Spatial Economic Analysis
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/rsea20

Testing for Unit Roots and Cointegration in Spatial Cross-Section Data

Michael Beenstock, Dan Feldman & Daniel Felsenstein

Available online: 20 Apr 2012

Spatial Economic Analysis, Vol. 7, No. 2, June 2012

Testing for Unit Roots and Cointegration in Spatial Cross-Section Data

MICHAEL BEENSTOCK, DAN FELDMAN & DANIEL FELSENSTEIN

(Received March 2011; accepted September 2011)

ABSTRACT *Spatial impulses are derived for SAR models containing a spatial unit root. Analytical solutions are obtained for lateral space where the number of spatial units tends to infinity. Numerical solutions are obtained for finite regular lattices where edge-effects are shown to influence spatial impulses, and for irregular lattices. Monte Carlo simulation methods are used to compute critical values for spatial unit root tests in SAR models estimated from spatial cross-section data for regular and irregular lattices. We also compute critical SAC values for spatial cointegration tests for cross-section data that happen to be spatially nonstationary. We show that parameter estimates in spatially cointegrated models are 'superconsistent'.*

Essais de racines unité et co-intégration dans des données transversales spatiales

RÉSUMÉ *On dérive des impulsions spatiales de modèles SAR contenant une racine unité spatiale. On obtient des solutions analytiques pour l'espace latéral lorsque le nombre d'unités spatiales tend vers l'infini. On obtient des solutions numériques pour des réseaux réguliers finis, où l'on relève l'influence d'« edge effects » sur les impulsions spatiales, et pour des réseaux irréguliers. Des méthodes de simulation Monte Carlo sont utilisées pour calculer des valeurs critiques pour des tests de racine unité spatiale dans des modèles SAR estimés sur la base de données transversales spatiales pour réseaux réguliers et irréguliers. Nous calculons également des valeurs critiques de SAC pour essais de co-intégration spatiale, concernant des données transversales qui s'avèrent être spatialement non stationnaires. Nous démontrons que les estimations de paramètres dans des modèles spatialement co-intégrés sont « ultra cohérentes ».*

Pruebas de raíces unitarias y cointegración en datos espaciales de corte transversal

EXTRACTO *Se derivan impulsos espaciales para modelos SAR que contienen una raíz unitaria espacial. Se obtienen soluciones analíticas para espacio lateral donde el número de unidades espaciales*

Michael Beenstock, Department of Economics, Hebrew University, Mount Scopus, Jerusalem 91905, Israel. Email: msbin@mscc.huji.ac.il. Dan Feldman, Department of Economics, Hebrew University, Jerusalem, Israel. Email: danmsf@gmail.com. Daniel Felsenstein (to whom correspondence should be sent), Department of Geography, Hebrew University, Mount Scopus, Jerusalem 91905, Israel. Email: msdfels@mscc.huji.ac.il. The authors would like to thank two referees for their comments.

ISSN 1742-1772 print; 1742-1780 online/12/020203-20 http://dx.doi.org/10.1080/17421772.2012.669491

tiende al infinito. Se obtienen soluciones numéricas para retículos finitos regulares que demuestran que los efectos de borde influyen sobre los impulsos espaciales, así como para retículos irregulares. Se utilizan métodos de simulación de Monte Carlo para computar valores críticos destinados a las pruebas espaciales de raíces unitarias en modelos SAR, estimados a partir de datos espaciales de corte transversal para retículos regulares e irregulares. También computamos valores SAC críticos destinados a pruebas de cointegración espacial para datos de corte transversal que no son espacialmente estacionarios. Mostramos que las estimaciones de parámetros en modelos espacialmente cointegrados son 'superconsistentes'.

空间截面数据的单位根和协整

摘要：空间脉冲源自包含空间单位根的SAR模型。空间单位数量趋于无穷大的横向空间获取分析解决方案。有限规则点阵获得数值解，证明边缘效应影响空间脉冲，以及不规则点阵。蒙特卡罗模拟方法用于计算SAR模型中的空间单位根检验的临界值，估计从规则和不规则点阵获取的空间截面数据。我们还计算刚好是空间非平稳的截面数据的空间协整分析测试的临界值。我们阐明空间协整模型中的参数估计量是"超相容"的。

KEYWORDS: *Spatial unit roots; spatial cointegration; spatial impulse responses*

JEL CLASSIFICATION: C23; C22

1. Introduction

Fingleton (1999) demonstrated that if the data generating processes (DGP) for spatial cross-section data happen to contain spatial unit roots, the estimated regression coefficients may be spurious.[1] He suggested that the concept of cointegration proposed by Engle & Granger (1987) to test for nonsense and spurious regression in time series data may be extended to spatial data. Therefore, if the DGPs of Y and X happen to embody spatial unit roots, the regression coefficient of Y on X will be genuine rather than nonsense provided Y and X are spatially cointegrated. If, on the other hand Y and X are not spatially cointegrated, the regression coefficient is nonsense or spurious. This happens when the residuals contain a spatial unit root.

Fingleton did not provide a spatial unit root test to determine whether the DGPs of spatial data contain spatial unit roots and are therefore spatially non-stationary. Nor did he provide a spatial cointegration test to determine whether parameter estimates are nonsense or not. Our main purpose is therefore twofold. First, we develop a test statistic for unit roots in spatial cross-section data. We derive the distribution of the SAR coefficient under the null hypothesis of a spatial unit root. We obtain critical values for the spatial counterpart to the well known Dickey–Fuller statistic. Second, we develop a test statistic for spatial cointegration in which the data happen to contain spatial unit roots. In doing so we apply to spatial cross-section data concepts used by Engle & Granger (1987) for nonstationary time series data. Specifically, spatial cointegration requires that the model residuals be stationary, and that their spatial autocorrelation coefficient (SAC) be less than one. Spurious and nonsense regression phenomena arise when this null hypothesis that SAC = 1 cannot be rejected. If, however, the null hypothesis is rejected the data are spatially cointegrated.

Lauridsen & Kosfeld have suggested two types of cointegration tests for nonstationary spatial data. Our approach is similar to Lauridsen & Kosfeld (2004) who test the null hypothesis that the model residuals contain a spatial unit root. They calculate the distribution of the Wald test under the null hypothesis that the residuals contain a spatial unit root.[2] In contrast, we calculate the spatial counterpart to the Dickey–Fuller statistic. Lauridsen & Kosfeld (2006, 2007) also suggested a two stage Lagrange multiplier (LM) test for spatial unit roots. In the first stage, the LM SAC statistic is calculated for the residuals. If the LM statistic is not significant the residuals must be stationary. If the LM statistic is significant,[3] the second stage is intended to determine whether SAC = 1 by estimating the model with spatially differenced data. If SAC = 1 the LM statistic for the residuals in the second stage should not be significant. The difficulty with this proposal is that if the model is cointegrated, the second stage is misspecified; it should be a spatial error correction model.[4]

At first, we assume that space is a square rook lattice and that the spatial connectivity matrix (W) is sparse with $w_{ij} = 1$ for contiguous spatial units and zero otherwise. We calculate critical values for this case. Subsequently, we vary the tesselation for oblong spaces, rook–queen lattices, and for specific locations, such as Columbus, Ohio and NUTS2. We do so because we show that spatial impulse responses depend on topology; they are stronger in square lattices than in oblong ones, and they are stronger in queen lattices than in rook ones. It turns out, however, that the critical values are quite close to unity. Typically, an estimated root of 0.95 is significantly less than 1.

