Journal of Econometrics 188 (2015) 264–280

Contents lists available at ScienceDirect

Journal of Econometrics

journal homepage: www.elsevier.com/locate/jeconom

Maximum likelihood estimation of a spatial autoregressive Tobit model

Xingbai Xu, Lung-fei Lee *

Department of Economics, The Ohio State University, Columbus, OH, United States

ARTICLE INFO

Article history:
Received 3 December 2013
Received in revised form
15 April 2015
Accepted 18 May 2015
Available online 9 June 2015

JEL classification:
C13
C21
C24
C63

Keywords:
Spatial autoregressive model
Tobit model
Censored data
Near-epoch dependence
Maximum likelihood

ABSTRACT

This paper examines a Tobit model with spatial autoregressive interactions. We consider the maximum likelihood estimation for this model and analyze asymptotic properties of the estimator based on the spatial near-epoch dependence of the dependent variable process generated from the model structure. We show that the maximum likelihood estimator is consistent and asymptotically normally distributed. Monte Carlo experiments are performed to verify finite sample properties of the estimator.

1. Introduction

The spatial autoregressive (SAR) model, $Y_n = \lambda W_n Y_n + X_n\beta + \epsilon_n$, has been extensively studied in spatial econometrics. Most of the early studies are summarized in Anselin (1988) and LeSage and Pace (2009). The two-stage least squares estimation is explored in Kelejian and Prucha (1998, 1999), while the generalized method of moments is studied in Lee (2007). The large sample properties of quasi-maximum likelihood estimation is considered in Lee (2004).

In recent years, there has been growing interest in nonlinear SAR models, as linear models cannot fully capture the characteristics of some types of data, such as censored or binary data. Jenish (2012) studies the nonparametric estimation of spatial near-epoch dependent (NED) random fields. An SAR model with a nonlinear transformation of the dependent variable is considered in Xu and Lee (2015). The smoothed maximum score estimation of binary choice panel models with spatial autoregressive errors can be found in Lei (2013). Qu and Lee (2015) investigate the estimation of an SAR model with an endogenous spatial weights matrix.

To study nonlinear SAR models and extend asymptotic properties of extreme estimators of nonlinear models with serial correlation (e.g. Gallant and White, 1988) to spatial correlation, some laws of large numbers (LLN) and central limit theorems (CLT) are needed. Jenish and Prucha (2009, 2012) have made fundamental contributions in this area by establishing LLN and CLT for spatial mixing and NED random fields.

In the microeconometric literature, the Tobit model has been widely studied since Tobin (1958) and Amemiya (1973). Some asymptotic properties of the maximum likelihood estimator (MLE) of the Tobit model are summarized in Amemiya (1985). Recently, an increasing number of studies have introduced spatial correlation into Tobit models. Some studies (e.g. Flores-Lagunes and Schnier, 2012) consider estimation or suggest tests of the spatial error Tobit model, but here we only review the literature on spatial autoregressive Tobit (SAR Tobit) models. In the literature, there are two types of SAR Tobit models (Qu and Lee, 2012): the simultaneous SAR Tobit model ($y_{i,n} = \max(0, \lambda_0 \sum_{j=1}^n w_{ij,n} y_{j,n} + x_{i,n}\beta_0 + \epsilon_{i,n})$), and the latent SAR Tobit model ($y_{i,n} = \max(0, y^*_{i,n})$, where $y^*_{i,n} = \lambda_0 \sum_{j=1}^n w_{ij,n} y^*_{j,n} + x_{i,n}\beta_0 + \epsilon_{i,n}$). The different interpretations for these two models are given in Qu and Lee (2012). So far, most studies have focused on the second type. LeSage (2000) and LeSage and Pace (2009) consider the Bayesian estimation of the latent SAR Tobit model. Marsh and Mittelhammer

* Corresponding author. Tel.: +1 614 292 5508.
E-mail addresses: xu.510@buckeyemail.osu.edu (X. Xu), lee.1777@osu.edu (L.-f. Lee).

http://dx.doi.org/10.1016/j.jeconom.2015.05.004
0304-4076/

(2004) study the performance of the generalized maximum entropy estimation of the SAR and the latent SAR Tobit models with Monte Carlo simulations. Testing of the existence of spatial correlation in the latent SAR Tobit model is carried out in Amaral and Anselin (2011) and Qu and Lee (2012). Donfouet et al. (2012) apply the latent SAR Tobit model to examine community-based health insurance using Bayesian estimation.

