JOURNAL OF REGIONAL SCIENCE

JOURNAL OF REGIONAL SCIENCE, VOL. 00, NO. 0, 2014, pp. 1–21

INTERPRETING SPATIAL ECONOMETRIC ORIGIN-DESTINATION FLOW MODELS

James P. LeSage

Department, of Finance and Economics, Texas State University, 601 University Drive, San Marcos, TX 78666. E-mail: jlesage@spatial-econometrics.com

Christine Thomas-Agnan

Toulouse School of Economics, Manufacture des Tabacs – Batiment F, 21 Allee de Brienne, 31000, Toulouse, France. E-mail: christine.thomas@tse-fr.eu

ABSTRACT. Spatial interaction or gravity models have been used to model flows that take many forms, for example population migration, commodity flows, traffic flows, all of which reflect movements between origin and destination regions. We focus on how to interpret estimates from spatial autoregressive extensions to the conventional regression-based gravity models that relax the assumption of independence between flows. These models proposed by LeSage and Pace (2008, 2009) define spatial dependence involving flows between regions. We show how to calculate partial derivative expressions for these models that can be used to quantify these various types of effect that arise from changes in the characteristics/explanatory variables of the model.

1. INTRODUCTION TO GRAVITY OR SPATIAL INTERACTION MODELS

Traditional gravity models assume independence among the dependent variable observations that reflect origin-destination (O-D) flows. LeSage and Pace (2008, 2009) provide specifics regarding spatial regression variants of traditional gravity or spatial interaction models. Smith (1975) and Sen and Smith (1995) use this label because the regional interaction is proportional to the product of regional size measures, and inversely related to distance between regions. In the case of interregional commodity flows, the measure of regional size is typically gross regional product or regional income. The model predicts more interaction in the form of commodity flows between regions of similar (economic) size than regions dissimilar in size. For the case of migration flows, population would be a logical measure of regional size, and in other contexts such as knowledge flows between regions, LeSage, Fischer, and Scherngell (2007) use regional knowledge stocks measured by patents to reflect size.

A complete exposition of econometric issues encountered in both traditional and spatial autoregressive (SAR) interaction models can be found in LeSage and Fischer (2010) and LeSage and Pace (2009, chapter 8). A history of the literature in this area is in Griffith (2007), and work on theoretical foundations in the context of international trade theory can be found in Anderson and van Wincoop (2003), with empirical applications in Behrens, Ertur, and Koch (2012), and Koch and LeSage (2012). Forecasting of these models is taken up in LeSage and Llano-Verduras (2014), and early work modeling spatial dependence in the disturbance structure can be found in Bolduc, Laferriere, and Santarossa (1992), with more recent work in Fischer and Griffith (2008).

Received: May 2013; revised: January 2014; accepted: January 2014.

 DOI: 10.1111/jors.12114

LeSage and Pace (2008, 2009) focus on maximum likelihood and Bayesian estimation of their spatial regression variants of traditional gravity or spatial interaction models. However, they leave the issue of interpreting the resulting coefficient estimates from these models as an issue for future research.

Spatial spillovers in these models can take the form of spillovers to both regions/observations neighboring the origin or destination in the dyadic relationships that characterize O-D flows. For example, in a sample involving n regions, a (ceteris paribus) decrease in taxes in a single region i would lead to: (1) inflows of population to this region from (potentially) all other regions, and (2) a decrease in outflows of population to (potentially) all other regions. Conventional interaction models assume independence between flows involving region i in the dyad and flows not involving i as an origin or destination. This means flows between other regions not involving i as an origin or destination would not be affected by this decrease in taxes for region i.

Spatial spillovers in the context of a spatial regression variant of the conventional interaction model reflect situations that violate this traditional assumption. LeSage and Pace (2008) introduced a specification where dependence is specified for flows involving an O-D dyad consisting of regions i, j. Their specification allows for: dependence of flows in/out of regions neighboring i to be influenced by changes in taxes at origin i (origin dependence), flows to regions nearby destination j to be impacted by tax changes in region i (destination dependence), and flows from regions neighboring the origin i to regions neighboring the destination j could also be impacted (O-D dependence).

Spatial dependence between conventional flow dyads greatly complicates the task of interpreting estimates from these models. We show how to calculate partial derivative expressions for these models that can be used to quantify: origin effects, destination effects, intraregional effects, and network/spillover effects arising from changes in the characteristics/explanatory variables in SAR interaction models. Changes in the characteristics of a single region i produce impacts on (potentially) all elements of the $n \times n$ flow matrix. Since we typically consider changes in characteristics of all $i = 1, \ldots, n$ regions, this produces a set of n different $n \times n$ matrices of partial derivatives associated with changes in each explanatory variable in the model. Scalar summary measures are proposed for the various types of effects that arise in these models allowing interpretation of estimates in a manner similar to that used in typical regression models.

Section 1 introduces the conventional interaction model that assumes independence between observed flows and relies on ordinary least squares estimation methods. In Section 2, we set forth the SAR extension to the conventional model introduced by LeSage and Pace (2008). The subject of interpreting estimates from independent and SAR interaction models is taken up in Section 2. We set forth an approach for these models that can be used to produce scalar expressions for the various types of impacts (origin, destination, intraregional, and spillover) that arise from changes in the explanatory variables of the model. Spatial spillovers in these models can take the form of network spillovers that impact regions not directly involved as origin or destination regions in the dyadic flow relationships that characterize dependent variable observations in these models.

Gravity or Spatial Interaction Models based on Independence

Regression models attempt to explain variation in the n^2 flows between the n regions in a closed network of regional flows. The $n \times n$ flow matrix Y is converted to an $n^2 \times 1$ vector by stacking columns. The flow matrix might be arranged so the i, jth element reflects a flow from region j to region i, which has been labeled an *origin-centric* flow arrangement by LeSage and Pace (2008). Many trade models rely on the convention that the i, jth element of the flow matrix represents a flow from region i to j, which would be a *destination-centric* arrangement of the flows. If we let y^o denote the origin-centric vector of flows and y^d a vector created by stacking columns from a destination-centric arrangement, there is a vec-permutation matrix P that can be used to relate these two different orderings. Specifically, $Py^o = y^d$, and using properties of permutation matrices, $y^o = P^{-1}y^d = P'y^d$.

A regression model that has been labeled a *gravity model* captures the notion that size of the two regions and distance are important factors that determine the magnitude of flows between regions. For example if one starts with the standard gravity model (c.f., equation (6.4) in Sen and Smith, 1995) shown in (1) and applies a log-transformation, the regression in (2) arises

$$\mu(i,j) = CX_d(i)X_o(j)H(i,j). \tag{1}$$

In (1), $\mu(i,j)$ represents the expected flows from region j to region i (assuming an origin-centric flow matrix), while $X_d(i)$, $X_o(j)$ denote sizes of the destination and origin and $H(i,j)$ represents resistance or deterrence to flows between the origin and destination, typically modeled using some function of distance between regions i and j. To facilitate the log-transformation, $X_o(j)$ can be specified using $X_o(j)^{\beta_o}$ and similarly, $X_d(i) = X_d(i)^{\beta_d}$, while $H(i,j)$ is some function of distance between regions i and j, for which we might use a power function, $D(i,j)^{\gamma}$, where $D(i,j)$ is the distance between regions i and j.

A point made by LeSage and Pace (2009) is that conventional work with these models has relied on mathematics emphasizing dyads i, j which has severe limitations for thinking about flows in the context of a network. Spatial dependence reflects relationships between observations, and is typically modeled using vectors and *spatial weight matrices* to express relations between observations. LeSage and Pace (2008) use the matrix/vector representation of the log-transformed dyad expression in (1) shown in (2), which more closely mirrors notation from conventional regression modeling

$$y = \alpha\iota_{n^2} + X_o\beta_o + X_d\beta_d + \gamma g + \varepsilon. \tag{2}$$

In (2), y is an $n^2 \times 1$ vector of (logged) flows constructed by stacking the columns of the $n \times n$ flow matrix Y, where we will assume an origin-centric organization throughout this paper. Similarly, applying the log transformation to the $n \times n$ matrix of distances $D(i,j)$ between the n destination and origin regions and stacking the columns results in a vector of logged distances g, with associated coefficient γ. LeSage and Pace (2008) show that $X_d = \iota_n \otimes X$, where X is an $n \times r$ matrix of characteristics for the n regions, $\otimes$ represents a *Kronecker product*, and ι_n is an $n \times 1$ vector of ones. In the simplest case, X might represent a vector with the appropriate size measure for each region, but without loss of generality this may be a matrix containing r characteristics of the regions that are thought to explain variation in flows. We note that this represents a general case, where the same set of explanatory variables is used for both origins and destinations. A special case might involve selection of a subset of variables in the matrix X for use as origin characteristics, and another subset of variables for the destination characteristics. However, the general case maybe the preferred approach, since inclusion of additional unimportant explanatory variables does not bias least-squares estimates, whereas exclusion of important explanatory variables can result in omitted variables bias.

