# A TWO-STEP ESTIMATOR FOR A SPATIAL LAG MODEL OF COUNTS: THEORY, SMALL SAMPLE PERFORMANCE AND AN APPLICATION

by

Dayton M. Lambert, Jason P. Brown,
Raymond J.G.M. Florax

Working Paper #10-5

March 2010

**Dept. of Agricultural Economics**

**Purdue University**

# A TWO-STEP ESTIMATOR FOR A SPATIAL LAG MODEL OF COUNTS: THEORY, SMALL SAMPLE PERFORMANCE AND AN APPLICATION

by

**Dayton M. Lambert**[a]

[a] Dept. of Agricultural Economics, University of Tennessee
321 Morgan Hall, Knoxville, TN 37996–4518, U.S.A.
E-mail: dmlambert@tennessee.edu

**Jason P. Brown**[b,*]

[b] USDA, Economic Research Service
1800 M St. NW, Washington, D.C., 20036–5831, U.S.A.
E-mail: jbrown@ers.usda.gov

**Raymond J.G.M. Florax**[c,d]

[c] Dept. of Agricultural Economics, Purdue University
403 W. State Street, West Lafayette, IN 47907–2056, U.S.A.
E-mail: rflorax@purdue.edu

[d] Dept. of Spatial Economics, VU University Amsterdam
De Boelelaan 1105, 1081 HV Amsterdam, The Netherlands

Working Paper#10-5
March 2010

**Abstract**

Several spatial econometric approaches are available to model spatially correlated disturbances in count models, but there are at present no structurally consistent count models incorporating spatial lag autocorrelation. A two-step, limited information maximum likelihood estimator is proposed to fill this gap. The estimator is developed assuming a Poisson distribution, but can be extended to other count distributions. The small sample properties of the estimator are evaluated with Monte Carlo experiments. Simulation results suggest that the spatial lag count estimator achieves gains in terms of bias over the aspatial version as spatial lag autocorrelation and sample size increase. An empirical example deals with the location choice of single-unit start-up firms in the manufacturing industry in the US between 2000 and 2004. The empirical results suggest that in the dynamic process of firm formation, counties dominated by firms exhibiting (internal) increasing returns to scale are at a relative disadvantage even if localization economies are present.

Keywords: count model, location choice, manufacturing, Poisson, spatial econometrics
JEL code: C21, C25, D21, R12, R30

---

* Corresponding author. The views expressed here are those of the authors, and may not be attributed to the Economic Research Service or the U.S. Department of Agriculture.

## 1. Introduction

In this paper, we formulate a two-step estimator for a spatial autoregressive lag model of counts. The development of this estimator was inspired by the lack of an interpretable spatial lag process in spatial autoregressive count models, and the pivotal role these models play in the firm investment location choice literature. Firm location studies typically focus on plant birth, and start from an enumeration of births observed in designated geographical areas over some specified time period. Guimarães et al. (2003, 2004) provided the theoretical foundations behind firm site selection in the context of a random profit maximization model. They showed that under specific assumptions, aggregate site selection decisions of firms representing a single sector could be modeled as discrete counts determined by local or regional factors. Their contribution continues to motivate the theoretical framework in terms of probabilistic models of many recent empirical studies of business establishment site selection.

Similar to the classical firm location literature of Weber (1929), optimal firm site selection is a trade-off between transport costs of inputs to production facilities, and outputs to product markets. Firms choose least cost sites to maximize profit ($\pi$). For example, firm $i$ chooses location $j$ over $k$ competing locations when $\pi_{ij} > \pi_{ik}$. From this relationship, the probability firm $i$ chooses location $j$ is $\Pr(\pi_{ij} > \pi_{ik}) = \Pr(z_j = 1)$. The probability that firm type $i$ selects location $j$ can be estimated as a conditional logit model, assuming the stochastic components follow a Weibull distribution and are independent and identically distributed (McFadden, 1974). But as Guimarães et al. (2003, 2004) demonstrate, the conditional logit model can be estimated by a Poisson regression,1 provided that the empirical analysis focuses on a single sector or firm type. This fact has provided the conceptual underpinning for a number of empirical studies explaining firm location events, many of which rely on aggregate cross-sectional data (e.g., Fotopoulos and Louri, 2000; Henderson and McNamara, 2000; Guimarães et al., 2004; Carod and Antolín, 2004; Davis and Schluter, 2005; Carod, 2005; Chong, 2006; Lambert et al., 2006a,b; Lambert and McNamara, 2009). However, these studies have rarely addressed the possibility that firm site selection in a given location may be simultaneously determined with firm location events in neighboring locations. To the extent that such spatial spillovers proxy localization economies, omitting such information may lead to biased or inconsistent estimates of the impacts local factors have on attracting or retaining business investment. In addition, modeling the correlation between firm investment flows may provide a richer, more detailed picture of the regional linkages supporting local growth, industry clustering, and economic development.

The limited empirical research investigating spatial interaction between firm location events as a *global* spatial autocorrelation process may be due to the dearth of research on cross-sectional count lag autoregressive process in general. The theory and application of spatial econometrics in discrete or limited dependent variable settings is relatively less

[1] Alternatively, with additional assumptions, other discrete distributions are feasible as well.

developed compared to the standard cross-sectional and panel data settings, although with some important exceptions. To the extent that such research exists, it has focused largely on the limited dependent variable class of estimators. For certain types of discrete spatial processes, such as binary or multinomial spatial logit or probit models, ordered versions of these models, or spatial tobit models, the spatial autoregressive lag specification has received more attention. Estimation is typically based on nonlinear generalized method of moments (GMM) techniques (Pinske and Slade, 1998; Klier and McMillen 2008), conditional autoregressive general linear models (Schabenberger and Pierce, 2002), and Bayesian simulation approaches (LeSage, 2000). Fleming (2004) and LeSage and Pace (2009, Chapter 10) provide comprehensive reviews of these methods. Previous work focusing explicitly on spatial dependence in count models includes the Kaiser and Cressie (1997) count model, conditional autoregressive models incorporating neighborhood contiguity (Rasmussen, 2004) or direct representation of error processes (Schabenberger and Pierce, 2002), geographically weighted regression in Fotheringham et al. (2002), information theoretic approaches as in Bhati (2005), spatial filtering by Griffith (2002, 2003; Haining et al., 2009), and Bayesian hierarchical methods (LeSage et al., 2007; Flores et al., 2009). However, these established methods for modeling spatial dependence in count outcomes do not clearly address *global* spatial autocorrelation arising from cross-sectional dependence between counts.

This paper attempts to fill this gap, suggesting a count estimator that models the response variable (or location events) as a function of neighboring counts. For reasons discussed below, the proposed Spatial Autoregressive Poisson (SAR-Poisson) model we investigate in a series of Monte Carlo experiments and apply in an empirical example is estimated using a two-step limited information maximum likelihood (LIML) approach. The SAR-Poisson model includes a spatially lagged dependent variable as a covariate, and can be seen as the counterpart of what is usually referred to as the "spatial lag model". The proposed estimator meets the double challenge of maintaining the distributional assumptions required by the probability-based firm location model of Guimarães et al. (2003, 2004) as well as theoretical consistency with the linear spatial autoregressive lag model typically applied in the spatial econometric literature. The latter enables one to decompose the effects of location determinants into direct, indirect, and induced effects by way of the "Leontief inverse" or "spatial multiplier", which provides information that may be important for understanding regional linkages.

## 2. Spatial autoregressive lag model of counts

Whittle's (1954) class of linear spatial process models provides the motivation behind the SAR-Poisson estimator proposed here. Whittle defined a family of linear spatial process models whereby an endogenous variable is specified to depend on spatial interactions between cross-sectional units plus a disturbance term. The interactions are modeled as a weighted average of neighboring observations. The endogenous variable comprising the interactions is referred to as the "spatially lagged dependent variable". The weights, grouped in a matrix identifying neighbor relations by means of contiguity or distance, form the distinctive core of the class of spatial process models.

The linear spatial autoregressive lag model, or SAR model, is given by:

(1) $y = \rho Wy + X\beta + \varepsilon,$

where $y$ is an $N \times 1$ stochastic variate, $X$ and $\beta$ have conformable dimensions for $k$ exogenous covariates including a constant, $\rho$ is an autoregressive parameter, $W$ an $N$ by $N$ non-stochastic and *a priori* defined weights matrix specifying the relationships between spatial units, and $\varepsilon$ a vector of independently and identically distributed disturbances.[2] The "reduced-form" version of the SAR model reads as:

(2) $y = A^{-1}X\beta + A^{-1}\varepsilon,$

where $A = I - \rho W$, $I$ a conformable identity matrix, and $A^{-1}$ the spatial multiplier or "Leontief inverse" (Anselin, 2002). The inverted $A$ matrix relays feedback/feed-forward effects of shocks between locations, thereby distinguishing this class of models from local regression models. Most empirical applications assume a linear relationship between the response and explanatory variables, but the system is flexible enough to accommodate nonlinear dynamic spatial processes.[3] This basic model is generalized to include count processes below.

The SAR count model suggested here is motivated by previous work on estimating *temporally* lagged count processes. Recall the data generating process for a random variable following a Poisson distribution, with the probability mass function $f$ and mean $\mu$ is,

(3) $f(y_i) = \frac{\mu_i^{y_i} e^{-\mu_i}}{y_i!}.$

The expected conditional mean of a random variable following a Poisson distribution is typically represented by the inverse of the logarithmic canonical link function:

(4) $\mathrm{E}(y_i) \equiv \mu_i = \exp\left(\beta' x_i\right),$

where $x_i$ is a $k \times 1$ vector of covariates containing measurements on observation $i = 1, 2, \ldots, N$ (Cameron and Trivedi, 1998, 2005). There are a number of methods available to incorporate temporal lags into the conditional mean function given in (4). Cameron and Trivedi (1998) provide an extensive review of this work. Two examples include Blundell et al.'s (1995) exponential feedback model, and Zeger and Qaqish's (1988) multiplicative AR model. We focus on these two models as candidate starting points in developing a SAR-Poisson structural model, as they provide intuitive links to the general linear SAR specification.

[2] Details for the SAR model with heteroskedastic errors are described in Anselin (1988, 2006). Whittle's SAR model was popularized and extended by Cliff and Ord (1981), who further developed models in which the disturbances follow a spatial autoregressive process. The count regression formulation of the error process model is not considered here. See Rasmussen (2004), Schabenberger and Pierce (2002), and Lambert et al. (2006a) for conditional spatially autoregressive error models of count processes.

[3] One should note that a polynomial expansion of the inverse in the reduced form shows that the SAR model is actually nonlinear in the parameters.

