Chapter 3
Constrained Variants of the Gravity Model and Spatial Dependence: Model Specification and Estimation Issues

Daniel A. Griffith and Manfred M. Fischer

Keywords Constrained gravity models • Count data • Patent citation flows • Poisson • Spatial dependence in origin-destination flows • Spatial econometrics • Spatial filtering

JEL Classifications: C13, C31, R15

3.1 Introduction

Traditionally, models that use origin-destination flow data to explain variation in the level of flows between origin and destination locations of interaction across some relevant geographic space are called gravity models,[1] in analogy with Newton's concept of gravity. Locations may be either regions or point units, and spatial interactions relate to movements of various kinds. Examples include not only migration, journey-to-work, traffic, commodity or trade flows, but also flows of less tangible entities such as capital, information and knowledge. By adopting a spatial

This paper has been previously published in the *Journal of Geographical Systems*. Special Issue on "Advances in the Statistical Modelling of Spatial Interaction Data", Vol. 15, Number 3/July 2013, ©Springer-Verlag Berlin Heidelberg, pp. 291–317.

[1]The terms gravity model and spatial interaction model are often used interchangeably. But they are not the same. Spatial interaction models not only include gravity models, but also similar models that have been derived using powerful methods of entropy maximisation from statistical mechanics (Wilson 1967), or utility maximization from economic theory (Niedercorn and Bechdolt 1969), and those based on intervening opportunities which can be derived heuristically.

D.A. Griffith (✉)
University of Texas at Dallas, Richardson, TX, USA
e-mail: dagriffith@utdallas.edu

M.M. Fischer
Vienna University of Economics and Business, Vienna, Austria
e-mail: manfred.fischer@wu.ac.at

R. Patuelli, G. Arbia (eds.), *Spatial Econometric Interaction Modelling*, Advances in Spatial Science, DOI 10.1007/978-3-319-30196-9_3

interaction perspective, attention is focused on interaction patterns at the aggregate rather than the individual level.

Gravity models[2] typically rely on three types of factors to explain mean interaction frequencies: Origin specific variables that characterise the ability of origin locations to produce or generate flows; destination specific variables that attempt to capture the attractiveness of destination locations; and, a separation function that reflects the way spatial separation of origins from destinations constrains or impedes the interaction (Fischer and Wang 2011). At larger scales of geographical inquiry spatial separation might be simply measured in terms of the great circle distance separating an origin from a destination location. In other cases, it might be transportation cost, perceived travel time or any other sensible measure such as political distance, language distance and cultural distance measured in terms of nominal or categorical attributes. One popular example of a separation or deterrence function is the exponential function that leads to gravity models known as exponential gravity models.

Alternative forms of the gravity model can be specified by imposing (exogenous) constraints on the mean interaction frequencies. These model variants include origin and/or destination specific balancing (normalising) factors that act as constraints to ensure that the origin and destination totals for spatial interactions are exactly predicted (see Wilson 1971). The model is said to be doubly constrained if both origin and destination constraints hold for each location. If either the origin or the destination constraints hold the model is singly constrained; otherwise it is unconstrained. It is worth noting that the doubly constrained gravity model has also become known as the trip distribution stage in the four-step transport planning approach.[3] One more recently recognized role of these constraints is their accounting for spatial autocorrelation effects in the geographic distribution attributes across origins and destinations.

The focus in this paper is on singly and doubly constrained exponential gravity model variants for situations involving flows taking the form of counts; for example, counts of persons commuting from home to work locations, or as in the example

[2]For a discussion of problems that plague empirical implementation of regression-based gravity models, and econometric extensions that have recently appeared in the literature, see LeSage and Fischer (2010). These new models replace the conventional assumption of independence among origin-destination flows with formal approaches that allow for spatial dependence in flow magnitudes. The econometric extensions are based on the assumption of a linear relation between the dependent and the independent variables, and this assumes the dependent variable to be normally distributed.

[3]Trip making is viewed as consisting of four components (see, for example, Fischer 2000): trip generation and attraction (the decision to make a trip and how often); trip distribution in a system of traffic zones; modal split (choice of mode of transport); and, trip assignment (choice of route through network). The gravity model is used for trip distribution, but is preceded by trip generation and attraction models that provide independent estimates of locational (zonal) trip origins and attractions that subsequently become the "mass" terms of the gravity model. Thus, the definition of the row and column sums of the predicted trip matrix coincides exactly with the definitions of the respective mass terms.

considered in this paper, the number of patent citations from one region to another. In such cases, current practice is to model origin-destination flow data with Poisson gravity model specifications. Under the assumption that the flows are independently distributed Poisson variables the constrained gravity model variants can be treated as particular cases of a generalised linear model (GLM) with fixed (or random) effects, employing a logarithmic link function and a Poisson mean flow. Maximum likelihood estimates of the model parameters can be achieved using an iterative re-weighted least squares algorithm, as implemented in statistical software packages such as GLIM (Generalised Linear Iterative Modelling).

Flows, however, are not strictly independent. Spatial (or network) dependence[4] is more likely than spatial independence when considering origin-destination flows. Spatial dependence in a flow setting refers to a situation where flows from nearby locations (either origins or destinations) are similar in magnitude. A failure to incorporate spatial dependence in model specifications leads to biased parameter estimates and incorrect conclusions. Eigenvector spatial filtering—described in Griffith (2003) for conventional regression models—offers an approach to dealing with spatial dependence in constrained gravity model variants. This approach relies on the decomposition of a spatial weight matrix into eigenvalues and eigenvectors and then uses a subset of the eigenvectors as additional explanatory variables in the singly and doubly constrained gravity model specifications to reduce potential bias in parameter estimates. A virtue of this approach is that existing software can be applied for the case of spatial dependence in constrained model variants involving flows taking the form of counts.

The purpose of this paper is twofold: first, to establish theoretical connections between the constrained gravity model versions with balancing factors, fixed effects represented by binary location specific indicator variables, and random effects; and second, to illustrate these connections with an empirical example while accounting for spatial dependence among flows during estimation. Fulfilling these goals reveals that fixed and random effects are identical and equal to the logarithm of the entropy maximisation derived factors, except for slight rounding/algorithm-convergence errors. This finding is the outcome of an equivalency between assigning a single fixed effects indicator variable to each origin/destination on the one hand, and estimating a single random effects value for an origin/destination while treating the corresponding destinations/origins as repeated measures, on the other. As with the unconstrained gravity model variant, adjusting for spatial dependence in origin-destination flows reduces bias in parameter estimates and improves model performance.

The rest of the paper is organised as follows. Section 3.2 describes unconstrained and constrained classes of the gravity model with a focus on doubly and singly constrained model variants that rely on a multiplicative adjustment scheme to

[4] Spatial dependence is also known as network autocorrelation (see Black 1992; Chun 2008; Griffith 2009; Chun and Griffith 2011) even though there are similar differences between both as between spatial dependence and spatial autocorrelation in general.

enforce satisfactorily the conservation rule. Section 3.3 presents the counterpart Poisson specifications that interpret/predict the level of flows as dependent on not only the explanatory variables (and the associated coefficient estimates), but also origin and destination specific effects coefficients. Section 3.4 describes spatial filtering as a way of filtering the sample origin-destination data for spatial dependence (i.e., transferring spatial autocorrelation effects from residuals to the mean/intercept parameter) in an effort to mimic independent data amenable to standard Poisson regression estimation procedures. Section 3.5 continues to establish theoretical connections between balancing factors, fixed effects and random effects (spatially filtered) model specifications. The results are illustrated in Sect. 3.6 with an empirical example involving knowledge flows between 257 European regions resulting in $257^2 = 66{,}049$ flow dyads. Section 3.7 concludes the paper.

3.2 Unconstrained and Constrained Classes of Gravity Models: The Classical View

Gravity models that describe mean interaction frequencies in a system of n locations can be written[5] as

$$E\left(Y_{ij}\right) = K_{ij}\, U_i\, V_j f\left(d_{ij}\right) \tag{3.1}$$

where Y_{ij} $(i, j = 1, \ldots, n)$ is a random variable denoting the level of flows from origin location i $(i = 1, \ldots, n)$ to destination location j $(j = 1, \ldots, n)$, U_i and V_j are appropriate origin and destination specific factors or functions reflecting locational propensities to emit or attract interactions, $f(d_{ij})$ is a separation function of some inter-location measure d that separates origin i from destination j, and K_{ij} is an origin-destination specific constant of proportionality, or scaling factor, which reduces to a constant scaling factor K for the unconstrained gravity model specification (which then is accompanied by attaching exponents of other than one to U_{i} and V_{j}). The role of this origin-destination specific constant of proportionality in the gravity model equation depends on how extensively the conservation rule (Ledent 1985) is enforced in the system of locations. Four alternative cases may be distinguished, giving rise to equally many classes of gravity models.

A gravity model is called unconstrained if the conservation principle is ignored altogether so that

$$K_{ij} = K \tag{3.2}$$

[5] An alternative formulation to that given in Eq. (3.1) is $Y_{ij} = K_{ij}\, U_i\, V_j f\left(d_{ij}\right)\, \eta_{ij} + \varepsilon_{ij}$ where ε_{ij} reflects the sample error and η_{ij} the specification error. In this case, the stochastic nature of Y_{ij} derives from assumptions made about the stochastic nature of ε_{ij} and η_{ij}.

where K is a constant scaling factor independent of all origins and destinations. If $Y_{\bullet\bullet}$ denotes the total number of flows in the spatial system, then

$$K = Y_{\bullet\bullet}\left[\sum_{\substack{i=1\\ i\neq j}}^{n}\sum_{j=1}^{n} U_i V_j f\left(d_{ij}\right)\right]^{-1} \tag{3.3}$$

where the summation is over the range $i = 1, \ldots, n$ and $j = 1, \ldots, n$. Although Eq. (3.1) has been developed by analogy with Newton's gravity equation, Isard (1960), and Sen and Smith (1995) developed versions of the unconstrained model using a probabilistic approach.