The Kronecker product repeats the same values of the n regions in a strategic fashion to create a vector (or matrix) of sizes associated with each destination region, hence use of the notation X_d to represent these explanatory variables reflecting destination characteristics. Ultimately, use of Kronecker products in conjunction with matrix algebra allowed LeSage and Pace (2008) to express simple estimation expressions that avoid storing multiple copies of the same numerical values, which is computationally inefficient. The matrix/vector $X_o = X \otimes \iota_n$, arranges the n regional characteristics to match the vector

y, producing explanatory variables associated with each origin region. The vectors β_d and β_o are $r \times 1$ parameter vectors associated with the destination and origin region characteristics, respectively. The scalar parameter γ reflects the effect of the vector of logged distances g on flows, which is traditionally thought to be negative. The parameter α denotes the constant term parameter, and the $n^2 \times 1$ vector ε represents zero mean, constant variance, zero covariance disturbances, consistent with the Gauss–Markov least squares assumptions. We note that the assumption of log-normally distributed disturbances in the mean expression (1) leads to a normal distribution for the logged dependent variable flows in (2), consistent with least squares. This is not consistent with some flows which represent count data. For example counts of persons migrating or commuters traveling from one region to another. However, the log transformation may help to produce more normally distributed flows. Our discussion of interpreting estimates for these models does not require consideration of these issues, so we assume estimates exist for the model parameters and associated variance-covariance matrix.

SAR Interaction Models

Intuitively, changes to the characteristics of a single region i will impact both inflows and outflows to all other regions engaged or connected with region i as either an origin or destination. For example, a (ceteris paribus) decrease in taxes in region i would lead to inflows of population to this region from (potentially) all other regions and a decrease in outflows of population to (potentially) all other regions.

LeSage and Pace (2008) suggest that flows across networks involving origins and destinations are likely to exhibit spatial dependence. They define spatial dependence in this type of setting to mean that larger observed flows from an origin region A to a destination region Z are accompanied by: (1) larger flows from regions nearby the origin A to the destination Z, say regions B and C that are neighbors to region A, which they label *origin-dependence*, (2) larger flows from the origin region A to regions neighboring the destination region Z, say regions X and Y, which they label *destination-dependence*, and (3) larger flows from regions that are neighbors to the origin (B and C) to regions that are neighbors to the destination (X and Y), which they label *origin-to-destination dependence*.

Casual observation of migration flows in a network of counties are consistent with this type of observation. If there are a large number of migrants moving away from a county A (say a county near the Detroit metropolitan area), we would expect to see migrants also moving away from other counties B and C near Detroit (presumably due to unfavorable regional labor market conditions). Similarly, if a large number of migrants are moving into a county Z (say a county in the Austin metropolitan area), we would expect to see migrants also moving into other counties X and Y in the Austin metropolitan area (presumably because of favorable regional labor market conditions).

LeSage and Pace (2008, 2009) propose a spatial autoregression extension of the independent empirical gravity model from (2) shown in (3)

$$
\begin{aligned}
(3) \qquad Ay &= \alpha\iota_{n^2} + X_d\beta_d + X_o\beta_o + g\gamma + \varepsilon \\
A &= (I_{n^2} - \rho_d W_d)(I_{n^2} - \rho_o W_o) \\
&= (I_{n^2} - \rho_d W_d - \rho_o W_o + \rho_w W_w) \\
W_d &= I_n \otimes W \\
W_o &= W \otimes I_n \\
W_w &= W_d \otimes W_o = W_o \otimes W_d = W \otimes W.
\end{aligned}
$$

The term A can be viewed as a spatial filter that captures: *origin-based dependence*, *destination-based dependence*, and *origin-to-destination based dependence*.[1] The model and associated data generating process (DGP) for the SAR interaction model take the forms shown in (4) and (5), respectively, where we rely on the definitions: $Z = (\,\iota_{n^2}\ X_d\ X_o\ g\,)$ and $\delta = (\,\alpha\ \beta_d\ \beta_o\ \gamma\,)'$

$$y = \rho_o W_o y + \rho_d W_d y + \rho_w W_w y + Z\delta + \varepsilon, \tag{4}$$

$$y = (I_{n^2} - \rho_o W_o - \rho_d W_d + \rho_w W_w)^{-1}(Z\delta + \varepsilon). \tag{5}$$

The spatial lag formed by the matrix product $W_d y$ extracts (logged) flows to neighbors of each destination region in the vector of origin-destination flow dyads to form a linear combination of logged flows to neighboring destinations. In the case where the $n \times n$ spatial weight matrix W represents a fixed number say m of equally weighted nearest neighbors, the spatial lag vector would contain an average of (logged) flows to the m neighboring destinations.[2] The matrix W is a conventional (row-normalized) spatial weight matrix of the type used in cross-sectional regressions involving n regions. This spatial lag captures destination-based dependence, with the parameter ρ_d measuring the strength of destination-based dependence.

A similar interpretation applies to the spatial lag formed by the product $W_o y$, which reflects a linear combination of (logged) flows from regions neighboring the origin, again for each O-D dyad in the flow vector. The scalar parameter ρ_o reflects the strength of origin-based dependence. The spatial lag $W_w y$ forms a linear combination of (logged) flows from neighbors to the origin and (logged) flows to neighbors of the destination, and the parameter ρ_w represents the magnitude of this type of dependence.

The stability restrictions for the spatial dependence parameters require that $1/\lambda_{\min} < \rho_o + \rho_d + \rho_w < 1$, where $\lambda_{\min}$ is the minimum eigenvalue of the matrix W.

LeSage and Pace (2008) provide details concerning maximum likelihood estimation for the SAR interaction model, and LeSage and Pace (2009) set forth a Bayesian Markov chain Monte Carlo (MCMC) estimation scheme. Both of these exploit computationally efficient moment matrices involving the sample data expressed using the smaller dimension $n \times n$ matrix W rather than forming the larger $n^2 \times n^2$ matrix A.

A Motivation for the SAR Interaction Model

An econometric motivation for the SAR interaction model specification is provided here, based on the notion that location decisions of economic agents are influenced by behavior of other agents in previous periods.[3] For example, when considering commuting to work flows as in our empirical application, residents might be influenced by nearby flows (congestion) resulting from past location decisions of other residents in neighboring

[1] The filter implies a restriction that $\rho_w = -\rho_o\rho_d$. This restriction need not be imposed during estimation, so we address the more general case here and allow for an unrestricted parameter ρ_w.

[2] For the case of a weight matrix that uses equal-weighting of neighboring observations, this implies a geometric average of the nonlogged flows.

[3] This motivation including the mathematical developments that follow is analogous to LeSage and Pace (2009, pp. 25–26.) for the case of a SAR model. Their development extends in a straightforward way to the case of the SAR interaction model considered here.

regions, or firms influenced by congestion arising from location decisions of nearby firms in the past. (We will use commuting-to-work flows in our discussion of the motivation set forth here for concreteness.)

We can formally express this type of dyadic O-D flow dependence of y_t at time t on past flows y_{t-1} as:

$$
\begin{aligned}
y_t &= Gy_{t-1} + Z\delta + \varepsilon_t, \\
G &= (\rho_d W_d + \rho_o W_o + \rho_w W_w), \\
Z &= \left(X_d \; X_o \; g \right), \\
\delta &= \left(\beta_d \; \beta_o \; \gamma \right)', \\
\varepsilon_t &\sim N(0, \sigma^2 I_n),
\end{aligned}
\tag{6}
$$

where we have assumed that underlying characteristics of the regions X remain relatively fixed over time or exhibit growth at a constant rate: $X_t = \phi^t X_0$, allowing us to write Z without a time subscript. Since the characteristics in flow models represent "size" measures of regions, this assumption seems reasonable. In the case of commuting flows, the assumption implies that regional growth of residents and jobs is constant over time, and dependent on initial period size. Of course, this assumption need only be approximately valid to justify the results that follow.

