TESTING FOR UNIT ROOTS AND COINTEGRATION IN SPATIAL CROSS SECTION DATA

Michael Beenstock

Department of Economics
Hebrew University of Jerusalem

Daniel Felsenstein

Department of Geography
Hebrew University of Jerusalem

February 12, 2009

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 lattices where edge-effects are shown to influence spatial impulses. Monte Carlo simulation methods are used to compute critical values for spatial unit root tests in SAR models estimated from spatial cross-section data. We also compute critical SAC values for spatial cointegration tests for cross-section data that happen to be spatially nonstationary.

Thanks to Dan Feldman who carried out the simulations.

TESTING FOR UNIT ROOTS AND COINTEGRATION IN SPATIAL CROSS SECTION DATA

1. Introduction

Fingleton (1999) has placed the issue of "nonsense" regression on the agenda of spatial econometrics. He showed that if the data generating processes (DGP) for spatial cross-section data happen to contain a spatial unit root, the estimated regression coefficients from such data may be "nonsense" as originally defined by Yule (1926) for time series data¹. Yule pointed out that uncorrelated driftless random walks for Y and X typically generate non-zero correlations between the levels of Y and X despite the fact that the true correlation is zero. Phillips (1986, 1987) proved that the nonsense regression phenomenon is induced by unit roots in the DGPs for Y and X which cause the variances of Y and X to increase linearly with the sample size.

Fingleton suggested that the concept of cointegration proposed by Engle and 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 between Y and X will not be nonsense if Y and X are spatially cointegrated. If, on the other hand Y and X are not spatially cointegrated, the regression coefficient is nonsense. This will happen when the residuals contain a spatial unit root.

Fingleton did not, however, provide a spatial unit root test to determine whether the DGPs of spatial data contain spatial unit roots and are therefore spatially nonstationary. Nor did he provide a spatial cointegration test to determine whether parameter estimates are nonsense or not. He pointed out that if Moran's I indicates that the residuals are highly spatially correlated, the parameter estimates are more likely to be nonsense.

Our main purpose is 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. In doing so we apply to spatial cross section data

¹ Fingleton in fact referred to this as the "spurious" rather than the "nonsense" regression phenomenon. The former was also discovered by Yule (1897), which 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. 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 and Trivez (2003) has spatial drift, so that the adjective spurious is fitting in their case.

methods used by Dickey and Fuller for time series data². 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 methods used by Engle and Granger (1987) for nonstationary time series data. We derive the distribution of the SAC coefficient under the null hypothesis that it is one. The nonsense regression phenomenon arises when this null hypothesis cannot be rejected. If, however, the null hypothesis is rejected the data are spatially cointegrated.

Our approach is very different to that of Lauridsen and Kosfeld (2006, 2007) who have also suggested unit root and cointegration tests in spatial cross section data. As explained below, they use the ordinary LM test for spatial autocorrelation in which the null hypothesis is that the SAC coefficient is zero instead of 1. They discuss but evade the option of testing, for example, that the SAR coefficient is less than one. They recognize that this resembles the Dickey – Fuller approach in which the distribution of the SAR coefficient is nonstandard. We derive the nonstandard distribution of the SAR coefficient for spatial unit roots and the nonstandard SAC coefficient for spatial cointegration.

A secondary contribution is to highlight the differences between space and time. Although we are drawing on ideas from time series the space-time analogy is far from perfect. Unit root and cointegration tests in spatial data are more complicated than their time series counterparts because space, unlike time, has no natural direction. Time only moves forwards whereas space can "move" forward and backwards as well as upward and downwards. What happens in a given spatial unit depends on what is happening to its immediate neighbors to its left and right as well as above and below it. Therefore a shock that originates in a particular spatial unit rebounds back via its neighbors since each unit is its neighbor's neighbor. This "spatial echo" or "boomerang effect" has no counterpart in time series data. The present may affect the future, but the future cannot affect the present. Whereas time is a one-way street, lateral space is a two-way street, bilateral space is a four – way street, and multilateral space is an 2M – way street.

² Fuller (1976).