At the other extreme of the spectrum is the doubly constrained case of spatial interaction that refers to a situation in which the conservation principle is enforced from both the viewpoint of origin and destination locations. The origin-destination specific constant of proportionality, K_{ij}, now depends on both origins and destinations. For simplicity, it is generally assumed[6] that

$$K_{ij} = A_i\, B_j \tag{3.4}$$

where the origin and destination specific constants, A_i and B_j, called balancing factors, are solutions of the equation system (Wilson 1967)

$$A_i = Y_{i\bullet}\left[U_i \sum_{\substack{j=1\\ j\neq i}}^{n} B_j\, V_j f\left(d_{ij}\right)\right]^{-1} \tag{3.5}$$

$$B_j = Y_{\bullet j}\left[V_j \sum_{\substack{i=1\\ i\neq j}}^{n} A_i\, U_i f\left(d_{ij}\right)\right]^{-1}. \tag{3.6}$$

These balancing factors act as constraints to ensure that the estimated inflows $\hat{Y}_{\bullet j}$ for $j = 1, \ldots, n$ and outflows $\hat{Y}_{i\bullet}$ for $i = 1, \ldots, n$ equal the observed inflow and outflow totals, respectively. Such doubly (or attraction-production) constrained models have

[6]The multiplicative form of the balancing factors A_i and B_j (Wilson 1967) ensures mathematical tractability in searching for an adequate estimation procedure. Alternatively, Tobler (1983) suggests an additive adjustment scheme, $K_{ij} = A_i + B_j$, to enforce satisfactorily the conservation rule. Ledent (1985) introduces a general functional form that subsumes both the multiplicative (Wilson) and the additive (Tobler) variants.

been extensively used as trip distribution models in transport modelling, and many variants of this model form exist for describing journey-to-work interactions.

Between these two extreme cases of unconstrained and doubly constrained spatial interaction lie many models that are subject to some constraints but not to others. Two important classes can be identified: the origin (or alternatively called production) constrained gravity model, and the destination (or alternatively called attraction) constrained gravity model. In the production constrained case the conservation principle is enforced from the viewpoint of origin locations[7] only. Hence

$$K_{ij} = A_i. \tag{3.7}$$

A_i is a factor dependent on the location of an origin, and is called an origin specific balancing factor. If $Y_{i\bullet}$ denotes the total number of outflows from location i,

$$A_i = Y_{i\bullet} \left[U_i \sum_{\substack{j=1 \\ j \neq i}}^{n} V_j f\left(d_{ij}\right) \right]^{-1}. \tag{3.8}$$

The origin constrained gravity model is useful in situations where the outflow totals are known or can be exogenously predicted for each origin location in the system. For an instructive example see Haynes and Fotheringham (1984, pp. 60–62).

The attraction constrained case of spatial interaction enforces the conservation principle from the viewpoint of destination locations. Thus

$$K_{ij} = B_j \tag{3.9}$$

where B_j is a factor dependent on the destination location. If $Y_{\bullet j}$ denotes the total number of inflows into location j,

$$B_j = Y_{\bullet j} \left[V_i \sum_{\substack{i=1 \\ i \neq j}}^{n} U_i f\left(d_{ij}\right) \right]^{-1}. \tag{3.10}$$

This model variant can be used to forecast total outflows from origin locations. Such a situation might arise, for example, in forecasting the effects of locating a new industrial park within a metropolitan area. The number of people to be

[7] In the origin constrained and the destination constrained models presented here, the constraints to which these models are subject refer to the full set of n origin or n destination locations. But it is possible to develop models that are only constrained over certain subsets of locations. Such models, which are not considered in this paper, may be found in Wilson (1970).

employed in the new development area is known, and the destination constrained gravity model can be used to forecast the demand for housing in particular locations of the metropolitan area that will result from the new employment (Haynes and Fotheringham 1984).

The models presented in Eqs. (3.1)–(3.10) are in a generalised form and no mention has yet been made to the particular set of parameters characterising such gravity models. Although the balancing factors are sometimes referred to as parameters, in this paper the term parameter is restricted to those constants that must be estimated statistically, rather than to those constants that imply the accounting constraints placed on the model.

Many different model formulations can be obtained from Eq. (3.1), despite its structural simplicity (see Baxter 1983). U_i and V_j can be treated as completely known, as parameters to be estimated (see Cesario 1973), or as simple power functions of some known variables (see Fotheringham and O'Kelly 1989). The separation function constitutes the very core of gravity models.[8] In this study we use the multivariate exponential deterrence function

$$f\left(d_{ij}\right) = \exp\left(-\theta\ d_{ij}\right) \tag{3.11}$$

in which d denotes a multivariate separation measure with an associated sensitivity parameter θ. This specification of the spatial separation function leads to the following three variants of the gravity model: the doubly constrained variant

$$E\left(Y_{ij}\right) = A_i\ B_j\ U_i\ V_j\ \exp\left(-\theta\ d_{ij}\right) \tag{3.12}$$

$$A_i = Y_{i\bullet}\left[U_i \sum_{\substack{j=1\\ j\neq i}}^{n} B_j\, V_j \exp\left(-\theta d_{ij}\right)\right]^{-1} \tag{3.13}$$

$$B_j = Y_{\bullet j}\left[V_j \sum_{\substack{i=1\\ i\neq j}}^{n} A_i\, U_i \exp\left(-\theta d_{ij}\right)\right]^{-1} \tag{3.14}$$

the origin constrained variant

$$E\left(Y_{ij}\right) = A_i\ U_i\ V_j\ \exp\left(-\theta\ d_{ij}\right) \tag{3.15}$$

[8]The notion that separation functions in conventional gravity models work to effectively capture spatial dependence in origin-destination flows has long been challenged. Griffith (2007) provides an historical review of the regional science literature about this topic in which he credits Curry (1972) as the first to conceptualise the problem of spatial dependence in flows.

$$A_i = Y_{i\bullet}\left[U_i \sum_{\substack{j=1\\ j\neq i}}^{n} V_j \exp\left(-\theta d_{ij}\right)\right]^{-1} \tag{3.16}$$

and, the destination constrained variant

$$E\left(Y_{ij}\right) = B_j\ U_i\ V_j\ \exp\left(-\theta\ d_{ij}\right) \tag{3.17}$$

$$B_j = Y_{\bullet j}\left[V_j \sum_{\substack{i=1\\ i\neq j}}^{n} U_i \exp\left(-\theta d_{ij}\right)\right]^{-1}. \tag{3.18}$$

The central concern in this paper is with the problem of estimating the model parameter θ rather than with the problem of determining appropriate values for the balancing factors.[9] A solution to this latter problem, for example in the case of Eqs. (3.12)–(3.14), involves using an iterative biproportional adjustment technique, known as the Deming-Stephan-Furness procedure[10] (see Sen and Smith 1995, p. 374). As Evans (1970) shows, convergence to a unique set of values for A_i and B_j is guaranteed for any non-trivial set of starting values.

3.3 Poisson Versions of the Constrained Gravity Models

Flows often take the form of counts such as numbers of migrants moving from one location to another. In such situations a common assumption is that the Y_{ij} $(i, j = 1, \ldots, n)$ follow independent[11] Poisson distributions,[12] $Y_{ij} \sim \mathcal{P}(\mu_{ij})$, where μ_{ij} is equated with the right hand side of Eq. (3.1). The mean and the variance of the distribution are equal to μ_{ij}. The Poisson specifications of the gravity model

[9]The constrained gravity model variants are intrinsically non-linear in their parameters, and thus the application of linear methods leads to biased estimates of these parameters.

[10]In the economics literature it is often called the RAS procedure.

[11]Independence means that the individual flows from origin i to destination j are independent from each other, and that origin-destination flows between any pair of locations are independent from flows between any other pair of locations.

[12]Closely related to this assumption are the assumptions that the set of observations for each origin location has a multinomial distribution, say $\mathcal{MN}(Y_{i1}, Y_{i2}, \ldots, Y_{in}; Y_{i\bullet})$, or that the set of all observations has a multinomial distribution $\mathcal{MN}(Y_{i1}, Y_{i2}, \ldots, Y_{nn}; Y_{\bullet\bullet})$, where $Y_{i\bullet}$ is the total flow from origin location i, $Y_{\bullet\bullet}$ is the overall flow, and n is the number of origin and destination locations. These multinomial distributions can be generated by assuming that the Y_{ij} $(i, j = 1, \ldots, n)$ are independent Poisson random variables sampled subject to the origin totals $Y_{i\bullet}$, or the overall total $Y_{\bullet\bullet}$, being fixed (Bishop et al. 1975).

variants interpret/predict the level of flows as dependent on not only the explanatory variables (and their associated coefficient estimates), but also origin and destination specific effects coefficients. The fixed effects version of the three constrained model variants of the gravity model can be described as in Eqs. (3.19)–(3.21), respectively.

$$E\left(Y_{ij}\right) = \mu_{ij} = U_i\ V_j \exp\left[\alpha + \sum_{h=1}^{n-1} I_{iho}\ \beta_{ho} + \sum_{k=1}^{n-1} I_{jkd}\ \beta_{kd} - \theta\ d_{ij}\right] \tag{3.19}$$

$$E\left(Y_{ij}\right) = \mu_{ij} = U_i\ \exp\left[\alpha + \sum_{h=1}^{n-1} I_{iho}\ \beta_{ho} - \theta\ d_{ij}\right] \tag{3.20}$$

$$E\left(Y_{ij}\right) = \mu_{ij} = V_j\ \exp\left[\alpha + \sum_{k=1}^{n-1} I_{jkd}\ \beta_{kd} - \theta\ d_{ij}\right] \tag{3.21}$$

with origin h and destination k specific effects coefficients $\exp(\beta_{ho})$ and $\exp(\beta_{kd})$, and corresponding binary indicator variables[13] I_{iho} and I_{jkd} for origins i and destinations j respectively,

$$I_{iho} = \begin{cases} 1 & \text{if } i = h \\ 0 & \text{otherwise} \end{cases} \tag{3.22}$$

$$I_{jkd} = \begin{cases} 1 & \text{if } j = k \\ 0 & \text{otherwise.} \end{cases} \tag{3.23}$$

The fixed effects parameters inflate or deflate the level of flows, depending on whether they are positive or negative. Of note is that one of the origin and one of the destination specific effects coefficients, β_{no} and β_{nd}, have to be set to zero to avoid perfect collinearity in the specifications, and these values are absorbed in the intercept term α.

