

BEMPS –

Bozen Economics & Management
Paper Series

NO 54/ 2018

Spatial Discrete Choice Models:
A Review Focused on
Specification, Estimation
and Health Economics applications

Giuseppe Arbia, Anna Gloria Billé

Spatial Discrete Choice Models: A Review Focused on Specification, Estimation and Health Economics applications

Giuseppe Arbia^a, Anna Gloria Billé^b

^a*Department of Statistical Science, Catholic University of the Sacred Heart, Rome, Italy*

^b*Faculty of Economics and Management, Free University of Bolzano–Bozen, Bolzano, Italy*

Abstract

Modeling individual choices is one of the main aim in microeconometrics. Discrete choice models has been widely used to describe economic agents' utility functions, and most of them play a paramount role in applied health economics. On the other hand, spatial econometrics collects a series of econometric tools which are particularly useful when we deal with spatially-distributed data sets. It has been demonstrated that accounting for spatial dependence can avoid inconsistency problems of the commonly used estimators. However, the complex structure of spatial dependence in most of the nonlinear models still precludes a large diffusion of these spatial techniques. The purpose of this paper is then twofold. The former is to review the main methodological problems and their different solutions in spatial discrete choice modeling as they have appeared in the econometric literature. The latter is to review their applications to health issues, especially in the last few years, by highlighting at least two main reasons why spatial discrete neighboring effects should be considered and then suggesting possible future lines of the development of this emerging field.

Keywords: Discrete Choice Modeling; Health Economics; Spatial Econometrics

1. Introduction

Discrete choice models with an explicit consideration of spatial neighboring effects have received less attention in the econometrics literature. Nevertheless, the role of space is becoming paramount in health economics as it is witnessed by the large and increasing amount of publications found in the literature on the subject in recent years, see Baltagi et al. (2012) Gravelle et al. (2014) Arbia et al. (2014) and Atella et al. (2014).

A plausible reason for the relatively scarce diffusion of spatial discrete choice (SDC) models in health economics and in all the other fields is certainly connected to their complexity, see Fleming (2004). Indeed, there are still a number of methodological problems to be solved connected with the computational burden

*Corresponding author: Anna Gloria Billé. Email: annagloria.bille@unibz.it, web page: <https://www.unibz.it/it/faculties/economics-management/academic-staff/person/38038-anna-gloria-bille>.

and accuracy of the various techniques as soon as we start dealing with large datasets which is often the case in health economics issues and microeconomic datasets. The problem becomes serious in a way that these models have experienced increasing attention in recent years by spatial econometricians, see Smirnov (2010) for a review, and the methodologies proposed are still largely unexplored, especially those related to health issues. This apparently might be not a serious problem, if only a-spatial discrete choices (DCs) and limited dependent variables (LDVs) were not widely used to solve health problems, see Jones (2000), Jones (2007). The estimation problems also preclude an easy extension to panel data applications, whose diffusion is experiencing a massive increase.

Modeling economic agent-based spatial relationships will be instead an approaching problem to be solved since that individual decisions usually depend upon neighboring agents' decisions. In the last 20 years, we have had an increased experience in econometric studies as basis of health policy. Most of them required the use of LDV or DC models to describe health care expenditures, treatment effects analysis, etc., see e.g. Varin and Czado (2009), Munkin and Trivedi (2008), Santos et al. (2017), Varkevisser et al. (2012), Lindeboom and Kerkhofs (2009), Deb et al. (2006) and Basu et al. (2007). However, a very limited number of papers take into account space and spatial structure of discrete health data sets. Some empirical works considered a distance variable or a spatial dummy variable to distinguish between districts/regions, see e.g. Geweke et al. (2003), Wolff et al. (2008) and Nketiah-Amponsah (2009), but none of them used spatial spillover effects by introducing autocorrelation coefficients.

The present paper is mainly focused on synchronic cross-sectional data analysis. Nevertheless, a brief review on spatial discrete choice panel data models is included, and for a review of spatial linear econometric models with panel health data sets the reader is referred to Moscone and Tosetti (2014). The purpose of this paper is twofold. The former is to review the main methodological problems and their different solutions in spatial discrete choice modeling as they have appeared in the econometric literature. The latter is to review their applications to health issues, especially in the last few years, by highlighting at least two main reasons why spatial discrete neighboring effects should be considered and then suggesting possible future lines of the development of this emerging field.

The paper is structured in the following way. Section 2 explain two limited dependent variable model specifications focusing on discrete choices. This section gives a brief overview of this type of models by suggesting possible guidelines. The proper definition of the marginal impacts for spatial binary probit models is also included. Section 3 a brief overview of spatial limited dependent variable models focusing on discrete choices with their main estimation problems and related solutions. In this section there is also a brief explanation of the inconsistency problem, a reason why spatial models seems to be preferred to a-spatial ones. Finally, a brief paragraph on the packages in R developed for spatial discrete choices is also included. Section 4 reviews the empirical applications in health economics which make use of SDC models and tries to stimulate the readers to adopt spatial econometric techniques at least as an alternative in comparison with standard econometric

approaches. Finally, Section 5 concludes.

2. Spatial discrete choice model specifications

2.1. Models' specifications

Spatial econometricians are usually interested in extending standard econometric models by assuming that it is likely the presence of some form of regional/social dependence when we deal with sample units observed over a space. In the following three subsections we explain the main discrete choice model specifications, the correct specifications of the marginal effects and the substantive correlation information inside the autoregressive parameter, respectively. In a discrete choice environment with binary outcomes a spatial autoregressive probit model with autoregressive disturbances, i.e. SARAR–probit, is rather general, see Billé and Leorato (2017).

Let \mathbf{y}_n be a n –dimensional stochastic vector of spatial binary variables located on a possibly unevenly spaced lattice $Z \subseteq \mathfrak{R}^n$. A spatial (first–order) autoregressive–regressive probit model with (first–order) autoregressive disturbances (SARAR(1,1)–probit) is defined as

$$\begin{aligned} \mathbf{y}_n^* &= \rho \mathbf{W}_n \mathbf{y}_n^* + \mathbf{X}_n \boldsymbol{\beta} + \mathbf{u}_n, & \mathbf{u}_n &= \lambda \mathbf{M}_n \mathbf{u}_n + \boldsymbol{\varepsilon}_n, & \boldsymbol{\varepsilon}_n &\sim \mathcal{N}_n(\mathbf{0}_n, \boldsymbol{\Sigma}_\varepsilon) \\ \mathbf{y}_n &= \mathbb{I}_n(\mathbf{y}_n^* > \mathbf{0}_n) \end{aligned} \tag{1}$$

where \mathbf{y}_n^* is the n –dimensional vector of latent continuous dependent variables, \mathbf{y}_n is the n –dimensional vector of observed binary dependent variables defined by the n –dimensional indicator function $\mathbb{I}_n(\mathbf{y}^* > \mathbf{0}) = (\mathbb{I}(y_1^* > 0), \dots, \mathbb{I}(y_n^* > 0))'$, \mathbf{X}_n is the n by k matrix of exogenous variables including a constant term, \mathbf{W}_n and \mathbf{M}_n are n –dimensional spatial weighting matrices of known constants, $\boldsymbol{\theta} = (\boldsymbol{\beta}', \rho, \lambda)'$ is a $(k + 2)$ –dimensional parameter vector with autoregressive coefficients ρ and λ , and $\boldsymbol{\varepsilon}_n$ is a multivariate normal vector of innovations with zero mean and finite variance $\sigma_\varepsilon^2 < \infty$, such that $\boldsymbol{\Sigma}_\varepsilon = \sigma_\varepsilon^2 \mathbf{I}_n$. Latent variables are then assumed to be linear functions of the regressors, but they are observed through the use of a binary variable that makes the overall model nonlinear in parameters. In the nonlinear case, σ_ε^2 is usually set to 1 for identification. Additional conditions are needed for the identification of (ρ, λ) in a SARAR(1,1)–probit model. Specifically, \mathbf{M}_n and \mathbf{W}_n are assumed to be different thus allowing for different mechanisms to govern spatial correlation between shocks affecting the latent model and spatial dependence of the latent variables themselves. Then, the entire spatial dependence can be easily disentangled. It is notable that, when $\mathbf{W}_n = \mathbf{M}_n$, then distinguishing among the two spatial effects may be difficult, with possible identification problems of the autoregressive parameters. In this particular case, sufficient conditions to ensure identifiability of the linear model is that the covariates make a material contribution towards explaining variation in the dependent variable.

