Almenar, S. and Sánchez Hormigo, A. 2004. Bibliografía de historia del pensamiento económico en España. In *Economía y economistas españoles*, vol. 9, ed. E. Fuentes Quintana. Barcelona: Galaxia Gutenberg-Funcas.

Azpilcueta, Martín de. 1556. *Comentario resolutorio de cambios*, ed. J.M. Pérez Prendes and L. Pereña. Madrid: CSIC, 1965.

Campomanes, Conde de. 1774–1775. *Discurso sobre el fomento de la industria popular y Discurso sobre la educación popular de los artesanos y su fomento*, ed. J. Reeder. Madrid: Instituto de Estudios Fiscales, 1975.

Cardoso, J.L. and Lluch, E. 1999. Las teorías económicas contempladas a través de una óptica nacional. In *Economía y economistas españoles*, vol. 1, ed. E. Fuentes Quintana. Barcelona: Galaxia Gutenberg-Funcas.

Colmeiro, M. 1861. *Biblioteca de economistas españoles de los siglos XVI, XVII y XVIII*, ed. L. Perdices y J. Reeder. Madrid: IEF, Fundación ICO y Real Academia de Ciencias Morales y Políticas, 2005.

Estapé, F. 1990. *Introducción al pensamiento económico. Una perspectiva española*. Madrid: Espasa Calpe.

Flórez Estrada, A. 1828. *Curso de Economía Política*, ed. S. Almenar, 2 vols. Madrid: Instituto de Estudios Fiscales, 1980.

Fuentes Quintana, E., ed. 1999–2004. *Economía y economistas españoles*, 9 vols. Barcelona: Galaxia Gutenberg-Funcas.

Grice-Hutchinson, M. 1993. *Economic thought in Spain: Selected Essays*. Aldershot: Edward Elgar.

Jovellanos, G.M. de. 1795. Informe en el Expediente de Ley Agraria. In *G.M. de Jovellanos, Escritos económicos*, ed. V. Llombart. Madrid: I.E.F., Fundación ICO, Real Academia de Ciencias Morales y Políticas, 2000.

Llombart, V. 2004. Traducciones españolas de economía política: 1700–1812: catálogo bibliográfico y una nueva perspectiva. *Cyber Review of Modern Historiography*, No. 9. Firenze: Firenze University Press.

Lluch, E. 1973. *El pensament econòmic a Catalunya (1760–1840). Els orígens ideològics del proteccionisme i la pressa de consciència de la burgesia catalana*. Barcelona: Edicions 62.

Lluch, E. 1980. Sobre la historia nacional del pensamiento económico. In *A. Flórez Estrada, Curso de Economía*, ed. S. Almenar. Madrid: I.E.F.

Lluch, E. 1999. Las historias nacionales del pensamiento económico y España. In *Economía y economistas españoles*, vol. 1, ed. E. Fuentes Quintana. Barcelona: Galaxia Gutenberg-Funcas.

Martín Rodríguez, M. 1989. La institucionalización de los estudios de Economía Política en la Universidad Española (1784–1857). In *Marqués de Valle Santoro, Elementos de Economía Política con aplicación particular a España*. Madrid: I.E.F.

Ortiz, L.E. 1558. *Memorial del contador Luis de Ortiz a Felipe II*, ed. J. Larraz. Madrid: Instituto de España, 1970.

Perdices, L. and Reeder, J. 2003. *Diccionario de Pensamiento Económico en España (1500–2000)*. Madrid: Síntesis, Fundación ICO.

Perrotta, C. 1993. Early Spanish mercantilism: the first analysis of underdevelopment. In *Mercantilism: The Shaping of an Economic Language*, ed. L. Magnusson. London: Routledge.

Schwartz, P. 2000. The Wealth of Nations censured. Early translations in Spain. In *Contributions to the History of Economic Thought in. Essays in Honour of R.D. Collison Black*, ed. E. Murphy y R. Prendergast. London: Routledge.

Smith, R.S. 1968. English economic thought in Spain, 1776–1848. In *The Transfer of Ideas: Historical Essays*, ed. C.D. Godwin et al. Durham, NC: South Atlantic Quarterly.

Smith, R.S. 1971. Spanish mercantilism: a hardy perennial. *Southern Economic Journal* 38, 1–11.

Uztáriz, Jerónimo de. 1724. *Theórica y práctica de comercio y de marina*. Madrid: Aguilar, 1968.

