A test strategy for spurious spatial regression, spatial nonstationarity, and spatial cointegration

Jørgen Lauridsen[1], **Reinhold Kosfeld**[2]

[1] Department of Business and Economics, University of Southern Denmark, Campusvej 55, DK–5230 Odense M, Denmark (e-mail: jtl@sam.sdu.dk)
[2] Department of Economics, University of Kassel, D-34109 Kassel, Germany (e-mail: kosfeld@wirtschaft.uni-kassel.de)

Received: 27 November 2002 / Accepted: 16 December 2006

Abstract. A test strategy consisting of a two-step Lagrange Multiplier test is suggested as a device to reveal spatial nonstationarity and spurious spatial regression. It is further illustrated how the test strategy can be used as a diagnostic for presence of a spatial cointegrating relationship between two variables. Using Monte Carlo simulations it is shown that the small-sample behaviour of the test strategy is as desired in these cases.

JEL classification: C21, C40, C51, J60,

Key words: Spatial nonstationarity, spatial cointegration, spurious regression

1 Introduction

Spatial regression has been discussed widely in books dedicated to developments in spatial econometrics, notably by Anselin (1988a), Anselin and Florax (1995), Griffith (2003), and Anselin et al. (2004). The consequenses for estimation and inference in the presence of stable spatial processes have been extensively investigated (Bivand 1980; Richardson and Hèmon 1981; Clifford and Richardson 1985; Clifford et al. 1989; Anselin 1988a; Haining 2003; Richardson 1990). A recent study (Fingleton 1999) makes the first steps in analysing the implications of spatial unit roots, spatial cointegration and spatial error correction models. A follow-up to this study is Mur and Trivez (2003), who develop the concept of spurious spatial regression in a framework of spatial trend (non)stationarity. Lauridsen (2006) investigates the estimation of spatial error-correction models using an IV approach.

© 2006 the author(s). Journal compilation © 2006 RSAI. Published by Blackwell Publishing, 9600 Garsington Road, Oxford OX4 2DQ, UK and 350 Main Street, Malden MA 02148, USA.

The topics studied in the present article may be viewed as generalisations of common topics studied in the general time series literature. For example, two survey papers on the subject of unit roots in economic time series data, Diebold and Nerlove (1990) and Campbell and Perron (1991), cite over 200 basic references to the subject. The literature on unit roots and cointegration is one of the most rapidly moving targets in econometrics. Stock's (1994) survey adds hundreds of references to those in the aforementioned surveys and brings the literature up to date. Useful basic references on the subjects are: Geweke (1984), Hendry et al. (1984), Judge et al. (1985), Harvey (1989, 1990), Mills (1990), Box et al. (1994), Hamilton (1994), Enders (1995), Granger and Newbold (1996), Granger and Watson (1984), and Patterson (2000).

The present article refines the suggestions of Fingleton (1999). Specifically, Fingleton suggests that "very high" values of the Moran test for spatial residual autocorrelation indicate spatial nonstationarity and spurious regression. It is, however, left as an open question how to distinguish between stationary positive autocorrelation and nonstationarity. The present investigation shows that a two-step LM test for positive residual autocorrelation can provide a better-founded basis to separate these two cases. It is further shown that the same procedure works as a diagnostic for spurious regression. Next, it is suggested that the test procedure works well as a test for spatial cointegration, using a specific two-variable data generating process. In all cases, the small-sample properties of the suggested procedures are derived using Monte Carlo simulation. It is concluded that the procedure works well in all cases, even for fairly small sample sizes.

2 Models with spatial dynamics

2.1 The regressive, spatially autoregressive model

The first order spatially autoregressive model or SAR(1) model was initially studied by Whittle (1954) and has been used extensively in works by Ord (1975), Cliff and Ord (1981), Ripley (1981), Upton and Fingleton (1985), Anselin (1988a), Haining (2003), Griffith (1992), Anselin et al. (1996), Florax et al. (2003), and Lauridsen (2006). For applied research the SAR(1) model is extended by explanatory variables (see the more recent references mentioned above). The regressive, spatially autoregressive model or SARX(1) model is given by:

$$\mathbf{y} = \rho \mathbf{W}\mathbf{y} + \mathbf{X}\boldsymbol{\beta} + \mathbf{v}, \tag{1}$$

in which $\mathbf{y}$ is an $n \times 1$ vector, $\mathbf{X}$ an $n \times K$ matrix of explanatory variables, ρ the autoregressive parameter, $\mathbf{I}$ the $n \times n$ identity matrix and $\mathbf{v}$ an $n \times 1$ vector of independently normally distributed errors with zero expectation and variances σ^2, that is $\mathbf{v} \sim \mathrm{N}(\mathbf{0}, \sigma^2\mathbf{I})$, and $\mathbf{W}$ denotes an $n \times n$ spatial weight matrix. The weights matrix is obtained by row-standardisation of the $n \times n$ contiguity matrix $\mathbf{W^*}$ which is defined by $w_{ij}^* = 1$ if observation j is assumed to impact observation i, and $w_{ij}^* = 0$

otherwise, so that $w_{ij} = w^*_{ij} / \Sigma_{j=1,\ldots,n} w^*_{ij}$. For alternative specifications of the spatial weight matrix, see for instance Cliff and Ord (1981) and Anselin (1988a). The weights may be multidirectional, which is not the case for the time-series case where $w_{ij} = 1$ if $j = i - 1$, for $i = 2, 3, \ldots, n$. For the general spatial case, the weights matrix is generally multidirectional. As proved by Anselin (1988a), multidirectionality of the weights matrix renders OLS estimation of the parameters inconsistent. Finally, for the general case, ρ is restricted to the interval between –1 and +1 and may thus take on positive as well as negative values. Although meriting interest in itself, the negative case is conceptually different from the usual positive case. We thus narrow our focus in the present analysis to the common case where ρ is positive.