The exponential feedback model for time-series data is given by:

5) $\mu_t = \exp\left(\rho y_{t-1} + \beta' x_t\right).$

The model can be estimated with maximum likelihood, eventually with hetero-skedasticity and autocorrelation consistent standard errors (Cameron and Trivedi, 1998). An intuitive extension of the exponential feedback model to the count version of the SAR process model is:

$$(6) \qquad \mu_i = \exp\left(\rho \sum_{j \neq i}^{N} w_{ij} y_j + \beta' x_i\right),$$

where $w_{ij}$ is the $ij$-th element of the spatial weights matrix $W$. This is in fact the specification used by Kaiser and Cressie (1997) and Griffith (2003), which leads to the conditional autoregressive (CAR) Poisson model with spatial dependence between counts.[4] However, the extent to which *global* feedback between locations is explicitly modeled is not clear in this specification. The algebraic equivalent of the Leontief inverse, which distinguishes spatial autoregressive lag models from their aspatial counterparts, is intractable. The usefulness of this particular specification to model global interactions between spatial neighbors is therefore rather limited. Cameron and Trivedi (1998) arrive at a similar conclusion, questioning the usefulness and applicability of the exponential feedback-type models in the context of modeling lagged time series counts.

The distinguishing feature of the SAR count process must lay in the specification of the expected mean of counts at location $i$ as a function of its $j$ neighbors. Again, work in modeling temporally lagged counts serves as a useful starting point. A more plausible autoregressive model for temporally lagged count responses specifies a multiplicative relationship between a predetermined count and future outcomes (Zeger and Qaqish, 1988), as in:

$$(7) \qquad E(y_t) = \mu_t = \exp\left(\beta' x_t\right) y_{t-1}^{\rho}.$$

Cameron and Trivedi (1998) suggest two approaches to overcome the obvious problem of zero counts on the right-hand side of this equation. Both approaches entail transforming the lagged count variable. The first approach adds a small, pre-fixed constant to the lagged outcome variable for zero-values of $y_{t-1}$. As a result, $y_{t-1}$ in equation (7) is replaced by $y_{t-1}^{*} = \max\{c, y_{t-1}\}$, $0 < c \leq 1$. The second approach is similar, but estimates the value of the constant $c$ simultaneously with the other model parameters. Effectively, this approach is based on the following model:

$$(8) \qquad \mu_t = \exp\left(\beta' x_t + \rho \ln y_{t-1}^{**} + c^{**} d_t\right), \text{ with}$$

$$(9) \qquad y_{t-1}^{**} = \begin{cases} y_{t-1} & y_{t-1} > 0 \\ 1 & y_{t-1} = 0 \end{cases} \text{ and } d_t = \begin{cases} 0 & y_{t-1} > 0 \\ 1 & y_{t-1} = 0 \end{cases},$$

[4] The conditional autoregressive (CAR) Poisson model is notorious in that the model can only accommodate autoregressive processes with negative feedback. To overcome this limitation, Kaiser and Cressie (1997) suggested a Winsorized version, which allows for positive conditional autocorrelation. Griffith (2003) extended their model using spatial filters.

where the constant $c$ is derived from the unrestricted estimate $c^{**} = \rho \ln c$ as $c = \exp\left(c^{**}/\rho\right)$. This model can be estimated using maximum likelihood or quasi-maximum likelihood techniques assuming a Poisson distribution (Greene, 2003, p. 188).

The situation is more complicated when outcomes are simultaneously determined as a single realization, as is typically observed in cross-sectional data with spatial dependence. The spatial analogue of the above multiplicative autoregressive time series model for count data is:

$$(10) \qquad \mathrm{E}(y_i) \equiv \mu_i = \exp\left(\beta' x_i\right) \prod_{j \neq i}^{N} \mathrm{E}(y_j)^{\rho w_{ij}}.$$

This result is similar to the case of binary lagged dependent variables, where the latent variable is a function of the simultaneous realization of its neighbors. As before, the elements of the weights matrix $w_{ij}$ are *a priori* determined exogenous constants, usually scaled over $j$ to sum to unity for each $i$. Moving the multiplicative component inside the exponential term leads to the structural model:

$$(11) \qquad \mu_i = \exp\left( \rho \sum_{j \neq i}^{N} w_{ij} \mu_j + \beta' x_i \right).$$

Expressing equation (11) compactly in matrix notation and including all spatial units leads to the reduced form of the conditional mean function:

$$(12) \qquad \mu_i^{SAR} = \exp\left(A_{i.}^{-1} X \beta\right),$$

where $A_{i.}^{-1}$ is the $i$-th row of the spatial multiplier term.

The usual aspatial Poisson marginal effects estimator $\partial \mu_i / \partial x_{ik} = \beta_k \exp(\beta' x_i)$ can be extended to the spatial Poisson count specification of (12). The marginal effects representing the change in the expected value at location $i$ following a change in covariate $x_k$ in one or more locations $j$ are derived as:

$$(13) \qquad \frac{\partial \mu_i^{SAR}}{\partial x_{ik}} = A_{i.}^{-1} \exp(A_{i.}^{-1} X \beta) \beta_k.$$

Because of the autoregressive nature of the model, the total marginal effects can be decomposed into direct and indirect effects (LeSage and Pace, 2009) as:

$$(14) \qquad \left( a_{ii}^{-1} \mu_i^{SAR} + \sum_{j \neq 1}^{N} a_{ij}^{-1} \mu_j^{SAR} \right) \beta_k,$$

where $a_{ii}^{-1}$ ($a_{ij}^{-1}$) refers to diagonal (off-diagonal) elements of the estimated spatial multiplier matrix $A^{-1}$. Average marginal effects can be calculated in different ways (Cameron and Trivedi, 1998, p. 80), but one can aggregate over $i$ and divide by $N$, arriving at:

$$(15) \qquad \left.\frac{\partial \mu^{SAR}}{\partial x_k}\right|_{\text{direct}} = \frac{\beta_k}{N} \sum_{j \neq i}^{N} a_{ii}^{-1} \mu_i^{SAR}, \text{ and}$$

(16) $$\left.\frac{\partial \mu^{SAR}}{\partial x_k}\right|_{\text{indirect}} = \frac{\beta_k}{N}\sum_{i=1}^{N}\sum_{j\neq i}^{N} a_{ij}^{-1}\mu_j^{SAR}.$$

Further research into the computation and interpretation of the marginal effects in a SAR Poisson model is warranted given their highly nonlinear nature.

Elasticities are slightly easier to calculate and interpret since the exponential term in (13) referring to the expected value cancels out, which leads to:

(17) $$\eta_{ik}^{SAR} = A_{i.}^{-1}x_k\beta_k,$$

which, due to the autoregressive nature of the model, can be decomposed into the "own" elasticity and "cross" elasticities. Sample-wide averages of these elasticities are estimated as:

(18) $$\left.\eta_k^{SAR}\right|_{\text{own}} = \frac{\beta_k}{N}\sum_{j\neq i}^{N} a_{ii}^{-1}x_{ik}, \text{ and}$$

(19) $$\left.\eta_k^{SAR}\right|_{\text{cross}} = \frac{\beta_k}{N}\sum_{i=1}^{N}\sum_{j\neq i}^{N} a_{ij}^{-1}x_{jk}.$$

## 3. Estimation of the SAR-Poisson count model

Below we derive the Full Information Maximum Likelihood (FIML) and a two-step, Limited Information Maximum Likelihood (LIML) estimator for the proposed SAR-Poisson model. Alternative approaches might include estimating this model using weighted nonlinear least squares or general method of moment-type estimators (Cameron and Trivedi, 1998; see also Hays and Franzese, 2009 for a recent exposition). These estimators may be flexible in terms of accommodating various types of cross-sectional lag autoregressive processes, but it is not clear whether the theoretical links with the random profit maximization framework frequently applied in location studies are maintained.[5]

### *3.1 Full Information Maximum Likelihood Estimation of the SAR-Poisson model*

The necessary and sufficient conditionals of the FIML approach are derived to provide a reference point for the LIML estimator developed below. Derivation of the SAR-Poisson FIML estimator begins with the Poisson probability density function:

(21) $$f\left(y \mid x, W; \beta, \rho\right) = \frac{(\mu_i^{SAR})^{y_i}\exp(-\mu_i^{SAR})}{y_i!},$$

with $\mu_i^{SAR} = \exp(A_{i.}^{-1}X\beta)$. The log-likelihood function is:

[5] The moment conditions of the FIML estimator could be used in the weighted nonlinear least squares or general method of moments estimation framework, but the underlying error distribution ensuring the theoretical link between the probability-based random profit maximization model and observed firm behavior may be compromised if the choice of spectral weights is incorrect. In the case of the usual Poisson estimator, the optimal weighting scheme for general method of moments estimator would be the Hessian of the maximum likelihood estimator (Cameron and Trivedi, 2005). In this case, general method of moments results are identical to full information maximum likelihood estimates. Other weighting schemes are suboptimal, and the extent to which the data generating process is still Poisson is unclear.

(22) $$\ln L = y'A^{-1}X\beta - \mu^{SAR} - \sum_i \ln y_i,$$

with the necessary conditions:

(23) $$\frac{\partial \ln L}{\partial \beta} = X'A^{-1}(y - \mu^{SAR}) = 0, \text{ and}$$

(24) $$\frac{\partial \ln L}{\partial \rho} = \beta' X'D[y - \mu^{SAR}] = 0,$$

with $D = A^{-1}WA^{-1}$. The Hessian of the system is:

(25) $$\frac{\partial^2 \ln L}{\partial \beta \partial \beta'} = -X'A^{-1}\Omega A^{-1}X, \text{with } \Omega = diag[\mu_i^{SAR}],$$

(26) $$\frac{\partial^2 \ln L}{\partial \beta \partial \rho} = \beta' X'D\Omega A^{-1}X - [y - \mu^{SAR}]'DX \text{ , and}$$

(27) $$\frac{\partial^2 \ln L}{\partial \rho^2} = 2\beta' X'A^{-1}WD[y - \mu^{SAR}] - \beta' X'D\Omega DX\beta .$$

While analytic expressions for the necessary and sufficient conditions of the FIML estimator may be derived, numerical solutions using a variety of algorithms were difficult to obtain consistently. The usual optimization algorithms were too frequently unsuccessful with respect to generating negative semi-definite Hessians over an adequate number of Monte Carlo replications to infer any meaningful conclusions. For this reason, we focus explicitly on the two-step SAR-Poisson LIML estimator developed below.

*3.2 Two-step LIML estimation of the SAR-Poisson model*

There are numerous examples of nonlinear models where endogenous parameters are estimated using two-step limited information procedures (see the examples in Murphy and Topel, 1985; Greene, 2003, p. 508; Cameron and Trivedi, 2005). A two-step, limited information maximum likelihood approach to estimate the SAR count model is proposed here. If the FIML numerical search algorithms had succeeded, then the problem of adjusting the count variable to accommodate the right-hand side logarithmic transformation in the structural equation would not be an issue. However, resorting to the LIML estimator requires some attention to the problem of zero counts, and transformation of the count variable for it to be consistent with a global autoregressive spatial process. Below, the SAR-Poisson LIML estimator using a Murphy-Topel (1985) covariance adjustment is developed first, assuming a suitable solution to the zero count problem has been made.