Velarde, J. 1974. *Introducción a la historia del pensamiento económico español en el siglo XX*. Madrid: Editora Nacional.

spatial econometrics

Spatial econometrics is concerned with models for dependent observations indexed by points in a metric space or nodes in a graph. The key idea is that a set of locations can characterize the joint dependence between their corresponding observations. Locations provide a structure analogous to that provided by the time index in time series models. For example, near observations may be highly correlated but, as distance between observations grows, they approach independence. However, while time series are ordered in a single dimension, spatial processes are almost always indexed in more than one dimension and not ordered. Even small increases in the dimension of the indexing space permit large increases in the allowable patterns of interdependence between observations. The primary benefit of this modelling strategy is that complicated patterns of interdependence across sets of observations can be parsimoniously described in terms of relatively simple and estimable functions of objects like the distances between them.

The fundamental ingredients in any spatial model are the index space and locations for the observations. In contrast to the typical time series situation where calendar observation times are natural indices and immediately available, the researcher will often need to decide upon an index space and acquire measurements of locations/distances. The role of measured locations/distances is to characterize the interdependence between economic agents' variables, particularly those that are unobservable – for example, regression error terms. The appropriate index space depends on the economic application, and its choice is inherently a judgement call by the researcher. Fortunately, the economics of the application often provide considerable guidance and the index space/metric(s) can be tailored to promote a good fit between the economic model and the empirical work. For example, when local spillovers or competition are the central economic features, obvious candidate metrics are measures of transaction/travel costs limiting the range of the spillovers or competition. If productivity measurement were the focus, distances between observed firms or sectors could be based upon economic mechanisms that might generate co-movement in productivity – for example, measures of similarity between production technologies. Index spaces are not limited to the physical space or times inhabited by the agents and can be as abstract as required by the economics of the application.

Locations/distances are almost never perfectly measured, and this puts a premium on empirical methods that are robust to their mismeasurement. Even if the ideal metric were physical distance, usually agents' physical locations are imprecise, known only within an area – for example, census tract or county. At best this will result in imprecise distance information between agents, and if inter-agent distances are approximated with measurements based on these areas, such as distance between centroids, errors result. Moreover, in the great majority of applications the ideal metric is *not* physical distance and must be either estimated or approximated, inevitably resulting in some amount of measurement error.

There are two main approaches to modelling a spatial data generation process (DGP). The first is to model explicitly a population residing in an underlying metric space and the process of drawing an observed sample from this population. The second is to model the data-set of observed agents' outcomes as being determined by a system of simultaneous equations. In the remainder of this article, I discuss each of these approaches in turn for the simplest case of cross-sectional data. It is important to note, however, that the methods in the following section – covariance and generalized method of moments (GMM) estimation, spatial correlation robust inference – can be directly applied to panel or repeated cross-section data by simply including time as one of the components in the spatial index (s defined below). Most if not all cluster/group effect models can be considered a special case of spatial models with a binary metric indicating common group/cluster membership. See Wooldridge (2003) for an excellent review of these models. I do not discuss them here because their associated empirical techniques and sampling schemes do not translate well to more general spatial models. I conclude with a brief discussion of areas of econometrics where links to spatial econometrics are perhaps underappreciated.

1. Models for samples from a population

This section discusses spatial econometric models that view the data as being a sample from some arbitrarily large population (see, for example, Conley, 1999, for a more formal treatment). The population of individuals is assumed to reside in some metric space, typically $\Re^k$ or an integer lattice, with each individual i located at a point s_i.

The basic model of dependence characterizes dependence between agents' random variables via their locations. The data are assumed to be weakly dependent (perhaps after de-trending). (Andrews, 2005, is an important exception that explicitly considers strong cross sectional dependence arising from common shocks.) If two agents' locations s_i and s_j are close, then their random variables ϕ_{s_i} and ϕ_{s_j} may be highly dependent. As the distance between s_i and s_j grows large, ϕ_{s_i} and ϕ_{s_j} become essentially independent. Notions of weak dependence can be formalized in essentially the same manner as for time series, for example, with mixing coefficients. Under regularity conditions limiting the strength of dependence, laws of large numbers and central limit results can be obtained for properly normalized averages of ϕ_s. See, for example, Takahata (1983) or Bolthausen (1982). These approximations almost always use what is called an increasing domain approach to limits, with the corresponding thought experiment being that, as the sample size grows, an envelope containing the locations would be growing without bound.