An alternative specification of equation (1) is the Spatial Durbin probit model (or the Spatial Durbin error probit model), which can be useful to avoid possible omitted variable biases and to inform about local

correlation effects, see e.g. Elhorst (2010) and LeSage (2014) for details in the linear case. The Spatial Durbin probit model can be written in the following way

$$\begin{aligned} \mathbf{y}_n^* &= \rho \mathbf{W}_{1,n} \mathbf{y}_n^* + \mathbf{X}_n \boldsymbol{\beta} + \mathbf{W}_{2,n} \mathbf{X}_n \boldsymbol{\theta} + \boldsymbol{\varepsilon}_n, \quad \boldsymbol{\varepsilon}_n \sim \mathcal{N}(\mathbf{0}_n, \boldsymbol{\Sigma}_\varepsilon) \\ \mathbf{y}_n &= \mathbb{I}_n(\mathbf{y}_n^* > \mathbf{0}_n) \end{aligned} \quad (2)$$

where $\mathbf{W}_{2,n} \mathbf{X}_n$ are the spatially lagged regressors with coefficients $\boldsymbol{\theta}$ which captures local spatial correlation effects, and the other terms are defined above.

The inclusion of spatially-lagged dependent variables $\mathbf{W}_n \mathbf{y}_n^*$ typically causes an endogeneity problem, which in turn produces inconsistency of least squares estimators. This problem is referred to the bi-directionality nature of spatial dependence in which each site, say i , is a second-order neighbor of itself, implying that spatial spillover effects have the important meaning of feedback/indirect effects also on the site where the shock may have had origin. The problem also makes the overall model a system of n *simultaneous* equations (one for each random variable in space), with the consequence that spatial autoregressive models cannot be viewed as simple extensions of natural *recursive* time-series econometric models, see Hamilton (1994). These type of spatial models are then multivariate by definition, with the peculiarity of having statistical information coming from one observation for each random variable in space in a cross-sectional framework.

Interesting should be also the case of more than two modalities. For instance, consider a multinomial probit model as the following one

$$\begin{aligned} u_{ij}^* &= \mathbf{x}'_{ij} \boldsymbol{\beta} + \varepsilon_{ij}, \quad \varepsilon_{ij} \sim \mathcal{N}(0, \sigma_{\varepsilon_j}^2) \\ y_{ij} &= 1 \quad \text{iff} \quad u_{ij}^* > u_{ik}^* \quad k \neq j \\ y_{ij} &= 0 \quad \text{otherwise} \end{aligned} \quad (3)$$

where the utility associated with the preference j of individual i is a linear function of some regressors x_{ij} for $j = 1, \dots, J$ alternatives and $i = 1, \dots, n$ individuals, the j -th alternative is chosen if its utility is a maximum respect to all the other alternatives, and usually $\sigma_{\varepsilon_j} = 1$ for model identification. An extension of the previous utility model can be considered by assuming that unobserved utility functions are autocorrelated, revealing that individuals' preferences depend also on the preferences of "neighboring" people (the problem in this case is to identify a reasonable \mathbf{W} matrix to define individuals' interactions). This may lead to the following spatial random utility maximization (SRUM) specification

$$\begin{aligned} u_{ij}^* &= \rho \sum_{i \neq h}^n w_{ih} u_{ij}^* + \mathbf{x}'_{ij} \boldsymbol{\beta} + \varepsilon_{ij}, \quad \varepsilon_{ij} \sim \mathcal{N}(0, \sigma_{\varepsilon_j}^2) \\ y_{ij} &= 1 \quad \text{iff} \quad u_{ij}^* > u_{ik}^* \quad \text{for} \quad k \neq j \\ y_{ij} &= 0 \quad \text{otherwise} \end{aligned} \quad (4)$$

where the term $\rho \sum_{i \neq h}^n w_{ih} u_{ij}^*$ summarizes the dependence structure between individuals' preferences. Recently, Smirnov and Egan (2012) have proposed a model of this type in order to measure unobserved spatial

interdependencies between households and establish if these interdependencies have a significant effect on the recreational travel choices. Unfortunately, the way in which they capture these unobserved spatial/social effects is based on an aggregation of the neighboring spatial units at a county level, losing the advantage of considering information at agent-based microeconomic data. Bolduc et al. (1996a) proposed a spatial autoregressive error (SAE) process (i.e. the model in (1) when $\rho = 0$) for the utility functions in order to allow for suspected interdependences among location choices (so among the alternatives or preferences: 18 regions) in a study of the choice of location by general practitioners in their initial work. Indeed, they argued that “spatial correlation is likely to be present in the data because of the similarity of unobserved attributes in neighboring regions”, finding that the spatial model was to be preferred. Some models of this type can be also found in land-use applications (e.g. Sidharthan and Bhat (2012) Chakir and Parent (2009)) in which individuals’ (land owners’) interactive decisions are associated with spatial correlation among the type of use of the land (i.e. parcel units, which corresponds to the alternatives). The problem of a knowledge diffusion of these type of models into the health field seem to be surely caused by the insufficient information on the individuals’ spatial locations, but sometimes also by the lack of applied economists’ information on more advanced econometric methods that usually comes from different literatures. For instance, the previous SAE model can be written in the following way

$$\begin{aligned}
u_{ij}^* &= \mathbf{x}'_{ij}\boldsymbol{\beta} + \varepsilon_{ij}, \quad \varepsilon_{ij} = \lambda \sum_{j \neq k}^J w_{jk} \varepsilon_{ij} + v_{ij}, \quad v_{ij} \sim \mathcal{N}(0, \sigma_{v_j}^2) \\
y_{ij} &= 1 \quad \text{iff} \quad u_{ij}^* > u_{ik}^* \quad \text{for} \quad k \neq j \\
y_{ij} &= 0 \quad \text{otherwise}
\end{aligned} \tag{5}$$

where $\lambda \sum_{j \neq k}^J w_{jk} \varepsilon_{ij}$ now summarizes the dependence structure between unobserved attributes or between selected alternatives. This model can be used for example in the context of patient hospital choices, see e.g. Varkevisser et al. (2012), in which individuals maximized their utilities in choosing among different hospitals and the choice cannot depend only on several hospital attributes (e.g. hospital’s quality) or travel time (which justifies the use of the a-spatial mixed logit model), but it is also likely the presence of spatial autocorrelation between those alternatives since a recent literature is for instance recognizing the importance of spatial competition between them, see Gravelle et al. (2014). Considering at least a spatial error structure of our model is more prominent when dealing with nonlinear models, since we generally have inconsistent estimates, see Section 3.3, rather than a loss of efficiency in the linear case.

2.2. Marginal effects

Billé and Leorato (2017) suggest the correct specifications of the marginal effects for spatial nonlinear autoregressive models. In the following we briefly explain how this marginal effects are defined.

In nonlinear regressions, the interpretation of the marginal effects in terms of the change in the conditional

mean of \mathbf{y} when regressors \mathbf{X} change by one unit is no longer possible. The effects arising from changes in the explanatory variables depend in a nonlinear way on the levels of these variables, i.e. changes in the explanatory variable near the mean have a very different impact on decision probabilities than changes in very low or high values. For spatial autoregressive probit models, the nonlinearity increases in the evaluation of the marginal effects, see Beron and Vijverberg (2004), LeSage et al. (2011). Recently, Billé (2014) has also pointed out the main consequences in evaluating marginal effects with and without the consideration of heteroskedasticity implied by the spatial autocorrelation coefficient.