The first step naturally entails replacing the lagged, log-transformed counts in the neighborhood of $y_i$ with their predicted values. Anselin (1988) and Kelejian and Prucha (1999) provide some guidance with respect to consistent estimation of the endogenous lagged variable *Wy* in the linear class of SAR estimators, and that approach is adopted here. Let the function $g(y_j^*)$ represent the logged-transformed values approximating neighboring counts. While it is not necessary to make any rigorous distributional assumptions about the first-stage linear predictor, it is useful to formulate the problem with reference to a log-likelihood function to define notation used in deriving the second-stage covariance adjustment used later. Denote the log-likelihood function of the (linear) first-stage estimator as:

$$\text{(28)} \qquad \ln L_1 = \sum_{i=1}^{n} f_1\left(W \cdot g\left(y_j^*\right) | Q_i ; \delta\right),$$

with the gradient vector $h_1 = Q'\upsilon_1$, $\upsilon_1$ a residual vector, $f_1$ the normal probability density function, and $\delta$ a vector of parameters that maximizes $L_1$. Given a set of appropriately defined instrumental variables, typically $Q$ = [$X$, $WX$, $WWX$], regressing the instruments on the transformation yields the vector of predicted values:

$$\text{(29)} \qquad Q\delta, \text{ with } \delta = Q(Q'Q)^{-1} Q'W \cdot g(y_j^*).$$

The corresponding vector of residuals is estimated as $\upsilon_1 = W \cdot g(y_j^*) - Q\delta$. A heteroskedasticity-robust covariance estimator, $V_1(\delta) = \left(Q'Q\right)^{-1} Q'\Omega Q \left(Q'Q\right)^{-1}$, may be applied to obtain the first-stage standard errors needed to adjust the second-stage covariance matrix (see below). The diagonal elements of $\Omega$ are the squares of the residuals from the first-stage regression, with off-diagonal elements of zero.

In the second step, the first-stage predicted values enter the Poisson probability density function as:

$$\text{(30)} \qquad f_2\left(y \mid x, W, \delta' Q_i ; \beta, \rho\right) = \frac{\exp\left(\beta' x_i + \rho \cdot \delta' Q_i\right)^{y_i} \exp\left(-\exp\left(\beta' x_i + \rho \cdot \delta' Q_i\right)\right)}{y_i !},$$

with the corresponding log-likelihood function:

$$\text{(31)} \qquad \ln L_2 = \sum_{i=1}^{N} y_i \left(\beta' x_i + \rho \cdot Q_i' \delta\right) - \exp\left(\beta' x_i + \rho \cdot Q_i' \delta\right) - \ln y_i ,$$

and gradient $h_2 = \left[X, Q\delta\right]' \upsilon_2$ based on the residual vector $\upsilon_{2i} = y_i - \exp\left(\beta' x_i - \rho \cdot \delta' Q_i\right)$. This procedure is essentially a Poisson regression with an endogenous covariate. Given consistent estimation of $\delta$, the usual necessary and sufficient conditions may be used in a variety of algorithms to maximize the log likelihood function of (31). In this application, we used the Newton-Raphson algorithm.

Because the predicted values of the transformed variables are used in the regression, standard errors of the second-stage parameters that maximize the likelihood function must be adjusted. It can be shown that the second-stage estimators are consistent if a rather undemanding set of regularity conditions are met for models (28) and (31) (Cameron and Trivedi, 2005). The distribution of the parameters in (31) is consistent and asymptotically normal with covariance:

$$\text{(32)} \qquad V_2^* = V_2 + V_2 \left[C V_1 C' - R V_1 C' - C V_1 R'\right] V_2,$$

where $V_2$ is the asymptotic covariance matrix of the log-likelihood equation $L_2$ (Murphy and Topel, 1985). The $C$ and $R$ matrices in (32) adjust the second-stage covariance matrix $V_2$ by including the covariance between the first-stage gradients and the second-stage likelihood function. Greene (2003, p. 510) provides their general form. Here, the derivations relevant to the LIML SAR-Poisson estimator are provided:

$$(33)\qquad R = \mathrm{E}\left( \frac{\partial \ln L_2}{\partial(\beta,\rho,c)} \frac{\partial \ln L_1}{\partial \delta'} \right) = \sum_{i=1}^{N} h_{2i} h_{1i}', \text{ and}$$

$$(34)\qquad C = \mathrm{E}\left( \frac{\partial \ln L_2}{\partial(\beta,\rho,c)} \frac{\partial \ln L_2}{\partial \delta'} \right) = \sum_{i=1}^{N} \hat{\rho}\hat{\upsilon}_{2i} h_{2i} Q_i'.$$

A negative binomial count model can be specified in the second stage with some modifications.

*3.3 The problem of zero counts*

In practice, the lagged, expected value of the count variable must be replaced by a transformation of the sample information. We outline three plausible transformations: (a) addition of an ad hoc constant *c*, (b) estimation of the constant using lagged dummy variables, and (c) an inverse hyperbolic sine transformation. The appeal of methods (b) and (c) is that they are data driven solutions, whilst the appeal of solution (a) is that it does not require more rigorous estimation procedures.

The first approach transforms neighboring counts using a pre-determined constant, leading to $y_j^* = \max\{c, y_j\}$, $0 < c \leq 1$. While this procedure succeeds in eliminating the problematic zeros, it may introduce some degree of bias into the estimation procedure. In the Monte Carlo experiments that follow, *c* is set to 0.5.

The second option is to estimate the constant *c* simultaneously with the structural parameters, as in the case of the time series analogue (see Cameron and Trivedi, 1998). This approach leads to the following SAR-Poisson model:

$$(35)\qquad \mu_i = \exp\left( \rho W_{i.} \cdot g(y_j^*) + c^* W_{i.} d + \beta' x_i \right) \text{, with}$$

$$(36)\qquad y_j^* = \begin{cases} y_j & y_j > 0 \\ 1 & y_j = 0 \end{cases}, \text{ and } d_j = \begin{cases} 1 & \text{if } y_j = 0 \\ 0 & \text{otherwise} \end{cases},$$

where $W_{i..}$ refers to the *i*-th row of the weights matrix. Clearly, the first two elements in the linear predictors $W \cdot g(y_j^*)$ and $Wd$ are endogenous. While this data-driven approach may be preferable to the *ad hoc* approach of adding a constant in terms of a priori judgment, there are additional computational difficulties that may arise. First, because of the bounded nature of $Wd$, an appropriate estimator that generates predicted values of $Wd$ in the [0, 1] interval must be selected (e.g., probit or logit regressions). Second, inclusion of the $Wd$ variable along with $W \cdot g(y^*)$ may introduce collinearity, which could complicate inference and possibly introduce bias. In fact, in a limited number of Monte Carlo trials (not reported here), this appeared to be the case.

A third option, which we evaluate in the Monte Carlo trials that follow, is to transform the neighboring counts using the inverse hyperbolic sine (IHS) transformation. Burbidge et al. (1988) give the concentrated log likelihood expression for this procedure. Pence

(2006) and El-Osta et al. (2007) provide recent applications. The advantage of this data-driven procedure is that it is more parsimonious as compared to the second approach above in terms of parameterization. In this approach, the IHS transformation function is parameterized as:

$$(37) \qquad g(y_j,\theta)=\sinh^{-1}(\theta y_j)=\left(\theta y_j+\sqrt{1+\theta^2 y_j^2}\right)\Big/\theta,$$

where $\theta \geq 0$ is a scaling parameter. As $\theta \rightarrow 0$, $\sinh^{-1}(\theta y_j)$ approaches $y_j$, and as $\theta \rightarrow \infty$, $y_j$ approaches 0. Pre- and post-multiplying the symmetric and idempotent "hat" matrix, $M = I - Q(Q'Q)^{-1}Q'$ (Hoaglin and Welsch, 1978) by the spatial lag operator, $W'MW$, helps concentrate the IHS log-likelihood function as:

$$(38) \qquad L_{IHS}=(\text{constant})-\frac{N}{2}\ln[g(y,\theta)'W'MWg(y,\theta)]-\frac{1}{2}\sum_{i=1}^{N}\ln[1+\theta^2 y_i^2].$$

Maximization of (38) can be accomplished by performing, for example, an appropriate search over the transformation parameter $\theta$. Given $\theta$, the IHS-transformed counts are used to generate the first-stage predicted values as in (29).

**4. Experimental design**

The small sample properties of the proposed SAR-Poisson estimator are investigated in this section. The estimator is based on limited information maximum likelihood with an adjusted Murphy-Topel covariance estimator, and is compared to the usual aspatial maximum likelihood based Poisson estimator in a Monte Carlo experiment. For the Poisson maximum likelihood estimator, the heteroskedasticity-robust quasi-ML covariance estimator is used, which attends to problems arising from overdispersion (Wooldridge, 2002, p. 653).

The experimental design closely follows the format frequently applied in the analysis of spatial regression models and model selection procedures (e.g., in Florax et al., 2003; Kelejian and Prucha, 2007; Fingleton and LeGallo, 2008). Random dependent variates were generated as $\tilde{y}_i \sim \text{Poisson}(\mu_i^{SAR})$, with $\mu_i^{SAR}=\exp(A_{i.}^{-1}X\beta_o)$ and $A = I - \rho_o W$, where the "true" or population values of the parameters were all set to 0.1 and $\rho_o$ takes on values of 0.0, 0.2, 0.4, 0.6, and 0.8. The sample sizes in the experiments included $N$ = 100, 256, 484, 729, and 1024. The design matrix $X$ included a constant ($x_0$), and two continuous random variables. The first ($x_1$) was generated randomly from the [0,2] uniform distribution. The second ($x_2$) was randomly generated from the normal distribution, $N(1, 2)$. The design matrix was held constant across 1,000 replications as well as across the different experiments.

Three spatial weighting matrices were used in the simulations. First, a length $N$ vector of random coordinates was generated from the uniform distribution. The coordinates were then used to construct a contiguity matrix ($W_1$) using a Delaunay algorithm with the *xy2cont.m* function in the MATLAB® Spatial Econometric Toolbox (http://www.spatial-econometrics.com). The average number of neighbors ranged between 5.68 (for $N$ = 100) and 5.96 (for $N$ = 1,024). The second weight matrix ($W_2$) was an inverse distance matrix, utilizing the Euclidian distance between contiguous neighbors identified in the first matrix. The third matrix ($W_3$) was a queen contiguity matrix, based on an $N$ by $N$ evenly spaced grid corresponding with the sample size under evaluation. For this contiguity matrix the average number of neighbors ranged between 7.27 (for $N$ = 100) and 7.63 (for $N$ = 1,024). All matrices were row-standardized such that the sum of the row elements across columns was one.

