)
SO2	0.094	0.103	0.143	0.246	0.009	0.004	0.005	0.010	0.000
	(0.002)	(0.002)	(0.003)	(0.005)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
NO2	0.427	0.468	0.650	1.118	0.041	0.018	0.024	0.044	0.001
	(0.004)	(0.003)	(0.005)	(0.008)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
CO	5.039	5.523	7.668	13.191	0.484	0.215	0.289	0.523	0.011
	(0.063)	(0.065)	(0.099)	(0.160)	(0.006)	(0.003)	(0.004)	(0.006)	(0.000)

Note: 1. Standard deviations in parentheses. 2. Due to space limit, the element-wise impacts for Langfang, Tianjin, Tangshan, and Shanghai are provided in the Supplementary Material.

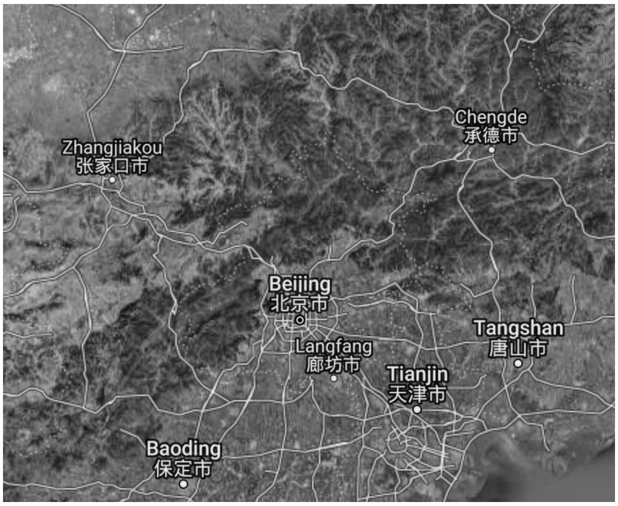

Figure 1. Map of Beijing and Its Neighboring Cities.

For the wind in Zhangjiakou, a one-unit increase in the wind speed from the NW or SW increases the pollution level in Beijing by 0.001 and 0.028, respectively, in summer. For other cities that are south of Beijing and are more polluted, these effects are more prominent. For example, in Baoding, a one-unit increase in wind speed from the NW or SW increases the $PM_{2.5}$ in Beijing by 0.025 and 0.067, respectively, in summer. In terms of air pollutants, we find that CO dominates other air pollutants in terms of affecting neighboring cities. Similarly to the case of wind speed, the cities to the south have a larger effect than those from the north.

We also report results that account for the barrier effect of mountains. Owing to space limitations, the details are provided in the Supplementary Material. With a more precise spatial weights matrix, the spatial effects increase from 0.618 to 0.756 in the levels regression. Intuitively, the misspecification of the positive weights from the true zero weights dilutes the spatial effect coefficient, because λ is a measure of the average effect given the spatial weights matrix. Consequently, the element-wise analysis shows that a one-unit change in the meteorological variables in neighboring cities has a larger effect on $PM_{2.5}$ in Beijing.

6. Conclusion

We have proposed an impact analysis for spatial models and investigated their statistical inference. With scalar effects defined, the effect of a one-unit change in the exogenous variables on the dependent variable can be better evaluated than in a structural SAR model representation. The proposed impact effects have satisfactory finite-sample performance. We applied the proposed analysis to study air pollution in China, finding that $PM_{2.5}$ is moderately persistent and strongly spatially correlated. All meteorological factors have significant effects on $PM_{2.5}$, especially wind direction and wind speed.

Two topics are left to future research. First, in obtaining the impact analysis, the spatial weights matrix is given exogenously. It would be of interest to extend the analysis to the setting of an endogenous spatial weights matrix, as in Qu and Lee (2015) and Qu, Lee and Yu (2017). Second, for the spatial panel data model, if the underlying regression coefficients are time varying, a corresponding impact analysis based on a time-varying approach is also needed.

Supplementary Material

The online Supplementary Material contains some algebra for the impact analysis, further simulation results, summary statistics, and additional empirical results.