Let $\mathbf{x}_{.h} = (x_{1h}, x_{2h}, \dots, x_{ih}, \dots, x_{nh})'$ an n -dimensional vector of units referred to the h -th regressor, $h = 1, \dots, k$, and $\mathbf{x}_i = (x_{i1}, x_{i2}, \dots, x_{ih}, \dots, x_{ik})'$ a k -dimensional vector of regressors referred to unit i . By correctly specifying the conditional expected value and covariance matrix of model in (1), the following specifications of the marginal effects has been proposed

$$\begin{aligned} \frac{\partial \text{P}(y_i = 1 \mid \mathbf{X}_n)}{\partial \mathbf{x}'_{.h}} \Big|_{\bar{\mathbf{x}}} &= \phi \left(\{ \Sigma_{\nu(\rho, \lambda)} \}_{ii}^{-1/2} \{ \mathbf{A}_\rho^{-1} \bar{\mathbf{X}} \}_i \cdot \boldsymbol{\beta} \right) \{ \Sigma_{\nu(\rho, \lambda)}^{-1/2} \}_{ii} \{ \mathbf{A}_\rho^{-1} \}_i \cdot \beta_h \\ \frac{\partial \text{P}(y_i = 1 \mid \mathbf{X}_n)}{\partial \mathbf{x}'_{.h}} \Big|_{\mathbf{x}} &= \phi \left(\{ \Sigma_{\nu(\rho, \lambda)} \}_{ii}^{-1/2} \{ \mathbf{A}_\rho^{-1} \mathbf{X} \}_i \cdot \boldsymbol{\beta} \right) \{ \Sigma_{\nu(\rho, \lambda)}^{-1/2} \}_{ii} \{ \mathbf{A}_\rho^{-1} \}_i \cdot \beta_h \end{aligned} \quad (6)$$

where $\Sigma_{\nu(\rho, \lambda)}$ is the variance-covariance matrix implied by the reduced form of a SARAR(1,1)-probit model and $\Sigma_{\nu(\rho, \lambda)}^{-1/2} = \{ \sigma_{\nu_i}^{-1} \}$, $\mathbf{A}_\rho^{-1} = (\mathbf{I} - \rho \mathbf{W})^{-1}$, $\bar{\mathbf{X}}$ is an n by k matrix of regressor-means, $(\cdot)_i$ considers the i -th row of the matrix inside, and $(\cdot)_{ii}$ the i -th diagonal element of a square matrix. Note that $\Sigma_{\nu(\rho, \lambda)}$ reduces to $\Sigma_{\mathbf{u}(\rho)}$ for a SAR(1)-probit specification with $\mathbf{u} = \mathbf{A}_\rho^{-1} \boldsymbol{\varepsilon}$.

The first specification of equations (6) explains the impact of a marginal change in the mean of the h -th regressor, i.e. $\bar{\mathbf{x}}_{.h}$, on the conditional probability of $\{y_i = 1\}$, i.e. $\text{P}(y_i = 1 \mid \mathbf{X}_n)$, setting $\bar{\mathbf{x}}_{.h'}$ for all the remaining regressors, $h' = 1, \dots, k - 1$. The second specification of equations (6) considers, instead, the marginal impact evaluated at each single value of $\mathbf{x}_{.h}$. The results are two n -dimensional square matrices for $\{y_1, y_2, \dots, y_n\}$. Both the specifications should be evaluated with consistent estimates of the spatial autocorrelation coefficients $(\hat{\rho}, \hat{\lambda})$.

Spatial marginal effects are then split into an *average direct impact* and an *average indirect impact*. The average of the main diagonal elements of the n -dimensional matrix, in both the equations, is the average direct effect (i.e., the impact from their own regions). The average of the cumulated off-diagonal elements is the average indirect effect – due to spatial spillover effects (i.e., the impact from other regions). Finally, the average total effects is the sum of them (LeSage and Pace, 2009). Changes in the value of an explanatory variable in a single observation (i.e. a spatial unit) i may influence all the $n - 1$ other observations. The scalar summary measure of indirect effects cumulates the spatial spillovers falling on all other observations, but the magnitude of impact will be greatest for nearby neighbors and declines in magnitude for higher-order neighbors. LeSage et al. (2011) pointed out the need to calculate measures of dispersion for these estimates. In Billé and Leorato (2017) there are some results on the marginal effects and their measures of dispersion based

on Monte Carlo simulations.

Observation-level total effects estimates, sorted from low-to-high values of each regressors, can be also viewed as an important measure of spatial variation in the impacts (Lacombe and LeSage, 2013). This kind of interpretation permits also to account for *spatial heterogeneity* due to the variation over space of the marginal impacts with respect to the spatial distribution of the regressors¹. Within nonlinear models, the possibility of evaluating a marginal impact with respect to a particular value \mathbf{x}_{ih} have the same meaning of considering a marginal impact in a particular region/site for regressor h . Finally, note that the specification of our marginal effects are different compared with those proposed by LeSage et al. (2011) and Beron and Vijverberg (2004).

2.3. A substantive correlation information

Apart from an omitted variable problem whose solving is a purely statistical purpose, it should be emphasized that the additional information deriving from the geographical location of data is of paramount importance in health economics for a number of reasons. For instance, it is relevant to describe regional inequalities in the geographical distribution of the total number of general practitioners, since this proxies inequalities in the access on health services, see Bolduc et al. (1996a). Two recent relevant papers in the spatial health econometrics field are those of Atella et al. (2014) and Gravelle et al. (2014). The former developed a spatial Durbin model (SDM) by partitioning the \mathbf{W} contiguity weighting matrix into two sub-matrices in order to take into account institutional constraints in a study of per-capita public health expenditure, finding that spatial effects plays a role mainly within entities belonging to the same institutional setting while the between effect is quite negligible. The latter used instead a spatial autoregressive-regressive (SAR) model in order to detect if a hospital's quality level depends on its rivals' quality levels in a competitive setting. The main finding was that hospitals' quality levels, in terms of the overall mortality rate, are positively autocorrelated.

From a substantive point of view, spatial parameters usually bear an important information content in a way that they cannot be thought as simply nuisance parameters². Indeed, spatial dependence not only means lack of independence between observations but also an underlying spatial structure, so that the autoregressive coefficient ρ should be interpreted as causal relationship information parameter between y^* and its neighboring values in a discrete context. This should be particularly relevant in all those cases in which we need to describe social interaction/dependence effects between economic agents over space. For instance, it might be interesting to evaluate the probability that a single person take the decision of choosing a particular health facility that has been affected by the decisions of neighboring economic agents. Moscone et al. (2012) have recently modeled peer effects between economic agents' hospital choices, but their interpretation is more related on a temporal dimension rather than a proximity in space.

¹See Billé et al. (2017) for a two-step approach specifically thought to account for unobserved discrete spatial heterogeneity in the beta's coefficients via iterated local estimation procedures.

²See Anselin (2002) for a brief discussion on differences between substantive and nuisance correlation parameters.

3. Estimation

3.1. Generalities

Traditionally, spatial regression models are estimated by maximum likelihood (ML) method. However, this approach can often become computationally unfeasible especially when dealing with discrete dependent variables. In order to solve this issue, some methodological and computational solutions have been recently proposed; and furthermore, in view of the possible computational advantages, many researchers seems to be increasingly incline to use Bayesian inference with the well-known MCMC and Gibbs sampling approaches (LeSage, 2000). At the same time, an emerging literature is seeing the development of semi- and non-parametric techniques (McMillen and McDonald, 2004). In the following of this section we provide a brief review of the main methodological innovations in the econometric subfield of discrete choice and limited dependent variable spatial modeling by distinguishing them according to the nature of the dependent variable, with the purpose to highlight the potential of the proposed solutions.

