Developments in the Analysis of Spatial Data

P. M. Robinson*
London School of Economics
August 31, 2007

The Suntory Centre
Suntory and Toyota International
Centres for Economics and Related
Disciplines
London School of Economics and
Political Science
Houghton Street
London WC2A 2AE
Tel: 020 7955 6679

Discussion paper
No. EM/2009/531
August 2008

* Research supported by ESRC Grant RES-062-23-0036.

Developments in the Analysis of Spatial Data

P. M. Robinson*
London School of Economics

August 31, 2007

Abstract

Disregarding spatial dependence can invalidate methods for analyzing cross-sectional and panel data. We discuss ongoing work on developing methods that allow for, test for, or estimate, spatial dependence. Much of the stress is on nonparametric and semiparametric methods.

1 Introduction

Issues of spatial dependence have arisen in several areas of research, such as the environmental sciences, economics and sociology, but may be more generally relevant in circumstances in which cross-sectional or panel data are collected. Book-length descriptions statistical methods of analyzing spatial data include Cliff and Ord (1981), Anselin (1988), Haining, R. (1990), Cressie (1993), Guyon (1995), Stein (1999), Arbia (2006). The present paper discusses some recent and ongoing developments, mainly from a semiparametric or nonparametric viewpoint.

It is helpful for the purposes of this introductory section to fix ideas by discussing a conventional setting of scalar (for simplicity) observations y_i, $i = 1, ..., n$, having representation

$$y_i = f(x_i) + \varepsilon_i, \quad i = 1, ..., n, \tag{1}$$

where x_i is a $p \times 1$ vector observation that can be deterministically or stochastically generated, f is a parametric, semiparametric or nonparametric function, and the unobservable zero-mean random variable ε_i is uncorrelated with $f(x_i)$. Particular issues arise in connection with the regression component $f(x_i)$, but these are partly due to the properties of the error ε_i and it is the modelling of the ε_i, and its implications for rules of statistical inference, that are central from our perspective.

The estimation of f or parameters that describe it which incorrectly assume independence across the ε_i are well understood in the wider statistical and

*Research supported by ESRC Grant RES-062-23-0036.

econometric literatures, and have been discussed in a specifically spatial context by numerous authors. The index i can be thought of as ranging over a cross-section or a panel, or indeed a time series and/or geographical space. In a cross-sectional setting, independence of the ε_i is often assumed, and can be extended to random-effects models which introduce within-group correlation. The cross-section dimension of a panel data set is often treated similarly, with dependence structure across the time dimension typically described by a dynamic model such as an autoregression (AR) (see e.g. Hsiao (2003)). Here we are concerned with the possibility that data collected like a cross-section or panel, or across geographical space, may have widespread dependence, that varies with some notion of relative location of, or distance between, observations, this might be due to omission of variables.

A basic question is the modelling of the dependence in such a way as to enable justification of rules of large sample inference on aspects of interest. This requires that information accumulate as $n \to \infty$ - at a modest rate for consistency to be achieved, faster for the central limit theorem; neither property can hold if the ε_i have a constant non-zero correlation. Sufficient conditions for such properties in case of time series observations, and observations on d-dimensional space with $d > 1$ (as random fields), have been developed. For stationary ε_i, asymptotic theory for parameters describing f in (1), or the autocovariance structure of the ε_i, is possible under a variety of weak dependence conditions, such as mixing with a suitable rate, or linear filters of independent innovations with summable coefficients, or even under long range dependence. To some degree the extension of theory for the time series case $d = 1$ to $d > 1$ is straightforward but particular features cause difficulty: for multilateral models, least squares (LS) tends to be inconsistent and use of a likelihood approximation is important, as first noted by Whittle (1954); the "edge-effect" is a source of bias in the central limit theorem when $d \geq 2$, and methods of overcoming it are discussed in Guyon (1982), Dahlhaus and Künsch (1987), Robinson and Vidal Sanz (2004). Under long range dependence, limit distributional behaviour may be nonstandard, without weighting of a type used more generally to achieve efficiency (see e.g., Fox and Taqqu (1986), Hidalgo and Robinson (2002)).