In the experiments the LIML SAR-Poisson estimator is implemented using the pre-defined transform $c$ = 0.5 and the IHS transformation. Four outcome measures for the experiments are used to investigate the small sample behavior of the estimator. For the spatial autoregressive parameter, a Likelihood Ratio (LR) test of the null hypothesis $\rho = 0$, and a Wald statistic with the null hypothesis $\rho = \rho_o$ are used. Both tests are $\chi^2$-variates with one degree of freedom. The advantage of the Wald statistic is that it uses covariance information, which helps discern the effectiveness of the Murphy-Topel covariance estimator relative to the aspatial ML estimator. For each combination of $\rho_o$, weights matrix and sample size, the frequency that the null hypothesis is rejected is reported. The nominal Type I error rate is set at 5%.

In terms of power the frequency that the two-step SAR estimator is preferred over the ML estimator is expected to increase as indicated by more frequent rejections of the null hypothesis with increasing sample size and stronger spatial autocorrelation. When $\rho_o = 0.0$, it is expected that the number of times the null hypothesis is rejected should fall around the nominal 5% level.

The root mean squared error (RMSE) of the aspatial ML and the two-step SAR-Poisson model were estimated at each replication for the coefficients associated with the $x_2$ and $x_3$ variables. For the $m$th Monte Carlo replicate and $k$th covariate the RMSE was estimated as $RMSE_m = \sqrt{Bias_{km}^2 + Var(\hat{\beta}_{km})}$, with the bias estimated as the difference between the true parameter value and the estimated parameters. The average of the RMSE's compares the relative precision of the aspatial ML and the two-step estimator under each scenario, while the average of the bias provides a measure of accuracy.

**5. Monte Carlo results**

Table 1 reports the results of the Likelihood Ratio and Wald tests, for power and size respectively, with respect to the spatial autoregressive parameter under the null hypothesis of no spatial autoregressive lag correlation. In general, the rejection rate pattern of the LR test corresponds with what would be expected: rejection frequencies monotonically increase as sample size and autoregressive spatial lag levels increase. The power results for the LR tests estimated with both transformation procedures were similar in terms of rejection frequencies. The size of the test is close to the 5% nominal level in most cases, although the rejection frequency appears slightly higher for the contiguity neighborhood definitions. The power of the LR test against the null hypothesis is rather weak. Relatively large sample sizes (≥ 729) and strong autocorrelation (≥ 0.6) are required for higher rejection frequencies. The inverse distance matrix specification ($W_2$) appears to slightly outperform the others in terms of detecting lag autocorrelation.

Turning to the size tests, the estimators employing the ad hoc constant of $c = 0.50$ appears to perform well, as the null hypothesis of $\rho = \rho_o$ was not rejected at least 95% of the time. However, at the highest AR levels, this pattern is slightly compromised. The SAR-Poisson estimator employing the IHS transformation performed poorly with respect to size at higher AR levels, but results were comparable with the ad hoc estimator at lower AR levels.[6]

<< Tables 1, 2 and 3 about here >>

Tables 2 and 3 summarize the Monte Carlo results for the RMSE and bias of the second and third covariates. In general, the aspatial ML estimator performed quite well relative to its spatial analogues. Neighborhood definition and the underlying distribution of the covariate appear to play some role with respect to RMSE performance of the SAR-Poisson estimator employing the ad hoc transformation, yet notable differences were only observed under the largest sample sizes (Table 2). For the normally distributed covariate, only for the largest sample sizes and highest autocorrelation levels was the RMSE of the spatial estimator lower than the aspatial ML results. Neighborhood and distributional definitions do not seem to be a factor in terms of RMSE for the SAR-Poisson model employing the IHS transformation. Only at the highest AR levels and largest sample sizes was the RMSE of this estimator lower than the aspatial ML estimates for the covariate parameters. On the other hand, the bias of the spatial estimators was generally lower than the aspatial ML estimator as sample size and autocorrelation levels were increased (Table 3).

Overall, the properties of the SAR-Poisson estimator are difficult to generalize. On the one hand, it remains a difficult task to detect AR-lag processes, given the test statistics used here. On the other hand, size tests suggest that the estimator performs relatively well in terms of correctly estimating the simulated AR levels. The RMSE results evaluating the precision of the AR estimator are encouraging as well, performing as one might expect: as sample size and autocorrelation levels increase, the overall RMSE decreased. In general, the Monte Carlo results are not too different form Klier and McMillen's (2008) investigation of the linearized spatial lag logit model.[7] They found that at higher AR levels, the linearized logit estimator associated with covariates had problems with respect to precision and accuracy. They concluded that the linearized logit estimator might be helpful in testing for spatial effects without biasing results towards rejecting the null of $\rho = 0$. A similar conclusion may be drawn here, in terms of confidence regarding AR lag parameters in empirical applications.

[6] The rejection frequencies of a Wald statistic testing the null hypothesis $\rho = 0$ performed similarly, but these results are not shown for reasons of space. They are available from the authors upon request.

[7] Klier and McMillen's (2008) linearization is attractive in that it applies a second order approximation of linear dependent variable models. By linearizing the selection model, computational issues that may arise from the repeated inversion of large $n$ by $n$ matrices are avoided.

<< Table 4 about here >>

**6. Empirical Application**

Given the satisfactory performance of the two-step SAR-Poisson estimator, we go on to demonstrate the usefulness of the estimator with an empirical example. The example is concerned with firm birth at the county level ($N$ = 3,078) in the lower 48 United States during the period 2000–2004. Clearly, firm birth, defined as cumulative new single-unit start-ups during the abovementioned time period (U.S. Census Bureau, Company Statistics Division, 2008), are discrete nonnegative counts. Guimarães et al. (2003) show how the conditional logit model reduces to a Poisson specification when location decisions are modeled as discrete-choice events, and only a single sector is considered. In this example, we focus on the manufacturing sector (NAICS 31–33). Manufacturing firms select sites based on company objectives potentially taking into account location factors related to agglomeration economies, market structure, the labor market, the availability of infrastructure, and fiscal policy characteristics. Below, we briefly describe these factors and the operationalizations we used.

Agglomeration economies may have a positive impact on firm birth, for instance because of knowledge spillovers between firms in similar market conditions when groups of firms in the same sector locate near each other (Glaeser and Kohlhase, 2004). Urbanization externalities may be associated with firms of different sectors co-locating. These types of effects are included in our model by means of the manufacturing share of employment (*msemp*), total establishment density (*tfdens*), and the percentage of manufacturing establishments with less than 10 (*pelt10*) and more than 100 employees (*pemt100*).[8]

The location literature finds that market structure is often the most important factor in location decisions of firms (Blair and Premus, 1987; Crone, 1997). Proximity to demand markets reduces transportation costs of inputs used in production as well as final output. Demand markets may also harbor a relatively larger stock of creative individuals capable of solving difficult supply issues or combining old ideas in novel ways, which may stimulate firm birth (Wojan and McGranahan, 2007). We use median household income (*mhhi*), population (*pop*), and the share of workers in creative occupations (*cclass*) to proxy market structure.[9]

Labor availability and labor cost are also important factors in location decisions (Schmenner et al., 1987; Henderson and McNamara, 2000). Finding skilled workers is more likely in locations with higher levels of educational attainment (Woodward, 1992; Coughlin and Segev, 2000). Low wages, labor availability, and an abundance of skilled

[8] Sector-specific employment data are from the U.S. Department of Transportation commuting patterns compiled by Research and Innovative Technology Administration (RITA), Bureau of Transportation Statistics. County Business Patterns data were used to calculate *tfdens*, *pelt10*, and *pemt100*.

[9] *Pop* and *mhhi* are scaled to be in thousands. The creative class share of employment was constructed by McGranahan and Wojan (2007) and is available at http://www.ers.usda.gov/Data/CreativeClassCodes/. The data for *mhhi* and *pop* are from County Business Patterns.

workers are expected to attract manufacturing investment. Average wage per job (*awage*), net flows of wages per commuter (*netflow*), unemployment rates (*uer*), and the percentage of adults with an associate's degree (*pedas*) measure cost, relative demand/supply of labor, labor availability, and skill of labor, respectively.[10]

Infrastructure contributes to regional economic development by improving connections to other regions, which is important for some manufacturing firms (Smith et al., 1978; Bartik, 1989). Maintaining infrastructure ensures that site accessibility is available in the future. Available land is also an important location factor (Carlson, 2000). Public road density (*proad*), interstate highway miles (*interst*), and government expenditures on highways per capita (*hwypc*) measure access, extent, and infrastructure quality.[11] The percentage of farmland to total county area measures available land (*avland*).[12]

Fiscal policy may impact the cost of conducting business in a region. Local governments walk a fine line between generating sufficient revenue to provide public goods and services, and supporting a favorable business climate (Gabe and Bell, 2004). Taxes may deter manufacturing investment (Wheat, 1986; Bartik, 1989), but local spending may constitute a benefit (Goetz, 1997). Obtaining detailed tax information at the county level is difficult. We use a composite measure of tax burden called the state tax business climate index (*bci*) (Hodge et al., 2003).[13] Higher index values indicate more favorable business climates. Government expenditures on education per capita (*educpc*) measure the level of public good services provided by local governments.[14]

Dummy variables identifying counties as belonging to metropolitan (*metro*) or micropolitan (*micro*) areas, as defined by the U.S. Office of Management and Budget (U.S. Census Bureau, 2008), are included as well. Non-core counties are the reference group.

<< Tables 5 and 6 around here >>

Table 5 provides the estimation results using the aspatial Poisson model. Table 6 shows the results for the SAR-Poisson estimator using a row-standardized inverse distance

---

[10] *Awage*, *netflow*, *uer*, and *pedas* are taken from the Bureau of Economic Analysis Regional Economic Accounts, Bureau of Labor Statistics, and U.S. Census Bureau, 2000 Decennial Census, respectively. *Awage* and *netflow* are scaled to be in thousands of dollars.

[11] Data on public road density and interstate miles are from the U.S. Department of Transportation (2000). Highway expenditures, scaled to hundreds, include federal, state, and local sources, which are measured using the U.S. Census Bureau, 1997 Census of Governments.

[12] This measure was calculated using a GIS database provided by ESRI in ArcGIS 9.2.

[13] The tax business climate index is only available at the state level starting from 2002. While a measure in 2000 would be preferable, the measures reported in subsequent years show that the index remains stable across time.

[14] Education expenditures include federal, state, and local sources, and they are measured using the 1997 Census of Governments. *Educpc* are scaled in hundreds of dollars.

matrix, utilizing the Euclidian distance between the eight nearest neighbors.[15] We include results using an exogenously defined transformation constant $c$, and a version in which the IHS transformation parameter $\theta$ is estimated. The results are very similar, and in our discussion, we predominantly attend to the former version. Each of the models explained about 80% of the variation in establishment birth. The autoregressive coefficient is statistically significant, although rather small, suggesting that start-ups in neighboring counties are important. Likelihood Ratio tests of $\rho = 0$ ($c$ = fixed at 0.5, LR = 2246, df = 1) and $\rho = \theta = 0$ (LR = 2280, df = 2) were rejected at the 1% level, suggesting the SAR-Poisson specifications are preferred over the aspatial model.

