The spatial autocorrelation problem in spatial interaction modelling: a comparison of two common solutions

Griffith, Daniel A. and Fischer, Manfred M. and LeSage, James P.

2016

THE SPATIAL AUTOCORRELATION PROBLEM IN SPATIAL INTERACTION MODELLING: A COMPARISON OF TWO COMMON SOLUTIONS

Daniel A. Griffith (University of Texas at Dallas)
Manfred M. Fischer (Vienna University of Economics and Business
James P. LeSage (Texas State University)

Abstract Spatial interaction models of the gravity type are widely used to describe origin-destination flows. They draw attention to three types of variables to explain variation in spatial interactions across geographic space: variables that characterize the origin region of interaction, variables that characterize the destination region of interaction, and variables that measure the separation between origin and destination regions. A violation of standard minimal assumptions for least squares estimation may be associated with two problems: spatial autocorrelation within the residuals, and spatial autocorrelation within explanatory variables. This paper compares a spatial econometric solution with the spatial statistical Moran eigenvector spatial filtering solution to accounting for spatial autocorrelation within model residuals. An example using patent citation data that capture knowledge flows across 257 European regions serves to illustrate the application of the two approaches.

1. Introduction

Spatial autocorrelation complicates the treatment of spatial interaction data (Griffith, 2007). Several spatial scientists address these complications by capturing spatial autocorrelation in the specification of spatial interaction models (e.g., Fischer and Griffith, 2008; Griffith and Chun, 2015; LeSage and Pace, 2008). Three features of these specifications are important: (i) the probability model employed; (ii) the term(s) included to account for spatial autocorrelation; and, (iii) the statistical reasoning perspective. The two most common probability models used are the log-normal and the Poisson. The former is a normal approximation for which zero flows become problematic. The latter deals directly with flows as counts (Flowerdew and Aitkin, 1982), and uses an offset variable in some of its implementations. Terms introduced into a specification to account for spatial autocorrelation include balancing factors, fixed effects, and random effects. Griffith and Fischer (2013) demonstrate equivalencies between these latter three conceptualizations.

What remains to explore is comparing the frequentist versus Bayesian approaches to accounting for spatial autocorrelation in spatial interaction models. In other words, are the spatial interaction parameters fixed but unknown, while the data are random [the probability of interest is $\Pr(\text{data}|H_0)$]? Or, are the parameters random while the data are fixed [the probability of interest is $\Pr(H_0|\text{data})$]? This paper summarizes comparisons between these two approaches. A principal motivation behind the study summarized here is to see which of the two models produces better estimates and inferences (particularly predictions) for empirical (in this case, actual knowledge flows) sample data. Because we do not know the true underlying data generating process (DGP) for these flows, we draw conclusions regarding which of the two model specifications is more consistent with the underlying empirical data based on estimation and prediction performance. Because the two model specifications are different, we address the question of whether the different aspects of the specifications make a material difference in estimates and description of the sample data. We thought that answering this question would prove useful to researchers and practitioners when they make a model selection. We view this as more valuable than conducting a simulation experiment[1], which requires knowledge of the DGP coupled with true parameters.[2]

2. Data description

We use patent citation data to compare spatial filter and Bayesian estimates and inferences for the Poisson regression model containing origin- and destination-specific random effects that follow a spatial autoregressive process. The data relate to citations between European high-technology patents, which are defined here to be patent applications at the European Patent Office assigned to high-technology firms located in Europe. Here high-technology includes the International Standard Industrial Classification (ISIC) sectors of aerospace (ISIC 3845), electronics-telecommunication (ISIC 3832), computers and office equipment (ISIC 3825), and pharmaceuticals (ISIC 3522). Self-citations (i.e., citations from patents assigned to the same firm) have been excluded, given our interest in pure externalities as evidenced by inter-firm knowledge spillovers. The observation of citations is subject to a truncation bias, a well-known phenomenon, because we observe citations for only a portion of the life of an invention. To avoid this bias in this analysis, we have established a five-year moving window (i.e., 1985-1989, 1986-1990, …, 1993-1997) to count citations to a patent.[3] The observation period is 1985-1997 with respect to cited patents, and 1990-2002 with respect to citing patents. The sample used in this analysis is restricted to inventors located in n = 257, generally NUTS-2[4], regions covering the EU-27 member states (excluding Cyprus), Norway, and Switzerland, resulting in N = 66,049 inter- and intra-regional flows. In the case of cross-regional inventor teams, we have used the procedure of multiple full counting, which, unlike fractional counting, does justice to the true integer nature of patent citations, and gives the inter-regional cooperative inventions greater weight.

