TESTING SPATIAL STATIONARITY AND SPATIAL COINTEGRATION

Michael Beenstock
Department of Economics
Hebrew University of Jerusalem
Daniel Felsenstein
Department of Geography
Hebrew University of Jerusalem

October 19, 2008

Abstract

Monte Carlo simulation methods are used to compute critical values for spatial unit root tests in the SAR model estimated from spatial cross-section data. We also compute critical values for residual-based spatial cointegration tests for cross-section data that happen to be spatially nonstationary. We show that edge effects in spatial lattices affect the asymptotics of the proposed tests.

We wish to thank Dan Feldman who carried out the computations. The critical values that we report are provisional.

1. Introduction

Fingleton (1999) has placed the issue of spurious regression on the agenda of spatial econometrics. He pointed out 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 spurious. Indeed, they will be spurious if the regression residuals contain a spatial unit root. If, on the other hand, the residuals do not contain a spatial unit root, the data are spatially cointegrated and the estimated regression coefficients are not spurious. This definition of spatial cointegration is in the spirit of Engle and Granger (1987) who first developed the concept for time eries data.

Strictly speaking if the residuals contain a spatial unit root, the estimated regression coefficients will be "nonsense" rather than "spurious". These terms were invented by Yule for time series data. The spurious regression phenomenon arises when time series happen to be independent random walks with drift, in which case the spurious regression coefficient is equal to the ratio of the drift parameters. The nonsense regression phenomenon (Yule 1926) arises when the time series are independent, driftless random walks. Spurious regression is induced by the fact that the means of the time series increase or decrease with time. Nonsense regression is induced by the fact that their variances increase with time[1].

The spatial DGPs investigated by Fingleton were in fact driftless random walks, so his topic was spatial nonsense regression rather than spatial spurious regression. The latter would arise if the DGPs contained spatial drift induced, for example, by the distance from the central business district in urban economic models. Spatial nonsense regression, on the other hand, is not induced by spatial drift, but by unit roots in the spatial DGPs as demonstrated by Fingleton.

Fingleton did not, however, 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. Taking up a suggestion of Fingleton's, Lauridsen and Kosfeld (2006, 2007) proposed a two-step LM strategy to test for spatial cointegration. We show that their strategy is based on a conceptual error and is therefore invalid.

[1] See e.g. Hendry (1995) pages 122-133 for further discussion.

Our main purpose is therefore to propose a valid spatial cointegration test, which may be used to determine whether parameter estimates obtained from spatial cross section data are nonsense. However, first it is necessary to propose a test statistic to determine whether the DGPs contain spatial unit roots. We begin by developing spatial "Dickey - Fuller" statistics where according to the null hypothesis there is a spatial unit root. We use Monte Carlo simulation to determine the distribution of the SAR coefficient when there is a spatial unit root. We obtain the spatial analogue to the Dickey - Fuller test statistic developed for time series data. This spatial unit root test may be used to determine the order of spatial integration of spatial data. If the data are spatially integrated to order zero, the nonsense regression issue does not arise. If, however, the data are spatially integrated to order 1, the nonsense regression issue arises.

We use Monte Carlo simulation to develop spatial cointegration test statistics, which parallel the cointegration tests developed by Engle and Granger (1987) for time series. The nonsense regression phenomenon implies that there is a spatial unit roots in the regression residuals. We calculate the distribution of the SAC coefficient for the regression residuals under the null hypothesis that they contain a spatial unit root.

2. The Lauridsen – Kosfeld Test

We begin by restating the spatial cointegration test proposed by Lauridsen and Kosfeld (2006). We show that the KL test is based on a conceptual error. Finally, we show how if the KL test was applied to time series, the test would be invalid.