The most direct approach to estimating the models is with maximum likelihood techniques. The likelihood function to be maximised is proportional to

$$\mathcal{L} = \prod_{i,j} \mu_{ij}^{y_{ij}} \exp\left(-\mu_{ij}\right) \tag{3.24}$$

[13]One advantage of the use of origin/destination indicator variables in a Poisson regression specification is that they yield individual rather than a single aggregate standard error, and null hypothesis probability estimates for each of the individual values in the two sets of balancing factors. One disadvantage is the amount of time necessary to estimate a GLM containing $2n - 2$ indicator variables.

where y_{ij} is the realisation of the random variable Y_{ij}. The Poisson distribution is a member of the exponential family of distributions. Hence parameter estimation can be achieved via GLMs (see McCullagh and Nelder 1983) so that the constrained gravity model variants (3.19)–(3.21) can be treated simply as particular cases of a GLM with a logarithmic link function[14] and a Poisson mean. Then for the doubly constrained case, for example, we get

$$\log\left[E\left(Y_{ij}\right)\right] = \log \mu_{ij} = \log U_i + \log V_j + \alpha + \sum_{h=1}^{n-1} I_{iho}\ \beta_{ko} + \sum_{k=1}^{n-1} I_{jkd}\ \beta_{kd} - \theta\ d_{ij} \tag{3.25}$$

where the term $\left(\log U_i + \log V_j\right)$ is included in the estimation procedure as an offset variable (that is, its coefficient is fixed to equal one).

The maximum likelihood estimates can be derived by means of an iterative re-weighted least squares procedure[15] that is implemented in many statistical software packages such as GLIM. A convenient property of the Poisson assumption along with the log-linear functional form assumed for μ_{ij} is that the resulting maximum likelihood estimates guarantee that the fitted flows $\hat{Y}_{ij}$ satisfy relationships that are consistent with the desirable origin and/or destination constraints of spatial interaction[16] (see Kirby 1974; Davies and Guy 1987, and Bailey and Gatrell 1995, pp. 353–354 for details). Hence, there is no need to modify standard maximum likelihood parameter estimation to incorporate explicit constraints on predicted flows. The goodness-of-fit of GLMs is assessed on the basis of the log-likelihood ratio statistic, known as the deviance.

Fixed effects model specifications allow the unobserved location specific effects to be correlated with the explanatory variables. If the individual effects are strictly uncorrelated with the regressors, then it might be appropriate to model the location specific constant terms as randomly distributed across the locations. The role of random effects terms in this context may be twofold: first, supporting inferences beyond the specific fixed values of covariates employed in an analysis, and, second, accounting for correlation in a non-random sample of data being analysed, in part due to missing variables, for which they function as a surrogate. Random effects may be used if the values of independent variables—which were not deliberately selected

[14]The logarithmic link function is best thought of as being an exponential conditional mean function.

[15]McCullagh and Nelder (1983) prove that the procedure converges to the maximum likelihood solution. Note that zero observed flows do not require any special treatment.

[16]The equivalence of maximum likelihood estimation with the Poisson assumption and the entropy maximisation solution for a doubly constrained gravity model with origin and destination specific balancing factors is well known (see Wilson and Kirkby 1980, p. 310). In the latter case, parameter estimation of a model such as Eq. (3.1) is obtained by maximising an objective function subject to sets of constraints on the origin and destination totals in combination with some constraint on a general measure of spatial separation in the system of locations (Baxter 1982).

by an experimenter—are thought to be a small subset of all possible values to which inferences are to be made, to account for heterogeneity/overdispersion/inter-observation correlation, or to handle observations that are not obtained by simple random sampling but come from a cluster or multi-level sampling design.

The random effects counterparts[17] of the fixed effects model specifications (3.19)–(3.21) may be formulated as in Eqs. (3.26)–(3.28).

$$E\left(Y_{ij}\right) = \mu_{ij} = U_i\ V_j \exp\left[\alpha + \xi_{io} + \xi_{jd} - \theta\ d_{ij}\right] \tag{3.26}$$

$$E\left(Y_{ij}\right) = \mu_{ij} = U_i\ \exp\left[\alpha + \xi_{io} - \theta\ d_{ij}\right] \tag{3.27}$$

$$E\left(Y_{ij}\right) = \mu_{ij} = V_j \exp\left[\alpha + \xi_{jd} - \theta\ d_{ij}\right] \tag{3.28}$$

with zero mean normally distributed origin and destination specific random effects[18] ξ_{io} and ξ_{jd}.

Finally, note that there is a constant of proportionality term, α, in the preceding Poisson gravity model specifications. This term is made explicit because the balancing factors that can be calibrated—as already mentioned – with the Deming-Stephan-Furness procedure, have a constant factored from them. This factorisation is achieved by: (i) setting the maximum A_i and/or the maximum B_j values to one at each iteration step in the procedure; (ii) arbitrarily removing one of the origin and/or destination indicator variables in the fixed effects specifications; and, (iii) imposing a mean of zero on the random effects prior distribution. This is equivalent to rewriting Eq. (3.12) as $E(Y_{ij}) = K\,A_i\,B_j\,U_i\,V_j\ \exp(-\theta\ \ d_{ij})$, where K is a constant.

3.4 Accounting for Spatial Dependence in the Model Specifications

origin-destination flows are not independent (Bolduc et al. 1995; Tiefelsdorf 2003), because flows are fundamentally spatial in nature (LeSage and Pace 2009). Spatial dependence in flows relates to correlations among flows between locations that are neighbouring a given origin-destination pair of locations.[19] Hence, a failure

[17]Whether the random effects model variants are appropriate model specifications in spatial research remains controversial. When the random effects gravity models are implemented, the spatial units of observation should be representative of a larger population, and n should potentially be able to get to infinity (see Elhorst 2010 for more details on this issue).

[18]Origin/destination specific spatial dependence in the random effects estimates motivated the gravity model set forth in LeSage et al. (2007) that formally incorporates spatially structured random effects in place of the zero mean, normally distributed independent random effects.

[19]This correlation differs from that latent in the geographic distributions of the origin and destination variables that are reflected in the balancing factors. Pace et al. (2011) show that spatial

to account for spatial dependence in model specifications may lead to biased parameter estimates and incorrect conclusions. One way to overcome this problem is by incorporating spatial dependence into the Poisson versions of the constrained gravity model variants.[20] Another way to address spatial dependence in origin-destination flows involves eigenvector spatial filtering[21] (see Chun 2008; Fischer and Griffith 2008; Griffith 2009; Chun and Griffith 2011). Eigenvector spatial filtering relies on a spectral decomposition[22] of an N-by-N spatial weight matrix W into eigenvalues and eigenvectors, and then uses a subset of these eigenvectors as additional explanatory variables in the model specifications.

Spatial filtering used here in this paper relies on a spectral decomposition of a transformed spatial weight matrix MWM, where W is an N-by-N spatial weight matrix[23]

$$W = W_n \otimes W_n \tag{3.29}$$

that captures spatial dependence between flows from locations neighbouring both the origins and destinations, labelled origin-to-destination dependence by LeSage and Pace (2008). W_n is a row-stochastic n-by-n spatial weight matrix that describes spatial neighbourhood relationships between the n locations. This matrix has—by convention—zeros in the main diagonal, and non-negative elements in the off-diagonal cells. Specifically the (i,j)th element of W_n is greater than zero if i and j are neighbouring[24] locations. $\otimes$ denotes the Kronecker product. M is the N-by-N

dependence in the explanatory variables decreases the ability of filtering to produce unbiased regression parameter estimates.

[20]In the fixed effects case of the doubly constrained gravity model, for example, this takes the form

$$E\left(Y_{ij}\right) = \mu_{ij} = U_i\, V_j \exp\left[\alpha + \sum_{h=1}^{n-1} I_{iho}\, \beta_{ho} + \sum_{k=1}^{n-1} I_{jkd}\, \beta_{kd} - \theta\, d_{ij}\right] \prod_{j \neq i}^{n} E\left(Y_{ij}\right)^{\rho W_{ij}}$$

where W_{ij} is the (i,j)th element of an N-by-N spatial weight matrix W and ρ is a scalar parameter that governs the degree of spatial dependence in origin-destination flows. Lambert et al. (2010) set forth a two-step maximum likelihood estimation approach for a spatial autoregressive Poisson model for count data which would need to be extended to the case of flows involving N observations.

[21]This is an especially valuable approach in situations where the flows are count data, because conventional spatial regression models and software tools are less developed for this data type.

[22]We assume that W is similar to a symmetric matrix so that it has real eigenvalues. If W is not symmetric, then $\frac{1}{2}\,(W + W')$, which is symmetric by construction, may be used.

[23]If intralocational flows are excluded from an analysis, the N-by-N spatial weight matrix reduces to an $n(n–1)$-by-$n(n–1)$ one, only marginally impacting upon these eigenvectors when $n > 100$.

[24]Neighbours may be defined using contiguity or measures of spatial proximity such as cardinal distance (for example, in terms of transportation costs) or ordinal distance (for example, the six nearest neighbours). In the illustrative example in Sect. 3.6, we use a binary contiguity matrix W_n to define W.

projection matrix

$$M = I - \iota\, \iota' \frac{1}{N} \tag{3.30}$$

where I is the N-by-N identity matrix, and ι the N-by-1 vector of ones.

The approach focuses on capturing correlations among flows with a spatial filter that is constructed as a linear combination of eigenvectors extracted from the matrix MWM. The eigenvalues scaled by $N/(\iota' W_n \iota)$ directly indicate Moran's I coefficient values of map patterns that are represented by the corresponding eigenvectors (Tiefelsdorf and Boots 1995). The first eigenvector, say E_1, is the set of real numbers that has the largest Moran's I value achievable by any set of real numbers for the spatial dependence structure defined by the spatial weight matrix W. The second eigenvector, E_2, is the set of real numbers that has the largest achievable Moran's I value by any set that is orthogonal to and uncorrelated with E_1. The third eigenvector is the third such set of values, and so on through E_N, the set of real numbers that has the largest negative Moran's I value achievable by any set that is orthogonal to and uncorrelated with the preceding $(N-1)$ eigenvectors. As such, these eigenvectors furnish distinct map pattern descriptions of latent spatial dependence in the origin-destination flow variable because they are both orthogonal and uncorrelated. Their Moran's Is corresponding eigenvalues index the nature and degree of spatial dependence portrayed by each eigenvector (Tiefelsdorf and Boots 1995), which can be standardised by the largest Moran's I value, $I_{\max}$.