The explanatory variables in our model contain the (logarithm of) stock of patents (x_d) in a knowledge producing region in the time period 1985-1997, the (logarithm of) stock of patents (x_o) in a knowledge receiving region in the time period 1990-2002, and the following three separation variables: geographical distance, a measure of technological proximity between origin and destination regions, and a binary indicator variable for the presence of a national border intervening between a pair of origin and destination regions. Geographical distance was measured in terms of great circle distance (in 1,000 km) between regions' economic centers, and

[1] Both estimation procedures are numerically intensive, making execution of a sufficient number of replications to take advantage of the Law of Large Numbers (e.g., 10,000) impractical.
[2] The two models we compare assume a different DGP for the underlying empirical (i.e., generation of knowledge flows) data. If, for example, we assume a DGP consistent with the eigenvector spatial filter specification, the results of a Monte Carlo study would, of course, show that model to produce superior estimates and inferences. In contrast, if we conduct a Monte Carlo experiment that assumes the Bayesian model DGP, of course, that model would produce superior estimates and inferences. We see little value in showing that each specification is superior when its DGP generates flows. Already existing Monte Carlo studies of both methods (i.e., spatial filtering and Bayesian spatially structured effects) show good performance when the true DGP is consistent with the specific model specifications. Another such study seems unneeded and unnecessary.
[3] Although the five-year horizon appears to be short, it does capture a significant amount of a typical patent's citation life. Note that the mean citation lag of all high-technology patent citations in 1985-2002 is 4.62 years.
[4] NUTS is an acronym of the French for the "nomenclature of the territorial units for statistics," which is a hierarchical system of regions used by the statistical office of the European Community for the produciton of regional statistics. At the top of the hierarchy are the NUTS-0 regions representing countries, with NUTS-1 regions below, representing regions within countries, and then NUTS-2 regions reflecting subdivisions of NUTS-1 regions.

technological proximity was measured in terms of an index defined by Maurseth and Verspagen (2002). The latter captures the extent to which the technological patenting structure of region r is likely to be cited by region s, given the technological profile of s and the technological citation linkages.

3. Bayesian prerequisites

LeSage generously furnished output based upon a Bayesian hierarchical Poisson regression model introduced in LeSage, Fischer and Scherngell (2007). This specification includes latent random effects that are structured to follow a spatial autoregressive process that allows spatial dependence to be modeled. As such, it addresses overdispersion arising from omitted covariates, and avoids the traditional practice of assuming a zero mean normally distributed effect vector, or of using fixed effects parameters. It utilizes formal matrix algebra notation LeSage and Pace (2008; Chapter 8, 2009) provide that allows gravity (or spatial interaction) models to be easily specified using spatial autoregressive processes described in the spatial regression literature.

The hierarchical Poisson regression model containing spatially structured random effects utilizes work by Frühwirth-Schnatter and Wagner (2004) and allows Bayesian Markov chain Monte Carlo (MCMC) techniques to be applied to these models based on data augmentation. The strategic introduction of two sequences of artificially missing data results in a partially Gaussian regression model, and allows MCMC sampling from conditional distributions for the parameters that belong to standard distribution families. This is in stark contrast to past MCMC estimation approaches for Poisson regression models that relied heavily on Metropolis-Hastings sampling from non-standard conditional distributions, and all of the incumbent difficulties involving suitable proposal densities and tuning required for successful Metropolis-Hastings estimation (see Chib et al., 1998).

One drawback of the Frühwirth-Schnatter and Wagner (2004) approach is its requirement to sample two sets of latent parameters whose length equals $\sum_{i=1}^{n^2}(n_i + 1)$, where n_i denotes the count for observation region i, creating considerable computational difficulties.

Suppose the n-by-n matrix **C** denotes the geographic weight matrix portraying the configuration of origins/destinations. Both the origin and destination random effects $\boldsymbol{\theta}_k$ are specified as spatial simultaneous autoregressive (SAR) variates of the matrix form

$$\boldsymbol{\theta}_k = \rho_k \mathbf{W} \boldsymbol{\theta}_k + \mathbf{u}_k,$$

where k denotes origin or destination, matrix **W** is the row-standardized version of matrix **C**, ρ_k is a spatial autocorrelation parameter, and $\mathbf{u}_k \sim \mathrm{N}(\mathbf{0}, \sigma_k^2 \mathbf{I})$, where **I** is the n-by-n identity matrix. This specification is posited as the proper Bayesian prior for the random effects parameters. Meanwhile, uniform priors are attached to the spatial autocorrelation parameters, and diffuse priors are attached to the other model parameters. Specifically, Gamma(a, b) priors are assigned to the precision parameters σ_k^{-2}, with a = b = 0.01, and the regression coefficients for origin and destination characteristics are assigned independent normal priors with a mean of zero and a standard deviation of 100,000.