The hypothesis of interest is assumed to be:

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

$$u_j = \rho \tilde{u}_j + v_j \qquad (2)$$

where j refers to region, $\tilde{u}_j$ denotes the average value of u among the neighbor's of j and ρ is the SAC coefficient. The DGPs for Y and X are assumed to be:

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

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

where ε and e are iid random variables. A variable Z is defined to be spatially integrated of order d when its d'th spatial difference is spatially stationary, hence Z ~ SI(d). Therefore Y ~ SI(1) and X ~ SI(1) when $\rho_Y = \rho_X = 1$. If u ~ SI(1) because ρ = 1 equation (1) is a nonsense regression, which happens if ε and e are independent for all spatial leads and lags. If, however, u ~ SI(0) because $\rho < 1$ equation (1) is not nonsense because ε and e must be dependent.

If $\rho = 1$ LK suggest that equation (1) be spatially differenced:

$$y_j = \beta' x_j + v'_j \qquad (5)$$

where

$$y_j = Y_j - \widetilde{Y}_j$$
$$x_j = X_j - \widetilde{X}_j$$

Their contention is that if v\` is spatially uncorrelated, equation (1) is nonsense. If v\` is spatially correlated then equation (1) is spatially cointegrated.

Notice that equation (5) is equivalent to a regression of ε_j on e_j since from equation (3) y = ε and from equation (4) x = e. If ε and e are not independent, equation (1) must be cointegrated regardless of whether or not v\` is spatially autocorrelated. Therefore KL's proposed test is incorrect. They are also incorrect in assuming that $\hat{\beta} = \hat{\beta}'$. If indeed equation (1) is a nonsense regression plim $\hat{\beta} \neq 0$ while plim $\hat{\beta}' = 0$.

The invalidity of KL's proposed test may be further highlighted by applying it to time series data. The data generating processes (DGP) for Y and X are assumed to be first-order driftless random walks, i.e. they are I(1) variables:

$$\Delta Y_t = \varepsilon_t \qquad (6)$$
$$\Delta X_t = e_t \qquad (7)$$
$$\varepsilon \approx iin(0,1) \quad e \approx iin(0,1) \quad E(\varepsilon_t e_{t-s}) = 0 \ all\ s$$

Y is regressed on X (or X on Y) as in equation (8):

$$Y_t = \theta X_t + u_t \qquad (8)$$

Since the DGPs are entirely independent the true value of θ is zero. However, OLS estimates of θ are "nonsense" because plim $\hat{\theta} \neq 0$ and the t-statistic of $\hat{\theta}$ tends to infinity[2] at the rate of root T. In nonsense regressions u ~ I(1). The cointegration test due to Engle and Granger (1987) tests whether u is I(1) or not.

Translating KL's proposal to time series is equivalent to applying the following two-stage procedure. First test the hypothesis that $\rho = 0$ (instead of $\rho = 1$) in equation (9):

[2] See e.g. Hendry (1995) p126. The nonsense regression phenomenon also arises when equations (6) and (7) are autoregressive.

$$u_t = \rho u_{t-1} + \omega_t \qquad (9)$$

If $\rho \neq 0$, estimate equation (8) in first differences:

$$\Delta Y_t = \theta`\Delta X_t + v_t \qquad (10)$$

KL make three claims. i) If u is serially uncorrelated then Y and X are cointegrated. ii) If u is serially correlated but v is serially independent, Y and X are not cointegrated. iii) If both u and v are serially correlated Y and X are cointegrated. The first claim is correct because if u is serially uncorrelated u ~ I(0). The other claims are incorrect because they are based on a conceptual error, which is to assume that $p\lim\hat{\theta} = p\lim\hat{\theta}^{`}$. The DGPs in equations (6) and (7) mean that in equation (8) $p\lim\hat{\theta} \neq 0$ due to the nonsense regression phenomenon, whereas in equation (10) plim $\hat{\theta}^{`}$ = 0 because ΔY and ΔX are stationary and ε and e are independent at all leads and lags. Indeed, the latter means that equation (10) is equivalent to v = ε. In short, v is not relevant to cointegration testing, which explains why cointegration tests for time series are based on u and not on v.

3. Spatial Difference Equations and Impulse Responses

In this section we introduce 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.