We also compute critical values for spatial cointegration. Suppose that our spatial unit root tests indicate that Y and X have spatial unit roots. For Y and X to be spatially cointegrated the residuals obtained from a regression of Y on X (or vice-versa) must be spatially stationary. Under the null hypothesis, these residuals are assumed to contain a spatial unit root. If the estimated SAC coefficient of the residuals is less than its critical value, we may reject the hypothesis that Y and X are not spatially cointegrated. We compute these critical values for different sample sizes and numbers of covariates.

We show that parameter estimates of spatially cointegrated models are super-consistent.[5] This means that the rate of asymptotic convergence is considerably more rapid if spatial data are nonstationary than for stationary data. In the latter case the rate of convergence is $1/N^{1/2}$. In the former case the rate of convergence is $1/N^{\delta}$ where $\delta > 1$.

A secondary contribution, related to the first, concerns the way in which spatial impulses propagate with and without spatial unit roots. What happens in a given spatial unit depends on what is happening in its neighbours. Since spatial units are mutual neighbours, shocks originating in a particular spatial unit rebound via its neighbours.[6] Spatial units on the edge of the lattice have fewer neighbours. We show that because space has edges, shocks originating at the epicentre propagate more strongly than shocks originating at the periphery. Also, even if there is a spatial unit root, edge effects create the misleading impression that spatial impulse responses die away with distance. Such edge effects must be taken into consideration in the design of spatial unit root tests.

In contrast to Fingleton (1999) and Mur & Trivez (2003), we do not assume that there is an unconnected spatial unit. Spatial units in the corners of the lattice have fewer neighbours, but this does not mean that they are unconnected. It only means that they are less connected.[7] Therefore our proposed spatial unit root tests take topology and edge effects into consideration without the artificial contrivance of unconnected spatial units. We note, in this context, that the standard practice of normalizing spatial weights to sum to unity ignores the fact that spatial units along the edges and in the corners of the lattice are less connected.

2. Data Generating Processes

In what follows there are N spatial units labelled by j and spatial lags are labelled with a tilde. The hypothesis of interest is that Y depends on X:

$$Y_j = \gamma + \beta X_j + u_j \tag{1a}$$

$$u_j = \rho \tilde{u}_j + v_j \tag{1b}$$

where, e.g. $\tilde{u}_j = \sum_{k \neq j}^{N} w_{jk} u_k$ denotes the spatially weighted[8] average value of u among the neighbours of j, ρ is the SAC coefficient of u. The v's are iid random variables.

The data generating processes (DGP) for Y and X are assumed to be first-order SAR models:

$$Y_j = \alpha_Y \tilde{Y}_j + \varepsilon_j \tag{2a}$$

$$X_j = \alpha_X \tilde{X}_j + e_j \tag{2b}$$

where ε and e are iid random variables with correlation r. A variable is defined to be strongly stationary if all its moments are finite and are independent of the sample size. It is weakly (covariance) stationary if the first two moments are finite and are independent of the sample size. In the case of spatial data means, variances and covariances should be independent of N. Spatial covariances should depend

> only upon the relative position of different locations, as determined by their relative orientation (angle) and respective distances. Since the orientation between two points in two (or more) dimensions still leaves a great number of different situations (potentially over a 360 degree rotation), the stricter notion of isotropy is imposed as well. (Anselin, 1988, p. 43)

Since regression does not use higher-order moments we focus on covariance stationarity.

2.1. The means of Y *and* X

A variable Z is spatially integrated of order d when its d'th spatial difference is spatially stationary. Therefore $Y \sim SI(1)$ and $X \sim SI(1)$ when $\alpha_Y = \alpha_X = 1$. If $u \sim SI(1)$ because $\rho = 1$, Equation (1) is a nonsense regression, which happens if ε and e are independent, i.e. $r = 0$. If $u \sim SI(0)$ because $\rho < 1$, Equation (1) is not nonsense as pointed out by Fingleton (1999). This happen because $r \neq 0$, i.e. because ε and e are jointly distributed random variables.

This suggests that the cointegration test statistic for spatial cross-section data should focus on rejecting the null hypothesis that $\rho = 1$. The same applies to spatial unit root tests for α_Y and α_X. In Section 4 we use Monte Carlo simulation to derive the distributions of α_Y and α_X under the null hypothesis that these parameters are equal to unity. These distributions provide critical values for rejecting the null hypothesis that the data contain a spatial unit root. In Section 5 we use Monte Carlo simulation to compute critical values for spatial cointegration tests when the data contain spatial unit roots and when $\rho = 1$ under the null hypothesis. If the calculated value of ρ is less than its critical value, we may reject the hypotheses that the result is spatially spurious.

Let y denote a column vector of length N with elements Y_j. W_N is an irregular but isotropic $N \times N$ spatial connectivity matrix in which connectivity may vary between spatial units. Equation (2a) may be vectorized as:

$$y = \alpha W_N y + \varepsilon \tag{3}$$

If $I_N - \rho W_N$ is invertible, the spatial Wold representation of Equation (3) is:

$$y = A_N \varepsilon \tag{4a}$$

$$A_N = (I_N - \rho W_N)^{-1} = I_N + \alpha W_N + \alpha^2 W_N^2 + \ldots \tag{4b}$$

Invertibility requires that the elements of A_N be finite which is satisfied when αW_N is convergent. The spatial variance–covariance matrix generated by Equation (4a) is expected to be:

$$\Sigma_N = E(yy') = \sigma_\varepsilon^2 B_N \tag{5a}$$

$$B_N = A_N A'_N \tag{5b}$$

According to Equation (4a) the first moment is finite and is independent of N since $E(y) = 0$. Matters are quite different in the case of second moments because according to Equation (5a), Σ may depend on N. Since $I_N - \alpha W_N$ is invertible det(A) must be finite, which guarantees that det(Σ) is finite. Therefore, the second moments are finite. However, the second moments will depend upon N if B_N varies with N. According to Equation (4b) this depends on W_N and α.

Normalizing $\sigma_\varepsilon = 1$, the variance of Y_k with N units is b_{kk} and its covariance with Y_j is b_{kj}. Suppose that the sample increases by 1 from N to M and location M is remote from k and j. Stationarity requires that b_{kk} and b_{kj} remain unchanged. These conditions require that shocks in remote spatial units have no repercussions on units k and j, e.g. $\alpha^{M-k} W^{M-k}$ tends to zero. Therefore stationarity requires that αW be convergent. Y is isotropic if adding a remote spatial unit makes no difference to the connectivity between incumbent units. Therefore, when Y is isotropic convergence does not depend on W, it depends on $\alpha < 1$. We do not consider the more general case in which the data are not isotropic. However, we investigate critical values for spatial unit roots for different topologies, including irregular topologies in which the number of neighbours is not the same for all spatial units.

3. Spatial Impulse Responses

In this section we show that spatial unit roots asymptotically induce spatial impulse responses that do not die out with distance. If spatial data are stationary, shocks occurring remotely from j have no effect on j. If, however, spatial data are nonstationary, such remote shocks affect j as if they occurred in j itself. This result is established analytically for lateral space, i.e. where space is an infinite line that has no beginning and end, and spatial units have neighbours on two sides only.

In multilateral space there are more than two neighbours. For example, in bilateral space (rook lattice) each spatial unit has four sides and neighbours. Unfortunately, we are unable to obtain analytical solutions for this more relevant case. We therefore simulate spatial impulse responses numerically. Because of computational constraints we are forced to assume that space is finite, which means that we cannot obtain asymptotic impulse responses as we do in the lateral case. We think that this may be advantageous because space is inherently finite. We show that in finite space impulse responses die out with distance even in the presence of a spatial unit root. This happens because peripheral spatial units are less connected than core spatial units. Indeed, this core–periphery effect may create the misleading impression that the data are spatially stationary when the opposite is true.

