Estimation of Sample Selection Models with Spatial Dependence

Alfonso Flores-Lagunes*
Food and Resource Economics Department, IFAS, University of Florida
and IZA, Bonn, Germany
P.O. Box 110240, Gainesville, FL, 32611-0240.
Phone: 352-392-1826 ext. 307. Fax: 352-392-9898.
Email: alfonsofl@ufl.edu

Kurt Erik Schnier
Department of Economics
Andrew Young School of Policy Studies
Georgia State University
P.O. Box 3992, Atlanta, GA 30302-3992.
Phone: 404-413-0159. Fax: 404-413-0145.
Email: kschnier@gsu.edu

This version: February 2009

Abstract

We consider the estimation of a sample selection model that exhibits spatial autoregressive errors (SAE). Our methodology employs a two-step strategy in which the first step is based on a spatial probit model and in the second step a consistently estimated inverse Mills ratio (IMR) is included as a regressor in the outcome equation to control for selection bias. Since the appropriate IMR under SAE depends on a parameter from the second step, both steps are jointly estimated employing the generalized method of moments. We explore the finite sample properties of our estimator using simulations and provide an empirical illustration.

Key words and phrases: sample selection, spatial autoregressive errors, generalized method of moments.

JEL classification: C13, C15, C24, C49

*Corresponding Author

1 Introduction

Econometric models taking into account spatial interactions among economic units have been increasingly used by economists over the last several years.[1] The different approaches for undertaking estimation and inference in linear regression models with spatial dependence are well developed and have been summarized in the work by Anselin (1988, 2001), Anselin and Bera (1998), and others. In contrast, the estimation of nonlinear models that include spatial interactions, in particular limited dependent variable models, is not as well developed. In fact, only recently have methods for estimating and conducting statistical inference in spatial models with limited dependent variables been proposed. This literature concentrates mainly on the probit and logit models with spatial effects, as in Beron and Vijverberg (2004), Case (1992), Fleming (2004), Klier and McMillen (2008), LeSage (2000), McMillen (1992), and Pinkse and Slade (1998). In this paper we contribute to this literature by introducing an estimation method for a sample selection model with spatial autoregressive errors (SAE). The type of sample selection model considered is the widely used heckit model (Heckman, 1976, 1979), also known as the Tobit type II model in the terminology of Amemiya (1985).

The pervasiveness of sample selection problems in spatial data is well documented. For instance, McMillen (1995) discusses the issue within urban economics and regional science. His main example deals with data on land use and values in the city of Chicago during the 1920s (see references therein). In this case, unobserved variables that make a parcel more likely to receive residential zoning may increase the value of residential land. Other illustrations include models of housing prices, rent and tenure choice (Goodman, 1988), office rents and lease provisions (Benjamin *et al.*, 1992), and home improvement choice (Montgomery, 1992) in urban economics; the choice between central city and suburban employment (McMillen, 1993), and analysis of earnings and migration (Borjas *et al.*, 1992) in labor economics. In general, the increasing availability of geo-coded data increases the relevance of methods to deal with sample selection when spatial dependence is present.

Despite the pervasiveness of selection bias in the economics literature, we know of only one other paper that attempts specifying and estimating a sample selection model with spatial dependence. McMillen (1995) specifies a model similar to the heckit model with SAE presented in Section 2 below and outlines an estimator based on the expectation-maximization (EM) algorithm. However, the estimator is infeasible as it requires knowledge

[1] Some examples are Case (1991), Fishback *et al* (2006), Topa (2001), among many others.

of the true SAE parameters, and it is fairly computationally intensive. Alternatively, he specifies and estimates an extension of the spatial expansion model of Casetti (1972) that is used in geography. Nevertheless, that model is not explicitly spatial since additional variables are required to indirectly control for the spatial dependence, and the consistency of the estimates depends heavily on correctly specifying the functional form of the underlying heteroskedasticity induced by the spatial dependence. Compared to the pioneering work by McMillen (1995), we propose a feasible estimator for the heckit model that explicitly accounts for and estimates the SAE parameters.

Our estimation strategy can be thought of as a two-step procedure analogous to the heckit estimator. The first step is based on the spatial probit estimator by Pinkse and Slade (1998), which yields consistent– although not fully efficient–estimates of the selection equation. These estimates are used to compute the inverse Mills ratio (IMR) to be included in the estimation of the outcome equation to correct for selectivity bias. Since in the presence of SAE the IMR depends upon unknown parameters from the outcome equation, we propose to estimate the model jointly employing a "pseudo" sequential estimator (Newey, 1984) using the generalized method of moments (GMM).

Maximum likelihood estimation of a probit model with SAE involves a non-spherical variance-covariance matrix that renders the simple probit estimator inconsistent. In turn, to obtain consistent and fully efficient estimates, one has to deal with multidimensional integrals. LeSage (2000) and Beron and Vijverberg (2004) employ simulation methods to approximate these multidimensional integrals. Unfortunately, this simulation is computationally intensive, restricting estimation to only moderate sample sizes. This limitation also applies to the estimation of sample selection models with SAE using simulation methods to approximate the multidimensional integrals in the corresponding likelihood function.[2]

To avoid approximating multidimensional integrals and still achieve consistency of the probit estimates (at the expense of efficiency), some authors propose to ignore the off-diagonal elements of the variance-covariance matrix and focus on the heteroskedasticity induced by the spatial dependence (e.g. Case, 1992; McMillen, 1992; and Pinkse and Slade, 1998). We use Pinkse and Slade's estimator in the first step of the sample selection model for the following reasons. First, it yields consistent estimates of the selection equation that are necessary to obtain consistent estimates of the parameters in the outcome equation.

[2] We are not aware of any applications that estimate a sample selection model with SAE using multidimensional integration methods.

Second, it is computationally simpler than both of the other estimators that approximate multidimensional integrals. Third, it has been developed within the framework of GMM, the same we employ in the joint estimation of our sample selection model.

The consistent estimates obtained in the first step are then used to construct the IMR to be included in the outcome equation to correct for selectivity bias (Heckman, 1979). When the SAE parameters in the selection and outcome equations are different–which is the likely situation in practice–the appropriate IMR turns out to be a function of the unknown spatial parameter in the outcome equation. Therefore, in order to increase the efficiency of the estimator and to obtain its variance-covariance matrix directly, we propose to estimate all parameters of the model simultaneously employing the sequential estimation framework of Newey (1984). Building on Pinkse and Slade's (1998) asymptotic results for their spatial probit model and standard GMM theory, our proposed estimator is consistent, asymptotically normally distributed, and its covariance matrix can be estimated. Lastly, we caution that our estimator has lower efficiency compared to a maximum likelihood estimator that requires computationally intensive simulation methods for multidimensional integration. However, given the fact that the multidimensional integration is in the order of the number of observations in the sample, the relative computational simplicity of our method is preferable in many relevant instances when the amount of data available to researchers is reasonably large.

The paper is organized as follows. Section 2 presents the sample selection model with SAE to be estimated. Section 3 introduces our proposed method of estimation (the "spatial heckit"), states its large-sample properties, and discusses some aspects of its implementation. Section 4 presents the results of a Monte Carlo experiment to analyze the finite sample properties of our estimator, while Section 5 presents an empirical example to illustrate its application. Concluding remarks are provided in the last section.

2 The Sample Selection Models with Spatial Autoregressive Errors

Our focus is on the estimation of a sample selection (Tobit type II) model with spatial autoregressive errors (SAE) in both the selection and the outcome equations:

$$y_{1i}^* = \alpha_0 + x_{1i}'\alpha_1 + u_{1i}, \qquad u_{1i} = \delta \sum_{j \neq i} c_{ij} u_{1j} + \varepsilon_{1i} \tag{1}$$

$$y_{2i}^* = \beta_0 + x_{2i}'\beta_1 + u_{2i}, \qquad u_{2i} = \gamma \sum_{j \neq i} c_{ij} u_{2j} + \varepsilon_{2i} \tag{2}$$

where y_{1i}^* and y_{2i}^* $(i = 1, ...N)$ are latent variables with the following relationship with the observed variables: $y_{1i} = 1$ if $y_{1i}^* > 0$ and $y_{1i} = 0$ otherwise, and $y_{2i} = y_{2i}^* * y_{1i}$. Therefore, (1) is the selection equation while (2) is the outcome equation. Note that each of these equations exhibit spatial dependence, as u_{1i} and u_{2i} depend on other u_{1j} and u_{2j} $(j = 1, ...N)$ through their location in space, as given by the spatial weights c_{ij} and the spatial autoregressive parameters δ and γ. Typically, the spatial weights are specified by the econometrician based on some function of contiguity or (economic) distance (Anselin, 1988; Anselin and Bera, 1998). Note also that, in general, one will specify different spatial autoregressive parameters for the selection and outcome equations.[3] It is assumed that the errors ε_{1i} and ε_{2i} are *iid* $N(\mathbf{0}, \mathbf{\Sigma})$ with $\mathbf{\Sigma} = [\sigma_1^2 \;\; \sigma_{12}, \sigma_{12} \;\; \sigma_2^2]$.

The model in (1)-(2) can also be presented in a reduced form:

$$y_{1i}^* = \alpha_0 + x_{1i}'\alpha_1 + \sum_j \omega_{ij}^1 \varepsilon_{1j} \tag{3}$$

$$y_{2i}^* = \beta_0 + x_{2i}'\beta_1 + \sum_j \omega_{ij}^2 \varepsilon_{2j} \tag{4}$$

where the weights ω_{ij}^1 and ω_{ij}^2 are the (i, j) elements of the inverse matrices $(I - \delta C)^{-1}$ and $(I - \gamma C)^{-1}$, respectively, with C the matrix of spatial weights c_{ij}. Note that both sets of weights, ω_{ij}^1 and ω_{ij}^2, depend upon the unknown parameters δ and γ, respectively.

In the absence of any sample selection, equation (2) is just a linear model with SAE, for which a number of estimation methods exist. Maximum likelihood (ML) was suggested by Ord (1975) and rigorously analyzed by Lee (2004). It relies on normality and its computational demands increase with the sample size as the Jacobian of the likelihood function requires the determinant of a full matrix of dimension equal to the sample size (Anselin and

[3] Without loss of generality, we specify the same spatial weights in each of the two equations.

Bera, 1998; Kelejian and Prucha, 1999). Another approach is the three-step FGLS procedure of Kelejian and Prucha (1998) that we refer to as KP-SAE. First, residuals obtained with OLS are used to estimate the SAE parameter employing the "generalized moments" estimator in Kelejian and Prucha (1999), followed by applying OLS to a model transformed with a Cochrane-Orcutt type procedure.[4] Compared to ML, KP-SAE is asymptotically less efficient, although has been found to be "virtually as efficient" in simulations (Kelejian and Prucha, 1999), while being computationally simpler and not relying on normality.

Equation (1) is a probit model with SAE that introduce a fully non-spherical variance-covariance matrix that renders the probit estimator inconsistent. To obtain ML estimates that are consistent and asymptotically efficient (relying on normality), multidimensional integration on the order of the sample size is necessary. This is possible using simulation algorithms (LeSage, 2000; Beron and Vijverberg, 2004), but becomes infeasible even for moderate-size samples. Alternatively, less efficient methods account for the induced heteroskedasticity while ignoring the off-diagonal elements of the variance-covariance matrix to obtain consistent estimates. For instance, Case (1992) transforms a particular model structure to obtain homoskedastic errors;[5] McMillen (1992) uses the EM algorithm to account for the heteroskedasticity induced by SAE; and Pinkse and Slade (1998) propose a GMM estimator that also takes into account the heteroskedasticity induced by SAE. In our method below, we use Pinkse and Slade's (1998) estimator since it is computationally simple and allows us to use the GMM framework to estimate the full model in (1)-(2).

In the absence of SAE, (1)-(2) reduces to the standard Heckman's (1979) sample selection model that can be estimated with a two-step procedure (heckit) or ML. In the presence of SAE, however, neither of these two methods results in consistent estimates since the model will be affected by the same problems discussed above in the context of the probit model. Consistent and fully efficient estimates can only be obtained with methods that account for the full non-spherical variance-covariance matrix, such as ML with multidimensional integrals that are on the order of the sample size. We introduce in the next section a feasible estimator for large samples that achieves consistency by accounting for the heteroskedasticity induced by the SAE while ignoring the off-diagonal elements.

[4] The original procedure in Kelejian and Prucha (1998) is more general since it allows for spatial lag dependence as well.

[5] Unfortunately, Case's (1992) method is constrained to situations in which the population can be partitioned into groups (e.g. "districts") whose errors can be assumed independent.