The analogy between space and time is, however, incomplete because time is unidirectional; it only moves forward. Space, on the other hand, is multidirectional because what happens in region j is affected by neighbors to the north, south, east and west. The econometric analysis of time series would be vastly more complicated if the past depended on the future and not just the future on the past. Also, whereas time may have a natural beginning, space does not have a natural starting point. In a flat world, regions on the edge have fewer neighbors, but this is a different matter. In a round world, however, there are no edge regions. Finally, if the world does not come to an end, the future is infinite, whereas space (on earth) is naturally finite.

We begin by assuming that space 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 relevant case in which space is four dimensional, 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

Spatial units are assumed to be located laterally (along a line representing west and east or north and south) so that each region has a neighbor on either side. Spillovers are assumed to occur between immediate neighbors. The SAR model in this case is:

$$Y_j = \alpha(Y_{j+1} + Y_{j-1}) + u_j \qquad (11)$$

Where α denotes the spatial spillover coefficient and u is an iid random variable with variance equal to σ_u^2. Equation (11) 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 $S^iY_j = Y_{j-i}$ where i may be positive (east of j) or negative (west of j). Multiplying equation (11) by S and rewriting the result in terms of the spatial lag operator gives:

$$(\alpha - S + \alpha S^2)Y_j = -u_{j-1} \qquad (12)$$

The characteristic equation of equation (12) is:

$$\alpha\lambda^2 - \lambda + \alpha = 0 \qquad (13)$$

If $\alpha > 0$ the real roots of equation (13), λ_1 and λ_2, are positive and are reciprocally related because $\lambda_1 = 1/\lambda_2$. The roots will be complex if $4\alpha^2 > 1$. When $\alpha = ½$ both roots are equal to unity. When $\alpha < ½$ one root is positive and less than one while the other is positive and greater than one. Therefore if $\lambda_1 < 1$ $\lambda_2 = 1/\lambda_1 > 1$. Since $(1 - \lambda_1 S)(1 - \lambda_2 S) = (\alpha - S + \alpha S^2)$ the general solution for Y_j to equation (12) is:

$$Y_j = -\frac{u_{j-1}}{(1-\lambda_1 S)(1-\lambda_2 S)} \qquad (14)$$

Using partial fractions we note that:

$$\frac{1}{(1-\lambda_1 S)(1-\lambda_2 S)} = \frac{1}{\lambda_1 - \lambda_2}\left[\frac{\lambda_1}{1-\lambda_1 S} - \frac{\lambda_2}{1-\lambda_2 S}\right] \qquad (15)$$

We also note that:

$$\frac{1}{1-\lambda_1 S} = \sum_{i=0}^{\infty} \lambda_1^i S^i \qquad (16)$$

$$\frac{1}{1-\lambda_2 S} = -\frac{(\lambda_2 S)^{-1}}{1-(\lambda_2 S)^{-1}} = -\frac{\lambda_1 S^{-1}}{1-\lambda_1 S^{-1}} = -\lambda_1 \sum_{i=1}^{\infty} \lambda_1^{i-1} S^{-i} \qquad (17)$$

Equation (16) operates eastwards since $i \geq 0$ and equation (17) operates westwards since $i < 0$. Substituting equations (15), (16) and (17) into equation (14) gives:

$$Y_j = \frac{1}{\lambda_1^{-1} - \lambda_1}\left[\sum_{i=1}^{\infty} \lambda_1^i u_{j-i} + \sum_{i=0}^{\infty} \lambda_1^i u_{j+i}\right] \qquad (18)$$

Equation (18) is the spatial Wold representation of equation (11) since it expresses Y_j in terms of the stochastic shocks in all spatial units to the east and west of j as well as in j itself. Equation (18) is also the spatial impulse response function. Because $\lambda_1 < 1$ equation (18) states that closer spatial units to j have a greater effect on j than more remote units. The spatial impulse responses are symmetric since u_{j+i} has the same effect on Y_j as u_{j-i}.

$$\frac{\partial Y_j}{\partial Y_{j-i}} = \frac{\partial Y_j}{\partial Y_{j+i}} = \frac{\lambda_1^i}{\lambda_1^{-1} - \lambda_1} \qquad (19)$$

Infinitely remotespatial units have asymptotically no effect on j. If, however, $\alpha = ½$, $\lambda_1 = 1$ in which case the spatial impulses do not die away with distance.

