University of Dundee

Estimation of the Spatial Weights Matrix under Structural Constraints

Bhattacharjee, Arnab; Jensen-Butler, Chris

Link to publication in Discovery Research Portal

SIRE DISCUSSION PAPER

SIRE-DP-2011-48

Estimation of the Spatial Weights Matrix under Structural Constraints

Arnab Bhattacharjee

University of Dundee

(Late) Chris Jensen-Butler

University of St Andrew

Estimation of the Spatial Weights Matrix under Structural Constraints[*]

Arnab Bhattacharjee[$] and (Late) Chris Jensen-Butler[#†]

June 11, 2011

Abstract

While estimates of models with spatial interaction are very sensitive to the choice of spatial weights, considerable uncertainty surrounds definition of spatial weights in most studies with cross-section dependence. We show that, in the spatial error model the spatial weights matrix is only partially identified, and is fully identified under the structural constraint of symmetry. For the spatial error model, we propose a new methodology for estimation of spatial weights under the assumption of symmetric spatial weights, with extensions to other important spatial models. The methodology is applied to regional housing markets in the UK, providing an estimated spatial weights matrix that generates several new hypotheses about the economic and socio-cultural drivers of spatial diffusion in housing demand.

Keywords: Spatial econometrics; Spatial autocorrelation; Spatial weights matrix, Spatial error model; Housing demand; Gradient projection.

JEL Classification: C14; C15; C30; C31; R21; R31.

1 Introduction

This paper considers inferences on an unknown spatial weights matrix using panel data on a given collection of spatial units. In the regional science and geography literatures, and increasingly in economic applications, spatial dependence is commonly modeled by a spatial weights matrix, whose elements represent the direction and strength of spillovers between each pair of units. The spatial weights matrix provides a convenient way to model structural spatial dependence.[1] The weights represent patterns of interaction and diffusion, and thereby provide a meaningful and easily interpretable representation of spatial interaction (spatial autocorrelation) in spatial dependence models.

The choice of appropriate spatial weights is a central component of spatial models as it assumes *a priori* a structure of spatial dependence, which may or may not correspond closely to reality. Typically, the spatial weights are interpreted as functions of relevant measures of economic or geographic distance (Anselin, 1988, 2002), and are therefore driven by the spatial structure of an application. The choice typically differs widely across applications, depending not only on the specific economic context but also on availability of data. Spatial contiguity (resting upon implicit assumptions about contagious processes) using a binary representation, is a frequent choice. Further, in many applications, there are multiple possible choices and substantial uncertainty regarding the appropriate choice of distance measures.

While the literature contains an implicit acknowledgment of these issues, most empirical studies treat spatial dependence in a superficial manner assuming inflexible diffusion processes in terms of *a priori* fixed spatial weights matrices (Giacomini and Granger, 2004). Unfortunately, the accuracy of spatial weights affects profoundly the estimation of spatial dependence models (Anselin, 2002; Fingleton, 2003). Therefore, the problem of choosing spatial weights is a key issue in most applications.[2]

---

[1]Here "structural" means that spatial dependence is exogenously determined by the spatial structure of the problem, that is, by the organisation of the observation units in space. Structural spatial dependence represented by spatial weights is conceptually quite different from the main alternative model, where spatial dependence is asumed to be driven by a finite number of unobserved common factors that affect all units (regions, economic agents, etc.). See Bhattacharjee and Holly (2011) for further discussion on the conceptual and methodological distinction between the two views.

[2]For example, ethnic networks and trade, spatial demand, risk-sharing in rural economies, convergence in growth models, and spillovers in housing markets (see Conley, 1999; Fingleton, 2003; Herander and Saavedra, 2005; Holly *et al.*, 2011). See Corrado and Fingleton (2011) for a recent discussion encompassing the regional studies, urban economics and economic geography perspectives.

However, the spatial econometrics literature tells us little about adequate foundations for these choices. Acknowledging uncertainty regarding specification of the weights, Conley (1999) used imperfectly measured economic distances to obtain asymptotic estimators of the implied spatial autocovariance matrix. Likewise, Pinkse *et al.* (2002) and Kelejian and Prucha (2004) developed methods accommodating uncertainty regarding distance measures and allowing for spatial nonstationarity. Some alternative approaches to construction of the spatial weights matrix have emerged from the geography and regional studies literature. Getis and Aldstadt (2004) use a local dependence statistic, while Aldstadt and Getis (2006) developed a multidirectional optimum eco-type-based algorithm. The latent variables approach of Folmer and Oud (2008) is conceptually similar to the multi-factor model of spatial dependence, but based more closely on latent variables implied by the specific applied context rather than those identified by statistical factor analysis. None of the above are formal estimators of spatial weights and their statistical properties remain largely unknown.

We study inferences on an unknown spatial weights matrix consistent with an observed (estimated) pattern of spatial autocovariances. We show that the estimation problem is partially identified if no structural assumptions are made, and fully identified under the assumption of symmetric spatial weights. Further, we develop an estimator of the spatial weights matrix under suitable structural constraints. Once these spatial weights have been estimated they can be subjected to interpretation in order to identify the true nature and strength of spatial dependence, representing a significant departure from the usual practice of assuming *a priori* the nature of spatial interactions. While our approach builds upon the spatial error model with autoregressive errors, we discuss extensions to other models of spatial dependence.

An illustration of the proposed methodology and its potential is presented through a study of regional housing markets in England and Wales, where the aim is to identify the economic and socio-cultural drivers of spatial diffusion in housing demand. Whilst the study of spatial diffusion of demand in the housing market is in itself not new (for example, in the case of the UK: Meen, 2001, 2003; Holly *et al.*, 2011; and for the US: Haurin and Brasington, 1996; Fik *et al.*, 2003), the *ex post* approach to identifying drivers of diffusion presented here is novel and the results highlight a striking departure from consideration exclusively of contiguity or distance measured *a priori*.

Section 2 describes the econometric framework and the methodology, estimators, their compuation and asymptotic properties. In section 3, we discuss bootstrap estimation of standard errors and hypothesis tests. Results of a Monte Carlo simulation study are reported in section 4, followed in section 5 by a real application. Finally, section 6 concludes.

2 Proposed Estimation Methodology

In this section, we propose methods for estimation of the spatial weights matrix when spatial dependence is in the form of a spatial error model. Much of our discussion is in the context of a model where the error process of spatial diffusion has an autoregressive structure; we also extend the methods to the moving average spatial error model and the spatial lag model. The proposed estimation takes as given a consistent and asymptotically normal estimator for elements of the spatial autocovariance matrix. In the following subsections, we discuss the spatial error model together with assumptions, followed by estimation of the spatial weights matrix implied by the given (estimated) autocovariance matrix of the spatial errors.

2.1 The Spatial Error Model

At the same time as spatial weights characterise structural spatial dependence in useful ways, their measurement has crucial effect on the estimation of spatial econometric models. While much of the literature assumes a prespecified spatial weights matrix, the problems with this assumption are well-known: the choice of weights is frequently arbitrary, there is substantial uncertainty regarding the choice, and empirical results vary considerably according to the choice of spatial weights. Given a particular choice of the spatial weights matrix, there are two important and distinct ways in which spatial interaction is modelled in spatial regression analysis – the spatial error model and the spatial lag model.

We describe the spatial lag model and two variants of the spatial error model focussing on an application, presented later in the paper, to spillovers in regional housing demand. We consider panel data on demand and other explanatory variables across several markets (regions) over several time periods. Within the context of estimated demand equations for each of several regions, we aim to understand how excess demand in any region diffuses over space to the other regions.

Our central model is a spatial error model where regional demand (D) is driven by the effect of several explanatory variables ($\boldsymbol{X}$) and diffusion of excess demand from other regions. The spatial externalities in the form of demand diffusion are driven by an unknown spatial weights matrix ($\boldsymbol{W}$) and the strength of spatial spillovers in each region described by spatial autocorrelation parameters ρ_K. The model and its reduced form are described as

follows:

$$
\begin{aligned}
\underline{D}_t &= \boldsymbol{X}_t.\underline{\beta} + \underline{u}_t, \quad t = 1, \ldots, n, \\
\underline{u}_t &= \boldsymbol{R}.\boldsymbol{W}.\underline{u}_t + \underline{\varepsilon}_t, \\
&\implies \underline{D}_t = \boldsymbol{X}_t.\underline{\beta} + (\boldsymbol{I} - \boldsymbol{R}.\boldsymbol{W})^{-1}.\underline{\varepsilon}_t,
\end{aligned} \tag{1}
$$

where there are n time periods ($t = 1, \ldots, n$) and K regions ($k = 1, \ldots, K$), $\underline{D}_t$ is the $K \times 1$ vector of regional demand in period t,
$\boldsymbol{W}$ is an unknown spatial weights matrix of dimension $K \times K$,
$\boldsymbol{R} = diag\,(\rho_1, \rho_2, \ldots, \rho_K)$ is a $K \times K$ diagonal matrix containing the spatial autoregression parameters for each region,[3] and
$\underline{\varepsilon}_t$ is the $K \times 1$ vector of independent but possibly heteroscedastic spatial errors. Since the spatial diffusion of errors is assumed to be autoregressive, the model is called a spatial error model with autoregressive errors (SEM-AR).

Alternatively, the spatial diffusion of errors can be modelled as a moving average process (Anselin, 1988), which we call the spatial error model with moving average errors (SEM-MA):

$$
\begin{aligned}
\underline{u}_t &= \boldsymbol{R}.\boldsymbol{W}.\underline{\varepsilon}_t + \underline{\varepsilon}_t, \\
&\implies \underline{D}_t = \boldsymbol{X}_t.\underline{\beta} + (\boldsymbol{I} + \boldsymbol{R}.\boldsymbol{W}).\underline{\varepsilon}_t.
\end{aligned} \tag{2}
$$

The above formulation of the SEM-AR and SEM-MA models represent several departures from the literature. First and foremost, we do not assume a given form of the spatial weights matrix. Indeed, our aim is to estimate the spatial weights matrix.[4]

Second, we allow the spatial autoregression coefficients to vary over the spatial units (regions). This flexibility in spatial spillovers captures an important dimension of heterogeneity.

Third, the description of the model also allows for heterogeneity in the regressors as well as their coefficients across the regions, a feature which is important in applications; that is, we assume a spatial regime model.

Fourth, we allow heteroscedasticity in the idiosyncratic errors across regions. However, all spatial autocorrelation under the model is driven by the

---

[3]Here, and henceforth in the paper, $diag(a_1, a_2, \ldots, a_k)$ denotes the diagonal matrix
$$
\begin{pmatrix}
a_1 & 0 & \ldots & 0 \\
0 & a_2 & \ldots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \ldots & a_K
\end{pmatrix}.
$$
[4]It is, however, obvious that the spatial weights matrix $\boldsymbol{W}$ cannot be separately identified from the diagonal matrix of autoregressive parameters $\boldsymbol{R}$, unless further assumptions are made. In other words, our aim is to identify and conduct inferences on $\boldsymbol{R}.\boldsymbol{W}$.

spatial weights matrix and the spatial autoregression parameters. This is the crucial feature in the setting that allows us to estimate the implied spatial weights from the observed spatial autocovariances.

Standard models of temporal variation, such as autoregressive or moving average processes in the temporal dimension, can also be accommodated in our framework. However, for reasons of expositional simplicity, we abstract from the issue of temporal dynamics in this paper. In other words, we assume independence of the joint distribution of the variables over time, but not over space.

An alternate model popular in the literature is the spatial lag model (Anselin, 1988):

$$
\begin{aligned}
\underline{D}_t &= \boldsymbol{R}.\boldsymbol{W}.\underline{D}_t + \boldsymbol{X}_t.\underline{\beta} + \underline{\varepsilon}_t, \quad t = 1, \ldots, n, \\
&\Rightarrow \underline{D}_t = (\boldsymbol{I} - \boldsymbol{R}.\boldsymbol{W})^{-1}.\boldsymbol{X}_t.\underline{\beta} + (\boldsymbol{I} - \boldsymbol{R}.\boldsymbol{W})^{-1}.\underline{\varepsilon}_t.
\end{aligned} \tag{3}
$$

The choice between the spatial error and the spatial lag models is largely context specific and often quite difficult. Though different in interpretation, the two kinds of models are difficult to distinguish empirically (Anselin, 1999, 2002). However, based on a given spatial weights matrix, there are tests that can help decide between the two models.[5]

In this paper, we focus on the spatial error model with autoregressive errors (1). In the context of our application, we find this model well suited to explain housing demand in terms of spatial diffusion of excess demand from neighbouring regions, where neighbourhood is defined in an abstract sense. We also discuss the spatial error model with moving average errors (2) and the spatial lag model (3).

The above setting is formalised in the following assumptions.

Assumption 1: *The number of time periods increases asymptotically ($n \to \infty$) while the number of spatial units is fixed ($K_n \equiv K < \infty$). The spatial errors, $\underline{\varepsilon}_t$, are iid (independent and identically distributed) across time, but is potentially heteroscedastic over spatial units. Thus, $\mathbb{E}\left(\underline{\varepsilon}_t.\underline{\varepsilon}_t^T\right) = \boldsymbol{\Sigma} = diag\left(\sigma_1^2, \sigma_2^2, \ldots, \sigma_K^2\right)$, and $\sigma_k^2 > 0$ for all $k = 1, \ldots, K$.*[6]

The uncorrelatedness of the idiosyncratic spatial errors across the regions is a crucial assumption. **Assumption 1** ensures that all spatial autocorrelation in the model is solely due to spatial diffusion described by the spatial

---

[5]See Baltagi *et al.* (2003) and Borg and Breitung (2009) for reviews of tests designed to aid model choice. However, theory is often the best guide to choice between these two models. This is because, in most empirical applications, spatial lag dependence tends to be insignificant once spatial error dependence is allowed for; and vice versa.

[6]Here and throughout the paper, ^{T} denotes transpose of a vector or matrix.

weights matrix and the autoregression coefficients. This feature of the model drives our identification and estimation strategy. Specifically, we use the observed (estimated) pattern of spatial autocovariances and variances to infer on the unknown spatial weights.

Assumption 2: *The spatial weights matrix $\boldsymbol{W}$ is unknown and possibly asymmetric. $\boldsymbol{W}$ has zero diagonal elements and the off-diagonal elements can be positive or negative.*

As discussed in the previous section, the literature on spatial modelling acknowledges substantial uncertainty in the specification of appropriate spatial weights. Practitioners are encouraged to exercise caution in the choice of the spatial weights matrix, and also to experiment with different choices. If the spatial weights are inversely related to some underlying metric distance between the regions, then the spatial weights matrix would be symmetric. At the moment, we retain the flexibility of a possibly asymmetric spatial weights matrix, though symmetry is assumed subsequently (**Assumption 4**).