Counties endowed with agglomeration economies, larger populations, and more persons employed in creative occupations were more likely to attract firm investment. Labor availability and skilled labor were also positive and statistically significant. Infrastructure factors were positive and statistically significant with the exception of available land, which had a negative effect on location activity across all models. The business tax climate index and government education expenditures per capita were positive and statistically significant location determinants.

The elasticities in Tables 5 and 6 are of similar magnitude, but they are slightly larger for the SAR-Poisson estimator due to spatial spillovers mediated through the spatial multiplier. The largest positive elasticities with respect to manufacturing establishment births are from local agglomeration, creative class employment, and educational attainment. Focusing on the pre-fixed constant transformation specification of the SAR-Poisson model, a one-percent increase in each of these measures increases the percentage of county level births by 0.484, 1.129, and 0.778%, respectively. A one-percent increase in the number of manufacturing establishments with more than 100 employees reduces the number of establishment births by 1.711%. This finding suggests that counties with higher shares of firms exhibiting internal economies of scale represent a significant barrier to new manufacturing establishments. In the dynamic process of firm formation, counties dominated by firms exhibiting (internal) increasing returns to scale are at a relative disadvantage even if localization economies are present.

## 7. Conclusions

Recent increases in data availability of business establishment site selection decisions have allowed researchers to model location determinants of manufacturing activity at lower levels of spatial aggregation. However, few studies have attempted to explicitly model the spatial lag processes coincident with establishment site selection, perhaps because research incorporating endogenous spatial lag processes into count regressions has been limited. To fill this gap, a structurally consistent spatial lag count regression model that accommodates global spatial spillovers was proposed.

---

[15] Results were robust against alternative definitions of the spatial weight matrix including first-order queen contiguity, inverse distance, and differences in $k$ for $k$-nearest neighbors. The "hybrid" matrix selected is a compromise between a between a spatial weight matrix based on contiguity and a full distance matrix.

The approach suggested here applied a two-step limited information maximum likelihood procedure to estimate count models with a spatially lagged dependent variable. Standard errors of the second-stage regression were adjusted using the Murphy-Topel covariance estimator. Small sample properties of the SAR-Poisson model were investigated with a Monte Carlo experiment under the assumption of different neighborhood structures. In terms of minimizing bias that may arise from AR lag processes, the proposed estimator appeared to perform reasonably well. However, in terms of RMSE, there appeared to be little gain over a conventional aspatial Poisson estimator. The power of the specification test used to detected AR levels was rather weak, but size tests suggest that the estimator performs well in terms of pinpointing the location of the autoregressive parameter. These results may not be too surprising given the two-stage estimator applied here, where gains in consistency comes at the cost of efficiency. With respect to the attending to zero counts of the lagged dependent variable, the Type I errors of the SAR-Poisson model with the lagged outcome variable transformed with the inverse hyperbolic sine method were excessive. While the pre-fixed constant transformation is *ad hoc*, it outperformed the data driven method. The inverse hyperbolic sine transformation is not recommended. Given these results, we recommend using the Wald statistic in empirical settings to verify the significance of the autoregressive parameter. Future directions might consider general moment estimators, or linearized versions of the SAR-Poisson model, as Klier and McMillen (2008) suggested for logit and probit autoregressive process models.

An empirical example estimated the impact of local determinants of manufacturing location decisions in the lower 48 United States, 2000–2004, as a discrete Poisson process. The SAR-Poisson models were preferred according to goodness of fit measures and likelihood ratio tests, suggesting that the likelihood of attracting a manufacturing establishment was dependent on the location activities of establishments in neighboring counties. Under the SAR-Poisson specification, the elasticities of the location determinants may be decomposed into direct and indirect effects providing a richer geographic understanding of location determinants and how neighboring resources may influence firm birth in certain locations. Such a perspective could have important policy implications with respect to leveraging or moderating externalities arising from neighborhood scale economies.

## References


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

Anselin, L., 2002. Under the hood: issues in the specification and interpretation of spatial regression models. Agricultural Economics 17 (3), 247–267.

Anselin, L., 2006. Spatial econometrics. In: Mills, T.C., Patterson, K. (Eds.). Palgrave Handbook of Econometrics: Volume 1, Econometric Theory. Basingstoke, Palgrave Macmillan, 901–969.

Bartik, T.J., 1989. Small business start–ups in the United States: estimates of the effects of characteristics of states. Southern Economic Journal 55(4), 1004–1018.

Bhati, A., 2005. Robust spatial analysis of rare crimes: an information-theoretic approach. Sociological Methodology 35, 239–301.

Blair, J.P., Premus, R., 1987. Major factors in industrial location: a review. Economic Development Quarterly 1(1), 72–85.

Blundell, R. R., Griffith, R., Windmeijer, F., 1995. Individual effects and dynamics in count data. Discussion Paper 95–03, Dept. of Economics, University College, London.

Burbidge, J.B., L. Magee, and A. L. Robb. 1988. Alternative Transformations to Handle Extreme Values of the Dependent Variable. Journal of the American Statistical Association 83(401): 123–127.

Bureau of Economic Analysis. Regional economic accounts. Available at: http://www.bea.gov/regional/index.htm. Accessed on 10/14/2008.

Bureau of Labor Statistics. Available at: http://www.bls.gov/lau/#tables. Accessed on 02/08/2009.

Cameron, A.C., Trivedi, P. K., 1998. Regression Analysis of Count Data. Econometric Society Monograph No.30, Cambridge, Cambridge University Press.

Cameron, A.C., Trivedi, P. K., 2005. Microeconometrics: Methods and Applications. Cambridge, Cambridge University Press.

Carlson, V., 2000. Studying firm locations: survey responses vs. econometric models. Journal of Regional Analysis & Policy 30(1), 1–22.

Carod, J.M.A., 2005. Determinants of industrial location: an application for Catalan municipalities. Papers in Regional Science 82(1), 105 – 120.

Carod, J. M. A., Antolín, M. C. M., 2004. Firm size and geographical aggregation: an empirical appraisal in industrial location. Small Business Economics 22, 299 – 312.

Chong, S., 2006. The role of labor cost in the location choices of Japanese investors in china. Papers in Regional Science 85(1), 121–138.

Cliff, A.D., Ord, J.K, 1981. Spatial Processes. London: Pion.

Coughlin, C.C., Segev E., 2000. Location determinants of new foreign–owned manufacturing plants. Journal of Regional Science 40, 323–351.

County Business Patterns. Available at: http://www.census.gov/econ/cbp/index.html. Accessed on 3/18/2009.

Crone, T.M., 1997. Where have all the factory jobs gone - and why? Business Review, Federal Reserve Bank of Philadelphia, May/June.

Davis, D.E., Schluter, G.E., 2005. Labor-force heterogeneity as a source of agglomeration economics an in empirical analysis of county-level determinants of food plant entry. Journal of Agricultural and Resource Economics 30(3), 480–501.

El-Osta, H., A. Mishra, and M. Morehart. 2007. Determinants of Economic Well-being among U.S. Farm Operator Households. Agricultural Economics 36: 291–304.

ESRI, GIS Software. 2008. Available at: http://www.esri.com/.

Fingleton, B, LeGallo, J., 2008. Estimating spatial models with endogenous variables, a spatial lag and spatially dependent disturbances: finite sample properties. Papers in Regional Science 87(3), 319–339.

Fleming, M.M., 2004. Techniques for estimating spatially dependent discrete-choice models. In: Anselin, L., Florax, R.J.G.M., Rey, S.J. (Eds.). Advances in Spatial Econometrics: Methodology, Tools and Applications. Berlin: Springer–Verlag.

Florax, R.J.G.M., Folmer, H., Rey, S. J., 2003. Specification searches in spatial econometrics: the relevance of Hendry's methodology. Regional Science and Urban Economics 33, 557–579.

Flores, O., Rossi, V., and Mortier, F., 2009. Autocorrelation Offsets Zero-Inflation in Models of Tropical Saplings Density. Ecological Modeling. *forthcoming*.

Fotheringham, A.S., Brunsdon, C., and Charlton, M. 2002. Geographically Weighted Regression: The Analysis of Spatially Varying Relationships. Wiley, West Sussex.

Fotopoulos, G., Louri, H., 2000. Location and survival of new entry. Small Business Economics 14, 311– 321.

Gabe, T., Bell , K.P., 2004. Tradeoffs between local taxes and government spending as determinants of business location. Journal of Regional Science 44(1), 21–41.

Glaeser, E.L., Kohlhase, J.E., 2004. Cities, Regions, and the Decline of Transport Costs. Papers in Regional Science 83, 197–228.

Goetz, S., 1997. State- and county-level determinants of food manufacturing establishment growth: 1987–1993. American Journal of Agricultural Economics 79(3), 838–850.

Greene, W. H., 2003. Econometric Analysis. Prentice Hall, Upper Saddle River, NJ.

Griffith, D.A., 2002. A Spatial Filtering Specification for the Auto-Poisson Model. Statistics and Probability 58, 245–251.

Griffith, D.A., 2003. Spatial Autocorrelation and Spatial Filtering: Gaining Understanding Through Theory and Scientific Visualization. Springer-Verlag, Berlin.

Guimarães, P., Figueiredo, O., Woodward, D., 2003. A tractable approach to the firm location decision problem. The Review of Economics and Statistics 85(1), 201–204.

Guimarães, P., Figueiredo, O., Woodward, D., 2004. Industrial location modeling: extending the random utility framework. Journal of Regional Science 44(1), 1–20.

Haining, R.P., Law, J., and Griffith, D.A., 2009. Modeling Small Area Counts in the Presence of Overdispersion and Spatial Autocorrelation. Computational Statistics and Data Analysis 53, 2923–2937.

Henderson, J.R., McNamara , K.T., 2000. The location of food manufacturing plant investments in Corn Belt counties. Journal of Agricultural and Resource Economics 25(2), 680–697.

Hays, J.C., and R.J. Franzese. 2009. A Comparison of the Small-Sample Properties of Several Estimators for Spatial-Lag Count Models. Unpublished Working Paper, University of Illinois Champaign-Urbana.

Hoaglin, D.C., and R.E. Welsch. 1978. The Hat Matrix in Regression and ANOVA. The American Statistician 32 (1): 17–22.

Hodge, S.A., Warcholik, W.P., Moody, J.S., 2003. State business tax climate index, first edition. Background Paper No. 41. The Tax Foundation, Washington D.C.

Kaiser, M.S. and Cressie, N., 1997. Modeling Poisson Variables with Positive Spatial Dependence. Statistics and Probability 35, 423–432.

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

Kelejian H.H., Prucha, I.R., 2007. HAC estimation in a spatial framework. Journal of Econometrics 140, 131–154.