Setting $i = 0$ in equation (18) solves for the effect of u_j on Y_j :

$$\frac{\partial Y_j}{\partial u_j} = \frac{1}{\lambda_1^{-1} - \lambda_1} = \frac{\lambda_1}{1-\lambda_1^2} > 0 \qquad (21)$$

which is positive and exceeds 1 if $\lambda_1 < 0.618$ which from equation (13) happens when $\alpha < 0.447$.

According to equation (18) $E(Y_j) = 0$ because the expected value of the u's are all zero by definition. Therefore, the first moment is independent of j, even in the case where $\lambda_1 = 1$ ($\alpha = ½$). However, the variance of Y is not independent of j when $\lambda_1 = 1$. From equation (18) the asymptotic variance is equal to:

$$\mathrm{var}(Y) = \frac{1+\lambda_1^2}{(1-\lambda_1^2)(\lambda_1^{-1}-\lambda_1)^2}\sigma_u^2 \qquad (21)$$

which is finite when $0 \leq \lambda_1 < 1$. If, however, $\lambda_1 = 1$ ($\alpha = ½$) the asymptotic variance is infinite. This happens because the spatial impulses in equation (18) do not vary inversely with distance (i) when $\lambda_1 = 1$. Therefore a remote shock has the same effect on region j as if it happened in region j itself. This parallels temporal nonstationarity where historically remote shocks have the same effect on time t as a current shock.

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 $\alpha^* = ½ = 1/n$. 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 so that space is bilateral each spatial unit has 4 neighbors in which case $\alpha^* = ¼$. If space is a queen lattice space would be trilateral. We focus here on bilateral space. Multilateral extensions require exponentially more computing power.

The bilateral counterpart to equation (11) may be written familiarly as the SAR model:

$$Y = \alpha WY + u \qquad (22)$$

where W is an NxN 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. Following Fingleton (1999), Lauridsen and Kosfeld normalize so that the sum of the weights (w) is unity, and normalize $\alpha = 1$. This means that at the corners of the lattice where there are two neighbors $w_{jk} = ½$ and at the edge where there are three neighbors $w_{jk} = ⅓$, which overstates the true weight. We therefore prefer to normalize $w_{jk} =1$ and $\alpha = ¼$ because it does not artificially increase spatial spillover at the corners and edge of the lattice.

The Wold representation of equation (22) is:

$$Y = Au \qquad (23)$$

$$A = (I - \alpha W)^{-1}$$

and the spatial impulse responses are:

$$\frac{\partial Y_j}{\partial u_j} = a_{jj} \qquad (24)$$

$$\frac{\partial Y_j}{\partial u_k} = a_{jk} \qquad (25)$$

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_{jk} to vary inversely with the distance between j and k. If, however, there is a spatial unit root, the impulses a_{jk} will not tend to zero as the distance between j and k 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 NxN lattices in which n = 4 and $\alpha^* = ¼$. To investigate asymptotics we ideally wish to let N tend to infinity, but this is not 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_{jk} 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_{jk} 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.