Also, whereas time may have a natural beginning, space does not have a natural starting point³. Another difference is that space on earth is finite⁴ whereas time is infinite. In a flat world, space comes to an end at the edge of the world. Regions on the edge have fewer neighbors. In a round world, however, there are theoretically no edge regions. However, the sea, mountains, deserts and political borders induce natural or man-made edges.

To show these differences we investigate spatial spillovers by obtaining the spatial impulse responses in which shocks in one spatial unit propagate between other spatial units and rebound back to their original source. We investigate these impulses in the absence and presence unit roots. We begin by assuming that space is two dimensional or lateral, i.e. there is east and west only. We are able in this case to obtain analytical solutions for spatial impulses and variances. Indeed, all the main intuitions are derived for lateral space. Subsequently, we consider the case of bilateral space where there is north and south as well as east and west. In this case we are no longer able to obtain analytical solutions for impulse responses and variances. Therefore, we compute them numerically from a finite square lattice. We draw attention to edge effects that inevitably arise when space is finite.

2. The Basic Problem

In what follows there are N spatial units labeled by j and spatial lags are labeled by tilde. The hypothesis of interest is assumed to be:

$$Y_j = \beta X_j + u_j \quad (1)$$

The SAC model for the residuals (u) is assumed to have a first - order SARMA representation:

$$u_j = \rho \tilde{u}_j + \phi \tilde{v}_j + v_j \quad (2)$$

where e.g. $\tilde{u}_j = \sum_{k \neq j}^N w_{jk} u_k$ denotes the spatially weighted⁵ average value of u among the neighbor's of j , ρ is the SAR coefficient of u and ϕ its SMA coefficient. The v 's are iid random variables.

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

³ Fingleton (1999) and Mur and Trivez (2003) assume that there is an unconnected spatial unit, which provides space with an artificial beginning. We drop this assumption.

⁴ Physicists disagree whether space itself is infinite.

⁵ The weights sum to unity.

$$Y_j = \rho_Y \tilde{Y}_j + \varepsilon_j \quad (3)$$

$$X_j = \rho_X \tilde{X}_j + e_j \quad (4)$$

where ε and e are iid random variables with correlation r . A variable Z is defined to be spatially integrated of order d when its d 'th spatial difference is spatially stationary, hence $Z \sim \text{SI}(d)$. Therefore $Y \sim \text{SI}(1)$ and $X \sim \text{SI}(1)$ when $\rho_Y = \rho_X = 1$. If $u \sim \text{SI}(1)$ because $\rho = 1$, equation (1) is a nonsense regression, which happens if ε and e are independent i.e. $r = 0$. The SMA coefficient ϕ has, of course, no implications for the spatial stationarity of u . If, however, $u \sim \text{SI}(0)$ because $\rho < 1$, equation (1) is not nonsense as pointed out by Fingleton (1999). This happens because $r \neq 0$, i.e. because ε and e are jointly distributed random variables.

This suggests that as in time series data, the cointegration test statistic for spatial cross section data should focus on rejecting the null hypothesis that $\rho \geq 1$. The same applies to spatial unit root tests for ρ_Y and ρ_X . However, Lauridsen and Kosfeld (2006) base their approach on information in which the null hypothesis is that ρ is zero instead of one. Just as in time series Dickey and Fuller and Engle and Granger hypothesized the existence of a unit root, so we contend that there is no alternative to obtaining the nonstandard distribution of the SAR and SAC coefficients under the null hypothesis that they are equal to one.

3. Spatial Difference Equations and Impulse Responses

In this section we discuss stochastic spatial difference equations. We show that spatial unit roots induce spatial impulse responses that do not die out with distance. If time series data are stationary, the impulse responses die out with time so that eventually history forgets itself. If, however, time series data are nonstationary, the impulse responses do not die out, history never forgets itself, and random shocks that occurred in the remote past affect the current value of the time series as if they had just occurred. If spatial data are stationary, shocks that occurred in the remote distance from region j have no effect on what happens in region j . If, however, spatial data are nonstationary, shocks that occurred in the remote distance affect region j as if they occurred in region j . In this case space "never forgets itself", and distance is of no consequence.