3.1.1. Binary variables

As well known from the econometric literature, discrete choice models can be distinguished according to the number of modalities of their discrete dependent variables. Nonlinear models like *binary probit/logit models* are useful to describe binary dependent variables and both of them have received particular attention in order to introduce spatial spillover effects, see McMillen (1992), Pinkse and Slade (1998), Fleming (2004), Beron and Vijverberg (2004). However, the spatial dependence structure adds complexity in the estimation of parameters, at least because of the implied heteroskedasticity. Solutions for inconsistency due to heteroskedastic variances in spatial probit/logit models have been proposed, see Case (1992) and Pinkse and Slade (1998). However, there is no consideration in these cases on the information coming from the off-diagonal elements of the variance-covariance matrix.

Due to the easier accessibility to computer-based solutions, a class of maximum simulated likelihood (MSL) estimators has been proposed to deal with both inconsistency and loss of efficiency, see McMillen (1992) and Beron et al. (2003). Nowadays, a major problem in maximizing this log-likelihood function with MSL approaches is represented by fact that it repeatedly involves the calculation of the determinant of n by n matrices whose dimension depends on the sample size, in which cases the use of sparse matrices is generally recommended, see also Pace and Barry (1997). The GMM is also affected by this problem in the nonlinear context. For this reason MSL/GMM approaches are still computationally unfeasible. Important contributions are those of Klier and McMillen (2008) in a GMM environment, Bhat (2011) and Mozharovskiy and Vogler (2016) in the realm of the composite ML estimation, and Martinetti and Geniaux (2017) for approximate ML estimation. However, the estimator proposed by Klier and McMillen (2008), a linearization of the GMM proposed by Pinkse and Slade (1998), has good properties has long has the true autocorrelation coefficient is

small. The other solutions are instead only approximations.

Differently from numerical approximation solutions, Bhat and Sener (2009) copula-based approach does not require simulation machinery and provides a simple closed-form solution, which is computationally feasible even with very large sample sizes. However, when dealing with discrete data there is no unique copula that can be defined and the interpretation of the correlation coefficient is different with respect to the autocorrelation coefficient ρ in model (1). Wang et al. (2013) proposed instead a partial maximum likelihood approach which is based on a trade-off solution between statistical efficiency and computational burden. Their limits are mainly relative to the model specification, i.e. a spatial error process which is less attractive for empirical applications, and the absence of a criterion for the partition of the spatial data into groups of pairs of random variables. Billé and Leorato (2017) overcome these limits.

Despite the above mentioned estimation limits, the literature is increasingly interested also in extending the above model specifications with panel data. In this context we recognise the recent work by Pinkse et al. (2006), Arduini (2016) and Baltagi et al. (2016). The first one specifies a dynamic model with a one-step GMM estimation procedure, whereas the second and the third proposed a semiparametric approach and a Bayesian pairwise approach, respectively.

3.1.2. Ordered and unordered variables

When we deal with more than two modalities ordered or unordered between them, *ordered-response probit/logit models* and *multinomial probit/logit models* are adopted, respectively.

In health economics, ordered response models are usually used to describe individual inequalities of self-assessed health (SAH) and its reporting heterogeneity (Lindeboom and Van Doorslaer, 2004), state-dependent reporting bias and justification bias (Lindeboom and Kerkhofs, 2009), or scale of reference bias problem (Groot, 2000)³. Although many databases require ordered discrete responses in a spatial context, few papers with spatial spillover effects have been found. Among these, two relevant papers are those of Ferdous and Bhat (2013) and Castro et al. (2013). The former developed a spatial panel ordered-response model with spatial dependence introduced in both the exogenous variables and the error terms, while accounting for unobserved spatial heterogeneity and accommodating time-varying dependency effects in a urban land-use application. The latter proposed a spatial random coefficient generalized ordered-response probit (SRC-GORP) model with a spatial intermediate formulation of the dependence structure to analyze injury severity of crashes occurring at urban intersections. Both the estimation procedures rely on the composite marginal likelihood proposed by Bhat (2011).

Multinomial probit/logit models are instead justified by the *random utility theory*, see McFadden (2001) and Manski (1981), and are usually used in health economics to describe individuals' choices and utilizations

³See Greene et al. (2014) for a recent review of ordered response models for this type of applications.

of health care services. These models define individual utility functions based on some features that only vary between individuals (i.e. effects through decision-makers) and some others that only vary among individual choices (i.e. effects among choice alternatives). See Weeks (1997) for a review on specification and estimation of this type of models. The IIA property is unlike to hold in spatial autoregressive models. The GEV class of models, see Hunt et al. (2004), Bhat and Guo (2004) Bekhor and Prashker (2008) and Pinjari (2011), relaxes the IID assumption of the MNL by allowing the random components of alternatives to be correlated, while maintaining the assumption that they are identically distributed, assuming a Gumbel distribution for the error terms. For instance, Bhat and Guo (2004) proposed a mixed spatially correlated logit (MSCL) model which utilized a GEV structure in order to consider utility correlation between spatial units, and they superimposed a mixed distribution on the GEV structure to capture the unobserved response heterogeneity in a housing choices study. Bekhor and Prashker (2008) examined several GEV models to discuss their adaptability on destination choice situations, with the object to determine the probability that a person from a given origin chooses a particular destination among different available alternatives. Pinjari (2011) has formally obtained the class of multiple discrete-continuous generalized extreme value (MDCGEV) models, and in particular he tested the existence and extracted the general form of the consumption probability in a closed-form, with an application in a household expenditure analysis. Finally, Bhat et al. (2015) developed a spatial multiple discrete-continuous probit (MDCP) model to specify and estimate a model of land-use change that is capable of predicting both the type and the intensity of urban development patterns over large geographic areas. The formulation also accommodates spatial heterogeneity and heteroskedasticity in the dependent variable, and should be applicable in a wide variety of fields where social and spatial dependencies between decision agents (or observation units) lead to spillover effects in multiple discrete-continuous choices (or states). The estimation procedures of the GEV class of models rely on maximum-simulated likelihood (MSL) estimation which is time-consuming as mentioned in the previous section, while Bhat et al. (2015) considers the composite marginal likelihood in Bhat (2011).

3.1.3. *Count data and limited dependent variables*

A different discussion can be made for count data variables. As already well-known, *count data models* are used when dependent variables consist in a count of positive integers. Due to the nature of these variables, data are usually affected by asymmetric distribution problems and high proportions of zero. In health economics these models have been subjected to a wide diffusion in order to analyze the demand for health care and the health care utilization. Empirical spatial econometric papers with count data dependent variables are still not many. Recent promising works are those of Lambert et al. (2010) and Castro et al. (2012). The former developed a two step limited information maximum likelihood (LIML) estimator for a spatial autoregressive Poisson model, with small sample properties evaluated using by Monte Carlo simulations. The latter proposed a spatial lag count model with temporal dependence in a generalized ordered response context, introducing

spatial dependences by using a spatial structure on the latent continuous variables and time-varying temporal correlation patterns by means of an appropriate structure for the error term of the latent variable. The estimation procedure is based on composite marginal likelihood in Bhat (2011).

Finally, it should be recognized the recent important contributions by Xu and Lee (2015), Qu and Lee (2012) and Qu and fei Lee (2012) on the estimation of Spatial Autoregressive Tobit models. In particular, the first one analysed the asymptotic properties based on the spatial near–epoch dependence of the dependent variable process, see Jenish (2012) Jenish and Prucha (2009) Jenish and Prucha (2012), of the maximum likelihood estimator. Finite sample properties of the estimator are also included. Whereas the second and the third focused on the asymptotic and finite sample properties of LM test statistics for the spatial simultaneous autoregressive Tobit model.