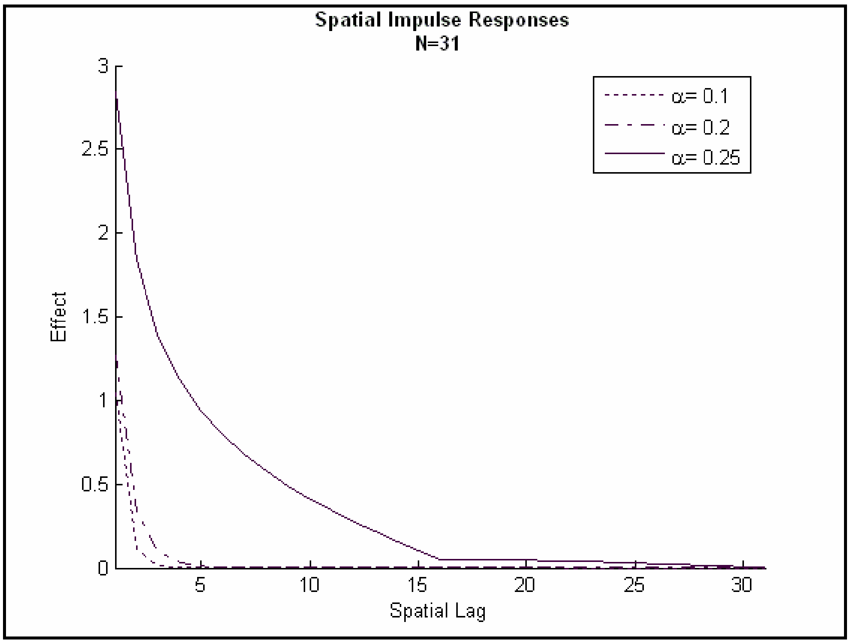

Figure 1

In Figure 1 we plot impulse responses for a_{j*k} 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 k is on the edge of the lattice and the 30th value is at the corner of the lattice. Figure 1 shows, as expected, that a_{jj} varies directly with α, a_{jk} varies inversely with k, and the impulse responses die away more slowly the larger is α. A perhaps surprising result is that when $\alpha = \alpha^* = ¼$ the impulse responses die away more slowly but do not reach zero. This happens because there is an edge to the lattice.

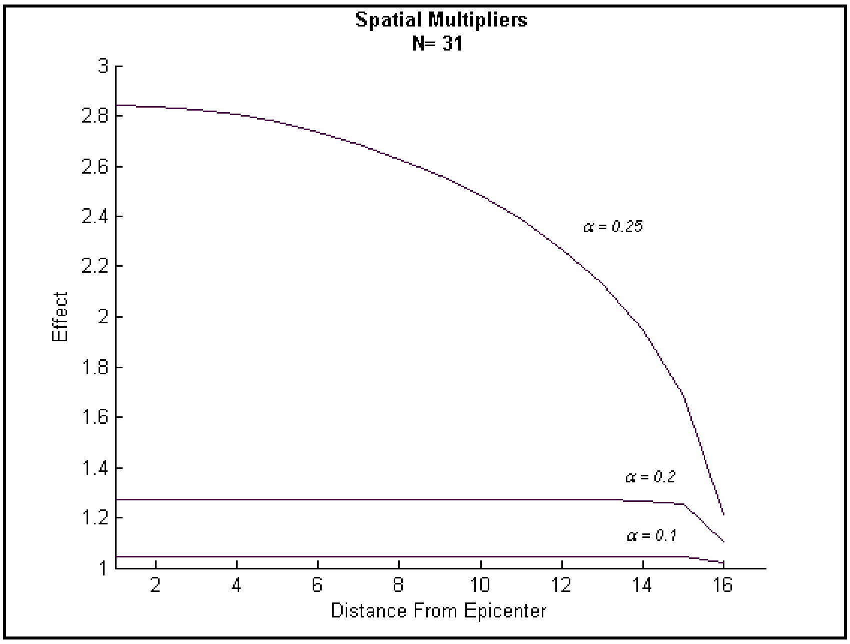

Figure 2

In Figure 2 we plot the diagonal of A to show, as expected, that a_{jj} varies directly with α. This is because a_{jj} reflects the feedback or echo from neighbors from shocks in a given region. The strength of this "echo" naturally varies directly with α. Figure 2 also shows that when $\alpha < \alpha^*$ the edge effect does not distort the force of the echo because in any case the echo tends to die away. Hence a_{jj} 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 a_{jj} varies inversely with distance from the epicenter. 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 (21). When space is bilateral equation (23) implies that the variance-covariance matrix of Y is equal to:

$$\Sigma = \sigma_u^2 A`A \qquad (26)$$

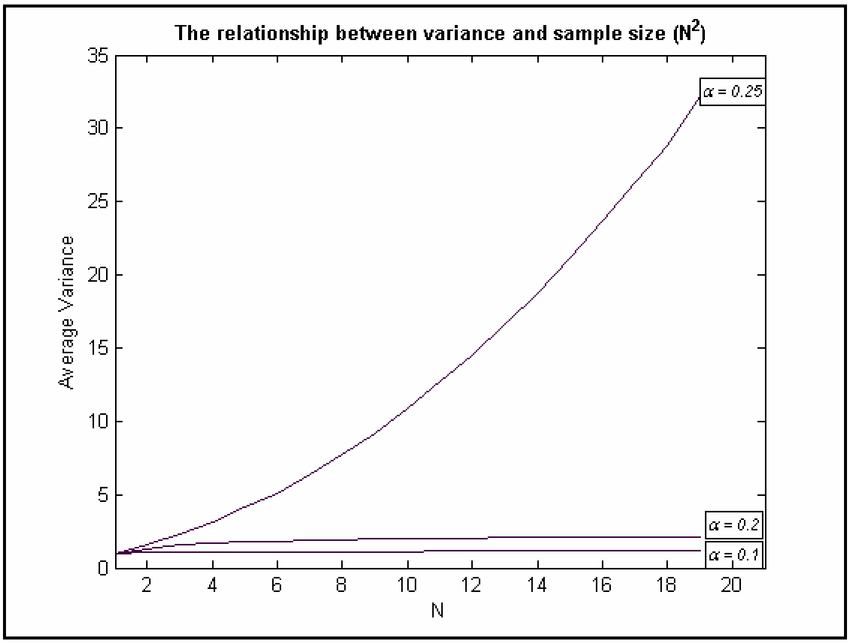

Figure 3

We follow Fingleton (1999) and calculate[3] the average variance of Y as N increases[4]. The results are plotted on Figure 3, which shows that as α increases towards $\alpha^* = ¼$, 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.

4. Spatial Unit Root Tests[5]

We set $\alpha = \alpha^* = ¼$ in equation (21) and generate 10,000 artificial data sets for Y by using the Meersen Twister[6] for drawing pseudo random numbers for the u's from a standard normal distribution for given N. We use these synthetic data sets to estimate by maximum likelihood 10,000 SAR models. The distribution of the 10,000 estimates of the SAR coefficient is plotted in Figure 4.

[3] 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 reduce w on the edge of the lattice.
[4] We calculate $N^{-2}\mathrm{trace}(\Sigma)$.
[5] Spatial unit root tests should not be confused with temporal unit root tests for spatially dependent data as in Baltagi, Bresson and Pirrote (2007)
[6] The Meersen Twister is the default in Matlab. We are currently investigating other pseudo random number generators such as Halton sequences. We are also investigating the sensitivity of the computations to the number of Monte Carlo trials.