Expression (6) indicates that (commuting-to-work) flows between O-D dyads at time t depend on past period flows observed by residents and firms in regions neighboring their origin ($W_o y_{t-1}$) and destination regions ($W_d y_{t-1}$), as well as flows between regions neighboring the origin to regions neighboring the destination ($W_w y_{t-1}$).

The dynamic relationship in (6) implies a relationship for time $t-1$ shown in (7), which can be used to replace y_{t-1} in (6), resulting in the expressions in (8) and (9)

$$y_{t-1} = Gy_{t-2} + Z\delta + \varepsilon_{t-1}, \tag{7}$$

$$y_t = Z\delta + G(Z\delta + Gy_{t-2} + \varepsilon_{t-1}) + \varepsilon_t, \tag{8}$$

$$y_t = Z\delta + GZ\delta + Gy_{t-1} + G\varepsilon_{t-1} + \varepsilon_t. \tag{9}$$

Recursive substitution of past values for the vector y_{t-r} on the right-hand side of (9) over q time periods leads to (10).

$$
\begin{aligned}
y_t &= (I_{n^2} + G + G^2 + \cdots + G^{q-1})Z\delta + G^q y_{t-q} + u, \\
u &= \varepsilon_t + G\varepsilon_{t-1} + G^2\varepsilon_{t-2} + \cdots + G^{q-1}\varepsilon_{t-(q-1)}.
\end{aligned}
\tag{10}
$$

Expression (10) can be simplified by noting that $E(\varepsilon_{t-r}) = 0, r = 0, \ldots, q-1$, which implies that $E(u) = 0$. In addition, the magnitude of $G^q y_{t-q}$ becomes small for large q, given the stability restrictions for the spatial dependence parameters ($\rho_o + \rho_d + \rho_w < 1$) and the fact that the matrix W is row-stochastic (has row-sums of unity), since row-stochastic matrices have a principle eigenvalue of one.

The implication of this development is that we can interpret the cross-sectional SAR interaction model as the outcome or expectation of a long-run equilibrium or steady state relationship, which is shown in (11)

$$\lim_{q\to\infty} E(y_t) = (I_{n^2} - G)^{-1} Z\delta \qquad (11)$$
$$= (I_{n^2} - \rho_d W_d - \rho_o W_o - \rho_w W_w)^{-1} Z\delta.$$

This is the expectation for the DGP of the SAR interaction model given in (5).

2. INTERPRETING SPATIAL INTERACTION MODELS

In Section 2, we consider how changes in the characteristics of regions impact flows in the case of the conventional (nonspatial) interaction model from (2). An important point to note is that changes in the rth characteristic of region i, ΔX_i^r will produce changes in flows into region i from (potentially) $(n-1)$ other regions, as well as flows out of region i to (potentially) $(n-1)$ other regions. This can be seen by noting that the matrices $X_d = \iota_n \otimes X$ and $X_o = X \otimes \iota_n$ repeat X_i^r n times. Unlike the situation in conventional regression models where a change ΔX_i^r leads to changes in only observation i of the dependent variable, y_i, we cannot change single elements of X_d^r, X_o^r, nor should we interpret the coefficient estimate $\hat{\beta}_o, \hat{\beta}_d$ as reflecting the impact of this change (averaged over all observations) on a single element of the dependent variable vector y. The fact that changes in characteristics of a single region give rise to numerous responses in the flow matrix rather than changes in a single observation (dyad) of the dependent variable (as in traditional regression) creates a challenge for drawing inferences about the partial derivative impacts of changing regional characteristics on flows. To address this challenge, Section 2 proposes scalar summary measures for the impact of changing regional characteristics on flows, that collapse the many changes in flows to a single number. These scalars average over the many changes that arise in the flow matrix from changing characteristics of the regions, as is typical of the way in which we interpret regression models.

Section 2 considers how changes in the characteristics of regions impact flows in the case of the SAR interaction model from (3). In this model specification, changes in the characteristics of a single region i can impact flows into and out of region i to its $2(n-1)$ dyad (i, j) partners (as described above), but also flows into and out of regions that neighbor the origin i and destination j regions that are not part of the dyad (i, j). This arises from the spatial dependence part of the SAR interaction model. An implication is that we should not interpret the coefficient estimates β_d, β_o as if they were regression estimates that reflect *partial derivative changes* in the dependent variable associated with changes in the explanatory variables. LeSage and Pace (2009) point out that this mistaken approach to simpler cross-sectional SAR models involving only n regions has been taken in much of the past spatial econometrics literature.

In section 2, we present a method that can be used to relate changes in characteristics of a single region i to flows across the $n \times n$ network of flows for the case of the SAR interaction model. This issue has not been tackled in the literature, yet it is essential for interpreting the coefficient estimates β_o, β_d in the SAR interaction model.

A Numerical Illustration for the Nonspatial Gravity Model

Prior to setting forth our method for quantifying how changes in the rth characteristic of region i impacts flows, we provide a simple numerical illustration to fix ideas. Using the DGP in (5), we generated a set of flows using $n = 8$ regions with $\beta_d = 1, \beta_o = -0.5, \delta = -0.5, \rho_d = 0, \rho_o = 0, \rho_w = -\rho_o \times \rho_d = 0$. Note that setting $\rho_d = \rho_o = \rho_W = 0$ produces a spatial interaction model with no spatial dependence between flows. No disturbance term was used, and the single vector $x' = \begin{pmatrix} 40 & 30 & 20 & 10 & 7 & 10 & 15 & 25 \end{pmatrix}$ was used, so we have the case where $r = 1$.

For use later with the model where spatial dependence exists so the parameters ρ_d, ρ_o, and ρ_W take nonzero values, a set of n latitude and longitude coordinates both equal to $(1, 2, \ldots, 8)$ were used. The systematic order of the latitude–longitude coordinates produces regions configured to lie on a line, with a simplified (logged) distance vector g and spatial weight matrix. The spatial weight matrix W implied by this spatial configuration and use of two nearest (distanced) neighbors is shown in (12). Only a single nearest neighbor was used for the first and last observations at the beginning and end of the line. The matrix shows that region 2 has regions 1 and 3 as neighbors, 3 has regions 2 and 4 as the two nearest neighbors, region 4 has regions 3 and 5 as the two neighbors, and so on. This greatly simplifies things relative to real world data. Of course, this matrix does not enter the conventional interaction model by virtue of zero values for the spatial dependence parameters ρ_d, ρ_o, and ρ_W.

$$W = \begin{pmatrix} 0.0 & 1.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.5 & 0.0 & 0.5 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.5 & 0.0 & 0.5 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.5 & 0.0 & 0.5 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.5 & 0.0 & 0.5 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.5 & 0.0 & 0.5 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.5 & 0.0 & 0.5 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 1.0 & 0.0 \end{pmatrix}. \tag{12}$$

For the conventional (nonspatial) model, a discrete change in each element/region $i = 1, \ldots, 8$ of the vector x by one unit was made and the discrete changes arising in the $n \times n$ flow matrix as a result of these perturbations were recorded. For each change in the value x_i for a single region i, a new flow matrix was generated and subtracted from the original flow matrix to illustrate how changes in the characteristics of a single region impact the matrix of flows.

An important point to note here is that unlike the regression model where the matrix X contains characteristics for each of the n regions, the matrices X_o, X_d in the interaction model strategically repeats values of the $n \times r$ matrix X to form $n^2 \times r$ matrices $X_d = \iota_n \otimes X, X_o = X \otimes \iota_n$. An implication is that when we change the characteristic/element of a single region i (which we denote using x_i), this produces a set of changes in n elements of the matrix X_d and changes in n elements of the matrix X_o. Together, this set of $2n$ altered values in the matrices X_d, X_o produce the change in flows that results from changing characteristics of the ith region from x_i to $x_i + 1$. This has computational implications for how we calculate the effects arising from changes in the explanatory variables of this model. Unlike the conventional regression model, we do not need to calculate changes in each of the n^2 elements of the vectors X_o and X_d to produce scalar summary measures of the impact of these changes on the flows. Although this approach is valid, it requires more computational effort. Instead we can consider only n changes in each observation i of the matrix/vector X as producing a total derivative response. There will be a vector of $n^2 \times 1$ responses in the flows (which can be viewed as a change in the $n \times n$ flows matrix Y) arising from a change in a single characteristic of the ith region, x_i. This single element *total derivative change* works through a series of $2n$ associated changes that arise in the $n^2 \times r$ model explanatory variables X_o, X_d.