Our most significant point of departure from the literature is in the assumption of an unknown spatial weights matrix. We do not impose a non-negativity constraint on the off-diagonal elements of $\boldsymbol{W}$; later we discuss how negative off-diagonal elements can be interpreted. In fact, in an application related to ours, Meen (1996) finds evidence of negative interaction between some pairs of regions.

We will show that the spatial weights matrix is identified by the spatial autocovariance matrix only upto an orthogonal transformation. Additional assumptions, such as symmetry of the spatial weights matrix, are therefore required for full identification and estimation of spatial weights.

Assumption 3: $(\boldsymbol{I} - \boldsymbol{R}.\boldsymbol{W})$ *is non-singular, where $\boldsymbol{I}$ is the identity matrix.*

This is a standard assumption in the literature. It is required for identification in the reduced form, and holds under the spatial granularity condition in Pesaran and Tosetti (2011).[7] Under **Assumption 3**, the reduced form representation of the SEM-AR model (1) is given by:

$$
\begin{aligned}
\underline{D}_t &= \boldsymbol{X}_t.\underline{\beta} + \underline{u}_t \\
&= \boldsymbol{X}_t.\underline{\beta} + (\boldsymbol{I} - \boldsymbol{R}.\boldsymbol{W})^{-1}.\underline{\varepsilon}_t,
\end{aligned}
$$

where (dropping the time subscript)

$$
\mathbb{E}\left(\underline{u}.\underline{u}^T\right) = (\boldsymbol{I} - \boldsymbol{R}.\boldsymbol{W})^{-1}.\boldsymbol{\Sigma}.(\boldsymbol{I} - \boldsymbol{R}.\boldsymbol{W})^{-1^T}.
$$

---

[7] Similar assumptions are needed for ML and GMM estimation of spatial econometric models with known $\boldsymbol{W}$; see, for example, Kelejian and Prucha (1999) and Lee (2004).

Since all spatial autocovariance under model (1) is driven by the spatial weights matrix $\boldsymbol{W}$, we ask the following question: whether, and under what conditions, would $\boldsymbol{W}$ be identified by the variances and autocovariances of spatial errors across all the units, that is, by $\boldsymbol{\Gamma} = \mathbb{E}\left(\underline{u}.\underline{u}^T\right)$? Typically, $\boldsymbol{\Gamma}$ is estimated from the data, using either a two-stage procedure, or ML, or GMM. Then, under suitable conditions that ensure identification, this paper develops methods to estimate spatial weights matrix implied by the estimate of the spatial autocovariance matrix $\boldsymbol{\Gamma}_n$.

2.2   Identification and Estimation of Spatial Weights

Meen (1996) studied spatial diffusion in housing starts across regions in England under a SEM-AR model by regressing OLS residuals of the regression relationship for each region on residuals from all the other regions

$$
\begin{aligned}
\underline{D}_t &= \boldsymbol{X}_{t.}\underline{\beta} + \underline{u}_t, \quad \text{(First stage regression)} \\
\widehat{\underline{u}}_t &= \boldsymbol{X}_{t.}\widehat{\underline{\beta}}, \\
\widehat{u}_{kt} &= \rho_k.\sum_{\substack{j=1 \\ j \neq k}}^{K} w_{kj}.\widehat{u}_{jt} + \varepsilon_{kt}, \quad \text{(Second stage regressions)} \quad (4)
\end{aligned}
$$

where the second stage regressions are estimated separately for each region.[8] The sign and statistical significance of the OLS regression estimates at the second stage were used to examine the sign and strength of spatial dependence between the regions. This approach was the first attempt towards estimating the unknown spatial weights matrix up to a factor of proportionality (the spatial autoregressive parameter).

This method, however, suffers from a serious problem of endogeneity. Specifically, the regression equations at the second stage constitute a system of simultaneous equations in the first stage residuals. Since these residuals are endogenous to the system, OLS will produce biased and inconsistent estimates of the regression coefficients, which are the spatial weights. Nonetheless, the approach of Meen (1996) provides the key insight that the spatial weights are the true partial effects of regression errors for each ordered pair of spatial units.

In contrast to Meen (1996), we begin with estimating the spatial error autocovariance matrix, $\boldsymbol{\Gamma} = \mathbb{E}\left(\underline{u}.\underline{u}^T\right)$, of the underlying structural model,

---

[8]Note that, Meen (1996) assumes homogeneity in the spatial autoregression parameter across the regions. For expositional clarity, we retain heterogeneity in the spatial autoregression parameter for the moment. In any case, the autoregression parameter is not separately identifiable from the spatial weights matrix in our setting.

potentially using the first stage residuals. Then, we use the reduced form of the second stage regression to estimate the spatial weights matrix (up to a factor of proportionality) from the spatial autocovariance matrix estimated in the first stage. Our method for estimation of the spatial weights matrix relies on the structure of the problem: most importantly, the zero diagonal elements of the spatial weights matrix (**Assumption 2**) and the diagonal structure of the covariance matrix ($\boldsymbol{\Sigma} = \mathbb{E}\left(\underline{\varepsilon}.\underline{\varepsilon}^T\right)$) of the spatial errors (**Assumption 1**). The proposed method does not impose *a priori* any structure on the drivers of spatial diffusion.

Before proceeding to estimation, we first discuss identification of $\boldsymbol{W}$ and uniqueness under the additional assumption of symmetric spatial weights.

Let $\boldsymbol{\Gamma} = \boldsymbol{E}.\boldsymbol{\Lambda}.\boldsymbol{E}^T$ be the spectral decomposition of $\boldsymbol{\Gamma} = \mathbb{E}\left(\underline{u}.\underline{u}^T\right)$, where $\boldsymbol{\Lambda} = diag\left(\lambda_1, \lambda_2, \ldots, \lambda_K\right)$ is the diagonal matrix of eigenvalues and the columns of $\boldsymbol{E} = [\underline{e}_1, \underline{e}_2, \ldots, \underline{e}_K]$ contain the corresponding eigenvectors.

Proposition 1 *(**Partial Identification**) Let **Assumptions 1–3** hold, and let $\boldsymbol{\Gamma}^{-1/2}$ denote the symmetric square root of $\boldsymbol{\Gamma} = \mathbb{E}\left(\underline{u}.\underline{u}^T\right)$, defined as*

$$\boldsymbol{\Gamma}^{-1/2} = \boldsymbol{E}.\boldsymbol{\Lambda}^{-1/2}.\boldsymbol{E}^T.$$

Then,

$$\boldsymbol{V} = (\boldsymbol{I} - \boldsymbol{R}.\boldsymbol{W})^T .diag\left(\frac{1}{\sigma_1}, \frac{1}{\sigma_2}, \ldots, \frac{1}{\sigma_K}\right) \tag{5}$$

is isomorphic to $\boldsymbol{\Gamma}^{-1/2}$, upto an arbitrary orthogonal transformation. That is,

$$\boldsymbol{\Gamma}^{-1/2} = \boldsymbol{V}.\boldsymbol{T}$$

for some arbitrary square orthogonal matrix $\boldsymbol{T}$. In other words, $\boldsymbol{R}.\boldsymbol{W}$ is partially identified by $\boldsymbol{\Gamma}^{-1/2}$ upto an arbitrary orthogonal transformation. Further, in the special case when the rows of $(\boldsymbol{I} - \boldsymbol{R}.\boldsymbol{W})$ are orthogonal to each other, there is exact identification, that is,

$$\boldsymbol{\Gamma}^{-1/2} = \boldsymbol{V}.$$

Proof. *By **Assumptions 1** and **3**, under the SEM-AR model (1), the population spatial autocovariance matrix, $\boldsymbol{\Gamma} = \mathbb{E}\left(\underline{u}.\underline{u}^T\right)$, is positive definite, and therefore so is $\boldsymbol{\Gamma}^{-1}$. Consider the spectral decomposition of $\boldsymbol{\Gamma}^{-1}$,*

$$\begin{aligned} \boldsymbol{\Gamma}^{-1} &= (\boldsymbol{I} - \boldsymbol{R}.\boldsymbol{W})^T .\boldsymbol{\Sigma}^{-1}.(\boldsymbol{I} - \boldsymbol{R}.\boldsymbol{W}) \\ &= \boldsymbol{V}.\boldsymbol{V}^T. \end{aligned} \tag{6}$$

This implies that

$$\boldsymbol{\Gamma}^{-1} = (\boldsymbol{V}.\boldsymbol{T}) . (\boldsymbol{V}.\boldsymbol{T})^T,$$

for any square orthogonal matrix $\boldsymbol{T}$.

Now, since $\boldsymbol{\Gamma}$ is positive definite, $\min(\lambda_1, \lambda_2, \ldots, \lambda_K) > 0$ and

$$\boldsymbol{\Lambda}^{-1/2} = diag\left(\frac{1}{\sqrt{\lambda_1}}, \frac{1}{\sqrt{\lambda_2}}, \ldots, \frac{1}{\sqrt{\lambda_K}}\right).$$

Since $\boldsymbol{\Gamma} = \boldsymbol{E}.\boldsymbol{\Lambda}.\boldsymbol{E}^T$, it follows that, the spectral decomposition of $\boldsymbol{\Gamma}^{-1}$ is given by:

$$
\begin{aligned}
\boldsymbol{\Gamma}^{-1} &= \boldsymbol{E}.\boldsymbol{\Lambda}^{-1}.\boldsymbol{E}^T \\
&= \boldsymbol{E}.\boldsymbol{\Lambda}^{-1/2}.\boldsymbol{E}^T.\boldsymbol{E}.\boldsymbol{\Lambda}^{-1/2}.\boldsymbol{E}^T \\
&= \left(\boldsymbol{E}.\boldsymbol{\Lambda}^{-1/2}.\boldsymbol{E}^T\right).\left(\boldsymbol{E}.\boldsymbol{\Lambda}^{-1/2}.\boldsymbol{E}^T\right)^T.
\end{aligned}
\tag{7}
$$

The first step holds because $\boldsymbol{E}^{-1} = \boldsymbol{E}^T$, and the final step follows since $\boldsymbol{\Lambda}^{-1/2}$ is a symmetric matrix.

The result follows, since (6) and (7) can both hold if and only if

$$\boldsymbol{E}.\boldsymbol{\Lambda}^{-1/2}.\boldsymbol{E}^T = \boldsymbol{\Gamma}^{-1/2} = \boldsymbol{V}.\boldsymbol{T},$$

for some orthogonal matrix $\boldsymbol{T}$.

Further, if the rows of $(\boldsymbol{I} - \boldsymbol{R}.\boldsymbol{W})$ are orthogonal to each other, then the spectral decomposition of $\boldsymbol{\Gamma}^{-1}$ is unique and $\boldsymbol{T}$ is the identity matrix; hence, the exact equality $\boldsymbol{\Gamma}^{-1/2} = \boldsymbol{V}$ holds. ∎

Since the rows of $(\boldsymbol{I} - \boldsymbol{R}.\boldsymbol{W})$ are orthogonal only in very special cases,[9] the spatial weights matrix is only partially identified by the spatial autocovariance matrix $\boldsymbol{\Gamma}$, because the orthogonal matrix $\boldsymbol{T}$ that links $\boldsymbol{R}.\boldsymbol{W}$ to $\boldsymbol{\Gamma}$ can be arbitrary. Additional structural assumptions are required to uniquely identify the corresponding orthogonal matrix $\boldsymbol{T}$.

Assumption 4: *The spatial weights matrix is symmetric and the spatial autoregression parameter is the same across all regions; in other words $\boldsymbol{R}.\boldsymbol{W} = \rho\boldsymbol{W}$ is symmetric.*

Note that a symmetric spatial weights matrix is typically assumed in the regional studies and geography literatures, with the weights as functions of metric distances between spatial units. However, this assumption may be unreasonable in some applications. Row-standardised spatial weights matrices

---

[9]Such as, when the spatial weights matrix is zero, implying no spatial dependence.

are usually asymmetric by construction; asymmetric spatial weights matrices are also important in the study of asymmetric shocks, network flows and core-periphery models. Bhattacharjee and Holly (2008, 2011) discuss alternate sets of assumptions that may be useful in such situations.

Proposition 2 *(**Uniqueness**) Let **Assumptions 1–3** hold. Then, the spatial weights matrix for which **Assumption 4** also holds is unique upto proportionality.*

Proof. *Suppose the symmetric spatial weights matrix is not unique. In other words, let there be two distinct matrices $\rho_1 \boldsymbol{W}_1$ and $\rho_2 \boldsymbol{W}_2$ consistent with **Assumptions 1-4**.*
Then,

$$
\begin{aligned}
\boldsymbol{\Gamma}^{-1} &= (\boldsymbol{I} - \rho_1 \boldsymbol{W}_1)^T .\boldsymbol{\Sigma}^{-1} .(\boldsymbol{I} - \rho_1 \boldsymbol{W}_1) \\
&= (\boldsymbol{I} - \rho_1 \boldsymbol{W}_1) .\boldsymbol{\Sigma}^{-1} .(\boldsymbol{I} - \rho_1 \boldsymbol{W}_1) ,
\end{aligned}
\tag{8}
$$

since $\boldsymbol{W}_1$ is symmetric. Likewise,

$$
\begin{aligned}
\boldsymbol{\Gamma}^{-1} &= (\boldsymbol{I} - \rho_2 \boldsymbol{W}_2)^T .\boldsymbol{\Sigma}^{-1} .(\boldsymbol{I} - \rho_2 \boldsymbol{W}_2) \\
&= (\boldsymbol{I} - \rho_2 \boldsymbol{W}_2) .\boldsymbol{\Sigma}^{-1} .(\boldsymbol{I} - \rho_2 \boldsymbol{W}_2)
\end{aligned}
\tag{9}
$$

Equations (8) and (9) imply that

$$
(\rho_1 \boldsymbol{W}_1 - \rho_2 \boldsymbol{W}_2)^T .\boldsymbol{\Sigma}^{-1} .(\rho_1 \boldsymbol{W}_1 - \rho_2 \boldsymbol{W}_2) = \boldsymbol{0}.
$$

But since the LHS is the variance covariance matrix of the random vector

$$
\underline{u}^* = (\rho_1 \boldsymbol{W}_1 - \rho_2 \boldsymbol{W}_2) \underline{u},
$$

the only way this covariance matrix can be zero is

$$
\rho_1 \boldsymbol{W}_1 = \rho_2 \boldsymbol{W}_2.
$$

That is, the symmetric weights matrix is unique upto proportionality. ∎

Proposition 2 states that, if **Assumption 4** holds, there exists a unique orthogonal matrix $\boldsymbol{T}$ such that

$$
\begin{aligned}
\boldsymbol{\Gamma}^{-1/2} .\boldsymbol{T} &= \left[(\boldsymbol{I} - \rho \boldsymbol{W})^T .diag \left(\frac{1}{\sigma_1}, \frac{1}{\sigma_2}, \ldots, \frac{1}{\sigma_K} \right) \right] \\
&= \boldsymbol{Q} = ((q_{ij}))_{i,j=1,\ldots,K} .
\end{aligned}
$$

11

Given $\boldsymbol{T}$, this relation provides a direct link between the spatial autocovariance matrix $\boldsymbol{\Gamma}$ and the spatial weights matrix $\rho\boldsymbol{W}$. Specifically, since (a) post-multiplication by the diagonal matrix $diag\left(\frac{1}{\sigma_1}, \frac{1}{\sigma_2}, \ldots, \frac{1}{\sigma_K}\right)$ transforms $(\boldsymbol{I} - \rho\boldsymbol{W})^T$ by multiplying each column by the corresponding diagonal element, and (b) by **Assumption 2**, the elements on the diagonal of $(\boldsymbol{I} - \rho\boldsymbol{W})^T$ are all unity, we have

$$(\boldsymbol{I} - \rho\boldsymbol{W})^T = \begin{bmatrix} 1 & q_{12}/q_{11} & \ldots & q_{1K}/q_{11} \\ q_{21}/q_{22} & 1 & \ldots & q_{2K}/q_{22} \\ \vdots & \vdots & \ddots & \vdots \\ q_{K1}/q_{KK} & q_{K2}/q_{KK} & \ldots & 1 \end{bmatrix}. \tag{10}$$

Therefore, the condition that the spatial weights matrix is symmetric (**Assumption 4**) implies that

$$\frac{q_{ij}}{q_{ii}} = \frac{q_{ji}}{q_{jj}}, \quad \text{for all} \;\; i \neq j; i, j = 1, \ldots, K. \tag{11}$$

Further, note that $\boldsymbol{T}$ consists of K^2 real values, and several restrictions apply to the elements of this matrix. Specifically, there are K normalization restrictions (one for each column of the matrix) and $K(K-1)/2$ orthogonality conditions (one for each distinct pair of columns). This leaves $K(K-1)/2$ free elements in the matrix.