3.1. Lateral Space

Spatial units are assumed to be located laterally (along an axis representing west and east or north and south) so that each unit has a neighbour on either side. Units continue to be labelled by $j = -\infty, \ldots, \infty$. There is an infinite number of spatial units and unit $j+1$ is unit j's neighbour to the left or west and unit $j-1$ is its neighbour to the right or east. Spillovers are assumed to be first-order, i.e. they occur between immediate neighbours. The SAR model in this case is:

$$Y_j = \alpha(Y_{j+1} + Y_{j-1}) + u_j \tag{6}$$

where α denotes the spatial spillover coefficient and u is an iid random variable with variance equal to σ_u^2. Equation (6) is a second-order stochastic spatial difference equation.

Let S denote a spatial lag operator such that $S^i Y_j = Y_{j-i}$ where i may be positive (west of j) or negative (east of j). Multiplying Equation (6) by S and rewriting the result in terms of the spatial lag operator gives:

$$(1 - \alpha^{-1}S + S^2)Y_j = -\alpha^{-1}u_{j-1} \tag{7}$$

The auxiliary equation of Equation (7) is:

$$\lambda^2 - \alpha^{-1}\lambda + 1 = 0 \tag{8}$$

If $0 < \alpha < ½$, the roots of Equation (8), denoted by λ_1 and λ_2, are real and positive and are reciprocally related because $\lambda_1 = 1/\lambda_2$. The roots will be complex if $4\alpha^2 > 1$ which arises when $\alpha > ½$. When $\alpha = ½$ both roots are equal to unity. When $0 < \alpha < ½$ one root is positive and less than one (λ_1) while the other (λ_2) is positive and greater than one, since the roots come in reciprocal pairs. Because $(1 - \lambda_1 S)(1 - \lambda_2 S) = (1 - \alpha^{-1}S + S^2)$, the general solution[9] to Equation (6) is:

$$Y_j = -\frac{\alpha^{-1}u_{j-1}}{(1 - \lambda_1 S)(1 - \lambda_2 S)} + A_1\lambda_1^i + A_2\lambda_2^i \tag{9}$$

where A_1 and A_2 are arbitrary constants of integration and i denotes distance from j. To obtain a particular solution it is necessary to determine A_1 and A_2 using data on Y for two spatial units. Since in what follows we have no interest in the particular solution we ignore it by setting these arbitrary constants to zero. Using partial fractions[10] we note that:

$$\frac{1}{(1 - \lambda_1 S)(1 - \lambda_2 S)} = \frac{1}{\lambda_1 - \lambda_2}\left[\frac{\lambda_1}{1 - \lambda_1 S} - \frac{\lambda_2}{1 - \lambda_2 S}\right] \tag{10}$$

We also note that[11]:

$$\frac{1}{1-\lambda_1 S} = \sum_{i=0}^{\infty} \lambda_1^i S^i \tag{11}$$

If $\lambda_1 < 1$, Equation (11) is convergent because λ_1^i tends to zero with i. Since the roots come in reciprocal pairs, applying Equation (11) to λ_2 would induce a divergent process since λ_2^i tends to infinity with i because $\lambda_2 > 1$ if $\lambda_1 < 1$. This would imply unreasonably that despite the fact that $\alpha < ½$ Y is explosive and divergent. The solution to this problem is to note that $(1-\lambda_2 S)^{-1}$ has two polynomial inversions, one that is the counterpart of Equation (11) and another which is:

$$\frac{1}{1-\lambda_2 S} = -\frac{(\lambda_2 S)^{-1}}{1-(\lambda_2 S)^{-1}} = -\frac{\lambda_1 S^{-1}}{1-\lambda_1 S^{-1}} = -\sum_{i=1}^{\infty} \lambda_1^i S^{-i} \tag{12}$$

Equation (12) is the 'forward' inversion whereas Equation (11) is a 'backward' inversion. Substituting $\lambda_2 = \lambda_1^{-1}$ into the forward inversion generates Equation (12) which is convergent because it depends on λ_1^i. Equation (11) operates 'westwards' since $i \geq 0$ and Equation (12) operates 'eastwards' since $i < 0$.

Substituting Equations (10), (11) and (12) into Equation (9) gives:

$$Y_j = \frac{\alpha^{-1}}{\lambda_1^{-1} - \lambda_1}\left[\sum_{i=1}^{\infty} \lambda_1^i u_{j-i} + \sum_{i=0}^{\infty} \lambda_1^i u_{j+i}\right] \tag{13}$$

Equation (13) is the spatial Wold representation of Equation (6) since it expresses Y_j in terms of the stochastic shocks in all units to the east and west of unit j as well as in unit j itself. Equation (13) is also the spatial impulse response function. If $\lambda_1 < 1$, Equation (13) states that closer units to j have a greater effect on j than more remote units. The spatial impulse responses are symmetric, as expected, since u_{j+i} has the same effect on Y_j as u_{j-i}:

$$\frac{\partial Y_j}{\partial u_{j-i}} = \frac{\partial Y_j}{\partial u_{j+i}} = \frac{\alpha^{-1}\lambda_1^i}{\lambda_1^{-1} - \lambda_1} \tag{14}$$

Equation (14) also shows that the impulses tend to zero as the distance between spatial units (i) tends to infinity. This means that shocks that occurred infinitely far from unit j have no effect on Y_j. In addition, Equation (14) shows, as expected, that the largest impulse is for shocks that occur in region j itself ($i = 0$).

Finally, λ_1 varies directly with α. For example, when $\alpha = 0.1$, $\lambda_1 = 0.1015$ in which case the impulse response from immediate neighbours ($i = 1$) according to Equation (14) is equal to 0.103 and the local impulse response ($i = 0$) is 1.0207. Notice that the local impulse exceeds 1 because there is a spatial echo; shocks in unit j propagate back onto it via other spatial units. When $\alpha = 0.2$, λ_1 increases as expected to 0.2087 and the impulse response from immediate neighbours increases as expected to 0.225, and the own impulse increases to 1.091 because the spatial echo varies directly with α. When $\alpha = 0.498$ these impulse responses jump to 10.24 and 11.19, respectively. As α approaches ½, λ_1 approaches 1 and the impulse responses approach infinity. When $\alpha = ½$ both roots equal 1, the impulse responses explode and no longer depend on distance.[12]

According to Equation (13), $E(Y_j) = 0$ because the expected value of the u's are all zero[13] by definition. Therefore, the first moment is independent of j. The variance of Y from Equation (13) is equal to:

$$\mathrm{var}(Y) = \frac{\alpha^{-2}\left(1+\lambda_1^2\right)}{\left(1-\lambda_1^2\right)\left(\lambda_1^{-1}-\lambda_1\right)^2}\sigma_u^2 \tag{15}$$

which is finite since $0 \leq \lambda_1 < 1$ and it does not depend on j. Therefore if $0 < \alpha < ½$ the first and second moments of Y_j are finite and independent of j. Matters are different when $\alpha = ½$. Since $\lambda_1 = 1$, the denominator of Equation (15) is zero and the variance of Y is therefore infinite.

3.2. Bilateral Space

When space is lateral and the number of neighbours (n) is two, we saw that the SAR coefficient inducing a spatial unit root is $\alpha^* = ½ = 1/n$. When space is multilateral and spatial units have more than two neighbours the critical value for α that gives rise to a spatial unit root is $\alpha^* = 1/n$. If space is a rook lattice each spatial unit has four neighbours in which case $\alpha^* = ¼$, and if it is a queen lattice $\alpha^* = 1/8$.