We begin by assuming that space is lateral, i.e. it is an infinite line that has no beginning and end. Spatial units have neighbors on two sides only. This turns out to

be equivalent to time series where the future and past are mutually dependent. Indeed, we use linear stochastic difference equation theory with "forward" and "backward" solutions to capture the two-way mutual dependence between neighbors. In this case we are able to obtain analytical solutions for the spatial impulse responses. However, we are unable to obtain analytical solutions for the more natural case in which space is four dimensional or bilateral, i.e. there is north and south, as well as east and west. In this case we use Matlab to calculate the spatial impulse responses. It turns out that these impulse responses are affected by how edge effects are handled.

3.1 Lateral Space

Regions are assumed to be located laterally (along an axis representing west and east or north and south) so that each region has a neighbor on either side. Regions continue to be labeled by $j = -\infty, \dots, \infty$. There is an infinite number of spatial units and unit $j + 1$ is unit j 's neighbor to the left or west and unit $j - 1$ is its neighbor to the right or east. Spillovers are assumed to occur between immediate neighbors. The SAR model in this case is:

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

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 2nd order stochastic spatial difference equation. The main difference between spatial and temporal difference equations is that whereas time only moves forward so that Y_t depends upon Y_{t-1} but not the other way round, in spatial data Y_j depends upon Y_{j-1} which in turn depends on Y_j .

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:

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

The characteristic equation of equation (7) is:

$$\lambda^2 - \alpha^{-1}\lambda + 1 = 0 \quad (8)$$

If $0 < \alpha < 1/2$ 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 $\alpha > 1/2$. When $\alpha = 1/2$ both roots are equal to unity. When $0 < \alpha < 1/2$ one

root is positive and less than one (λ_1) while the other (λ_2) is positive and greater than one. Since $(1 - \lambda_1 S)(1 - \lambda_2 S) = (1 - \alpha^{-1}S + S^2)$ the general solution⁶ to equation (6) is:

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

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 in two regions. Since in what follows we have no interest in the particular solution we ignore it by setting these arbitrary constants to zero. Using partial fractions⁷ we note that:

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

We also note that⁸:

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

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

where distance from j is measured by 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:

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

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

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

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

⁶ See e.g. Sargent (1979) chapter 9.

⁷ Ibid p 179.

⁸ Ibid page 176.

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 neighbors ($i = 1$) 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 neighbors 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 $\frac{1}{2}$ λ_1 approaches 1 and the impulse responses approach infinity. When $\alpha = \frac{1}{2}$ both roots equal 1, the impulse responses explode and no longer depend on distance⁹.

According to equation (13) $E(Y_j) = 0$ because the expected value of the u 's are all zero¹⁰ by definition. Therefore, the first moment is independent of j . The variance of Y from equation (13) is equal to:

$$\text{var}(Y) = \frac{\alpha^{-2}(1 + \lambda_1^2)}{(1 - \lambda_1^2)(\lambda_1^{-1} - \lambda_1)^2} \sigma_u^2 \quad (15)$$

which is finite since $0 \leq \lambda_1 < 1$ and it does not depend on j . Therefore if $0 < \alpha < \frac{1}{2}$ the first and second moments of Y_j are finite and independent of j .

3.2 Bilateral Space

When space is lateral and the number of neighbors (n) is two, we saw that the SAR coefficient inducing a spatial unit root is $\alpha^* = \frac{1}{2} = 1/n$. When space is multilateral and spatial units have more than two neighbors 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 4 neighbors in which case $\alpha^* = \frac{1}{4}$.

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

$$Y = \alpha WY + u \quad (17)$$

⁹ 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.