Acknowledgments

We would like to thank the two anonymous referees and associate editor of this journal for their helpful comments. Yu gratefully acknowledges the financial support from the National Natural Science Foundation of China (No. 71925006 and No. 92046021) and China's National Key Research Special Program Grant (No. 2016YFC0207705).

References

Anselin, L., LeGallo, J. and Jayet, H. (2008). Spatial panel econometrics. In *The Econometrics of Panel Data* (Edited by L. Matyas and P. Sevestre), 625–660. Springer-Verlag, Berlin, Heidelberg.

Baltagi, B., Song, S. H. and Koh, W. (2003). Testing panel data regression models with spatial error correlation. *Journal of Econometrics* **117**, 123–150.

Baltagi, B., Song, S. H., Jung, B. C. and Koh, W. (2007). Testing for serial correlation, spatial autocorrelation and random effects using panel data. *Journal of Econometrics* **140**, 5–51.

Cliff, A. D. and Ord, J. K. (1973). *Spatial Autocorrelation.* Pion Ltd., London.

Debarsy N., Ertur, C. and LeSage, J. P. (2012). Interpreting dynamic space-time panel data models. *Statistical Methodology* **9**, 158–171.

Elhorst, J. P. (2005). Unconditional maximum likelihood estimation of linear and log-linear dynamic models for spatial panels. *Geographical Analysis* **37**, 85–106.

Kapoor, M., Kelejian, H. H. and Prucha, I. R. (2007). Panel data models with spatially correlated error components. *Journal of Econometrics* **140**, 97–130.

Kelejian, H. H. and Prucha, I. R. (1998). A generalized spatial two-stage least squares procedure for estimating a spatial autoregressive model with autoregressive disturbance. *Journal of Real Estate Finance and Economics* **17**, 99–121.

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

Lee, L. F. (2004). Asymptotic distributions of quasi-maximum likelihood estimators for spatial econometric models. *Econometrica* **72**, 1899–1925.

Lee, L. F. (2007). GMM and 2SLS estimation of mixed regressive, spatial autoregressive models. *Journal of Econometrics* **137**, 489–514.

Lee, L. F. and Yu, J. (2010a). A spatial dynamic panel data model with both time and individual fixed effects. *Econometric Theory* **26**, 564–597.

Lee, L. F. and Yu, J. (2010b). Estimation of spatial autoregressive panel data models with fixed effects. *Journal of Econometrics* **154**, 165–185.

Lee, L. F. and Yu, J. (2012). Spatial panels: Random components vs. fixed effects. *International Economic Review* **53**, 1369–1412.

Pope III, C. A., Burnett, R. T., Thun, M. J. Calle, E. E., Krewski, D., Ito, K. et al. (2002). Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution. *JAMA* **287**, 1132–1141.

Qu, X. and Lee, L. F. (2015). Estimating a spatial autoregressive model with an endogenous spatial weight matrix. *Journal of Econometrics* **184**, 209–232.

Qu, X., Lee, L. F. and Yu, J. (2017). QML estimation of spatial dynamic panel data models with endogenous time varying spatial weights matrices. *Journal of Econometrics* **197**, 173–201.

Schwartz, J. and Neas, L. M. (2000). Fine particles are more strongly associated than coarse particles with acute respiratory health effects in schoolchildren. *Epidemiology* **11**, 6–10.

Su, L. and Yang, Z. (2015). QML estimation of dynamic panel data models with spatial errors. *Journal of Econometrics* **185**, 230–258.

Yu, J., de Jong, R. and Lee, L. F. (2008). 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.

Yu, J. and Lee, L. F. (2010). Estimation of unit root spatial dynamic panel data models. *Econometric Theory* **26**, 1332–1362.

Yu, J., de Jong, R. and Lee, L. F. (2012). Estimation for spatial dynamic panel data with fixed effects: The case of spatial cointegration. *Journal of Econometrics* **167**, 16–37.

Hsuan-Yu Chang

The Center for Green Economy, Chung-Hua Institution for Economic Research, Taipei 10672.

E-mail: hychang@cier.edu.tw

Jihai Yu

Guanghua School of Management, Peking University, Beijing 100871.

E-mail: jihai.yu@gsm.pku.edu.cn

(Received March 2020; accepted April 2022)