3 Estimation of the Sample Selection Models with SAE

Our proposed estimation method for the sample selection model with SAE follows the intuition of the two-step procedure of Heckman (1976, 1979) but is estimated jointly employing the GMM. The selection equation is estimated using Pinkse and Slade's (1998) GMM estimator for the spatial probit model, while the outcome equation is estimated with Kelejian and Prucha's (1998) FGLS estimator (KP-SAE), although other methods can be employed as well. An estimate of the inverse Mills ratio is included in the outcome equation to correct for selectivity bias. To estimate these two parts simultaneously, the corresponding moment conditions are stacked and a GMM criterion function is minimized with respect to all parameters in the model.

From the model in (1)-(2), we start with the following calculations (McMillen, 1995):

$$var(u_{1i}) = \sigma_1^2 \sum_j (\omega_{ij}^1)^2 \tag{5}$$

$$var(u_{2i}) = \sigma_2^2 \sum_j (\omega_{ij}^2)^2 \tag{6}$$

$$E(u_{1i}, u_{2i}) = \sigma_{12} \sum_j \omega_{ij}^1 \omega_{ij}^2. \tag{7}$$

In the heckit model, a probit is employed to estimate the probability of each observation being included in the observed sample. The presence of SAE, however, induces heteroskedasticity in the error terms in (5), resulting in inconsistent probit estimates. Pinkse and Slade (1998) propose a consistent estimator by taking into account the known form of the induced heteroskedasticity.

Define $\theta_1 = \{\alpha_0, \alpha_1', \delta\}$ as the parameters in the spatial probit model, and $\psi_i(\theta_1) = \frac{\alpha_0 + x_{1i}'\alpha_1}{\sqrt{var(u_{1i})}}$ the index function of the probit model weighted by the standard deviation of the residual. The corresponding "generalized residuals" of this model are:

$$\tilde{u}_{1i}(\theta_1) = \sqrt{\sigma_1^2 \sum_j (\omega_{ij}^1)^2} \cdot \{y_{1i} - \Phi[\psi_i(\theta_1)]\} \cdot \frac{\phi[\psi_i(\theta_1)]}{\Phi[\psi_i(\theta_1)]\{1 - \Phi[\psi_i(\theta_1)]\}}. \tag{8}$$

The GMM estimates for θ_1 can be obtained as follows:

$$\hat{\theta}_{1,GMM} = \underset{\theta_1 \in \Theta_1}{\arg\min} \quad S_N(\theta_1)' M_N S_N(\theta_1) \tag{9}$$

where $S_N(\theta_1) = \frac{1}{N} z_N' \tilde{u}_{1N}(\theta_1)$, z_N is the data matrix of regressors, $\tilde{u}_{1N}(\theta_1)$ is the vector of generalized residuals with elements as shown in (8), and M_N is a positive definite matrix such that $M_N \xrightarrow{p} M$. Pinkse and Slade (1998) show that this estimator is consistent and asymptotically normal. One interpretation of the minimization problem in (9) is that $\hat{\theta}_{1,GMM}$ employs "optimal instruments" which are weighted averages of nonlinear functions of the regressors (e.g. Hansen, 1985; Chamberlain, 1987). These optimal instruments are given by the derivative of the moment conditions $S_N(\theta_1)$ with respect to θ_1.[6]

The consistent estimates of θ_1 are used in the construction of the inverse Mills ratio (IMR). Note that the conditional regression function for the outcome equation (2) has the following form (McMillen, 1995):

$$\begin{aligned} E[y_{2i}|y_{1i} &= 1] = \beta_0 + x_{2i}'\beta_1 + E[u_{2i}|u_{1i} > -(\alpha_0 + x_{1i}'\alpha_1)] \\ &= \beta_0 + x_{2i}'\beta_1 + \frac{E(u_{1i}, u_{2i})}{\sqrt{var(u_{1i})}} \cdot \frac{\phi\left[-\psi_i(\theta_1)\right]}{\{1 - \Phi\left[-\psi_i(\theta_1)\right]\}} \\ &= \beta_0 + x_{2i}'\beta_1 + \frac{\sigma_{12}\sum\limits_j \omega_{ij}^1\omega_{ij}^2}{\sqrt{\sigma_1^2\sum\limits_j(\omega_{ij}^1)^2}} \cdot \frac{\phi\left[-\psi_i(\theta_1)\right]}{\{1 - \Phi\left[-\psi_i(\theta_1)\right]\}} \\ &= \beta_0 + x_{2i}'\beta_1 + \frac{\sigma_{12}}{\sigma_1} \cdot \frac{\sum\limits_j \omega_{ij}^1\omega_{ij}^2}{\sqrt{\sum\limits_j(\omega_{ij}^1)^2}} \cdot \frac{\phi\left[-\psi_i(\theta_1)\right]}{\{1 - \Phi\left[-\psi_i(\theta_1)\right]\}} . \end{aligned}$$

Therefore, the selectivity correction implies the following "adjusted" IMR:

$$\lambda_i \equiv \frac{\sum\limits_j \omega_{ij}^1\omega_{ij}^2}{\sqrt{\sum\limits_j[\omega_{ij}^1]^2}} \cdot \frac{\phi\left[-\psi_i(\theta_1)\right]}{\{1 - \Phi\left[-\psi_i(\theta_1)\right]\}}. \tag{10}$$

Once estimated ($\hat{\lambda}_i$), the "adjusted" IMR is included as an additional regressor in the outcome equation, which in turn can be estimated with any of the spatial methods developed for this linear equation. However, note that the "adjusted" IMR in (10) depends on a parameter that is not estimated in the first step: γ, which is included in the weights ω_{ij}^2. In order to increase the efficiency of the estimator and directly obtain its variance-covariance matrix, we simultaneously estimate all parameters of the model by rewriting it as a sequential GMM estimator (Newey, 1984) composed of the Pinkse and Slade (1998) and the KP-SAE estimators. More specifically, we stack their corresponding moment conditions:

$$g(z_N, \theta) = [s(z_{1N}, \theta)', m(z_{2N}, \theta)']', \quad \theta = \{\alpha_0, \alpha_1', \delta, \beta_0, \beta_1', \mu, \gamma\}$$

[6] For this nonlinear model, the derivative of the moment conditions with respect to δ is non-zero, and thus this parameter is identified. This is contrary to the linear model as demonstrated in Kelejian and Prucha (1997).

with

$$\begin{aligned} s(z_{1N},\theta) &= z_{1N}'\tilde{u}_{1N}(\theta), \quad \tilde{u}_{1N}(\theta) \text{ as in (8)}, \\ m(z_{2N},\theta) &= [y_{1N}\cdot z_{2N}]'\tilde{u}_{2N}(\theta), \quad \tilde{u}_{2N}(\theta) = y_{2N} - \beta_0 - x_{2N}'\beta_1 - \mu\widehat{\lambda}_N(\delta,\gamma,\alpha_0,\alpha_1') \end{aligned}$$

where the subscript N denotes the corresponding vector or matrix of data, we have let $z_N' \equiv (z_{1N}', [y_{1N}\cdot z_{2N}]')'$, z_{1N} includes the regressors of the selection equation, and z_{2N} includes the regressors of the outcome equation plus the estimated "adjusted" IMR, which is represented as $\widehat{\lambda}_N(\delta,\gamma,\alpha_0,\alpha_1')$ to make explicit its dependence on other parameters of the model.[7]

Defining $\tilde{u}_N(\theta) \equiv (\tilde{u}_{1N}'(\theta), \tilde{u}_{2N}'(\theta))'$ then all parameters of the SAE sample selection model can be estimated as:

$$\hat{\theta}_{GMM} = \underset{\theta\in\Theta}{\arg\min} \;\; g_N(\theta)'M_N g_N(\theta) \tag{11}$$

where $g_N(\theta) = \frac{1}{N}z_N'\tilde{u}_N(\theta)$, for a conformable positive definite matrix M_N such that $M_N \xrightarrow{p} M$. We call $\hat{\theta}_{GMM}$ the "spatial heckit" estimator for the sample selection model with SAE.

Denote $g(\theta) \equiv \lim_{N\to\infty} E[g_N(\theta)]$ and let θ_0 be the true parameter vector. Under conditions similar to those in Pinkse and Slade (1998), $\hat{\theta}_{GMM}$ is consistent and asymptotically normal:[8]

$$\sqrt{N}(\hat{\theta}_{GMM} - \theta_0) \xrightarrow{d} N(\mathbf{0}, [\Psi_2(\theta_0)]^{-1}[\partial g'(\theta_0)/\partial\theta]M\Psi_1(\theta_0)M[\partial g(\theta_0)/\partial\theta'][\Psi_2(\theta_0)]^{-1})$$

where $\Psi_1(\theta_0) = \lim_{N\to\infty} E\{Ng_N(\theta_0)g_N(\theta_0)'\}$, and $\Psi_2(\theta_0) = [\partial g'(\theta_0)/\partial\theta]M[\partial g(\theta_0)/\partial\theta']$.

Furthermore, the asymptotic variance of $\hat{\theta}_{GMM}$ can be estimated:

$$\Psi_{1N}(\hat{\theta}_{GMM}) \xrightarrow{p} \Psi_1(\theta_0) \;\text{ and }\; \Psi_{2N}(\hat{\theta}_{GMM}) \xrightarrow{p} \Psi_2(\theta_0), \;\text{ where}$$

$$\begin{aligned} \Psi_{1N}(\hat{\theta}_{GMM}) &= NE\{g_N(\hat{\theta}_{GMM})g_N(\hat{\theta}_{GMM})'\} \text{ and} \\ \Psi_{2N}(\hat{\theta}_{GMM}) &= [\partial g_N'(\hat{\theta}_{GMM})/\partial\theta]M_N[\partial g_N(\hat{\theta}_{GMM})/\partial\theta']. \end{aligned}$$

The asymptotically efficient $\hat{\theta}_{GMM}$ within the class of GMM estimators is obtained when the optimal GMM weighting matrix is used. That is, if $M_N = [\Psi_{1N}(\hat{\theta}_{GMM})]^{-1}$ is chosen,

[7] It is well-known that in a linear model like the outcome equation, the first order conditions with respect to the SAE parameter γ are equal to zero (e.g. Kelejian and Prucha, 1997). Therefore, to estimate γ, the set of three "generalized moments" introduced by Kelejian and Prucha (1999) are included among the moment conditions for the outcome equation–although they are not explicitly shown in $m(z_{2N},\theta)$ to simplify the exposition. These moment conditions are based on second-order moments of the residuals: $E[\varepsilon_2'\varepsilon_2] = N\sigma_2^2$; $E[\varepsilon_2'C'C\varepsilon_2] = \sigma_2^2 tr(C'C)$; and $E[\varepsilon_2'C\varepsilon_2] = 0$; with $\varepsilon_2 = u_2 - \gamma Cu_2$ from (2). Also, note that it is possible to employ similar moments in the estimation of the spatial probit using the corresponding generalized residuals. However, given evidence from a trial set of simulations indicating that there are no considerable gains from doing so, we conduct our simulations below without them.

[8] An appendix with the derivation of the formal asymptotic properties of $\hat{\theta}_{GMM}$ under assumptions similar to those in Pinkse and Slade (1998) is available from the JAE Data Archive website.

such that $M_N \xrightarrow{p} [\Psi_1(\theta_0)]^{-1}$, then the asymptotic distribution of the optimal $\hat{\theta}_{GMM}$ simplifies to:

$$\sqrt{N}(\hat{\theta}_{GMM} - \theta_0) \xrightarrow{d} N(\mathbf{0}, [\Psi_2(\theta_0)]^{-1}). \quad (12)$$

Nevertheless, it is well documented (e.g. Altonji and Segal, 1996) that in finite samples the optimal GMM estimator can be biased and that the equally-weighted GMM estimator (setting $M_N = I_N$) may be preferred.[9] We explore this practical issue in the context of simulations in the next section.

Lastly, an alternative estimation method to the sequential GMM approach–that we do not explore here–employs a two-step procedure. In the first step consistent estimates of the parameters in θ_1 are obtained from (9), and in the second step nonlinear least squares (NLLS) is employed to estimate the parameters in the outcome equation where the parameter γ enters nonlinearly in the "adjusted" IMR ($\widehat{\lambda}_i$). While this procedure preserves the two-step intuition of the heckit model, extra calculations are necessary to estimate the correct standard errors for the second-step estimates, in addition to employing a heteroskedasticity-consistent variance-covariance estimator for NLLS since the corresponding disturbance is non-spherical.