Such theory crucially regards the locations, s_i, of the ε_i, as diverging without bound as $i \to \infty$, so that, say, the distance between s_1 and s_n tends to infinity, as $n \to \infty$, and correspondingly ε_1 and ε_n approach independence. This kind of setup is commonly accepted in time series research. It can also seem plausible in some spatial applications in, say, environmental science or astronomy, or even field experiments, but less so when the domain of possible observation is bounded, for physical, historical or administrative reasons. This may not be of major concern in that the practitioner is typically faced with a given, fixed sample, useful asymptotic theory is motivated by approximation rather than extrapolation, and there is an element of artificiality in any regime that caters for an arbitrarily large sample size.

In economics, geographical distance can also be relevant in modelling dependence, but a more general and relevant notion is "economic distance" (see e.g., Conley and Dupor (2003)). Here, unit (economic agent) i is endowed with a vector of characteristics z_i (which may overlap with x_i in (1)). The economic distance between agents i and j is defined by the distance between z_i and z_j, such as the Euclidean norm $\|z_i - z_j\|$ (where there are advantages to parsimonious modelling in assuming isotropy). If z_i has infinite support, so do such distances. Conley (1999) approximated the locations z_i by regularly-spaced (lattice) points and applied mixing conditions in deriving asymptotic theory for certain estimates in a random fields setting (see e.g. also Pinkse, Shen and Slade, 2004). Conley (1999), Conley and Molinari (2004) also examined robustness to stationary point process errors in the lattice approximation. However this kind of assumption on the locations implies a degree of regularity that might not always be plausible. For example, if the z_i are identically distributed the density will be small in tail regions so observations at remote locations will be insufficient to take advantage of weak dependence conditions on ε_i. In spatio-temporal settings where spatial size remains fixed while temporal length increases (for econometric examples see e.g., Chen and Conley (2001), Giacomini and Granger (2004)) this is not a problem because weak dependence conditions over time do the work. However, the time series may be short or non-existent.

In some of the spatial statistics literature, "infill asymptotics" has been developed, assuming increasingly fine observations over a bounded domain as n increases (see e.g. Stein (1991), Cressie (1993), Lahiri (1996)). Typically, a continuous model for dependence across the domain is assumed, but observations form a triangular array when, say, the observations are on a lattice whose mesh decreases as n increases. However, due to the fixed dependence, non-classical asymptotic properties tend to result, e.g. inconsistency due to convergence to a non-degenerate random variable. While these sort of findings are interesting, they are not of much practical use. Useful, standard, asymptotic theory has resulted from analyses in many cross-sectional, panel data and time series problems, and we cannot see persuasive grounds for pursuing a theory under cross-sectional dependence in spatial data that loses this. It seems difficult in general to model the falling off of dependence as $n \to \infty$ in a way that will produce useful asymptotics without a triangular array prescription $\varepsilon_i = \varepsilon_{i,n}$, $1 \leq i \leq n$, with ε_i (even for small i), and thence all elements of the covariance matrix of $\varepsilon_{(n)} = (\varepsilon_1, ..., \varepsilon_n)'$, changing subtly as n increases. Such devices aimed at producing useful distributional approximations are familiar, e.g. in Pitman efficiency theory, structural change modelling, fixed-design nonparametric regression, and local-to-unit roots.

Our discussion of recent and ongoing work on developing statistical methods that allow for, or estimate, spatial dependence, principally in a semiparametric setting, will avoid technical details. Section 2 describes spatial autoregressive (SAR) models, that have become a major feature of spatial statistics and econometrics. Section 3 discusses adaptive estimation of such models, where efficient estimation is possible despite lack of knowledge of innovation distributional form. Section 4 considers nonparametric regression, where f in (1) is of unknown form. Section 5 discusses the testing of spatial independence in data or unobservable errors.