Hence, if the weights matrix is symmetric, these free elements can be pinned down by the $K(K-1)/2$ constraints relating to the assumption of symmetry (11).[10] This identification is upto a sign transformation on the columns and rows of $\boldsymbol{T}$ that preserves the orthogonality condition while at the same time ensuring that the diagonal elements of the transformed matrix are all positive.

Note that, the implication of **Proposition 2** runs only one way. Specifically, given a spatial autocovariance matrix $\boldsymbol{\Gamma}$, no symmetric spatial weights matrix consistent with this may exist. However, in the 2-region case ($K = 2$), the implication runs both ways and the unique identification of $\boldsymbol{T}$ with any given $\boldsymbol{\Gamma}$ can verified by construction. In this case, any orthogonal matrix can be expressed as

$$\begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix},$$

---

[10]This constraint is seen even better from the result (Cayley, 1846; Reiersöl, 1963) that any orthogonal matrix is uniquely associated with a corresponding skew-symmetric matrix: $\boldsymbol{T} = (\boldsymbol{I} - \boldsymbol{K})(\boldsymbol{I} + \boldsymbol{K})^{-1}$; a skew symmetric matrix has exactly $K(K-1)/2$ free elements that can potentially be fixed by the corresponding symmetry assumptions.

which involves a single unknown angle ϕ. While ϕ can take values in $[0, 2\pi]$, all unique cases are completely covered in $\phi \in [0, \pi]$. Denoting the elements of the symmetric square root $\mathbf{\Gamma}^{-1/2}$ as $\left((\gamma_{ij}) \right), i, j = 1, 2$ where $\gamma_{12} = \gamma_{21}$, **Assumption 4** can be expressed as

$$\frac{q_{12}}{q_{11}} = \frac{q_{21}}{q_{22}} \implies \left[\tan \phi - \frac{1}{\tan \phi} \right] = \frac{2 \left(\gamma_{11} \cdot \gamma_{22} + \gamma_{12}^2 \right)}{\gamma_{12} \cdot \left(\gamma_{11} - \gamma_{22} \right)},$$

which has a unique solution within the range $\phi \in [0, \pi]$.

Next, we turn to estimation. Since $\mathbf{W}$ is identified only upto proportionality, we assume without loss of generality that $\rho \equiv 1$. Based on **Proposition 2** and (11), we propose the following natural estimator for the orthogonal matrix $\mathbf{T}$ corresponding to $\mathbf{W}$:

$$
\begin{aligned}
\mathbf{T}_n &= \underset{\mathbf{T}:\mathbf{T}\mathbf{T}^T=\mathbf{I}}{\arg\min} \sum_{i<j} \left(\frac{q_{ij}}{q_{ii}} - \frac{q_{ji}}{q_{jj}} \right)^2, \\
\mathbf{Q} &= \left((q_{ij}) \right)_{i,j=1,\dots,K} = \mathbf{\Gamma}_n^{-1/2} \mathbf{T}
\end{aligned}
\tag{12}
$$

where $\mathbf{\Gamma}_n$ denotes the estimated spatial autocovariance matrix of the spatial errors $\underline{u}$, and the criterion function is optimised over all orthogonal matrices $\mathbf{T}$. From (10), the corresponding estimators for the heteroscedastic idiosyncratic errors $(\sigma_{i,n}, i = 1, \dots, K)$ and symmetric spatial weights matrix $(\mathbf{W}_n)$ and are uniquely given by:

$$
\begin{aligned}
\mathbf{\Gamma}_n^{-1/2} \cdot \mathbf{T}_n &= \mathbf{Q}_n = \left((q_{ij,n}) \right)_{i,j=1,\dots,K}, \\
\mathbf{W}_n &= \begin{bmatrix} 0 & -q_{21,n}/q_{22,n} & \dots & -q_{K1,n}/q_{KK,n} \\ -q_{12,n}/q_{11,n} & 0 & \dots & -q_{K2,n}/q_{KK,n} \\ \vdots & \vdots & \ddots & \vdots \\ -q_{1K,n}/q_{11,n} & -q_{2K,n}/q_{22,n} & \dots & 0 \end{bmatrix}, \\
\sigma_{k,n} &= \frac{1}{q_{kk}}, \quad k = 1, \dots, K.
\end{aligned}
\tag{13}
$$

Since the criterion function is (implicitly) expressed in terms of moments, the estimator is a generalised M-estimator based on complicated moment conditions. Next, we show consistency and weak convergence of the above estimator.

The covariance matrix $\mathbf{\Gamma}$ contains $p = K(K+1)/2$ distinct elements, which we combine into a vector $\underline{l}_{(p \times 1)}$. Likewise, the spatial weights matrix $\mathbf{W}$ together with the heteroscedastic variances $(\sigma_1^2, \sigma_2^2, \dots, \sigma_K^2)$ constitute p

unknown and free parameters, which we combine into $\underline{w}_{(p \times 1)}$. Then, corresponding to (12), the true spatial weights matrix and idiosyncratic error variances $(\underline{w})$ are implicitly defined by the equation:

$$
\begin{aligned}
G\left(\underline{l}, \underline{w}\right) &= \sum_{i<j}\left(\frac{q_{ij}}{q_{ii}} - \frac{q_{ji}}{q_{jj}}\right)^2 = 0, \tag{14} \\
\boldsymbol{Q} &= \boldsymbol{\Gamma}^{-1/2}\boldsymbol{T} = \left[(\boldsymbol{I} - \boldsymbol{W})^T . diag\left(\frac{1}{\sigma_1}, \frac{1}{\sigma_2}, \ldots, \frac{1}{\sigma_K}\right)\right], \\
\boldsymbol{\Gamma} &= \mathbb{E}\left(\underline{u}.\underline{u}^T\right), \tag{15}
\end{aligned}
$$

where the parameter space constitutes all positive definite matrices $\boldsymbol{\Gamma}$, symmetric spatial weights matrices $\boldsymbol{W}$ satisfying **Assumptions 2–4**, and diagonal positive definite covariance matrices $\boldsymbol{\Sigma}$.[11] Corresponding to $G\left(., .\right)$, we define the vector of partial derivatives

$$
g = \begin{bmatrix} g_1 \\ g_2 \\ \vdots \\ g_p \end{bmatrix} = \begin{bmatrix} \partial G/\partial w_1 \\ \partial G/\partial w_2 \\ \vdots \\ \partial G/\partial w_p \end{bmatrix},
$$

the Jacobian matrix

$$
\mathbf{J}_{(p\times p)} = \begin{bmatrix} \partial^2 G/\partial s_1^2 & \partial^2 G/\partial s_1 \partial s_2 & \ldots & \partial^2 G/\partial s_1 \partial s_p \\ \partial^2 G/\partial s_2 \partial s_1 & \partial^2 G/\partial s_2^2 & \ldots & \partial^2 G/\partial s_2 \partial s_p \\ \vdots & \vdots & \ddots & \vdots \\ \partial^2 G/\partial s_p \partial s_1 & \partial^2 G/\partial s_p \partial s_2 & \ldots & \partial^2 G/\partial s_p^2 \end{bmatrix},
$$

and the cross partial derivatives matrix

$$
\mathbf{H}_{(p\times p)} = \begin{bmatrix} \partial^2 G/\partial w_1 \partial l_1 & \partial^2 G/\partial w_1 \partial l_2 & \ldots & \partial^2 G/\partial w_1 \partial l_p \\ \partial^2 G/\partial w_2 \partial l_1 & \partial^2 G/\partial w_2 \partial l_2 & \ldots & \partial^2 G/\partial w_2 \partial l_p \\ \vdots & \vdots & \ddots & \vdots \\ \partial^2 G/\partial w_p \partial l_1 & \partial^2 G/\partial w_p \partial l_2 & \ldots & \partial^2 G/\partial w_p \partial l_p \end{bmatrix}.
$$

We make the following assumptions on the moment conditions underlying (12) and the underlying estimator $\boldsymbol{\Gamma}_n \equiv \underline{l}_n$ for the spatial error variances and autocovariances.

---

[11]Equivalently, the implicit function $G\left(., .\right)$ can also be stated in terms of $\boldsymbol{\Gamma}$ and all orthogonal matrices $\boldsymbol{T}$, or $\boldsymbol{\Gamma}$ and all skew symmetric matrices $\boldsymbol{S}$, with $\boldsymbol{T} = (\boldsymbol{I} - \boldsymbol{S})(\boldsymbol{I} + \boldsymbol{S})^{-1}$.

Assumption 5 (estimator $\mathbf{\Gamma}_n$ and moment conditions)**:**

(a) $G\left(.,.\right)$ *is a totally differentiable function such that* $G\left(\underline{l},\underline{w}\right)=0$ *defines* $\underline{w}$ *uniquely for each* $\underline{l}$.

(b) *The gradient functions* g_i *are continuous with continuous first partial derivatives in an open set containing* $(\underline{l},\underline{w})$.

(c) $n^{1/2}\left(\underline{l}_n-\underline{l}\right)\xrightarrow{D} N\left(0,\mathbf{\Omega}\right)$.

(d) *Evaluated at the true parameters* $(\underline{l},\underline{w})$, *the Jacobian is nonsingular:* $|\mathbf{J}|\neq 0$, *and each row of* $\mathbf{J}^{-1}\mathbf{H}$ *contains at least one nonzero element.*

Assumption 5 is standard. Since the spaial weights matrix is symmetric (**Assumption 4**), **Assumption 5(a)** holds for the true parameters $\underline{l}$ by **Proposition 2** on uniqueness. In addition, we assume that the objective function is minimised at a unique $\underline{w}$ corresponding to each $\underline{l}$.[12] **Assumption 5(b)** follows from the assumed constraint in (14). **Assumption 5(c)** is a property of the underlying estimator of the spatial autocovariance matrix, and will be generally satisfied by any estimator based on a two-stage procedure, maximum likelihood or method of moments. **Assumption 5(d)** is a standard regularity condition on the smoothness of the implicit function in the neighbourhood of the true parameter values.

Proposition 3 *(**Consistency and Asymptotic Normality**) Let **Assumptions 1−5** hold. Then, as n increases, to each* $\mathbf{\Gamma}_n$ *there corresponds a unique solution* $\underline{w}_n$ *to the system of equations* $g=0$. *Further,* $\underline{w}_n\xrightarrow{P^*}\underline{w}$ *and* $n^{1/2}\left(\underline{w}_n-\underline{w}\right)$ *converges to a p-variate normal distribution with mean zero and covariance matrix* $\mathbf{J}^{-1}\mathbf{H}\mathbf{\Omega}\mathbf{H}^T\left(\mathbf{J}^{-1}\right)^T$, *where* $\mathbf{H}$ *and* $\mathbf{J}$ *are evaluated at the true parameter vector* $(\underline{l},\underline{w})$ *and* $\xrightarrow{P^*}$ *denotes convergence in outer probability.*

Proof. *By **Assumption 5(b)**, the gradient functions* g_i *are continuous with continuous first partial derivatives in an open set containing the true parameter vector* $(\underline{l},\underline{w})$. *Also by **Assumptions 5(a)** and **5(d)**,* $g=0$ *and* $|\mathbf{J}|\neq 0$ *at* $(\underline{l},\underline{w})$.

---

[12]The symmetric spatial weights matrix assumption may be unrealistic in some applications. Alternatively, **Assumption 5(a)** can be implied by other structural constraints on spatial weights and/or idiosyncratic variances (as in Bhattacharjee and Holly, 2008) or moment conditions (Bhattacharjee and Holly, 2010); for further discussions on appropriate identifying restrictions, see Bhattacharjee and Holly (2011).

Also, by the implicit function theorem (Taylor and Mann, 1983, p.225), there exists an open rectangular region $L \times \Psi$ around $(\underline{l}, \underline{w})$ (with $\underline{l} \in L$ and $\underline{w} \in \Psi$) and a set of p real functions f_i such that

$$\underline{w} = (w_1, w_2, \ldots, w_p) = (f_1(\underline{l}), f_2(\underline{l}), \ldots, f_p(\underline{l}))$$

and

$$g_i(\underline{l}, \underline{w}) = 0, \quad i = 1, 2, \ldots, p,$$

whenever $(\underline{l}, \underline{w}) \in L \times \Psi$. Further, the functions f_i are continuous and have continuous first order derivatives $(\partial f_i / \partial l_j)$, which are elements of the matrix $-\mathbf{J}^{-1}\mathbf{H}$.