The bilateral counterpart to Equation (6) may be written familiarly as the SAR model:

$$Y = \alpha WY + u \tag{16}$$

where W is a sparse $N \times N$ matrix with elements $w_{jk} = 1$ if j and k are neighbours and $w_{jk} = 0$ otherwise, and Y and u are vectors of length N. In Section 3.1 N is infinite and space has no edge. However, in the present section N is assumed to be finite, in which case space has an edge. In any case, spatial datasets typically have edges because they refer to specific geopolitical entities. Fingleton (1999) normalized the row-sum weights (w) to unity, and normalized $\alpha = 1$. This means that at the corners of the lattice where there are two neighbours $w_{jk} = ½$ instead of ¼, and at the edge of the lattice where there are three neighbours $w_{jk} = 1/3$; instead of ¼, which overstates the true weight. We therefore prefer to normalize $w_{jk} = 1$ in which case the sum of weights is four and $\alpha^* = ¼$ because this does not artificially inflate spatial spillover at the corners of the lattice and along its edges. Another difference is that unlike Fingleton we do not assume the existence of an 'unconnected spatial unit'; in our lattice all units are mutually connected.

The Wold representation of Equation (16) is:

$$\begin{aligned} Y &= Au \\ A &= (I - \alpha W)^{-1} = I_N + \sum_{i=1}^{\infty} \alpha^i W^i \end{aligned} \tag{17}$$

where $W^i u$ denotes u among i'th order neighbours. If α is normalized to 1 and the restriction of row-sum = 1 is applied to W, the matrix $I–\alpha W$ is not invertible regardless of N. If instead $\alpha = 1/n$ and the row-sum restriction is not applied, $I–\alpha W$ is invertible if N is finite. In fact, $\det(I–\alpha W)$ is $O(1/N)$ so that as space becomes infinitely large edge effects are asymptotically unimportant.

The spatial impulse responses are:

$$\frac{\mathrm{d}Y_j}{\mathrm{d}u_j} = a_{jj} \tag{18}$$

$$\frac{\mathrm{d}Y_j}{\mathrm{d}u_i} = a_{ji} \tag{19}$$

We expect a_{jj} to vary directly with the number of spatial units because this gives rise to more scope for spatial spillover, and we expect a_{ji} to vary inversely with the distance between j and i. If, however, there is a spatial unit root, the impulses a_{ji} will not tend to zero as the distance between j and i tends to infinity. We use Matlab to calculate A for square lattices containing N spatial units in which $n = 4$ and $\alpha^* = ¼$. To investigate asymptotics we ideally wish to let N tend to infinity, but this is not computationally feasible. We therefore make N as large as practically possible (approximately 1,000) given computing constraints.

Because N is finite A is inevitably affected by edge effects. Spatial units on the edge are less exposed to spatial spillover because they have only three neighbours instead of four. Spatial units in the four corners of the lattice only have two neighbours. We expect a_{jj} and a_{ji} to be greater the closer j is to the epicentre j^* of the lattice because there is more scope for spatial interaction in the centre than at the periphery. We do not expect a_{ji} to be symmetrical unless $j = j^*$ because only at the epicentre is the distance to the edge of the square lattice the same in all four directions.

In Figure 1 we plot impulse responses for a_{j^*i} when $N = 961$. Since the lattice is square it is 31×31, and the distance from its epicentre to the edge is therefore 15. The impulse responses are measured along the vertical and Euclidean distance

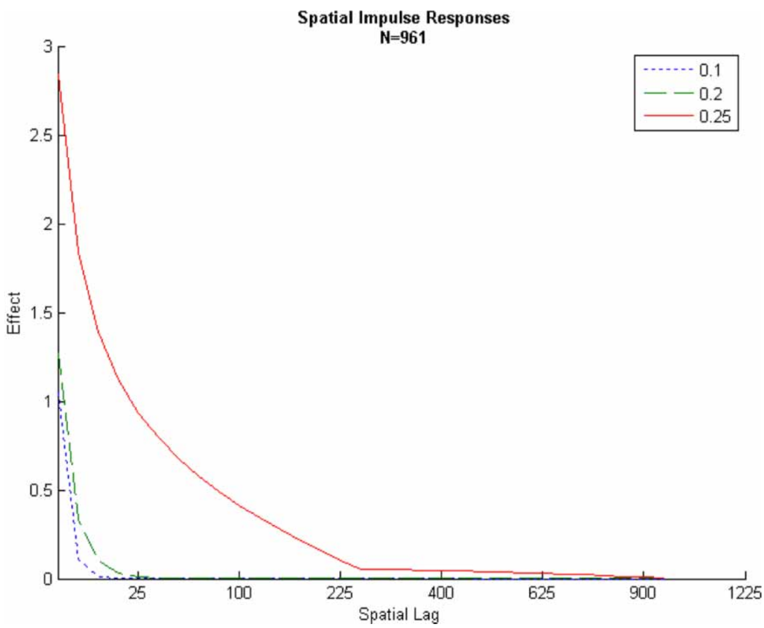

Figure 1. Spatial impulses.

from j^* is measured along the horizontal. The 15th value of i is on the edge of the lattice and the 30th value is at the corner of the lattice, hence the dog-leg at spatial lag 15. Figure 1 shows, as expected, that a_{ji} varies directly with α, and varies inversely with k. The impulse responses die away more slowly for larger α. It might have been expected that when $\alpha = \alpha^* = ¼$ the impulse responses should lie on a straight line and fail to die away. We expect this to happen as N tends to infinity, as in the case of lateral space. It does not happen in Figure 1 because N is finite and space has an edge. Nevertheless, Figure 1 clearly shows that when $\alpha = ¼$ there is a qualitative difference and the impulses linger longer in space than when $\alpha = 0.2$.

As indicated in Equation (18) the diagonal of A measures the direct and indirect effect of a shock in j on itself. We refer to these as 'local impulses', which are plotted in Figure 2 for different values of α according to the distance of j from the epicentre of the lattice. The local impulses exceed unity because when a shock occurs in j it affects j's neighbours which feedback onto j. This echo or boomerang effect naturally varies directly with α. Figure 2 also shows that when $\alpha < \alpha^*$ the edge effect does not affect local impulses because the echo does not extend very far. Hence the local impulse does not depend upon distance from the epicentre, except at the edge of the lattice.

When $\alpha = \alpha^*$, however, matters are quite different. Figure 2 shows that in this case the local impulses vary inversely with distance from the epicentre. This happens because the echo resounds far, so that close to the epicentre there is much more echo than at the edges. Had there been no edge, the effect at the epicentre would have been infinity because the echo resounds forever, and the effect elsewhere would have been infinite too. Indeed, when N is infinity there is no meaning to the epicentre because the lattice has no borders.

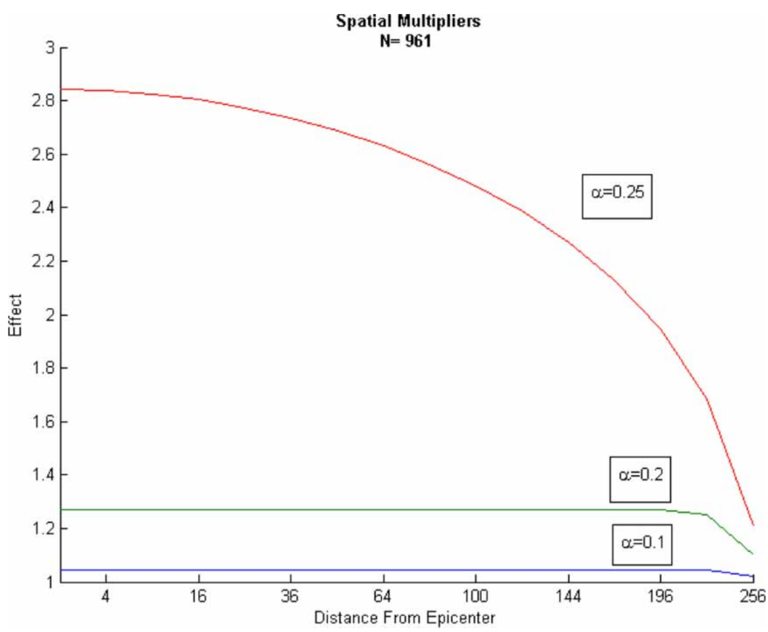

Figure 2. Local impulses.