4 Monte Carlo Experiment

We conduct a Monte Carlo experiment to explore the finite-sample performance of the spatial heckit (spheck) estimator for the sample selection model with SAE. The spheck estimator is compared to three other estimators: the Kelejian and Prucha (1998) estimator for the linear model that ignores sample selection but accounts for SAE (KP-SAE); the heckit estimator that accounts for sample selection but ignores SAE (heckit); and the ordinary least squares (OLS) estimator that ignores both sample selection and SAE. We compare these estimators in terms of their finite sample bias and root-mean square error (RMSE).

[9] In our implementation below, the equally-weighted GMM estimator is obtained as a preliminary estimator used to compute the optimal GMM weighting matrix employed by the optimal GMM estimator.

4.1 Experimental design

Our data generating process (DGP) is as follows:[10]

$$
\begin{aligned}
y_{1i}^{*} &= \alpha_0 + \alpha_1 x_{1i} + \alpha_2 x_{2i} + u_{1i}, \quad u_{1i} = \delta \sum_{j \neq i} c_{ij} u_{1j} + \varepsilon_{1i} &\quad (13) \\
y_{2i}^{*} &= \beta_0 + \beta_1 x_{3i} + \beta_2 x_{1i} + u_{2i}, \quad u_{2i} = \gamma \sum_{j \neq i} c_{ij} u_{2j} + \varepsilon_{2i}. &\quad (14)
\end{aligned}
$$

Each of our models consists of three independent exogenous variables, one of which is common to both equations, as in Cosslett's (1991) experimental design. These exogenous variables are generated as $x_k \sim U(0,1)$, $k = 1, 2, 3$. The innovations ε_{1i} and ε_{2i} are generated bivariate normal as follows:

$$
\begin{bmatrix} \varepsilon_{1i} \\ \varepsilon_{2i} \end{bmatrix} \backsim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix} \right) \quad (15)
$$

where we allow the correlation between the disturbances in each of the two equations to take on values $\rho = \{0.5, 0.75\}$.[11] The parameters of the model that are not related to the spatial dependence feature are set to $\alpha_1 = \alpha_2 = \beta_1 = \beta_2 = 1$ and $\beta_0 = 0$. The parameter α_0 is used to control the amount of sample selection, for which we consider two cases: 25% censoring ($\alpha_0 = -0.3$) and 40% censoring ($\alpha_0 = -.77$). We consider three different sample sizes: $N = 324$, 625, and 900 observations.[12] Importantly, these sample sizes refer to the uncensored sample, implying that the average number of observations available for the estimation of the outcome equation is 243, 468, and 675, respectively, for the case of 25% sample selection; and 194, 375, and 540, respectively, for the case of 40% sample selection.

We consider six pairs of values for the SAE parameters. In the first four pairs they are equal to each other and take values in $\{0, 0.25, 0.5, 0.75\}$. In the last two pairs they take values $(\delta, \gamma) = \{0.75, 0.25\}$ or $\{0.25.0.75\}$. We analyze the relative performance of two versions of the spatial heckit estimator. The first is an equally-weighted GMM estimator that sets $M_N = I_N$ in (11) which we refer to as "spheck-E". The second is the optimal GMM estimator in (12) that sets $M_N = [\Psi_{1N}(\hat{\theta}_{GMM})]^{-1}$ and is referred to as "spheck-O".

[10] Given that we combine features of a sample selection model with a SAE specification, we pay close attention to previous simulation studies in specifying each of the two features of our models, such as Cosslett (1991) and Leung and Yu (1996) for the sample selection model, and Beron and Vijverberg (2004) and Kelejian and Prucha (1999) for spatially dependent models.

[11] In our design, the coefficient on the (adjusted) IMR is equal to ρ.

[12] A set of simulations with $N = 225$ were also obtained. The results are qualitatively similar to the sample with $N = 324$.

For the models with $N = 324$ we undertake $1,000$ replications, whereas 500 replications are undertaken for the models with $N = 625$ and with $N = 900$.[13]

The matrix of spatial weights has to be specified. For this, we create three grids of 18 by 18, 25 by 25 and 30 by 30 for the observation matrices of 324, 625 and 900, respectively. Each grid is assigned an X-Y coordinate centered on the grid such that the bottom left corner of the grid has a value of $(0.5, 0.5)$. We use these grids to create a weighting matrix that is based on the square of the inverse Euclidean distance between any two points. After creating the location specific weights for each grid, a band is used to determine the number of observations that may influence a centered observation. Such band is set with a lower bound of 0 and an upper bound equal to $\sqrt{5}$.[14] Finally, the weighting matrix is row standardized so that its diagonal elements are all zeros and the sum of any one row is equal to 1.[15] This way of specifying the spatial weighting matrix is widely used within the literature, see, for instance, Anselin (1988).

4.2 Results

The results from the Monte Carlo experiment are summarized in eight tables, six of which correspond to the outcome equation and two to the selection equation. The results for the outcome equation are divided in three pairs, each corresponding to a particular sample size, with each table in the pair corresponding to a level of sample selection. To save space, we omit here the two tables for the sample size of 625–although we make reference to them below–and make them all available in an Internet Appendix through the JAE Data Archive at www.econ.queensu.ca/jae.

4.2.1 Outcome Equation

Tables 1 through 4 present results for the bias and RMSE of five estimators of the outcome equation: OLS, heckit, KP-SAE, Spheck-E, and Spheck-O.[16] The first column in

[13] The reason for the smaller number of replications for the larger sample sizes is that the computing time increases dramatically. Representative computing time (using a 2.1 GHz processor with 2.0 GB of RAM) for one replication is (in seconds) 30 for $N = 324$, 300 for $N = 625$, and 470 for $N = 900$.

[14] The resulting number of neighbors varies between 10 and 12, depending on the sample size. We note that Kelejian and Prucha (1999) find that controlling or not for the number of neighbors per unit when specifying a weighting matrix does not lead to significantly different results in their simulation study.

[15] For the estimation of the outcome equation using the spatial heckit and KP-SAE the censored observations are dropped from the weighting matrix before row-standardizing.

[16] The spatial heckit estimators require starting values. Both in the simulations and in the empirical illustration below, we employ starting values that are available in practice. In particular, the starting values

each table indicates the model specification in terms of spatial error dependence (δ and γ) and the correlation between equation disturbances (ρ); while the second column indicates the parameters being estimated.

The first estimator reported in the tables is OLS, which ignores both features of the data: sample selection and spatial dependence. As a result, OLS is theoretically inconsistent, reflected in the fact that it shows large biases across model specifications. In a number of models and parameters–mostly in the smallest sample size–OLS has relatively low RMSE compared to the other estimators. However, this advantage vanishes as the sample size increases or the model specification contains more sample selection or higher ρ.

The estimator KP-SAE accounts for spatial dependence but ignores sample selection, resulting in theoretically inconsistent parameter estimates. In agreement with this notion, both its bias and RMSE increase as the amount of sample selection or spatial dependence increases. Compared to OLS, KP-SAE has higher bias and RMSE in all of their common coefficients, although KP-SAE's measures deteriorate more rapidly as the model specification contains higher sample selection or spatial dependence. This said, the KP-SAE does produces an estimate of the SAE parameter γ. Its bias tends to increase with the amount of spatial dependence, but its RMSE tends to be lower under higher spatial dependence. Both measures improve as the sample size grows larger. These observations about γ hold regardless of whether $\delta = \gamma$ or not.

The three remaining estimators–which generally perform better than the previous two–are discussed and compared in turn.[17] The heckit estimator, which accounts for sample selection but ignores spatial dependence, is theoretically inconsistent since the probit estimated in the first step is heteroskedastic due to the SAE (except in the models with $\delta = \gamma = 0$). Perhaps surprisingly, then, the heckit estimator fares very well in our simulations across models and parameters, both in terms of bias and RMSE. Its edge relative to the spheck estimators, however, decreases dramatically as the sample size increases. In the models with $N = 900$ (Tables 3 and 4), their performance is fairly similar. One obvious disadvantage of heckit is that it does not produce an estimate of the SAE parameter γ.

The spheck estimators (equally-weighted and optimal) are theoretically consistent for all

employed for all parameters except δ and γ are the heckit parameter estimates. The starting values employed for δ and γ are equal to the KP-SAE estimate of γ.

[17] In general, the instances where OLS and/or KP-SAE perform comparably in terms of RMSE to the other three estimators occur in the models with $N = 324$ and low sample selection. This excludes the parameter γ, which will be discussed separately below.

parameters across model specifications. In our simulations, however, they show finite sample biases for β_0 and β_2 that sometimes are high relative to the heckit estimator in the models with $N = 324$ (Tables 1 and 2). Nevertheless, their biases are considerably smaller than OLS and KP-SAE even for this sample size. The parameter β_1 is typically estimated with very low bias by both spheck estimators.[18] For this smallest sample size, the spheck's RMSE is somewhat larger than that of other estimators for all "beta" coefficients, but not drastically so, and in general the relative performance of the spheck estimators–both in terms of bias and RMSE–improves for models with higher sample selection (Table 2). The coefficient ρ–which is the coefficient on the IMR in this equation–is estimated very well by the spheck estimators across all model specifications, with bias and RMSE close to or better than the corresponding one for heckit, except for the RMSE in the models with $N = 324$ where they are always higher.

In general, the parameter best estimated by the spheck estimators in the outcome equation is β_1, followed by ρ and γ. The parameters β_2 and β_0 are more difficult to estimate by the spheck estimators, but this improves with the sample size. In the larger sample sizes (Tables 3 and 4 for $N = 900$, and Internet Appendix Tables A1 and A2 for $N = 625$), both the bias and RMSE of β_2 and β_0 for the spheck estimators improve, getting closer to–and sometimes surpassing–the performance of the heckit estimator. When $N = 900$ and there is 25% sample selection (Table 3), both the bias and RMSE of β_2 for the spheck estimators are low (e.g. typically less than 5% for the bias), although they deteriorate in the models with $\gamma = 0.75$, which surprisingly does not happen to the same extent with heckit. A similar pattern holds under 40% selection (Table 4), although under this higher selection the performance of the spheck estimators is markedly better relative to that in Table 3. Similar patterns hold for $N = 625$ (Internet Appendix Tables A1 and A2) although with lower performance relative to $N = 900$ due to the sample size. The estimation of the intercept β_0 by the spheck estimators appears more challenging in our simulations, especially when the models have high spatial dependence ($\gamma = 0.75$). In general, though, the estimation of β_0 by the spheck estimators follows a similar pattern as that described for the slope β_2.

Both spheck estimators as well as the KP-SAE estimator obtain estimates of γ. These

[18] This is partly due to the fact that, in the experimental design, the variables x_k $(k = 1, 2, 3)$ are generated independently, and thus there is little effect of the sample selection on the coefficient on x_3 (β_1). However, there is no apriori reason why the spheck estimators should estimate this coefficient with less bias compared to the other estimators.

estimates are all very similar in terms of bias and RMSE across all model specifications.[19] In general, the bias of the estimate of γ tends to be higher in models with high spatial dependence ($\gamma = 0.75$), when there is asymmetry in the specification of the SAE parameter of both equations (i.e. $\gamma \neq \delta$), and also when ρ increases. In terms of RMSE a different pattern emerges in which (with a few exceptions) it is slightly higher for models with no spatial dependence and $\gamma = 0.25$.

Finally, we compare the two versions of the spheck estimators: spheck-E and spheck-O. Theoretically, both are consistent but spheck-O is a more efficient estimator than spheck-E. The evidence in this simulation exercise for the outcome equation supports a very similar performance of the two estimators with a slight edge of the spheck-O over spheck-E.

4.2.2 Selection Equation

Tables 5 and 6 present simulation results for the selection equation for models with 25% and 40% selection, respectively. This equation is only estimated in the heckit (using a probit model) and the two versions of spheck.[20] Recall that the same number of observations are employed in each case of amount of censoring (25% and 40%) in the selection equation but the value of the constant term is different to be able to generate the required amount of censoring in the outcome equation.

In general, the performance of the spheck estimators relative to the heckit in the selection equation is similar to that in the outcome equation. The heckit typically has smaller bias and RMSE, although the performance of the spheck estimators becomes increasingly similar to the heckit as the sample size increases–especially in terms of bias. The relative performance of the spheck estimators also shows a marked improvement when the estimation environment is more challenging–e.g. higher amount of spatial dependence, higher sample selection, and higher correlation between the errors in each equation. As it was the case before, the spheck estimators have more difficulty estimating the constant term (α_0) than the other parameters. The parameter δ, which is only estimated by the spheck estimators, is estimated successfully except when it takes a value of 0.75, where a substantial bias is present. Finally, although the performance of the two spheck estimators is similar for this equation, a tenuous pattern emerges in which spheck-E performs slightly better in terms of bias and spheck-O in terms

[19] This may not be surprising as the identification of this parameter comes from the same source in each of the estimators–the Kelejian and Prucha (1999) moments.