When one works within this framework, it is often useful to approach an empirical problem in two steps. First, decide upon a (small) set of metrics based on the economics of the application, and then consider statistical modelling of dependence as a function of the metrics. It is much easier to conduct statistical modelling given a metric than to try to simultaneously vary both the model specification and the metric itself.

Statistics that describe spatial correlation patterns are simple to construct. Any statistic relating co-variation of ϕ_{s_i} and ϕ_{s_j} to some measure of their proximity could be used to characterize patterns in dependence. Classic references are Moran (1950) and Geary (1954), and the text by Cliff and Ord (1981) contains a good treatment. One useful approach is based on nonparametric estimation of a covariance function (see for example Conley and Topa, 2002, or Conley and Ligon, 2002). The ϕ_s process is covariance stationary if its expectation is the same at all locations and $cov(\phi_s, \phi_{s+h})$ depends only on the relative displacement h. For high-dimensional h, it is useful to consider a special case called isotropy where covariances depend only on the length of h; covariance depends upon distance but not direction. Take an isotropic covariance stationary ϕ_s with expectation zero for simplicity. Its covariance function f can be expressed in a regression equation involving distances $d_{i,j}$:

$$E(\phi_{s_i}\phi_{s_j}|s_i, s_j) = f(d_{i,j}). \qquad (1)$$

The function f in eq. (1) can be estimated parametrically or, as is particularly useful in preliminary data analysis, via a nonparametric regression of $\phi_{s_i}\phi_{s_j}$ on $d_{i,j}$. Investigation of correlation patterns when there is more than one candidate metric can by done by simply letting f be a function of more than one distance measure.

In cases where ϕ_s is not isotropic or non-stationary, f can still be interpretable as a measure of average co-movement. If the process is covariance stationary but not isotropic, an estimate at a given distance d_0, call it $\hat{f}(d_0)$, will converge to a weighted average of $cov(\phi_s, \phi_{s+h})$ for displacements h that have length d_0. The relative weights of different directions h will depend on their frequency of sampling. An analogous interpretation of f is available when ϕ_s is non-stationary, $cov(\phi_s, \phi_{s+h})$ depends on s, but still weakly dependent with averages of $cov(\phi_s, \phi_{s+h})$ across s remaining convergent. In this case, $\hat{f}(d_0)$ will converge to a weighted average of $cov(\phi_s, \phi_{s+h})$ across those h with length d_0 and across all s. Typically, this is still a valuable measure of co-movement. If non-stationarity is suspected, it is also very useful to construct localized versions of measures of spatial correlation. Localized f estimates for subregions of the locations can easily be constructed by just confining the observations used to estimate (1); see Anselin (1995) for extensive treatment of localized versions of Moran (1950) and Geary (1954) measures of spatial correlation.

Estimates of f can also be viewed directly as test statistics for the null hypothesis of independence. Under the null hypothesis of independence, the sampling distribution of an f estimator can be approximated and compared to the realized value of f estimates to test the hypothesis of independence. Such tests for independence remain valid even with measurement errors in distances (see Conley and Ligon, 2002).

Parameter estimation via moment conditions

In most econometric applications, the parameters of interest can be estimated using GMM. GMM estimation with weakly spatially dependent data is straightforward, and the spatial dependence is relevant for inference and efficiency (see Conley, 1999). Consider instrumental variables (IV) estimation in the linear model with outcome y_{s_i}, regressors x_{s_i}, and instruments z_{s_i}

$$y_{s_i} = x'_{s_i}\beta + u_{s_i} \quad \text{and} \quad Ez_{s_i}u_{s_i} = 0 \tag{2}$$

The IV estimator is identified by the moment condition (2): that the instruments are not correlated with the error term. Since this is a moment condition with respect to the marginal distribution of the data across agents, it is valid with or without spatial dependence. The familiar solution remains: $\beta = \left(Ez_{s_i}x'_{s_i}\right)^{-1} Ez_{s_i}y_{s_i}$. Consistent estimates of β can be obtained using sample averages to approximate these expectations since a law of large numbers applies to weakly dependent spatial data. Thus, the usual IV estimator, $\hat{\beta}_N = (\frac{1}{N}\sum_{i=1}^{N} z_{s_i}x'_{s_i})^{-1}\frac{1}{N}\sum_{i=1}^{N} z_{s_i}y_{s_i}$, remains consistent with weak spatial dependence. It is of course feasible to construct $\hat{\beta}_N$ without any knowledge of locations/distances, so it is trivially robust to measurement error in them. The impact of such spatial dependence is only upon inference, getting correct standard errors or testing.