3.2. Bayesian analysis

Because of the apparent computational advantages, Bayesian techniques have received an increasing attention in several applied research fields, especially those related to agricultural and land use issues. For instance, some of the followings made use of binary variables (Holloway et al., 2002), (Holloway et al., 2007), ordered responses (Wang and Kockelman, 2009b), (Wang and Kockelman, 2009a), unordered responses (Chakir and Parent, 2009) and count data (Rathbun and Fei, 2006), (Ver Hoef and Jansen, 2007). Anyway, it should be clarified that the use of Bayesian inference should not be generally preferred to a frequentist approach without a justified statistical reason. In order to briefly explain, in all cases in which a non-correct prior distribution is chosen, the estimates may give misleading results, and in most empirical cases we generally do not have sufficient information to define a proper prior. Many uninformative or diffuse priors have been surely proposed, but for those priors we generally expect that “*the likelihood will dominate the prior*” (i.e. the likelihood function will provide the significant part of information). Being Bayesian inference a different “way” to view the estimation of parameters, a comparison with MSL and other kind of estimators is necessary to make us sure of our results. For example, in a comparison between MSL estimator and the Gibbs sampling approach Bolduc et al. (1997), no significant differences were found. Moreover, recently in LeSage and Pace (2009) it has emerged that Bayesian MCMC requires extensive simulation, is time-consuming, is not straightforward to implement, it can create convergence assessment problems and, therefore, it seems does not have particular advantages on MSL-based estimators. However, as LeSage (2014) stressed, a Bayesian approach can be used in many situations where a prior knowledge (for example on the well-known W matrix) is required.

3.3. The problem of inconsistency

Although a long list of reasons would justify the use of spatial autoregressive models, the one we are dealing with is the inconsistency problem of the standard probit estimators.

The error term in a simple probit model summarizes the unknown information coming from other regressors (i.e. omitted variables) which we assume to be uncorrelated with those in \mathbf{X}_n . In this case, extremum estimators, such as likelihood based estimators, are consistent, see Amemiya (1977), Amemiya (1978) and Amemiya (1985). However, unknown forms of misspecification of the functional form (Yatchew and Griliches, 1985), for example when heteroskedastic errors are incorrectly assumed to be homoskedastic, lead to inconsistency of the maximum likelihood estimators in a nonlinear setting (Poirier and Ruud, 1988). Indeed, MLE is consistent if the conditional density of $\mathbf{y}_n|\mathbf{X}_n$ is correctly specified. Misspecification of the functional form in a probit context is equivalent to have a misspecification of the Bernoulli probability for each $y_i, 1 \leq i \leq n$.

In a SAE(1)–probit setting, heteroskedasticity will arise whenever the weights \mathbf{M}_n induce non–constant diagonal terms of the matrix $\boldsymbol{\Sigma}_{\mathbf{u}} = [\mathbf{B}'_{\lambda}\mathbf{B}_{\lambda}]^{-1}$. Indeed, this usually happens even for rather *simple* choices of \mathbf{M}_n , such as a k -nearest neighbor matrix. Heteroskedastic probit estimators (Case, 1992) that explicitly consider the diagonal elements of the variance-covariance matrix, i.e. $\text{diag}(\boldsymbol{\Sigma}_{\mathbf{u}}) = \text{diag}[\mathbf{B}'_{\lambda}\mathbf{B}_{\lambda}]^{-1}$, remain consistent. However, the form of heteroskedasticity is generally unknown if it is implied by the spatial autocorrelation coefficient, see McMillen (1995) and Pinkse and Slade (1998).

In the general case, let $\mathbf{A}_{\rho} = (\mathbf{I}_n - \rho\mathbf{W}_n)$ and $\mathbf{B}_{\lambda} = (\mathbf{I}_n - \lambda\mathbf{M}_n)$. So we get

$$\begin{aligned} \mathbf{y}_n^* &= \rho\mathbf{W}_n\mathbf{y}_n^* + \mathbf{X}_n\boldsymbol{\beta} + \mathbf{u}_n, & \mathbf{B}_{\lambda}\mathbf{u}_n &= \boldsymbol{\varepsilon}_n \\ \mathbf{B}_{\lambda}\mathbf{y}_n^* &= \rho\mathbf{B}_{\lambda}\mathbf{W}_n\mathbf{y}_n^* + \mathbf{B}_{\lambda}\mathbf{X}_n\boldsymbol{\beta} + \boldsymbol{\varepsilon}_n \\ \mathbf{y}_n^* &= \lambda\mathbf{M}_n\mathbf{y}_n^* + \rho\mathbf{B}_{\lambda}\mathbf{W}_n\mathbf{y}_n^* + \mathbf{B}_{\lambda}\mathbf{X}_n\boldsymbol{\beta} + \boldsymbol{\varepsilon}_n, & \boldsymbol{\varepsilon}_n &\sim \mathcal{N}_n(\mathbf{0}_n, \boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}) \end{aligned} \quad (7)$$

which is known as the Cochrane–Orcutt type transformation (Cochrane and Orcutt, 1949), a model in which the resulting disturbances are innovations. Even after the Cochrane–Orcutt transformation, both $\mathbf{W}_n\mathbf{y}_n^*$ and $\mathbf{M}_n\mathbf{y}_n^*$ are correlated with $\boldsymbol{\varepsilon}_n$ because

$$\mathbb{E}[\mathbf{y}_n^*\boldsymbol{\varepsilon}_n'] = \mathbf{A}_{\rho}^{-1}\mathbb{E}[\mathbf{u}_n\boldsymbol{\varepsilon}_n'] = \mathbf{A}_{\rho}^{-1}\mathbf{B}_{\lambda}^{-1} \quad (8)$$

and these correlations rule out the use of nonlinear least squares methods due to their inconsistency. For the SARAR(1,1)–probit model in equation (1), and its sub–specification SAR(1)–probit by letting $\lambda = 0$, we have $\mathbb{E}((\mathbf{W}_n\mathbf{y}_n^*)\boldsymbol{\varepsilon}_n') \neq \mathbf{0}_n$ where $\mathbf{u}_n = \mathbf{B}_{\lambda}^{-1}\boldsymbol{\varepsilon}_n$ and $\mathbb{E}((\mathbf{W}_n\mathbf{y}_n^*)\boldsymbol{\varepsilon}_n') \neq \mathbf{0}_n$, respectively, see Kelejian and Prucha (1998) and Kelejian and Prucha (1999) in the linear case. Therefore, consistency can only be achieved by correctly specifying the conditional expected value of model in equation (1).

3.4. Estimation procedures in R

The estimation procedures developed in spatial econometrics are gradually spreading out in the R language, see e.g. Arbia (2014). With respect to discrete choice models, we found the **McSpatial** package to be useful in estimating spatial binary probit models with both the MLE and the Linearized GMM proposed by

Klier and McMillen (2008). The **mvProbit** package used the GHK algorithm to numerically approximate the multidimensional integral, which is unfortunately computationally unfeasible. A fast approximated ML procedure is proposed by Martinetti and Geniaux (2017) with their package **ProbitSpatial**. Finally, within the Bayesian estimation, we recognise the **spatialprobit** package.

4. Spatial Discrete Choice models in Health

Although a large number of papers dealing with limited dependent variable (LDV) and discrete choice (DC) models with empirical applications in health economics can be found, those models with an explicit reference to space and spatial relationships are not so common in the literature mainly because of the peculiarities and the micro-scale of health data. Modeling economic agent-based spatial relationships will be instead an approaching problem to be solved since that individual decisions usually depend upon neighboring agents' decisions. As an example, in the realm of land use modeling, both Sidharthan and Bhat (2012) and Ferdous and Bhat (2013) have recently stressed that spatial dependence among land development intensity levels is justified by the interactions between land owners of the corresponding spatial units. That is, land owners of proximately located spatial units, acting as profit-maximizing economic agents, are likely to be influenced by each other's perceptions of net stream of returns from land use development. The same dynamics can occur between economic agents in the demand for health care utilization.

Observational data are though vulnerable to biases in estimating effects due to non-random selection and confounding that are avoided in randomized experimental data. In most cases the above-mentioned peculiarities of health data make us unable to correctly use the econometric techniques which then differ according to different observed data and they are continuously subject to criticisms and improvements by researchers⁴.