3.3. Variances in Bilateral Space

When space is lateral there is an analytical expression for the variance of Y, as seen in Equation (15). When space is bilateral Equation (17) implies that the variance–covariance matrix of Y is equal to:

$$\Sigma = \sigma_u^2 AA' \tag{20}$$

We follow Fingleton (1999) and calculate[14] the average variance of Y as N increases.[15] The results are plotted in Figure 3, which shows that as α increases towards $\alpha^* = 1/4$, the variance of Y varies directly with N. Figure 3 clearly establishes that the variance depends upon N as α increases towards α^*. Indeed, the variance increases nonlinearly with N. This happens for two reasons. First, as N increases ε_N increases the variance through a natural scale effect. Secondly, as N increases there is more scope for spatial interaction. The latter increases the variance among the incumbent $N-1$ units.[16] However, when $\alpha < \alpha^*$, the variance does not depend upon N, as should be the case if the data are stationary. Stationarity requires that the variance should not depend on the sample size.

3.4. Irregular Lattices

In regular lattices the number of neighbours is the same for all spatial units except along the edges and in the corners. In irregular lattices the number of neighbours varies between spatial units even if they are located in the core of the lattice. It is conceptually unclear how to define spatial unit roots in irregular lattices. One intuitive conjecture is that if $w_{ji} = 1$ for contiguous units, α should be normalized to equal the reciprocal of the average number of neighbours. This conjecture is based on the principle that $\alpha^* = 1/n$ in regular lattices. For example, in NUTS2 the number of contiguous neighbours ranges between 1 and 11 with a mean of 4.714. Setting $\alpha^* = 0.212$ turns out to be incorrect because the mean of α in the Monte Carlo simulations is 0.186 and none of the 10,000 estimates exceeds 0.212. Therefore this conjecture is incorrect.

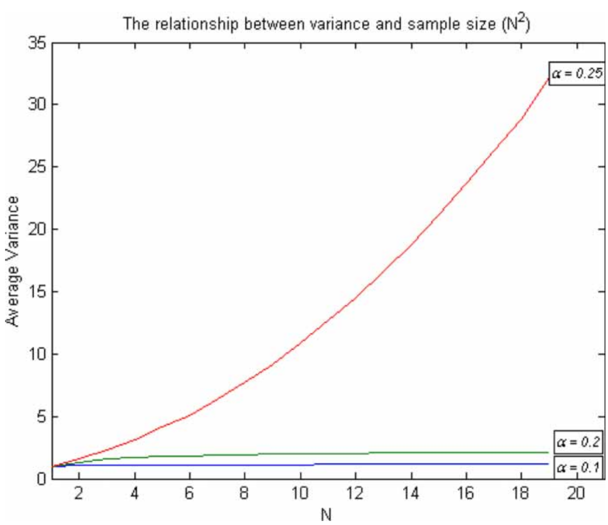

Figure 3. Relationship between variance and sample size.

An obvious solution to this problem is to set $\alpha^* = 1$ and to restrict W to be row-sum $= 1$. In this case the number of neighbours does not matter from a technical point of view. This solution assumes unreasonably that spatial connectedness is independent of the number of neighbours, so that the NUTS2 unit with one neighbour is just as spatially connected as the unit with 11 neighbours. This solution also ignores edge effects, as noted.

Nonstationarity implies that spatial impulses do not die out asymptotically with distance. We have shown that in regular lattices with $w_{ij} = 1$ for contiguous units, spatial impulses die out because of edge effects even when a spatial unit root is present. Figure 1 showed that impulses die out more slowly when there is a unit root. However, as N tends to infinity spatial impulses cease to die out. The same applies in irregular lattices. If W_0 is an irregular (sparse) spatial weight matrix for N_0 units and α^* denotes the spatial unit root, edge effects guarantee that $I_N - \alpha^* W_0$ is invertible and spatial impulses die out. If α^* is correctly defined these impulses should cease to die out as N tends to infinity.

Let W_1 be an irregular weights matrix when $N_1 > N_0$. Because N has increased, some or all of the edge units in W_0 will lose their edge status. Suppose that W_1 is irregular in the same way that W_0 was irregular (otherwise it is difficult to make asymptotic arguments). Impulses induced by $\alpha^* W_1$ should die away more slowly than those induced by $\alpha^* W_0$, and so on for W_2 etc. If $\alpha < \alpha^*$ these impulses will die out too rapidly. Therefore, α^* is selected to ensure that impulses do not die out at all as N tends to infinity.

Just as we saw in Figure 1 that α^* induced a qualitative change in the persistence of spatial impulses, and in Figure 2 it induced explosive tendencies in own impulses, so we suggest that α^* may be calculated for each irregular lattice. For example, in the case of NUTS2 where the number of neighbours ranges between 1 and 11 we find, by simulation, that $\alpha^* = 0.167$. For $\alpha < 0.167$ spatial impulses do not persist and own impulses are damped. However, if $\alpha = 0.167$ spatial impulses become persistent and own impulses cease to be damped. For Columbus, Ohio we find by simulation that $\alpha^* = 0.17$.

3.5. Spatial Superconsistency

If in Equation (1) Y and X are stationary and independent, the OLS estimate of β $\left(\hat{\beta}\right)$ is obviously root N-consistent. If they are dependent, consistency additionally requires that Y and X be ergodic so that the data are asymptotically independent. This means that the effect of stochastic shocks in spatial unit $z(\varepsilon_z)$ on Y_j tends to zero sufficiently strongly with the distance between j and z.

If Y and X are nonstationary SI(1) but they are cointegrated, the OLS estimate of β is 'superconsistent'. This means that the rate of asymptotic convergence of $\hat{\beta}$ to β is faster than root N. In fact it is N^{ω}-consistent where $\omega > 1$. If the data are stationary $\omega = ½$.

To show this let $\hat{\beta} = \beta + b$ where:

$$b = \frac{\frac{1}{N}\sum_{j=1}^{N} X_j u_j}{\frac{1}{N}\sum_{j=1}^{N} X_j^2} \qquad (21)$$

The denominator of Equation (21) is the variance of *X*, which according to Figure 3 is a power function of *N*, where the power (*s*) clearly exceeds 1. Therefore the denominator of *b* is $O_p(N^s)$; its asymptotic order in probability is *s*. The numerator of *b* is the covariance between *X* and *u*. Because Equation (1) is cointegrated, *u* is stationary by definition. The covariance between *X* which is nonstationary and *u* which is stationary is a random variable with asymptotic order (*q*) that must be less than *s*. Therefore the asymptotic order of *b* is $q-s=-\omega$. If $s=1½$ as suggested[17] by Figure 3 and $q=0$ then $\omega=1½$. This means that *b* tends to zero at a rate of $1/N^{\omega}$, in which case $\hat{\beta}$ is N^{ω}-consistent.[18]

Superconsistency implies that we learn more about β from a sample of *N* observations when spatial data are nonstationary and cointegrated than when they are stationary. Indeed, even if *X* is not exogenous and plim(*Xu*) is not zero, *b* still tends to zero with N^{ω}. Therefore superconsistency has the further advantage of eliminating simultaneous equations bias that would arise if the data were stationary, and $\text{plim}\left(\hat{\beta}\right)=\beta$. This, of course, is an asymptotic argument that might not be valid in finite samples. Finally, since the asymptotic order of the variance of *Y* is *s*, the asymptotic order of the variance of the residuals is 0 (because they are stationary). Since $R^2=1-\text{var}(u)/\text{vary}(Y)$, $\text{plim}(R^2)=1$ because the ratio between the variances tends to zero[19] at the rate of $1/s$.