Then, since $f_i(.)$'s are continuous functions in a neighbourhood of $(\underline{l}, \underline{w})$, for every sequence of random variables $x_n \in L$ such that $x_n \to x \in L$,

$$(f_1(x_n), f_2(x_n), \ldots, f_p(x_n)) \to (f_1(x), f_2(x), \ldots, f_p(x)).$$

*Further, since by **Assumption 5(c)**, $n^{1/2}(\underline{l}_n - \underline{l})$ converges to a Gaussian distribution with mean zero, we also have $\underline{l}_n \xrightarrow{P^*} \underline{l}$ (Serfling, 1980, p.26, Application C). By the extended continuous mapping theorem (Theorem 1.11.1 and Problem 1.11.1, van der Vaart and Wellner, 1996, p.67 and p.70), we therefore have*

$$\underline{w}_n \xrightarrow{P^*} \underline{w}.$$

Similarly, by a combination of the extended continuous mapping theorem for weak convergence together with the Delta method and Slutsky's Theorem (Serfling, 1980, p.122-124, and p.19), we also have

$$n^{1/2}(\underline{w}_n - \underline{w}) \xrightarrow{D} N\left(0, \mathbf{J}^{-1}\mathbf{H}\mathbf{\Omega}\mathbf{H}^T\left(\mathbf{J}^{-1}\right)^T\right). \tag{16}$$

■

Some qualifications are required for **Proposition 3**. First, to avoid measurability concerns, consistency considered here is in terms of outer probability.[13] If there are no measurability issues, outer probability (P^*) can be replaced by probability (P), and correspondingly convergence in outer probability can be replaced by convergence in probability.

Second, if $\mathbf{\Omega}_n$ is a consistent estimator for $\mathbf{\Omega}$, the above covariance matrix can be consistently estimated by substituting $\mathbf{J}^{-1}\mathbf{H}$ evaluated at $(\underline{l}_n, \underline{w}_n)$ in place of the population quantity. However, since our estimator is based on a generic estimated spatial autocovariance matrix $\mathbf{\Gamma}_n$, we acknowledge that

---

[13]The outer probability of an arbitrary set A is defined as $P^*(A) = \inf\{E(b) : b \text{ is measurable and } 1(A) \leq b\}$.

such an estimator for $\boldsymbol{\Omega}$ may not be readily available. In the following section, we discuss a bootstrap method to estimate the covariance matrix.

Third, the implicit relation (14) is based on functions of moments, which suggests that efficient GMM-type M-estimation may be possible. However, this is not straightforward, because the moment equations are complicated and only implicitly defined. While this paper focusses on exploring whether the spatial weights matrix is identified by the spatial autocovariance matrix and how the spatial weights can be estimated whenever a suitable estimator for the spatial autocovariance matrix is available, efficient estimation is retained as a topic for future research.

With regard to the objectives of this paper, **Propositions 1–3** show that the above estimation problem is partially identified, but a consistent and aymptotically normal estimator can be obtained whenever the estimator for the corresponding spatial autocovariance matrix is asymptotically Gaussian.

Next, we turn to computation of the optimal solution in (12) within the space of all orthogonal matrices $\boldsymbol{T}$. Jennrich (2001, 2004) proposed a "gradient projection" algorithm for optimising any objective function over the group of orthogonal transformations of a given matrix. The conditions necessary for implementing the algorithm are satisfied in our case: namely that (a) the objective function is differentiable (**Assumption 5(a)**), and (b) there exists a stationary point of the objective function within the class of orthogonal transformations (**Proposition 2**). Therefore, we adapt this algorithm to our case.

For a candidate estimate $\widehat{\boldsymbol{V}}$ and any orthogonal matrix $\boldsymbol{T}$, compute $\boldsymbol{Q} = \widehat{\boldsymbol{V}}.\boldsymbol{T}$ and define
(i) a scalar function of $\boldsymbol{T}$ (our objective function):

$$f(\boldsymbol{T}) = \sum_{i=1}^{K-1} \sum_{j=i+1}^{K} \left(\frac{q_{ij}}{q_{ii}} - \frac{q_{ji}}{q_{jj}} \right)^2, \tag{17}$$

(ii) a gradient matrix

$$\frac{df}{d\boldsymbol{T}} = \widehat{\boldsymbol{V}}^T.\boldsymbol{G}^*; \quad \boldsymbol{G}^* = \left((g_{ij}^*) \right)_{i,j=1,\ldots,K}, \tag{18}$$

$$\text{where} \quad g_{ii}^* = -\frac{2}{q_{ii}^2}.\sum_{\substack{k=1 \\ k \neq i}}^{K} q_{ik}.\left[\frac{q_{ik}}{q_{ii}} - \frac{q_{ki}}{q_{kk}} \right]$$

$$g_{ij}^* = \frac{2}{q_{ii}}.\left[\frac{q_{ij}}{q_{ii}} - \frac{q_{ji}}{q_{jj}} \right] \quad \text{if} \ i \neq j.$$

17

and (iii) a scalar constant

$$s = \left\| skm\left(\boldsymbol{T}^T.\boldsymbol{G}^*\right)\right\|, \tag{19}$$

where $skm(\boldsymbol{B}) = \frac{1}{2}\left(\boldsymbol{B} - \boldsymbol{B}^T\right)$ and $\|\boldsymbol{B}\| = tr\left(\boldsymbol{B}^T.\boldsymbol{B}\right)$ denote the skew symmetric part and the Frobenius norm respectively of a square matrix $\boldsymbol{B}$.

Algorithm: *Choose $\alpha > 0$ and a small $\varepsilon > 0$, and set $\boldsymbol{T}$ to some arbitrary initial orthogonal matrix.[14]*
(a) Compute s. If $s < \varepsilon$, stop.
(b) Compute $\boldsymbol{G}^*$.
(c) Find the singular value decomposition $\boldsymbol{A}.\boldsymbol{D}.\boldsymbol{C}^T$ of $\boldsymbol{T} - \alpha\boldsymbol{G}^*$; set $\widetilde{\boldsymbol{T}} = \boldsymbol{A}.\boldsymbol{C}^T$.
(d) Compute $\boldsymbol{Q} = \widehat{\boldsymbol{V}}.\widetilde{\boldsymbol{T}}$. If any diagonal element of $\boldsymbol{Q}$ is negative, multiply the corresponding column of $\widetilde{\boldsymbol{T}}$ by -1.
(e) If $f(\widetilde{\boldsymbol{T}}) \geq f(\boldsymbol{T})$, replace α by

$$\widetilde{\alpha} = \left(s^2\alpha^2\right) / \left[2.\left(f(\widetilde{\boldsymbol{T}}) - f(\boldsymbol{T}) + s^2\alpha\right)\right]$$

and go to (c). If $f(\widetilde{\boldsymbol{T}}) < f(\boldsymbol{T})$, replace $\boldsymbol{T}$ by $\widetilde{\boldsymbol{T}}$ and go to (a).
The estimate of $\boldsymbol{T}$ at convergence, $\boldsymbol{T}_n$, is such that all diagonal elements of the matrix

$$\boldsymbol{Q}_n = \widehat{\boldsymbol{V}}.\boldsymbol{T}_n = ((q_{ij,n}))_{i,j=1,\ldots,K}$$

are positive. Given this $\boldsymbol{T}_n$, the corresponding estimators of the spatial weights matrix and spatial error variances are given in (13).

Proposition 4 (*Computation*) *The algorithm is strictly monotone and stops when it is sufficiently close to a stationary point. The return to step (c) from (e) occurs only a finite number of times.*
*Under **Assumptions 1–5**, the final estimate, $\boldsymbol{T}_n$, of the relevant orthogonal matrix is such that $\boldsymbol{Q}_n$ has positive diagonal elements, and the estimated spatial weights matrix $\boldsymbol{W}_n$ is symmetric.*

Proof. *Note that, reversing the sign of each element in a specific column of $\boldsymbol{T}$ reverses the sign of the corresponding diagonal element of $\boldsymbol{Q} = \widehat{\boldsymbol{V}}.\boldsymbol{T}$, but*

---

[14]In our implementations, we choose a set of random orthogonal matrices, including the identity matrix, to ensure that the algorithm does not get stuck in a local stationary point that does not optimise the objective function.

preserves the orthogonality of $\boldsymbol{T}$. Therefore, step (d) ensures that the diagonal elements of $\boldsymbol{Q}$ are positive in each iteration of the algorithm. Effectively, our objective function is

$$f(\boldsymbol{T}) = \sum_{i=1}^{K-1} \sum_{j=i+1}^{K} \left(\frac{q_{ij}}{|q_{ii}|} - \frac{q_{ji}}{|q_{jj}|} \right)^2,$$

*which is not a differentiable function when $\boldsymbol{Q}$ has any zero diagonal element. However, by **Assumption 1** the spatial errors have positive variances, $\min\left(\sigma_1^2, \sigma_2^2, \ldots, \sigma_K^2\right) > 0$, and therefore in large samples the diagonal elements of $\boldsymbol{Q}$ will be away from zero.*[15]

Therefore, close to each candidate $\boldsymbol{T}$, $f(.)$ is locally a smooth differentiable function of $\boldsymbol{T}$ which is bounded below by zero. Hence, a stationary point exists.

*It can be verified easily that $\boldsymbol{G}^* = df/d\boldsymbol{T}$. Following arguments in Jennrich (2001), for sufficiently large α, the algorithm is convergent from any starting value and is converges to a stationary point of $f(.)$ over the group of all orthogonal matrices. Further, under **Assumption 4** (symmetry), the stationary point is unique and value at the stationary point is zero (**Proposition 2**). Therefore, under the maintained conditions on $\boldsymbol{\Gamma}_n$ (**Assumption 5**), $\boldsymbol{Q}_n = \widehat{\boldsymbol{V}}.\boldsymbol{T}_n \xrightarrow{P} \left[(\boldsymbol{I} - \rho\boldsymbol{W})^T .diag\left(\frac{1}{\sigma_1}, \frac{1}{\sigma_2}, \ldots, \frac{1}{\sigma_K} \right) \right].$*

*Note that, since the spatial weights matrix $\boldsymbol{W}$ has zero diagonal elements (**Assumption 2**), the elements on the diagonal of $(\boldsymbol{I} - \rho\boldsymbol{W})$ must all be unity. Hence, under the model assumptions, the diagonal elements of $\boldsymbol{Q}_n$ converge in probability to $\frac{1}{\sigma_1}, \frac{1}{\sigma_2}, \ldots, \frac{1}{\sigma_K}$ respectively. In other words, $\sigma_{k,n}^2 > 0$ for $k = 1, \ldots, K$. Further, since the value of the objective function at the stationary point is zero, the corresponding estimator $\boldsymbol{W}_n$ is symmetric.* ∎

The algorithm is simple to implement. There are, however, two important issues. First, the choice of α can be quite critical. If α is too large, the convergence is very slow; on the other hand, if α is too small, there will be many returns to step (c) from step (e). However, in practice this does not turn out to be a problem.[16] Second, there could be a problem with the implementation if Step (d) is executed too many times. This would mean that we are starting too far from the stationary point, and every time that one of the negative diagonal elements of $\boldsymbol{Q}$ is corrected for, we are moving

---

[15] This argument can be made more precise by imposing a boundedness assumption on the spatial error variances.

[16] In our implementation, we did some trial and error to choose a suitable α before starting the iterations. Once this choice was made, the convergence was very fast in about 95 per cent of our monte carlo experiments.

away from the stationary point instead of taking an optimal step along the gradient; in our implementation, we reduce this problem by choosing multiple starting values.

The assumption of a symmetric spatial weights matrix (**Assumption 4**) may be too restrictive in some applications. First, it is often convenient to work with row-standardised spatial weights matrices (Anselin, 1999), which are asymmetric by construction. Second, there are applications where it is reasonable to expect asymmetric strength of diffusion between regions. Third, if we do not assume homogeneity in the autoregression parameter ρ_k across the regions, the matrix $\boldsymbol{R}.\boldsymbol{W}$ will be asymmetric; such heterogeneity may be quite natural in many situations, like in a core-periphery structure or in relation to flows in a network.

However, the framework developed here is flexible and can admit many other sets of conditions. Useful constraints in this context could include setting the row sums of the estimated spatial weights matrix equal to each other,[17] or homoscedasticity in all or some of the spatial error variances, or constraints on specific spatial autocorrelations, or even a combination of several constaints; see Bhattacharjee and Holly (2008, 2011) for further discussion on suitable structural and moment restrictions. Alternatively, some of the optimisation criteria commonly used in factor analysis (quartimax, orthomax, etc.) could be useful in some spatial applications.

Another interesting feature of the estimator described here is that some of the estimated spatial weights in $\boldsymbol{W}_n$ may be negative. This happens if $q_{ij,n} > 0$ for some $i \neq j$. When the underlying spatial weights matrix has positive off-diagonal elements, we can have negative estimates of spatial weights because of sampling variations. However, spatial weights can be negative even in other situations. For example, in the context of housing demand across regions, such negative weights would imply that the excess demand in the index region is negatively related to some of the other regions. This can happen because of asynchronicity of the underlying housing market cycles in these regions, as would be expected for example if there are ripple effects (Meen, 1999), or if the two regions provide substitute housing markets. Meen (1996) explains negative interaction between regions in a study of housing starts as arising from planning restrictions in certain regions.

Finally, Meen (1996) emphasized an useful interpretation of the spatial dynamics within the spatial error model is in terms of an endogenous system of simultaneous equations where residuals for each region are regressed on those of the other regions (4). However, unlike usual simultaneous equations

---

[17]This would be relevant when all spatial weights are positive and one is interested in estimating the row standardised spatial weights matrix.

problems, we consider situations where there are no exogenous variables. In this sense, this paper is related to prior research on identification in simultaneous equations systems based on covariance restrictions (see, for example, Hausman and Taylor, 1983; Hausman *et al.*, 1987).

3 Extensions and Further Inference

3.1 Row-standardized Spatial Weights Matrix

A notable feature of our estimator $\boldsymbol{W}_n$ (or in the more general case, when **Assumption 4** is not imposed, $\boldsymbol{R}.\boldsymbol{W}$) is that the covariance pattern in the errors is determined solely by the product $\boldsymbol{R}.\boldsymbol{W}$. This observation has implications for the estimation of the spatial weights. A property of this estimator, which derives from the underlying spatial error model, is that the autoregression parameters ρ_k are in general not identifiable separately from the spatial weights matrix $\boldsymbol{W}$, so that only the product $\boldsymbol{R}.\boldsymbol{W}$ is usually estimable, and not the individual components $\boldsymbol{R}$ and $\boldsymbol{W}$.