Klier, T. and McMillen, D.P., 2008. Clustering of auto supplier plants in the United States: generalized method of moments spatial logit for large samples. Journal of Business and Economic Statistics 26, 460–471.

Lambert, D.M., Garret, M.I., McNamara, K.T., 2006a. An application of spatial Poisson models to manufacturing investment location analysis. Journal of Agricultural and Applied Economics 38(1), 105–121.

Lambert, D.M., Garret, M.I., McNamara, K.T., 2006b. Food industry investment flows: implications for rural development. Review of Regional Studies 36(2), 140–162.

Lambert, D.M., McNamara, K.T., 2009. Location determinants of food manufacturers in the U.S., 2000 - 2004: are nonmetropolitan counties competitive? Agricultural Economics, *forthcoming*.

LeSage, J.P., 2000. Bayesian estimation of limited dependent variable spatial autoregressive models. Geographical Analysis 32, 19–35.

LeSage, J.P., Fischer, M.M., Scherngell, T., 2007. Knowledge spillovers across Europe: evidence from a Poisson spatial interaction model with spatial effects. Papers in Regional Science 86(3), 393–421.

LeSage, J.P., Pace, R.K., 2009. Introduction to Spatial Econometrics. CRC Press, Taylor and Francis Group.

Pence, K.M., 2006. The role of wealth transformations: an application to estimating the effect of tax incentives on saving. Contributions to Economic Analysis & Policy 5(1), Article 20, 1-24.

Pinske, J., Slade, M. E., 1998. Contracting in space: an application of spatial statistics to discrete-choice model. Journal of Econometrics 85, 125–154.

MATLAB® Spatial Econometric Toolbox. Available at http://www.spatial-econometrics.com.

McFadden, D., 1974. Conditional Logit Analysis of Qualitative Choice Behavior. In: Zarembka, P. (Ed.). Frontiers in Econometrics. New York: Academic Press.
Murphy, K., Topel, R., 1985. Estimation and inference in two step econometric models. Journal of Business and Economic Statistics 3, 370–379.
Research and Innovative Technology Administration (RITA), Bureau of Transportation Statistics. http://www.transtats.bts.gov. Accessed on 5/18/2008.
Rasmussen, S., 2004. Modeling of discrete spatial variation in epidemiology with SAS using GLIMMIX. Computer Methods and Programs in Biomedicine 76(1), 83–89.
Schabenberger, O., Pierce, F.J., 2002. Contemporary Statistical Models for the Plant and Soil Sciences. New York: CRC Press.
Schmenner, R.W., Huber, J.C., Cook, R.L., 1987. Geographic differences and the location of new manufacturing facilities. Journal of Urban Economics 21, 83–104.
Smith, E.D., Deaton, B.J., Kelch, D.R., 1978. Location determinants of manufacturing industry in rural areas. Southern Journal of Agricultural Economics 10, 23–32.
U.S. Census Bureau, 2008. Metropolitan and Micropolitan Statistical Areas. Available at: http://www.census.gov/population/www/metroareas/metrodef.html.
U.S. Census Bureau, Company Statistics Division, 2008. Statistics of U.S. Businesses: Dynamic Data. Available at: http://www.census.gov/csd/susb/susbdyn.htm. Accessed on 5/23/2008.
U.S. Census Bureau, 1997 Census of Governments. Available at: http://www.census.gov/govs/www/cog97.html. Accessed on 10/21/2008.
U.S. Census Bureau, U.S. Census 2000. Available at: http://www.census.gov/main/www/cen2000.html. Accessed on 11/1/2008.
U.S. Department of Transportation, Federal Highway Administration. Highway Statistics 2000. Available at: http://www.fhwa.dot.gov/ohim/hs00/re.htm. Accessed on 3/11/2008.
Weber, A. 1929. Theory of the Location of Industries. Chicago: The University of Chicago Press, 1929. [translated by Carl J. Friedrich].
Wheat, L., 1986. The Determinants of 1963–77 regional manufacturing growth: why the South and West grow. Journal of Regional Science 26(4), 635–660.
Whittle, P., 1954. On the stationary process in the plane. Biometrika 41, 434–439.
Wojan, T.R., McGranahan, D.A., 2007. Ambient returns: creative capital's contribution to local manufacturing competitiveness. Agricultural and Resource Economics Review 36(1), 133–148.
Woodward, D.P., 1992. Location determinants of Japanese manufacturing start–ups in the United States. Southern Economics Journal 58, 690–708.
Wooldridge, J., 2002. Econometric Analysis of Cross Section and Panel Data. Cambridge, Cambridge University Press.
Zeger, S. L., Qaqish, B., 1988. Markov regression models for time series: a quasi-likelihood approach. Biometrics 44, 1019–1031.

Table 1. Rejection probability for the null hypotheses H$_0$: $\rho = 0$ and H$_0$: $\rho = \rho_o$

**Method 1: *ad hoc* constant, $c$ = 0.50**

Likelihood Ratio test, H$_0$: $\rho = 0$

| | $W_1$ | | | | | $W_2$ | | | | | $W_3$ | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $\rho_o$ | 100 | 256 | 484 | 729 | 1024 | 100 | 256 | 484 | 729 | 1024 | 100 | 256 | 484 | 729 | 1024 |
| **0.0** | 0.06 | 0.08 | 0.07 | 0.05 | 0.06 | 0.06 | 0.07 | 0.07 | 0.05 | 0.05 | 0.07 | 0.08 | 0.08 | 0.07 | 0.06 |
| **0.2** | 0.08 | 0.11 | 0.11 | 0.13 | 0.14 | 0.06 | 0.10 | 0.10 | 0.10 | 0.14 | 0.08 | 0.09 | 0.11 | 0.11 | 0.12 |
| **0.4** | 0.13 | 0.20 | 0.24 | 0.37 | 0.49 | 0.10 | 0.19 | 0.19 | 0.40 | 0.57 | 0.09 | 0.17 | 0.22 | 0.30 | 0.32 |
| **0.6** | 0.30 | 0.50 | 0.66 | 0.90 | 0.97 | 0.33 | 0.56 | 0.56 | 0.94 | 0.99 | 0.15 | 0.51 | 0.56 | 0.83 | 0.88 |
| **0.8** | 0.96 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 0.50 | 1.00 | 1.00 | 1.00 | 1.00 |

Wald test, H$_0$: $\rho = \rho_o$

| | $W_1$ | | | | | $W_2$ | | | | | $W_3$ | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $\rho_o$ | 100 | 256 | 484 | 729 | 1024 | 100 | 256 | 484 | 729 | 1024 | 100 | 256 | 484 | 729 | 1024 |
| **0.0** | 0.07 | 0.06 | 0.06 | 0.05 | 0.05 | 0.05 | 0.06 | 0.04 | 0.04 | 0.04 | 0.07 | 0.07 | 0.06 | 0.06 | 0.05 |
| **0.2** | 0.04 | 0.04 | 0.04 | 0.04 | 0.05 | 0.03 | 0.04 | 0.03 | 0.03 | 0.04 | 0.05 | 0.04 | 0.04 | 0.04 | 0.04 |
| **0.4** | 0.03 | 0.03 | 0.02 | 0.02 | 0.03 | 0.03 | 0.03 | 0.02 | 0.02 | 0.02 | 0.03 | 0.02 | 0.02 | 0.02 | 0.03 |
| **0.6** | 0.02 | 0.01 | 0.02 | 0.00 | 0.02 | 0.03 | 0.02 | 0.02 | 0.02 | 0.01 | 0.03 | 0.01 | 0.01 | 0.01 | 0.01 |
| **0.8** | 0.05 | 0.02 | 0.04 | 0.07 | 0.08 | 0.08 | 0.03 | 0.06 | 0.12 | 0.13 | 0.04 | 0.04 | 0.03 | 0.06 | 0.09 |

**Method 2: first-stage regression, IHS transformation of neighboring counts**

Likelihood Ratio test, H$_0$: $\rho = 0$

| | $W_1$ | | | | | $W_2$ | | | | | $W_3$ | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $\rho_o$ | 100 | 256 | 484 | 729 | 1024 | 100 | 256 | 484 | 729 | 1024 | 100 | 256 | 484 | 729 | 1024 |
| **0.0** | 0.06 | 0.08 | 0.07 | 0.05 | 0.06 | 0.05 | 0.07 | 0.07 | 0.05 | 0.05 | 0.07 | 0.08 | 0.08 | 0.07 | 0.06 |
| **0.2** | 0.08 | 0.10 | 0.11 | 0.13 | 0.14 | 0.06 | 0.10 | 0.10 | 0.10 | 0.14 | 0.08 | 0.09 | 0.11 | 0.11 | 0.12 |
| **0.4** | 0.13 | 0.19 | 0.25 | 0.37 | 0.49 | 0.10 | 0.19 | 0.19 | 0.39 | 0.57 | 0.09 | 0.17 | 0.22 | 0.30 | 0.32 |
| **0.6** | 0.30 | 0.51 | 0.66 | 0.89 | 0.96 | 0.35 | 0.56 | 0.56 | 0.94 | 0.99 | 0.16 | 0.51 | 0.56 | 0.83 | 0.89 |
| **0.8** | 0.96 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 0.51 | 1.00 | 1.00 | 1.00 | 1.00 |

Wald test, H$_0$: $\rho = \rho_o$

| | $W_1$ | | | | | $W_2$ | | | | | $W_3$ | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $\rho_o$ | 100 | 256 | 484 | 729 | 1024 | 100 | 256 | 484 | 729 | 1024 | 100 | 256 | 484 | 729 | 1024 |
| **0.0** | 0.06 | 0.06 | 0.06 | 0.05 | 0.05 | 0.05 | 0.06 | 0.04 | 0.04 | 0.04 | 0.06 | 0.07 | 0.06 | 0.06 | 0.05 |
| **0.2** | 0.04 | 0.04 | 0.03 | 0.03 | 0.04 | 0.03 | 0.03 | 0.03 | 0.04 | 0.04 | 0.05 | 0.03 | 0.04 | 0.04 | 0.03 |
| **0.4** | 0.03 | 0.02 | 0.02 | 0.02 | 0.02 | 0.05 | 0.03 | 0.03 | 0.04 | 0.04 | 0.03 | 0.02 | 0.02 | 0.02 | 0.03 |
| **0.6** | 0.18 | 0.09 | 0.11 | 0.20 | 0.32 | 0.24 | 0.17 | 0.18 | 0.32 | 0.47 | 0.07 | 0.14 | 0.10 | 0.22 | 0.26 |
| **0.8** | 0.96 | 1.00 | 1.00 | 1.00 | 1.00 | 0.96 | 1.00 | 1.00 | 1.00 | 1.00 | 0.86 | 1.00 | 1.00 | 1.00 | 1.00 |

Notes: Shaded entries denote cases where the null hypothesis was rejected for the likelihood ratio test and accepted for the Wald test at the five percent level.