4. Eigenvector spatial filter frequentist prerequisites

Eigenvector spatial filtering furnishes an approximation to a SAR process in such a way that it replaces the spatial autoregressive structure governing origin- and destination-specific effects in a Poisson model specification. Having done this, standard Poisson regression routines can be used to estimate the remaining parameters of the model, conditional on the (approximated) spatially structured effects parameters. This methodology may be summarized as follows (Griffith, 2009, p. 123):

> The modified geographic weights matrix, say $(\mathbf{I} - \mathbf{11}^T/n)\mathbf{C}(I - \mathbf{11}^T/n)$, where $\mathbf{1}$ is an n-by-1 vector of 1s, and superscript T denotes the matrix transpose operation, from which eigenfunctions are extracted to construct spatial filters appears in the numerator of the Moran Coefficient (MC) spatial autocorrelation index. The eigenvectors of this matrix exhibit the following property: when mapped spatially, the first eigenvector, say $\mathbf{E}_1$, is the set of real numbers that has the largest MC value achievable by any set of real numbers for the spatial arrangement defined by the geographic connectivity matrix $\mathbf{C}$; the second eigenvector is the set of real numbers that has the largest achievable MC value by any set that is uncorrelated with $\mathbf{E}_1$; the third eigenvector is the third such set of values; and so on through $\mathbf{E}_n$, the set of real numbers that has the largest negative MC value achievable by any set that is uncorrelated with the preceding (n – 1) eigenvectors. As such, these eigenvectors furnish distinct map pattern descriptions of latent spatial autocorrelation in georeferenced variables, because they are both orthogonal and uncorrelated. Their corresponding eigenvalues quantify the nature and degree of spatial autocorrelation portrayed by each eigenvector.

The constructed spatial filter is the approximation to a SAR process.

Chun and Griffith (2013, pp. 59-60) describe how eigenvectors are concatenated and combined with Kronecker product matrix operations to capture spatial autocorrelation effects not only in origin and destination geographic distributions, but also in spatial interaction flows network structure linking origins and destinations. These constructions are similar to those in LeSage and Pace (2008, 2009).

5. The spatial interaction model

Expanding upon Wilson (1967), a doubly-constrained gravity model may be written as

$$F_{ij} = O_i D_j \exp(\alpha + \sum_{h=1}^{n} I_{i,h,o}\beta_{h,o} + \sum_{k=1}^{n} I_{j,k,d}\beta_{k,d} - \gamma_d d_{ij} - \gamma_t t_{ij} + \beta I_{Bij} + \rho_{ij}) , \quad (1)$$

where F_{ij} is the patent citation flow between origin region i and destination region j,

O_i and D_j are the total number, respectively, of patents created in origin i and of patent citations in destination j that constitute flows within the system, and enter the Poisson regression model specification as the offset variable $LN(O_i)+LN(D_j)$,

$\exp(\beta_{h,o})$ and $\exp(\beta_{k,d})$ respectively are the balancing factors for origin h (conventionally denoted by A_i; ensuring that the predicted number exactly equals the observed number of patents created in each origin h) and for destination k (conventionally denoted by B_j; ensuring that the predicted number exactly equals the observed number of patent citations in each destination k), where the n^{th} areal unit is assigned $\beta_{n,o}$ (i.e., $A_i = 1$) and $\beta_{n,d}$ (i.e., $B_j = 1$) values of zero (to avoid perfect multicollinearity in the specification)—the sum of these two values is absorbed into the intercept term,

$I_{i,h,o}$ and $I_{j,k,d}$ denote the binary 0-1 indicator variables for, respectively, the origin and destination areal units associated with areal unit i [i.e., indicator variables with n values of one, and n(n-1) values of zero, such that $I_{i,h,o} = 1$ when i = h, and $I_{j,k,d} = 1$ when j = k],

I_{Bij} denotes a binary indicator variable for the presence of a national border intervening between origin (region) i and destination (region) j,

d_{ij} is the distance separating origin (region) i and destination (region) j (measured here between regional economic centers, in 1,000 km),

t_{ij} is the technological proximity between origin (region) i and destination (region) j,

γ_d is the geographic distance decay parameter,

γ_t is the technological distance decay parameter,

ρ_{ij} is the spatial autocorrelation term representing dependencies between flows from nearby regions of i to nearby regions of j (this component is captured by the eigenvector network autocorrelation filter), and

$\exp(\alpha)$ and β respectively are the constant of proportionality and the intervening borders.