2 Spatial autoregressive models

A very familiar parametric model for spatial data, due to Cliff and Ord (1968), and extensively applied since, is the SAR model. Introduce an $n \times n$ matrix W_n with non-negative elements, where diagonal ones are zero. Strategies for specifying W_n have been discussed in the literature, e.g. its (i, j)-th element $w_{ij,n}$ might be formed as $c_{ij} / \sum_{k=1}^{n} c_{ik}$, where c_{ij} is inversely proportional to some measure of distance between s_i and s_j (perhaps depending on z_i and z_j), though W_n need not be symmetric. A consequence of the ratio specification just introduced is that rows of W_n all sum to 1. This has the advantage of securing a stability property, analogous to that familiar in stationary AR time series models, in the SAR model

$$\varepsilon_{(n)} = \rho W_n \varepsilon_{(n)} + \eta_{(n)}, \tag{2}$$

where ρ is an unknown scalar in $(-1, 1)$ and $\eta_{(n)} = (\eta_1, ..., \eta_n)'$ is a vector of independent and identically distributed random variables. This has been considered in a linear regression setting, where (1) becomes

$$y_{(n)} = X_n \beta + \varepsilon_{(n)}, \tag{3}$$

with $y_{(n)} = (y_1, .., y_n)'$, $X_n = (x_1, ..., x_n)'$, and β an unknown $p \times 1$ vector. A related model that has also been extensively studied, analogous to the dynamic regression model of time series econometrics, is

$$y_{(n)} = \lambda W_n y_{(n)} + X_n \beta + \eta_{(n)}, \tag{4}$$

for unknown scalar $\lambda \in (-1, 1)$.

Asymptotic properties of various estimates of the parameters ρ, λ and β in (2)-(4) have been developed. In particular as noted by e.g. Anselin (1988), under some asymptotic (as $n \to \infty$) conditions on W_n, least squares (LS) estimates of λ and β in (4) are inconsistent, for similar reasons as found by Whittle (1954) in the fixed lattice case. Kelejian and Prucha (1998, 1999) and Lee (2003, 2004) showed other estimates such as instrumental variables (IV) and (Gaussian) maximum likelihood (ML) to be consistent and asymptotically normal, under (2), (4) and a generalized model. On the other hand, under certain other asymptotic conditions on W_n, Lee (2002) showed that LS can be consistent in (4). Panel data versions of these models have also been studied (e.g., Baltagi, Song and Koh (2003), Case, Rosen and Hines (1993)).

Under Gaussianity, (2) can principally be viewed as a model for the covariance matrix of $\varepsilon_{(n)}$, i.e. $(I_n - \rho W_n)^{-1} \Omega (I_n - \rho W_n)^{-1'}$, where I_n is the $n \times n$ identity matrix and Ω is the unknown covariance matrix of $\eta_{(n)}$. Viewed in this light, (2) appears very arbirary, because any number of alternative positive definite parametric matrices could serve as a model for the covariance matrix of $\varepsilon_{(n)}$. On the other hand, the stress on (2) is understandable due to the intuitive appeal of a "lag" structure analogous to that in represenions of time series. Stationary time series imply a Toeplitz matrix, whence standard asymptotics tend to result if the eigenvalues are bounded and bounded away from zero and infinity for large n. Though $(I_n - \rho W_n)^{-1}\Omega(I_n - \rho W_n)^{-1'}$ is not generally Toeplitz, it can share the same kind of eigenvalue property, and lead to similarly nice asymptotics. Moreover (2) is very parsimonious, an advantage in smallish samples, and is a convenient basis for testing the null hypothesis of cross-sectional independence (see e.g. Baltagi and Dong Li (2001). A major feature of (2) and (4) is the specification of W_n , to which parameter estimates will of course be sensitive. As an earlier remark implies, this rests upon a satisfactory determination of "distance" measures between each pair of locations. One can consider related models to (2) and (4), such as a spatial moving average, or an extension of (2) with ρW_n replaced by $\sum_{i=1}^{r} \rho_i W_{in}$ for $r > 1$, $\sum_{i=1}^{r} W_{in} = W_n$, and unknown ρ_i, as may be plausible when the sample can be naturally split between sub-samples. Of course the W_{in} must at least be distinct for the ρ_i to be identifiable (see e.g. Anselin (2001)).