The construction of a spatial filter involves a stepwise selection process. Griffith (2003) suggests identifying a set of candidate eigenvectors first, based on a critical value for the corresponding eigenvalues, a value that indicates a specific minimum spatial autocorrelation level[25] such as 0.5 measured in terms of the statistic $I/I_{\max}$. From these candidate eigenvectors, a subset of Q eigenvectors then can be selected with standard model selection criteria such as the Akaike information criterion. In the doubly constrained case of spatial interaction, for example, this yields the following spatial filter versions of the model specifications (3.12), (3.19) and (3.26), respectively:

$$E\left(Y_{ij}\right) = A_i\, B_j\, U_i\, V_j \exp\left(\alpha - \theta\, d_{ij} + \sum_{q=1}^{Q} E_q\, \phi_q\right) \tag{3.31}$$

$$E\left(Y_{ij}\right) = \mu_{ij} = U_i\, V_j \exp\left[\alpha + \sum_{h=1}^{n-1} I_{iho}\, \beta_{ho} + \sum_{k=1}^{n-1} I_{jkd}\, \beta_{kd} - \theta\, d_{ij} + \sum_{q=1}^{Q} E_q\, \phi_q\right] \tag{3.32}$$

[25]The criterion $I/I_{\max} = 0.5$ suggests a restriction of the search over eigenvectors with moderate to high spatial autocorrelation.

$$E\left(Y_{ij}\right) = \mu_{ij} = U_i\ V_j \exp\left[\alpha + \xi_{io} + \xi_{jd} - \theta\ d_{ij} + \sum_{q=1}^{Q} E_q\ \phi_q\right] \tag{3.33}$$

where E_q denotes the qth eigenvector and ϕ_q its associated coefficient. The term $\exp\left(\sum_q E_q\ \phi_q\right)$ is called a spatial filter.

The approach provides a simple way of filtering the sample flow data for spatial dependence in an effort to mimic independent data amenable to standard estimation procedures, and hence to reduce potential bias in the estimation of coefficients associated with the explanatory variables. Spatial filtering, however, also faces computational challenges in situations involving a large sample of observations.[26]

3.5 Equivalency Relationships Between the Balancing Factors, Fixed Effects, and Random Effects

This section compares the three different model variants of constrained spatial interaction with each other. First, attention is shifted toward comparisons between model specifications with balancing factors and with fixed effects, and then between model specifications with balancing factors and with random effects, for both the doubly and singly constrained cases of spatial interaction.

3.5.1 Comparisons Between Balancing Factors and Fixed Effects

The first comparison is between the doubly constrained model with balancing factors and its corresponding fixed effects model specification, and hence focuses on the relationship between Eqs. (3.31) and (3.32).

Theorem 1 *If $Y_{ij} \sim$ Poisson with mean $\mu_{ij} = \exp(\log U_i + \log V_j + \alpha + \alpha_{io} + \alpha_{jd} - \theta\ d_{ij} + + \sum_q E_q\ \phi_q)$, where α denotes the global Poisson regression intercept term, α_{io} the Poisson regression origin location intercept term, and α_{jd} the Poisson regression destination location specific intercept term, then the balancing factors for the doubly constrained gravity model are given by $A_i = \exp(\alpha_{io})$ and $B_j = \exp(\alpha_{jd})$.*

[26] Pace et al. (2011) demonstrate how using iterative eigenvalue routines on sparse matrices such as W can make filtering feasible for data sets involving a million or more observations, and empirically estimate an operation count on the order of $N^{1.1}$.

Proof Because Eqs. (3.31) and (3.32) posit the expectation for the same random variable, Y_{ij}, for $i, j = 1, \ldots, n$

$$A_i\, B_j\, U_i\, V_j \exp\left(\alpha - \theta\, d_{ij} + \sum_{q=1}^{Q} E_q\, \phi_q\right)$$

$$= U_i\, V_j \exp\left(\alpha + \beta_{io} + \beta_{jd} - \theta\, d_{ij} + \sum_{q=1}^{Q} E_q\, \phi_q\right) \tag{3.34}$$

$$A_i\, B_j = \exp\left(\beta_{io} + \beta_{jd}\right) = \exp\left(\beta_{io}\right) \exp\left(\beta_{jd}\right) \tag{3.35}$$

$$\therefore A_i = \exp\left(\beta_{io}\right) = \exp\left(\alpha_{io}\right) \text{ for all } i = 1, \ldots, n \tag{3.36}$$

$$B_j = \exp\left(\beta_{jd}\right) = \exp\left(\alpha_{jd}\right) \text{ for all } j = 1, \ldots, n \;\square \tag{3.37}$$

Remarks The equivalencies $\alpha_{io} = \beta_{io}$ and $\alpha_{jd} = \beta_{jd}$ relate these results not only to the doubly, but also to the singly constrained cases. Furthermore, $\exp(\alpha)$ is the constant of proportionality, which frequently is set to one (i.e., $\alpha = 0$) for the traditional entropy maximising solution, and other than one for the conventional gravity model solution. Allowing α to deviate from one in the Deming-Stephan-Furness procedure helps to stabilise convergence for large flow matrices, and may be achieved by setting the largest A_i and the largest B_j values to one during each iteration. This adjustment is equivalent to setting one of the $\alpha_{io} = \beta_{io}$ and one of the $\alpha_{jd} = \beta_{jd}$ to zero in the fixed effects specification in order to avoid perfect multicollinearity between the location specific indicator variables and the global mean (which is a coefficient times an n-by-1 vector of ones). Estimates of β_{io} and β_{jd} are obtained with Poisson regression.

This result relates the log-balancing factors, $\log(A_i)$ and $\log(B_j)$, for the doubly constrained gravity model to their counterpart origin and destination fixed effects, α_{io} and α_{jd}. Hence fixed effects take on a particular meaning because they can be interpreted as balancing factors. Cesario (1977) characterises the meaning of the origin and destination balancing factors as follows: $1/A_i$ indexes the accessibility of all destination locations vis-à-vis origin i, and $1/B_j$ indexes the accessibility of all origin locations vis-à-vis destination j.

The next comparisons are between the model specifications with balancing factors and fixed effects in the singly constrained cases of spatial interaction. To this end, **Theorem 1** suggests the following two corollaries pertaining to the singly constrained spatial filter model specifications.

Corollary 1 *If $Y_{ij} \sim$ Poisson with mean $\mu_{ij} = \exp(\log U_i + \alpha + \alpha_{io} - \theta\, d_{ij} + \sum_q E_q\, \phi_q)$, where α denotes the global Poisson regression intercept term, and α_{io} the Poisson regression origin location specific intercept term, then the balancing*

factors for the origin constrained gravity model are given by $A_i = \exp(\alpha_{io})$ *for* $i = 1, \ldots, n$.

Corollary 2 *If* $Y_{ij} \sim$ *Poisson with mean* $\mu_{ij} = \exp(\log V_j + \alpha + \alpha_{jd} - \theta\ d_{ij} + \sum_q E_q\ \phi_q)$, *where* α *denotes the global Poisson regression intercept term, and* α_{jd} *the Poisson regression destination location specific intercept term, then the balancing factors for the destination constrained gravity model are given by* $B_j = \exp(\alpha_{jd})$ *for* $j = 1, \ldots, n$.

These two corollaries relate the log-balancing factors for the singly constrained gravity models to their counterpart origin or destination fixed effect model specifications.

3.5.2 Comparisons Between Balancing Factors and Random Effects

Finally, comparisons can be made between the preceding results and the random effects model specifications. In this context, a specification includes ξ_{io} and/or ξ_{jd}, normal random variables, which respectively denote the random effects for origin i and/or destination j, whose stochastic quantities are added to the global intercept term. The next theorem relates to the relationship between a doubly constrained model with balancing factors and its corresponding random effects model specification, and hence focuses on the relationship between Eqs. (3.31) and (3.33).

Theorem 2 *If* $Y_{ij} \sim$ *Poisson with mean* $\mu_{ij} = \exp(\log U_i + \log V_j + \alpha + \xi_{io} + \xi_{jd} - \theta\ d_{ij} + \sum_q E_q\ \phi_q)$, *where* α *denotes the global Poisson regression intercept term,* ξ_{io} *the Poisson regression origin location random effect, and* ξ_{jd} *the Poisson regression destination location random effect, such that* $\xi_{io} \sim \mathcal{N}(0, \sigma^2_{\xi_o})$ *and* $\xi_{jd} \sim \mathcal{N}(0, \sigma^2_{\xi_d})$, *where* $\sigma^2_{\xi_o}$ *and* $\sigma^2_{\xi_d}$ *denote the origin and the destination location random effects variances respectively, then the balancing factors for the doubly constrained gravity model are given by* $A_i = \exp(\xi_{io})$ *for* $i = 1, \ldots, n$, *and* $B_j = \exp(\xi_{jd})$ *for* $j = 1, \ldots, n$.