[20] As in the outcome equation, the results for $N = 625$ are omitted from the paper but available in the Internet Appendix.

of RMSE.

4.2.3 Summary

Despite the necessary limited range of specifications employed in this simulation experiment, we view the results as encouraging with respect to the finite sample properties of the spatial heckit estimator. In particular, while lagging the heckit in estimation performance on the common parameters in the smaller sample sizes, the fact that the advantages of our estimator become evident in the simulations suggest that the spatial heckit has adequate finite sample properties for the larger sample sizes analyzed. It is likely that even better performance can be realized in larger samples given the documented trend over the sample size. The simulations also show that the two versions of the spatial heckit estimator have similar performance. Finally, the spheck estimators adequately recover the SAE parameters–which is important since they are the only estimators that identify both parameters–except in the selection equation in the models with highest spatial dependence. We leave this issue for future research. Another feature that deserves further investigation is the surprising performance of the heckit estimator in the presence of SAE within our simulations. While this may come as good news to previous work that have used this estimator ignoring spatial dependence (e.g. Jud and Seaks, 1994), we caution that this evidence is far from conclusive.

5 Empirical Application

Our application in this section is in the area of natural resource economics, in particular fisheries economics. We employ the sample selection model with SAE to estimate spatial production within a fishery using data that is censored for reasons of confidentiality, comparing the performance of the spatial heckit estimator to OLS, heckit, and KP-SAE.

Within fisheries management there has been a strong push to incorporate the spatial structure of the bioeconomic model (e.g., Wilen, 2004). An initial interest is to investigate the production process within fisheries over the spatial region defined by the distribution of the metapopulation harvested. The catch-per-unit-effort (CPUE) has been traditionally employed to analyze the productivity and efficiency of production within fisheries, where CPUE is defined as the catch per a "haul" executed.[21] Previous work has investigated non-spatially defined production in an effort to determine the factors that explain deviations

[21] A "haul" represents the technology used such as a trawl device, pot vessel, hook-and-line, jig, etc.

from the production frontier (Kirkley et al. 1995, 1998; Squires and Kirkley 1999; Pascoe and Coglan 2002). However, the results of these studies do not incorporate the spatial process into the model and thus are likely inappropriate for spatial fisheries management. For instance, closing a particular region within the fishery with a low level of spatial technical efficiency will displace fishing effort into more efficient surrounding areas, forcing fishermen to more exhaustively push the frontier of their production capabilities to capture the same amount of the target species. This will invariably yield a higher cost of harvesting and lower rents for the fishermen, more so than if a high efficiency area is closed instead. Estimating spatial efficiency is beyond the scope of this application, but investigating spatial production is a necessary first step in the process of facilitating fisheries policy.

Our analysis is conducted on the Pacific cod fishery within the Eastern Bering Sea of Alaska for the year 1997 using publicly available data collected by the National Marine Fisheries Service (NMFS).[22] Determining the spatial rates of production requires a very fine spatial resolution of data, which is often screened to preserve the privacy of the fishermen within the fleet. The NMFS data is censored by not reporting the CPUE within a location unless 4 or more vessels with similar characteristics fish within that region. Therefore, the use of this data is limited unless a method is available that can account for this censorship, justifying the use of the spatial sample selection model. The data contains 320 observations of which 207 are uncensored, resulting in a sample selection rate of 35%. Unfortunately, these data do not include vessel identifiers, which would allow an analysis focusing on inter and intra vessel differences in the spatial distribution of production. Nevertheless, it is still possible to determine the overall level of the fleet's spatial production using these data and test for spatial heterogeneity at this level.[23]

To conduct the analysis, the spatial resolution (i.e. "locations") utilized are the Alaska Department of Fish and Game's (ADF&G) statistical reporting units. These units divide the Eastern Bering Sea into a grid with each cell being one-half degree latitude by one degree longitude. For the year analyzed, the fishery is divided into 90 spatially different locations. The CPUE is defined as the metric tons of fish caught during the year within the ADF&G

[22] Estimating the fleets' spatial production with regard to this species is beneficial due to its broad distribution within the Eastern Bering Sea which makes it susceptible to recent regulations targeted to protect the Stellar sea lion rookeries and the concerns of essential fish habitat (EFH) management.

[23] The vessel-level model is undobtedly more interesting. However, given that a researcher will invest a substantial amount of time and effort in obtaining the necessary information to estimate that model, it is beneficial to first analyze fleet performance and test for spatial heterogeneity at this level.

statistical reporting regions. This measure is the average of all vessels with similar characteristics that fished within the region. Vessels were grouped according to the size of the vessel, gear utilized, and type of vessel (catcher-processor vs. catcher-vessel). Therefore, each observation represents a relatively homogeneous micro-fleet within the ADF&G region. Given that these observations are spatially defined, it is likely that they are spatially correlated, and therefore a spatial econometric method must be utilized to obtain appropriate estimates. Indeed, the Moran-I test statistic using the OLS residuals soundly rejects the null hypothesis of zero spatial autocorrelation on the data with a p-value of 0.00.

A production model is estimated with each of the four estimators. Given the documented spatial dependence and the censoring in the data, a sample selection model with SAE is likely the more appropriate one. Thus, the sample selection model with spatial dependence in (1)-(2) is estimated with y_2 as the natural logarithm of CPUE and x_2 containing the log-transformed bathymetric measurements corresponding with the maximum and minimum depth within the location, the stock assessment data resulting from the NMFS annual biomass trawl survey, and indicator variables for the vessel characteristics of the fleet: catcher-vessel (CV), hook-and-line gear (HAL), non-pelagic trawl gear (NPT), and vessel at least 125 feet long (Large). As for the selection equation, x_1 contains the same variables as x_2 plus the one-year lagged biomass trawl survey observation, under the assumption that this lagged variable influences the probability that four or more vessels will fish in a given location but not the amount of "hauls" that will be conducted.

For the spatial weighting matrix we use a common specification to assign spatial weights among the locations: $c_{ij} = \frac{1}{d_{ij}^f}$, where c_{ij} is the spatial weight assigned to the distance between location i and location j, d_{ij} is the Euclidian distance between locations i and j, and f is a "friction" parameter.[24] To control the number of neighbors per statistical reporting unit a band is chosen. Finally, the spatial weights, c_{ij}, are row standardized such that the diagonal elements of the spatial weighting matrix are all zero and the sum of any one row is one. We employ a band of 7 and a friction parameter of 2. The results from all estimators are presented in Table 7.[25]

[24] The spatial weighting matrix was constructed by superimposing the ADF&G on to the $X-Y$ coordinate plane.

[25] We note that the numerical optimizations needed to estimate the spatial heckit make it computationally intensive relative to the other three estimators. In the currrent application, using a 2.1 GHz processor with 2.0 GB of RAM, the times of each estimator (in seconds) are 0.03 for OLS, 0.79 for KP-SAE, 0.61 for heckit, 448 for spheck-E and 905 for spheck-O.

The estimated coefficients for the outcome equation from each of the models (left panel) present a number of important differences. First, the OLS and KP-SAE estimators often differ in magnitude, sign and statistical significance compared to the estimates yield by the other estimators. Second, the spheck estimators and the heckit largely agree in the magnitude of the estimates in most coefficients, although not always in their statistical significance. Third, the spheck estimators are very similar in the magnitude of their estimated coefficients but not in their statistical significance, with spheck-O showing more precise estimates. Fourth, the estimate of the SAE parameter (γ) is high and largely agrees across the KP-SAE and spheck estimators, although it is only statistically significant in spheck-O and KP-SAE. Importantly, most of these features of the outcome equation are qualitatively consistent with the simulation results described in the previous section, with the exception perhaps of the magnitude in the differences in precision among the estimators.

Interestingly, despite the relatively high amount of selection in the sample (35%), the IMR is not statistically significant except in spheck-O, although it is estimated to be positive and of similar magnitude in all three models. This feature could be due to a low correlation between the equation errors (ρ) or the result of using a small sample. The statistical significance of the spheck-O coefficients agrees with expectations for these data. For instance, in the case of "Max. Depth", it is expected that vessels fishing in deeper areas would obtain a larger CPUE. Similarly, it is also expected that areas with high "Biomass" signal would attain more productivity per haul executed, if such signal is accurate. Finally, large vessels ("Dum Large") are also expected to be more productive. Summarizing the results for the outcome equation in this empirical illustration, the spheck estimators and heckit yield more sensible results, but only spheck-O achieves sufficiently small standard errors that result in a number of statistically significant coefficients.

The right panel of Table 7 presents the estimated coefficients for the selection equation. The estimators yield somewhat similar parameter estimates albeit with some noteworthy differences. For instance, the estimated coefficient on "Max. Depth" in heckit is smaller than that of the spheck estimators although it is not statistically significant in spheck-E; while the estimated coefficient on "Min. Depth" is statistically significant in spheck-O and of higher magnitude compared to both the heckit and spheck-E. The coefficient on "Dum HAL" is somewhat smaller in the spheck estimators while being statistically significant in the heckit and spheck-O. The spheck estimators yield estimates of δ that are between 0.4 and

0.2 but are, perhaps surprisingly, not statistically significant at conventional levels. Finally, a feature worth noting–in both equations–is the relatively large standard errors of spheck-E relative to spheck-O, which is not necessarily at odds with theory but it is somewhat so with the simulation results of the previous section. Nevertheless, we note that our sample size in this application is in the low end of our simulation sample sizes, where the differences in RMSE between the spheck estimators were more marked.

Lastly, although a more complete analysis would be required before any concrete policy recommendations are made from this exercise, our results suggest that there is some degree of heterogeneity in the spatial production rates within the Pacific cod fishery. This could be attributed to a number of different factors such as climatic conditions, skipper ability, and interactions with other fisheries (to name a few).

6 Conclusions

This paper proposes a method of estimation for a sample selection model with spatial autoregressive errors (SAE). The method of estimation is analogous to the popular heckit model–and thus we call our estimator the "spatial heckit"–in which consistent estimates of the probability of observing a particular unit (selection equation) are estimated following the probit model of Pinkse and Slade (1998). Then, the odds of observing each unit are calculated–the inverse Mills ratio–and used as an additional regressor that controls for the selection bias in the equation of interest (outcome equation). Importantly, the appropriate inverse Mills ratio depends on the SAE parameter of the outcome equation. Therefore, we propose to estimate the model jointly by nesting the two equations into a sequential GMM framework (Newey, 1984), obtaining directly the variance-covariance matrix of the estimator.

We explore the properties of the spatial heckit by stating its asymptotic properties, conducting simulations, and applying it to actual data. The estimator is consistent and asymptotically normally distributed and it is shown to have acceptable finite sample properties. The simulations also show the potential biases incurred by other estimators that ignore sample selection, or both sample selection and spatial dependence. An interesting finding is that the heckit estimator has good finite sample properties in our simulations despite being theoretically inconsistent. Finally, the empirical application section illustrates that our estimator is both feasible and valuable to use in practice.

To our knowledge, the proposed estimator is the first to account for sample selection and spatial dependence simultaneously. Nevertheless, some shortcomings are worth mentioning. First, our estimator relies on a distributional assumption (joint normality) of the error terms in selection and outcome equations, just as the heckit estimator does. This indicates an area for future research. Second, is the relatively greater computational intensity of our estimator compared to the available methods for linear spatial models without sample selection. However, our estimator still compares favorably in this respect with full-blown maximum likelihood, which requires approximation of multidimensional integrals on the order of the sample size.

References

[1] Altonji JG, Segal LM. 1996. Small-Sample Bias in GMM Estimation of Covariance Structures. *Journal of Business and Economic Statistics* **14:** 353-366.

[2] Amemiya T. 1985. *Advanced Econometrics*. Cambridge, Mass.: Harvard University Press.

[3] Anselin L. 1988. *Spatial Econometrics: Methods and Models*. Dordrecht: Kluwer Academic Publishers.

[4] Anselin L. 2001. Spatial Econometrics. In *A Companion to Theoretical Econometrics,* Baltagi BH (ed.). Oxford: Blackwell.

[5] Anselin L, Bera A. 1998. Spatial Dependence in Linear Regression Models with an Introduction to Spatial Econometrics. In *Handbook of Applied Economic Statistics*, Ullah A, Giles DEA (eds.). New York: Marcel Dekker.