In the last 20 years, we have had an increased experience in econometric studies as basis of health policy. As already said, most of them required the use of LDV or DC models to describe health care expenditures, treatment effects analysis, etc., see e.g. Varin and Czado (2009), Munkin and Trivedi (2008), Santos et al. (2017), Varkevisser et al. (2012), Lindeboom and Kerkhofs (2009), Deb et al. (2006) and Basu et al. (2007). However, a very limited number of papers take into account space and spatial structure of discrete health data sets. Some empirical works considered a distance variable or a spatial dummy variable to distinguish between districts/regions, see e.g. Geweke et al. (2003), Wolff et al. (2008) and Nketiah-Amponsah (2009), but none of them used spatial spillover effects by introducing autocorrelation coefficients.

Bolduc et al. (1996a) yet developed a spatial autoregressive multinomial probit model. The hybrid MNP model approximated the correlation among the utilities of the different locations using a first-order spatial autoregressive [SAR(1)] process based on a distance decaying relationship. They used a maximum simulated likelihood (MSL) estimation to describe the effect of various incentive measures introduced in Quebec (Canada)

⁴See for example Madden (2008) and Norton et al. (2008) for comprehensive debates.

to influence the geographical distribution of physicians across 18 regions. Other more recent contributions are those of Bukenya et al. (2003) who proposed a spatial ordered probit model to examine the relationship existing between quality of life (QOL), health and several socioeconomic variables; Bhat and Sener (2009) instead used a spatial binary logit model to study teenagers' physical activity participation levels, a subject of considerable interest in the Public Health as well as in other fields. Similarly, Sener et al. (2010) proposed a spatial ordered-response model to estimate physical activity participation levels by including unobserved dependences inside clusters of observations (i.e. family units) which affects those participation levels and, in the same way, Sener and Bhat (2012) extended a multinomial model to take spatial effects into the error terms with the motivation that it is likely the presence of unobserved residential urban form factors (such as good bicycle and walk path continuity) which may increase participation tendencies in specific activity and unobserved lifestyle perspectives (such as physically active lifestyle attitudes) that affect activity participation decisions based on the proximity of teenagers' residences.

In studies on the demand of health care, health care utilizations, and more in general in all cases in which we need to describe the individual choices of the health services among different alternatives, it is generally reasonable to assume that there are unobserved factors, which are correlated among geographical units or among individuals who are proximal in space, as it is frequently the case in health. For instance, in Nketiah-Amponsah (2009) it is likely that unobserved factors such as tobacco and alcohol consumptions are correlated over space since it is almost certain that individuals, especially among the youngest, have social interaction effects on the others who live in the neighbor areas. This correlation information can be taken into account by specifying a simple spatial error structure of the discrete choice model, which can be used to improve coefficient estimates by avoiding inconsistent estimates⁵ and leading to a correct inference approach. In Bolduc et al. (1996b), individuals' utility functions, which are generally described by individuals' choices of health care services, can be autocorrelated in space due to social interaction effects between those individuals who are proximal in space, that is individuals' choices are also likely to be determined by neighbor individual opinions. Also in child labor (Wolff et al., 2008) and child mortality (Iram and Butt, 2008) studies, it is not to exclude the possibility of specifying a spatial autoregressive discrete choice structure, since it is reasonable to assume the presence of autocorrelation between health status among children who lie within the same neighborhood. In the same way, it could be interesting to see if there are spatial interaction effects among child labor choices taken by individuals (i.e. parents' choices) who are in the closeness. Gravelle et al. (2014) used a spatial autoregressive-regressive (SAR) model in order to detect if a hospital's quality level depends on its rivals' quality levels in a competitive setting. The main finding was that hospitals' quality levels are positively autocorrelated over space and then geographical proximity plays a paramount role in describing hospitals' competition. One way in which hospitals can raise their quality is surely the adoption of advanced technologies. In this context,

⁵The problem of inconsistency in spatial binary nonlinear models is referred to Section 3.3.

it is then reasonable to assume that hospitals closed in space, which are competitive in terms of technology adoptions, share information about the quantity and the quality of their qualified technologies, which in turn have an impact on the hospitals' attractive potential of patients.

Spatial dependence is indeed inherent in many aspects of human-decision-making, with the choice decisions of one individual being affected by those of other individuals who are proximal in space. The importance of such spatial dependence has been recognized in a variety of disciplines. In recent years, it has become more common to include social interactions or neighborhood effects (i.e. social network effects) also in discrete choice models (see Goetzke and Andrade (2010) Li and Lee (2009) Páez et al. (2008) Brock and Durlauf (2007)). In particular, Goetzke and Andrade (2010) stressed the need to include social interactions and correlated effects in mode choice models as one combined spatial spillover variable for two reasons: spatial spillover serves the purpose to avoid a possible omitted variable bias, and, in addition, the spatial spillover variable can be seen as a proxy for the mode-friendliness in the neighborhood. As also Rosenquist and Lehrer (2014) stressed, if such influences are ignored estimates of the impact of policy interventions will in many cases be biased because they neglect the indirect pathway that occurs due to spillovers or what is known as the social multiplier effects. This should be a tempting prospect also in applied health economics, even more so microeconomic geo-referred data will become more and more available in the near future.

5. Conclusions

Accounting for spatial autocorrelation in the discrete or limited dependent variable is a fundamental challenge in the econometrics literature. One of the most important reasons for the relatively scarce diffusion of these models is certainly their complexity, often requiring MSL or Bayesian algorithms to estimate them. To this purpose, some methodological and computational solutions have been proposed, but the aim of developing "best" estimators is still unreached. As expected, we have found that only a small number of papers use the above-mentioned models with the purpose to solve health economics issues, emerging the need to first understand the potential of these models in this applied field. This was one of the aims of the present paper.

Most of the sample sizes used in health empirical applications are of the order of millions of observations because of their micro-scale nature. Indeed, Bell and Dalton (2007) highlighted the problem of specifying a weighting matrix for micro-scale or individual data, in which the difficulties are related to correctly describe all the relationships among economic agents. Data of this kind will become more and more available in the near future and probably we will not be ready to manage them or, even worse, to correctly use the whole amount of information. Filling this gap in the literature will surely lay the foundations for the development of *Spatial Microeconometrics*.