Intuitively, changing a single region i's characteristic (say increasing jobs available in region i) means: (1) this region will attract inflows of workers as a destination from all n regions (including itself), and (2) will produce decreased outflows of workers to all n regions (including itself). This facet of changes in the characteristic of a single region is what accounts for the model repeating the same altered value of x_i (the new jobs for region i) n times in the vector/matrix X_d, and n times in X_o. Given this, it is computationally inefficient to consider conventional partial derivatives that would independently change each of the n^2 elements in X_d or X_o and examine their impact on the flow matrix. Changes to individual elements of X_d and X_o need not be considered given the structure of the model (and associated *DGP*).

There may be applied modeling situations where different explanatory variables are used to model the origin and destination characteristics of the regions that are thought to be important for explaining variation in flows. In these situations, the argument used above regarding changes to a single explanatory variable x_i for each region will not be valid. The more computationally inefficient approach of using conventional partial derivatives that would independently change each of the n^2 elements in X_d and X_o would be required to examine the impact of these changes on the flow matrix. We discuss this type of situation when providing a numerical illustration.

Results showing the changes in the $n \times n$ flow matrix associated with a change in the third region's characteristic, x_3, by one unit for the case of the independent (nonspatial) gravity model in (2) are shown in (13). These were produced by setting $\rho_o = \rho_d = \rho_w = 0$ in the SAR interaction model from (5), which results in the independent nonspatial model from (2)

$$\Delta Y/\Delta X_3 = \begin{pmatrix} 0.0 & 0.0 & -0.5 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & -0.5 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 1.0 & 1.0 & 0.5 & 1.0 & 1.0 & 1.0 & 1.0 & 1.0 \\ 0.0 & 0.0 & -0.5 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & -0.5 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & -0.5 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & -0.5 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & -0.5 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \end{pmatrix}. \tag{13}$$

The role of the independence assumption is clear in (13), where we see from column 3 that the change of outflows from region 3 to all other regions equals -0.5, which is the value of the coefficient $\beta_o = -0.5$ in our example. Similarly, row 3 exhibits changes in inflows to region 3, taking the value of the coefficient $\beta_d = 1$ in our example. The diagonal (3,3) element reflects a response equal to $\beta_o + \beta_d = 0.5$, the sum of the changes in flows in and out of region 3, reflecting the change in intraregional flows arising from the change in x_3. We have only $2n$ nonzero changes in flows by virtue of the independence assumption. All changes involving flows in and out of regions other than those in the dyads involving region 3 are zero.

Scalar Summary Estimates for the Nonspatial Gravity Model

Our method for producing scalar summary measures of the impacts arising from changes in characteristics of the regions involves averaging over the *cumulative flow impacts* associated with changes in all regions, $i = 1, \dots, n$. Scalars summaries are consistent with how coefficient estimates for the parameters in a conventional regression model are interpreted, and cumulating the impacts makes intuitive sense in our flow setting.

To calculate scalar summaries, note that we can express the partial derivatives for this model as shown in (14), where we record the $n \times n$ matrices of changes in (logged) flows arising from changing variable X_i^r using Y_i.

$$TE = \begin{pmatrix} \partial Y_1/\partial X_1^r \\ \partial Y_2/\partial X_2^r \\ \vdots \\ \partial Y_n/\partial X_n^r \end{pmatrix} = \begin{pmatrix} Jd_1\beta_d^r + Jo_1\beta_o^r \\ Jd_2\beta_d^r + Jo_2\beta_o^r \\ \vdots \\ Jd_n\beta_d^r + Jo_n\beta_o^r \end{pmatrix}. \tag{14}$$

In (14), Jd_i is an $n \times n$ matrix of zeros with the ith row equal to $\iota_n'\beta_d$, and Jo_i is an $n \times n$ matrix of zeros with the ith column equal to $\iota_n\beta_o$. We have n sets of $n \times n$ outcomes (one for each change in $X_i^r, i = 1, \ldots, n$) resulting in an $n^2 \times n$ matrix of partial derivatives reflecting the total effect on flows from changing the rth characteristic of all n regions, hence the label TE.

The total effects on flows can be cumulated and then averaged to produce a scalar summary measure of the total impact of changes in the *typical* region's rth characteristic. This takes the form: $te = (1/n^2)\iota_{n^2}' \cdot TE \cdot \iota_n$, where we use lower case te to represent the scalar summary measure of the $n^2 \times n$ matrix TE. This scalar summary is consistent with the way that regression coefficient estimates are interpreted as averaging over changes in all observations of an explanatory variable.

Empirical applications of the interaction model focus on *origin effects* associated with changes in the rth characteristic of the origin region in the O-D dyad, and on *destination effects* arising from changes in the rth characteristic of the destination region in the (O-D) dyad. We can produce scalar summary measures for these effects along with a scalar representing the impact on intraregional flows, which we label the *intraregional effect*.

A scalar measure of the origin effect can be based on $oe = (1/n^2)\iota_{n^2}' \cdot OE \cdot \iota_n$, where OE is the $n^2 \times n$ matrix defined in (15).

$$OE = \begin{pmatrix} \tilde{J}o_1\beta_o^r \\ \tilde{J}o_2\beta_o^r \\ \vdots \\ \tilde{J}o_n\beta_o^r \end{pmatrix}. \tag{15}$$

In (15), $\tilde{J}o_i$ is the $n \times n$ matrix Jo_i adjusted to have a zero in the i, ith row and column element. This adjustment separates out the intraregional effect, allowing for only $(n-1)$ changes in flows to origins in the O-D dyad, since flows involving cases where region i is both an origin and destination reflect intraregional flows.

Similarly, we can produce a scalar summary measure of *destination effects* using: $de = (1/n^2)\iota_{n^2}' \cdot DE \cdot \iota_n$, where DE is the $n^2 \times n$ matrix defined in (16).

$$DE = \begin{pmatrix} \tilde{J}d_1\beta_d^r \\ \tilde{J}d_2\beta_d^r \\ \vdots \\ \tilde{J}d_n\beta_d^r \end{pmatrix}. \tag{16}$$

The $n \times n$ matrix $\tilde{J}d_i$ is Jd_i adjusted to have a zero in the i, ith row and column element. This adjustment separates out the intraregional effect.

TABLE 1: Scalar Summary Measures of Effects for the Nonspatial Model from a Change in the (Single) rth Characteristic X^r Averaged over All Regions

Origin effects	−0.4375	$\beta_o = -0.5$
Destination effects	0.8750	$\beta_d = 1.0$
Intraregional effects	0.0625	
Total effects	0.5000	$\beta_o + \beta_d = 0.5000$

Finally, a scalar summary for the *intraregional effect* can be constructed using: $ie = (1/n^2)\iota'_{n^2} \cdot IE \cdot \iota_n$, where IE is the $n^2 \times n$ matrix defined in (17).

$$IE = \begin{pmatrix} Ji_1(\beta_d^r + \beta_o^r) \\ Ji_2(\beta_d^r + \beta_o^r) \\ \vdots \\ Ji_n(\beta_d^r + \beta_o^r) \end{pmatrix}. \tag{17}$$

The matrix Ji_i is an $n \times n$ matrix of zeros with a one in the i, i row and column position. We could express $\tilde{J}d_i = Jd_i - Ji_i$, and also $\tilde{J}o_i = Jo_i - Ji_i$.

Doing this produces scalar summaries that sum up to the scalar summary total effect. For our numerical example, the effects summaries calculated this way are shown in Table 1. In addition to our proposed scalar summary effects estimates, we present the parameters β_o, β_d whose estimates are typically interpreted as origin and destination effects, and whose sum is considered the total effect arising from a change in the rth explanatory variable.

An important point to note is that this approach differs from the conventional interpretation of nonspatial gravity models where the coefficient β_o is interpreted as a partial derivative reflecting the impact of changes in origin characteristics and β_d that associated with changing destination characteristics. Although the conventional approach that used the coefficient sum $\beta_o + \beta_d$ as a measure of the total effect on flows arising from changes in origin and destination characteristics would produce a correct inference, the appropriate decomposition into origin, destination, and intraregional effects has been missing from this literature.