2.2 *Spurious regression and nonstationarity*

If **y** and one or more of the **x** variables are generated according to SAR schemes with positive autoregressive parameters and **y** is regressed on **X**:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}, \tag{2}$$

with $\boldsymbol{\varepsilon}$ as the error term, a risk of spurious regression occurs. Especially, in the case of spatial near nonstationarity, where **y** and one or more of the **x** variables have autoregressive parameters close to 1, the risk of spurious regression is alarmingly high. It manifests in the OLS residuals **e** of the regression tending to be highly spatially autocorrelated. This is demonstrated in Fingleton (1999), where extremely high values of the test statistics of the Moran test for spatial autocorrelation (Whittle 1954; Anselin 1988a) have been found. In this setting high values of Moran's I can be viewed as the counterpart of low values of the Durbin-Watson statistic, which is common in spurious time-series regression. In both cases the behaviour of the test statistics is used as an indication of nonstationarity.

The stochastic process generating the OLS residual vector **e** in equation (1) usually has to be inferred from inspecting the behaviour of the residuals. Fingleton (1999) leaves it as an open question how to separate the case of stationary positive autocorrelation ($0 < \rho < 1$) from the nonstationarity case ($\rho = 1$). This means that the implicitly assumed error process:

$$\boldsymbol{\varepsilon} = \rho_\varepsilon \mathbf{W}\boldsymbol{\varepsilon} + \boldsymbol{\mu}, \ \boldsymbol{\mu} \sim \mathrm{N}\left(\mathbf{0}, \sigma^2\mathbf{I}\right), \tag{3}$$

is considered with $\rho_\varepsilon = 0$ under the null hypothesis of independently and identically distributed (i.i.d.) disturbances, and with $\rho_\varepsilon > 0$ under the alternative hypothesis of spatially autocorrelated errors. In general, the error process (3) is not the spatial analogue of the Markov process underlying the Durbin-Watson test in time series analysis (see Haining 2003, p. 299). In both cases, though, only first-order error autocorrelation is taken into account as the alternative. Note that spatial autocorrelation can be caused by both a spatially autoregressive SAR(1) and a spatial moving average or SMA(1) process (see, e.g., Kelejian and Robinson 1995;

Hepple 1995a,b). However, manifestation of spatial nonstationarity can only be attributed to a SAR process. Moreover, Fingleton (1999) does not address the well-known power of the Moran *I* test towards misspecifications in the form of, for instance, spatial heterogeneity (see Anselin 1988a). Being an advantage in some circumstances, this feature of the Moran *I* coefficient is not necessarily an advantage when investigating specific features of the data generating processes underlying the model that is being considered.

In order to account for both shortcomings, the present study suggests a two-step Lagrange Multiplier test for spatially autocorrelated errors. The LM error statistic (LME) developed in Anselin (1988a,b):

$$\mathrm{LME} = \left(\mathbf{e}'\mathbf{W}\mathbf{e}/\sigma^2\right)^2 \Big/ \mathrm{tr}\left(\mathbf{W}^2 + \mathbf{W}'\mathbf{W}\right), \tag{4}$$

is asymptotically χ^2 distributed with 1 degree of freedom under H_0: $\rho_\varepsilon = 0$. Therefore, a large LME value indicates either spatial nonstationarity or stationary spatial error autocorrelation. This result corresponds to the suggestions of Fingleton (1999) with the Moran's *I* test replacing the LM test. Next, under spatial nonstationarity, $\rho_\varepsilon = 1$,

$$\Delta\boldsymbol{\varepsilon} = \boldsymbol{\mu} \tag{5}$$

follows from the spatial error process (3) with $\boldsymbol{\Delta} = \mathbf{I} - \mathbf{W}$ as the spatial difference operator. By employing $\boldsymbol{\Delta}$ to equation (2), the transformed regression equation:

$$\Delta\mathbf{y} = \Delta\mathbf{X}\boldsymbol{\beta} + \boldsymbol{\mu} \tag{6}$$

is obtained. Equation (6) implies that a regression of $\boldsymbol{\Delta}\mathbf{y}$ on $\boldsymbol{\Delta}\mathbf{X}$ provides i.i.d. errors, so that the LM error test statistic for this spatially differenced model (DLME) will be close to zero. On the other hand, if the null of nonstationarity, H_0: $\rho_\varepsilon = 1$. In the spatially differenced model (6) does not hold, then the spatial differencing will bring about an error term on the form:

$$\Delta\boldsymbol{\varepsilon} = (\mathbf{I} - \mathbf{W})(\mathbf{I} - \rho_\varepsilon\mathbf{W})^{-1}\boldsymbol{\mu} \Leftrightarrow \boldsymbol{\mu} = (\mathbf{I} - \rho_\varepsilon\mathbf{W})\boldsymbol{\varepsilon}. \tag{7}$$

The spatially autocorrelated errors resulting from spatial "overdifferencing" are expected to go along with a positive DLME value. In sum, the test strategy consists of calculating and inspecting the LME and the DLME values, and will lead to one of the four conclusions detailed in Table 1. In Table 1, a test result is designated

Table 1. Outcomes of two-stage LM tests

	DLME zero	DLME positive
LME zero	—	Absence of spatial autocorrelation
LME positive	Spatial nonstationarity (spurious regression)	Stationary spatial autocorrelation

"positive" if the LM test statistic differs significantly from zero, and it is referred to as "zero" otherwise.

As a specific detail, one should be aware that the conclusion from the DLME test is conditioned on the conclusion from the LME test. Thus, a further precision of the two-step procedure might be obtained by incorporating for example a Bonferroni type correction to the level of the combined test procedure, although it should be kept in mind that this correction would be too conservative due to the correlation among the LME and the DLME tests.