4. Spatial Unit Root Tests

We set $\alpha=\alpha^*=¼$ in Equation (17) and generate 10,000 artificial datasets for *Y* by drawing the *u*'s using pseudo random numbers from a standard normal distribution for given *N*. We use these synthetic datasets to estimate by maximum likelihood[20] 10,000 SAR models. The distribution of the 10,000 estimates of the SAR coefficient is plotted in Figure 4 for the case when $N=400$ spatial units in a square lattice.

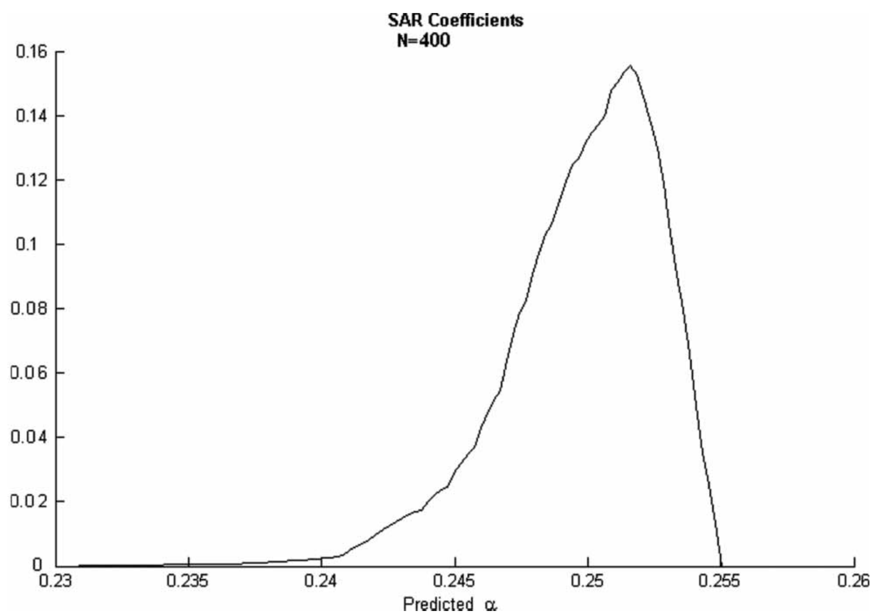

Figure 4. The distribution of SAR coefficient when $\alpha=¼$.

Not surprisingly the mean estimate of the SAR coefficient is almost 0.25 (0.2498) and the mode is around 0.25 (0.2520). However, some estimates exceed 0.25. The distribution is clearly skewed to the left.[21] According to Figure 4, when $N = 400$ there is a 95% chance of getting a SAR coefficient that is greater than $SAR^* = 0.243$. Therefore, the critical value for the SAR coefficient is 0.243 at $p = 0.05$.

In Table 1 we report SAR* for different values of N and p. If the estimated SAR coefficient is greater than SAR* the spatial cross-section data contain a spatial unit root. For example, when $N = 100$ and $p = 0.05$, SAR* is 0.225. If SAR is greater than SAR* we cannot reject the null hypothesis of a spatial unit root. Therefore, if the SAR estimate is, for example, 0.2 we may reject the hypothesis of a spatial unit root. SAR* naturally varies inversely with p and it varies directly with N, or the sample size.

The computations reported in Table 1 are inherently random because they depend on the seed used to generate the pseudo random numbers, which is chosen randomly. To obtain some impression of the degree of randomness we reseeded the case for $N = 100$, 500 times and reduced[22] the number of Monte Carlo trials from 10,000 to 5,000 (see Table 2). For example, when $p = 0.05$ one standard deviation of the critical value of SAR* is about ½% and SAR* is bounded between 0.224 and 0.226. Therefore, the critical values reported in Table 1 are reliable.

We also investigated the sensitivity of the computations reported in Table 1 to the number of Monte Carlo trials, which in Table 1 is 10,000. Here too $N = 100$. Table 3 shows, as expected, that the critical values are not sensitive to the number of trials when p is relatively large. Indeed, in this case even 1,000 trials would have been sufficient. However, matters are quite different when p is relatively small, e.g. $p = 0.01$. In this case the critical value of SAR* varies directly with the number of trials. Clearly, to dig into the tail of the distribution requires increasing the number of trials.

In Section 3 we saw that spatial impulses are affected by topology especially when there is a spatial unit root. This suggests that spatial unit root tests might vary by topology. In Table 4 we therefore report critical values for different topologies when the number of spatial units is 400. We also report critical values for irregular topologies in Columbus, Ohio and the European Union[23] (NUTS2). Case 1 in Table 4 is identical to the case in Table 1. We normalize the spatial unit root to unity for ease of comparison. For example, in case 1 each unit has four neighbours, therefore the normalized critical value is $4 \times 0.243 = 0.972$ at $p = 0.05$. Changing the lattice from a square to an oblong reduces the critical value slightly to 0.944 from 0.972. Since in an oblong the average distance between spatial units is greater than in the case of a square, there is correspondingly less spatial interaction. This makes it more difficult to estimate the SAR coefficient, and as a result, its critical value is less.

Table 1. Spatial unit root test statistics for square rook lattice: SAR*

	N		
P	25	100	400
0.01	0.071	0.209	0.24
0.05	0.139	0.225	0.243
0.1	0.161	0.231	0.244

Table 2. Confidence intervals for Table 1 ($N = 100$)

	Mean	Variance	Std.	Min	Max	Mode
Mean	0.2456	1.68E-08	1.30E-04	0.2452	0.246	0.2456
Mode	0.249	1.81E-08	1.35E-04	0.2489	0.252	0.2489
1%	0.2086	4.32E-06	0.0021	0.204	0.215	0.209
5%	0.2249	2.88E-07	5.37E-04	0.224	0.226	0.225
10%	0.2308	1.10E-06	1.00E-03	0.228	0.233	0.231

Table 3. Sensitivity of Table 1 to the number of trials ($N = 100$)

	Trials				
	500	1,000	5,000	10,000	15,000
Mean	0.2454	0.2454	0.2456	0.2452	0.2456
Mode	0.25	0.248	0.249	0.244	0.249
1%	0.193	0.208	0.2059	0.209	0.211
5%	0.22	0.225	0.224	0.225	0.225
10%	0.228	0.23	0.231	0.2309	0.2319
Truncated	32.2	30.3	32.04	32.74	31.2667

In case 3 each unit has eight neighbours instead of four, which increases the scope for spatial interaction. However, despite the fact that there is more spatial interaction in case 3 than in case 1 it is harder to reject the null hypothesis of a spatial unit root (the critical value in case 3 is 0.96 whereas in case 1 it is 0.972). The reason for this is that edge effects are stronger in case 3 than in case 1. In case 3, corner and edge units have three and five neighbours, respectively, instead of eight neighbours, whereas in case 1 they have two and three neighbours, respectively, instead of four.

Cases 4 and 5 differ to the previous cases in that they do not refer to artificial topologies in which the lattices are regular. Case 4 refers to Anselin's (1988) spatial connectivity matrix for Columbus, Ohio, and case 5 refers to NUTS2. Using the methodology described in Section 3.4 for determining unit roots when W is irregular, through simulation we find that the unit root for Columbus is 0.17 and for NUTS2 it is 0.167. If, for example, in the case of Columbus, $\alpha = 0.168$ the spatial impulses are damped. However, when $\alpha = 0.17$ these impulses become explosive. The critical values have been normalized to unity for purposes of comparison. Because the critical value is higher for NUTS2 than for Columbus, it is easier to reject the null hypothesis in the case of NUTS2 than in the case of Columbus. The critical values at $p = 0.01$ are 0.964 for NUTS2 and 0.853 for Columbus. This shows that the MC distribution for the SAR coefficients is tighter for NUTS2 than it is for Columbus.