¹⁰ Strictly speaking $E(Y_j)$ is constant since we are ignoring the particular solution.

where W is an $N \times N$ matrix with elements $w_{jk} = 1$ if j and k are neighbors and $w_{jk} = 0$ otherwise. Y and u are vectors of length N^2 . Fingleton (1999) normalized the sum of weights (w) to unity, and normalized $\alpha = 1$. This means that at the corners of the lattice where there are two neighbors $w_{jk} = 1/2$ and at the edge where there are three neighbors $w_{jk} = 1/3$, which overstates the true weight. We therefore prefer to normalize $w_{jk} = 1$ in which case the sum of weights is 4 and $\alpha^* = 1/4$ because this does not artificially increase spatial spillover at the corners and along the edges of the lattice. 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 (17) is:

$$Y = Au \quad (18)$$

$$A = (I - \alpha W)^{-1} = I_N + \sum_{i=1}^{\infty} \alpha^i W^i$$

where $W^i u$ denotes u among i 'th order neighbors. The spatial impulse responses are:

$$\frac{\partial Y_j}{\partial u_j} = a_{jj} \quad (19)$$

$$\frac{\partial Y_j}{\partial u_i} = a_{ji} \quad (20)$$

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. Unfortunately when space is bilateral analytical expressions for the spatial impulse responses are unobtainable. We therefore use Matlab to calculate A for $N \times N$ lattices in which $n = 4$ and $\alpha^* = 1/4$. 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 given computing constraints.

Because N is finite A is inevitably distorted by edge effects. Spatial units on the edge are less exposed to spatial spillover because they have only three neighbors instead of four. Spatial units in the four corners of the lattice only have two neighbors. These edge effects inevitably distort the calculation of A . We expect a_{jj} and a_{ji} to be greater the closer is j to the epicenter j^* because there is more scope for spatial interaction in the center than at the periphery. We do not expect a_{ji} to be symmetrical

unless $j = j^*$ because only at the epicenter is the distance to the edge of the square lattice the same in all four directions.

Figure 1 Spatial Impulses

In Figure 1 we plot impulse responses for a_{j^*i} when $N = 31$. The impulse responses are measured along the vertical and Euclidean distance 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 the larger is α . It might have been expected that when $\alpha = \alpha^* = 1/4$ 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 $\alpha = 1/4$ there is a qualitative difference and the impulses linger longer in space than when $\alpha = 0.2$.

Figure 2 Local Impulses

As indicated in equation (19) 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 epicenter of the lattice. The local impulses exceed unity because when a shock occurs in j it affects j 's neighbors 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 distort local impulses because the echo does not extend very far. Hence the local impulse does not depend upon distance from the epicenter, 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 epicenter. This happens because the echo extends far; so that close to the epicenter there is much more echo than at the edges. Had there been no edge the effect at the epicenter would have been infinity because the echo carries on for ever, and the effect elsewhere would have been infinite too. Indeed, when N is infinity there is no meaning to the epicenter because the lattice has no borders.

3.3. Variances in Bilateral Space

When space is lateral there is an analytical expression for the variance of Y , see equation (15). When space is bilateral equation (18) implies that the variance-covariance matrix of Y is equal to:

$$\Sigma = \sigma_u^2 A' A \quad (21)$$

Figure 3 Relationship between Variance and Sample Size

We follow Fingleton (1999) and calculate¹¹ the average variance of Y as N increases¹². The results are plotted on 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^* . 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.

4. Spatial Unit Root Tests¹³

We set $\alpha = \alpha^* = 1/4$ in equation (17) and generate 10,000 artificial data sets for Y by drawing the u 's using pseudo random numbers from a standard normal distribution for given N . We use these synthetic data sets to estimate by maximum likelihood¹⁴ 10,000 SAR models. The distribution of the 10,000 estimates of the SAR coefficient is plotted in Figure 4 for the case when $N = 20$, i.e. there are 400 spatial units in a square lattice.

¹¹ Unlike Fingleton we do not fix Y_{j^*} at the epicenter, 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.

¹² We calculate $N^{-2}\text{trace}(\Sigma)$.

¹³ Spatial unit root tests should not be confused with temporal unit root tests for spatially dependent data as in Baltagi, Bresson and Pirrote (2007)

¹⁴ Lauridsen and Kosfeld (2006) note that this is, "... in principle doable although hardly practical in simulation studies..." (page 367).