This logic carries over to any GMM estimator of a parameter θ_0 that is identified from a moment condition involving a (potentially nonlinear) function g:

$$Eg(\phi_s; \theta_0) = 0.$$

The majority of econometric models with nonlinearity or limited dependent variables can be estimated via some choice for g. Under mild regularity conditions, θ_0 can be consistently estimated by minimum distance methods using $\frac{1}{N}\sum_{i=1}^{N} g(\phi_{s_i}; \cdot)$ to approximate $Eg(\phi_s; \cdot)$. A GMM estimator is the argument minimizing the criterion function, $J_N(\theta)$, which takes the same form as with time series or independent data:

$$J_N(\theta) = \left[\frac{1}{N}\sum_{i=1}^{N} g(\phi_{s_i}; \theta)\right]' \Omega \left[\frac{1}{N}\sum_{i=1}^{N} g(\phi_{s_i}; \theta)\right],$$

where Ω is some positive definite matrix. Just as for the time series case (Hansen, 1982), an efficient GMM estimator can be obtained by taking Ω to be a consistent estimator of the limiting variance-covariance matrix of $\frac{1}{\sqrt{N}}\sum_{i=1}^{N} g(\phi_{s_i}; \theta_0)$, whose form depends on the spatial covariance structure of the data. One such covariance matrix estimator is described in the following subsection.

Inference

The usual approach to inference using large sample approximations can be employed with weakly spatially dependent data. Returning to the IV model, the typical approximation for the distribution for $\hat{\beta}_N$ is based on the expression:

$$\sqrt{N}(\hat{\beta}_N - \beta) = \left(\frac{1}{N}\sum_{i=1}^{N} z_{s_i}x'_{s_i}\right)^{-1} \left[\frac{1}{\sqrt{N}}\sum_{i=1}^{N} z_{s_i}u_{s_i}\right]. \tag{3}$$

Under regularity conditions, the first term in the product converges to the matrix $Ez_{s_i}x'_{s_i}$. The second term in brackets has a limiting normal distribution:

$$\frac{1}{\sqrt{N}}\sum_{i=1}^{N} z_{s_i}u_{s_i} \Rightarrow N(0, V) \tag{4}$$

where V is the limiting variance-covariance matrix of

$$\frac{1}{\sqrt{N}}\sum_{i=1}^{N} z_{s_i}u_{s_i}.$$

V contains terms of the form $Ez_{s_i}u_{s_i}z'_{s_i}u_{s_i}$ and cross-covariance terms, $Ez_{s_i}u_{s_i}z'_{s_j}u_{s_j}$, that will be non-zero for at least some i,j pairs. With weak dependence, the covariance between variables indexed i and j will eventually vanish as the distance between s_i and s_j grows.

In some cases, V has a nice form. For example, suppose locations were on an integer lattice, Z^k; samples consist of all integer coordinates in a region (assumed to grow as $N \to \infty$); and variables are covariance stationary. In this case, V can be expressed as an infinite sum of a covariance function:

$$V = \sum_{h \in Z^k} Cov(z_s u_s, z_{s+h}u_{s+h}). \qquad (5)$$

With integer locations on the line, this expression coincides with its analog for covariance stationary time series.

With a consistent estimate of V, call it $\hat{V}_N$, the approximate distribution implied by (3) and (4) can be used for inference:

$$\sqrt{N}(\hat{\beta}_N - \beta) \overset{Approx}{\sim} N\left(0, \left(\frac{1}{N}\sum_{i=1}^{N} z_{s_i}x'_{s_i}\right)^{-1} \times \hat{V}_N \left(\frac{1}{N}\sum_{i=1}^{N} z_{s_i}x'_{s_i}\right)^{-1'}\right).$$

There are of course many ways V could be estimated. If it were assumed to have a parametric form, for example, by parameterizing the covariance function in (5), then consistent estimates could be obtained by GMM. Perhaps the most popular approach has been nonparametric estimation of V following Conley (1996; 1999). This approach is analogous to time series heteroskedasticity and autocovariance (HAC) consistent covariance matrix estimation, and can be viewed as a smoothed periodogram spectral density estimator. (See Priestley, 1981, for a discussion of the vast literature on spectral methods in time series and some extensions to spatial processes. Spectral methods for spatial processes date back to at least the 1950s; for example, Whittle, 1954; Bartlett, 1955; Grenander and Rosenblatt, 1957; Priestley, 1964). With the use of residuals $\hat{u}_{s_i}$ to approximate u_{s_i}, V can be estimated as a weighted sum of cross products $z_{s_i}\hat{u}_{s_i}z'_{s_j}\hat{u}_{s_j}$:

$$\hat{V}_N = \frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{N} K_N(s_i, s_j) \cdot z_{s_i}\hat{u}_{s_i}z'_{s_j}\hat{u}_{s_j}.$$