It may also be relevant to investigate whether $\mathbf{y}$ or any of the $\mathbf{x}$ variables are spatially nonstationary. This may be revealed by using the suggested procedure for a regression of the variable in question (i.e., $\mathbf{z}$ being one of $\mathbf{y}$, $\mathbf{x}_1$, $\mathbf{x}_2$, . . .) on a constant term. Specifically, the regressions:

$$\mathbf{z} = \alpha\mathbf{i} + \boldsymbol{\varepsilon}, \tag{8}$$

and

$$\Delta\mathbf{z} = \alpha\Delta\mathbf{i} + \Delta\boldsymbol{\varepsilon} = \Delta\boldsymbol{\varepsilon}, \tag{9}$$

easily provide the LME and DLME test statistics, which lead to one of three conclusions: $\mathbf{z}$ is spatially nonstationary (LME positive, DLME zero); $\mathbf{z}$ represents a stationary SAR scheme (LME positive, DLME positive); or $\mathbf{z}$ is free of any spatial pattern (LME zero, DLME positive). According to the data generating process $\mathbf{z} = \rho\mathbf{Wz} + \mathbf{v}$, the $\mathbf{z}$ variables are spatially integrated of order one, SI(1), in the case of nonstationarity. It should be kept in mind that this relies on the data generating process being well specified. For example, the difference stationary spatial random walk with drift defined by $\mathbf{z} = \mu\mathbf{i} + \mathbf{Wz} + \mathbf{v}$ and the spatial trend stationary process $\mathbf{z} = \mu\mathbf{i} + \mathbf{C}\boldsymbol{\beta} + \mathbf{v}$, where $\mathbf{C}$ is a matrix of coordinates for the spatial units, are both nonstationary but may be diagnosed as stationary.

An intuitively appealing alternative to the LM test procedure suggested might be to estimate the SAR model and test the hypothesis $\rho = 1$ using a Wald or Likelihood Ratio test. At least two objections may be raised against this proposal. First, efficient estimation of the SAR model requires maximum likelihood estimation, which is in principle doable although hardly practical for simulation studies aimed at deriving finite-sample properties of the test. Second, the proposal resembles the Dickey-Fuller approach applied to the time series case. Even for this special case, it is known that $(1 - \rho)/\sigma_\rho$ does not adhere to a standard normal or t distribution. This disclaimer may well pertain to the application of the Moran I test, which can be viewed as a generalisation of the Dickey-Fuller test when applied to the differenced model.

A further advantage of the LM test strategy is that it is quite flexible. Thus, it is possible to control for omitted model features insofar that these can be incorporated as part of the likelihood function. For example, it is straightforward to account for omitted heterogeneity and an omitted autoregression in the dependent variable, along the lines suggested in Anselin (1988b).

2.3 Spatial cointegration

Spatial cointegration denotes the case where two or more variables in a regression are nonstationary, while the errors are stationary. Let **x** and **y** be both spatially integrated of order one. Then in general any linear combination of **x** and **y** is also SI(1). If, however, a linear combination $\mathbf{y} - \beta\mathbf{x}$ exists which is stationary, **x** and **y** are said to be spatially cointegrated. In this case the cointegrating vector is given by $(1 - \beta)$. The linear combination $\mathbf{y} - \beta\mathbf{x}$ which renders the errors in a regression setting is then spatially integrated of order zero, SI(0) (Fingleton 1999).

A simple data generating process which generates two nonstationary but possibly cointegrating series is the following system, which is a spatial generalisation of a time series specification presented in Banerjee et al. (1993):

$$\mathbf{x} + \beta\mathbf{y} = \mathbf{u}, \mathbf{u} = \mathbf{W}\mathbf{u} + \mathbf{e}_1, \tag{10}$$

$$\mathbf{x} + \alpha\mathbf{y} = \mathbf{e}_2, \tag{11}$$

where $\mathbf{e}_1$ and $\mathbf{e}_2$ are white noise processes. From these definitions:

$$\mathbf{x} = \alpha(\alpha - \beta)^{-1}\mathbf{u} - \beta(\alpha - \beta)^{-1}\mathbf{e}_2, \tag{12}$$

$$\mathbf{y} = -(\alpha - \beta)^{-1}\mathbf{u} + (\alpha - \beta)^{-1}\mathbf{e}_2, \tag{13}$$

from which it is clear that **x** and **y** are SI(1) but that they are cointegrated for any α different from 0 and certain β values, because $\mathbf{x} + \alpha\mathbf{y}$ is I(0). Specifically, the relation will be non-integrated if (i) $\alpha = 0$ or (ii) $\alpha > 0$ and $\beta > \alpha$.

We suggest that the above LM strategy may apply to this situation. Specifically, a regression of **y** on **X** represents a cointegrating relation if LME is zero and DLME is positive, or a non-cointegrating relation if LME is positive and DLME is zero. The limiting case of "near cointegration" ($\alpha > 0$, $\beta > \alpha$) will also be revealed, specifically if LME and DLME are positive.

3 Monte Carlo simulation studies: designs and results

In this section, the finite-sample properties of the above suggested test strategies will be investigated using Monte Carlo simulation studies. The chosen Monte Carlo designs are outlined together with the results. All calculations are done using SAS/IML, including the software's standard routines for random number generation.

3.1 Spurious regression

To investigate the finite-sample properties of the suggested LME test strategy for spurious regression, the following Monte Carlo design is investigated:

For a specific sample size n, we perform 10,000 iterations where $\mathbf{e_x}$ and $\mathbf{e_y}$ are generated as independent N(0,1) series, $\mathbf{x} = (\mathbf{I} - \rho_x\mathbf{W})^{-1}\mathbf{e_x}$, $\mathbf{y} = (\mathbf{I} - \rho_y\mathbf{W})^{-1}\mathbf{e_y}$, and we regress $\mathbf{y}$ on $\mathbf{X} = [\mathbf{i}\ \mathbf{x}]$ and $\Delta\mathbf{y}$ on $\Delta\mathbf{X}$ to obtain LME and DLME. We report the percentage of cases out of 10,000 where LME, respective DLME exceeds the 5 percent critical value of $\chi^2(1) = 3.84$. The parameters ρ_x and ρ_y are varied over a grid of the values (0.0, 0.1, 0.2, . . . , 0.8, 0.9, 0.99, 1.00). To investigate the impact of the contiguity matrix type, we make use of the rook and queen type of regular contiguity matrices based on an $r \times r$ checker board (so that $n = r^2$) with r assumed to take on the values 5, 10, 15, and 20. The rook matrix represents a square tesselation with a connectivity of 4 for the inner fields on the chessboard and 2 and 3 for the corner and border fields, respectively. The queen matrix represents an octogonal tesselation with a connectivity of 8 for the inner fields and 3 and 5 for the corner and border fields. Thus, these tesselations represent extremes for a number of patterns, including the hexagonal tesselation, which is of importance due to its application for empirical maps in vector and raster based GIS (Boots and Tiefelsdorf 2000; Tiefelsdorf 2000). Actually, the hexagonal tesselation can be constructed from the queen tesselation by deleting connections from any field to the fields vertically above and below this. Moreover, most empirically observed regional structures in spatial econometrics are made up of regions with connectivity within the range of the rook and queen tesselations. Further, irregular matrices based on the 275 Danish municipalities are applied: a $n = 36$ matrix based on the municipalities located on the island of Funen, a $n = 97$ matrix made up of the municipalities located on Seeland together with the adjacent islands Lolland and Falster, a $n = 141$ matrix created from the municipalities located on the peninsula of Jutland, and the full matrix of $n = 275$ Danish municipalities, which consists of the above municipalities plus 5 municipalities located on the island of Bornholm. The map of the 275 municipalities, together with the above partitioning, is shown in Figure 1.

The behaviour of the strategy under spatial nonstationarity as well as stationarity (including the case of near nonstationarity) is investigated by assuming ρ_y to take on the values (0.0, 0.1, 0.2, . . . , 0.8, 0.9, 0.99, 1.00). For each of these, ρ_x is assumed to take on the values (0.0, 0.1, 0.2, . . . , 0.8, 0.9, 0.99, 1.00). For the cases of nonstationarity, we use the Moore-Penrose generalised inverse $(\mathbf{I} - \mathbf{W})^+$ instead of $(\mathbf{I} - \mathbf{W})^{-1}$.

Figures 2 and 3 show that the procedure performs well, and that the performance of the procedure is acceptable even for fairly small sample sizes, and for regular as well as irregular matrices. For very small samples, the results for the $n = 25$ regular matrices and the results for the $n = 36$ irregular matrix show that the power of the tests are not convincing, as the increase in the power function for the LME and the DLME tests is relatively slow. This is especially apparent for the queen case. Generally, the size of the tests are as desired, being close to the true value of 0.05. That the case of near nonstationarity causes problems in identifying the "true" data generating process is well known from time series analysis. The performance of the procedure seems to be unaffected by the type of contiguity matrix, as the regular rook and queen and the irregular cases provide similar results. It should be noticed that this observation does not guarantee robustness of

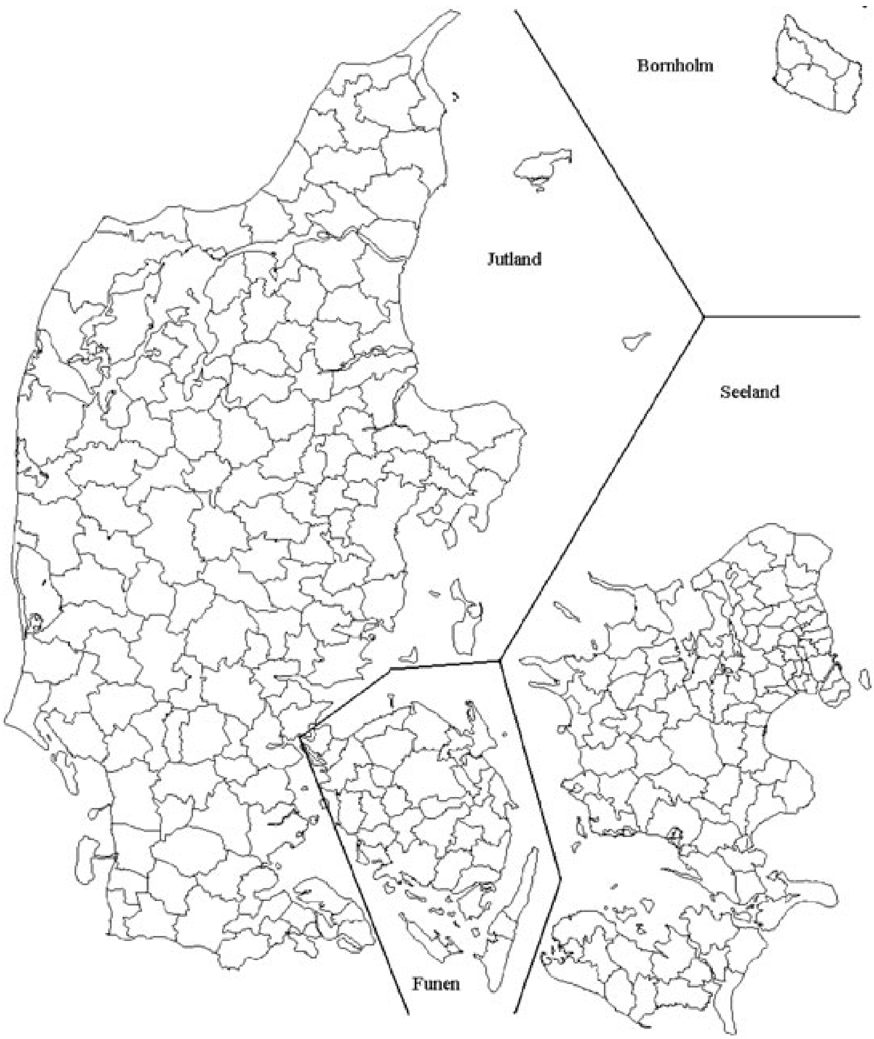

Fig. 1. The Danish municipalities

the test procedure against misspecification of the contiguity matrix. Such misspecifications might occur in, for example, geostatistical studies employing an automatically generated contiguity structure from raster based GIS and remote sensing. Further, it is noticeable that only the SAR process in the **y** variable matters, while the power function for any value of ρ_x seems to be similar.