Figure 4 The Distribution of SAR Coefficients when $\alpha = \frac{1}{4}$

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 of the estimates exceed 0.25, while others are less than 0.25. The distribution is clearly skewed to the left. Indeed, this result is qualitatively similar to its time series counterpart¹⁵. We truncate the distribution from the right since SAR estimates that are equal to or greater than 0.25 imply a spatial unit root, and calculate the percentiles for the SAR coefficient from the truncated distribution. Since the probability of obtaining an estimate of α greater or equal to 0.25 is 0.4776 we truncate Figure 4 from the right accordingly. According to Figure 4, when $N = 20$ there is a 95 percent 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 $N = 10$ 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.

Table 1 Spatial Unit Root Test Statistics

¹⁵ See Hendry (1995) page 104.

	N		
P	5	10	20
0.01	0.071	0.209	0.24
0.05	0.139	0.225	0.243
0.1	0.161	0.231	0.244

The critical values in Table 1 are conceptually similar to the $Z = T(\rho - 1)$ statistic that is (infrequently) used for testing unit roots in time series data, where T is the sample size and ρ is the AR coefficient, which parallels the SAR coefficient α . The number of spatial units in the sample is N^2 in which case $Z = N^2(4\alpha^* - 1)$. For example, when $N = 10$ and $p = 0.05$ Table 1 implies that $Z = -10$. Note that when $T = 100$ the critical value of Z for a driftless random walk is -7.9 . We compare with the driftless case because there is no spatial drift in the SAR model.

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 $N = 10$ 500 times and reduced¹⁶ the number of Monte Carlo trials from 10000 to 5000. For example when $p = 0.05$ one standard deviation of the critical value of ρ^* is about $\frac{1}{2}\%$ and α^* is bounded between 0.224 and 0.226. Therefore, the critical values reported in Table 1 are reliable.

Table 2 Confidence Intervals for Table 1 (N = 10)

	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

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 we set $N = 10$ for reasons of economy. 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 1000 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 α^* varies directly with

¹⁶ The computer time increases exponentially and multiplicatively with the number of trails, the number of seedings and N .

the number of trials. Clearly, to dig into the tail of the distribution requires increasing the number of trials.

Table 3 Sensitivity of Table 1 to the Number of Trials (N = 10)

	Trials				
	500	1000	5000	10000	15000
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 time series the Dickey-Fuller statistic has proved much more popular than the Z statistic even though the two statistics are equivalent. The DF statistic is equal to $\rho - 1$ divided by its standard error and consequently has the dimension of a t-statistic. We may therefore express the critical values reported in Table 1 as “spatial Dickey-Fuller” equivalents, or SDF statistics, which are reported in Table 4.

Table 4 Spatial Dickey-Fuller Statistics

	N		
P	5	10	20
0.1	-1.04	-0.98	-1.21
0.05	-1.22	-1.14	-1.32
0.01	-1.62	-1.49	-1.61

The critical value of the DF statistic when $T = 100$ and $p = 0.05$ is -1.95. Table 4 suggests that the spatial counterpart to the DF statistic is smaller in absolute value than its temporal counterpart. Of course, these statistics are not strictly comparable because time only moves forward whereas space moves in all four directions in a square lattice.

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 $Y \sim \text{SI}(1)$ then it must be the case that $u \sim \text{SI}(1)$ 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 $\alpha = 1/4$. 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 $N = 20$. The distribution in Figure 5 is qualitatively similar¹⁷ to its time series counterpart insofar it is approximately normal and the mean is not exactly zero.

Figure 5 The Distribution of the Spatial Nonsense Regression Coefficient

Not surprisingly, the mode is close to zero (-0.0088) and the mean is close to zero (0.006) as well. However, there are positive as well as negative estimates of β . The residuals from these nonsense regressions are used to estimate¹⁸ 10,000 estimates of ρ , which are plotted on Figure 6.