Proof Equation (3.33) implies

$$\log\left[E\left(Y_{ij}\right)\right] = \log \mu_{ij} = \log U_i + \log V_j + \alpha + \xi_{io} + \xi_{jd} - \theta\ d_{ij} + \sum_{q=1}^{Q} E_q\ \phi_q. \tag{3.38}$$

Because Eqs. (3.31) and (3.38) posit the expectation for the same random variable, Y_{ij}, for $i, j = 1, \ldots, n$

$$A_i\, B_j\, U_i\, V_j \exp\left(\alpha - \theta\, d_{ij} + \sum_{q=1}^{Q} E_q\, \phi_q\right)$$

$$= U_i\, V_j \exp\left(\alpha + \xi_{io} + \xi_{jd} - \theta\, d_{ij} + \sum_{q=1}^{Q} E_q\, \phi_q\right) \tag{3.39}$$

$$A_i B_j = \exp(\xi_{io} + \xi_{jd}) = \exp(\xi_{io})\ \exp(\xi_{jd}) \tag{3.40}$$

$\therefore A_i = \exp(\xi_{io})$ for all $i = 1, \ldots, n$ and $B_j = \exp(\xi_{jd})$ for all $j = 1, \ldots, n$ ■

Remarks Both ξ_{io} and ξ_{jd} have a mean of zero, which is achieved by having the global mean, α, in the model specification. In other words, the individual origin and destination location means deviate from the global mean by random quantities. **Theorems 1** and **2** together imply: $A_i = \exp(\xi_{io}) = \exp(\beta_{io}) = \exp(\alpha_{io})$ for all $i = 1, \ldots, n$, and $B_j = \exp(\xi_{jd}) = \exp(\beta_{jd}) = \exp(\alpha_{jd})$ for all $j = 1, \ldots, n$. Estimates of ξ_{io} and ξ_{jd} are obtained by integrating them out of the likelihood function.

Singly constrained models are obtained by setting $\xi_{jd} = 0$ for all j, yielding the origin constrained specification, or $\xi_{io} = 0$ for all i, yielding the destination constrained specification. Accordingly, **Theorem 2** suggests the following two corollaries pertaining to the singly constrained model specifications.

Corollary 3 *If $Y_{ij} \sim$ Poisson with mean $\mu_{ij} = \exp(\log U_i + \alpha + \xi_{io} - \theta\ d_{ij} + \sum_q E_q\, \phi_q)$, where α denotes the global Poisson regression intercept term, and ξ_{io} the Poisson regression origin location random effect, such that $\xi_{io} \sim \mathcal{N}(0, \sigma^2_{\xi_o})$, where $\sigma^2_{\xi_o}$ denotes the origin location finite random effects variance, then the balancing factors for the origin constrained gravity model are given by $A_i = \exp(\xi_{io})$ for $i = 1, \ldots, n$.*

Corollary 4 *If $Y_{ij} \sim$ Poisson with mean $\mu_{ij} = \exp(\log V_j + \alpha + \xi_{jd} - \theta\ d_{ij} + \sum_q E_q\, \phi_q)$, where α denotes the global Poisson regression intercept term, and ξ_{jd} the Poisson regression destination location random effect, such that $\xi_{id} \sim \mathcal{N}(0, \sigma^2_{\xi_d})$, where $\sigma^2_{\xi_d}$ denotes the destination location finite random effects variance, then the balancing factors for the destination constrained gravity model are given by $B_j = \exp(\xi_{jd})$ for $j = 1, \ldots, n$.*

3.6 An Illustrative Example

In this section we use knowledge flows as captured by patent citation data to numerically illustrate the relationships between the aforementioned balancing factors, fixed effects and random effects in the cases of singly and doubly constrained variants of the gravity model. The origin-destination data relate to citations between European high-technology patents. By European patents we mean patent applications at the European Patent Office assigned to high-technology firms located in Europe. High-technology is defined to include 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 (that is, citations from patents assigned to the same firm) have been excluded, given our interest in pure externalities as evidenced by interfirm knowledge spillovers.

Experts acknowledge that observations of patent citations are subject to a truncation bias, because we observe citations for only a portion of the life of an invention. To avoid this bias in the analysis, we have established a 5-year window (that is, 1985–1989, 1986–1990, . . . , 1993–1997) to count citations to a patent.[27] The observation period is 1985–1997 with respect to cited patents, and 1990–2002 with respect to citing patents. The sample used in this section is restricted to inventors located in $n = 257$ European NUTS-2 regions, covering the EU-27 member states (excluding Cyprus and Malta) plus Norway and Switzerland. In case of cross-regional inventor teams, we have used the procedure of multiple full counting that—unlike fractional counting—does justice to the true integer nature of patent citations, but gives interregional cooperative inventions greater weight.

Subject to caveats relative to the relationship between patent citations and knowledge spillovers, the sample data allow us to identify and measure spatial separation effects for interregional knowledge spillovers in the spatial system of 257 regions. We use a binary 257-by-257 contiguity matrix to specify the 66,049-by-66,049 spatial weight matrix W that captures spatial dependence between patent citation flows from locations neighbouring both the origins and the destinations. Our interest is focused on the following three measures of separation: geographical distance, measured in terms of the great circle distance (in km), technological proximity, measured in terms of an index (for details see Fischer et al. 2006), and a dummy variable that represents border effects measured in terms of the presence of country borders between the regions. The product $U_i V_j$ may be interpreted simply as the number of distinct (i, j) interactions that are possible. Thus, a reasonable way to measure the origin factor U_i is in terms of the number of patents in knowledge producing region i in the time period 1985–1997, and the destination factor V_j in terms of the number of patents in knowledge absorbing region j in the time period

[27] For details about the data construction, see Fischer et al. (2006).

1990–2002 (Fischer and Griffith 2008). Accordingly, we have 66,049 observations, five (four) covariates and an intercept term in the doubly (singly) constrained cases of spatial interaction.

3.6.1 Model Specifications Ignoring Spatial Dependence in origin-destination Flows

Empirical experiments were conducted to numerically illustrate relationships between the aforementioned balancing factors, fixed effects, and random effects. The preceding theorems and corollaries indicate that these should be perfectly straight trend line relationships (using the log-balancing factors) with a slope of one, but not necessarily an intercept of zero. The intercept term represents an arbitrary multiplicative factor (i.e., a constant of proportionality).

Theorem 1 together with **Corollaries 1** and **2** indicate that the scatterplots of the log-balancing factors versus their concatenated Poisson regression indicator variable coefficients (augmented by zero for the arbitrarily removed indicator variables) form a perfect straight line [see Fig. 3.1]. The accompanying linear regression equations[28] relating these two pairings of values are as follows[29]: for the origin constrained case of spatial interaction: $\log(A_i) = 0.00051 + 0.99997\ \alpha_{io}$ ($R^2 = 1.0000$), the destination constrained case of spatial interaction: $\log(B_j) = 0.00001 + 0.99999\ \alpha_{jd}$ ($R^2 = 1.0000$), and the doubly constrained case of spatial interaction: $\log(A_i) = 0.00018 + 1.00114\ \alpha_{io}$ ($R^2 = 1.0000$) and $\log(B_j) = 0.00099 + 1.00110\ \alpha_{jd}$ ($R^2 = 1.0000$). Furthermore, these log-balancing factors strongly covary [see Fig. 3.2(a) and (b)], and all deviate somewhat from a normal frequency distribution as indicated by Fig. 3.2(c)–(f).

Model comparison results for the fixed effects versions of the constrained models are presented in Table 3.1. Inclusion of the origin and/or destination balancing factors as fixed effects covariates reduces overdispersion as indicated by the deviance statistic,[30] noticeably changes the three separation function component parameter estimates (especially that for the geographical distance decay), and remarkably increases the pseudo-R^2 value (measured in terms of a linear relationship between the predicted and observed counts). The last column in Table 3.1 presents estimation results for the doubly constrained spatially filtered gravity model specification, for

[28] A_{257} and B_{257} are the arbitrarily selected balancing factors set to one in each case, to avoid perfect multicollinearity with the intercept term, resulting in an expected intercept of zero and an expected slope of one.

[29] The regression equations describe each set of log-balancing factors as a function of the corresponding fixed effects indicator variables. Error terms are not included here.

[30] A deviance statistic exceeding one indicates that overdispersion is present; that is, the Poisson variance is greater than its mean. Although the existence of overdispersion does not affect the unbiased character of the parameter estimates, their standard errors are underestimated, and hence their significance is unrealistically increased.

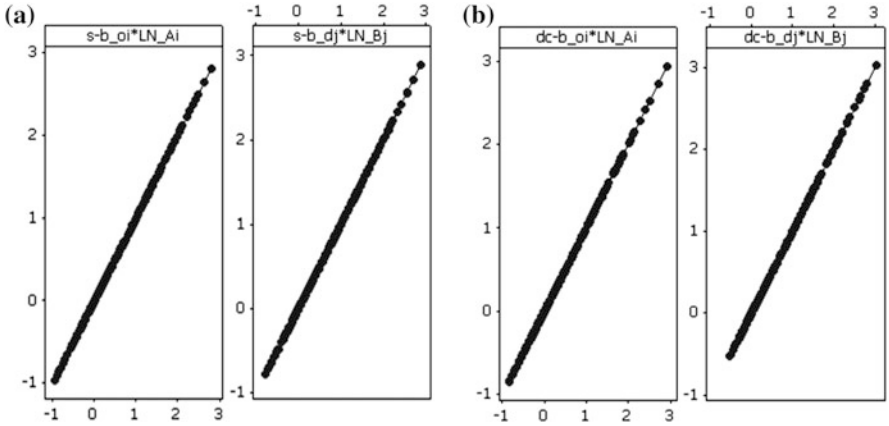

Fig. 3.1 Scatterplots of the log-balancing factors [$\log(A_i)$ and $\log(B_j)$] versus the vector of the Poisson regression indicator variable coefficients: (**a**) the singly-constrained cases: origin and destination balancing factor plots; and (**b**) the doubly constrained case: the origin and the destination balancing factor plots

comparative purposes. The elimination of spatial dependence in the flows triggers a change of estimated parameter values and generates a decrease in the estimated overdispersion, compared with the standard doubly constrained model specification. That is, a part of the overdispersion, caused by spatial dependence, is eliminated by including eigenvectors, which are the proxy variables of the spatial dependence embedded in the standard model.

Theorem 2 together with **Corollaries 3** and **4** address the random effects model specifications for the three constrained variants of the gravity model. Treated as particular cases of a GLM with a logarithmic link function and a Poisson mean, these specifications yield the following expected values: $\log\left[E(Y_{ij})\right] = \log(\mu_{ij}) = \log(U_i) + \log(V_j) + \alpha + \xi_{io} + \xi_{jd} - \theta\ d_{ij}$ in the doubly constrained case of spatial interaction, $\log\left[E(Y_{ij})\right] = \log(\mu_{ij}) = \log(U_i) + \alpha + \xi_{io} - \theta\ d_{ij}$ in the origin constrained case, and $\log\left[E(Y_{ij})\right] = \log(\mu_{ij}) = \log(V_j) + \alpha + \xi_{jd} - \theta\ d_{ij}$ in the destination constrained case. The log terms on the right-hand side of the equations are the offset variables. Bolduc et al. (1995) argue that estimating origin and destination specific random effects in the gravity model specification is very difficult. But the implication from **Theorem 2** for the doubly constrained specification supports a numerical demonstration for it, too.