$K_N(\cdot,\cdot)$ is a kernel used to weight pairs of observations, with close observations receiving a weight near 1 and those far apart receiving weights near zero. $K_N(s_i,s_j)$ is commonly specified to be uniform kernel that is 1 if s_i and s_j are within a cut-off distance and zero otherwise. (This indicator function K_N is not guaranteed to provide positive definite (PD) covariance matrix estimates; however, this is very rarely a problem in practice. PD estimates can be insured by an alternate choice of kernel; see Conley, 1999.) $\hat{V}_N$ will be consistent if as $N \to \infty$, $K_N(s, s+h) \to 1$ for any given displacement h, but slowly enough so that the variance of $\hat{V}_N$ collapses to zero.

In practice, this estimator will require a decision about the exact form of $K_N(\cdot,\cdot)$. With a uniform kernel, this is just an operational definition of which observations are near and which are far. A conservative distinction between near and far observations can be made even with multiple candidate metrics by assigning a far classification only when all metrics agree. There is no need for the data to be covariance stationary, nor is the specific sampling framework here necessary. Analogous HAC methods can be applied to weakly dependent but non-stationary data, including that generated by simultaneous equations DGPs like those discussed in the following section 2 (see Pinkse, Slade and Brett, 2002; Kelejian and Prucha, 2007).

The main reason nonparametric estimators like $\hat{V}_N$ are often preferred to parametric models for V is their robustness to measurement errors in locations/distances. Parametric V estimators are generally inconsistent with such errors, while $\hat{V}_N$ remains consistent under mild conditions. $\hat{V}_N$ can be consistent with spatially correlated and even endogenous errors; a sufficient condition is simply that they be bounded (Conley, 1999). With location/distance errors, the weight assigned to pair i,j can be altered relative to the weight $K_N(\cdot,\cdot)$ would assign with exact locations. But $\hat{V}_N$ remains consistent, because the altered weights will still satisfy the necessary conditions for consistency of $\hat{V}_N$: the weight on observations at any true displacement will still converge to 1, slowly enough. Even when working with parametric models of V, $\hat{V}_N$ remains of interest since the discrepancy between it and a parametric V estimator can provide a useful joint test for the absence of location/distance errors and proper parametric specification (Conley and Molinari, 2007).

More important than $\hat{V}_N$ remaining consistent is its robustness in practice to moderate amounts of location error. Consider the impact of introducing location error for $\hat{V}_N$ defined with a kernel $K_N(s_i,s_j)$ equal to 1 if s_i and s_j are within L_N units, and zero otherwise. If the magnitude of measurement error is moderate relative to L_N, then the weights on most pairs of points would be unchanged if erroneously measured locations were used in place of true locations. Changes in weights occur only for those points whose true distance is near enough to the cut-off L_N that location errors result in measured and true distances being on opposite sides of L_N. With moderate amounts of location error, these pairs of observations with true distance near L_N will usually not be a large portion of the sample, so $\hat{V}_N$ will tend to be close to its value with true locations. Similar results obtain for other kernels as weights arising from moderately mismeasured locations remain close to those received with true locations (see Conley and Molinari, 2007).

2. Population simultaneous equation models

The second approach to modelling spatial data is with a simultaneous equations model, most directly interpretable as a model for a *population* of N agents. This approach explicitly specifies a joint model for the population, in contrast to typical models in Section 1, where the joint determination of outcomes in the population is not explicitly treated. These simultaneous equation models are directly applicable to situations where the entire population of agents is observed, like all US states or counties or even all firms in an industry. Typical applications include studies of games being played among these agents or of spillovers across agents; see, for example, Case, Hines and Rosen (1993) and Pinkse, Slade and Brett (2002).