Table 4. Critical values for spatial unit roots for different tesselations ($p = 0.05$)

Case	Critical value: SAR*
1. Square (20 × 20) rook lattice	0.972
2. Oblong (10 × 40) rook lattice	0.944
3. Square (20 × 20) queen lattice	0.960
4. Columbus, Ohio	0.953
5. NUTS2	0.988

Table 4 shows that critical values are not sensitive to topography. It is tempting to say that if estimated SAR coefficients are less than 0.9, one may be reasonably confident in rejecting the hypothesis of a spatial unit root regardless of topography.[24]

5. Spatial Cointegration Tests

Fingleton (1999) observed that if the DGPs for Y and X contain spatial unit roots, estimates of β in Equation (1) may be 'nonsense'. If $\beta = 0$ in Equation (1) and $Y \sim \mathrm{SI}(1)$ then it must be the case that $u \sim \mathrm{SI}(1)$ so that $\rho = 1$ in Equation (2). If, however, Y and X are spatially cointegrated, u must be stationary in which case $\rho < 1$.

We generate 10,000 artificial datasets for Y and X with $\alpha = ¼$. The random numbers used to generate Y and X are independent, hence we expect $\beta = 0$ in each draw. We use these datasets to generate 10,000 OLS estimates of β. The distribution of these estimates is plotted in Figure 5 for $N = 400$. The distribution in Figure 5 is[25] approximately normal, but the mean is a non-zero random variable. In Figure 5 the mean is 0.006 and the mode is -0.0088. There are positive as well as negative estimates of β. The residuals from these nonsense regressions are used to estimate[26] 10,000 estimates of SAC (ρ), which are plotted in Figure 6.

The mode in Figure 6 is 0.2527 and the mean is 0.2482. However, there are estimates that are below and above 0.25. When $p = 0.05$, Figure 6 implies that $\mathrm{SAC}^* = 0.241$. If $\mathrm{SAC} < \mathrm{SAC}^*$ the OLS estimate of β is not nonsense, in which event Y and X are spatially cointegrated. If, on the other hand, $\mathrm{SAC} > \mathrm{SAC}^*$ we cannot reject the hypothesis that the residuals contain a unit root, in which event the estimate of β is 'nonsense' and Y and X are not spatially cointegrated.[27]

Table 5 records the critical value of SAC^* in the bivariate case ($k = 2$). Note that SAC^* in Table 5 is typically smaller than SAR^* in Table 1. This discrepancy reflects the loss in degrees of freedom since SAC^* is based on estimated residuals rather than true residuals. As expected, SAC^* varies directly with N and p.

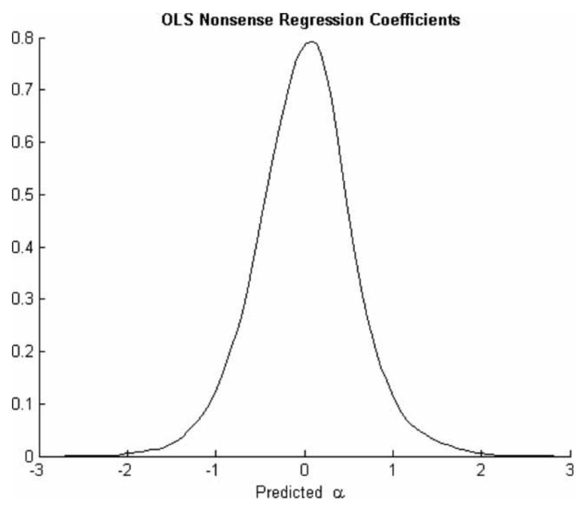

Figure 5. The distribution of the spatial nonsense regression coefficient.

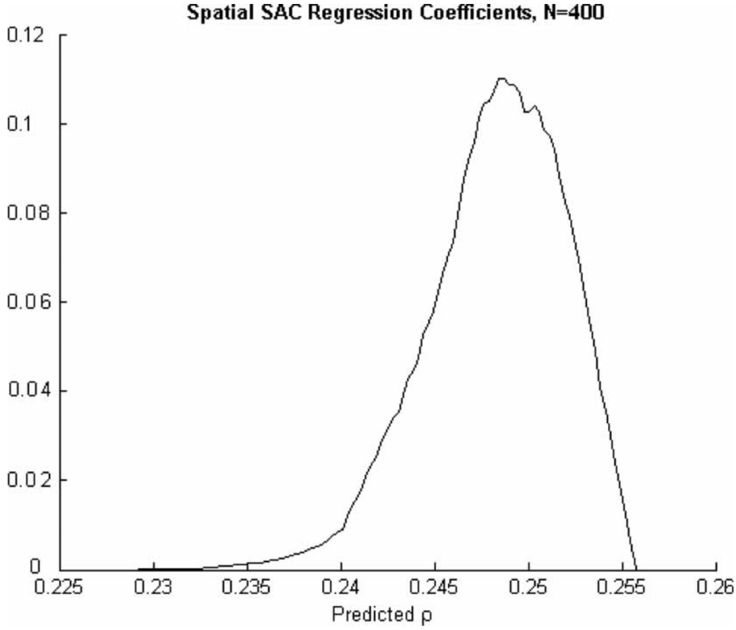

Figure 6. The distribution of the SAC coefficient for spatial nonsense regressions.

Table 6 records critical values for SAC* for different values of k when $p = 0.05$. As expected, SAC* varies inversely with k, especially the smaller N is, because there are fewer degrees of freedom.

6. Conclusion

This paper makes two contributions. We investigate spatial impulse responses for the SAR model in the presence and absence of spatial unit roots. We show that asymptotically space has an 'infinite spatial memory' when there is a spatial unit root such that infinitely remote shocks impact on spatial units as if distance did not matter. By contrast, in finite space, which has a natural edge, spatial memory ceases to be infinite and spatial impulses dissipate. However, there is a qualitative difference in the presence of unit roots; spatial impulses tend to linger. In contrast to time series, spatial impulses 'echo' and 'boomerang' because each unit is its neighbour's neighbour. This induces forward and backward linkages between spatial units. Here too, we show that there is a qualitative difference between SAR processes with and without unit roots.

Table 5. Spatial cointegration test statistics: SAC* ($k = 2$)

	N		
p	25	100	400
0.01	0.0450	0.198	0.239
0.05	0.107	0.215	0.243
0.1	0.1400	0.224	0.245

Table 6. Spatial cointegration test statistics: SAC* ($p = 0.05$)

	N		
k	25	100	400
2	0.107	0.215	0.243
3	0.073	0.205	0.240
4	0.034	0.197	0.238

Our second and main contribution is to report critical values for SAR coefficients when under the null hypothesis there is a unit root in spatial cross-section data. Our Monte Carlo computations follow procedures previously used by Dickey and Fuller who computed critical values for temporal unit roots.

Critical values for spatial unit roots tend to be larger than their time series counterparts. Indeed, they are very close to unity. This qualitative difference is explained by the fact that because there is more scope for interaction in spatial data than in time series data, it is easier to reject the null hypothesis of a spatial unit root than a temporal unit root. Indeed, this intuition is brought out in the critical values calculated for different symmetric topologies. For example, the critical value is smaller for oblong topologies than square topologies because there is more scope for interaction in squares than in oblongs that have the same number of spatial units.