An alternative approach to the introduction of weak dependence that leads to standard asymptotics is essentially nonparametric, involving mixing conditions, which have been popular in time series asymptotic theory for many years, and they have been employed in the random fields probabilistic literature, as well as in some spatial econometric settings. They desirably avoid parametric descriptions of dependence, and permit a degree of non-trending heterogeneity, and can be imposed, say, on $\varepsilon_{(n)}$ in (3) in order to establish asymptotic normality of LS and other estimates of β. In the time series literature, mixing conditions, which deliver asymptotic normality, have sometimes featured alongside consistent estimates of the limiting covariance matrix (of parameter estimates, such as regression coefficients) which are analogous to smoothed nonparametrc estimates of the spectral density of a stationary process at frequency zero (see e.g. Hannan (1957), Brillinger (1979) and many subsequent econometric references). Analogous estimates have also been developed in spatial settings, see e.g. Kelejian and Prucha (2006). However, mixing conditions require some sort of ordering of the data, and in moderate sample sizes the essentially nonparametric covariance matrix estimates can be imprecise, while heterogeneity can still cause a problem in finite samples because the inference rules used were originally developed for stationary data.

3 Adaptive estimation

Most work on models such as (2)-(4) has been motivated by Gaussianity, not in the sense that it is really needed for basic asymptotic theory for estimates of ρ, λ, β, but in that estimates are based on second moment statistics and may be asymptotically efficient under Gaussianity. On the other hand they are not asymptotically efficient for non-Gaussian populations, and if n is not very large it is desirable to try to improve precision.

The representations in (2) and (4) introduce a useful whitening of the data $y_{(n)}$, and it is possible to establish desirable properties ($n^{1/2}-$consistency and asymptotic normality) of ML estimates for some specified non-normal parametric distribution of the elements of $\eta_{(n)}$. But if this is misspecified the estimates might be less efficient than "Gaussian-based" ones, and may even be inconsistent. There is thus interest in "adaptive" semiparametric estimates, that achieve the same asymptotic efficiency as ML but without knowing the distribution, and also lead to more powerful tests, for example for the hypothesis of cross-sectional independence. In other settings, the nonparametric method most often used to estimate the distribution (or more precisely its score function) has been kernels. An attractive alternative is series estimates, which have definite advantages, in terms of the regularity conditions on the model for asymptotic theory that they entail. They have been developed by Beran (1976), Newey (1988), Robinson (2005). (Series estimation has also been used for different purposes in a spatial context by Chen and Conley (2001), Pinkse, Slade and Brett (2002).)

Robinson (2006a) developed series adaptive estimates of ρ and β in (4), and justified them under conditions on W_n similar to those found by Lee (2002) to allow consistency and asymptotic normality of LS (where all $w_{ij,n}$ at least tend to zero as n increases). A simple example satisfying such conditions (see Case, 1992) considers n_1 districts, each of which has n_2 farmers ($n = n_1 n_2$) and there is uniform weighting within districts and zero-weighting across; Lee (2002) lets n_1 and n_2 both diverge as $n \to \infty$. (Robinson (2006a) also established analogous results for the fully parametric case, where a parametric distribution of the elements of $\eta_{(n)}$ is correctly specified.)

4 Nonparametric regression

Modern practice with cross-sectional and time series data leads us to question the standard linear regression setting of (3) or (4). If these are misspecified, invalid inferences are liable to result. Nonparametric regression has become a standard statistical tool, at least in large data sets, due to a recognition that there can be little confidence that the functional form is linear (as in (3) or (4)), or of a specific nonlinear type. Thus we revert to the model (1) with f a nonparametric, albeit smooth, function.

Estimates of the nonparametric regression function f are typically obtained at several fixed points by some method of smoothing. The most popular smoothed nonparametric regression estimate, when x_i is stochastically generated, is perhaps still the Nadaraya-Watson kernel estimate. When the errors ε_i and the regressors x_i in (1) are independent and identically distributed, under additional conditions this estimate is consistent, and moreover under further conditions it is asymptotically normally distributed with asymptotic variance of simple form, and indeed estimates of f at distinct fixed points are desirably asymptotically independent. Corresponding properties are enjoyed by the kernel estimate of the probability density of x_i .