3.2 Test for nonstationarity

To investigate the finite-sample properties of the suggested LME test strategy for nonstationaity of a single variable, the following Monte Carlo design is investi-

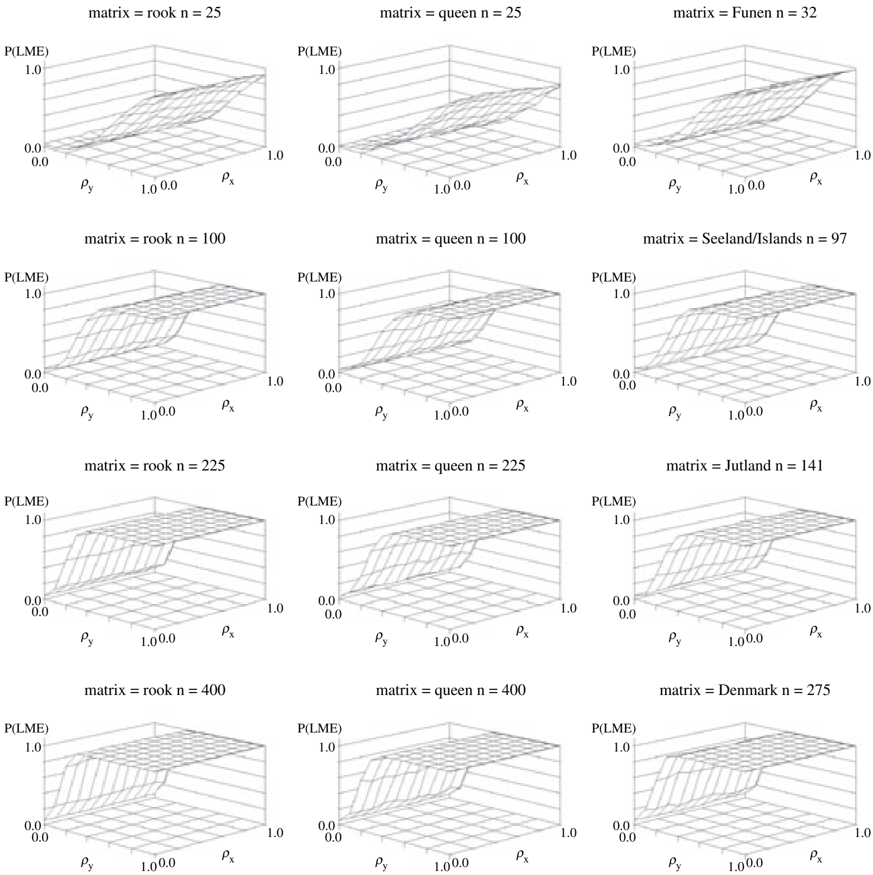

Fig. 2. Monte Carlo results for spurious regression; proportion of cases where LME value rejects H_0 at the 5 percent level

gated. For a specific sample size n, we perform 10,000 iterations, where $\mathbf{e}$ is generated as independent N(0,1) series, $\mathbf{y} = (\mathbf{I} - \rho\mathbf{W})^{-1}\mathbf{e}$, and we regress $\mathbf{y}$ on $\mathbf{X} = \mathbf{i}$ and $\mathbf{\Delta y}$ on $\mathbf{\Delta X}$ in order to obtain LME and DLME results. We report the percentage of cases out of 10,000 where LME, respective DLME exceeds the 5 percent critical value of $\chi^2(1) = 3.84$. Again, we use the rook and queen type contiguity matrices based on an $r \times r$ board with r assumed to take the values 5, 10, 15, and 20 and the irregular empirical matrices. Further, the behaviour of the strategy under spatial: nonstationarity as well as stationarity (including the case of near nonstationarity) is investigated by varying ρ across the values (0.0, 0.1, 0.2, . . . , 0.8, 0.9, 0.99, 1.0). For the cases where ρ equals 1, we replace $(\mathbf{I} - \mathbf{W})^{-1}$ by the Moore-Penrose generalised inverse $(\mathbf{I} - \mathbf{W})^{+}$.

Figures 4 and 5 show that the procedure performs well even for fairly small sample sizes, and for the regular as well as the irregular matrices. For very small samples, the results for the $n = 25$ regular cases and the $n = 36$ irregular case show