The most common type of model is a simultaneous spatial autoregression (SAR). Its simplest formulation for an $N \times 1$ outcome vector Y_N is:

$$Y_N = \rho W_N Y_N + \varepsilon_N, \tag{6}$$

with scalar parameter ρ and IID shocks ε_N (typically Gaussian). The $N \times N$ matrix W_N is commonly referred to as a 'spatial weights' matrix and assumed known. W_N has zero main-diagonal elements, and its off-diagonal elements reflect some notion of interaction. Typical W_N contain (i,j) elements that are non-zero only if locations i and j are adjacent on a graph or elements inversely related to distances between locations. W_N is usually row-standardized so that its rows sum to 1. The parameter space is restricted so that $(I - \rho W_N)^{-1}$ exists and the model has reduced form:

$$Y_N = (I - \rho W_N)^{-1} \varepsilon_N.$$

Thus Y_N is a linear combination of the ε_N IID shocks. Though SAR models are finite (usually) irregular lattice models, their origins date to at least the infinite regular lattice models of Whittle (1954). Textbook treatment of SARs can be found in Anselin (1988).

Typical specifications for W_N imply a great deal of heterogeneity across observations. Variances will typically differ across the elements of Y_N by construction unless $\rho = 0$. Unconditional heteroskedasticity is thus coupled with spatial dependence. Covariances between pairs of agents will differ in patterns that are of course determined by W_N but will depend on the entire structure of this matrix and will not generally follow a simple pattern in terms of some metric. For example, with W_N defined based upon a graph, covariance between agents i and j will not be a function of their graph distance, though it can be characterized in terms of properties of the graph (Martellosio, 2004). A given graph will 'hard-wire' patterns in correlations across agents. For example Wall (2004) notes, with model (6) for US states with W_N based on adjacency, that Missouri and Tennessee are constrained to be the least spatially correlated states, while relative correlations between other pairs of states change depending on ρ. Even with a more flexible parameterization – for example, specifying the elements of W_N to be flexible functions of distance, as in Pinkse, Slade and Brett (2002) – there is still a tendency for heterogeneity in the implied joint distribution to be difficult to anticipate. While this complicates their use as statistical models, as discussed below, it is in my view likely to be a desirable property in a structural model. For example, if the model's joint distribution is to be taken seriously as capturing equilibrium outcomes for N asymmetric agents playing a game, then one would expect 'hard-wired' heterogeneity depending on the exact structure of the game.

Though the population of agents is observed, large sample approximations taking limits as $N \to \infty$ are still potentially useful. However, the requisite limit theorems technically differ from those referenced in Section 1. Since the DGP is changing as N grows, triangular array limit results are required. Consistency and distribution results for Gaussian maximum likelihood estimators (MLEs) with spatial dependence have existed at least since Mardia and Marshall (1984). An extensive set of SAR limiting distribution results is obtained by Lee (2004a) for likelihood-based estimators under a variety of conditions upon 'spatial weights' matrices like W_N. Quite useful limit theorem results can also be found in Kelejian and Prucha (2001). Correct specification of W_N is essential for these results, as SAR estimators will generally be inconsistent when there is measurement error in locations/distances used to specify this matrix (the same holds true for other parametric models of dependence structure).

A great deal of the literature has focused on computational issues involving MLEs. Non-trivial W_N matrices make computation of normalizing constants challenging. Substantial progress has been made in techniques for computing MLEs by exploiting sparseness or specific structure of 'spatial weights' matrices and re-parameterization to facilitate computation (see Pace and Barry, 1997; Barry and Pace, 1999; LeSage and Pace, 2007). These numerical techniques allow likelihood-based inference for even very large data-sets in certain applications or specifications. It is also feasible, of course, to estimate SAR parameters without computing MLEs, by using only a subset of the implications of the model to obtain method of moments estimates (see Kelejian and Prucha, 1999, and Lee, 2007, and subsequent work by these authors). This literature has been successful in addressing most computational issues with SAR models.

The key remaining difficulties in using SAR models are in terms of model specification and interpretation. Even for the simplest SAR model (6), it is hard to characterize implications of different ρ without explicitly calculating their implied joint distributions. The parameter ρ is not a simple correlation coefficient; in general it is not comparable across different specifications for W_N. In my experience, explicit calculations of descriptive measures of the implied joint distributions for many different ρ are required to understand whether varying this parameter will trace out a useful path through the space of joint distributions.

Unless one has access to virtually complete data on a population, SAR models are very difficult to properly specify as structural models. To take an optimistic case, suppose model (6) with Gaussian ε applied to a population of N agents, but a subset of agents were sampled. The likelihood of such a sample is well-defined, and in principle its form could be found by integrating out all the unobserved variables. But this calculation requires the exact form of W_N, which depends on all the unobserved agents, a full structure which will rarely be observed if only a small fraction of the agents are sampled. Proper specification of W_N is perhaps feasible only if the vast majority of the population is sampled – for example, if only a few states or counties are missing.