Kernel regression and density estimates were originally motivated by independent and identically distributed observations, but their asymptotic statistical behaviour has long been studied in the presence of stationary time series dependence. Predictably, they remain consistent in the presence of even quite strong forms of dependence. More surprisingly, under forms of (mixing or linear) weak dependence, it has been found that not only do the kernel estimates retain their basic consistency property, but they have the same limit distributional properties as just referred to under independence (see, e.g. Roussas (1969), Rosenblatt (1971, Robinson (1983)). This finding contrasts with that in parametric forms of the regression model (1), where dependence in errors ε_i generally changes the asymptotic variance, and causes a loss in efficiency of estimates such as least squares.

With long range dependence, however, asymptotic distributional properties of kernel estimates are liable to be affected. For the kernel probability density estimate, Robinson (1991) found that while under some conditions asymptotic joint normality of a vector of estimates at distinct fixed points may still obtain, the asymptotic variance matrix, far from being diagonal as in the previous circumstances mentioned, is is singular, of rank 1. In other circumstances the limit distribution can be nonstandard. In the fixed-design nonparametric regression form of (1), where the x_i are nonstochastic and, for example, equally-spaced on the unit interval (so they get closer as n increases) Hall and Hart (1990), Robinson (1997) found that long range dependence in ε_i affects the rate of convergence of estimates of f, though asymptotic normality can still hold.

Robinson (2007) establishes consistency and asymptotic distribution theory for the Nadaraya-Watson kernel regression estimate in a framework that applies to various kinds of spatial data. The general triangular array setting covers, for example, stationary ε_i on a lattice of arbitrary dimension, and ε_i generated by a spatial autoregression. Unlike in the bulk of the time series literature mixing conditions are not employed. Instead, a linear (in independent random variables) process representation for ε_i is assumed, covering both weak dependence and long range dependence. Moreover, the ε_i can be conditionally (on x_i) heteroscedastic. A mild falling-off of dependence in the x_i is imposed, but unusually for the kernel literature, they are not necessarily assumed identically distributed, but satisfy a milder kind of homogeneity condition. Asymptotic normality of the Nadaraya-Watson kernel regression estimate is established, where the limit distribution is the same as under independence in case of weak dependence in ε_i, but not under long range dependence. The implications of the conditions are examined in various spatial settings. A consistency result for the kernel probability density estimate is established, under different conditons from those previously employed by Hallin, Lu, and Tran (2004). As always with smoothed nonparametric regression there is a curse-of-dimensionality problem if x_i is vector-valued; a semiparametric approach is considered by Gao, Lu and Tjostheim (2006), is designed to alleviate this.

5 Testing for spatial independence

Previous sections have heavily emphasized the influence of spatial dependence. Spatial dependence can invalidate inferences based on parametric models, and is likely to impair finite-sample properties in inference on nonparametric models. Moreover, developing procedures that take account of spatial dependence, in observations or disturbances, can be very complicated, the procedures can be computationally onerous, and derivation of asymptotic statistical properties of such procedures under comprehensible conditions can be problematic. For example this is liable to be the case when observations are irregularly-spaced, as experience fom the time series literature suggests.

Cressie (1993) has suggested that much spatial data can often be satisfactorily modelled in terms of the conditional mean, in the sense that little correlation in errors will remain. This desirable outcome cannot be taken for granted, however, but it is often desirable to commence analysis of spatial data by a test for independence of observations or unobservable errors. If the evidence for independence is strong then simple rules of inference on the parameters of interest have validity.

Testing for independence has been a major, long-standing theme of the spatial literature. By "independence" here we really mean "lack-of-correlation", though these concepts are generally identical only under Gaussianity. A key early contribution is Moran (1950), which indeed preceded the bulk of the time series literature on independence testing. The literature hs been further developed by Cliff and Ord (1968, 1972), Anselin (2001), Baltagi and Dong Li (2001), Baltagi, Song and Koh. (2003), Pinkse (2004), though settings have been fairly specific.