The sets of balancing factors $\exp(\beta_{h,o})$ and $\exp(\beta_{k,d})$ respectively ensure that $\sum_{i=1}^{n} \hat{F}_{ij} = O_i$ and that $\sum_{j=1}^{n} \hat{F}_{ij} = D_j$, where $\hat{F}_{ij}$ denotes the predicted flow between origin region i and destination region j. These values appear only when $I_{i,h,o}$ and $I_{j,k,d}$ equal one; otherwise, when $I_{i,h,o}$ and $I_{j,k,d}$ equal zero, the factors become exp(0) = 1, which disappear in the specification because any value multiplied by one equals itself.

The terms $\sum_{h=1}^{n} I_{i,h,o}\beta_{h,o}$ and $\sum_{k=1}^{n} I_{j,k,d}\beta_{k,d}$, the concatenated balancing factors, which are similar to fixed effects, are the counterparts to the Bayesian random effects terms $\boldsymbol{\theta}_k$. Conventional thinking is that this difference in perspective results in 2n–2 degrees of freedom being associated with this component of the frequentist model, and two degrees of freedom (i.e., mean and variance) being associated with the Bayesian model. Accordingly, the Bayesian model specification is not equivalent to an unconstrained gravity model. The two model specifications also differ by the network autocorrelation term ρ_{ij}, which is not in the Bayesian specification but is in the frequentist specification. If both model specifications would be identical, then both models would produce identical estimates. MCMC estimation (Bayesian) is merely a different way of maximizing the objective/likelihood function (à la maximum likelihood techniques; frequentist).

6. Frequentist estimation results

Estimation results from a Poisson regression with only the distance covariates include the following negative exponential function distance exponent parameters:

separation	exponent estimate	standard error	t-statistic
geographical distance	0.4750	0.0247	19.2
technological proximity	2.0407	0.0344	59.3
intervening border	–1.1386	0.0175	–65.1

As geographical distance between two regions increases, the number of patent citations tends to decrease. When borders intervene between two regions, the number of patent citations tends to decrease. And, as technological proximity between two regions increases, the number of patent citations tends to increase. The ranking of these estimates by the absolute values of their corresponding t-statistics coincides with their order of importance according to their partial pseudo-R^2 values: intervening borders accounts for roughly 14.6%, technological proximity for roughly 15.4%, and geographical distances for roughly 2.4% of the variance in the flows. These results are conditional on the offset variable, which indexes the size of origins and of destinations, and accounts for roughly 25.9% of the variance in patent citation flows.

The first two steps in the data analysis were: (i) to calculate the eigenfunctions; and, (ii) to estimate a doubly-constrained gravity model. Eigenfunctions were extracted from the matrix $\mathbf{C} = (\mathbf{B} + \mathbf{B}^T)/2$ because matrix **B** is an asymmetric binary 0-1 contiguity matrix defined as follows:

$b_{ij} = 1$, if j is one of the eight nearest neighbors (based upon economic centroid great circle distances) to i; and,
$b_{ij} = 0$, otherwise.

Because this definition of contiguity results in an asymmetric matrix, it needs to be converted to a symmetric matrix in order to work with real numbers; a standard adjustment is adding it to its transpose. Division by two avoids double counting, and ensures that if matrix **B** is symmetric, $\mathbf{C} = \mathbf{B}$. Using a selection threshold value of 0.25 for the absolute value of the ratio of eigenvalues to the largest eigenvalue (which yields a Moran Coefficient of 1.07 for this matrix **C**) yields 42 candidate eigenvectors representing various levels of weak-to-strong positive spatial autocorrelation, and 34 candidate eigenvectors representing various levels of weak-to-strong negative spatial autocorrelation. Meanwhile, a Poisson regression routine was used to estimate the doubly-constrained gravity model. Figure 1a portrays the goodness-of-fit for this model: the open circles denote the data points, and the linear alignment of the xs denotes the perfect fit line. This scatterplot exhibits a typical V-shaped variance relationship with size: the variance of a Poisson random variable increases as its expected value increases. The accompanying pseudo-R^2 value is 0.7323 (i.e., the balancing factors account for an additional roughly 15.0% of the variance in flows); the bivariate regression trend line (Figure 1a) has an intercept (–0.2947) that is significantly different from zero, and a slope (1.0706) that is significantly different from one; given that these values do not appear to be substantially different from their null hypothesis values, these significance results may be exaggerated because of the considerably large sample size involved. And, the Poisson regression results indicate overdispersion, with a deviance statistic of 2.5924. Adjusting for spatial autocorrelation shrinks the overdispersion (Figure 1b) to