If the DGPs for spatial cross-section data happen to be spatially nonstationary, the nonsense regression phenomenon arises in spatial cross-section data as pointed out by Fingleton (1999). We report critical values for spatial cointegration in spatial cross-section data. These critical values are designed to distinguish between genuine and nonsense regressions. Here too we follow procedures already developed for time series data. Specifically, we derive critical SAC values since we follow Engle & Granger in using residual-based cointegration test statistics.

Just as a variety of cointegration tests have been developed for time series data, we do not wish to suggest that residual-based test statistics are uniquely suited for testing spatial cointegration. We see no reason why Johansen-type tests and error correction tests for cointegration cannot be developed for spatial cross-section data. Johansen's reduced rank regression methodology could be applied to spatially filtered data, and spatial error correction models could form a basis for testing spatial cointegration. However, we hope the present paper has made a useful start by considering residual-based spatial cointegration tests.

Notes

1. Strictly speaking, Fingleton referred to 'nonsense' regression rather than 'spurious' regression. The latter, discovered by Yule (1897), arises when Y and X are independent random walks with drift. Drift causes the means of Y and X to increase over time, which induces spurious correlation. Nonsense regression, also discovered by Yule (1926), arises when Y and X are driftless random walks, and is induced by the fact that the variances of Y and X increase over time. The spatial DGPs studied by Fingleton have zero spatial drift, hence the correct adjective is nonsense rather than spurious. By contrast, the DGP in Mur & Trivez (2003) has spatial drift, so that the adjective spurious is appropriate in their case.
2. Strictly speaking, they calculate critical values for cointegration tests when there are two variables in the model.
3. The LM statistic is only valid for stationary residuals. Shin (1994) has suggested an LM statistic to test the null hypothesis of stationary time series residuals. We do not pursue this idea here.

4. If the model is $y = \beta x + u$ and u is stationary, the spatial ECM in the second stage should be $\Delta y = \delta \Delta x + \lambda W u + e$ where λ is the spatial error correction coefficient, $\Delta = I - W$ is the spatial difference operator and e is iid. Lauridsen & Kosfeld (2006) ignore λ. Their test implicitly assumes that Δx and u are independent.
5. Parameter estimates obtained from stationary data are root N-consistent. Stock (1987) showed that if time series models are cointegrated, OLS parameters estimated from difference stationary data are $N^{3/2}$-consistent.
6. This 'spatial echo' or 'boomerang effect' has no counterpart in time series data since the present may affect the future, but the future cannot affect the present. Space is multi-directional whereas time only moves forward.
7. This is another way in which space differs from time. Time has a natural beginning whereas space does not. Also time goes on forever whereas space is inherently finite. Indeed, spatial dynamics are conceptually different from temporal dynamics.
8. The weights sum to unity.
9. See on this, e.g., Sargent (1979, ch. 9).
10. *Ibid*., p. 179.
11. *Ibid*., p. 176.
12. In time series the impulse responses do not explode and tend to 1 because Y_{t+1} cannot feedback onto Y_t whereas in lateral space Y_{j+1} and Y_j feedback onto each other, i.e. time series data are uni-directional, whereas spatial data are multi-directional.
13. Strictly speaking $E(Y_j)$ is constant since we are ignoring the particular solution.
14. Unlike Fingleton we do not fix Y_{j^*} at the epicentre, we use 10,000 Monte Carlo simulations instead of 1,000, and we do not arbitrarily inflate spatial weights at the corners and edges of the lattice.
15. We calculate $N^{-2}\text{trace}(\Sigma)$.
16. The scale effect increases the variance linearly, as it does in time series. The scope effect makes the variance a nonlinear function of N. This constitutes a further difference between spatial and temporal nonstationarity.
17. In time series asymptotic orders may be calculated analytically. In spatial data analytical solutions are not available because spatial dependence is multi-directional. Therefore, asymptotic orders have to be calculated numerically as in Figure 3. In time series the asymptotic order of the numerator is zero. The covariance between X and u is asymptotically zero because u is stationary while X is nonstationary. We conjecture here that the speed (q) at which cov(Xu) tends to zero is similar to its time series counterpart.
18. In time series data $\omega = 1$. In spatial data $\omega > 1$ because of multi-directionality.
19. This result is well known for time series.
20. Lauridsen & Kosfeld (2006) note that this is, '. . . in principle doable although hardly practical in simulation studies' (p. 367). We use the ML procedure in Matlab's Econometric Toolbox, but allow the SAR coefficient to range between -2 and 2 instead of the default of -1 to 1.
21. Since the null hypothesis is $\alpha = 0.25$, it is no surprise that some estimates exceed 0.25. This also happens in time series, see Hendry (1995, p. 104). Indeed, the general shape of Figure 4 is similar to its time series counterpart, but it is less diffuse.
22. The computer time increases exponentially and multiplicatively with the number of trails, the number of seedings and N.
23. We thank Bernard Fingleton for supplying these data.
24. Critical values for spatial unit roots are larger than their time series counterparts. For example, if there are 100 observations the critical value for ρ at $p = 0.05$ is 0.863 according to Dickey and Fuller. Therefore it is easier to reject unit roots in spatial data than in time series data. We conjecture that this is because space is multi-directional whereas time only moves forward.
25. The distribution in Figure 5 is qualitatively similar to its time series counterpart (see Hendry, 1995, p. 124).
26. In Equation (2) ρ is normalized to ¼ instead of 1.
27. Lauridsen & Kosfeld (2004) calculate critical values for the Wald statistic under the null hypothesis that the residuals are nonstationary. They assume in Equations (1) and (2) that $\gamma = \beta = 1$, $X \sim U(0,1)$, $\rho = 1$, and $v \sim N(0,1)$. Based on 1,000 trials they calculate (rook case, $p = 0.05$) $\chi^2(1) = 4.83$ for $N = 25$. The Wald statistic must exceed this critical value to reject the null hypothesis of no cointegration. Surprisingly, their critical values increase with N. One would think that with more data it would be easier to reject the null hypothesis. Since ours is not a Wald test, it is difficult to compare our results with theirs.

References

Anselin, L. (1988) *Spatial Econometrics: Methods and Models*, London, Kluwer.

Engle, R. F. & Granger, C. W. J. (1987) Cointegration and error correction: representation, estimation and testing, *Econometrica*, 55, 251–276.

Fingleton, B. (1999) Spurious spatial regression: some Monte Carlo results with spatial unit roots and spatial cointegration, *Journal of Regional Science*, 39, 1–19.

Hendry, D. F. (1995) *Dynamic Econometrics*, Oxford, Oxford University Press.

Lauridsen, J. & Kosfeld, R. (2004) A Wald test for spatial nonstationarity, *Estudios de Economia Aplicada*, 22, 475–486.

Lauridsen, J. & Kosfeld, R. (2006) A test strategy for spurious spatial regression, spatial nonstationarity, and spatial cointegration, *Papers in Regional Science*, 85, 363–377.

Lauridsen, J. & Kosfeld, R. (2007) Spatial cointegration and heteroscedasticity, *Journal of Geographical Systems*, 9, 253–265.

Mur, J. & Trivez, F. J. (2003) Unit roots and deterministic trends in spatial econometric models, *International Regional Science Review*, 26, 289–312.

Sargent, T. J. (1979) *Macroeconomic Theory*, New York, Academic Press.

Shin, Y. (1994) A residual-based test of the null of cointegration against the alternative of no cointegration, *Econometric Theory*, 10, 91–115.

Stock, J. (1987) Asymptotic properties of least squares estimates of cointegrating vectors, *Econometrica*, 55, 1035–1056.

Yule, G. U. (1897) On the theory of correlation, *Journal of the Royal Statistical Society*, 60, 812–854.

Yule, G. U. (1926) Why do we sometimes get nonsense correlations between time series? A study in sampling and the nature of time series (with discussion), *Journal of the Royal Statistical Society*, 89, 1–64.