Robinson (2006b) presents a general approach which can be applied in a variety of spatial circumstances. As with the earlier work of Moran (1950) and others, the tests are based on quadratic functions, in particular of least squares residuals in linear regeessions. A general class of statistics is developed that has a chi-square limit distribution under the null hypothesis of independence of errors. It is found that special cases of the statistic can be interpreted as Lagrange multiplier statistics directed against specified alternatives, where they should have good power. Indeed the Lagrange multiplier tests maximize local power, as expected. Under Gaussianity, modified versions of the statistics are developed which exactly have the mean of the relevant chi-square distribution, and even both the mean and the variance, and should thus have better finite-sample properties. The principal focus takes homoscedasticity of errors for granted, but the tests are also robustified to nonparametric heteroscedasticity, in the sense that a valid null asymptotic (chi-square) distribution.results. The conditions are illustrated in tests in specific spatial settings, including lattice, SAR and irregularly-spaced ones. With respect to pure, distrbution-free, independence, as distinct from "lack-of-correlation", various existing tests can in principal be applied to spatial data, and Brett and Pinkse (1997) provided a specific test in a spatial context.

REFERENCES

Anselin, L. (1988) Spatial Econometrics: Methods and Models. Kluwer Academic Publishers, The Netherlands.

Anselin, L. (2001) Rao's score test in spatial econometrics. Journal of Statistical Planning and Inference 97, 113-139.

Arbia, G. (2006) Spatial Econometrics: Statistical Foundations and Applications to Regional Analysis. Springer-Verlag, Berlin.

Baltagi, B. and Dong Li (2001) LM tests for functional form and spatial error correlation. International Regional Science Review 24, 194-225. Baltagi, B., Song, S.H. and Koh, W. (2003) Testing panel data regression models with spatial error correlation. Journal of Econometrics 117, 123-150.

Beran, R. (1976) Adaptive estimates for autoregressive processes. Annals of the Institute of Statistical Mathematics 26, 77-89.

Brett, C. and Pinkse, J. (1997) The taxes are all over the map! A test for spatial independence of municipal tax rates in British Columbia. International Regional Science Review 20, 131-151.

Brillinger, D.R. (1979) Confidence intervals for the cross-covariance function. Select. Statistica Canadiana 5, 3-16.

Case, A.C. (1992) Neighbourhood influence and technological change. Regional Science and Urban Economics 22, 491-508.

Case, A., Rosen, H.S. and Hines, J.R. (1993) Budget spillovers and fiscal policy interdependence: evidence from the States. Journal of Public Economics 52, 285-307.

Chen, X. and Conley, T.G. (2001) A new semiparametric spatial model for panel data time series. Journal of Econometrics 105, 59-83.

Cliff, A. and Ord, J.K. (1968). The problem of spatial autocorrelation. Joint Discussion Paper, Universty of Bristol: Department of Economics No. 26, Department of Geography, Series A, No. 15.

Cliff, A. and Ord, J.K. (1972). Testing for spatial autocorrelation among regression residuals. Geographical Analysis 4, 267-284.

Cliff, A.and Ord, J.K.(1981). Spatial Processes: Models & Applications, Pion, London.

Conley, T.G. (1999) GMM estimation with cross sectional dependence. Journal of Econometrics 92, 1-45.

Conley, T.G. and Molinari, F. (2004) Spatial correlation robust inference with errors in location or distance. Paper presented at ceMMAP meeting on Spatial Econometrics, November 2004.

Conley, T.G. and Dupor, B. (2003) A spatial analysis of sectoral complementarity. Journal of Political Economy 111, 311-352.

Cressie, N. (1993) Statistics for Spatial Data. Wiley, New York.

Dahlhaus, R. and Künsch, H. (1987) Edge effects and efficient parameter estimation for stationary random fields. Biometrika 74, 877-882.

Fox, R. and Taqqu, M.S. (1986) Large sample properties of parameter estimates for strongly dependent stationary Gaussian time series. Annals of Statistics 14, 517-532.

Giacomini, R. and Granger, C.W.J. (2004) Aggregation of space-time process. Journal of Econometrics 118, 7-26.

Gao, J., Lu, Z. and Tjostheim, D. (2006) Estimation in semiparametric spatial regression. Annals of Statistics 34, 1395-1435.

Guyon, X. (1982) Parameter estimation for a stationary process on a d-dimensional lattice. Biometrika 69, 95-106.

Guyon, X. (1995) Random Fields on a Network, Springer-Verlag, New York.

Haining, R. (1990) Spatial Data Analysis in the Social and Environmental Sciences. Cambridge University Press, Cambridge.

Hall, P. and Hart, J.D. (1990) Nonparametric regression with long-range dependence. Stochastic Processes and their Applications 36, 339-351.