Note, however, that the row-standardised spatial weights matrix, $\boldsymbol{W}^{(RS)}$, can be uniquely estimated as:

$$
\boldsymbol{W}_n^{(RS)} = \begin{pmatrix} 0 & \dfrac{\frac{q_{21,n}}{K}}{\sum\limits_{k=2}^{K} q_{k1,n}} & \cdots & \dfrac{\frac{q_{K1,n}}{K}}{\sum\limits_{k=2}^{K} q_{k1,n}} \\ \dfrac{\frac{q_{12,n}}{K}}{\sum\limits_{k=1,k\neq 2}^{K} q_{k2,n}} & 0 & \cdots & \dfrac{\frac{q_{K2,n}}{K}}{\sum\limits_{k=1,k\neq 2}^{K} q_{k2,n}} \\ \vdots & \vdots & \ddots & \vdots \\ \dfrac{\frac{q_{1K,n}}{K-1}}{\sum\limits_{k=1}^{K} q_{kK,n}} & \dfrac{\frac{q_{2K,n}}{K-1}}{\sum\limits_{k=1}^{K} q_{kK,n}} & \cdots & 0 \end{pmatrix} \tag{20}
$$

Further, $\boldsymbol{R}.\boldsymbol{W}$ is obtained by premultiplying the spatial weights matrix by a diagonal matrix whose diagonal elements are the spatial autoregressive parameters for each region. Hence, the estimates of the spatial autoregressive parameters corresponding to the above row-standardised spatial weights matrix are given by:

$$
\rho_{k,n}^{(RS)} = \sum_{l=1,l\neq k}^{K} q_{lk,n}, \quad k = 1,\ldots,K. \tag{21}
$$

3.2 Confidence Intervals

The asymptotic distribution given by (16) can be used to compute standard errors and confidence intervals for elements of the estimated spatial weights matrix. This is done by first consistently estimating the asymptotic covariance matrix of the variances and autocovariances of the spatial errors ($\mathbf{\Omega}$), and then substituting $\mathbf{J}^{-1}\mathbf{H}$ by its value at the estimated ($\underline{l}_n, \underline{w}_n$). However, these standard error estimates tend to have slow convergence and may be quite poor in small samples. As an alternative, we propose the bootstrap to construct confidence intervals around the elements of $\boldsymbol{W}_n$. It is well known that the bootstrap is valid when the statistics are smooth functions of the sample moments and the model can be consistently estimated.

While the bootstrap has also been found useful in simultaneous equations models (Freedman, 1984), its use here is substantially more complicated as compared to an equation involving only exogenous variables. This is because the bootstrap DGP (data generating process) has to provide a method to generate realisations of all the endogenous variables. Further, since our estimates of spatial weights are based on the estimated spatial autocovariance matrix, we require the bootstrap to be valid not only for the estimates of the regression function but also for estimates of all the spatial variances and autocovariances. Hence, in addition to the moment and smoothness conditions specified in the literature on bootstrap for SURE, 2SLS and 3SLS estimators, the proposed procedure would require additional conditions necessary for bootstrapping the spatial autocovariance matrix. In particular, following Beran and Srivastava (1985), we require the spatial weights matrix and the region-specific spatial variances to be such that the reduced form spatial autocovariance matrix is nonsingular and all eigenvalues of the matrix have unit multiplicity.[18] Therefore, we make the following additional assumption.

Assumption 6: *We assume the moment and smoothness condtions required for the validity of the bootstrap under SURE (Rilstone and Veall, 1996) or under 3SLS estimation (Fair, 2003). Following Beran and Srivastava (1985), these are conditions that ensure (a) weak convergence of the vector of reduced form spatial residuals $(\widehat{u}_1, \widehat{u}_2, \dots, \widehat{u}_K)$ to the corresponding vector of the error terms $(u_1, u_2, \dots, u_K)$, and (b) existence of finite fourth order moments of the error vector (i.e., $\left|\mathbb{E}\left(\prod_{k=1}^{K} u_k^{r_k}\right)\right| < \infty$ for every set of non-negative integers r_k such that $\sum_{k=1}^{K} r_k = 4$). In addition, we assume that the spatial weight matrix, $\boldsymbol{W}$, and the spatial variances of the structural*

---

[18]See Beran and Srivastava (1985) for consequences of the violation of this assumption, and description of a valid bootstrap procedure in the case of repeated eigenvalues.

equations, $\boldsymbol{\Sigma}$, are such that the reduced form spatial autocovariance matrix $\mathbb{E}\left(\underline{u}.\underline{u}^T\right) = (\boldsymbol{I} - \boldsymbol{R}.\boldsymbol{W})^{-1}.\boldsymbol{\Sigma}.(\boldsymbol{I} - \boldsymbol{R}.\boldsymbol{W})^{-1^T}$ *has distinct (non-zero) eigenvalues.*

Note that, since $\sigma_k^2 > 0$ for all $k = 1, \ldots, K$ (**Assumption 1**) and since the eigenvalues and eigenvectors of the spatial autocovariance matrix $\mathbb{E}\left(\underline{u}.\underline{u}^T\right) = (\boldsymbol{I} - \boldsymbol{R}.\boldsymbol{W})^{-1}.\boldsymbol{\Sigma}.(\boldsymbol{I} - \boldsymbol{R}.\boldsymbol{W})^{-1^T}$ are continuously differentiable with respect to $\mathbb{E}\left(\underline{u}.\underline{u}^T\right)$, the estimator of the spatial weights matrix $\boldsymbol{R}.\boldsymbol{W}_n$ (or $\rho\boldsymbol{W}_n$) is also a continuously differentiable function of $\mathbb{E}\left(\underline{u}.\underline{u}^T\right)$. It follows from Beran and Srivastava (1985) that, under the conditions of **Assumption 6**, the eigenvalues and eigenvectors of $\mathbb{E}\left(\underline{u}.\underline{u}^T\right)$ possess pivotal statistics which will validate a bootstrap procedure to estimate the sampling distribution of spatial weights estimates.

3.3 Testing for a given driver of spatial diffusion

We have motivated the estimators proposed in this paper based on uncertainty regarding the choice of spatial weights matrices and the arbitrariness regarding such choice in practice. It is, therefore useful to test the hypothesis that the observed pattern of spatial autocovariances has been generated by a hypothesized spatial weights matrix, $\boldsymbol{W}_0$:

$$\mathbb{H}_0 : \boldsymbol{W} = \boldsymbol{W}_0 \quad \text{versus} \quad \mathbb{H}_1 : \boldsymbol{W} \neq \boldsymbol{W}_0. \tag{22}$$

Under $\mathbb{H}_0$, the spatial weights matrix is known. Therefore one can use standard spatial econometric methods to estimate the unknown parameters of the spatial error model (Equation 1) and compute the spatial autocovariance matrix $\widehat{\boldsymbol{\Gamma}}_{\boldsymbol{W}_0} = \left(\boldsymbol{I} - \widehat{\boldsymbol{R}}\boldsymbol{W}_0\right)^{-1}.\widehat{\boldsymbol{\Sigma}}.\left(\boldsymbol{I} - \widehat{\boldsymbol{R}}\boldsymbol{W}_0\right)^{-1^T}$ consistent with the given spatial weights matrix. Since our proposed estimator is a unique transformation of the estimated spatial autocovariance matrix $\widehat{\boldsymbol{\Gamma}}$, the above test of hypothesis is equivalent to testing that $\widehat{\boldsymbol{\Gamma}}_{\boldsymbol{W}_0}$ is the same as $\boldsymbol{\Gamma}_n$.

Therefore, under the assumption of normal spatial errors, we can follow Ord (1975) and Mardia and Marshall (1984) to obtain MLEs of the unknown parameters by maximising the log-likelihood

$$\ln L\left(\boldsymbol{B}, \boldsymbol{\Sigma}, \boldsymbol{R} \,|\, \boldsymbol{W}_0\right) = \text{const.} - \frac{T}{2}\sum_{k=1}^{K}\ln\sigma_k^2 + T.\left|\boldsymbol{I} - \boldsymbol{R}\boldsymbol{W}_0\right| - \frac{1}{2}\sum_{t=1}^{T}\underline{\varepsilon}_t^T.\boldsymbol{\Sigma}^{-1}.\underline{\varepsilon}_t,$$

where $\underline{\varepsilon}_t = (\boldsymbol{I} - \boldsymbol{R}\boldsymbol{W}_0).(\underline{D}_t - \boldsymbol{X}_t\boldsymbol{B})$, $\boldsymbol{B} = \left(\underline{\beta}_1 : \underline{\beta}_2 : \ldots : \underline{\beta}_K\right)$ is a matrix whose columns correspond to regression coefficients for each region, and

$\boldsymbol{\Sigma} = \sigma^2 \boldsymbol{I}$ and $\boldsymbol{R} = diag\left(\rho_1, \rho_2, \ldots, \rho_K\right)$ are diagonal matrices containing the spatial error variances and spatial autocorrelation parameters respectively for each region. Maximum likelihood estimation in this setup is, in general, computation intensive unless the spatial autocorrelation parameters, ρ_k, are known in advance. Alternative GMM based estimation procedures are described in Kelejian and Prucha (1999) and Bell and Bockstael (2000). These GMM procedures are computationally simpler and, for reasonable sample sizes, almost as efficient as the MLE.

Having obtained the MLE or GMM estimates $\widehat{\boldsymbol{R}}$ and $\widehat{\boldsymbol{\Sigma}}$, we can construct the spatial covariance matrix $\widehat{\boldsymbol{\Gamma}}_{\boldsymbol{W}_0} = \left(\boldsymbol{I} - \widehat{\boldsymbol{R}}\boldsymbol{W}_0\right)^{-1}.\widehat{\boldsymbol{\Sigma}}.\left(\boldsymbol{I} - \widehat{\boldsymbol{R}}\boldsymbol{W}_0\right)^{-1^T}$ under $\mathbb{H}_0$, and then use a wide variety of tests available in the statistical literature for testing equality of two covariance matrices. In particular, we suggest the test statistic proposed by Ledoit and Wolf (2002); this test is valid when the estimated spatial covariance matrix is not full rank and when the number of regions increases with sample size. This situation may be relevant in many microeconomic applications where the number of agents increase asymptotically, or if we are interested in finer spatial aggregation as we accumulate more data.

The additional assumptions regarding the distribution of the errors and the nature of asymptotics are as follows.

Assumption 7: *The spatial errors, $\underline{\varepsilon}_t$, are normally distributed.*

Under **Assumption 7**, we use the Cholesky decomposition of $\widehat{\boldsymbol{\Gamma}}_{\boldsymbol{W}_0} = \boldsymbol{D}_{\boldsymbol{W}_0}^T.\boldsymbol{D}_{\boldsymbol{W}_0}$, and restate the null and alternative hypotheses (22) as:

$$\mathbb{H}_0 : \widehat{\boldsymbol{\Gamma}}_{\boldsymbol{W}_0}^* = \boldsymbol{I} \quad \text{versus} \quad \mathbb{H}_1 : \widehat{\boldsymbol{\Gamma}}_{\boldsymbol{W}_0}^* \neq \boldsymbol{I}, \tag{23}$$

where $\widehat{\boldsymbol{\Gamma}}_{\boldsymbol{W}_0}^* = \left(\boldsymbol{D}_{\boldsymbol{W}_0}^T\right)^{-1}.\widehat{\boldsymbol{\Gamma}}.\left(\boldsymbol{D}_{\boldsymbol{W}_0}\right)^{-1}$. For the above hypotheses, the Ledoit-Wolf test statistic is given by:

$$
\begin{aligned}
LW \quad = \quad & \frac{1}{K}.tr\left(\widehat{\boldsymbol{\Gamma}}_{\boldsymbol{W}_0}^* - \boldsymbol{I}\right)^2 - \frac{K}{n}.\left[\frac{1}{n}.tr\left(\widehat{\boldsymbol{\Gamma}}_{\boldsymbol{W}_0}^*\right)\right]^2 + \frac{K}{n}, \\
& \frac{Kn}{2}.LW \sim \chi^2\left(\frac{K(K+1)}{2}\right) \qquad \text{under } \mathbb{H}_0 \text{ as } T \to \infty
\end{aligned}
\tag{24}
$$

where $tr(.)$ denotes trace of a square matrix. As demonstrated by Ledoit and Wolf (2002), the test has very good small sample performance.

3.4 Extensions to other spatial models

3.4.1 Spatial error model with moving average errors

Under the spatial error model with moving average errors (2), the spatial weights matrix can be estimated in a very similar manner. In this model, the spatial autocovariance matrix is given by:

$$
\begin{aligned}
\boldsymbol{\Gamma} &= \mathbb{E}\left(\underline{u}.\underline{u}^T\right) = \left(\boldsymbol{I} + \boldsymbol{R}.\boldsymbol{W}\right).\boldsymbol{\Sigma}.\left(\boldsymbol{I} + \boldsymbol{R}.\boldsymbol{W}\right)^T \\
&= \boldsymbol{Z}.\boldsymbol{Z}^T \\
&= \left(\boldsymbol{Z}.\boldsymbol{T}\right).\left(\boldsymbol{Z}.\boldsymbol{T}\right)^T,
\end{aligned}
$$

where

$$
\boldsymbol{Z} = \left(\boldsymbol{I} + \boldsymbol{R}.\boldsymbol{W}\right).diag\left(\sigma_1, \sigma_2, \ldots, \sigma_K\right),
$$

and $\boldsymbol{T}$ is any orthogonal matrix. Therefore, under the symmetric spatial weights matrix assumption (**Assumption 4**), we estimate $\boldsymbol{Q}_n$ exactly in the same way as the SEM-AR model. The estimator for the spatial weights matrix (assuming, without loss of generality, that $\rho \equiv 1$) is now given by:

$$
\boldsymbol{W}_n^{(MA)} = \begin{bmatrix}
0 & q_{21,n}/q_{22,n} & \cdots & q_{K1,n}/q_{KK,n} \\
q_{12,n}/q_{11,n} & 0 & \cdots & q_{K2,n}/q_{KK,n} \\
\vdots & \vdots & \ddots & \vdots \\
q_{1K,n}/q_{11,n} & q_{2K,n}/q_{22,n} & \cdots & 0
\end{bmatrix}. \tag{25}
$$

3.4.2 Spatial lag model

Estimation of the spatial weights matrix for the spatial lag model (3) is in principle similar to our main case – the spatial error model with autoregressive errors. However, in this case, one would require a priori an estimator of the spatial autocovariance matrix of the response variable ($\underline{D}$) conditional on the exogenous regressors ($\boldsymbol{X}$). This can be obtained by a matrix of partial variances and covariances: $\widehat{\Sigma}_{\underline{D}.\boldsymbol{X}}$. Estimation of this partial autocovariance matrix is standard in the multivariate statistical analysis literature; see for example, McDonald (1978).

4 Monte Carlo Study

This section investigates the performance of the proposed estimators of the spatial weights matrix under autoregressive (13) and moving average (25) spatial error processes, and compares the performance with that of the residual regression estimator (Meen, 1996). The design of the Monte Carlo study

is similar to Baltagi *et al.* (2003) and calibrated to the study of β-convergence across the 9 census regions in the U.S.