[6] Benjamin JD, Sa-Aadu J, Shilling JD. 1992. Influence of Rent Differentials on the Choice Between Office Rent Contracts with and without Relocation Provisions. *Journal of American Real Estate and Urban Economics Association* **20:** 289-302.

[7] Beron K, Vijverberg W. 2004. Probit in a Spatial Context: A Monte Carlo Analysis. In *Advances in Spatial Econometrics: Methodology, Tools and Applications*, Anselin L, Florax R, Rey S (eds.). Berlin: Springer-Verlag.

[8] Borjas GJ, Bronars SG, Trejo S. 1992. Self-Selection and Internal Migration in the United States. *Journal of Urban Economics* **32:** 159–85.

[9] Case AC. 1991. Spatial Patterns in Household Demand. *Econometrica* **59:** 953-965.

[10] Case AC. 1992. Neighborhood Influence and Technology Change. *Regional Sciences and Urban Economics* **22:** 491-508.

[11] Casetti E. 1972. Generating Models by the Expansion Method: Applications to Geographical Research. *Geographical Analysis* **4:** 81-91.

[12] Chamberlain G. 1987. Asymptotic Efficiency in Estimation with Conditional Moment Restrictions. *Journal of Econometrics* **34:** 305-334.

[13] Cosslett SR. 1991. Semiparametric Estimation of a Regression Model with Sample Selectivity. In *Nonparametric and Semiparametric Methods in Econometrics and Statistics*, Barnett WA, Powell J, Tauchen G (eds.). Cambridge: Cambridge University Press.

[14] Fishback PV, Horrace WC, Kantor S. 2006. The Impact of New Deal Expenditures on Mobility During the Great Depression. *Explorations in Economic History* **43:**179-222.

[15] Fleming M. 2004. Techniques for Estimating Spatially Dependent Discrete Choice Models. In *Advances in Spatial Econometrics: Methodology, Tools and Applications*, Anselin L, Florax R, Rey S (eds.). Berlin: Springer-Verlag.

[16] Goodman AC. 1988. An Econometric Model of Housing Price, Permanent Income, Tenure Choice, and Housing Demand. *Journal of Urban Economics* **23:** 327-53.

[17] Hansen LP. 1985. A Method for Calculating Bounds on the Asymptotic Covariance Matrices of Generalized Method of Moments Estimators. *Journal of Econometrics* **30:** 203-238.

[18] Heckman J. 1976. The Common Structure of Statistical Models of Truncation, Sample Selection, and Limited Dependent Variables and a Simple Estimator for Such Models. *The Annals of Economic and Social Measurement* **5:** 475-492.

[19] Heckman J. 1979. Sample Selection Bias as a Specification Error. *Econometrica* **47:** 153-61.

[20] Jud GD, Seaks TG. 1994. Sample Selection Bias in Estimating Housing Sales Prices. *Journal of Real Estate Research* **9:** 289-298.

[21] Kelejian HH, Prucha IR. 1997. Estimation of Spatial Regression Models with Autoregressive Errors by Two-Stage Least Squares Procedures: A Serious Problem. *International Regional Science Review* **20:** 103-111.

[22] Kelejian HH, Prucha IR. 1998. A Generalized Spatial Two-Stage Least Squares Procedure for Estimating a Spatial Autoregressive Model with Autoregressive Disturbances. *Journal of Real Estate Finance and Economics* **17:** 99-121.

[23] Kelejian HH, Prucha IR. 1999. A Generalized Moments Estimator for the Autoregressive Parameter in a Spatial Model. *International Economic Review* **40:** 509-33.

[24] Kelejian HH, and Prucha IR. 2009. Specification and Estimation of Spatial Autoregressive Models with Autoregressive and Heteroskedastic Disturbances. Forthcoming.

[25] Kirkley J, Squires D, Strand IE. 1995. Assessing Technical Efficiency in Commercial Fisheries: The Mid-Atlantic Sea Scallop Fishery. *American Journal of Agricultural Economics* **77:** 689-697.

[26] Kirkley J, Squires D, Strand IE. 1998. Characterizing Managerial Skill and Technical Efficiency. *Journal of Productivity Analysis* **9:**145-160.

[27] Klier T, McMillen DP. 2008. Clustering of Auto Supplier Plants in the United States: Generalized Method of Moments Spatial Logit for Large Samples. *Journal of Business and Economic Statistics* **26:** 460-471.

[28] Lee LF. 2004. Asymptotic Distributions of Quasi-Maximum Likelihood Estimators for Spatial Autoregressive Models. *Econometrica* **72:** 1899-1925.

[29] LeSage JP. 2000. Bayesian Estimation of Limited Dependent Variable Spatial Autoregressive Models. *Geographical Analysis* **32:** 19-35.

[30] Leung SF, Yu S. 1996. On the Choice Between Sample Selection and Two-Part Models. *Journal of Econometrics* **72:** 197-229.

[31] McMillen DP. 1992. Probit with Spatial Autocorrelation. *Journal of Regional Science* **32:** 335-48.

[32] McMillen DP. 1993. Can Blacks Earn More in the Suburbs? Racial Differences in Intrametropolitan Earnings Variation. *Journal of Urban Economics* **33:** 167-88.

[33] McMillen DP. 1995. Selection Bias in Spatial Econometric Models. *Journal of Regional Science* **35:** 417-36.

[34] Montgomery C. 1992. Explaining Home Improvement in the Context of Household Investment in Residential Housing. *Journal of Urban Economics* **32:** 326-50.

[35] Newey W. 1984. A Method of Moments Interpretation of Sequential Estimators. *Economics Letters* **14:** 201-206.

[36] Ord JK. 1975. Estimation Methods for Models of Spatial Interaction. *Journal of the American Statistical Association* **70:**120-126.

[37] Pascoe S, Coglan L. 2002. The Contribution of Unmeasureable Inputs to Fisheries Production: An Analysis of Technical Efficiency of Fishing Vessels in the English Channel. *American Journal of Agricultural Economics* **84:**585-597.

[38] Pinkse J, Slade ME. 1998. Contracting in Space. An Application of Spatial Statistics to Discrete-Choice Models. *Journal of Econometrics* **85:** 125-54.

[39] Squires D, Kirkley J. 1999. Skipper skill and panel data in fishing industries. *Canadian Journal of Fisheries and Aquatic Science* **56:** 2011-2018.

[40] Topa G. 2001. Social Interactions, Local Spillovers, and Unemployment. *Review of Economic Studies* **2:** 133-138.

[41] Wilen J. 2004. Spatial Management of Fisheries. *Marine Resource Economics* **19:** 7-20.

Table 1. Simulation Results for N=324, 25% Sample Selection

(δ,γ,ρ)		BIAS					RMSE				
		OLS	KP-SAE	Heckit	Spheck-E	Spheck-O	OLS	KP-SAE	Heckit	Spheck-E	Spheck-O
(0,0,0.5)*	β_0	0.316	0.318	-0.024	-0.084	-0.089	0.364	0.366	0.411	0.591	0.557
	β_1	0.021	0.020	-0.002	0.000	-0.003	0.223	0.225	0.222	0.228	0.254
	β_2	-0.235	-0.236	0.014	0.060	0.048	0.322	0.324	0.355	0.554	0.457
	ρ			0.042	0.126	0.030			0.567	0.797	0.644
	γ		-0.065		-0.069	-0.065		0.252		0.266	0.271
(0.25,0.25, 0.5)	β_0	0.320	0.325	-0.024	-0.092	-0.084	0.370	0.386	0.410	0.616	0.564
	β_1	0.021	0.019	-0.003	-0.002	-0.006	0.221	0.250	0.221	0.223	0.248
	β_2	-0.235	-0.241	0.016	0.076	0.046	0.324	0.350	0.356	0.595	0.463
	ρ			0.047	0.131	0.023			0.564	0.800	0.647
	γ		-0.005		0.003	0.004		0.220		0.229	0.236
(0.5,0.5, 0.5)	β_0	0.332	0.364	-0.030	-0.172	-0.122	0.394	0.538	0.462	0.805	0.679
	β_1	0.030	0.033	0.007	0.005	-0.001	0.236	0.485	0.235	0.228	0.248
	β_2	-0.240	-0.294	0.016	0.131	0.072	0.337	0.591	0.381	0.710	0.525
	ρ			0.068	0.178	0.027			0.628	0.958	0.703
	γ		0.071		0.087	0.086		0.184		0.197	0.202
(0.75,0.75, 0.5)	β_0	0.370	0.613	0.019	-0.392	-0.186	0.474	1.412	0.559	1.329	1.049
	β_1	0.036	0.014	0.014	0.009	0.004	0.278	1.535	0.276	0.230	0.251
	β_2	-0.260	-0.661	-0.043	0.169	0.091	0.368	1.898	0.398	0.991	0.704
	ρ			0.006	0.239	-0.047			0.669	1.291	0.880
	γ		0.083		0.105	0.101		0.123		0.142	0.142
(0.25,0.75, 0.5)	β_0	0.332	0.530	0.028	-0.408	-0.176	0.446	1.459	0.480	1.225	0.885
	β_1	0.033	0.021	0.012	0.002	-0.002	0.271	1.601	0.270	0.227	0.247
	β_2	-0.263	-0.593	-0.042	0.196	0.104	0.363	2.002	0.382	0.955	0.638
	ρ			-0.022	0.230	-0.095			0.558	1.154	0.701
	γ		0.113		0.133	0.132		0.137		0.157	0.158
(0.75,0.25, 0.5)	β_0	0.311	0.317	-0.049	-0.108	-0.125	0.367	0.379	0.498	0.646	0.667
	β_1	0.032	0.028	0.011	0.013	0.006	0.234	0.252	0.232	0.234	0.260
	β_2	-0.207	-0.212	0.022	0.066	0.061	0.311	0.332	0.374	0.568	0.488
	ρ			0.047	0.123	0.047			0.690	0.859	0.754
	γ		-0.067		-0.060	-0.062		0.241		0.252	0.255
(0,0,0.75)*	β_0	0.475	0.476	-0.032	-0.107	-0.178	0.503	0.504	0.408	0.600	0.566
	β_1	0.033	0.031	-0.001	0.000	-0.006	0.210	0.212	0.208	0.212	0.233
	β_2	-0.352	-0.352	0.020	0.075	0.090	0.406	0.407	0.352	0.580	0.460
	ρ			0.056	0.164	0.092			0.561	0.793	0.638
	γ		-0.066		-0.072	-0.063		0.256		0.272	0.276
(0.25,0.25, 0.75)	β_0	0.481	0.487	-0.027	-0.079	-0.164	0.510	0.521	0.407	0.574	0.562
	β_1	0.033	0.031	-0.001	-0.001	-0.007	0.208	0.225	0.206	0.208	0.231
	β_2	-0.353	-0.360	0.016	0.051	0.073	0.409	0.429	0.348	0.548	0.449
	ρ			0.057	0.136	0.087			0.562	0.767	0.659
	γ		-0.036		-0.031	-0.024		0.227		0.235	0.239
(0.5,0.5, 0.75)	β_0	0.503	0.538	-0.035	-0.167	-0.203	0.538	0.631	0.453	0.743	0.659
	β_1	0.040	0.052	0.006	0.000	-0.005	0.221	0.395	0.218	0.213	0.232
	β_2	-0.362	-0.424	0.019	0.117	0.102	0.422	0.596	0.371	0.663	0.500
	ρ			0.094	0.191	0.087			0.621	0.880	0.693
	γ		0.025		0.043	0.043		0.178		0.190	0.197
(0.75,0.75, 0.75)	β_0	0.575	0.772	0.015	-0.547	-0.358	0.630	1.304	0.553	1.299	1.038
	β_1	0.048	0.132	0.013	0.004	0.000	0.252	1.242	0.248	0.221	0.239
	β_2	-0.383	-0.786	-0.032	0.237	0.167	0.453	1.686	0.385	0.966	0.654
	ρ			0.073	0.376	0.025			0.658	1.231	0.841
	γ		0.058		0.090	0.086		0.117		0.142	0.144
(0.25,0.75, 0.75)	β_0	0.500	0.681	0.038	-0.530	-0.282	0.570	1.427	0.466	1.225	0.848
	β_1	0.045	0.131	0.012	-0.004	-0.008	0.252	1.498	0.249	0.211	0.229
	β_2	-0.381	-0.759	-0.046	0.244	0.146	0.449	1.971	0.368	0.889	0.592
	ρ			-0.024	0.232	-0.150			0.526	1.074	0.627
	γ		0.114		0.142	0.142		0.135		0.164	0.165
(0.75,0.25, 0.75)	β_0	0.472	0.473	-0.042	-0.100	-0.189	0.504	0.507	0.488	0.590	0.635
	β_1	0.038	0.036	0.007	0.008	-0.002	0.218	0.223	0.215	0.219	0.241
	β_2	-0.309	-0.310	0.019	0.060	0.078	0.377	0.383	0.367	0.515	0.452
	ρ			0.028	0.112	0.084			0.657	0.792	0.742
	γ		-0.188		-0.188	-0.184		0.306		0.319	0.320

Note: Simulation results are based on 1000 replications. * In these models with no spatial dependence Heckit is theoretically consistent.