Hallin, M., Lu, Z. and Tran, L.T. (2004) Kernel density estimation for spatial processes: the L_1 theory. Journal of Multivariate Analysis 88, 61-75.

Hannan, E.J. (1957) The variance of the mean of a stationary process. Journal of the Royal Statistical Society Series B 19, 282-285.

Hidalgo, F.J. and Robinson, P.M. (2002) Adapting to unknown disturbance autocorrelation with long memory. Econometrica 70, 1545-1581.

Hsiao, C. (2003) Analysis of Panel Data (2nd Edition). Cambridge University Press, Cambridge.

Kelejian, H.H. and Prucha, I.R. (1998) A generalized spatial two-stage least squares procedure for estimating a spatial autoregressive model with autoregressive disturbances. Journal of Real Estate Finance and Economics 17, 99-121.

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

Kelejian, H.H. and Prucha, I.R. (2006) HAC estimation in a spatial framework.Journal of Econometrics, forthcoming.

Lahiri, S.N. (1996) On inconsistency of estimates based on spatial data under infill asymptotics. Sankhya Series A, 58, 403-417.

Lee, L.F. (2002) Consistency and efficiency of least squares estimation for mixed regressive, spatial autoregressive models. Econometric Theory 18, 252-277.

Lee, L.F. (2003) Best spatial two-stage least squares estimators for a spatial autoregressive model with autoregressive disturbances. Econometric Reviews 22, 307-335.

Lee, L.F. (2004) Asymptotic distributions of quasi-maximum likelihood estimates for spatial autoregressive models. Econometrica.

Newey, W.K. (1988) Adaptive estimation of regression models via moment restrictions. Journal of Econometrics 38, 301-339.

Pinkse, J., Slade, M.E. and Brett, C. (2002) Spatial price competition: a semiparametric approach. Econometrica 70-3, 1111-1153.

Pinkse, J. (2004) Moran-flavoured tests with nuisance parametrics: examples. In Anselin, L., Florax, R.G.M., Rey, S. (eds.) Advances in Spatial Econometrics: Methodology, Tools and Applications. Springer-Verlag, Berlin, pp.67-77.

Pinkse, J., Shen, L. and Slade, M.E. (2004) Dynamic spatial probit with fixed effects using one step GM1: an application to mine operating decisions. Paper presented at cemmap meeting on Spatial Econometrics, November 2004.

Robinson, P.M. (1983) Nonparametric estimators for time series. Journal of Time Series Analysis 4, 185-207.

Robinson, P.M. (1991) Nonparametric function estimation for long-memory time series". In Barnett, W., Powell, J., and Tauchen, G. (eds.) Nonparametric and Semiparametric Methods in Econometrics and Statistics.

Cambridge University Press, Cambridge, 437-457.

Robinson, P.M. (1997) Large-sample inference for nonparametric regression with dependent errors. Annals of Statistics 28, 2054-2083.

Robinson, P.M. (2005) Efficiency improvements in inference on stationary and nonstationary fractional time series. Annals of Statistics. 33, 1800-1842.

Robinson, P.M. (2006a) Efficient estimation of the semiparametric spatial autoregressive model. Journal of Econometrics, forthcoming.

Robinson, P.M. (2006b) Correlation testing in time series, spatial and cross-sectional data. Preprint.

Robinson, P.M. (2007) Nonparametric regression with spatial data. Preprint.

Robinson, P.M. and Vidal Sanz, J. (2006) Modified Whittle estimation of multilateral models on a lattice. Journal of Multivariate Analysis 97, 1090-1120.

Rosenblatt, M. (1971) Curve estimates. Annals of Mathematical Statistics 42, 1815-1842.

Roussas, G. (1969) Nonparametric estimation of the transition distribution function of a Markov process. Annals of Mathematical Statistics 40, 1386-1400.

Stein, M.L. (1991) Fixed-domain asymptotic for spatial periodograms. Journal of the American. Statistical Association 90, 1277-1288.

Stein, M.L. (1999) Interpolation of Spatial Data: Some Theory for Kriging. Springer-Verlag, New York.

Whittle, P. (1954) On stationary processes in the plane. Biometrika 41, 434-449.