1.7784; the accompanying bivariate regression has an intercept of –0.2150 and a slope of 1.052 (both improvements vis-à-vis the non-network autocorrelation adjusted results). The network autocorrelation term accounts for roughly 16% of the variability in flows.

The origin and destination log-balancing factors each can be decomposed into spatially structured (i.e., spatial filters) and spatially unstructured variates. These decompositions result in the following equations, with the eigenvectors chosen to construct the spatial filters being selected with a stepwise regression procedure (level of significance = 0.10):

$$\hat{\boldsymbol{\eta}} = 0.49887\,\mathbf{1} + {}_{22}\mathbf{E}_c\mathbf{b}_o + \mathbf{E}_{15}b_{15,o} + \mathbf{e}_o \text{ , and}$$

$$\hat{\boldsymbol{\pi}} = 0.65772\,\mathbf{1} + {}_{22}\mathbf{E}_c\mathbf{b}_d + \mathbf{E}_{11}b_{11,d} + \mathbf{E}_{62}b_{62,d} + \mathbf{e}_d \text{ ,}$$

where ${}_{22}\mathbf{E}_c$ denotes the set of 22 common eigenvectors, and $\mathbf{b}_o$ and $\mathbf{b}_d$ respectively denote the 22-by-1 vectors of estimated regression coefficients for the origin and destination parts of the network autocorrelation. The spatial filters (the first contains 23 and the second contains 24 positive, and each respectively contains three and four negative spatial autocorrelation eigenvectors; the negative spatial autocorrelation eigenvectors account for less than 2% of the variance in the flows) in these two equations respectively account for roughly 71% and 69% of the variance in the origin and destination log-balancing factors; the respective MCs for these spatial filters are 0.92 and 0.91. These two spatial filters have 22 eigenvectors in common. The Shapiro-Wilk diagnostic statistic indicates that neither $\mathbf{e}_o$ nor $\mathbf{e}_d$ conforms closely to a normal distribution. Both $\mathbf{e}_o$ and $\mathbf{e}_d$ have only trace spatial autocorrelation ($z_{MC} = 0.3$).