Descriptive statistics for the random effects estimates are given in Table 3.2. A frequentist approach requires integration of these effects out of the likelihood function under study. As n increases, the multidimensional integration involved becomes increasingly difficult. Here, with $n = 257$, the SAS procedure, called SAS PROC NLMIXED, fails to correctly calculate about 10 % of the random effects (see the Appendix). This complication resulted in the design of an indirect demonstration of **Theorem 2** as follows. Each balancing factor was introduced into the model

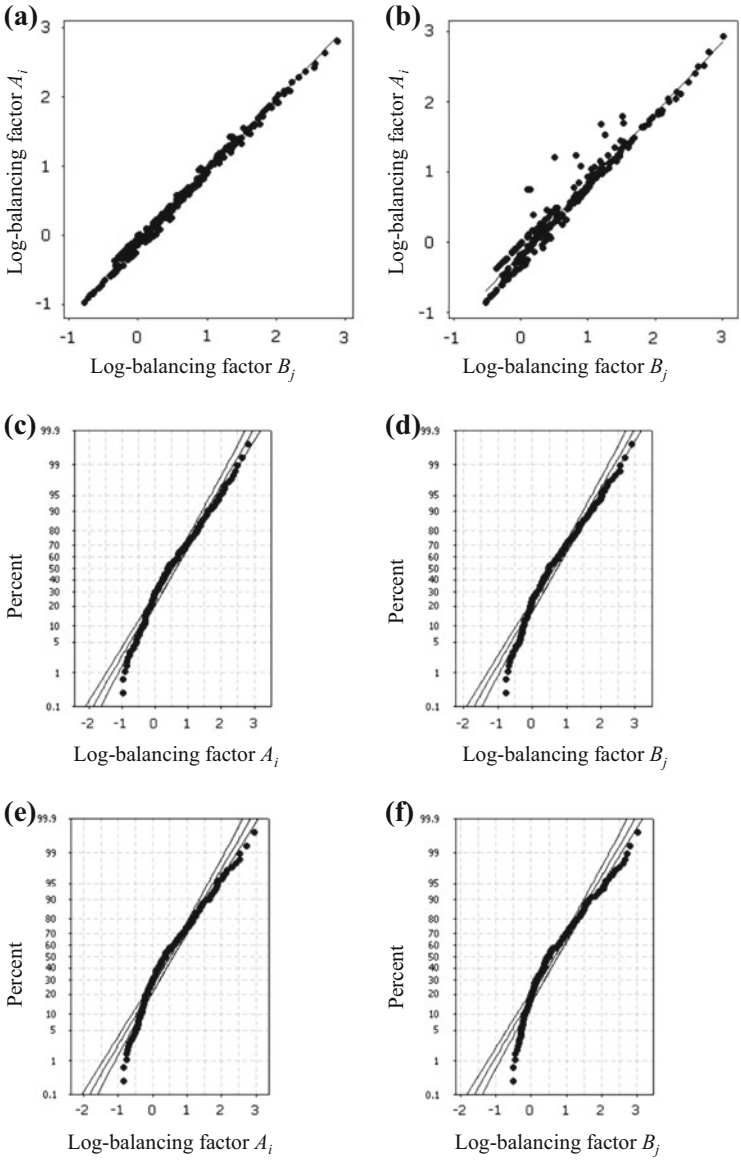

Fig. 3.2 Log-balancing factors for the constrained model variants: (**a**) scatterplot of the singly constrained origin and destination log-balancing factor pairs; (**b**) scatterplot of the doubly constrained origin and destination log-balancing factor pairs; (**c**) normal quantile plot of $\log(A_i)$ values in the origin-constrained case, with its 95 % confidence interval (CI); (**d**) normal quantile plot of $\log(B_j)$ values in the destination-constrained case, with its 95 % CI; (**e**) normal quantile plot of $\log(A_i)$ values in the doubly-constrained case, with its 95 % CI; and (**f**) normal quantile plot of $\log(B_j)$ values in the doubly-constrained case, with its 95 % CI

Table 3.1 Summary statistics for estimation of the fixed effects constrained variants of the gravity model (standard deviations in parentheses)

Statistics	Model specifications ignoring spatial dependence				Spatially filtered model specification
	Unconstrained[a]	Origin-constrained	Destination-constrained	Doubly-constrained	Doubly-constrained
Separation					
Geog. distance	−0.4750(0.0070)	−0.8317(0.0080)	−0.7539(0.0079)	−0.9092(0.0080)	−1.4992(0.0263)
Technology	−2.3407(0.0098)	−2.6041(0.0117)	−2.6845(0.0118)	−2.9788(0.0118)	−2.4919(0.0210)
Border	−1.1386(0.0049)	−1.3029(0.0058)	−1.2585(0.0057)	−1.2628(0.0057)	−0.9091(0.0136)
Deviance	4.0261	3.0574	3.2158	2.5924	1.8135
Residual mean	−0.0000	−0.0000	−0.0000	−0.0000	0.0000
Residual P(K-S)[b]	<0.01	<0.01	<0.01	<0.01	<0.01
Pseudo-R^2	0.5828	0.7221	0.7152	0.7323	0.8914

[a]The row and column totals were included with a power exponent of one
[b]K-S denotes the Kolmogorov–Smirnov statistic, a diagnostic test for normality here

Table 3.2 Summary statistics for random effects estimations: the origin-constrained, the destination-constrained and the doubly-constrained cases

	Model specifications given by Eqs. (3.14)–(3.16)			
			Doubly-constrained	
Statistics	Origin-constrained	Destination-constrained	Origin	Destination
Minimum	$-17.2 \cdot 10^{-13}$	$-15.5 \cdot 10^{-13}$	$-10.2 \cdot 10^{-13}$	$-10.5 \cdot 10^{-13}$
Mean	$-2.5 \cdot 10^{-13}$	$-2.7 \cdot 10^{-13}$	$-83.6 \cdot 10^{-13}$	$-2.3 \cdot 10^{-13}$
Median	$-2.9 \cdot 10^{-13}$	$-3.0 \cdot 10^{-13}$	$-0.1 \cdot 10^{-13}$	$-3.1 \cdot 10^{-13}$
Maximum	$10.9 \cdot 10^{-13}$	$9.8 \cdot 10^{-13}$	$11.0 \cdot 10^{-13}$	$10.4 \cdot 10^{-13}$
Standard deviation	$4.1 \cdot 10^{-13}$	$4.1 \cdot 10^{-13}$	$4.5 \cdot 10^{-13}$	$4.2 \cdot 10^{-13}$
P(Shapiro–Wilk)	<0.0001	0.0274	0.0077	<0.0001

GLMM estimation utilized a Newton–Raphson optimization procedure. The Shapiro–Wilk statistic furnishes a diagnostic statistic for normality

specification, and then a random effects term was estimated. If a balancing factor is equivalent to a random effects term, then all of the estimated random effects are approximately zero. This expectation characterises the findings summarised in Table 3.2. In other words, the estimated fixed and random effects display no consequential differences. The generalised linear mixed model (GLMM) random effects estimates are nearly identical to their balancing factor fixed effects counterparts.

Spatial filter descriptions of these variates are nearly identical,[31] as shown in Table 3.3, and comprise 17–20 of the 42 candidate eigenvectors depicting at least weak positive spatial autocorrelation map patterns. These filters allow the balancing factors to be deconstructed into spatially structured (SSRE) and spatially unstructured (SURE) random effects: the linear combination of eigenvectors constitutes the SSRE, and the (remaining) residuals constitute the SURE. The SSREs account for roughly two-thirds of the variance displayed by the total random effects terms. This spatial structuring represents moderate-to-strong positive spatial autocorrelation, and is one reason the individual terms deviate from a normal frequency distribution [all P(S-W) statistics increase, but still indicate marked deviation from a normal distribution]. These linear spatial filters account for virtually all of the spatial autocorrelation latent in the spatial distribution of these balancing factor variates, which differ from the spatial dependence latent in the flows between the regions.

[31] Because balancing factors are autoregressive specifications [see Eqs. (3.13)–(3.14)], they contain marked spatial dependence by construction. The spatial filter descriptions of these balancing factors rely on eigenvectors of the transformed spatial weight matrix $M_n W_n M_n$ where W_n is the n-by-n binary contiguity matrix and M_n is the n-by-n projection matrix defined by $M_n = I_n - \iota_n \ \iota_n' \ n^{-1}$. Forty-two candidate eigenvectors (for which $I/I_{\max} > 0.25$) are available for constructing spatial filters portraying positive spatial autocorrelation across the European regions. Of these, subsets have been selected with a stepwise regression procedure for constructing spatial filters describing the two sets of balancing factors. The criteria used for selection were statistically significant coefficients at the 10 % level associated with minimisation of the log-likelihood function, which is standard practice.

Table 3.3 Summary statistics for the balancing factors and the decomposition

	Log-balancing factor		Spatially structured random effects			Spatially unstructured random effects	
Constraint	I[a]	P(S-W)[b]	I[a]	# vectors	R^2	z_I[c]	P(S-W)[b]
Singly-constrained origin	0.65	0.0002	0.93	18	70	0.36	0.0009
Singly-constrained destination	0.61	<0.0001	0.93	17	67	0.13	0.0007
Doubly-constrained origin	0.63	<0.0001	0.93	20	69	−0.29	0.0006
Doubly-constrained destination	0.61	<0.0001	0.93	20	67	−0.29	0.0032

[a] I denotes the Moran's coefficient
[b] S-W denotes the Shapiro–Wilk statistic
[c] The asymptotic standard error for the Moran's I was used to compute the z-scores

Consequently, these particular singly constrained gravity model results confirm **Corollaries 3** and **4**, and as such indirectly demonstrate **Theorem 2**. They also illustrate that $A_i = \exp(\xi_{io}) = \exp(\beta_{io}) = \exp(\alpha_{io})$ for $i = 1, \ldots, n$, and $B_j = \exp(\xi_{jd}) = \exp(\beta_{jd}) = \exp(\alpha_{jd})$ for $j = 1, \ldots, n$. In other words, the model specifications with balancing factors, fixed effects and random effects, respectively, yield identical estimation results for the production constrained and the attraction constrained cases of spatial interaction. These findings imply that the same results hold for the doubly constrained case (Fig. 3.3).