Barro and Sala-i-Martin (1992) studied convergence in the U.S. over the period 1963 to 1986 using data on per capita gross state product for 48 U.S. states They estimated a cross-section regression of the form

$$y_i = \alpha + \beta.x_i + \varepsilon_i, \;\; i = 1, \ldots, 48,$$

where y denotes the average growth rate in per capita income over the 23 years, x denotes the log of per capita income in 1963, and $\beta < 0$ measures the strength of convergence. While estimating similar regression equations across different time periods, Rey and Montouri (1999) observe significant spatial autocorrelation among the states.

Using data on per capita state domestic product reported in Barro and Sala-i-Martin (1992), we estimated separate regression equations for states within the 9 different census regions and find evidence of heterogeneity in the rate of β-convergence across the regions; there is stronger convergence in the contiguous regions South Atlantic, East South Central, East North Central and West North Central as compared to the other regions. Further, there is heterogeneity in the initial levels of per capita income between the southern states and the non-southern states. Given this descriptive analysis, we set up a simulation model incorporating heterogeneity across the census regions, both in the rate of convergence and in initial income, and model the spatial autocorrelation using a spatial weights matrix approximately consistent with first-order contiguity.

The model is set up as follows. Spatial panel data y_{it} and x_{it} are generated for the 9 census regions for T periods of time according to the DGP

$$y_{it} = \alpha_i + \beta_i.x_{it} + u_{it}, \;\; i = 1, \ldots, 9, t = 1, \ldots, T,$$

where the regression coefficients vary across the regions. The regressor x is independently distributed as $N\left(\mu_i, 0.15^2\right)$ with different means across the regions. We choose the parameter values based on our descriptive analysis and regression results (Table 1).

The spatial errors, $\underline{u}_t$, in each period of time t are modelled as (a) a spatial autoregressive process $\left(\underline{u}_t = \rho.\boldsymbol{W}.\underline{u}_t + \underline{\varepsilon}_t\right)$ and (b) a spatial moving average process $\left(\underline{u}_t = \rho.\boldsymbol{W}.\underline{u}_t + \underline{\varepsilon}_t\right)$. In both cases, the pattern of spatial interactions is generated by a spatial weights matrix (approximately a first-order contiguity matrix) between the 9 census regions (Table 2A). This spatial weights matrix $(\rho\boldsymbol{W})$ is chosen to ensure that both $(\boldsymbol{I} - \rho\boldsymbol{W})$ and $(\boldsymbol{I} + \rho\boldsymbol{W})$ are strictly diagonally dominant and has no repeated eigenvalues.

Table 1: Region-specific parameters[19]

Regions	α_i	β_i	μ_i
NENG	0.047	-0.011	2.25
MATL	0.047	-0.011	2.25
SATL	0.073	-0.024	2.00
ESC	0.073	-0.024	2.00
WSC	0.047	-0.011	2.00
ENC	0.073	-0.024	2.25
WNC	0.073	-0.024	2.00
MTN	0.047	-0.011	2.25
PAC	0.047	-0.011	2.25

The idiosyncratic error term, ε_{it}, is independently and identically distributed as $N\left(0, \sigma_\varepsilon^2\right)$, where the parameter value for σ_ε^2 ($= 3.0e$-9) is chosen to approximately match the trace of the resulting spatial autocovariance matrix to that of the independent error OLS regression estimates.

We generate data from the above DGP for various sample sizes ($n = 25, 50, 100$) and estimate the parameters using maximum likelihood SURE estimates. The estimates of the spatial weights matrix are computed using both the above SURE estimator of the spatial autocovariance matrix (our proposed estimator), and regression of the SURE residuals (Meen, 1996). Since SURE performs poorly in many applications where the number of equations is large, we also repeat the analysis for the three contiguous regions Middle Atlantic, South Atlantic and East South Central; each of these regions is first order contiguous with the other two. The reported results are based on 1000 Monte Carlo replications of the above simulation scheme.

The estimated spatial weights matrix for the spatial error autoregressive model, $\boldsymbol{W}_n$, based on $n = 100$ and averaged over the 1000 Monte Carlo replications are presented in the bottom panel of Table 2 (Table 2 B), along with 95 per cent confidence intervals (2.5 and 97.5 bootstrap percentiles).

None of the zero spatial weights in Table 2 A have significantly positive or negative estimates in Table 2 B; the estimates of all the true positive weights are significantly positive at least at the 10 per cent level (in fact, most are significant at 1 per cent level). The average bias across the 81 elements in the spatial weights matrix is 0.0011 and the average RMSE (root mean squared error) is 0.0511. This is very good, particularly since we are estimating a large number of spatial weights.

---

[19] $NENG$: New England; $MATL$: Middle Atlantic; $SATL$: South Atlantic; ESC: East South Central; WSC: West South Central; ENC: East North Central; WNC: West North Central; MTN: Mountain; PAC: Pacific.

Table 2: Spatial Weights Matrix (Actual and Simulations)[20]

<div align="center">A. Spatial Weights Matrix (Actual), $\rho\boldsymbol{W}$</div>

	NENG	MATL	SATL	ESC	WSC	ENC	WNC	MTN	PAC
NENG	0	0.25	0	0	0	0.167	0	0	0
MATL	0.25	0	0.25	0.125	0	0.125	0	0	0
SATL	0	0.25	0	0.25	0	0	0	0	0
ESC	0	0.125	0.25	0	0.125	0.125	0	0	0
WSC	0	0	0	0.125	0	0	0.125	0.167	0.167
ENC	0.167	0.125	0	0.125	0	0	0.125	0	0
WNC	0	0	0	0	0.125	0.125	0	0.167	0.167
MTN	0	0	0	0	0.167	0	0.167	0	0.167
PAC	0	0	0	0	0.167	0	0.167	0.167	0

<div align="center">B. Estimated Symmetric Spatial Weights Matrix, $\rho\boldsymbol{W}_n$</div>

	NENG	MATL	SATL	ESC	WSC	ENC	WNC	MTN	PAC
NENG	**0**								
MATL	0.261** (.17,.34)	**0**							
SATL	−0.003 (-.10,.09)	0.244** (.16,.32)	**0**						
ESC	0.001 (-.09,.10)	0.129** (.03,.23)	0.252** (.14,.35)	**0**					
WSC	0.001 (-.11,.10)	0.001 (-.10,.12)	0.000 (-.10,.10)	0.127** (.03,.23)	**0**				
ENC	0.168** (.06,.27)	0.123* (.01,.22)	0.002 (-.09,.11)	0.125* (.02,.23)	−0.002 (-.10,.10)	**0**			
WNC	0.000 (-.08,.10)	0.001 (-.10,.12)	0.000 (-.10,.11)	0.002 (-.11,.10)	0.128* (.03,.23)	0.122** (.02,.21)	**0**		
MTN	0.002 (-.10,.10)	−0.004 (-.09,.09)	−0.002 (-.08,.10)	−0.001 (-.10,.10)	0.167** (.07,.27)	−0.001 (-.09,.09)	0.164** (.06,.26)	**0**	
PAC	0.001 (-.08,.10)	0.001 (-.10,.09)	0.006 (-.08,.12)	−0.004 (-.10,.09)	0.171** (.06,.26)	0.000 (-.12,.10)	0.173** (.07,.28)	0.170** (.08,.27)	**0**

In Table 3, we report average bias, standard deviation and root mean squared errors (RMSE) for the two estimators of the spatial weights matrix. Similar statistics for the SURE regression estimates are also reported. Results are presented both for the spatial error model with autoregressive errors (see also Table 2) and for the spatial error model with moving average errors.

---

[20]Abbreviations for the regions are as in Table 1.

Estimates reported are averages based on 1000 Monte Carlo simulations with $T = 100$.

Figures in parentheses are 95 per cent confidence intervals based on percentiles from the Monte Carlo simulations.

**,*, +: Significant at 1 per cent, 5 per cent and 10 per cent level respectively.

The results show that the residual regression based estimator is biased and inconsistent. The proposed estimator based on the SURE-estimated spatial autocovariance matrix performs quite well even for reasonably small sample sizes; even with 9 regions (36 unique elements in the spatial weights matrix) the average bias and RMSE for a sample size of $n = 50$ are quite reasonable. The performance improves quite substantially with sample size.

Table 3: Monte Carlo Results –
Performance of the Proposed Estimator

	9 REGIONS			3 REGIONS		
	$T = 25$	$T = 50$	$T = 100$	$T = 25$	$T = 50$	$T = 100$
SEM − AR Model						
REGRESSION COEFF.						
– Average bias	$-1.39e$-7	$7.11e$-6	$-3.75e$-6	$2.07e$-5	$7.51e$-5	$3.87e$-5
– Average RMSE	0.0081	0.0047	0.0030	0.0063	0.0045	0.0028
SPAT.ERR.STD.DEV.						
– Average bias	$5.68e$-4	$2.33e$-4	$1.14e$-4	$1.70e$-4	$6.95e$-5	$4.82e$-5
– Average RMSE	$7.05e$-4	$3.67e$-4	$2.31e$-4	$4.47e$-4	$2.95e$-4	$2.15e$-4
SPAT. WTS. MATRIX						
Proposed estimator						
– Average bias	$-7.24e$-3	$-3.49e$-3	$-5.33e$-4	$-5.14e$-3	$-2.30e$-3	$-3.07e$-5
– Average std. dev.	0.1391	0.0753	0.0489	0.0897	0.0585	0.0410
– Average RMSE	0.1393	0.0754	0.0489	0.0898	0.0586	0.0410
Residual regression est.						
– Average bias	0.0271	0.0284	0.0292	0.0850	0.0876	0.0887
– Average std. dev.	0.2236	0.1388	0.0937	0.1307	0.0939	0.0661
– Average RMSE	0.2326	0.1507	0.1101	0.1568	0.1285	0.1107
SEM − MA Model						
REGRESSION COEFF.						
– Average bias	$-1.39e$-7	$7.11e$-6	$-3.75e$-6	$2.07e$-5	$7.51e$-5	$3.87e$-5
– Average RMSE	0.0081	0.0047	0.0030	0.0063	0.0045	0.0028
SPAT.ERR.STD.DEV.						
– Average bias	$1.17e$-4	$7.49e$-5	$3.66e$-5	$8.36e$-5	$4.62e$-5	$2.92e$-5
– Average RMSE	$4.29e$-4	$3.00e$-4	$2.07e$-4	$4.17e$-4	$2.25e$-4	$2.04e$-4
SPAT. WTS. MATRIX						
Proposed estimator						
– Average bias	$-4.51e$-3	$-2.21e$-3	$-1.45e$-3	$-8.06e$-3	$-3.08e$-3	$-2.52e$-3
– Average std. dev.	0.1113	0.0697	0.0470	0.1126	0.0876	0.0391
– Average RMSE	0.1114	0.0697	0.0470	0.1127	0.0879	0.0391

5 Spillovers in Regional Housing Demand

Substantial literature on the UK housing market has accumulated over the past three decades; research highlights sluggish growth in supply, mismatch between demand and supply at least in a localised context (by locality and type of housing), an extremely low and declining price-elasticity of supply, low response of demand to price signals, geographically varying price effects (spatial heterogeneity) and spatial dependence (Meen, 2003; Barker, 2003). Attempts have been made to explain spatial diffusion by neighbourhood conditions[21] (Meen, 2001; Cheshire and Sheppard, 2004; Gibbons, 2004; Gibbons and Machin, 2005), social interaction and segregation (Meen and Meen, 2003) and heterogenous effect of common shocks (Holly *et al.*, 2011).

The above literature abounds in implicit acknowledgement of the strong spatio-temporal dependence in features of regional or local housing markets. However, barring some notable exceptions,[22] what is distinctly missing in the literature is adequate understanding of the reasons behind spatial or spatio-temporal interactions. To examine this issue, we employ an economic model, combining a traditional 'supply and demand' model of housing markets with a micro-founded model of search and bargaining in local housing markets. Our focus lies in applying the methods proposed in this paper to estimate the implied spatial weights matrix for housing demand, and thereby provide new inferences regarding the nature of spatial diffusion across the different regions of England and Wales.

5.1 The Data

The empirical analysis covers housing markets in England and Wales over a 48-month period April 2001 to March 2005. The spatial units of analysis are the ten government office regions in England and Wales (Figure 1). Monthly data on local housing markets at 3-digit postcode level were obtained from Hometrack, an independent property research and database company in the UK.[23] The variables included are:

---

[21] Such as crime rates, schooling, transport infrastructure and quality of public services.

[22] See, for example, Meen (1996) and Holly *et al.* (2011).

[23] The Hometrack data are based on monthly responses to a questionnaire by about 3,500 major estate agents in the UK. The data are rather unique in providing information on time on the market and degree-of-overpricing, both of which have important roles in our analysis. We benchmark this information with data from other sources. In particular, we augment the Hometrack data with quarterly information on sales price and number of sales by type of property, for each county and local/ unitary authority, collected from HM Land Registry of England and Wales.

Figure 1: **Government Office Regions (GORs) in England and Wales**

- Average number of views;

- Average time on the market (TOM); and

- Average final to listing price ratios (reciprocal of degree-of-overpricing).

Additional regional spatio-temporal data on supply, demand, neighbour-hood characteristics and market conditions were collected (see Appendix 1). Data were also collected for other variables useful in interpreting the estimated spatial weights.

5.1.1 Structural Equations

The structural equations of our model of housing markets[24] include four relationships in the four endogenous variables – namely, prices, demand, degree-

---

[24]For further details on the model, see Bhattacharjee and Jensen-Butler (2005).

of-overpricing and time on the market. Demand is endogenously determined but supply is exogenous.[25] In equilibrium, supply (S_t) is related to demand (D_t) as

$$D_t \equiv (1 - \nu_t).S_t, \tag{26}$$

where ν_t denotes the vacancy rate.

The realised value (price) (V_t) of housing properties follows a rental adjustment model (Hendershott, 1996), relating change in realised value (price) (V_t) of housing properties to deviations of the vacancy rate from the natural vacancy rate and deviations of the realised value from its natural (equilibrium) level. Then, assuming that natural value is fixed in the short run, we have:

$$\begin{aligned}
\Delta \ln V_t &= \gamma_1 \Delta \ln(1 - \nu_{t-1}) + \gamma_2 \Delta \ln V_{t-1} + \epsilon_{1t} \\
&= \gamma_1 \Delta \ln D_{t-1} + \gamma_2 \Delta \ln V_{t-1} - \gamma_3 \Delta \ln S_{t-1} + \epsilon_{1t}, \tag{27}
\end{aligned}$$

$$0 < \{\gamma_1, \gamma_2\} < 1, \gamma_3 = \gamma_1.$$

Demand is modelled as a function of realised value, housing market conditions and neighbourhood characteristics. The market conditions include economic activity (Y_t; local and economy-wide income, unemployment, productivity and interest rates) and the neighbourhood characteristics include socio-economic variables (X_t; quality of education and public services, demographics, *etc.*). Hence, change in demand is explained by change in local (neighbourhood characteristics), change in price, and change in (local) income or other indicators of local market conditions.

$$\Delta \ln D_t = \lambda_1 \Delta X_t + \lambda_2 \Delta \ln V_t + \lambda_3 \Delta \ln Y_t + \epsilon_{2t}, \tag{28}$$

where $\lambda_2 < 0$ is the price elasticity and $\lambda_3 > 0$ may be regarded as the income elasticity.