Table 2. Simulation Results for N=324, 40% Sample Selection

(δ,γ,ρ)		BIAS					RMSE				
		OLS	KP-SAE	Heckit	Spheck-E	Spheck-O	OLS	KP-SAE	Heckit	Spheck-E	Spheck-O
(0,0,0.5)*	β_0	0.445	0.445	-0.042	-0.081	-0.120	0.491	0.492	0.491	0.600	0.636
	β_1	0.041	0.041	0.008	0.010	0.003	0.254	0.257	0.251	0.258	0.286
	β_2	-0.289	-0.290	0.022	0.054	0.052	0.382	0.384	0.382	0.490	0.459
	ρ			0.039	0.071	0.032			0.482	0.584	0.546
	γ		-0.068		-0.074	-0.073		0.250		0.270	0.273
(0.25,0.25, 0.5)	β_0	0.447	0.454	-0.049	-0.087	-0.122	0.497	0.511	0.516	0.595	0.629
	β_1	0.045	0.043	0.010	0.011	0.004	0.257	0.280	0.254	0.256	0.283
	β_2	-0.291	-0.299	0.026	0.057	0.052	0.384	0.407	0.392	0.485	0.446
	ρ			0.048	0.080	0.035			0.502	0.578	0.531
	γ		-0.068		-0.061	-0.060		0.226		0.235	0.241
(0.5,0.5, 0.5)	β_0	0.463	0.488	-0.045	-0.142	-0.142	0.523	0.614	0.553	0.760	0.726
	β_1	0.052	0.060	0.018	0.015	0.011	0.269	0.434	0.266	0.257	0.283
	β_2	-0.297	-0.340	0.016	0.096	0.061	0.399	0.575	0.404	0.599	0.497
	ρ			0.065	0.106	0.036			0.546	0.671	0.577
	γ		-0.030		-0.011	-0.012		0.185		0.190	0.197
(0.75,0.75, 0.5)	β_0	0.531	0.682	-0.001	-0.405	-0.184	0.618	1.344	0.686	1.297	0.994
	β_1	0.049	0.092	0.017	0.009	0.007	0.307	1.368	0.305	0.267	0.283
	β_2	-0.312	-0.596	-0.014	0.215	0.092	0.423	1.706	0.461	0.929	0.593
	ρ			0.084	0.243	-0.003			0.652	1.067	0.714
	γ		0.026		0.053	0.051		0.127		0.140	0.145
(0.25,0.75, 0.5)	β_0	0.499	0.653	-0.015	-0.461	-0.187	0.595	1.411	0.619	1.294	0.872
	β_1	0.050	0.031	0.015	0.008	0.005	0.308	1.482	0.308	0.265	0.284
	β_2	-0.335	-0.559	-0.008	0.252	0.089	0.442	1.764	0.446	0.921	0.548
	ρ			0.065	0.191	-0.074			0.564	0.911	0.532
	γ		0.059		0.085	0.084		0.114		0.134	0.139
(0.75,0.25, 0.5)	β_0	0.424	0.426	-0.050	-0.095	-0.141	0.476	0.485	0.607	0.698	0.765
	β_1	0.040	0.038	0.013	0.012	0.007	0.257	0.278	0.255	0.258	0.283
	β_2	-0.252	-0.252	0.017	0.054	0.053	0.356	0.373	0.417	0.531	0.504
	ρ			0.032	0.075	0.045			0.616	0.712	0.684
	γ		-0.107		-0.099	-0.096		0.258		0.269	0.273
(0,0,0.75)*	β_0	0.670	0.671	-0.050	-0.088	-0.213	0.696	0.697	0.477	0.561	0.639
	β_1	0.058	0.057	0.009	0.011	0.003	0.235	0.237	0.227	0.234	0.257
	β_2	-0.433	-0.434	0.028	0.060	0.078	0.488	0.490	0.371	0.473	0.444
	ρ			0.048	0.077	0.087			0.461	0.536	0.545
	γ		-0.071		-0.079	-0.073		0.256		0.276	0.276
(0.25,0.25, 0.75)	β_0	0.676	0.680	-0.056	-0.083	-0.226	0.702	0.710	0.495	0.532	0.642
	β_1	0.060	0.059	0.010	0.010	0.001	0.235	0.250	0.228	0.231	0.254
	β_2	-0.435	-0.441	0.032	0.053	0.085	0.491	0.505	0.377	0.428	0.436
	ρ			0.060	0.082	0.105			0.480	0.534	0.557
	γ		-0.102		-0.097	-0.093		0.245		0.252	0.255
(0.5,0.5, 0.75)	β_0	0.703	0.727	-0.056	-0.140	-0.243	0.735	0.789	0.547	0.704	0.709
	β_1	0.065	0.076	0.015	0.011	0.005	0.246	0.352	0.238	0.233	0.256
	β_2	-0.443	-0.484	0.026	0.085	0.090	0.503	0.617	0.396	0.547	0.465
	ρ			0.091	0.132	0.114			0.531	0.665	0.575
	γ		-0.081		-0.061	-0.062		0.205		0.207	0.212
(0.75,0.75, 0.75)	β_0	0.802	0.970	-0.023	-0.391	-0.277	0.849	1.343	0.699	1.147	0.952
	β_1	0.064	0.160	0.017	0.001	-0.001	0.276	1.064	0.270	0.243	0.262
	β_2	-0.460	-0.810	0.007	0.208	0.114	0.530	1.499	0.457	0.790	0.543
	ρ			0.166	0.220	0.014			0.670	0.931	0.684
	γ		-0.011		0.025	0.020		0.136		0.147	0.153
(0.25,0.75, 0.75)	β_0	0.755	0.940	0.008	-0.504	-0.294	0.812	1.502	0.568	1.236	0.884
	β_1	0.066	0.117	0.014	0.002	0.000	0.283	1.331	0.277	0.236	0.254
	β_2	-0.495	-0.819	-0.018	0.270	0.125	0.561	1.807	0.408	0.871	0.535
	ρ			0.074	0.125	-0.127			0.513	0.824	0.547
	γ		0.059		0.098	0.098		0.113		0.143	0.147
(0.75,0.25, 0.75)	β_0	0.638	0.638	-0.057	-0.095	-0.241	0.667	0.669	0.589	0.645	0.774
	β_1	0.049	0.048	0.009	0.008	-0.001	0.238	0.246	0.234	0.238	0.260
	β_2	-0.369	-0.368	0.023	0.048	0.082	0.436	0.441	0.399	0.469	0.476
	ρ			0.032	0.071	0.098			0.592	0.653	0.675
	γ		-0.218		-0.216	-0.210		0.327		0.345	0.343

Note: Simulation results are based on 1000 replications. * In these models with no spatial dependence Heckit is theoretically consistent.

Table 3. Simulation Results for N=900, 25% Sample Selection

(δ,γ,ρ)		BIAS					RMSE				
		OLS	KP-SAE	Heckit	Spheck-E	Spheck-O	OLS	KP-SAE	Heckit	Spheck-E	Spheck-O
(0,0,0.5)*	β_0	0.306	0.306	0.003	-0.025	-0.050	0.324	0.324	0.214	0.329	0.336
	β_1	0.020	0.021	0.007	0.006	0.012	0.127	0.127	0.127	0.128	0.141
	β_2	-0.216	-0.216	-0.008	0.018	0.014	0.252	0.253	0.188	0.304	0.276
	ρ			-0.006	0.028	-0.031			0.294	0.428	0.396
	γ		-0.017		-0.017	-0.016		0.149		0.152	0.158
(0.25,0.25, 0.5)	β_0	0.309	0.314	0.003	-0.026	-0.038	0.329	0.336	0.217	0.295	0.246
	β_1	0.022	0.021	0.008	0.006	0.012	0.130	0.144	0.129	0.129	0.141
	β_2	-0.217	-0.224	-0.009	0.020	0.004	0.254	0.268	0.188	0.282	0.208
	ρ			-0.003	0.026	-0.049			0.293	0.393	0.295
	γ		0.050		0.054	0.056		0.132		0.135	0.139
(0.5,0.5, 0.5)	β_0	0.327	0.366	0.004	-0.065	-0.055	0.350	0.422	0.237	0.436	0.311
	β_1	0.021	0.022	0.008	0.004	0.010	0.137	0.276	0.136	0.131	0.141
	β_2	-0.226	-0.289	-0.011	0.049	0.018	0.265	0.395	0.196	0.363	0.238
	ρ			0.017	0.044	-0.063			0.309	0.533	0.331
	γ		0.122		0.129	0.131		0.152		0.158	0.161
(0.75,0.75, 0.5)	β_0	0.385	0.712	0.010	-0.344	-0.101	0.424	1.273	0.318	0.844	0.491
	β_1	0.016	-0.122	0.002	0.002	0.006	0.153	1.350	0.153	0.130	0.141
	β_2	-0.236	-0.678	-0.010	0.130	0.052	0.287	1.460	0.226	0.487	0.302
	ρ			0.069	0.250	-0.107			0.386	0.942	0.465
	γ		0.149		0.163	0.167		0.155		0.171	0.174
(0.25,0.75, 0.5)	β_0	0.348	0.555	0.027	-0.271	-0.040	0.390	1.293	0.260	0.769	0.337
	β_1	0.018	0.008	0.004	-0.001	0.002	0.154	1.384	0.154	0.136	0.144
	β_2	-0.244	-0.585	-0.024	0.135	0.020	0.291	1.606	0.213	0.505	0.251
	ρ			0.019	0.110	-0.192			0.300	0.779	0.312
	γ		0.171		0.182	0.186		0.174		0.187	0.191
(0.75,0.25, 0.5)	β_0	0.309	0.310	0.001	-0.021	-0.045	0.329	0.332	0.251	0.330	0.290
	β_1	0.016	0.016	0.004	0.002	0.008	0.131	0.141	0.130	0.130	0.143
	β_2	-0.192	-0.193	-0.007	0.009	0.005	0.235	0.242	0.194	0.283	0.217
	ρ			-0.026	0.001	-0.059			0.343	0.432	0.342
	γ		-0.006		-0.002	0.000		0.127		0.128	0.133
(0,0,0.75)*	β_0	0.459	0.459	-0.002	-0.017	-0.113	0.468	0.468	0.209	0.237	0.268
	β_1	0.030	0.030	0.010	0.009	0.012	0.118	0.118	0.115	0.116	0.131
	β_2	-0.322	-0.323	-0.006	0.009	0.028	0.344	0.344	0.185	0.240	0.213
	ρ			0.000	0.016	0.001			0.286	0.305	0.290
	γ		-0.013		-0.012	-0.006		0.142		0.145	0.149
(0.25,0.25, 0.75)	β_0	0.464	0.470	-0.003	-0.026	-0.130	0.474	0.482	0.212	0.260	0.280
	β_1	0.030	0.028	0.010	0.008	0.011	0.120	0.130	0.117	0.117	0.131
	β_2	-0.324	-0.332	-0.006	0.017	0.036	0.346	0.357	0.185	0.261	0.217
	ρ			0.008	0.025	0.014			0.287	0.334	0.299
	γ		0.021		0.024	0.029		0.121		0.124	0.127
(0.5,0.5, 0.75)	β_0	0.489	0.536	-0.001	-0.057	-0.144	0.502	0.563	0.230	0.338	0.320
	β_1	0.031	0.024	0.011	0.009	0.010	0.127	0.220	0.123	0.120	0.130
	β_2	-0.333	-0.401	-0.007	0.044	0.047	0.357	0.455	0.193	0.309	0.248
	ρ			0.036	0.036	0.002			0.303	0.439	0.340
	γ		0.078		0.086	0.087		0.120		0.128	0.129
(0.75,0.75, 0.75)	β_0	0.575	0.912	0.000	-0.460	-0.215	0.596	1.183	0.307	0.919	0.576
	β_1	0.027	-0.027	0.006	0.005	0.006	0.143	1.028	0.141	0.126	0.136
	β_2	-0.355	-0.879	-0.009	0.132	0.078	0.385	1.258	0.223	0.515	0.325
	ρ			0.130	0.323	-0.096			0.399	0.981	0.537
	γ		0.120		0.147	0.147		0.130		0.159	0.160
(0.25,0.75, 0.75)	β_0	0.513	0.802	0.034	-0.443	-0.174	0.538	1.339	0.260	0.951	0.512
	β_1	0.033	-0.004	0.013	0.007	0.009	0.145	1.396	0.142	0.119	0.129
	β_2	-0.354	-0.805	-0.025	0.143	0.073	0.387	1.598	0.222	0.576	0.335
	ρ			0.023	0.228	-0.202			0.303	1.028	0.498
	γ		0.166		0.189	0.190		0.170		0.193	0.195
(0.75,0.25, 0.75)	β_0	0.464	0.464	-0.001	-0.015	-0.120	0.475	0.475	0.243	0.286	0.309
	β_1	0.024	0.024	0.007	0.005	0.008	0.122	0.126	0.120	0.120	0.133
	β_2	-0.286	-0.286	-0.007	0.003	0.024	0.313	0.313	0.191	0.250	0.221
	ρ			-0.033	-0.012	-0.016			0.332	0.397	0.338
	γ		-0.133		-0.130	-0.125		0.188		0.188	0.187

Note: Simulation results are based on 500 replications. * In these models with no spatial dependence Heckit is theoretically consistent.