3.6.2 Spatial Filter Model Specifications Accounting for Spatial Dependence in Flows

Estimating the balancing factors for singly and doubly constrained model specifications accounts for spatial dependence in the origin and destination factors of the gravity model, but not for spatial dependence in flows. Because only one set of indicator variables is involved in singly constrained model specifications, the intercept term can be added to each factor, forcing α to zero in the origin constrained model specification, and in the destination constrained model specification, respectively. This simple adjustment is not possible for the doubly constrained model, for which the intercept term includes the sum of the two arbitrarily selected indicator variable coefficients set to zero. Estimating random effects in the doubly constrained case also overlooks spatial dependence in flows, and treats the n origin flow recipients as repeated measures for each destination, and the n destination

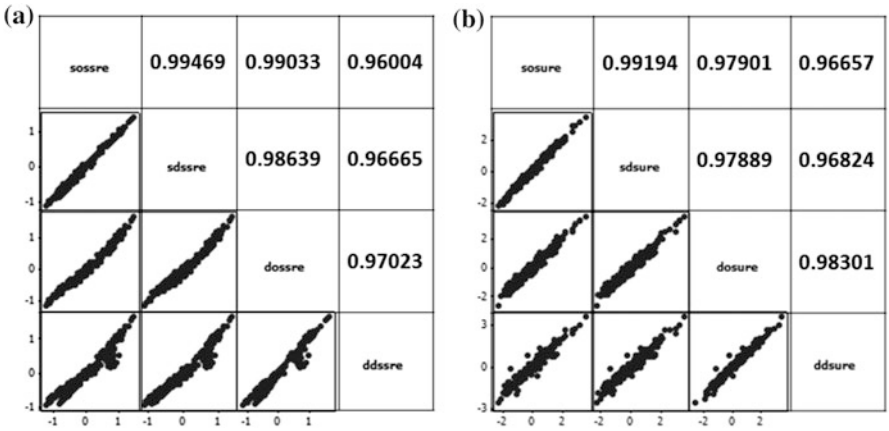

Fig. 3.3 Matrix *lower triangular* scatterplots, and *upper triangular* correlations: (**a**) spatially structured random effects (SSREs): sossre—singly-constrained origin, sdssre—singly-constrained destination, dossre—doubly-constrained origin, and ddssre—doubly-constrained destination; and (**b**) spatially unstructured random effects (SUREs): sossre—singly-constrained origin, sdssre—singly-constrained destination, dossre—doubly-constrained origin, and ddssre—doubly-constrained destination

flow sources as repeated measures for each origin, respectively. All of these specifications posit a unique value for each origin/destination for the $N = n^2$ flow data.

Origin and destination balancing factors must be estimated simultaneously (not sequentially) with the spatial filters, in order to preserve the row and column constraining totals. For the current case study, the spatial filters represent moderate-to-strong positive spatial autocorrelation ($I \approx 0.70$), decrease overdispersion by a third or more beyond the reduction attributable to the balancing factors (Table 3.1), produce a modest increase in the pseudo-R^2 value, induce a marked decrease in the distance decay parameter (for example, the confidence interval does not overlap with those for the other specifications), and comprise $Q = 221$ of the 576 candidate eigenvectors[32] of matrix W.

Figure 3.4 reports the scatterplots of observed versus predicted flows for the unconstrained gravity model specification, and the doubly constrained gravity model specification with and without accounting for origin-to-destination depen-

[32]Of note is that for n larger than about 100, current computer resources do not allow direct calculation of the eigenvectors of W. In order to reduce computational intensity we, followed Griffith (2009) to construct the spatial filter with a linear combination of Kronecker products of pairs of origin and destination eigenvectors. The result of this adjustment is $24^2 = 576$ candidate eigenvectors identified as Kronecker products of the 24 eigenvectors with an $I > 0.5$ extracted from matrix $(I - \iota\,\iota'\,n^{-1})\,W_n\,(I - \iota\,\iota'\,n^{-1})$. With 66,049 observations, five covariates and an intercept term, and 576 candidate eigenvectors, the numerical intensity of the problem solution becomes feasible but is still high.

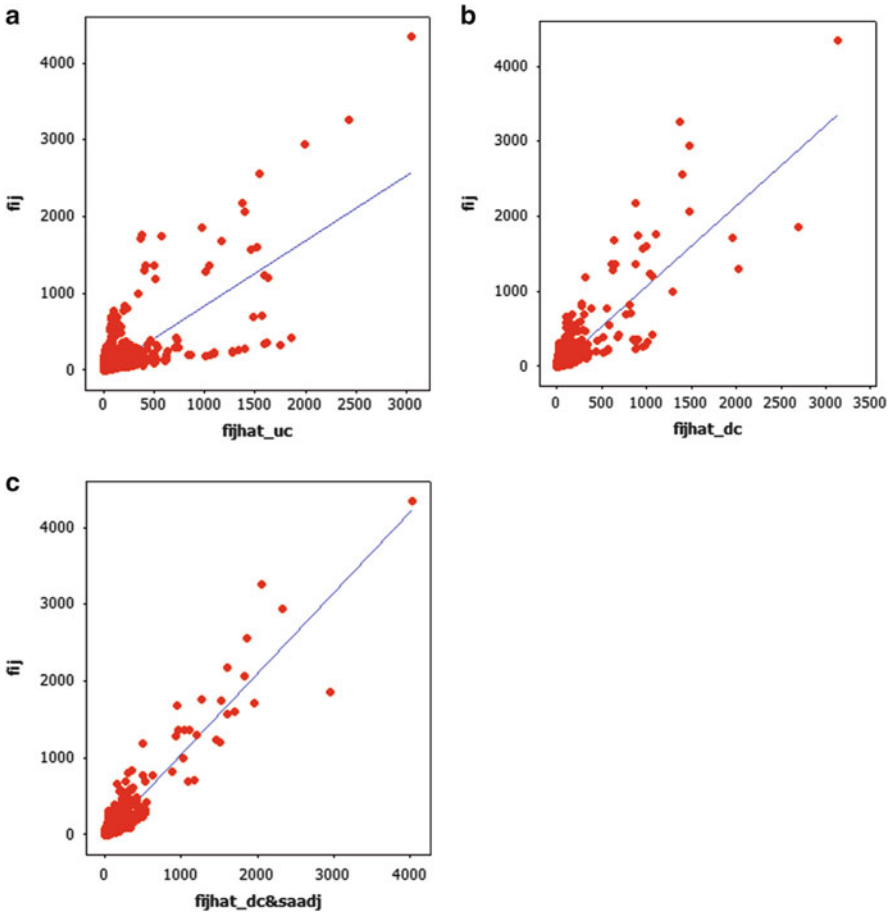

Fig. 3.4 Scatterplots of observed (vertical axis) versus predicted (horizontal axis) flows; *grey lines* denotes the line of perfect prediction. (**a**) Unconstrained model specification; (**b**) doubly constrained model specification; and (**c**) doubly constrained model specification adjusting for spatial dependence

dence in the flows. The scatterplots display a standard Poisson random variable plot of increasing variance with increasing amount of flow, and indicate a sequentially improved alignment of predicted with observed values. Imposing flow data matrix row and/or column total constraints coupled with inclusion of a spatial filter capturing spatial dependence between flows from locations neighbouring both the origins and destinations during estimation, shrinks especially the larger predicted flow values toward the perfectly straight trend line.

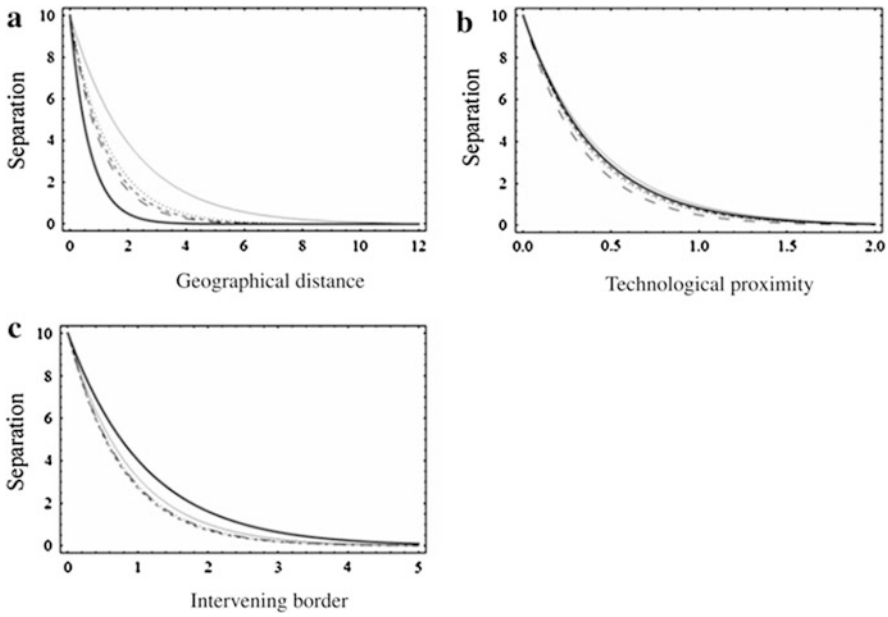

Fig. 3.5 Separation decay effects for the various model specifications: unconstrained (*thin line*), destination-constrained (*dotted line*), origin-constrained (*short dash line*), doubly-constrained (*long dash line*), doubly-constrained adjusted for network spatial autocorrelation (*thick line*). (**a**) geographical distance; (**b**) technological distance; and (**c**) intervening border

Figure 3.5 portrays the three individual separation effects. An expected finding is that the geographical distance decay parameter estimate adjusted for spatial dependence is less than in the model specifications that ignore spatial dependence in flows. And, it differs substantially from its unadjusted counterparts [see Fig. 3.5a]. The pairs of values do not have overlapping confidence intervals (CIs), in part because of the large sample size. The CI for the unconstrained case of spatial interaction is (−0.4887, −0.4613), the origin constrained case (−0.8473, −0.8161), the destination constrained case (−0.7695, −0.7384), the doubly constrained case (−0.9249, −0.8934), and the spatially filtered doubly constrained case (−1.5507, −1.4476). The technological separation decay parameter estimate exhibits little difference across the specifications [see Fig. 3.5b]. And, ignoring spatial dependence appears to exaggerate border separation effects [see Fig. 3.5c]. Of note is that the geographical distance parameter estimate has the largest spread across the model specifications.