Another point is that one can use changes in each element of the $n^2 \times 1$ vectors X_o and X_d to arrive at the same scalar summary measures as shown in Table 1. However, this would require that we sequence through changes in n^2 individual elements of X_o and also n^2 elements of X_d, recording the change in the $n \times n$ matrix of flows that arise from this sequence of $2n^2$ changes, which is computationally much more difficult. We would also need to aggregate the changes in flows arising from changes in both X_o and X_d to produce final results. To avoid this, our approach exploits the special structure of the $n^2 \times r$ matrices X_o, X_d as they relate to the underlying $n \times r$ matrix X.

As noted above, there could be applied modeling situations where practitioners choose to include a specific characteristic only in the X_o or X_d vector, but not in both. As an example, consider a model for *commuting-to-work flows*. The number of residents might be used as a size measure for origin regions whereas the number of business establishments might be used as a size measure for the destination regions. In this case, it might be more appropriate for interpretative purposes to report separate scalar effects summaries arising from the calculations involving changing all elements in the vector X_o and X_d. We will have more to say about this in Section 3 where we provide an empirical application.

A Numerical Illustration for the SAR Interaction Model

Using the same numerical values set forth in the previous section, but setting $\rho_o = 0.5$, $\rho_d = 0.4$, and $\rho_w = -\rho_o\rho_d = -0.2$, we carried out the same experiment where each value of $x_i, i = 1, \ldots, 8$ was changed by one unit. The resulting changes in the flow matrix were recorded, with the total flow effects arising from the change in x_3 shown in (18).

$$\frac{\Delta Y}{\Delta X_3} = \begin{pmatrix} 0.052 & -0.086 & -0.777 & -0.069 & 0.121 & 0.171 & 0.185 & 0.187 \\ 0.337 & 0.199 & -0.492 & 0.216 & 0.406 & 0.457 & 0.470 & 0.473 \\ 2.048 & 1.910 & 1.219 & 1.927 & 2.117 & 2.168 & 2.181 & 2.184 \\ 0.318 & 0.180 & -0.511 & 0.197 & 0.387 & 0.438 & 0.451 & 0.454 \\ -0.043 & -0.181 & -0.872 & -0.164 & 0.026 & 0.077 & 0.090 & 0.093 \\ -0.118 & -0.256 & -0.947 & -0.239 & -0.050 & 0.001 & 0.015 & 0.017 \\ -0.134 & -0.272 & -0.963 & -0.255 & -0.065 & -0.014 & -0.001 & 0.002 \\ -0.136 & -0.275 & -0.965 & -0.257 & -0.068 & -0.017 & -0.004 & -0.001 \end{pmatrix}. \tag{18}$$

One difference between this SAR model result and the nonspatial model is the presence of network effects, shown by the nonzero elements in rows and columns other than 3. This means that a change in say the attractiveness of region 3 impacts flows throughout the network. This arises because the SAR model specification allows for dependence of flows neighboring regions A, Z as a dyad, e.g., regions B, C that are neighbors to region A and X, Y who are neighbors to Z.

Of course, the largest impacts still tend to reside in the third row and column, since the change in attractiveness of region 3 has the largest impact on flows involving region 3 in the O-D dyads. The magnitude of impact as we move further from the (3,3) element in the up/down or left/right direction in column and row 3 depends in a nonlinear way on a number of factors. The spatial dependence structure reflected in the matrix W, strength of origin and destination dependence represented by ρ_o, ρ_d as well as O-D dependence ρ_w.

For this example, where regions were arranged in a line, moving to row and column elements further from the 3,3 position should reflect more distant neighbors. This implies an increase in the number of paths through which the flows must pass to reach the (8,8) and (1,1) dyads. We see smaller network effects in the flow matrix for these dyads, which arises from decay with *higher-order neighbors* typical of SAR processes.

Scalar Summary Estimates for the SAR Interaction Model

We can modify the approach set forth in the previous section for calculating scalar summary measures of the impacts arising from changes in characteristics of the regions. The partial derivatives reflecting total effects for this model are shown in (19), where we see that the spatial dependence parameters and weight matrix W come into play.

$$TE = \begin{pmatrix} \partial Y_1/\partial X_1^r \\ \partial Y_2/\partial X_2^r \\ \vdots \\ \partial Y_n/\partial X_n^r \end{pmatrix} = (I_{n^2} - \rho_d W_d - \rho_o W_o - \rho_w W_W)^{-1} \begin{pmatrix} Jd_1\beta_d^r + Jo_1\beta_o^r \\ Jd_2\beta_d^r + Jo_2\beta_o^r \\ \vdots \\ Jd_n\beta_d^r + Jo_n\beta_o^r \end{pmatrix}. \tag{19}$$

Scalar summary measures of the (average) origin, destination, and intraregional effects can proceed in the obvious way, while the scalar summary for (cumulated) network effects (ne) can be calculated as: $ne \equiv te - de - oe - ie$. We follow LeSage and Pace (2009) and cumulate these spatial spillover impacts falling on all regions not involved in the O-D dyads, then average over all observations. Of course, the network effects could also

TABLE 2: Scalar Summary Measures of Effects for the Spatial Autoregressive Interaction Model Arising from a Change in a Single Characteristic X^r Averaged over All Regions

	Spatial	Nonspatial
Origin effects	−0.6217	−0.4375
Destination effects	1.2178	0.8750
Intraregional effects	0.0636	0.0625
Network effects	1.0560	
Total effects	1.7158	0.5000

be calculated using matrices: $Jn_i = J_i - Jd_i - Jo_i$, where J_i is an $n \times n$ matrix of ones, and Jd_i, Jo_i are defined as previously, resulting in: $ie = (1/n^2)\iota_{n^2}' \cdot NE \cdot \iota_n$, where NE is an $n^2 \times n$ matrix analogous to those previously defined.

We used this approach for our example to produce scalar summary measures of the total effects, as well as the decomposition into origin, destination, intraregional, and network effects. These are shown in the first column of Table 2.

The second column of the table reproduces the scalar summaries from the nonspatial gravity model for contrast. We note that in practice SAR and nonspatial models are not likely to produce the same parameter estimates for β_d, β_o as was assumed here. Nonetheless, it should be clear that allowing for spatial dependence in flows produces larger impacts, ceteris paribus. Origin and destination effects are enhanced by feedback effects arising from changes in network flows, those associated with regions neighboring each O-D dyad. Network effects are zero by assumption in the nonspatial gravity model.

One could calculate equivalent flow matrix responses to changes in each element of the $n^2 \times 1$ vectors X_o, X_d, but this is computationally inefficient and ignores the special structure of the model specification.

Consider again the example involving commuting-to-work flows, where the number of residents is used as a size measure for origin regions and the number of business establishments as a size measure for the destination regions, so X_o and X_d are distinct. This type of specification would lead to a slight change in interpretation, where changes in X_o (residents at the origin) lead to origin, destination, intraregional, network, and total effects on flows, as do changes in X_d (business establishments at the destination). This type of model specification could be viewed as an *a priori* zero restriction on the coefficient for the characteristic residents at the *destination* as well as a zero restriction on the coefficient for business establishments at the *origin*. It should be possible to include the full set of explanatory variables (residents and business establishments) in the set of model characteristics for both origins and destinations, and then test the validity of the *a priori* zero restrictions. This would involve a test for significant differences between the full and nested model scalar summary effects estimates. If there are no differences in conclusions regarding the size and significance of the scalar summaries, then the restrictions are consistent with the sample data.

Empirical Measures of Dispersion for the Scalar Summary Effects Estimates

As with any estimates we would like a measure of dispersion for the scalar summary point estimates proposed in the previous two sections. These can be constructed by simulating the expressions set forth with draws for the parameters that come from either the posterior distribution based on Bayesian MCMC estimation of the model, or draws simulated using maximum likelihood estimates and the associated variance-covariance matrix for the parameters. LeSage and Pace (2009, chapter 8) describe both maximum likelihood as well as Bayesian MCMC estimation methods for the SAR interaction model.

Processing a sample of, say 1,000 draws, using the $n^2 \times n$ matrix expressions from (14) through (17) requires premultiplication by the $n^2 \times n^2$ matrix inverse: $(I_{n^2} - \rho_d W_d - \rho_o W_o - \rho_w W_W)^{-1}$, for the case of the SAR interaction model. This requires inversion of the $n^2 \times n^2$ matrix for each set of (1,000) drawn values for ρ_d, ρ_o, ρ_w.