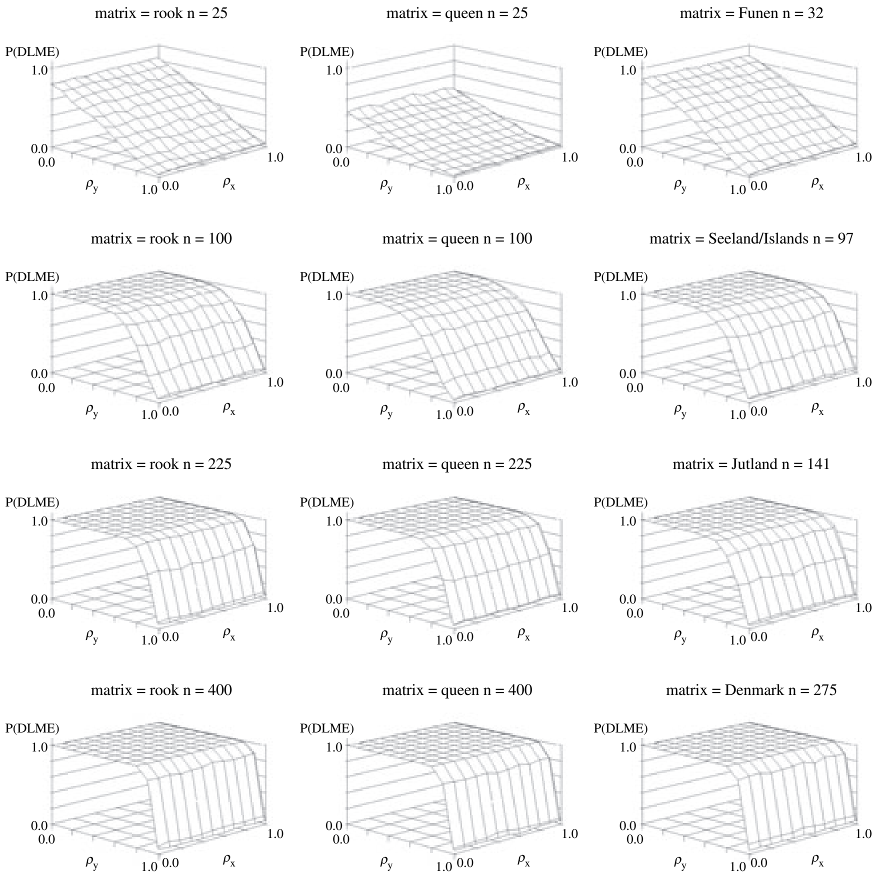

Fig. 3. Monte Carlo results for spurious regression; proportion of cases where DLME value rejects H_0 at the 5 percent level

that the power of the tests is very flat and thus not convincing. This holds true under the assumption of nonstationarity as well as different stationarity cases. The case of near nonstationarity is again included for comparative purposes. Note that performance is independent of the type of contiguity matrix.

3.3 Test for cointegration

To investigate the finite-sample properties of the suggested LME test strategy for cointegration using the suggested example, the following Monte Carlo design is investigated. For a specific sample size *n*, we perform 10,000 iterations, where we generate $\mathbf{e}_1$, $\mathbf{e}_2$ as independent N(0,1) series, $\mathbf{u} = (\mathbf{I} - \mathbf{W})^{+}\mathbf{e}_1$, $\mathbf{x} = \alpha(\alpha - \beta)^{-1}\mathbf{u} - \beta(\alpha - \beta)^{-1}\mathbf{e}_2$ and $\mathbf{y} = -(\alpha - \beta)^{-1}\mathbf{u} + (\alpha - \beta)^{-1}\mathbf{e}_2$, and we regress $\mathbf{y}$ on $\mathbf{X} = [\mathbf{i}\ \mathbf{x}]$ and $\Delta\mathbf{y}$ on $\Delta\mathbf{X}$, in order to obtain LME and DLME results. We again

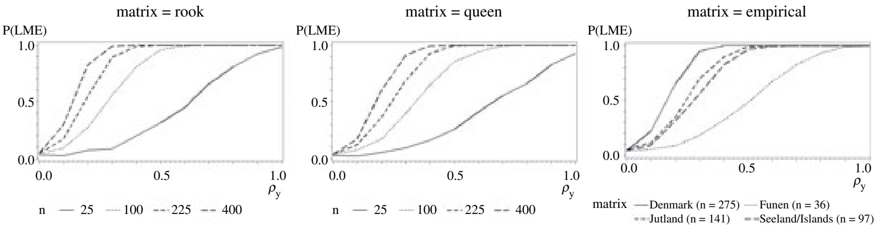

Fig. 4. Monte Carlo results for nonstationary variables; proportion of cases where LME value rejects H_0 at the 5 percent level

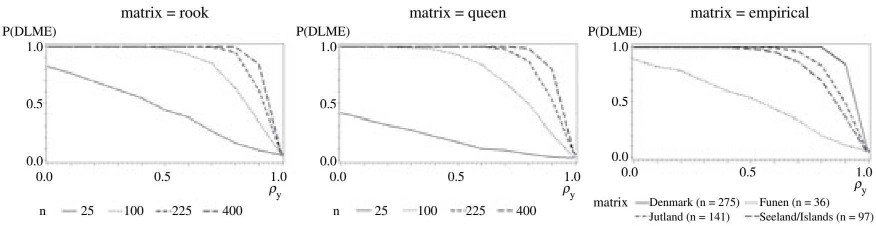

Fig. 5. Monte Carlo results for nonstationary variables; proportion of cases where DLME value rejects H_0 at the 5 percent level

report the percentage of cases out of 10,000 where LME, respective DLME exceeds the 5 percent critical value of $\chi^2(1) = 3.84$. In order to investigate the impact of contiguity matrix type, we again use the regular rook and queen type contiguity matrices based on an $r \times r$ board with r assumed to take the values 5, 10, 15, and 20 and the irregular matrices based on the Danish case. Further, the behaviour of the strategy under spatial nonstationarity as well as stationarity (including the case of near nonstationarity) is investigated for varying α and β. Specifically, α was varied across the values (0, 0.1, 0.2, . . . , 0.8, 0.9, 1.0). For each of these, β was varied across the same values, with an exception for the cases when $\alpha = \beta$. For these, β was set to $(\alpha + 0.01)$, except for the $\alpha = 1.0$ cases, where β was set to 0.99. The results are shown in Figures 6 and 7.

Figures 6 and 7 show that the procedure performs well, especially for fairly large n, and that the performance of the procedure is acceptable, even for fairly small sample sizes. For very small samples, the results for the $n = 25$ regular cases and the $n = 36$ irregular case show that the power of the test is not convincing. Especially, in the case of cointegration ($\alpha = 1$) and non-integration ($\alpha = 0$), the procedure works excellently, while the grey area case of near-integration ($0 < \alpha < 1$, $\beta > \alpha$) is characterised by inconclusive test sizes. These conclusions hold for both types of contiguity matrices, with a single exception for the cases where α and β are close to each other. In such cases, the rejection percentages are much higher for the queen than for the rook case.