Table 4. Simulation Results for N=900, 40% Sample Selection

(δ,γ,ρ)		BIAS					RMSE				
		OLS	KP-SAE	Heckit	Spheck-E	Spheck-O	OLS	KP-SAE	Heckit	Spheck-E	Spheck-O
(0,0,0.5)*	β_0	0.448	0.447	0.008	0.005	-0.043	0.464	0.463	0.261	0.271	0.306
	β_1	0.019	0.019	0.002	0.002	0.009	0.146	0.146	0.146	0.148	0.161
	β_2	-0.276	-0.275	-0.011	-0.010	-0.005	0.311	0.310	0.207	0.223	0.230
	ρ			-0.006	-0.004	-0.035			0.258	0.261	0.265
	γ		-0.006		-0.004	-0.004		0.142		0.145	0.150
(0.25,0.25, 0.5)	β_0	0.450	0.454	-0.002	-0.003	-0.057	0.467	0.472	0.268	0.269	0.315
	β_1	0.021	0.020	0.004	0.002	0.010	0.147	0.157	0.147	0.146	0.161
	β_2	-0.275	-0.279	-0.006	-0.006	0.003	0.311	0.319	0.211	0.215	0.233
	ρ			0.007	0.002	-0.028			0.263	0.262	0.266
	γ		-0.007		-0.003	0.000		0.123		0.125	0.129
(0.5,0.5, 0.5)	β_0	0.467	0.496	0.002	-0.022	-0.059	0.484	0.530	0.287	0.365	0.338
	β_1	0.022	0.017	0.005	0.002	0.010	0.152	0.244	0.152	0.146	0.161
	β_2	-0.277	-0.317	-0.008	0.012	0.010	0.315	0.393	0.220	0.267	0.243
	ρ			0.022	0.005	-0.041			0.279	0.335	0.296
	γ		0.034		0.041	0.045		0.105		0.110	0.113
(0.75,0.75, 0.5)	β_0	0.525	0.710	0.008	-0.290	-0.090	0.558	1.051	0.365	0.846	0.546
	β_1	0.018	-0.018	0.001	-0.002	0.002	0.171	1.022	0.170	0.148	0.159
	β_2	-0.278	-0.547	-0.003	0.146	0.041	0.326	1.073	0.245	0.479	0.321
	ρ			0.083	0.153	-0.095			0.350	0.721	0.385
	γ		0.086		0.100	0.105		0.106		0.119	0.125
(0.25,0.75, 0.5)	β_0	0.494	0.691	0.014	-0.253	-0.061	0.529	1.066	0.330	0.788	0.418
	β_1	0.026	-0.003	0.008	0.005	0.011	0.172	1.041	0.172	0.154	0.164
	β_2	-0.299	-0.588	-0.010	0.127	0.030	0.350	1.205	0.259	0.482	0.288
	ρ			0.038	0.060	-0.163			0.299	0.631	0.297
	γ		0.108		0.121	0.127		0.121		0.134	0.141
(0.75,0.25, 0.5)	β_0	0.421	0.420	-0.002	-0.018	-0.066	0.438	0.438	0.303	0.353	0.399
	β_1	0.017	0.015	0.003	0.003	0.006	0.146	0.151	0.145	0.148	0.162
	β_2	-0.226	-0.221	-0.002	0.005	0.007	0.266	0.266	0.212	0.262	0.255
	ρ			-0.022	-0.007	-0.035			0.308	0.351	0.359
	γ		-0.045		-0.037	-0.034		0.138		0.136	0.139
(0,0,0.75)*	β_0	0.670	0.669	0.000	0.000	-0.146	0.678	0.678	0.251	0.252	0.337
	β_1	0.030	0.030	0.004	0.004	0.007	0.133	0.134	0.131	0.132	0.145
	β_2	-0.411	-0.410	-0.007	-0.012	0.021	0.430	0.429	0.202	0.207	0.230
	ρ			0.002	0.002	0.026			0.249	0.252	0.271
	γ		-0.001		0.003	0.007		0.141		0.144	0.150
(0.25,0.25, 0.75)	β_0	0.674	0.680	-0.006	-0.007	-0.164	0.683	0.689	0.256	0.259	0.349
	β_1	0.031	0.029	0.005	0.004	0.007	0.134	0.141	0.132	0.132	0.146
	β_2	-0.409	-0.415	-0.004	-0.006	0.030	0.429	0.436	0.203	0.210	0.233
	ρ			0.014	0.008	0.037			0.253	0.260	0.278
	γ		-0.032		-0.028	-0.023		0.126		0.127	0.130
(0.5,0.5, 0.75)	β_0	0.701	0.737	-0.001	-0.026	-0.163	0.710	0.753	0.276	0.293	0.360
	β_1	0.032	0.024	0.007	0.005	0.008	0.140	0.199	0.136	0.132	0.148
	β_2	-0.414	-0.461	-0.007	0.009	0.032	0.435	0.498	0.213	0.226	0.242
	ρ			0.038	0.015	0.030			0.272	0.291	0.296
	γ		-0.011		-0.002	-0.001		0.102		0.105	0.107
(0.75,0.75, 0.75)	β_0	0.788	1.025	0.005	-0.283	-0.204	0.804	1.173	0.348	0.720	0.587
	β_1	0.032	-0.006	0.007	0.005	0.005	0.157	0.758	0.154	0.137	0.147
	β_2	-0.418	-0.767	-0.001	0.125	0.064	0.444	1.020	0.239	0.405	0.322
	ρ			0.135	0.125	-0.060			0.357	0.621	0.420
	γ		0.045		0.072	0.068		0.084		0.105	0.105
(0.25,0.75, 0.75)	β_0	0.750	1.033	-0.005	-0.284	-0.177	0.770	1.301	0.330	0.679	0.463
	β_1	0.037	0.016	0.008	0.003	0.008	0.162	1.057	0.158	0.134	0.144
	β_2	-0.461	-0.891	-0.009	0.112	0.047	0.486	1.317	0.241	0.359	0.282
	ρ			0.097	-0.027	-0.211			0.312	0.486	0.347
	γ		0.118		0.145	0.148		0.128		0.156	0.160
(0.75,0.25, 0.75)	β_0	0.632	0.631	-0.007	-0.013	-0.154	0.641	0.640	0.293	0.297	0.390
	β_1	0.025	0.024	0.004	0.004	0.004	0.136	0.136	0.133	0.135	0.149
	β_2	-0.338	-0.334	0.001	0.001	0.025	0.361	0.358	0.207	0.217	0.236
	ρ			-0.026	-0.020	0.006			0.300	0.304	0.324
	γ		-0.151		-0.139	-0.134		0.205		0.199	0.197

Note: Simulation results are based on 500 replications. * In these models with no spatial dependence Heckit is theoretically consistent.

Table 5. Simulation Results for the Selection Equation, 25% Sample Selection

		N=324						N=900					
		BIAS			RMSE			BIAS			RMSE		
(δ,γ,ρ)		**Heckit**	**Spheck-E**	**Spheck-O**	**Heckit**	**Spheck-E**	**Spheck-O**	**Heckit**	**Spheck-E**	**Spheck-O**	**Heckit**	**Spheck-E**	**Spheck-O**
(0,0,0.5)*	α_0	-0.008	-0.021	-0.240	0.195	0.355	0.329	-0.002	-0.006	-0.216	0.120	0.170	0.252
	α_1	0.016	0.055	0.069	0.279	0.651	0.350	0.003	0.017	0.031	0.171	0.315	0.191
	α_2	0.022	0.052	0.076	0.281	0.410	0.335	0.003	0.004	0.031	0.152	0.184	0.176
	δ		-0.057	-0.055		0.313	0.301		-0.016	-0.016		0.153	0.150
(0.25,0.25, 0.5)	α_0	-0.009	-0.030	-0.256	0.199	0.351	0.349	-0.002	0.000	-0.224	0.122	0.176	0.263
	α_1	0.007	0.083	0.098	0.275	0.650	0.397	-0.006	0.028	0.044	0.172	0.330	0.224
	α_2	0.019	0.073	0.104	0.281	0.437	0.365	0.000	0.006	0.051	0.152	0.208	0.199
	δ		-0.041	-0.073		0.322	0.338		0.046	0.020		0.167	0.160
(0.5,0.5, 0.5)	α_0	0.005	0.009	-0.264	0.206	0.374	0.378	0.010	0.042	-0.233	0.128	0.202	0.281
	α_1	-0.033	0.104	0.118	0.273	0.663	0.472	-0.043	0.032	0.047	0.176	0.324	0.243
	α_2	-0.021	0.028	0.100	0.277	0.458	0.415	-0.038	-0.014	0.044	0.148	0.232	0.219
	δ		-0.089	-0.198		0.405	0.495		0.083	-0.012		0.193	0.208
(0.75,0.75, 0.5)	α_0	0.012	0.117	-0.224	0.252	0.487	0.403	0.026	0.177	-0.175	0.164	0.380	0.274
	α_1	-0.160	-0.110	-0.030	0.305	0.719	0.444	-0.143	-0.190	-0.097	0.226	0.459	0.266
	α_2	-0.127	-0.235	-0.079	0.293	0.562	0.450	-0.143	-0.306	-0.133	0.205	0.463	0.261
	δ		-0.402	-0.588		0.612	0.838		-0.386	-0.572		0.510	0.745
(0.25,0.75, 0.5)	α_0	-0.026	0.077	-0.252	0.199	0.454	0.370	-0.013	0.122	-0.213	0.120	0.319	0.257
	α_1	0.007	0.052	0.122	0.277	0.712	0.436	-0.003	-0.039	0.040	0.172	0.396	0.220
	α_2	0.026	-0.088	0.069	0.274	0.557	0.436	0.014	-0.156	0.024	0.156	0.399	0.202
	δ		0.077	-0.141		0.418	0.590		0.099	-0.083		0.275	0.435
(0.75,0.25, 0.5)	α_0	0.050	0.032	-0.190	0.256	0.389	0.340	0.045	0.052	-0.171	0.170	0.199	0.244
	α_1	-0.153	-0.079	-0.072	0.297	0.636	0.374	-0.154	-0.149	-0.124	0.229	0.313	0.229
	α_2	-0.136	-0.098	-0.068	0.302	0.379	0.338	-0.147	-0.145	-0.113	0.209	0.222	0.208
	δ		-0.586	-0.601		0.670	0.686		-0.506	-0.523		0.530	0.544
(0,0,0.75)*	α_0	-0.008	-0.018	-0.218	0.194	0.334	0.312	-0.003	-0.007	-0.191	0.119	0.172	0.230
	α_1	0.017	0.054	0.058	0.277	0.605	0.357	0.003	0.017	0.010	0.170	0.324	0.186
	α_2	0.018	0.044	0.055	0.280	0.381	0.329	0.003	0.006	0.013	0.152	0.173	0.164
	δ		-0.049	-0.050		0.322	0.324		-0.017	-0.019		0.152	0.150
(0.25,0.25, 0.75)	α_0	-0.008	-0.013	-0.234	0.197	0.339	0.339	-0.002	-0.001	-0.199	0.122	0.182	0.243
	α_1	0.008	0.052	0.092	0.272	0.599	0.431	-0.005	0.021	0.026	0.172	0.342	0.228
	α_2	0.016	0.073	0.106	0.281	0.400	0.409	0.000	0.008	0.033	0.152	0.192	0.209
	δ		-0.060	-0.083		0.331	0.367		0.025	0.014		0.166	0.158
(0.5,0.5, 0.75)	α_0	0.004	0.023	-0.252	0.206	0.361	0.391	0.009	0.042	-0.221	0.129	0.207	0.279
	α_1	-0.032	0.072	0.153	0.273	0.617	0.555	-0.041	0.039	0.090	0.176	0.338	0.329
	α_2	-0.022	0.020	0.139	0.277	0.413	0.522	-0.037	-0.013	0.086	0.148	0.212	0.296
	δ		-0.082	-0.169		0.380	0.494		0.077	0.031		0.198	0.239
(0.75,0.75, 0.75)	α_0	0.015	0.159	-0.192	0.250	0.485	0.402	0.031	0.259	-0.141	0.164	0.440	0.275
	α_1	-0.153	-0.126	-0.036	0.301	0.680	0.455	-0.146	-0.237	-0.101	0.225	0.515	0.347
	α_2	-0.120	-0.267	-0.101	0.285	0.540	0.447	-0.145	-0.362	-0.158	0.206	0.494	0.321
	δ		-0.341	-0.555		0.546	0.821		-0.280	-0.547		0.422	0.751
(0.25,0.75, 0.75)	α_0	-0.032	0.125	-0.221	0.198	0.474	0.353	-0.017	0.225	-0.161	0.118	0.429	0.228
	α_1	0.006	0.002	0.098	0.279	0.682	0.426	0.005	-0.143	0.009	0.170	0.472	0.220
	α_2	0.042	-0.133	0.055	0.272	0.547	0.453	0.015	-0.252	-0.033	0.162	0.459	0.240
	δ		0.107	-0.147		0.393	0.612		0.078	-0.143		0.268	0.487
(0.75,0.25, 0.75)	α_0	0.044	0.034	-0.171	0.253	0.365	0.328	0.044	0.046	-0.151	0.169	0.200	0.229
	α_1	-0.150	-0.109	-0.100	0.298	0.567	0.363	-0.153	-0.149	-0.143	0.229	0.308	0.239
	α_2	-0.132	-0.098	-0.082	0.299	0.363	0.362	-0.147	-0.145	-0.131	0.208	0.215	0.213
	δ		-0.692	-0.696		0.764	0.773		-0.630	-0.635		0.647	0.654