Compared to the latent SAR Tobit model, there are even fewer studies on the simultaneous SAR Tobit model. Autant-Bernard and LeSage (2011) consider the Bayesian estimation of this model and apply it to study knowledge spillovers. Qu and Lee (2012, 2013) examine the existence of spatial correlation in the simultaneous SAR Tobit model. To the best of our knowledge, there have been no formal studies on asymptotic properties of estimators of the simultaneous SAR Tobit model (we will call it SAR Tobit model for short, since we do not study the latent one). In this paper, we first show the NED properties of important variables generated by this model. Next, we establish the consistency and asymptotic normality of the MLE via the LLN and CLT developed in Jenish and Prucha (2012).

The structure of this paper is as follows: In Section 2, we introduce the SAR Tobit model and discuss its model coherency. In Section 3, we derive NED properties of the dependent variable and some other relevant functions of random variables. In Section 4, the identification of the SAR Tobit model and the consistency of its MLE are discussed. In Section 5, we establish the asymptotic normality of the estimator. In Section 6, we study finite sample properties and the robustness of the estimator using Monte Carlo experiments. All of the proofs for propositions and theorems are presented in the Appendices.

2. The spatial autoregressive Tobit model

The SAR Tobit model is motivated by two branches of literature in economics. One branch is concerned with peer effects from an exogenous social network. In such studies, each player chooses his/her effort, which is usually assumed to be nonnegative, and the Nash Equilibrium is exactly our model. See the detailed discussion below Eq. (2). The other branch is concerned with econometrics studies where a significant fraction of nonnegative data can be zero. Our model captures both features with a perfect information game framework in which an individual maximizes utility by choosing effort subject to a nonnegative constraint. A few empirical studies in the existing literature might appropriately be implemented by our model: (1) Rupasingha et al. (2004) investigate the environmental Kuznets curve for US counties. In the data set, facilities do not need to report to relevant environmental agents when they manufacture or process less than 25,000 pounds of a listed chemical during a year. Some counties do not have firms that individually meet these criteria. Thus, the pollutant data of these counties are censored. (2) Direct agriculture disaster payment relief for different states in US is studied in Marsh and Mittelhammer (2004), where some states do not receive any such payment (in certain years) and some spatial correlation exists. (3) LeSage (2009) examines origin–destination (OD) commuting flows from 60 districts in Toulouse, France. About 15% of the 3600 OD flows have zero values. (4) As is pointed out in Donfouet et al. (2012), community-based health insurance (CBHI) has increasing demand in rural areas in developing countries and households are more likely to pay for CBHI if other households in the same village are willing to do so. However, some households do not pay for the CBHI. (5) Spatial correlation among school district income tax rates in various school districts in Iowa is examined in Qu and Lee (2012, 2013).[1]

Let $\{(y_{i,n}, x_{i,n})\}_{i=1}^n$, where $y_{i,n}$ is censored such that $y_{i,n} \geqslant 0$ and $x_{i,n} \in \mathbb{R}^K$, be the sample we observe. We denote the position of individual (spatial unit) i as $s_i \in \mathbb{R}^d$, a point in the d-dimensional Euclidean space. For simplicity of notation, we also use i to represent s_i. As there are interactions among different individuals, we use an $n \times n$ matrix $W_n = (w_{ij,n})$ to represent their relative strength of direct interactions. If there is a potential direct interaction between individuals i and j, then $w_{ij,n} \neq 0$, or $w_{ji,n} \neq 0$ or both; zero, otherwise. As usual, a proper normalization has $w_{ii,n} = 0$ for all i.