Table 2. RMSE: Aspatial Poisson ML and SAR-Poisson LIML $\beta_1$, $\beta_2$ estimators

| | Aspatial $\beta_1$ - ML | | | | | $\beta_1$ - SAR-Poisson LIML [1/] | | | | | $\beta_1$ - SAR-Poisson LIML [2/] | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $W_1$ | | | | | $W_1$ | | | | | $W_1$ | | | | |
| **ρ/*n*** | **100** | **256** | **484** | **729** | **1024** | **100** | **256** | **484** | **729** | **1024** | **100** | **256** | **484** | **729** | **1024** |
| **0.0** | 0.206 | 0.124 | 0.092 | 0.075 | 0.062 | 0.220 | 0.128 | 0.094 | 0.076 | 0.063 | 0.220 | 0.128 | 0.094 | 0.076 | 0.063 |
| **0.2** | 0.196 | 0.120 | 0.087 | 0.073 | 0.060 | 0.210 | 0.123 | 0.089 | 0.075 | 0.061 | 0.209 | 0.123 | 0.089 | 0.075 | 0.061 |
| **0.4** | 0.187 | 0.110 | 0.081 | 0.068 | 0.056 | 0.198 | 0.114 | 0.082 | 0.070 | 0.057 | 0.198 | 0.114 | 0.083 | 0.070 | 0.057 |
| **0.6** | 0.161 | 0.098 | 0.072 | 0.064 | 0.051 | 0.171 | 0.100 | 0.074 | 0.064 | 0.051 | 0.172 | 0.101 | 0.075 | 0.064 | 0.052 |
| **0.8** | 0.121 | 0.068 | 0.049 | 0.068 | 0.042 | 0.124 | 0.070 | 0.050 | 0.045 | 0.036 | 0.125 | 0.071 | 0.051 | 0.045 | 0.036 |
| | $W_2$ | | | | | $W_2$ | | | | | $W_2$ | | | | |
| **0.0** | 0.206 | 0.124 | 0.092 | 0.075 | 0.062 | 0.224 | 0.129 | 0.093 | 0.076 | 0.063 | 0.227 | 0.129 | 0.098 | 0.073 | 0.064 |
| **0.2** | 0.196 | 0.120 | 0.087 | 0.073 | 0.061 | 0.209 | 0.124 | 0.088 | 0.074 | 0.061 | 0.218 | 0.124 | 0.093 | 0.070 | 0.061 |
| **0.4** | 0.187 | 0.110 | 0.081 | 0.068 | 0.057 | 0.198 | 0.114 | 0.083 | 0.070 | 0.058 | 0.202 | 0.117 | 0.091 | 0.067 | 0.058 |
| **0.6** | 0.163 | 0.097 | 0.072 | 0.067 | 0.053 | 0.173 | 0.101 | 0.075 | 0.064 | 0.052 | 0.183 | 0.103 | 0.081 | 0.061 | 0.053 |
| **0.8** | 0.129 | 0.072 | 0.050 | 0.082 | 0.055 | 0.130 | 0.071 | 0.051 | 0.046 | 0.037 | 0.135 | 0.073 | 0.057 | 0.044 | 0.037 |
| | $W_3$ | | | | | $W_3$ | | | | | $W_3$ | | | | |
| **0.0** | 0.206 | 0.124 | 0.087 | 0.075 | 0.062 | 0.223 | 0.129 | 0.094 | 0.076 | 0.063 | 0.232 | 0.123 | 0.095 | 0.074 | 0.063 |
| **0.2** | 0.197 | 0.120 | 0.080 | 0.073 | 0.060 | 0.214 | 0.124 | 0.088 | 0.073 | 0.061 | 0.224 | 0.119 | 0.092 | 0.072 | 0.060 |
| **0.4** | 0.187 | 0.110 | 0.071 | 0.067 | 0.056 | 0.203 | 0.114 | 0.082 | 0.068 | 0.057 | 0.212 | 0.111 | 0.086 | 0.067 | 0.058 |
| **0.6** | 0.162 | 0.098 | 0.048 | 0.061 | 0.050 | 0.176 | 0.101 | 0.073 | 0.061 | 0.051 | 0.185 | 0.097 | 0.076 | 0.060 | 0.052 |
| **0.8** | 0.120 | 0.068 | 0.075 | 0.045 | 0.039 | 0.124 | 0.070 | 0.049 | 0.043 | 0.036 | 0.131 | 0.066 | 0.055 | 0.042 | 0.037 |

| | Aspatial $\beta_2$ - ML | | | | | $\beta_2$ - SAR-Poisson LIML [1/] | | | | | $\beta_2$ - SAR-Poisson LIML [2/] | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **100** | **256** | **484** | **729** | **1024** | **100** | **256** | **484** | **729** | **1024** | **100** | **256** | **484** | **729** | **1024** |
| **ρ/*n*** | $W_1$ | | | | | $W_1$ | | | | | $W_1$ | | | | |
| **0.0** | 0.052 | 0.036 | 0.026 | 0.021 | 0.018 | 0.058 | 0.038 | 0.027 | 0.022 | 0.019 | 0.058 | 0.038 | 0.027 | 0.022 | 0.019 |
| **0.2** | 0.050 | 0.034 | 0.026 | 0.021 | 0.018 | 0.056 | 0.036 | 0.027 | 0.021 | 0.018 | 0.056 | 0.036 | 0.027 | 0.021 | 0.018 |
| **0.4** | 0.047 | 0.032 | 0.024 | 0.020 | 0.017 | 0.052 | 0.035 | 0.025 | 0.020 | 0.018 | 0.052 | 0.035 | 0.025 | 0.020 | 0.018 |
| **0.6** | 0.043 | 0.029 | 0.022 | 0.020 | 0.018 | 0.048 | 0.030 | 0.022 | 0.018 | 0.016 | 0.048 | 0.031 | 0.022 | 0.019 | 0.016 |
| **0.8** | 0.046 | 0.026 | 0.025 | 0.028 | 0.027 | 0.036 | 0.022 | 0.016 | 0.014 | 0.011 | 0.037 | 0.023 | 0.016 | 0.014 | 0.011 |
| | $W_2$ | | | | | $W_2$ | | | | | $W_2$ | | | | |
| **0.0** | 0.052 | 0.036 | 0.026 | 0.021 | 0.018 | 0.058 | 0.038 | 0.027 | 0.022 | 0.019 | 0.070 | 0.037 | 0.026 | 0.022 | 0.019 |
| **0.2** | 0.050 | 0.034 | 0.026 | 0.021 | 0.018 | 0.055 | 0.036 | 0.026 | 0.021 | 0.018 | 0.068 | 0.035 | 0.026 | 0.021 | 0.018 |
| **0.4** | 0.047 | 0.032 | 0.024 | 0.020 | 0.017 | 0.051 | 0.035 | 0.025 | 0.021 | 0.018 | 0.064 | 0.034 | 0.026 | 0.020 | 0.018 |
| **0.6** | 0.043 | 0.030 | 0.023 | 0.021 | 0.019 | 0.048 | 0.031 | 0.023 | 0.019 | 0.016 | 0.058 | 0.030 | 0.023 | 0.019 | 0.017 |
| **0.8** | 0.052 | 0.034 | 0.032 | 0.034 | 0.036 | 0.037 | 0.023 | 0.017 | 0.014 | 0.012 | 0.044 | 0.023 | 0.017 | 0.013 | 0.013 |
| | $W_3$ | | | | | $W_3$ | | | | | $W_3$ | | | | |
| **0.0** | 0.052 | 0.036 | 0.026 | 0.021 | 0.018 | 0.058 | 0.039 | 0.027 | 0.022 | 0.019 | 0.061 | 0.040 | 0.028 | 0.022 | 0.019 |
| **0.2** | 0.050 | 0.034 | 0.026 | 0.021 | 0.018 | 0.055 | 0.037 | 0.026 | 0.021 | 0.018 | 0.059 | 0.038 | 0.027 | 0.021 | 0.018 |
| **0.4** | 0.046 | 0.032 | 0.024 | 0.020 | 0.017 | 0.052 | 0.035 | 0.024 | 0.020 | 0.017 | 0.056 | 0.036 | 0.026 | 0.019 | 0.017 |
| **0.6** | 0.042 | 0.031 | 0.021 | 0.019 | 0.015 | 0.046 | 0.031 | 0.022 | 0.018 | 0.015 | 0.049 | 0.032 | 0.023 | 0.018 | 0.015 |
| **0.8** | 0.031 | 0.034 | 0.019 | 0.023 | 0.017 | 0.034 | 0.022 | 0.015 | 0.013 | 0.011 | 0.036 | 0.023 | 0.017 | 0.013 | 0.011 |

Notes: Shaded entries denote cases where the RMSE of the SAR-Poisson estimators (center and right-most columns) were less than the aspatial ML Poisson estimator. Root mean squared errors (RMSE) are estimated as $\sqrt{Bias^2 + Var(\beta)}$ for each Monte Carlo replicate, with the bias estimated as $\beta_0 - \hat{\beta}$; the difference between the true parameter value and its estimate. Entries are averaged over 1,000 simulations. [1/] To account for zeros in the first stage estimation, neighboring counts were transformed using the inverse hyperbolic sine estimation procedure of Burbidge et al. (1988). [2/] The *ad hoc* constant $c$ = 0.5 was added to neighboring counts of zero to facilitate the first-stage regression.