The mode in Figure 6 is 0.2527 and the mean is close to 0.25 (0.2482). However, there are estimates that are below and above 0.25. As in section 4 we truncate Figure 6 from the right at 0.25 and calculate the probability that $\rho < 0.25$. The area that is cut off is 30.14% of the overall area. When $p = 0.05$ Figure 6 implies that $\rho^* = 0.241$. If $\rho < \rho^*$ the OLS estimate of β is not nonsense in which event Y and X are spatially cointegrated. If, on the other hand, $\rho > \rho^*$ 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.

¹⁷ See Hendry (1995) p 124.

¹⁸ In equation (2) ρ is normalized to $1/4$ instead of 1 and ϕ is assumed to be zero.

Figure 6 The Distribution of the SAC Coefficient for Spatial Nonsense Regressions

Table 5 Spatial Cointegration Test Statistics

	N		
P	5	10	20
0.01	0.038	0.198	0.238
0.05	0.101	0.217	0.241
0.1	0.132	0.224	0.243

$k = 2$

Table 5 records the critical value of ρ^* in the bivariate case ($k = 2$). Note that ρ^* in Table 5 is typically smaller than α^* in Table 1. This discrepancy parallels its temporal counterpart where, due to the loss in degrees of freedom, the univariate DF statistic is more negative than its multivariate counterpart for cointegration. Therefore because $k = 2$ ρ^* must be less than α^* . For example when $N = 10$ and $p = 0.05$ ρ^* is 0.217 whereas α^* is 0.225.

Table 6 records critical values of ρ^* for larger values of k when $p = 0.05$. Not surprisingly ρ^* varies directly with N because there are more degrees of freedom and it varies inversely with k because there are fewer degrees of freedom. The SDF counterparts to Table 6 are given in Table 7. The SDF statistics in Table 7 are

naturally more negative than their univariate counterparts in Table 4 due to the reduction in degrees of freedom.

Table 6 Spatial Cointegration Test Statistics ($k > 2$)

	N		
K	5	10	20
3	0.069	0.205	0.239
4	0.035	0.197	0.238

$p = 0.05$

Table 7 SDF Cointegration Test Statistics ($k > 2$)

K	5	20
3	-1.70	-1.61
4	-1.93	-1.77

$p = 0.05$

6. Conclusion

This paper makes two contributions. We investigate spatial impulse responses from the SAR model in the presence and absence of spatial unit roots. We show that asymptotically space has an "infinite 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 bounded 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 neighbor's neighbor. 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.

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. In contrast to Lauridsen and Kosfeld (2006) we consider it inappropriate to base unit root tests on test statistics, such as the LM test, which under the null assume that the root is zero instead of one.

If the DGPs for spatial cross-section data happen to be 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. Here too we follow procedures already developed for time series data. Specifically, we derive critical SAC values since we follow Engle and Granger in using residual-based 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.

References

- Baltagi B.H., G. Bresson and A. Pirotte (2007) Panel unit root tests and spatial dependence, *Journal of Applied Econometrics*, 22:339-60.
- Engle R.F. and C.W.J. Granger (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.
- Fuller W.A. (1976) *Introduction to Statistical Time Series*, New York, John Wiley.
- Hendry D.F. (1995) *Dynamic Econometrics*, Oxford University Press.
- Lauridsen J. and R. Kosfeld (2006) A test strategy for spurious spatial regression, spatial nonstationarity, and spatial cointegration, *Papers in Regional Science*, 85: 363-377.
- Lauridsen J. and R. Kosfeld (2007) Spatial cointegration and heteroscedasticity, *Journal of Geographical Systems*, 9: 253-265.
- Mur J. and F.J. Trivez (2003) Unit roots and deterministic trends in spatial econometric models, *International Regional Science Review*, 26: 289-312.
- Phillips P.C.B. (1986) Understanding spurious regressions in econometrics, *Journal of Econometrics*, 33: 311-40.
- Phillips P.C.B. (1987) Time series regression with a unit root, *Econometrica*, 55: 277-301.
- Sargent T.J. (1979) *Macroeconomic Theory*, New York, Academic Press.
- Yule G.U. (1897) On the theory of correlation, *Journal of the Royal Statistical Society*, 60: 812-54.
- 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.