$F(x) \equiv \max(0, x)$ is both a non-decreasing convex function and a Lipschitz function such that $|F(x_1) - F(x_2)| \leqslant |x_1 - x_2|$. The SAR Tobit model in this paper is specified as

$$y_{i,n} = F(\lambda_0 w_{i\cdot,n} Y_n + x_{i,n}\beta_0 + \epsilon_{i,n}), \tag{1}$$

where $w_{i\cdot,n}$ is the ith row of W_n. With generalized notations $\max(0, (x_1, \ldots, x_n)') = F((x_1, \ldots, x_n)') \equiv (\max(0, x_1), \ldots, \max(0, x_n))'$, the model can be written as

$$Y_n = \max(0, Y_n^*) = F(\lambda_0 W_n Y_n + X_n\beta_0 + \epsilon_n), \tag{2}$$

where $Y_n^* \equiv \lambda_0 W_n Y_n + X_n\beta_0 + \epsilon_n$. The model can be derived as a complete information game with each spatial unit (an agent) maximizing its linear–quadratic utility function subject to nonnegative constraints, given the actions of its links (see Ballester et al., 2006; Calvó-Armengol et al., 2009). Assume individual i's utility is $u(y_{1,n}, \ldots, y_{n,n}) = -y_{i,n}^2 + 2(\lambda_0 w_{i\cdot,n} Y_n + x_{i,n}\beta_0 + \epsilon_{i,n}) y_{i,n}$. When $\lambda_0 w_{i\cdot,n} Y_n + x_{i,n}\beta_0 + \epsilon_{i,n} > 0$, u is maximized at $y_{i,n} = \lambda_0 w_{i\cdot,n} Y_n + x_{i,n}\beta_0 + \epsilon_{i,n}$. When $\lambda_0 w_{i\cdot,n} Y_n + x_{i,n}\beta_0 + \epsilon_{i,n} \leqslant 0$, $\partial u/\partial y_{i,n} = -2y_{i,n} + 2(\lambda_0 w_{i\cdot,n} Y_n + x_{i,n}\beta_0 + \epsilon_{i,n}) < 0$ when $y_{i,n} > 0$, i.e., u is strictly decreasing when $y_{i,n} > 0$. Thus, u is maximized at $y_{i,n} = 0$. Hence, the Nash equilibrium is Eq. (2).

After a slight modification, this model can deal with the case where censoring points $c_{i,n}$'s are known and nonzero. With $c_{i,n}$ being the censoring points for i, an extended model can be $\tilde{Y}_n = \max\{C_n, \tilde{Y}_n^*\}$ where $\tilde{Y}_n^* = \lambda_0 W_n \tilde{Y}_n + X_n\beta_0 + \epsilon_n$. We can do a transformation, $Y_n = \tilde{Y}_n - C_n$, then the extended model can be rewritten as Eq. (1) with a trivial modification on the regressors, $Y_n^* = \lambda_0 Y_n + \lambda_0 W_n C_n - C_n + X_n\beta_0 + \epsilon_n$. But this paper cannot deal with the case studied in Nelson (1977) where $c_{i,n}$'s are unknown. An example of such a model is the female labor supply where $y_{i,n}$ is a market wage and $c_{i,n}$ is a reservation wage. The reservation wage can be modeled as another regression equation based on an individual's unobserved utility. We can observe market wage $y_{i,n}$ when $y_{i,n} > c_{i,n}$, but we will not observe $y_{i,n}$ while $y_{i,n} \leqslant c_{i,n}$ because those females will not work. Such a model is in the category of sample selection models. The censored model in this paper cannot handle sample selection models.

Because our model (2) is a system of nonlinear equations with censored dependent variables, it is necessary to discuss conditions for model coherency (Amemiya, 1974). Before doing so, we list our assumptions.

Assumption 1. Individual units in the economy are located or living in a region $D_n \subset D \subset R^d$, where the cardinality of D_n satisfies $\lim_{n\to\infty} |D_n| = \infty$. The distance $d(i, j)$ between any two different individuals i and j is larger than or equal to a specific positive constant, without loss of generality, say, 1.

Note that the space D can be a space of economic characteristics, a geographical space or a mixture of both economic and physical spaces. Correspondingly, the distance may refer to economic and/or physical distance induced from any norm on $\mathbb{R}^d$. Assumption 1 uses the increasing domains asymptotic and rules out the scenario of infilled asymptotic.[2] This setting is introduced in Jenish and Prucha (2009, 2012) for spatial mixing and NED processes.

[1] We complete their studies with the asymptotic normality and variance of the MLE, presented in a supplementary online file (see Appendix C).

[2] Under infilled asymptotic, even some popular estimators, such as the least squares and the method of moments may not be consistent, as noted in Lahiri (1996).