- Amemiya, T. (1977). The maximum likelihood and the nonlinear three-stage least squares estimator in the general nonlinear simultaneous equation model. *Econometrica: Journal of the Econometric Society*, pages 955–968.
- Amemiya, T. (1978). The estimation of a simultaneous equation generalized probit model. *Econometrica: Journal of the Econometric Society*, pages 1193–1205.
- Amemiya, T. (1985). *Advanced Econometrics*. Harvard university press.
- Anselin, L. (2002). Under the hood: Issues in the specification and interpretation of spatial regression models. *Agricultural economics*, 27(3):247–267.
- Arbia, G. (2014). *A primer for spatial econometrics: with applications in R*. Springer.
- Arbia, G., Espa, G., Giuliani, D., and Dickson, M. M. (2014). Spatio-temporal clustering in the pharmaceutical and medical device manufacturing industry: A geographical micro-level analysis. *Regional Science and Urban Economics*, 49:298–304.
- Arduini, T. (2016). Distribution free estimation of spatial autoregressive binary choice panel data models.
- Atella, V., Belotti, F., Depalo, D., and Mortari, A. P. (2014). Measuring spatial effects in the presence of institutional constraints: The case of italian local health authority expenditure. *Regional Science and Urban Economics*, 49:232–241.
- Baltagi, B. H., Egger, P. H., and Kesina, M. (2016). Bayesian spatial bivariate panel probit estimation. In *Spatial Econometrics: Qualitative and Limited Dependent Variables*, pages 119–144. Emerald Group Publishing Limited.
- Baltagi, B. H., Moscone, F., and Tosetti, E. (2012). Medical technology and the production of health care. *Empirical Economics*, 42(2):395–411.
- Basu, A., Heckman, J. J., Navarro-Lozano, S., and Urzua, S. (2007). Use of instrumental variables in the presence of heterogeneity and self-selection: An application to treatments of breast cancer patients. *Health economics*, 16(11):1133–1157.
- Bekhor, S. and Prashker, J. N. (2008). Gev-based destination choice models that account for unobserved similarities among alternatives. *Transportation Research Part B: Methodological*, 42(3):243–262.
- Bell, K. P. and Dalton, T. J. (2007). Spatial economic analysis in data-rich environments. *Journal of Agricultural Economics*, 58(3):487–501.
- Beron, K. J., Murdoch, J. C., and Vijverberg, W. P. (2003). Why cooperate? public goods, economic power, and the montreal protocol. *Review of Economics and Statistics*, 85(2):286–297.
- Beron, K. J. and Vijverberg, W. P. M. (2004). *Advances in Spatial Econometrics: Methodology, Tools and Applications*, chapter Probit in a Spatial Context: A Monte Carlo Analysis, pages 169–195. Springer Berlin Heidelberg, Berlin, Heidelberg.
- Bhat, C. R. (2011). The maximum approximate composite marginal likelihood (macml) estimation of multinomial probit-based unordered response choice models. *Transportation Research Part B: Methodological*, 45(7):923–939.
- Bhat, C. R., Dubey, S. K., Alam, M. J. B., and Khushefati, W. H. (2015). A new spatial multiple discrete-continuous modeling approach to land use change analysis. *Journal of Regional Science*, 55(5):801–841.
- Bhat, C. R. and Guo, J. (2004). A mixed spatially correlated logit model: formulation and application to residential choice modeling. *Transportation Research Part B: Methodological*, 38(2):147–168.
- Bhat, C. R. and Sener, I. N. (2009). A copula-based closed-form binary logit choice model for accommodating spatial correlation across observational units. *Journal of Geographical Systems*, 11(3):243–272.
- Billé, A. G. (2014). Computational issues in the estimation of the spatial probit model: A comparison of various estimators. *The Review of Regional Studies*, 43(2, 3):131–154.
- Billé, A. G., Benedetti, R., and Postiglione, P. (2017). A two-step approach to account for unobserved spatial heterogeneity. *Spatial Economic Analysis*, 0(0):1–20.
- Billé, A. G. and Leorato, S. (2017). Quasi-ml estimation, marginal effects and asymptotics for spatial autoregressive nonlinear models. Technical report, Faculty of Economics and Management at the Free University of Bozen.
- Bolduc, D., Fortin, B., and Fournier, M.-A. (1996a). The effect of incentive policies on the practice location of doctors: a multinomial probit analysis. *Journal of labor economics*, 14(4):703–732.

- Bolduc, D., Fortin, B., and Gordon, S. (1997). Multinomial probit estimation of spatially interdependent choices: an empirical comparison of two new techniques. *International Regional Science Review*, 20(1-2):77–101.
- Bolduc, D., Lacroix, G., and Muller, C. (1996b). The choice of medical providers in rural benin: a comparison of discrete choice models. *Journal of health economics*, 15(4):477–498.
- Brock, W. A. and Durlauf, S. N. (2007). Identification of binary choice models with social interactions. *Journal of Econometrics*, 140(1):52–75.
- Bukenya, J. O., Gebremedhin, T. G., and Schaeffer, P. V. (2003). Analysis of rural quality of life and health: A spatial approach. *Economic Development Quarterly*, 17(3):280–293.
- Case, A. (1992). Neighborhood influence and technological change. *Regional Science and Urban Economics*, 22(3):491–508.
- Castro, M., Paleti, R., and Bhat, C. R. (2012). A latent variable representation of count data models to accommodate spatial and temporal dependence: Application to predicting crash frequency at intersections. *Transportation research part B: methodological*, 46(1):253–272.
- Castro, M., Paleti, R., and Bhat, C. R. (2013). A spatial generalized ordered response model to examine highway crash injury severity. *Accident Analysis & Prevention*, 52:188–203.
- Chakir, R. and Parent, O. (2009). Determinants of land use changes: A spatial multinomial probit approach. *Papers in Regional Science*, 88(2):327–344.
- Cochrane, D. and Orcutt, G. H. (1949). Application of least squares regression to relationships containing auto-correlated error terms. *Journal of the American Statistical Association*, 44(245):32–61.
- Deb, P., Munkin, M. K., and Trivedi, P. K. (2006). Bayesian analysis of the two-part model with endogeneity: application to health care expenditure. *Journal of Applied Econometrics*, 21(7):1081–1099.
- Elhorst, J. P. (2010). Applied spatial econometrics: raising the bar. *Spatial Economic Analysis*, 5(1):9–28.
- Ferdous, N. and Bhat, C. R. (2013). A spatial panel ordered-response model with application to the analysis of urban land-use development intensity patterns. *Journal of Geographical Systems*, 15(1):1–29.
- Fleming, M. M. (2004). Techniques for estimating spatially dependent discrete choice models. In *Advances in spatial econometrics*, pages 145–168. Springer.
- Geweke, J., Gowrisankaran, G., and Town, R. J. (2003). Bayesian inference for hospital quality in a selection model. *Econometrica*, 71(4):1215–1238.
- Goetzke, F. and Andrade, P. M. (2010). Walkability as a summary measure in a spatially autoregressive mode choice model: an instrumental variable approach. In *Progress in Spatial Analysis*, pages 217–229. Springer.
- Gravelle, H., Santos, R., and Siciliani, L. (2014). Does a hospital’s quality depend on the quality of other hospitals? a spatial econometrics approach. *Regional science and urban economics*, 49:203–216.
- Greene, W., Harris, M. N., Hollingsworth, B., and Weterings, T. A. (2014). Heterogeneity in ordered choice models: A review with applications to self-assessed health. *Journal of Economic Surveys*, 28(1):109–133.
- Groot, W. (2000). Adaptation and scale of reference bias in self-assessments of quality of life. *Journal of health economics*, 19(3):403–420.
- Hamilton, J. D. (1994). *Time series analysis*, volume 2. Princeton university press Princeton.
- Holloway, G., Lapar, M., and Lucila, A. (2007). How big is your neighbourhood? spatial implications of market participation among filipino smallholders. *Journal of Agricultural Economics*, 58(1):37–60.
- Holloway, G., Shankar, B., and Rahman, S. (2002). Bayesian spatial probit estimation: a primer and an application to hyv rice adoption. *Agricultural Economics*, 27(3):383–402.
- Hunt, L. M., Boots, B., and Kanaroglou, P. S. (2004). Spatial choice modelling: new opportunities to incorporate space into substitution patterns. *Progress in Human Geography*, 28(6):746–766.
- Iram, U. and Butt, M. S. (2008). Socioeconomic determinants of child mortality in pakistan: Evidence from sequential probit