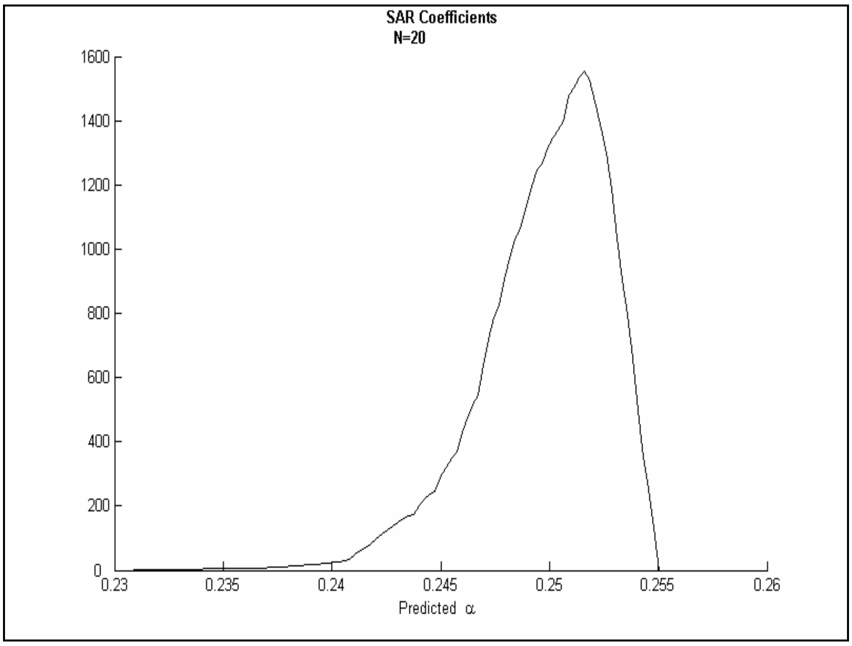

Figure 4

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 counterpart[7]. 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 truncate Figure 4 from the right accordingly, i.e. the area that is cut off is A=47.76% of the total area. 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[8]. 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 spacial 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.

[7] See Hendry (1995) page 104.
[8] The critical values reported in Tables 1 – 5 are seed dependent and are therefore random. We have yet to calculate their standard errors, which involves repeated seeding of the 10,000 trials.

Table 1 Spatial Unit Root Test Statistics

	N		
p	**5**	**10**	**20**
0.01	0.071	0.206	0.24
0.05	0.139	0.225	0.243
0.1	0.161	0.23	0.244

The critical values in Table 1 are conceptually similar to the Z = T(ρ - 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.

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 ρ - 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 2.

Table 2 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 2 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 ~SI(1) then it must be the case that u ~ 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 $\alpha = ¼$ with $\beta =0$ in each random 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.

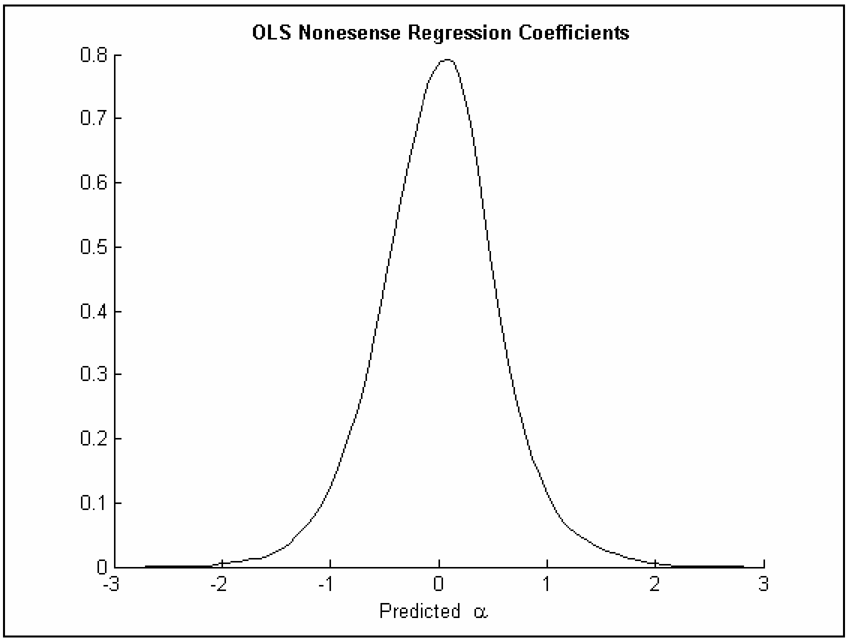

Figure 5

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. Our 10,000 regressions generate residuals which are used to generate 10,000 estimates of ρ, which are plotted on Figure 6.