The matrices Jd_i, Jo_i, Ji_i can all be calculated once prior to processing the sample of draws, since they contain nonstochastic elements. These $n^2 \times n$ matrices would be multiplied by scalar draws for the parameters β_d, β_o on each of the (1,000) passes, which is relatively fast.

This makes the discussion of computational efficiency quite relevant. In most applications that would rely on the SAR interaction model in conjunction with maximum likelihood or Bayesian estimation, the sample size is likely to be relatively small. This is because excessive zero values begin to appear in flow matrices based on regions constructed using finer spatial scales. For example, migration flows between the 3,109 U.S. counties in the lower 48 states result in a flow matrix that contains over 90 percent zero values. This type of situation would not be consistent with normality of the dependent variable needed for maximum likelihood or Bayesian estimation. It would likely require a Poisson or negative Binomial estimation procedure (Gourieroux, Monfort, and Trognon, 1984), and these have not been developed for the SAR interaction model on which we focus here. However, recent work by Lambert, Brown, and Florax (2010) proposes a Poisson estimation procedure for the cross-sectional SAR model.

For the application described in the next section, we used a 60×60 flow matrix that required less than two seconds time per draw, or around 30 minutes to process 1,000 draws on a laptop computer. The application took advantage of sparse matrix inversion, since that matrix $(I_{n^2} - \rho_d W_d - \rho_o W_o - \rho_w W_W)^{-1}$ contains many zero values.

3. AN EMPIRICAL ILLUSTRATION

We illustrate calculation and interpretation of the scalar summary measures using commuting-to-work flows for 60 regions (Quartiers) in Toulouse France, taken from the 1999 census from The National Institute for Statistics and Economic Studies, France. Flows were constructed from the census home and work addresses provided by the actively employed population. These were aggregated from individual level information to the regional level. Workers in the defense sector and workers moving to variable sites or working at home were excluded from the individuals used in aggregating to the regional level. The distance matrix was formed using distance between centroids of the spatial units.

For simplicity, we use only two explanatory variables (plus distance): Workers residing in each district and employment (jobs) located in each district. Workers are defined as residents who are working, and thus commuting to work on a regular basis. We use the conventional log transformation of flows as well as the two explanatory variables. Given the log–log specification where effects estimates can be interpreted as elasticities, we would expect *a priori* that flows are directly proportional to workers located at the origin and jobs located at the destination, having effects estimates near unity. It also seems reasonable that effects associated with workers located at employment districts (destinations), and jobs located in residential districts (origins) should have estimates close to zero. Intuitively, workers residing in the same district as their place of employment should have no impact on interregional flows.

For the 60 region area studied, 52 percent of workers in the region come from outside the region and 19 percent of residents in the region work outside. These were excluded

TABLE 3: Coefficient Estimates for the Least-Squares and Spatial Autoregressive Interaction Models

	Least-Squares		
Variable	Coefficient	t-Statistic	p-Level
Constant	−9.4337	−40.7805	0.0000
β_d residents	−0.1330	−6.7896	0.0000
β_d jobs	1.0460	66.2664	0.0000
β_o residents	0.8767	44.7489	0.0000
β_o jobs	−0.0335	−2.1259	0.0335
log(distance)	−0.1372	−23.1712	0.0000

	Spatial Autoregressive		
Variable	Lower 0.05	Coefficient	Upper 0.95
Constant	−5.4366	−3.8507	−3.0349
β_d residents	−0.1000	−0.0493	0.0010
β_d jobs	0.1986	0.2610	0.4179
β_o residents	0.2701	0.3391	0.4308
β_o jobs	−0.0535	−0.0139	0.0261
log(distance)	−0.0314	−0.0164	−0.0017
ρ_d	0.5324	0.6208	0.6577
ρ_o	0.4695	0.6661	0.7114
ρ_w	−0.3751	−0.3013	−0.0021

to form a system of flows between districts that includes only persons who both live and work in one of the 60 regions.

There were 15 percent of the 3,600 flows with zero values. Ranjan and Tobias (2007) motivate zero flows as arising from specific O-D dyads where transportation and other costs associated with trade exceed a threshold making trade unprofitable. They treat zero flows using a Tobit model, noting that our use of $\ln(1 + y)$ as the dependent variable ignores the mixed discrete/continuous nature of flows and arbitrarily adds unity to the dependent variable to avoid taking the log of zero. They describe a Bayesian MCMC sampling approach to implementing Tobit procedures, which have been extended to the case of SAR interaction models in LeSage and Pace (2009, chapters 8 and 10). Since our application is merely illustrative with a focus on interpretation issues that arise given estimates of the underlying model parameters, we did not do this.

The restriction implied by the O-D dependence filter view of the SAR variant of the model was not imposed, meaning a parameter ρ_w was estimated.

The matrix W was constructed using the six nearest neighboring regions, which was the average number of contiguous neighbors for the 60 regions. The estimates and inferences are insensitive to reasonable changes in the number of neighbors, or use of a first-order contiguity weight matrix. (See LeSage and Pace (2010) for an exploration of the sensitivity of inferences from SAR models to changes in the weight matrix used.)

Least-squares and SAR interaction model estimates are shown in Table 3 for the model parameters, not the scalar summary effects estimates. The Bayesian MCMC SAR estimates were based on 5,500 draws with 2,500 omitted for start-up. Lower and upper 0.05 and 0.95 credible intervals are reported based on the retained 3,000 MCMC draws.

The R-squared statistic for the regression model was 0.6774, and the distance variable has a negative and significant coefficient as expected. We note that the coefficients and t-statistics reported in the table should not be interpreted as reflecting the partial

TABLE 4: Effects Estimates for the Nonspatial and Spatial Autoregressive Interaction Models

	Least-Squares			
Variables	Lower 0.05	Mean	Median	Upper 0.95
Origin residents	0.8391	0.8700	0.8700	0.9005
Origin jobs	−0.0585	−0.0326	−0.0318	−0.0072
Destination residents	−0.1665	−0.1349	−0.1353	−0.1036
Destination jobs	1.0273	1.0539	1.0538	1.0784
Intraregional residents	0.0117	0.0125	0.0125	0.0132
Intraregional jobs	0.0167	0.0173	0.0173	0.0179
Total residents	0.7020	0.7476	0.7471	0.7945
Total jobs	1.0042	1.0386	1.0390	1.0741

	Spatial Autoregressive			
Variables	lower 0.05	mean	median	upper 0.95
Origin residents	1.1193	1.4036	1.4006	1.6893
Origin jobs	0.1710	0.4068	0.3988	0.6668
Destination residents	−0.0321	0.1763	0.1728	0.3952
Destination jobs	0.6692	0.8743	0.8677	1.1149
Intraregional residents	0.0167	0.0224	0.0224	0.0280
Intraregional jobs	0.0094	0.0141	0.0139	0.0193
Network residents	7.5857	15.2276	14.8881	26.0464
Network jobs	18.0505	26.6604	25.9431	39.1400
Total residents	8.7048	16.8298	16.4664	26.0464
Total jobs	18.9739	27.9557	27.2426	39.1400

derivative impact of changes in origin and destination characteristics, as is typical in applied practice. We need to calculate the scalar summary measures of origin, destination, and intraregional effects to draw valid inferences regarding how changes in regional characteristics impact flows.

Estimates for the SAR interaction model point to a high level of origin- and destination-based spatial dependence. The coefficients ρ_d, ρ_o were 0.62, and 0.66, respectively, and the 95 percent credible intervals suggest they are different from zero and were estimated with a great deal of precision. The estimated dependence parameter $\rho_w = -0.30$, which is not near the restricted value $-\rho_d\rho_o = 0.40$ suggested by the filter interpretation. The coefficient for the distance variable is negative and much closer to zero (based on the 0.05 and 0.95 credible intervals) than the least squares estimate. This type of result where the importance of distance diminishes after taking spatial dependence into account often occurs for spatial variants of gravity models (see LeSage et al., 2007). It suggests that distance may be acting as a proxy for spatial dependence, or that spatial dependence plays a role similar to that of distance in nonspatial models.

The other estimates are not directly comparable to those from least squares, which is the motivation for calculating scalar summary effects estimates. The effects estimates for the nonspatial interaction model should be comparable to those from the SAR interaction model, since they both measure partial derivative impacts associated with changes in regional characteristics on flows.