Note: Simulation results are based on 1000 replications for N=324 and 500 for N=900. * In these models with no spatial dependence Heckit is theoretically consistent.

Table 6. Simulation Results for the Selection Equation, 40% Sample Selection

		N=324						N=900					
		BIAS			RMSE			BIAS			RMSE		
(δ,γ,ρ)		**Heckit**	**Spheck-E**	**Spheck-O**	**Heckit**	**Spheck-E**	**Spheck-O**	**Heckit**	**Spheck-E**	**Spheck-O**	**Heckit**	**Spheck-E**	**Spheck-O**
(0,0,0.5)*	α_0	-0.010	-0.027	-0.251	0.190	0.259	0.340	-0.003	-0.003	-0.216	0.116	0.129	0.248
	α_1	0.009	0.041	0.046	0.253	0.408	0.316	0.006	0.004	0.015	0.156	0.199	0.166
	α_2	0.021	0.032	0.057	0.255	0.296	0.292	0.005	0.010	0.021	0.146	0.150	0.158
	δ		-0.057	-0.054		0.286	0.284		-0.006	-0.007		0.141	0.134
(0.25,0.25, 0.5)	α_0	-0.005	-0.035	-0.267	0.196	0.272	0.371	0.005	0.006	-0.215	0.119	0.132	0.250
	α_1	0.002	0.058	0.060	0.249	0.429	0.335	-0.006	-0.008	0.009	0.152	0.192	0.162
	α_2	0.015	0.041	0.071	0.257	0.314	0.317	-0.004	0.004	0.019	0.143	0.148	0.156
	δ		-0.066	-0.089		0.282	0.295		-0.004	-0.025		0.138	0.128
(0.5,0.5, 0.5)	α_0	0.031	-0.025	-0.300	0.208	0.304	0.462	0.031	0.015	-0.245	0.133	0.173	0.317
	α_1	-0.046	0.078	0.095	0.252	0.486	0.460	-0.040	0.012	0.036	0.161	0.251	0.250
	α_2	-0.023	0.029	0.101	0.252	0.354	0.435	-0.039	-0.007	0.036	0.150	0.216	0.253
	δ		-0.093	-0.163		0.348	0.422		0.032	-0.035		0.171	0.161
(0.75,0.75, 0.5)	α_0	0.115	0.110	-0.234	0.275	0.460	0.500	0.120	0.160	-0.162	0.201	0.336	0.295
	α_1	-0.156	-0.029	0.010	0.296	0.597	0.438	-0.150	-0.092	-0.063	0.219	0.324	0.250
	α_2	-0.134	-0.171	-0.021	0.279	0.512	0.488	-0.149	-0.208	-0.101	0.206	0.385	0.269
	δ		-0.366	-0.526		0.614	0.813		-0.242	-0.369		0.480	0.583
(0.25,0.75, 0.5)	α_0	-0.014	0.002	-0.362	0.199	0.400	0.534	0.001	0.052	-0.280	0.124	0.290	0.348
	α_1	0.003	0.126	0.177	0.251	0.581	0.489	0.001	0.046	0.085	0.158	0.363	0.268
	α_2	0.025	-0.044	0.142	0.262	0.510	0.525	0.000	-0.058	0.051	0.153	0.354	0.244
	δ		0.155	-0.008		0.502	0.590		0.286	0.110		0.464	0.454
(0.75,0.25, 0.5)	α_0	0.138	0.112	-0.116	0.287	0.325	0.317	0.125	0.124	-0.095	0.204	0.208	0.204
	α_1	-0.157	-0.104	-0.100	0.295	0.452	0.370	-0.153	-0.149	-0.137	0.221	0.241	0.230
	α_2	-0.142	-0.123	-0.095	0.290	0.309	0.319	-0.153	-0.148	-0.135	0.211	0.214	0.221
	δ		-0.619	-0.633		0.683	0.704		-0.548	-0.562		0.565	0.577
(0,0,0.75)*	α_0	-0.011	-0.029	-0.232	0.190	0.247	0.334	-0.003	0.001	-0.185	0.116	0.126	0.222
	α_1	0.009	0.049	0.038	0.253	0.399	0.336	0.007	-0.002	-0.008	0.156	0.185	0.163
	α_2	0.022	0.029	0.044	0.255	0.279	0.294	0.004	0.007	-0.001	0.146	0.149	0.156
	δ		-0.057	-0.049		0.293	0.320		-0.002	-0.003		0.141	0.137
(0.25,0.25, 0.75)	α_0	-0.006	-0.034	-0.257	0.196	0.261	0.391	0.005	0.007	-0.191	0.119	0.137	0.242
	α_1	0.003	0.053	0.067	0.249	0.394	0.387	-0.006	-0.010	-0.006	0.152	0.189	0.177
	α_2	0.016	0.040	0.073	0.255	0.296	0.358	-0.004	0.001	0.006	0.143	0.154	0.192
	δ		-0.098	-0.098		0.298	0.334		-0.025	-0.032		0.138	0.157
(0.5,0.5, 0.75)	α_0	0.031	-0.012	-0.299	0.206	0.296	0.477	0.032	0.023	-0.257	0.133	0.170	0.370
	α_1	-0.041	0.062	0.119	0.251	0.466	0.511	-0.040	0.004	0.070	0.160	0.234	0.328
	α_2	-0.026	0.019	0.116	0.251	0.331	0.470	-0.040	-0.019	0.070	0.148	0.174	0.331
	δ		-0.092	-0.139		0.326	0.433		0.025	0.007		0.162	0.195
(0.75,0.75, 0.75)	α_0	0.121	0.129	-0.234	0.277	0.443	0.535	0.120	0.183	-0.184	0.201	0.355	0.391
	α_1	-0.155	-0.039	0.034	0.293	0.540	0.509	-0.149	-0.103	-0.023	0.218	0.334	0.366
	α_2	-0.132	-0.172	-0.006	0.280	0.459	0.525	-0.150	-0.220	-0.071	0.207	0.368	0.382
	δ		-0.260	-0.413		0.497	0.720		-0.125	-0.280		0.324	0.518
(0.25,0.75, 0.75)	α_0	-0.021	0.034	-0.347	0.199	0.396	0.542	0.001	0.106	-0.267	0.120	0.310	0.396
	α_1	0.006	0.093	0.176	0.253	0.539	0.515	-0.001	-0.001	0.093	0.156	0.337	0.368
	α_2	0.032	-0.072	0.129	0.262	0.495	0.559	-0.003	-0.119	0.047	0.145	0.346	0.330
	δ		0.222	0.046		0.484	0.584		0.263	0.096		0.484	0.504
(0.75,0.25, 0.75)	α_0	0.136	0.121	-0.094	0.286	0.312	0.306	0.125	0.124	-0.070	0.204	0.206	0.186
	α_1	-0.158	-0.124	-0.122	0.294	0.400	0.353	-0.153	-0.153	-0.158	0.221	0.236	0.236
	α_2	-0.140	-0.127	-0.109	0.290	0.303	0.330	-0.153	-0.148	-0.152	0.211	0.212	0.223
	δ		-0.713	-0.710		0.770	0.781		-0.652	-0.655		0.667	0.670

Note: Simulation results are based on 1000 replications for N=324 and 500 for N=900. * In these models with no spatial dependence Heckit is theoretically consistent.

Table 7. Estimated Coefficients for the Empirical Application

	Outcome Equation[1]					Selection Equation[3]		
	OLS	KP-SAE[2]	Heckit	Spheck-E	Spheck-O	Heckit	Spheck-E	Spheck-O
Constant	7.562 ***	7.342 ***	5.003 *	5.011	5.063 **	-0.104	-0.135	-0.098
	(0.451)	(0.332)	(2.592)	(22.350)	(2.080)	(0.648)	(19.720)	(0.756)
Max. Depth	0.071	0.025	0.316	0.326	0.340 *	0.179 *	0.197	0.232 **
	(0.068)	(0.052)	(0.284)	(0.864)	(0.201)	(0.092)	(1.201)	(0.116)
Min. Depth	0.011	-0.058	-0.108	-0.069	-0.088	-0.093	-0.088	-0.133 *
	(0.050)	(0.044)	(0.160)	(0.352)	(0.090)	(0.068)	(0.572)	(0.073)
Biomass	0.201 ***	0.202 ***	0.181	0.192	0.158 *	0.005	0.005	0.007
	(0.055)	(0.052)	(0.114)	(0.500)	(0.091)	(0.078)	(0.842)	(0.072)
Dum CV	1.316 ***	2.514 ***	0.013	0.019	0.029	-0.739 ***	-0.696	-0.736 ***
	(0.214)	(0.207)	(1.244)	(5.350)	(0.481)	(0.183)	(1.883)	(0.257)
Dum HAL	0.102	0.793	1.074	0.715	0.921	0.650 ***	0.475	0.581 **
	(0.205)	(0.179)	(0.966)	(10.400)	(0.752)	(0.202)	(3.281)	(0.269)
Dum NPT	-0.542 **	0.210 **	-0.339	-0.411	-0.354	0.073	0.078	0.080
	(0.237)	(0.262)	(0.466)	(8.051)	(0.537)	(0.261)	(4.654)	(0.313)
Dum Large	0.600 ***	0.665 ***	0.470	0.452	0.604 **	-0.078	-0.098	-0.088
	(0.131)	(0.165)	(0.303)	(3.204)	(0.257)	(0.176)	(2.063)	(0.177)
IMR			2.909	2.113	2.493 *			
			(2.676)	(13.870)	(1.502)			
Lag Biomass						-0.043	-0.043	-0.069
						(0.080)	(1.064)	(0.066)
SAE parameter (γ)		0.912 *		0.950	0.917 ***			
		(0.509)		(1.165)	(0.090)			
SAE parameter (δ)							0.392	0.203
							(1.020)	(0.133)

Notes: Sample size is 320 with 35% selection. Standard errors in parentheses; *, **, *** significant at the 10%, 5%, and 1% level, respectively.
[1] Dependent variable is the log of the average catch-per-unit-effort (CPUE) for the statistical reporting regions in the Eastern Bering Sea
[2] The standard errors for KP-SAE are computed following Kelejian and Prucha (2009)
[3] Dependent variable is whether or not CPUE is observed for that unit