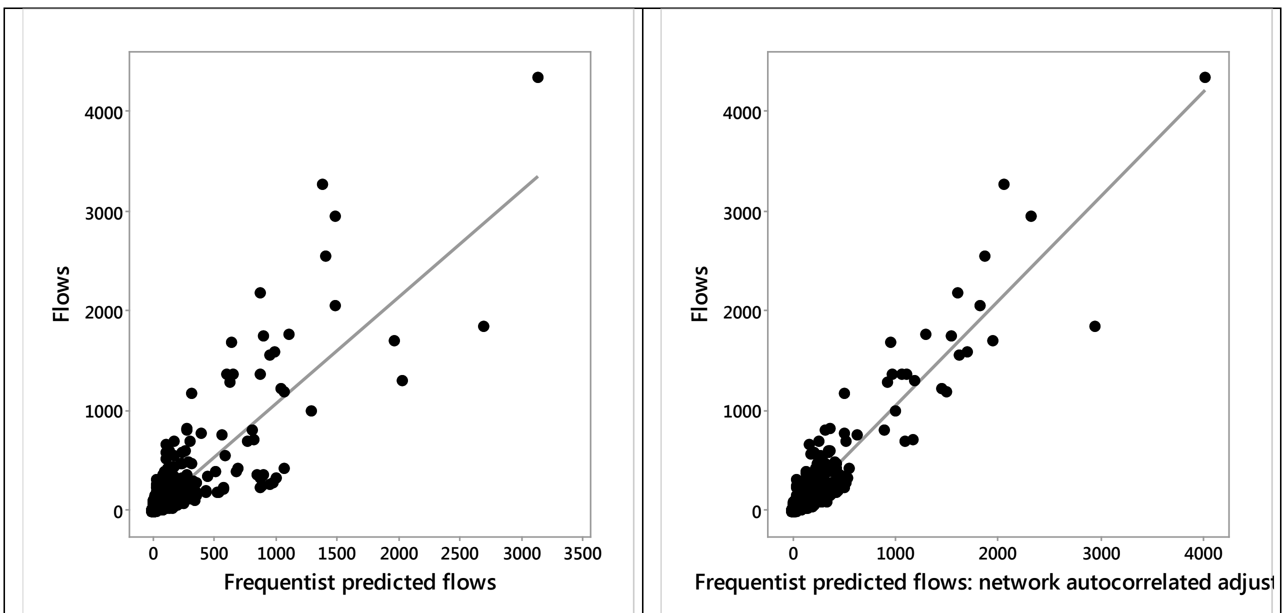

Figure 1. Scatterplot of observed (vertical axis) versus predicted flows (horizontal axis). Left (a): frequentist doubly-constrained results. Right (b): frequentist network autocorrelation adjusted results.

The geographic distributions of the log-balancing factor components appear in Figure 2. Both spatial filters display considerable spatial pattern, suggesting concentric zones centering on Central Europe (a swath from southern Germany through southern England), and radiating outward toward the periphery of Europe. Balancing factors index the propensity to send and to receive patent citations: Central Europe has the highest propensity to both send and receive, whereas the periphery of Europe has the lowest propensity to both send and receive patents. In contrast, the spatially unstructured components display a haphazard pattern.

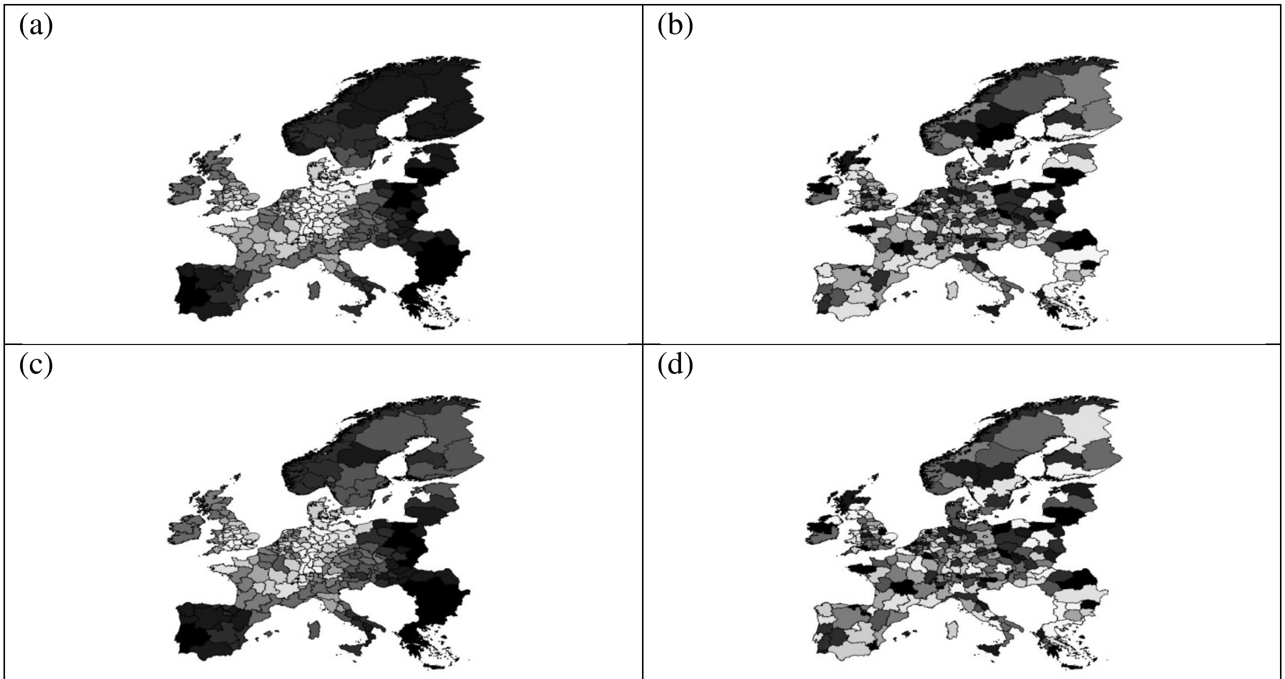

Figure 2. Top left (a): spatial filter for the origin log-balancing factor. Top right (b): spatially unstructured part of the origin log-balancing factor (e_o). Bottom left (c): spatial filter for the destination log-balancing factor. Bottom right (d): spatially unstructured part of the destination log-balancing factor (e_d).