Effects estimates for the conventional regression and SAR models are reported in Table 4. The nonspatial regression estimates indicate that a 1 percent increase in (working) residents at the *typical* origin region would lead to a 0.87 percent increase in commuter outflows, while a 1 percent increase in jobs at the typical destination would lead to a 1.05 percent increase in commuting inflows. Both of these estimates are close to unity as we would expect. The impact of increasing jobs at the origin (where residents are living) by 1 percent would be to decrease commuting inflows by 0.032 percent, and a 1 percent increase in residents at the destination (living in the region where their jobs are located) would decrease commuting outflows by -0.134 percent.

The intraregional impacts of increasing residents or jobs in the typical region are positive and small (0.0125, 0.0173, respectively), but significantly different from zero based on the 0.05 and 0.95 credible intervals. This suggests that increasing residents or jobs leads to more commuting within the region (commuting to work trips that start and end in the same region). The small magnitude of these estimates suggests that only a small number of workers/jobs are located in the same region as their jobs/workers.

Total effects reflect the sum of the origin, destination, and intraregional effects/impacts on commuting to work flows in and out of regions as well as within regions arising from changes in working residents (*total residents effect*) and jobs (*total jobs effect*). The total residents effect (that of increasing working residents) by 1 percent is an increase in commuting to work flows (in and out of regions as well as within regions) by around 0.75 percent. The total effect of increasing jobs by 1 percent is a 1.038 percent increase in commuting to work flows (in and out of regions as well as within regions).

An interesting point is that the scalar effects estimates for the origin effects are close to $\hat{\beta}_o$ for least squares presented in Table 3, as are the destination effects and $\hat{\beta}_d$. This suggests that the common practice of interpreting conventional regression coefficients *as if* they represent the partial derivative impact of changes in origin and destination characteristics on flows would not lead to erroneous inferences in this example. One reason for this result is that intraregional effects take small values in this application, and these are related to the bias that arises when treating $\hat{\beta}_o, \hat{\beta}_d$ as partial derivatives.

It should be emphasized that the nonspatial regression model (and the resulting effects estimates) assumes that flows are not systematically related to flows to or from nearby regions. Expressed differently, the nonspatial model assumes the endogenous interaction parameters ρ_d, ρ_o, ρ_w take zero values. Of course, this is inconsistent with the nonzero estimates for these parameters found for the SAR interaction model.

Turning attention to the effects estimates for the SAR interaction model, we see origin effects for residents and destination effects for jobs with values of 1.40 and 0.87, near unity, as in the case of least squares. However, for this model a 1 percent increase in residents produces a 1.4 percent increase in outflows, and a 1 percent increase in jobs leads to a 0.87 increase of inflows. The spatial model attributes a greater (relative) impact to origin residents and a lesser impact to destination jobs.

The origin effect of jobs is positive, with a 1 percent increase in jobs located in residential regions (origins) leading to a 0.40 percent increase in commuting-to-work outflows from the region. This is a contrast with the nonspatial model estimate which is negative and significant. We note that least squares estimates for the nonspatial model can be shown to be biased and inconsistent in the presence of spatial dependence, so this difference in inference from the two models likely reflects bias in the nonspatial model.

The destination effect of residents is not significantly different from zero based on the 0.05 and 0.95 credible intervals, implying that an increase in residents in regions where their jobs are located would not impact commuting-to-work flows. It seems intuitively plausible that locating residences in places of their employment should not create commuting flows outside the region. Again, this contrasts with the negative and significant effect estimate from the nonspatial model.

Intraregional effects are positive and small, but different from zero for both variables, consistent with the nonspatial model estimates. This suggests that increasing residents or jobs would produce a small increase in flows within regions. This is consistent with the notion that few residents live and work within the same region.

The big difference between the nonspatial and spatial models resides in the network effects, which cumulate the impact on commuting to work flows across all regions that would arise from increasing residents or jobs (on average across all regions). The strictest interpretation of the network effects would rely on our theoretical development in Section 1, which relates changes at one point in time to the changes in flows cumulated over the network needed to bring about a new steady-state equilibrium at some future time period.

From the perspective of a transportation planner, cumulating the network effects would be the most useful measure of how expected traffic flows in the region would respond over time to changes in residents or jobs.

Given the DGP for the SAR interaction model, expressed as in (20), the (expected) outcome from a change in say residents living in a typical region takes a form involving the infinite series in (21), where ΔX^r represents a change in the number of residents and β_d^r, β_o^r are the coefficients associated with destination and origin variables: $X_d^r = \iota_n \otimes X^r$, $X_o^r = X^r \otimes \iota_n$.

(20)
$$\begin{aligned}\lim_{q\to\infty} E(y_t) &= (I_{n^2} - G)^{-1} Z\delta \\ &= (I_{n^2} + G + G^2 + G^3 + \ldots)Z\delta\end{aligned}$$

(21)
$$\lim_{q\to\infty} E(\Delta y_{t+q}) = (I_{n^2} + G + G^2 + G^3 + \ldots)(\iota_n \otimes \Delta X^r \beta_d^r + \Delta X^r \otimes \iota_n \beta_o^r).$$

The term $G = \rho_d W_d + \rho_o W_o + \rho_w W_w$ works on the change in residents to produce impacts on neighbors to the destination, origin, and neighbors to both, while the power G^2 reflects diffusion of the impact on flows (from the change in residents) to regions that are neighbors to these neighbors. G^3 reflects further diffusion from the change in residents to neighbors to these neighbors of the neighbors, and so on for higher powers of G. Since the coefficients $\rho_d + \rho_o + \rho_w < 1$, the impact of the change in residents becomes smaller as we move to higher order neighbors reflected in higher powers of G. That is, the influence of increasing residents on commuting-to-work flows falls most heavily on regions identified by lower orders of G and decays in magnitude as we move to regions identified by higher order powers of G. Nonetheless, our scalar summary measure of network effects cumulates these diffusive changes required to reach the new steady-state over all regions. Of course the accumulation process assigns the appropriate smaller weight to regions associated with higher order powers of G during the process. Since we are working with cross-sectional sample data, there is no explicit role for time in our interpretation of the results. That is, we have the conventional comparative static experiment, where we can calculate changes needed to arrive at a new steady-state equilibrium, but not how long this would take.

If the spatial configuration of our observations was expressed as a network (which is not the case here), then G would represent first-order nodes having a direct path/segment connection to the region whose residents changed, and G^2 would reflect network nodes connected to the region in question that require passing through two paths/segments in the network, with higher powers associated with an increasing number of paths between the region in question and those impacted by the diffusion.

A 1 percent increase in residents (in the typical region) would lead to a 15.2 percent (cumulative) increase in commuting to work flows throughout the network of regions (excluding flows involving the O-D dyads that are already captured by the origin and destination effects and within region flows associated with intraregional effects). This might seem surprisingly large, given that spillover or diffusion effects might be thought of as "second-order" impacts. One might suppose that origin and destination effects directly involving O-D dyad regions would be "first-order" effects having much larger magnitudes than the "second-order" network effects. The explanation for the large size of the cumulative network effects estimates lies in the cumulative nature of this scalar summary measure. This means we are changing the focus from impact on a "typical single region" to cumulative impacts on the network which fall on many regions, so the scale of these impacts is different.

To see this, consider that we used six nearest neighboring regions when forming the matrix W, so that $W_d = I_n \otimes W$ contains six neighboring regions (nonzero elements in each of the 3,600 rows) as does $W_o = W \otimes I_n$. The matrix $W_w = W \otimes W$ contains 36 neighbors in each row. The second-order powers of W_d and W_o contain 36 neighbors (nonzero elements in each row), and the second-order power of W_w contains 1,296 nonzero elements in each of the 3,600 rows. Higher order powers of these matrices become even denser with more and more zero elements being filled-in by nonzero elements.

Of course, most impacts of any magnitude will fall on regions associated with low orders of G, with regions identified by higher order G impacted far less because of the decay of influence arising as we move to higher powers of ρ_d, ρ_o, ρ_w.

One might suppose that the large number of neighbors that come into play for the SAR interaction model specification is problematical. After all if W_w^2 contains 1,296 nonzero elements in each of the 3,600 rows, higher order powers must surely produce so much "fill-in" that these represent situations where we have reached the boundaries of the 60 region Toulouse sample, and effects start bouncing off borders, moving back into interior regions. In fact, the nature of decay of influence with powers of the matrix inverse $(I_n^2 - G)^{-1}$ in this example is such that the infinite series expansion represented by the inverse contains only 13.66 percent of the 12.96 million elements that are greater than 0.01 in value. So, the increase in substantial network flows are relatively localized to regions associated with low orders of G.