Even with complete data on a population, SARs are difficult to specify because they are inherently fragile. Changing a single element of W_N will in general influence the entire joint distribution of Y and it is difficult to intuitively understand the impact of a given change in W_N. Increasing flexibility by parameterizing W_N by taking its elements to be a series expansion in distance(s), as in Pinkse, Slade and Brett (2002), is of limited help. There remains only an indirect link between the series expansion and the implied joint distribution. It is hard to see how much additional flexibility in, for example, allowed covariance structure is gained by adding another term in the expansion.

I think these difficulties should be considered a consequence of modelling a large-dimensional system of structural simultaneous equations rather than SAR-specific problems. It seems likely to be difficult to anticipate changes in equilibrium outcomes resulting from changes in individual agents' decision rules or best-response functions in any modelling framework. In my view, SARs remain a useful first step towards the goal of constructing good large-dimensional structural simultaneous equation models.

Of course SAR models need not be intended as structural models; they can be viewed, for instance, as tools to incorporate spatial dependence into forecasting models. A mis-specified but parsimonious model might still forecast well. However, the cumbersome relation between specification of 'spatial weights' and the implied joint distribution makes it hard to fashion parsimonious SAR models. This seems ample reason to avoid their use in forecasting. Directly specifying measures of dependence like covariances as a parsimonious function of distance appears far easier, even if the true DGP were an SAR.

3. Links between spatial econometrics and other areas

Work on interactions-based models has much in common with simultaneous equations-style spatial models (see Brock and Durlauf, 2001, for an extensive review). In these models, the behaviour of individuals is influenced by the characteristics and/or behaviour of others. Insofar as the relevant set of 'others' can be described in a spatial framework, they can be thought of as spatial econometric models. Much of this work is theory, taking the approach of specifying conditional probability measures to capture individuals' behaviours and then deriving the implied properties of the compatible joint distribution(s). Empirical work with these models has just begun and will share many of the same challenges described above; some can even be cast directly as SARs (see Lee, 2004b).

Spatial models are potentially very useful in modelling high-dimensional vector time series. Limited degrees of freedom with typical length samples require substantial restrictions upon the DGP to make progress. The potential of spatial models to capture complicated interdependence with a small number of parameters (given auxiliary location/distance information) makes them well suited for use in characterizing a variety of restrictions upon high-dimensional vector DGPs. Good examples of the benefits of spatial approaches to this type of time series modelling are Chen and Conley (2001), Giacomini and Granger (2004), and Bester (2005a; 2005b).

TIMOTHY G. CONLEY

See also **generalized method of moments estimation; heteroskedasticity and autocorrelation corrections; social interactions (empirics); spectral analysis; stratified and cluster sampling; statistical mechanics.**

Bibliography

Andrews, D. 2005. Cross-section regression with common shocks. *Econometrica* 73, 1551–85.

Anselin, L. 1988. *Spatial Econometrics: Methods and Models.* Boston: Kluwer Academic Publishers.

Anselin, L. 1995. Local indicators of spatial association. *Geographical Analysis* 27, 93–115.

Barry, R. and Pace, R. 1999. A Monte Carlo estimator of the log determinant of large sparse matrices. *Linear Algebra and its Applications* 289, 41–54.

Bartlett, M. 1955. *An Introduction to Stochastic Processes.* Cambridge: Cambridge University Press.

Bester, C. 2005a. Random field and affine models for interest rates: an empirical comparison. Working paper, University of Chicago.

Bester, C. 2005b. Bond and option pricing in random field models. Working paper, University of Chicago.

Bolthausen, E. 1982. On the central limit theorem for stationary mixing random fields. *Annals of Probability* 10, 1047–50.