7. Comparisons of Bayesian and frequentist results

Table 1 summarizes estimation results for the two approaches[5]. Specification differences to keep in mind include the network autocorrelation term ρ_{ij}, which is not included in analyses reported here so that results are more comparable, and the inclusion of an offset variable that is the sum of the logarithms of origin and destination flows totals.

Table 1. Results for the frequentist and Bayesian estimations.

separation	doubly-constrained gravity model		gravity model with Bayesian random effects	
	exponent estimate	standard error	exponent estimate	standard error
geographical distance	0.9092	0.0163	0.6170	0.0054

[5] The frequentist model specification includes fixed effects, which are the origin/destination balancing factors. The Bayesian model is not an unconstrained version of the gravity model. This latter specification includes random effects, which involve 4 degrees of freedom. These two specifications reflect the common fixed versus random effects debate. Furthermore, if we were to make both model specifications identical, then both models would produce identical estimates. MCMC estimation (Bayesian) is merely a different way of maximizing the objective/likelihood function (a la maximum likelihood techniques; frequentist).

technological proximity	2.9788	0.0239	3.0830	0.0113
intervening border	–1.2628	0.0117	–1.6666	0.0062
LN(A)			0.3167	0.0021
LN(B)			0.3468	0.0043
$LN(O_T)$	1	0		
$LN(D_T)$	1	0		

Both specifications include the same distance variates. The estimate for technological proximity and the intervening border indicator variable are similar, with the Bayesian standard errors being smaller. A principal difference here is for the geographical distance decay parameter, with the Bayesian estimate lying between the frequentist estimates ignoring and adjusting for spatial autocorrelation.

Figure 3 presents scatterplots portraying the relationships between the Bayesian random effects terms and the frequentist balancing factor terms (estimates and standard errors). Although some trend exists in the estimates themselves, it is not very pronounced.

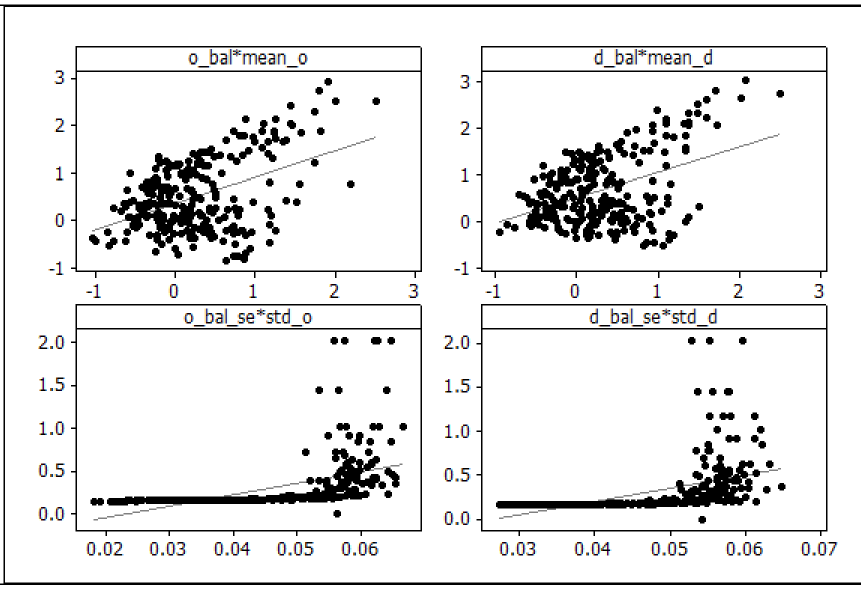

Figure 3. Scatterplots for model comparisons. Top left (a): origin log-balancing factors versus mean origin latent unobservables. Top right (b): destination log-balancing factors versus mean destination latent unobservables. Bottom left (c): standard error of origin log-balancing factors versus standard deviation of origin latent unobservables. Bottom right (d): standard error of destination log-balancing factors versus standard deviation of destination latent unobservables.

Figure 1a presents the scatterplot of actual versus predicted flows for the frequentist model specification without a network autocorrelation term; this specification is more directly comparable with the Bayesian model specification (Figure 4). The frequentist results are superior to the Bayesian results in this case. The total number of predicted flows for the frequentist analysis is 275,578, matching the observed number of flows. The total number of predicted flows for the Bayesian analysis is 1,571,100, nearly six times the observed number of flows. A bivariate regression of the observed flows on the frequentist predicted flows has an intercept of

–0.3 and a slope of 1.1; those for the Bayesian predicted flows are, respectively, 0.0 and 0.2. Both scatterplots exhibit the funnel shaped scatter of points that is typical of a Poisson random variable.