- model. *International Journal of Social Economics*, 35(1/2):63–76.
- Jenish, N. (2012). Nonparametric spatial regression under near-epoch dependence. *Journal of Econometrics*, 167(1):224 – 239.
- Jenish, N. and Prucha, I. R. (2009). Central limit theorems and uniform laws of large numbers for arrays of random fields. *Journal of Econometrics*, 150(1):86 – 98.
- Jenish, N. and Prucha, I. R. (2012). On spatial processes and asymptotic inference under near-epoch dependence. *Journal of Econometrics*, 170(1):178 – 190.
- Jones, A. (2007). *Applied econometrics for health economists: A practical guide*, volume 2. Radcliffe Publishing.
- Jones, A. M. (2000). Health econometrics. In *Handbook of health economics*, volume 1, pages 265–344. Elsevier.
- 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. *The Journal of Real Estate Finance and Economics*, 17(1):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(2):509–533.
- Klier, T. and McMillen, D. P. (2008). Clustering of auto supplier plants in the united states: generalized method of moments spatial logit for large samples. *Journal of Business & Economic Statistics*, 26(4):460–471.
- Lacombe, D. J. and LeSage, J. P. (2013). Use and interpretation of spatial autoregressive probit models. *The Annals of Regional Science*, pages 1–24.
- Lambert, D. M., Brown, J. P., and Florax, R. J. (2010). A two-step estimator for a spatial lag model of counts: Theory, small sample performance and an application. *Regional Science and Urban Economics*, 40(4):241–252.
- LeSage, J. (2014). What regional scientists need to know about spatial econometrics. *The Review of Regional Studies*, 44(1):13–32.
- LeSage, J. and Pace, R. K. (2009). Introduction to spatial econometrics. *Boca Raton, FL: Chapman & Hall/CRC*.
- LeSage, J. P. (2000). Bayesian estimation of limited dependent variable spatial autoregressive models. *Geographical Analysis*, 32(1):19–35.
- LeSage, J. P., Kelley Pace, R., Lam, N., Campanella, R., and Liu, X. (2011). New orleans business recovery in the aftermath of hurricane katrina. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 174(4):1007–1027.
- Li, J. and Lee, L.-f. (2009). Binary choice under social interactions: an empirical study with and without subjective data on expectations. *Journal of Applied Econometrics*, 24(2):257–281.
- Lindeboom, M. and Kerkhofs, M. (2009). Health and work of the elderly: subjective health measures, reporting errors and endogeneity in the relationship between health and work. *Journal of Applied Econometrics*, 24(6):1024–1046.
- Lindeboom, M. and Van Doorslaer, E. (2004). Cut-point shift and index shift in self-reported health. *Journal of health economics*, 23(6):1083–1099.
- Madden, D. (2008). Sample selection versus two-part models revisited: The case of female smoking and drinking. *Journal of health economics*, 27(2):300–307.
- Manski, C. (1981). *Alternative Estimators and Sample Designs for Discrete Choice Analysis*. The MIT Press.
- Martinetti, D. and Geniaux, G. (2017). Approximate likelihood estimation of spatial probit models. *Regional Science and Urban Economics*, 64:30 – 45.
- McFadden, D. (2001). Economic choices. *American Economic Review*, pages 351–378.
- McMillen, D. P. (1992). Probit with spatial autocorrelation. *Journal of Regional Science*, 32(3):335–348.
- McMillen, D. P. (1995). Selection bias in spatial econometric models. *Journal of Regional Science*, 35(3):417–436.
- McMillen, D. P. and McDonald, J. F. (2004). Locally weighted maximum likelihood estimation: Monte carlo evidence and an application. In *Advances in spatial econometrics*, pages 225–239. Springer.
- Moscone, F. and Tosetti, E. (2014). Spatial econometrics: Theory and applications in health economics. In *Encyclopedia of Health Economics*, pages 329–334. Elsevier.
- Moscone, F., Tosetti, E., and Vittadini, G. (2012). Social interaction in patients hospital choice: evidence from italy. *Journal of*

- the Royal Statistical Society: Series A (Statistics in Society)*, 175(2):453–472.
- Mozharovskiy, P. and Vogler, J. (2016). Composite marginal likelihood estimation of spatial autoregressive probit models feasible in very large samples. *Economics Letters*, 148:87–90.
- Munkin, M. K. and Trivedi, P. K. (2008). Bayesian analysis of the ordered probit model with endogenous selection. *Journal of Econometrics*, 143(2):334–348.
- Nketiah-Amponsah, E. (2009). Demand for health insurance among women in Ghana: Cross sectional evidence. *International Research Journal of Finance and Economics*, 33(1).
- Norton, E. C., Dow, W. H., and Do, Y. K. (2008). Specification tests for the sample selection and two-part models. *Health Services and Outcomes Research Methodology*, 8(4):201–208.
- Pace, R. K. and Barry, R. (1997). Sparse spatial autoregressions. *Statistics & Probability Letters*, 33(3):291–297.
- Páez, A., Scott, D. M., and Volz, E. (2008). A discrete-choice approach to modeling social influence on individual decision making. *Environment and Planning B: Planning and Design*, 35(6):1055–1069.
- Pinjari, A. R. (2011). Generalized extreme value (gev)-based error structures for multiple discrete-continuous choice models. *Transportation Research Part B: Methodological*, 45(3):474–489.
- Pinkse, J., Slade, M., and Shen, L. (2006). Dynamic spatial discrete choice using one-step gmm: an application to mine operating decisions. *Spatial Economic Analysis*, 1(1):53–99.
- Pinkse, J. and Slade, M. E. (1998). Contracting in space: An application of spatial statistics to discrete-choice models. *Journal of Econometrics*, 85(1):125–154.
- Poirier, D. J. and Ruud, P. A. (1988). Probit with dependent observations. *The Review of Economic Studies*, 55(4):593–614.
- Qu, X. and fei Lee, L. (2012). Lm tests for spatial correlation in spatial models with limited dependent variables. *Regional Science and Urban Economics*, 42(3):430 – 445. Special Section on Asian Real Estate Market.
- Qu, X. and Lee, L.-F. (2012). Lm tests for spatial correlation in spatial models with limited dependent variables. *Regional Science and Urban Economics*, 42(3):430 – 445. Special Section on Asian Real Estate Market.
- Rathbun, S. L. and Fei, S. (2006). A spatial zero-inflated poisson regression model for oak regeneration. *Environmental and Ecological Statistics*, 13(4):409.
- Rosenquist, J. and Lehrer, S. (2014). Peer effects, social networks, and healthcare demand.
- Santos, R., Gravelle, H., and Propper, C. (2017). Does quality affect patients choice of doctor? evidence from England. *The Economic Journal*, 127(600):445–494.
- Sener, I. N. and Bhat, C. R. (2012). Flexible spatial dependence structures for unordered multinomial choice models: formulation and application to teenagers activity participation. *Transportation*, 39(3):657–683.
- Sener, I. N., Eluru, N., and Bhat, C. R. (2010). On jointly analyzing the physical activity participation levels of individuals in a family unit using a multivariate copula framework. *Journal of Choice Modelling*, 3(3):1–38.
- Sidharthan, R. and Bhat, C. R. (2012). Incorporating spatial dynamics and temporal dependency in land use change models. *Geographical Analysis*, 44(4):321–349.
- Smirnov, O. A. (2010). Modeling spatial discrete choice. *Regional science and urban economics*, 40(5):292–298.
- Smirnov, O. A. and Egan, K. J. (2012). Spatial random utility model with an application to recreation demand. *Economic Modelling*, 1(29):72–78.
- Varin, C. and Czado, C. (2009). A mixed autoregressive probit model for ordinal longitudinal data. *Biostatistics*, 11(1):127–138.
- Varkevisser, M., van der Geest, S. A., and Schut, F. T. (2012). Do patients choose hospitals with high quality ratings? empirical evidence from the market for angioplasty in the Netherlands. *Journal of Health Economics*, 31(2):371–378.
- Ver Hoef, J. M. and Jansen, J. K. (2007). Spacetime zero-inflated count models of harbor seals. *Environmetrics*, 18(7):697–712.
- Wang, H., Iglesias, E. M., and Wooldridge, J. M. (2013). Partial maximum likelihood estimation of spatial probit models. *Journal of Econometrics*, 172(1):77–89.

- Wang, X. and Kockelman, K. M. (2009a). Application of the dynamic spatial ordered probit model: Patterns of land development change in austin, texas. *Papers in Regional Science*, 88(2):345–365.
- Wang, X. and Kockelman, K. M. (2009b). Bayesian inference for ordered response data with a dynamic spatial-ordered probit model. *Journal of Regional Science*, 49(5):877–913.
- Weeks, M. (1997). The multinomial probit model revisited: A discussion of parameter estimability, identification and specification testing. *Journal of Economic Surveys*, 11(3):297–320.
- Wolff, F.-C. et al. (2008). Evidence on the impact of child labor on child health in indonesia, 1993–2000. *Economics & Human Biology*, 6(1):143–169.
- Xu, X. and Lee, L.-F. (2015). Maximum likelihood estimation of a spatial autoregressive tobit model. *Journal of Econometrics*, 188(1):264 – 280.
- Yatchew, A. and Griliches, Z. (1985). Specification error in probit models. *The Review of Economics and Statistics*, pages 134–139.