The implication of this type of interpretation is that a 1 percent increase in residents in the typical region will produce a long-run response of 15 percent in commuting-to-work flows cumulated over the network of 60 regions. However, in practice the 15 percent increase in cumulated commuting-to-work flows would be located mostly in the 6 or 36 regions that neighbor a typical region. This increase would of course come about after enough time elapsed to produce a new steady-state equilibrium between locations of working resident commuters and jobs/employers.

The network effect for a 1 percent increase in jobs in the typical region leads to a 26.66 percent cumulative increase in commuting to work flows throughout the region (excluding flows involving O-D dyads already captured by the origin, destination, and intraregional effects). Again, the largest magnitude of impact would fall on regions nearby the region where jobs increased, and this scalar summary is interpreted to be a cumulated average over all regions/observations. This means we could view the 26.6 percent increase as the percentage of (cumulated) flows that would be observed in 6 or 36 neighbors to the typical region in the new steady-state equilibrium.

Spatial clustering of residents around locations with road and metro access points that lead to similar destinations, versus clustering of jobs in the central city, might explain the inferred difference in the network residents effect and the network jobs effect. Residents have an easier task when it comes to optimizing their location to minimize commuting to work, leading to a smaller network residence effect. An employer's objective function likely pertains more to location decisions that optimize customer patronage and facilitate face-to-face interactions with other firms, producing a larger network jobs effect.

4. CONCLUSION

SAR specifications for the spatial interaction model along with maximum likelihood and Bayesian estimation methods set forth by LeSage and Pace (2008, 2009) hold a great deal of promise for regional modeling of flows. Beyond the issue of estimating model parameters, there is also a need to carefully consider how the parameter estimates are interpreted.

In the case of the independent nonspatial interaction model, changes in characteristics of a single region can exert impacts on: inflows from $n-1$ other regions, outflows to $n-1$ regions, as well as intraregional flows. These impacts can be measured by considering the changes that arise in appropriate rows and columns of the flow matrix from changes in characteristics of each region. We propose scalar summary measures of these impacts that average over changes applied to a single explanatory variable (regional characteristic) for all regions, analogous to how regression model estimates are interpreted. The proposed approach allows separation of impacts by row/column and diagonal matrix elements, which we label: origin, destination, and intraregional effects. Past applications of regression-based spatial interaction models interpret the coefficient estimates associated with origin and destination characteristics *as if* they represent partial derivative impacts arising from changes in regional characteristics. This assumes that one can change characteristics of origin regions while holding constant characteristics of destination regions involved in the O-D dyad which forms the basis for nonspatial gravity models. We derive the correct partial derivative expressions that take into account that changing characteristics of a single region will have an impact on flows involving this region in multiple O-D dyads, where the region appears as either an origin, destination, or both.

For the case of the SAR interaction model, interpretation of the model estimates in terms of their partial derivative impacts on flows is more complicated. Changes to a single region's characteristics can impact not only inflows from $n-1$ other regions, outflows to $n-1$ other regions, and intraregional flows, but also other flows that occur throughout the network of regions. This is because spatial dependence allows for regions not directly involved as origins or destinations associated with a single i, j dyad to come into play. These impacts can be measured using changes taking place in rows, columns, the diagonal and off-diagonal elements of the flow matrix as a result of a change to a single region's characteristic. We propose a scheme for calculating scalar summary measures for these impacts that we label: origin, destination, intraregional, and network effects, which was derived from the partial derivative expressions arising from this type of model specification.

Specifics regarding simulation of the partial derivative impact estimates based on the estimated distribution for the model parameters was also discussed. In cases where least squares or maximum likelihood SAR estimation procedures are used, the point estimates and estimated variance-covariance matrix for the model parameters can be used to generate draws for each model parameter. These are then used in conjunction with the approach proposed here to produce a distribution of the scalar estimates for the various types of impacts. These empirically derived distributions serve as the basis for inference regarding significance of the various types of impacts. Bayesian MCMC estimation based on uninformative priors can be used to produce estimates equivalent to least squares or maximum likelihood spatial procedures, and these provide a set of draws as a by-product of model estimation that can be used in the approach proposed here for calculating scalar summary measures of dispersion for the various types of impacts.

An illustration of our interpretative methods was provided that examined commuting to work flows for 60 regions in the Toulouse France. Estimates and inference from a nonspatial model were compared to those from the SAR interaction model. Bias in the nonspatial model estimates was shown to lead to different inferences, and the nonspatial model vastly underestimates the total impact of increasing residents or jobs on commuting flows. This underestimate arises from ignoring network (or spatial spillover) effects that arise when flow dyads exhibit spatial dependence on flows from nearby regions. A theoretical economic motivation for this type of dependence was provided.

Interpretation of scalar summary measures of origin, destination, and intraregional effects based on the model partial derivatives seems relatively straightforward and intuitive. On the other hand, interpreting the cumulative network effects scalar summary is a bit more complicated. This seems a good area for future research.

REFERENCES

Anderson, James E. and Eric van Wincoop. 2003. “Gravity with Gravitas: A Solution to the Border Puzzle,” *American Economic Review*, 93(1), 170–192.

Behrens, Kristen, Cem Ertur, and Wilfried Koch. 2012. “‘Dual’ gravity: Using Spatial Econometrics to Control for Multilateral Resistance,” *Journal of Applied Econometrics*, 27(5), 773–794.

Bolduc, Denis, Richard Laferriere, and Gino Santarossa. 1992. “Spatial Autoregressive Error Components in Travel Flow Models,” *Regional Science and Urban Economics*, 22(3), 371–385.

Fischer, Manfred M. and Daniel A. Griffith. 2008. “Modeling Spatial Autocorrelation in Spatial Interaction Data: An Application to Patent Citation Data in the European Union,”*Journal of Regional Science*, 48(5), 969–989.

Gourieroux, Christian, Alain Monfort, and Alain Trognon. 1984. “Pseudo Maximum Likelihood Methods: Applications to Poisson Models,” *Econometrica*, 52(3), 701–720.

Griffith, Daniel A. 2007. “Spatial Structure and Spatial Interaction: 25 Years Later,” *The Review of Regional Studies*, 37(1), 28–38.

Koch, Wilfried and James P. LeSage. 2012. “Latent Multilateral Trade Resistance Indices: Theory and Evidence,” *Scottish Journal of Political Economy*. Forthcoming.

Lambert, Dayton M., Jason P. Brown, and Raymond J. G. M. Florax. 2010. “A Two-Step Estimator for a Spatial Lag Model of Counts: Theory, Small Sample Performance and an Application,” *Regional Science and Urban Economics*, 40(4), 241–252.

LeSage, James P. and Manfred M. Fischer. 2010. “Spatial Econometric Modeling of Origin-Destination Flows,” in M. M. Fischer and A. Getis (eds.), *Handbook of Applied Spatial Analysis*. Berlin, Heidelberg and New York: Springer, pp. 409–433.

LeSage, James P. and Ronald K. Pace. 2008. “Spatial Econometric Modeling of Origin-Destination Flows,” *Journal of Regional Science*, 48(5), 941–967.

———. 2009. *Introduction to Spatial Econometrics*, Boca Raton, FL: CRC Press, Taylor & Francis Group.

———. 2010. “The Biggest Myth in Spatial Econometrics,” SSRN Working Paper: papers.ssrn.com/sol3/papers.cfm?abstract_id=1725503

LeSage, James P., Manfred M. Fischer, and Thomas Scherngell. 2007. “Knowledge Spillovers across Europe, Evidence from a Poisson Spatial Interaction Model with Spatial Effects,” *Papers in Regional Science*, 86(3), 393–421.

LeSage, James P. and Carlos Llano-Verduras. 2014. “Forecasting Spatially Dependent Origin and Destination Commodity Flows,” *Empirical Economics*. Forthcoming.

Ranjan, Priya and Justin L. Tobias. 2007. “Bayesian Inference for the Gravity Model,” *Journal of Applied Econometrics*, 22(4), 817–838.

Sen, Ashish K. and Tony E. Smith. 1995. *Gravity Models of Spatial Interaction Behavior*, Heidelberg: Springer-Verlag.

Smith, Tony E. 1975. “A Choice Theory of Spatial Interaction,” *Regional Science and Urban Economics*, 5(2), 137–176.