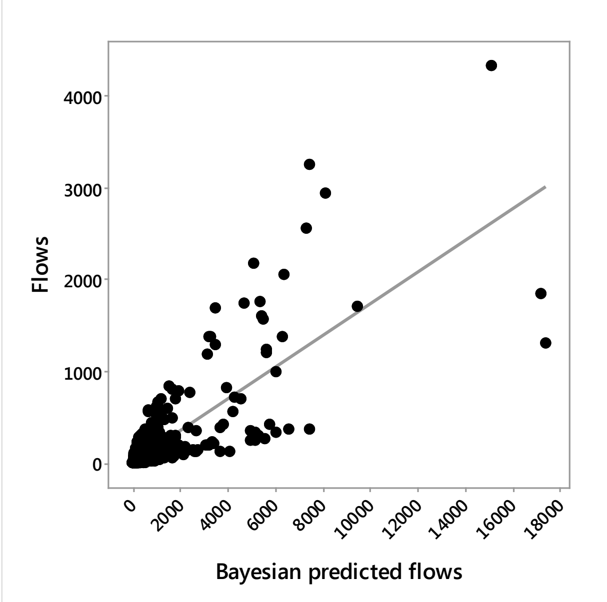

Figure 4. Scatterplot of observed (vertical axis) versus predicted flows (horizontal axis: Bayesian results.

8. Conclusions

A primary expectation is that the frequentist and Bayesian spatial interaction results would be very similar, if not the same. Comparisons summarized in this paper suggest otherwise. In addition, the frequentist specifications outperform the Bayesian specification, in terms of prediction accuracy, reduction of extra-Poisson variation, and observed-predicted value alignment. Although the technological proximity and intervening border variate relationship estimates are very similar, the geographical distance decay parameters are not; presumably the frequentist estimate is more reliable.

9. References

Chib, S., E. Greenberg, and R. Winkelmann. 1998. Posterior simulation and Bayes factors in panel count data models, *J. of Econometrics*, 86: 335-344.

Chun, Y., and D. Griffith. 2013. *Spatial Statistics & Geostatistics*. Thousand Oaks, CA: SAGE.

Fischer, M., and D. Griffith. 2008. Modelling spatial autocorrelation in spatial interaction data: An application to patent data in the European Union, *J. of Regional Science*, 48: 969-989.

Flowerdew, R. and M. Aitkin. 1982. A method of fitting the gravity model based on the Poisson distribution, *Journal of Regional Science*, 22, 191–202.

Frühwirth-Schnatter, S., and H. Wagner. 2004. Data augmentation and Gibbs sampling for regression models of small counts, *IFAS Research Paper Series* # 2004-04. Linz, Austria: Johannes Kepler University Linz, Department for Applied Statistics.

Griffith, D. 2007. Spatial structure and spatial interaction: 25 years later, *The Review of Regional Studies*, 37, #1: 28-38.

Griffith, D. 2009. Modeling spatial autocorrelation in spatial interaction data: empirical evidence from 2002 Germany journey-to-work flows, *J. of Geographical Systems*, 11: 117-140.

Griffith, D., and Y. Chun. 2015. Spatial autocorrelation in spatial interactions models: geographic scale and resolution implications for network resilience and vulnerability, *Networks and Spatial Economics*, 15: 337-365.

Griffith, D., and M. Fischer. 2013. Constrained variants of the gravity model and spatial dependence: Model specification and estimation issues, *J. of Geographical Systems*, 15: 291-317.

LeSage, J., and R. Pace. 2008. Spatial econometric modelling of origin-destination flows, forthcoming in the *Journal of Regional Science*, 48, 941-967.

LeSage, J., and R. Pace. 2009. *Introduction to Spatial Econometrics*. Boca Raton, FL: Taylor & Francis.

LeSage, J., M. Fischer and T. Scherngell. 2007. Knowledge spillovers across Europe: Evidence from a Poisson spatial interaction model with spatial effects, *Papers in Regional Science*, 86: 393-421.

Maurseth, P., and B. Verspagen. 2002. Knowledge spillovers in Europe: A patent citation analysis, *Scandinavian J. of Economics*, 104: 531-545.

Wilson, A. (1967). A statistical theory of spatial distribution models, *Transportation Research*, 1, 253-269.