If vacancy rates were perfectly observed, the above three relationships (Equations (26), (27) and (28)) would become recursive and the structural relationships could be easily estimated. However, data on vacancy rates for the residential housing market in the UK are not available at a high level of temporal disaggregation. Therefore, we use microeconomic features of the housing markets, specifically the degree-of-overpricing (DOP_t) and time-on-the-market (TOM_t), to identify the wedge between demand and supply.

---

[25]Housing supply has changed only marginally over time for each region under study. We assume that supply in each region is exogenously determined by planning, land use and other local constraints.

Following Anglin *et al.*(2003),

$$\Delta \ln DOP_t = \alpha_1 \Delta X_t + \alpha_2 \Delta \ln Y_t + \alpha_3.\Delta \ln D_t - \alpha_4 \Delta \ln V_t + \epsilon_{4t}, \qquad (29)$$

$$\alpha_4 = 1.$$

and

$$\Delta TOM_t = \beta_1 \Delta \ln Y_t + \beta_2 \Delta \ln DOP_t - \beta_3 \Delta \ln D_t + \beta_4 \Delta \ln S_t + \epsilon_{5t}, \qquad (30)$$

$$\beta_4 = \beta_3.$$

The system comprising the above four simultaneous equations (Equations (27), (28), (29) and (30)) is overidentified. The four endogenous variables ($\Delta \ln V_t$, $\Delta \ln D_t$, $\Delta \ln DOP_t$ and ΔTOM_t) are measured in first differences, and tests indicate that each of these variables are stationary over time in each of the 10 regions under consideration, as well as in a panel. Hence, we estimate the relationships simultaneously using a 3SLS method, where the spatial autocovariance matrix estimated from the residuals incorporates unrestricted spatial interaction across the different regions.

In the first stage of our estimation procedure, we obtain predictions of the endogenous variables in the model from estimated regression models including the exogenous variables (supply, neighbourhood characteristics and market conditions) and lagged endogenous variables (lagged demand and prices). Following Bound *et al.* (1995), we check the F-statistics of the first stage regressions for each of the endogenous variables in our model, and verify that the instruments in our estimated model are well-specified; all these F-statistics are in excess of 200.

At the second stage, we use these predictions to obtain estimates of the four structural equations individually for each region, using measures of demand (average number of views per week), realised value (price), degree-of-overpricing, time on the market, neighbourhood characteristics (unfit houses, access to education, and crime detection rates) and indicators of market conditions (claimant counts and average household income). This allows for heterogeneity in the relationships across the regions, both in the sense of intercept (spatial fixed effects) and slope heterogeneity, and in the choice of indicators for neighbourhood characteristics and market conditions.

In the third stage, we estimate our structural equations separately for each region using 3SLS, and estimate the spatial autocovariance matrix for the demand equation from the second-stage residuals.

The estimates of the structural equations are consistent with *a priori* expectation and qualitatively very similar to the second stage SURE estimates

reported in Bhattacharjee and Jensen-Butler (2005). The results are not reported here since we are more interested in analysing the spatial errors from the demand equation; however, we briefly discuss the estimates.

We find substantial heterogeneity in the coefficients and in the specification of the relationships across the various regions. For the demand relationship, which is our main focus in this paper, the coefficient of the price variable is negative (significantly for most regions), but with substantial slope heterogeneity. Neighbourhood characteristics have an important effect on demand, where share of unfit houses is negatively related to demand in most of the regions, and access to education and crime detection have positive effects. Similarly, market conditions (claimant counts and average household income) have the expected signs (negative and positive respectively).

5.2   Spatial Diffusion of Demand

The spatial autocovariance matrix of demand is estimated from the second-stage residuals of the demand equation. Separate equations are estimated for each region, and we allow spatial errors in demand to be contemporaneously correlated across the regions. This is consistent with the spatial regime model allowing heterogeneity in the demand relationship. Further, we allow for spatial autocorrelation in the regression errors, and use the estimated spatial autocovariance matrix to estimate the implied matrix of spatial weights, $\boldsymbol{W}_n$ (13).

The spatial autocovariance matrix across the 10 regions (Table 4 A) and the corresponding spatial autocorrelation matrix (Table 4 B) indicate strong spatial spillovers. In particular, housing demand in the three spatially contiguous regions Greater London, the South East and East of England are strongly correlated. Likewise, the strong correlation between North West and North East appear to be related to geographical distances and contiguity. However, there are also significant correlations between Yorkshire and Humberside and the South East, and similarly between East Midlands and the South West. These patterns indicate that spatial patterns in demand are not necessarily related to simple notions of contiguity and geographical distance.

Table 4: Estimated Error Spatial Autocorrelation Matrix
(autocovariances in parentheses)[26]

	E	EM	L	NE	NW	SE	SW	W	WM	YH
E	**1.00** (0.0102)									
EM	0.575 (0.0047)	**1.00** (0.0066)								
L	0.568 (0.0036)	0.471 (0.0024)	**1.00** (0.0039)							
NE	0.673 (0.0129)	0.436 (0.0067)	0.356 (0.0042)	**1.00** (0.0360)						
NW	0.583 (0.0064)	0.417 (0.0037)	0.351 (0.0024)	0.799 (0.0166)	**1.00** (0.0120)					
SE	0.673 (0.0051)	0.513 (0.0031)	0.602 (0.0028)	0.503 (0.0071)	0.370 (0.0030)	**1.00** (0.0056)				
SW	0.428 (0.0025)	0.499 (0.0024)	0.366 (0.0013)	0.451 (0.0050)	0.436 (0.0028)	0.392 (0.0017)	**1.00** (0.0034)			
W	0.420 (0.0048)	0.299 (0.0027)	0.288 (0.0020)	0.450 (0.0097)	0.474 (0.0059)	0.412 (0.0035)	0.343 (0.0023)	**1.00** (0.0128)		
WM	0.465 (0.0054)	0.353 (0.0033)	0.274 (0.0019)	0.479 (0.0104)	0.333 (0.0042)	0.517 (0.0044)	0.409 (0.0027)	0.483 (0.0062)	**1.00** (0.0130)	
YH	0.487 (0.0065)	0.470 (0.0051)	0.461 (0.0038)	0.449 (0.0113)	0.482 (0.0070)	0.560 (0.0056)	0.467 (0.0036)	0.514 (0.0077)	0.452 (0.0069)	**1.00** (0.0177)

In order to understand the nature of spatial diffusion in terms of spatial weights, we use our estimator $\boldsymbol{W}_n$ (13) to estimate the symmetric spatial weights matrix under the assumption of a spatial error model with autoregressive errors (1). These estimates, reported in the top panel of Table 5 (Table 5A), measure the spatial contribution of demand in other regions to the determination of housing demand in each individual region. Put differently, the elements in a row corresponding to the index region represent, within the context of the estimated spatial autoregressive model, the contributions of idiosyncratic excess demand from other regions to excess demand in the index region. The corresponding estimated row-standardised spatial weights matrix is reported in the lower panel of Table 5 (Table 5B). Bootstrap 95 percent confidence intervals for each element of the spatial weights matrix, based on 200 bootstrap resamples, and the estimates of region-specific standard deviations of the spatial errors are also reported in Table 5A. The bootstrap confidence limits are used to identify statistically significant spatial weights in Table 5A.

---

[26]E: East of England; EM: East Midlands; L: Greater London; NE: North East; NW: North West; SE: South East; SW: South West; W: Wales; WM: West Midlands; YH: Yorks & Humberside.

Table 5: Estimated Spatial Weights Matrix[27]

A. Symmetric Spatial Weights Matrix, $\rho \boldsymbol{W}_n$

	E	EM	L	NE	NW	SE	SW	W	WM	YH
E	**0**									
EM	.145** (.08,.36)	**0**								
L	.148* (.01,.28)	.079 (-.10,.16)	**0**							
NE	.191* (.05,.33)	.009 (-.15,.17)	−.039 (-.17,.07)	**0**						
NW	.095[+] (-.02,.30)	.044 (-.08,.25)	.041 (-.18,.18)	.412** (.28,.57)	**0**					
SE	.201** (.10,.30)	.078 (-.14,.19)	.194** (.08,.35)	.094 (-.08,.24)	−.085* (-.30,-.01)	**0**				
SW	.007 (-.15,.10)	.147* (.01,.29)	.054 (-.12,.20)	.061 (-.15,.16)	.062 (-.08,.22)	.004 (-.17,.21)	**0**			
W	.031 (-.14,.18)	−.017 (-.20,.12)	.001 (-.13,.22)	.022 (-.06,.20)	.119* (.01,.30)	.045[+] (-.03,.16)	.022 (-.15,.16)	**0**		
WM	.056 (-.06,.12)	.018 (-.14,.19)	−.045 (-.12,.13)	.126** (.02,.24)	−.066 (-.20,.09)	.132[+] (-.01,.28)	.092 (-.06,.21)	.146* (.02,.26)	**0**	
YH	.004 (-.32,.26)	.079[+] (-.06,.20)	.083 (-.08,.19)	−.014 (-.14,.21)	.113[+] (-.01,.31)	.141** (.03,.25)	.095 (-.03,.26)	.143 (-.10,.32)	.074 (-.09,.33)	**0**
$\widehat{\sigma_k}$	.063 (.05,.07)	.062 (.04,.07)	.048 (.04,.05)	.109 (.08,.11)	.067 (.04,.07)	.049 (.04,.06)	.047 (.03,.05)	.090 (.07,.11)	.089 (.06,.09)	.098 (.06,.11)

B. Row-Standardised Spatial Weights Matrix, $\boldsymbol{W}_n^{(RS)}$[28]

|	E	EM	L	NE	NW	SE	SW	W	WM	YH
E	0	**0.16**	**0.17**	**0.22**	0.11	**0.23**	0.01	0.04	0.06	0.00
EM	**0.25**	0	0.14	0.02	0.08	0.13	**0.25**	−0.03	0.03	0.14
L	**0.29**	0.15	0	−0.08	0.08	**0.38**	0.10	0.00	−0.09	0.16
NE	**0.22**	0.01	−0.04	0	**0.48**	0.11	0.07	0.03	**0.15**	−0.02
NW	0.13	0.06	0.06	**0.56**	0	**−0.12**	0.08	**0.16**	−0.09	0.15
SE	**0.25**	0.10	**0.24**	0.12	**−0.11**	0	0.01	0.06	0.16	**0.18**
SW	0.01	**0.27**	0.10	0.11	0.11	0.01	0	0.04	0.17	0.17
W	0.06	−0.03	0.00	0.04	**0.23**	0.09	0.04	0	**0.28**	0.28
WM	0.10	0.03	−0.09	**0.24**	−0.13	0.25	0.17	**0.27**	0	0.14
YH	0.01	0.11	0.11	−0.02	0.16	**0.20**	0.13	0.20	0.10	0

Several elements in the estimated symmetric spatial weights matrix (Table 5A) are significantly different from zero, either positive or negative. The

---

[27] Abbreviations for the regions are as in Table 4.

+, *, **: Significant at 10 per cent, 5 per cent and 1 per cent level respectively.

[28] Figures in bold in the estimated row-standardised spatial weights matrix correspond to the symmetric spatial weights that are statistically significant at the 5% level.

Figure 2: **Spatial network of GORs on implied topological plane**
(Distances inversely proportional to absolute values of estimated spatial weights)
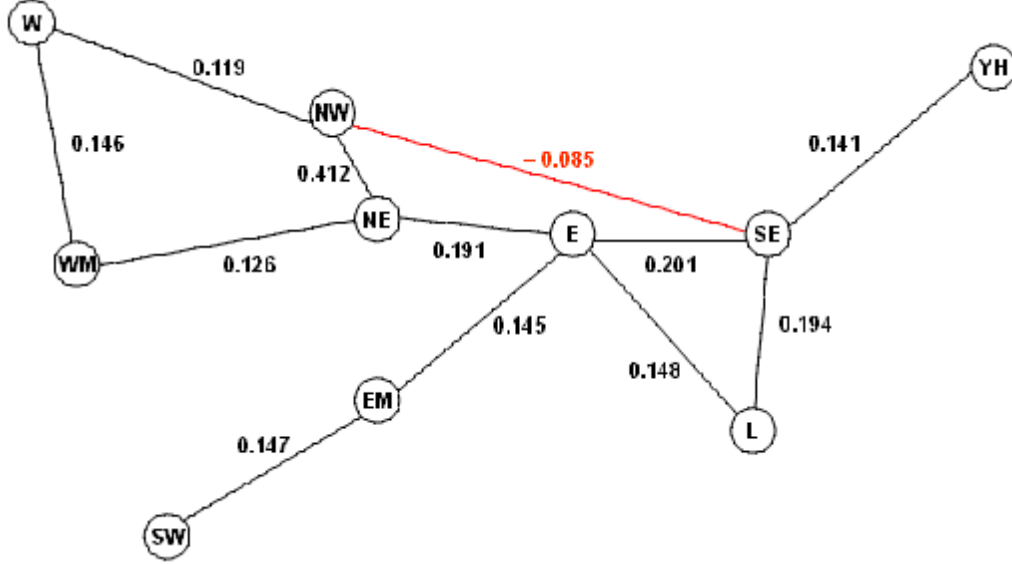
corresponding elements in the row-standardized matrix (Table 5 B) are also numerically large, suggesting a significant contribution from some regions to demand in the index region. Our main finding is that the nature of spatial diffusion in regional housing markets is more complicated than what can be modeled using a single and simple notion of distance. In particular, we find several explanations, singly or in combination, for the pattern of observed spatial diffusion; these include geographic contiguity or distance, inter-regional economic interactions, and social interactions and segregation. Thus, the topographical representation implied by these estimated spatial weights (Figure 2) look very different from the map of England and Wales (Figure 1).

First, contiguity or distance explains a number of the significant positive spatial weights. These include: spatial weights between Greater London, South East and the East of England; the East of England and East Midlands; the North West with North East and Wales; and Wales and West Midlands. Tentative explanations can be offered. Greater London, the South East and East can be regarded as substitutes in the choice of housing location. Hence, idiosyncratic shocks to housing demand in one of these regions can also affect demand in the other two regions. It is, however, interesting to note that the spatial weights between several pairs of contiguous regions are not statistically significant, and some other significant weights relate to non-

37

contiguous regions. In other words, many significant spatial weights appear to be driven by reasons other than geographic distance or contiguity. The Ledoit-Wolf test (24) for the first-order contiguity spatial weights matrix also rejects the specified driver of spatial diffusion in demand at the 1 per cent level of significance. This evidence also supports the view that there is more to the nature of spatial dependence than what can be explained by simple geographic distances.