3.7 Concluding Remarks

This paper suggests a number of interesting conclusions and implications for the statistical analysis of origin-destination data. Foremost, and quite counterintuitive, fixed effects and random effects are identical and equal the logarithm of the entropy maximisation derived balancing factors, except for slight rounding/algorithm-convergence errors. This finding is the outcome of an equivalency between assigning a single fixed effects indicator variable to each origin/destination, on the one hand, and estimating a single random effects (which is a mean) value for an origin/destination while treating the corresponding n destinations/origins as repeated measures, on the other hand. This finding also indicates that the number of degrees of freedom associated with the random effects term in this context may well be closer to $n - 1$ than to two (i.e., for estimating the mean and the variance of a random effects term) for the origins as well as the destination stochastic variable.

As with the unconstrained gravity model, adjusting for spatial dependence in flows improves the performance of the constrained variants of the gravity model in terms of both the pseudo-R^2 and the deviance statistic, and has a substantial impact on the separation parameter estimates that is in line with Curry (1972). The cost in degrees of freedom is modest. On average, at least 90 degrees of freedom are available for each parameter estimated in this case study. The eigenvectors successfully capture origin-to-destination dependence in flows. Hence, eigenvector spatial filtering provides a useful way of filtering spatial dependence in the sample origin-destination data. A virtue of this approach is that standard model specifications of the constrained gravity models and existing software can be applied to origin-destination data samples. This proves especially useful when dealing with flows taking the form of counts. However, the difficulty of computing eigenvalues and eigenvectors when dealing with a large number of locations limits the ability of filtering to capitalise on these advantages.

Appendix: Results for the Estimation of Singly Constrained Random Effects Specifications

Because of the large dimensionality of the calculus problem, multivariate integration struggles to properly estimate the random effects terms. Largest values appear to introduce the greatest difficulties. Figure 3.6 A reveals that integration is completely successful between the minimum and roughly 0.5 in our case study. Integration is only partially successful beyond 0.5. Incorrectly calculated random effects constitute about 10 % of the total number of random effects in this case study.

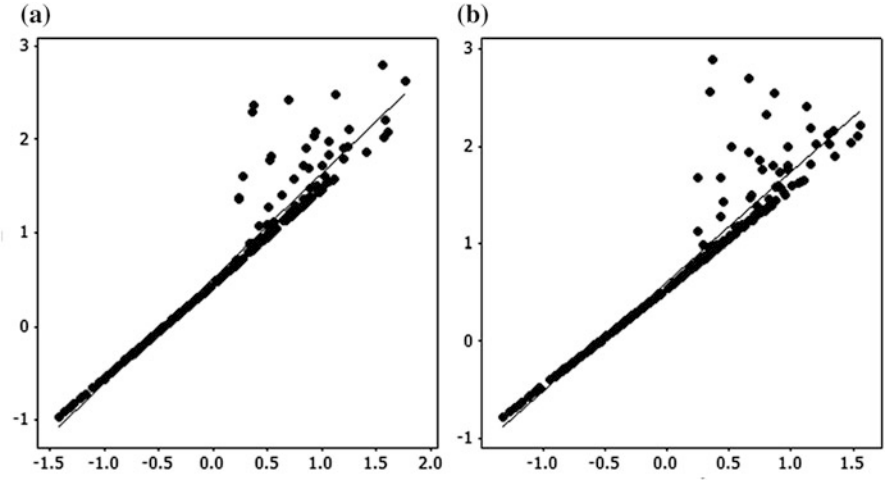

Fig. 3.6 Scatterplot of (**a**) the origin log-balancing factor (*vertical axis*) versus the Poisson regression origin location random effects (*horizontal axis*); (**b**) the destination log-balancing factor (*vertical axis*) versus the Poisson regression destination location random effects (*horizontal axis*)

References

Bailey TC, Gatrell AC (1995) Interactive spatial data analysis. Longman, Harlow

Baxter M (1982) Similarities in methods of estimating spatial interaction models. Geogr Anal 14(3):267–272

Baxter M (1983) Estimation and inference in spatial interaction models. Prog Hum Geogr 7:40–59

Bishop YMM, Fienberg SE, Holland PW (1975) Discrete multivariate analysis. MIT Press, Cambridge

Black WR (1992) Network autocorrelation in transport network and flow systems. Geogr Anal 24(3):207–222

Bolduc D, Laferrière R, Santarossa G (1995) Spatial autoregressive error components in travel flow models: an application to aggregate mode choice. In: Anselin L, Florax R (eds) New directions in spatial econometrics. Springer, Berlin/Heidelberg/New York, pp 96–108

Cesario FJ (1973) A generalized trip distribution model. J Reg Sci 13(2):233–248

Cesario FJ (1977) A new interpretation of the "normalizing" or "balancing factors" of gravity type spatial models. J Socio Econ Plan Sci 11(3):131–136

Chun Y (2008) Modeling network autocorrelation within migration flows by eigenvector spatial filtering. J Geogr Syst 10(4):317–344

Chun Y, Griffith DA (2011) Modeling network autocorrelation in space-time migration flow data: an eigenvector spatial filtering approach. Ann Assoc Am Geogr 101(3):523–536

Curry L (1972) A spatial analysis of gravity flows. Reg Stud 6(2):131–137

Davies RB, Guy CM (1987) The statistical modelling of flow data when the Poisson assumption is violated. Geogr Anal 19(4):300–314

Elhorst JP (2010) Spatial panel data models. In: Fischer MM, Getis A (eds) Handbook of applied spatial analysis. Springer, Berlin/Heidelberg/New York, pp 377–407

Evans AW (1970) Some properties of trip distribution models. Transp Res 4(1):19–36

Fischer MM (2000) Travel demand – theory. In: Polak JB, Heertje A (eds) Analytical transport economics. Edward Elgar, Cheltenham, pp 51–78

Fischer MM, Griffith DA (2008) Modeling spatial autocorrelation in spatial interaction data: a comparison of spatial econometric and spatial filtering specifications. J Reg Sci 48(5):969–989
Fischer MM, Wang J (2011) Spatial data analysis: models, methods and techniques [Springer briefs in regional science]. Springer, Berlin/Heidelberg/New York
Fischer MM, Scherngell T, Jansenberger E (2006) The geography of knowledge spillovers between high-technology firms in Europe: evidence from a spatial interaction modeling perspective. Geogr Anal 38(3):288–309
Fotheringham AS, O'Kelly ME (1989) Spatial interaction models: formulations and applications. Kluwer Academic Publishers, Dordrecht
Griffith DA (2003) Spatial autocorrelation and spatial filtering. Springer, Berlin/Heidelberg/New York
Griffith DA (2007) Spatial structure and spatial interaction: 25 years later. Rev Reg Stud 37(1):28–38
Griffith DA (2009) Modeling spatial autocorrelation in spatial interaction data: empirical evidence from 2002 Germany journey-to-work flows. J Geogr Syst 11(2):117–140
Haynes KE, Fotheringham AS (1984) Gravity and spatial interaction models. Sage, Bevery Hills
Isard W (1960) Methods of regional analysis. MIT Press, Cambridge
Kirby HR (1974) Theoretical requirements for calibrating gravity models. Transp Res 8(1):97–104
Lambert DM, Brown JP, Florax RJGM (2010) A two-step estimator for a spatial lag model of counts: theory, small sample performance and an application. Reg Sci Urban Econ 40(4):241–252
Ledent J (1985) The doubly constrained model of spatial interaction: a more general formulation. Environ Plan A 17(2):253–262
LeSage JP, Fischer MM (2010) Spatial econometric methods for modeling origin-destination flows. In: Fischer MM, Getis A (eds) Handbook of applied spatial analysis. Springer, Berlin/Heidelberg/New York, pp 409–433
LeSage JP, Pace RK (2008) Spatial econometric modeling of origin-destination flows. J Reg Sci 48(5):941–968
LeSage JP, Pace RK (2009) Introduction to spatial econometrics. CRC Press, Boca Raton
LeSage JP, Fischer MM, Scherngell T (2007) Knowledge spillovers across Europe. Evidence from a Poisson spatial interaction model with spatial effects. Pap Reg Sci 86(3):393–421
McCullagh P, Nelder JA (1983) Generalized linear models. Chapman and Hall, London
Niedercorn J, Bechdolt B (1969) An economic derivation of the "gravity law" of spatial interaction. J Reg Sci 9(2):273–282
Pace RK, LeSage JP, Zhu S (2011) Interpretation and computation of estimates from regression models using spatial filtering. Paper presented at the Fifth World Conference of the Spatial Econometrics Association, Toulouse, July 6–8, 2011
Sen A, Smith T (1995) Gravity models of spatial interaction behavior. Springer, Berlin/Heidelberg/New York
Tiefelsdorf M (2003) Misspecification in interaction model distance decay relations: a spatial structure effect. J Geogr Syst 5(1):25–50
Tiefelsdorf M, Boots B (1995) The exact distribution of Moran's I. Environ Plan A 27(6):985–999
Tobler W (1983) An alternative formulation for spatial interaction modeling. Environ Plan A 15(5):693–703
Wilson AG (1967) A statistical theory of spatial distribution models. Transp Res 1(3):253–269
Wilson AG (1970) Entropy in urban and regional modelling. Pion, London
Wilson AG (1971) A family of spatial interaction models and associated developments. Environ Plan 3(1):1–32
Wilson AG, Kirkby MJ (1980) Mathematics for geographers and planners, 2nd edn. Clarendon, Oxford