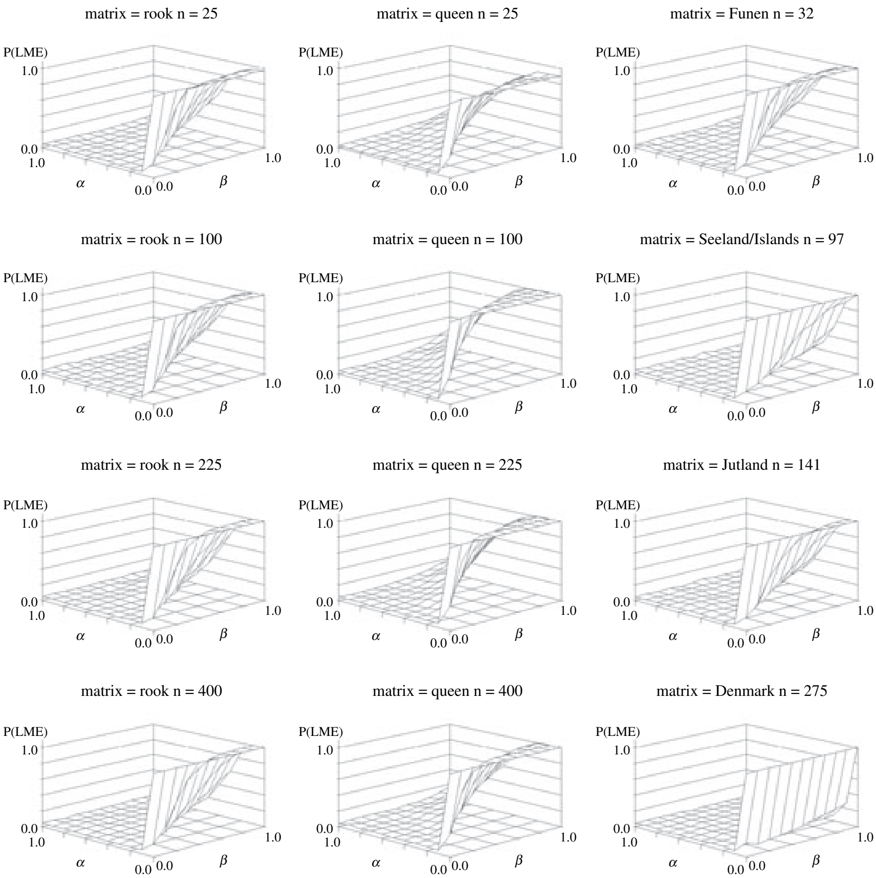

Fig. 6. Monte Carlo results for nonstationary cointegration; proportion of cases where LME value rejects H_0 at the 5 percent level

4 Conclusion

Until now, it has not been well established how to separate the case of spatial nonstationarity from the case of stationary positive autocorrelation. As a consequence, reliable diagnostics for spurious spatial regression and for the existence of spatial cointegrating relations have not been available. The present study aims to contribute to closing these gaps by proposing a strategy for detecting spatial nonstationarity. It is shown that the test strategy consisting of a two-step Lagrange Multiplier test provides adequate diagnostics for both spurious spatial regression and the presence of spatial cointegrating relations. By means of Monte Carlo simulations it is demonstrated that the finite sample properties of the suggested methodology are as desired.

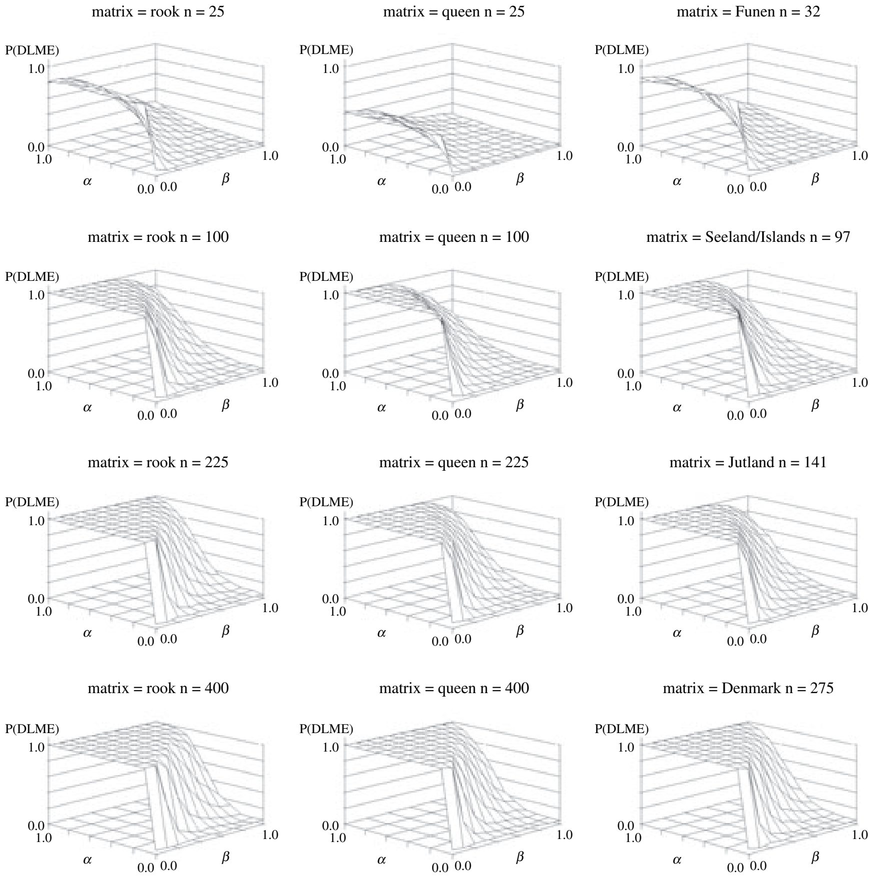

Fig. 7. Monte Carlo results for nonstationary cointegration; proportion of cases where DLME value rejects H_0 at the 5 percent level

References

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

Anselin L (1988b) Lagrange Multiplier Tests Diagnostics for Spatial Dependence and Spatial Heterogeneity. *Geographical Analysis* 20: 1–17