Second, some of the strong positive spatial interactions operate along major inter-region transport links. Significant positive spatial weights between Yorkshire and Humberside and the South East, and between the North East and both West Midlands and the East of England, appear to be related to fast transport links. This is in line with evidence on the effect of commuting time on housing demand (Gibbons and Machin, 2005).

Third, the housing markets in some regions, like East Midlands and the South West appear to be related to the core-periphery relationship. These two regions lie on the periphery of the two most prominent urban housing markets – (Greater) London and Birmingham (West Midlands). Thus, an external shock may affect the periphery differently, and in some senses uniformly, compared to the core regions. Also, these two regions are socially and culturally quite closely related, having very similar per capita income levels and deprivation, ethnicity and political views (voting patterns).

Finally, Table 5 A indicates a significant negative spatial interaction between the South East and the North West, which suggests two possibilities. First, the housing markets in these two regions could be segmented along social or ethnic dimensions, implying that while one region may be attractive for certain social groups, these groups may be less attracted to housing market in the other region. Second, they may be related to the "ripple effects" phenomenon, whereby sharp changes in housing markets in the South East and London slowly spread over time to other parts of the country. This could imply that different regions may be on different (and possibly asynchronous) housing cycles.

In summary, the application identifies significant and interpretable spatial relationships in demand between government office regions in England and Wales, based on an estimate of the symmetric spatial weights matrix. We find that while contiguity and geographic distance explain the strength of inter-region interactions to some extent, other factors such as socio-cultural distances and transport infrastructure are also important. More complete analysis of these effects would require several candidate spatial weights matrices reflecting the drivers suggested by the current work, and then to formally examine the explanatory power of corresponding spatial weights.

6 Conclusion

The paper shows that estimation of an unknown spatial weights matrix consistent with an observed pattern of spatial autocovariances is a partially identified problem. Further, we propose a methodology for estimation of the spatial weights matrix, based on a given estimator for the spatial autocovariance matrix and suitable structural constraints. The methodology has the important advantage that it does not assume any specific distance measure or spatial weights, neither do we make any *a priori* assumption about the nature of spatial diffusion. This flexible approach to studying spatial diffusion represents a departure from the literature. We discuss various features of the estimator including its large sample properties, outline a bootstrap procedure for computing standard errors, and propose a test for a specified driver of spatial diffusion. Monte Carlo simulations demonstrate the superior small-sample performance of the proposed estimators.

The proposed methodology is used to study spatial diffusion in housing demand between government office regions in England and Wales. The spatial weights matrix is estimated under an assumption of symmetric spatial weights, and tentative explanations for significant spatial weights are advanced.

The nature and strength of diffusion in the different regions appears to be driven by a combination of factors including contiguity or distance, peripherality, as well as social, ethnic, and national composition representing cultural distance, none of which are *a priori* obvious. Besides, some spatial weights are negative, implying substitution effects or "ripple effects". Thus, the proposed approach provides substantial information about the nature of spatial diffusion and the functioning of regional housing markets, and is potentially useful for the evaluation of region-specific housing policies.

The work suggests several extensions and paths of future research. In the context of our application to housing demand, further analyses using spatial weights matrices which incorporate the drivers of spatial diffusion identified in this paper may be useful, and likewise analyses at a lower level of spatial disaggregation. At the methodological level, potential future work include: efficient M-estimation under structural constraints expressed in terms of moments, and extending our methodology to the context of spatio-temporal dependence. An alternative estimation strategy based on moment restrictions has been proposed in Bhattacharjee and Holly (2008), and Bhattacharjee and Holly (2010) extend the methodology to the case where spatial dependence may be driven by both structural reasons and unobserved factors. Bhattacharjee and Holly (2011) also demonstrate the use of all the three methods in an application to decision making in a monetary policy committee. Exten-

sion of the methodology to other models of spatial interaction will enhance the usefulness of the methodology. Similar analyses can help understand spatial interaction in other applications, such as the study of convergence across regions, and more generally in applied microeconomic studies of cross-sectional dependence.

Appendix 1: Sources of spatio-temporal data

Other than Hometrack, sources for regional spatio-temporal data were:

- **Supply**: Housing stock (Source: Office of the Deputy Prime Minister (ODPM) and the Office of National Statistics (ONS);

- **Demand**: Proportion of Local Authority and RSL dwellings having low demand (Source: ODPM); Property transactions (Source: HM Land Registry and Inland Revenue); Supply minus vacant housing (Source: ODPM); Average number of views per week (Source: Hometrack);

- **Neighbourhood characteristics**: Percentage of unfit houses (Source: ODPM); Crime rates (Source: ODPM); Crime detection rates (Source: Home Office); Percentage of university acceptances to applications (Source: Universities and Colleges Admissions Service (UCAS)); Percentage of population of 16-24 year olds attending university (Source: UCAS); Best value performance indicators (Source: ODPM); and

- **Market conditions**: Average weekly household income (Source: ONS); unemployment rate (Source: Labour Force Survey (LFS)); Proportion of population claiming income support (Source: ONS).

References

[1] Aldstadt, J. and Getis, A. (2006). Using AMOEBA to create a spatial weights matrix and identify spatial clusters. *Geographical Analysis* **38**, 327–343.

[2] Anglin, P.M, Rutherford, R. and Springer, T.M. (2003). The trade-off between selling price of residential properties and time-on-the-market: The impact of price setting. *Journal of Real Estate Finance and Economics* **26**, 95–111.

[3] Anselin, L. (1988). *Spatial Econometrics: Methods and Models*. Kluwer Academic Publishers: Dordrecht, The Netherlands.

[4] Anselin, L. (1999). Spatial econometrics. In Baltagi, B.H. (Ed.), *A Companion to Theoretical Econometrics*. Oxford: Basil Blackwell, 310–330.

[5] Anselin, L. (2002). Under the hood: Issues in the specification and interpretation of spatial regression models. *Agricultural Economics* **27**, 247–267.

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

[7] Barker, K. (2003). *Review of Housing Supply.* Interim Report, HM Treasury: London.

[8] Barro, R.J. and Sala-i-Martin, X. (1992). Convergence. *Journal of Political Economy* **100**, 223–251.

[9] Bell, K.P. and Bockstael, N.E. Applying the generalized-moments estimation approach to spatial problems involving microlevel data. *The Review of Economics and Statistics* **82**, 72–82.

[10] Beran, R. and Srivastava, M.S. (1985). Bootstrap tests and confidence regions for functions of a covariance matrix. *The Annals of Statistics*, **13**, 95–115.

[11] Bhattacharjee, A. and Holly, S. (2008). Taking personalities out of monetary policy decision making? Interactions, heterogeneity and committee decisions in the Bank of England's MPC. *Mimeo.* Revised version of Working Paper **0612**, Centre for Dynamic Macroeconomic Analysis, University of St Andrews, UK.

[12] Bhattacharjee, A. and Holly, S. (2010). Understanding interactions in social networks and committees. Working Paper **1004**, Centre for Dynamic Macroeconomic Analysis, University of St Andrews, UK.

[13] Bhattacharjee, A. and Holly, S. (2011). Structural interactions in spatial panels. *Empirical Economics* **40**, 69–94.

[14] Bhattacharjee, A. and Jensen-Butler, C. (2005). A model of regional housing markets in England and Wales. *CRIEFF Discussion Paper* No. **0508**, Centre for Research into Industry, Enterprise, Finance and the Firm, University of St. Andrews.

[15] Born, B. and Breitung, J. (2009). Simple regression based tests for spatial dependence. Bonn Economics Discussion Papers **23_2009**, University of Bonn.

[16] Bound, J., Jaeger, D. and Baker, R. (1995). Problems with instrumental variables estimation when the correlation between instruments and the endogenous explanatory variable is weak. *Journal of American Statistical Association* **90**, 443–450.

[17] Cayley, A. (1846). Sur quelques propriétés des déterminants gauches. *Journal für die Reine und Angewandte Mathematik (Crelle's Journal)* **32**, 119–123. Reprinted in the *Collected Mathematical Papers of A. Cayley*, Cambridge University Press 1889-1898, Vol. 1, 332–336.

[18] Cheshire, P. and Sheppard, S. (2004). Capitalising the value of free schools: The impact of supply constraints and uncertainty. *Economic Journal* **114**, F397–F424.

[19] Conley, T.G. (1999). GMM estimation with cross sectional dependence. *Journal of Econometrics* **92**, 1–45.

[20] Corrado, L. and Fingleton, B. (2011). Where is the economics in spatial econometrics?. Journal of Regional Science 20(10), 1-30.

[21] Fair, R.C. (2003). Bootstrapping macroeconometric models. *Studies in Nonlinear Dynamics and Econometrics* **7**, 4, Article 1.

[22] Fik, T.J., Ling, D.C. and Mulligan, G.F. (2003). Modeling spatial variation in housing prices: A variable interaction approach. *Real Estate Economics* **31**, 623–646.

[23] Fingleton, B. (2003). Externalities, economic geography and spatial econometrics: conceptual and modeling developments. *International Regional Science Review* **26**, 197–207.

[24] Folmer, H. and Oud, J. (2008). How to get rid of W: a latent variables approach to modelling spatially lagged variables. *Environment and Planning A* **40**, 2526–2538.

[25] Freedman, D. (1984). On bootstrapping two-stage least squares estimates in stationary linear models. *The Annals of Statistics* **12**, 827–842.

[26] Getis, A. and Aldstadt, J. (2004). Constructing the spatial weights matrix using a local statistic. *Geographical Analysis* **36**, 90–104.

[27] Giacomini, R. and Granger, C.W.J. (2004). Aggregation of space-time processes. *Journal of Econometrics* **118**, 7 – 26.

[28] Gibbons, S. (2004). The costs of urban property crime. *Economic Journal* **114**, F441–F463.

[29] Gibbons, S and Machin, S. (2005). Valuing rail access using transport innovations. *Journal of Urban Economics* **57**, 148–169.

[30] Haurin, D.R. and Brasington, D. (1996). School quality and real house prices: intra- and interjurisdictional effects. *Journal of Housing Economics* **5**, 351–368.

[31] Hausman, J.A., Newey, W.K. and Taylor, W.E. (1987). Efficient estimation and identification of simultaneous equation models with covariance restrictions. *Econometrica* **55**, 849–874.

[32] Hausman, J.A. and Taylor, W.E. (1983). Identification in simultaneous equation systems with covariance restrictions. *Econometrica* **51**, 1527–1549.

[33] Hendershott, P. (1996). Rental adjustment and valuation in overbuilt markets: Evidence from the Sydney office market. *Journal of Urban Economics* **39**, 51–67.

[34] Herander, M.G. and Saavedra, L.A. (2005). Exports and the structure of immigrant-based networks: the role of geographic proximity. *Review of Economics and Statistics* **87**, 323–335.

[35] Holly, S., Pesaran, M.H. and Yamagata, T. (2011). The spatial and temporal diffusion of house prices in the UK. *Journal of Urban Economics* **69**(1), 2–23.

[36] Jennrich, R.I. (2001). A simple general procedure for orthogonal rotation. *Psychometrika* **66**, 289–306.

[37] Jennrich, R.I. (2004). Derivative free gradient projection algorithms for rotation. *Psychometrika* **69**, 475–480.

[38] Kelejian, H.H. and Prucha, I.R. (1999). A generalized moments estimator for the autoregressive parameter in a spatial model. *International Economic Review* **40**, 509–533.

[39] Kelejian, H.H. and Prucha, I.R. (2004). Estimation of systems of spatially interrelated cross sectional equations. *Journal_of_Econometrics* **118**, 27–50.

[40] Ledoit, O. and Wolf, M. (2002). Some hypothesis tests for the covariance matrix when the dimension is large compared to the sample size. *The Annals of Statistics* **30**, 1081–1102.

[41] Lee, L.-F. (2004). Asymptotic distribution of quasi-maximum likelihood estimators for spatial autoregressive models. *Econometrica* **72**, 1899–1925.

[42] Mardia, K.V. and Marshall, R.J. (1984). Maximum likelihood estimation of models for residual covariance in spatial regression. *Biometrika* **71**, 135–146.

[43] McDonald, R.P. (1978). A note on the factor analysis of partial covariance matrices. *Psychometrika* **43**(1), 121.

[44] Meen, G. (1996). Spatial aggregation, spatial dependence and predictability in the UK housing market. *Housing Studies* **11**, 345–372.

[45] Meen, G. (1999). Regional house prices and the ripple effect: A new interpretation. *Housing Studies* **14**, 733–753.

[46] Meen, G. (2001). *Modelling Spatial Housing Markets: Theory, Analysis and Policy*. Kluwer Academic Publishers: Boston.

[47] Meen, G. (2003). Regional housing supply elasticities in England. Report prepared for the Barker Review of Housing Supply. Centre for Spatial and Real Estate Economics, University of Reading, October 2003.

[48] Meen, D. and Meen, G. (2003). Social Behaviour as a Basis for Modelling the Urban Housing Market: A Review. *Urban Studies* **40**, 917–935.

[49] Muirhead, R.J. (1982). *Aspects of Multivariate Statistical Theory*. Wiley: New York.

[50] Ord, J.K. (1975). Estimation methods for models of spatial interaction. *Journal of the American Statistical Association* **70**, 120-126.

[51] Pesaran, M.H. and Tosetti, E. (2011). Large panels with common factors and spatial correlation. *Journal of Econometrics* **161**, 182–202.

[52] Pinkse, J., Slade, M.E. and Brett, C. (2002). Spatial price competition: a semiparametric approach. *Econometrica* **70**, 1111–1155.

[53] Reiersöl, O. (1963). Rational orthogonal matrices. *American Mathematical Monthly* **70**, 63–65.

[54] Rey, S.J. and Montouri, B.D. (1999). US regional income convergence: a spatial econometric perspective. *Regional Studies* **33**, 143–156.

[55] Rilstone, P. and Veall, M. (1996). Using bootstrapped confidence intervals for improved inferences with seemingly unrelated regression equations. *Econometric Theory* **12**, 569–580.

[56] Serfling, R.J. (1980). *Approximation Theorems of Mathematical Statistics.* Wiley: New York.

[57] Taylor, A.E. and Mann, W.R. (1983). *Advanced Calculus*, 3rd Edition. Wiley: New York.

[58] van der Vaart, A.W. and Wellner, J.A. (1996). *Weak Convergence and Empirical Processes.* Springer: New York.