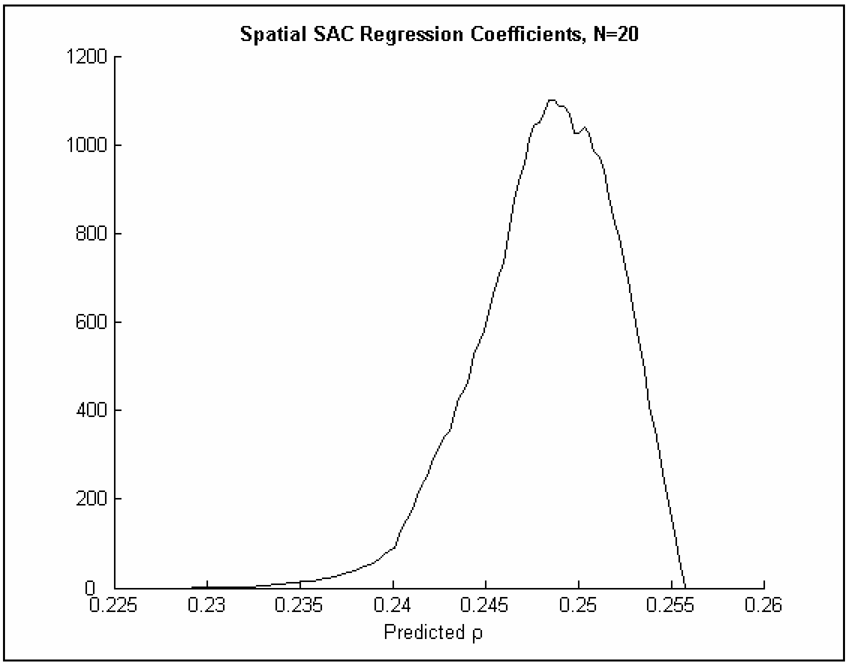

Figure 6

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 A=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.

Table 3 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 3 records the critical value of ρ^* in the bivariate case (k = 2). Note that ρ^* in Table 3 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 4 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 4 are given in Table 5. The SDF statistics in Table 5 are naturally more negative than their univariate counterparts in Table 2 due to the reduction in degrees of freedom.

Table 4 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 5 SDF Cointegration Test Statistics (k > 2)

K	**N = 5**	**N = 20**
3	-1.70	-1.61
4	-1.93	-1.77

p = 0.05

6. Conclusions

We report for the first time critical values for spatial unit root and spatial cointegration test statistics in spatial cross-section data. For example, if there are 100 spatial units the critical value of the normalized SAR coefficient is 0.9 for a spatial unit root. Our impression is that estimated SAR coefficients are typically smaller than these critical values in which case spatial cross-section data are typically stationary. If, however, spatial data happen to be nonstationary and they are regressed upon each other the critical value for the normalized SAC coefficient of the residuals is 0.87. If the estimated SAC coefficient is less than 0.87 the variables are spatially cointegrated and the estimated relationship between the variables is genuine. If, however, the estimated SAC coefficient exceeds 0.87, the estimated relationship is nonsense or spurious.

The concept of cointegration used here follows the residual-based tests of Engle and Granger (1987). A number of alternative cointegration tests have been developed for time series data, including Johansen's vector error correction methodology (VECM). This kind of test framework is particularly suitable for situations in which the observation period is insufficiently long so that the long-run relationship in the data, if it exists, is harder to detect. It is difficult to find a parallel in spatial cross-section data to VECM. Whereas time series are always partially observed because historic data are typically lacking and the future is not observable, spatial cross section data are usually completely observed. Spatial samples such as NUTS 2 are well defined. Therefore, the problem of unrepresentative observation periods in time series data does not have a parallel in spatial cross section data. This makes residual-based cointegration tests more appropriate to spatial cross sections than to time series data.

Matters would be different in spatial panel data because the observation period in such data may be too short to detect the long-run hypotheses that are to be tested. We leave the issue of spatial and temporal cointegration in spatial panel data to another occasion.

References

Baltagi B.H., G. Bresson and A. Pirotte (2007) Panel Unit Root Tests and Spatial Dependence, *Journal of Applied Econometrics*,

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.

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 Coinegration and Heteroscedasticity, *Journal of Geographical Systems*, 9: 253-265.

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.