Anselin L, Bera A, Florax RJGM, Yoon M (1996) Simple Diagnostic Tests for Spatial Dependence. *Regional Science and Urban Economics* 26: 77–104

Anselin L, Florax RJGM (1995) *New Directions in Spatial Econometrics*. Springer-Verlag, Berlin

Anselin L, Florax RJGM, Rey SJ (2004) *Advances in Spatial Econometrics*. Springer-Verlag, Berlin

Banerjee A, Dolado J, Galbraith JW, Hendry DF (1993) *Co-Integration, Error-Correction, and the Econometric Analysis of Non-Stationary Data*. Oxford University Press, NY

Bivand R (1980) A Monte-Carlo Study of Correlation Coefficient Estimation with Spatially Autocorrelated Observations. *Quaestiones Geographicae* 6: 5–10

Box G, Jenkins G, Reinsel GC (1994) *Time-Series Analysis: Forecasting and Control*, 3rd ed. Holden Day, San Francisco

Boots BN, Tiefelsdorf M (2000) Global and Local Spatial Autocorrelation in Bounded Regular Tesselations. *Journal of Geographical Systems* 2: 319–348

Campbell J, Perron P (1991) *Pitfalls and Opportunities: What Macroeconomists Should Know About Unit Roots*. National Bureau of Economic Research, Macroeconomics Conference, Cambridge, Massachusetts

Cliff AD, Ord JK (1981) *Spatial Processes: Models and Applications*. Pion, London

Clifford P, Richardson S (1985) Testing the Association between Two Spatial Processes. *Statistics and Decisions*, Supplement no. 2: 155–160

Clifford P, Richardson S, Hémon D (1989) Assessing the Significance of the Correlation between Two Spatial Processes. *Biometrics* 45: 123–134

Diebold F, Nerlove M (1990) Unit Roots in Economic Time Series: A Selective Survey. In: Bewley T (ed) *Advances in Econometrics*, Vol. 8. JAI Press, NY: 3–70

Enders W (1995) *Applied Econometric Time Series*. John Wiley and Sons, NY

Fingleton B (1999) Spurious Spatial Regression: Some Monte Carlo Results with a Spatial Unit Root and Spatial Cointegration. *Journal of Regional Science* 39: 1–19

Florax RJGM, Folmer H, Rey SJ (2003) Specification Searches in Spatial Econometrics: The Relevance of Hendry's Methodology. *Regional Science and Urban Economics* 33: 557–579

Geweke J (1984) Inference and Causality in Econometric Time Series Models. In: Griliches Z, Intriligator M (eds) *Handbook of Econometrics*, Vol. 2. North Holland, Amsterdam: 1101–1144

Granger C, Newbold P (1996) *Forecasting Economic Time Series*. 2nd Ed. Academic Press, NY

Granger C, Watson M (1984) Time Series and Spectral Methods in Econometrics. In: Griliches Z, Intriligator M (eds) *Handbook of Econometrics*, Vol. 2. North Holland, Amsterdam: 979–1022

Griffith DA (1992) Simplifying the Normalizing Factor in Spatial Autoregressions for Irregular Lattices. *Papers in Regional Science* 71: 71–86.

Griffith DA (2003) *Spatial Autocorrelation and Spatial Filtering*. Springer-Verlag, Berlin

Haining RP (2003) *Spatial Data Analysis: Theory and Practice*. Cambridge University Press, Cambridge, UK

Hamilton J (1994) *Time Series Analysis*. Princeton University Press, Princeton

Harvey A (1989) *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge University Press, NY

Harvey A (1990) *The Econometric Analysis of Time Series*. 2nd Ed. MIT Press, Cambridge

Hendry D, Pagan A, Sargan J (1984) Dynamic Specification. In: Griliches Z, Intriligator M (eds) *Handbook of Econometrics*, Vol. 2. North Holland, Amsterdam: 1231–1303

Hepple LW (1995a) Bayesian Techniques in Spatial and Network Econometrics 1. Model Comparison and Posterior Odds. *Environment and Planning A* 27: 447–469

Hepple LW (1995b) Bayesian Techniques in Spatial and Network Econometrics 2. Computational Methods and Algorithms. *Environment and Planning A* 27: 615–644

Judge G, Griffiths W, Hill C, Lee T (1985) *The Theory and Practice of Econometrics*. John Wiley and Sons, NY

Kelejian HH, Robinson DP (1995) Spatial Correlation: A Suggested Alternative to the Autoregressive Model. In: Anselin L, Florax RJGM (eds) *New Directions in Spatial Econometrics*. Springer-Verlag, Berlin: 75–95

Lauridsen J (2006) Spatial Autoregressively Distributed Lag Models: Equivalent Forms, Estimation, and an Illustrative Commuting Model. *Annals of Regional Science* (forthcoming)

Mills T (1990) *Time Series Techniques for Economists*. Cambridge University Press, NY

Mur J, Trívez FJ (2003) Unit Roots and Deterministic Trends in Spatial Econometric Models. *International Regional Science Review* 26: 289–312

Ord JK (1975) Estimation Methods for Models of Spatial Interaction. *Journal of the American Statistical Association* 70: 120–126

Patterson K (2000) *An Introduction to Applied Econometrics*. St. Martin's Press, NY

Richardson S (1990) Some Remarks on the Testing of Association between Spatial Processes. In: Griffith DA (ed) *Spatial P* *ed.slF* *F*

Stock J (1994) Unit Roots, Structural Breaks, and Trends. In: Engle R, McFadden D (eds) *Handbook of Econometrics*, Vol. 4. North Holland, Amsterdam: 2739–2841

Tiefelsdorf M (2000) *Modelling Spatial Processes: The Identification and Analysis of Spatial Relationships in Regression Residuals by Means of Moran's I*. Springer-Verlag, Berlin

Upton GJG, Fingleton B (1985) *Spatial Data Analysis by Example*, Vol. 1. John Wiley and Sons, NY

Whittle P (1954) On Stationary Processes in the Plane. *Biometrika* 41: 434–449