Brock, W. and Durlauf, S. 2001. Interactions-based models. In *Handbook of Econometrics* 5, ed. J. Heckman and Leamer. Amsterdam: North-Holland.
Case, A., Hines, J. and Rosen, H. 1993. Budget spillovers and fiscal policy interdependence: evidence from the states. *Journal of Public Economics* 52, 285–307.
Chen, X. and Conley, T. 2001. A new semiparametric spatial model for panel time series. *Journal of Econometrics* 105, 59–83.
Cliff, A. and Ord, J. 1981. *Spatial Processes*. London: Pion Limited.
Conley, T. 1996. Econometric modeling of cross-sectional dependence. Ph.D. thesis, University of Chicago.
Conley, T. 1999. GMM estimation with cross sectional dependence. *Journal of Econometrics* 92, 1–45.
Conley, T. and Ligon, E. 2002. Economic distance, spillovers, and cross country comparisons. *Journal of Economic Growth* 7, 157–87.
Conley, T. and Molinari, F. 2007. Spatial correlation robust inference with errors in location or distance. *Journal of Econometrics* 140(1), 76–96.
Conley, T. and Topa, G. 2002. Socio-economic distance and spatial patterns in unemployment. *Journal of Applied Econometrics* 17, 303–27.
Geary, R. 1954. The contiguity ratio and statistical mapping. *Incorporated Statistician* 5, 115–45.
Giacomini, F. and Granger, C. 2004. Aggregation of space–time processes. *Journal of Econometrics* 118, 7–26.
Grenander, U. and Rosenblatt, M. 1957. Some problems in estimating the spectrum of a time series. *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability* 7, 77–93.
Hansen, L. 1982. Large sample properties of generalized method of moments estimators. *Econometrica* 50, 1029–54.
Kelejian, H. and Prucha, I. 1999. A Generalized moments estimator for the autoregressive parameter in a spatial model. *International Economic Review* 40, 509–33.
Kelejian, H. and Prucha, I. 2001. On the asymptotic distribution of the Moran I test statistic with applications. *Journal of Econometrics* 104, 219–57.
Kelejian, H. and Prucha, I. 2007. HAC estimation in a spatial framework. *Journal of Econometrics* 140(1), 131–54.
Lee, L. 2004a. Asymptotic distributions of quasi-maximum likelihood estimators for spatial autoregressive models. *Econometrica* 72, 1899–925.
Lee, L. 2004b. Identification and estimation of spatial econometric models with group interactions, contextual factors and fixed effects. Working paper, Ohio State University.
Lee, L. 2007. GMM and 2SLS estimation of mixed regressive, spatial autoregressive models *Journal of Econometrics* 140(1), 155–89.
LeSage, J. and Pace, R. 2007. A matrix exponential spatial specification. *Journal of Econometrics* 140(1), 190–214.
Mardia, K. and Marshall, R. 1984. Maximum likelihood estimation of models for residual covariance in spatial regression. *Biometrika* 71, 135–46.
Martellosio, F. 2004. The correlation structure of spatial autoregressions. Working paper, University of Southampton.
Moran, P. 1950. Notes on continuous stochastic phenomena. *Biometrika* 37, 17–23.
Pace, R. and Barry, R. 1997. Quick computation of regressions with a spatially autoregressive dependent variable. *Geographical Analysis* 29, 232–47.
Pinkse, J., Slade, M. and Brett, C. 2002. Spatial price competition: a semiparametric approach. *Econometrica* 70, 1111–53.
Priestley, M. 1964. Analysis of two-dimensional processes with discontinous spectra. *Biometrika* 51, 195–217.
Priestley, M. 1981. *Spectral Analysis and Time Series*, 2 vols. New York: Academic Press.
Takahata, H. 1983. On the rates in the central limit theorem for weakly dependent random fields. *Zeitschrift fur Wahrscheinlichkeitstheorie und verwandte Gebiete* 64, 445–56.
Wall, M. 2004. A close look at the spatial structure implied by the CAR and SAR models. *Journal of Statistical Planning and Inference* 121, 311–24.
Whittle, P. 1954. On stationary processes on the plane. *Biometrika* 2, 434–49.
Wooldridge, J. 2003. Cluster-sample methods in applied econometrics. *American Economic Review* 93, 133–8.

spatial economics

What is spatial economics? In a nutshell, spatial economics is concerned with the allocation of (scarce) resources over space and the location of economic activity. Depending on how this definition is read, the realm of spatial economics may be either extremely broad or rather narrow. On the one hand, economic activity has to take place somewhere so that spatial economics may be concerned with anything that economics is concerned about. On the other hand, location analysis focuses mostly on one economic question, namely, location choice. This is only one decision among a large number of economic decisions.

Which boundaries for spatial economics?

In practice, we can distinguish three sets of questions for which the importance of the spatial dimension is very different. Consider first the core questions of spatial economics. For example, why are there cities? Why do some regions prosper while others do not? Why do we observe residential segregation? Why do firms from the same industry cluster? These are intrinsically 'spatial' questions, that is, questions in which the spatial dimension plays a dominant role. For instance, it would be difficult to speak meaningfully about the existence and growth of cities without some explicit consideration of