Table 3. Bias: Aspatial Poisson ML and SAR-Poisson LIML $\beta_1$, $\beta_2$ estimators

| | Aspatial $\beta_1$ - ML | | | | | $\beta_1$ - SAR-Poisson LIML [1/] | | | | | $\beta_1$ - SAR-Poisson LIML [2/] | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | ----- $W_1$ ----- | | | | | ----- $W_1$ ----- | | | | | ----- $W_1$ ----- | | | | |
| **ρ/*n*** | **100** | **256** | **484** | **729** | **1024** | **100** | **256** | **484** | **729** | **1024** | **100** | **256** | **484** | **729** | **1024** |
| **0.0** | -0.007 | -0.002 | 0.000 | 0.002 | -0.002 | -0.011 | -0.004 | -0.002 | -0.001 | -0.004 | -0.011 | -0.004 | -0.002 | -0.001 | -0.004 |
| **0.2** | -0.007 | -0.003 | 0.000 | 0.003 | 0.000 | -0.014 | -0.005 | -0.003 | -0.001 | -0.003 | -0.014 | -0.005 | -0.003 | -0.001 | -0.003 |
| **0.4** | -0.003 | -0.001 | 0.004 | 0.008 | 0.003 | -0.013 | -0.004 | -0.002 | 0.000 | -0.003 | -0.013 | -0.004 | -0.002 | 0.000 | -0.003 |
| **0.6** | 0.008 | 0.001 | 0.010 | 0.015 | 0.010 | -0.008 | -0.005 | 0.000 | 0.000 | -0.002 | -0.009 | -0.005 | 0.000 | 0.000 | -0.003 |
| **0.8** | 0.030 | 0.015 | 0.024 | 0.034 | 0.025 | -0.004 | -0.001 | 0.000 | 0.002 | 0.000 | -0.005 | 0.000 | 0.001 | 0.002 | -0.002 |
| | ----- $W_2$ ----- | | | | | ----- $W_2$ ----- | | | | | ----- $W_2$ ----- | | | | |
| **0.0** | -0.007 | -0.002 | 0.000 | 0.002 | -0.002 | -0.012 | -0.006 | -0.003 | 0.000 | -0.004 | -0.011 | -0.006 | -0.003 | 0.000 | -0.004 |
| **0.2** | -0.007 | -0.001 | 0.001 | 0.003 | 0.000 | -0.013 | -0.007 | -0.003 | 0.000 | -0.003 | -0.012 | -0.007 | -0.003 | 0.000 | -0.003 |
| **0.4** | -0.004 | 0.003 | 0.006 | 0.007 | 0.006 | -0.013 | -0.006 | -0.001 | 0.001 | -0.003 | -0.013 | -0.006 | -0.001 | 0.001 | -0.003 |
| **0.6** | 0.011 | 0.011 | 0.015 | 0.016 | 0.017 | -0.005 | -0.005 | 0.001 | 0.002 | -0.002 | -0.005 | -0.005 | 0.001 | 0.002 | -0.002 |
| **0.8** | 0.045 | 0.037 | 0.035 | 0.036 | 0.041 | 0.000 | -0.003 | 0.002 | 0.002 | -0.001 | -0.001 | 0.000 | 0.002 | 0.000 | -0.004 |
| | ----- $W_3$ ----- | | | | | ----- $W_3$ ----- | | | | | ----- $W_3$ ----- | | | | |
| **0.0** | -0.007 | -0.002 | -0.002 | 0.002 | -0.002 | -0.001 | -0.004 | 0.000 | 0.000 | -0.004 | 0.000 | -0.004 | 0.000 | 0.000 | -0.004 |
| **0.2** | -0.010 | -0.001 | -0.002 | 0.002 | -0.002 | -0.004 | -0.002 | -0.001 | 0.000 | -0.003 | -0.003 | -0.002 | -0.001 | 0.000 | -0.003 |
| **0.4** | -0.015 | 0.000 | -0.003 | 0.004 | 0.000 | -0.008 | -0.002 | -0.001 | 0.000 | -0.003 | -0.007 | -0.002 | -0.001 | 0.000 | -0.003 |
| **0.6** | -0.019 | 0.003 | -0.004 | 0.010 | 0.004 | -0.005 | -0.001 | 0.000 | 0.001 | -0.003 | -0.006 | -0.001 | 0.000 | 0.001 | -0.003 |
| **0.8** | -0.029 | 0.005 | 0.002 | 0.027 | 0.016 | -0.006 | 0.001 | 0.000 | -0.001 | -0.002 | -0.007 | -0.001 | 0.000 | 0.001 | -0.002 |

| | Aspatial $\beta_2$ - ML | | | | | $\beta_2$ - SAR-Poisson LIML [1/] | | | | | $\beta_2$ - SAR-Poisson LIML [2/] | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **100** | **256** | **484** | **729** | **1024** | **100** | **256** | **484** | **729** | **1024** | **100** | **256** | **484** | **729** | **1024** |
| **ρ /*n*** | ----- $W_1$ ----- | | | | | ----- $W_1$ ----- | | | | | ----- $W_1$ ----- | | | | |
| **0.0** | -0.002 | 0.001 | 0.001 | -0.001 | -0.001 | -0.006 | -0.002 | -0.001 | -0.002 | -0.002 | -0.007 | -0.002 | -0.001 | -0.002 | -0.002 |
| **0.2** | 0.000 | 0.001 | 0.001 | -0.001 | 0.000 | -0.007 | -0.003 | -0.001 | -0.002 | -0.002 | -0.007 | -0.003 | -0.001 | -0.002 | -0.002 |
| **0.4** | 0.005 | 0.003 | 0.002 | 0.001 | 0.002 | -0.006 | -0.002 | -0.001 | -0.002 | -0.001 | -0.007 | -0.002 | -0.001 | -0.002 | -0.001 |
| **0.6** | 0.015 | 0.008 | 0.005 | 0.005 | 0.007 | -0.004 | -0.001 | 0.000 | -0.001 | -0.001 | -0.004 | -0.001 | -0.001 | -0.001 | -0.001 |
| **0.8** | 0.037 | 0.021 | 0.013 | 0.015 | 0.019 | -0.001 | 0.002 | 0.001 | 0.000 | -0.001 | -0.002 | 0.001 | 0.001 | -0.001 | -0.001 |
| | ----- $W_2$ ----- | | | | | ----- $W_2$ ----- | | | | | ----- $W_2$ ----- | | | | |
| **0.0** | -0.002 | 0.001 | 0.001 | -0.001 | -0.001 | -0.006 | -0.002 | -0.001 | -0.003 | -0.002 | -0.007 | -0.002 | -0.001 | -0.003 | -0.002 |
| **0.2** | 0.000 | 0.001 | 0.001 | -0.001 | 0.000 | -0.006 | -0.003 | -0.002 | -0.003 | -0.002 | -0.006 | -0.003 | -0.002 | -0.003 | -0.002 |
| **0.4** | 0.004 | 0.004 | 0.003 | 0.002 | 0.003 | -0.005 | -0.002 | -0.001 | -0.002 | -0.001 | -0.005 | -0.002 | -0.001 | -0.002 | -0.001 |
| **0.6** | 0.016 | 0.010 | 0.007 | 0.009 | 0.011 | -0.001 | 0.000 | 0.000 | -0.001 | -0.001 | -0.001 | 0.000 | 0.000 | -0.001 | -0.001 |
| **0.8** | 0.045 | 0.026 | 0.019 | 0.023 | 0.029 | 0.001 | 0.002 | 0.001 | 0.000 | -0.001 | 0.001 | 0.001 | 0.001 | -0.001 | -0.002 |
| | ----- $W_3$ ----- | | | | | ----- $W_3$ ----- | | | | | ----- $W_3$ ----- | | | | |
| **0.0** | -0.002 | 0.001 | 0.001 | -0.001 | -0.001 | -0.005 | -0.004 | -0.002 | -0.003 | -0.002 | -0.005 | -0.004 | -0.002 | -0.003 | -0.002 |
| **0.2** | -0.002 | 0.003 | 0.002 | -0.001 | 0.000 | -0.005 | -0.003 | -0.002 | -0.003 | -0.002 | -0.005 | -0.003 | -0.002 | -0.003 | -0.002 |
| **0.4** | -0.001 | 0.008 | 0.004 | 0.002 | 0.002 | -0.006 | -0.001 | -0.001 | -0.002 | -0.001 | -0.006 | -0.001 | -0.001 | -0.002 | -0.001 |
| **0.6** | 0.001 | 0.016 | 0.007 | 0.007 | 0.006 | -0.005 | 0.000 | 0.000 | -0.001 | -0.001 | -0.004 | 0.000 | 0.000 | -0.001 | -0.001 |
| **0.8** | 0.006 | 0.035 | 0.015 | 0.018 | 0.015 | -0.002 | -0.001 | 0.000 | 0.000 | 0.000 | -0.002 | 0.000 | 0.000 | 0.000 | 0.000 |

Notes: Bias was estimated as $\beta_0 - \hat{\beta}$ ; the difference between the true parameter value and its estimate.

Entries are calculated as the average of 1,000 simulations.

1/ To account for zeros in the first stage estimation, neighboring counts were transformed using the inverse hyperbolic sine estimation procedure of Burbidge et al. (1988).

2/ The *ad hoc* constant of $c$ = 0.5 was added to neighboring counts of zero to facilitate first stage.

Table 4. Root mean squared error of the AR estimator

| | ad hoc constant, $c$ = 0.5 | | | | | first-stage IHS estimator | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **ρ /*n*** | **100** | **256** | **484** | **729** | **1024** | **100** | **256** | **484** | **729** | **1024** |
| | $W_1$ | | | | | $W_1$ | | | | |
| **0.0** | 1.440 | 1.049 | 0.831 | 0.660 | 0.550 | 2.167 | 1.109 | 0.761 | 0.572 | 0.476 |
| **0.2** | 1.208 | 0.894 | 0.724 | 0.559 | 0.463 | 1.446 | 0.783 | 0.603 | 0.460 | 0.379 |
| **0.4** | 0.919 | 0.696 | 0.564 | 0.428 | 0.348 | 0.818 | 0.539 | 0.436 | 0.325 | 0.261 |
| **0.6** | 0.587 | 0.443 | 0.364 | 0.263 | 0.212 | 0.459 | 0.343 | 0.294 | 0.242 | 0.222 |
| **0.8** | 0.264 | 0.193 | 0.167 | 0.130 | 0.107 | 0.564 | 0.509 | 0.525 | 0.536 | 0.534 |
| | $W_2$ | | | | | $W_2$ | | | | |
| **0.0** | 1.211 | 0.895 | 0.718 | 0.541 | 0.456 | 1.867 | 0.934 | 0.647 | 0.473 | 0.398 |
| **0.2** | 1.040 | 0.763 | 0.623 | 0.467 | 0.389 | 1.278 | 0.659 | 0.515 | 0.391 | 0.323 |
| **0.4** | 0.804 | 0.604 | 0.487 | 0.362 | 0.299 | 0.707 | 0.475 | 0.379 | 0.288 | 0.232 |
| **0.6** | 0.495 | 0.377 | 0.312 | 0.227 | 0.184 | 0.422 | 0.323 | 0.278 | 0.235 | 0.215 |
| **0.8** | 0.217 | 0.169 | 0.150 | 0.122 | 0.101 | 0.556 | 0.516 | 0.528 | 0.518 | 0.516 |
| | $W_3$ | | | | | $W_3$ | | | | |
| **0.0** | 1.575 | 1.086 | 0.947 | 0.728 | 0.655 | 2.489 | 1.150 | 0.865 | 0.633 | 0.569 |
| **0.2** | 1.433 | 0.914 | 0.803 | 0.601 | 0.549 | 1.773 | 0.795 | 0.666 | 0.494 | 0.451 |
| **0.4** | 1.227 | 0.673 | 0.615 | 0.443 | 0.411 | 1.088 | 0.525 | 0.478 | 0.339 | 0.313 |
| **0.6** | 0.915 | 0.393 | 0.391 | 0.263 | 0.248 | 0.698 | 0.333 | 0.316 | 0.253 | 0.249 |
| **0.8** | 0.488 | 0.167 | 0.182 | 0.125 | 0.129 | 0.641 | 0.527 | 0.535 | 0.536 | 0.548 |

Table 5. Single–Unit Manufacturing Establishment Births (2000–2004), Aspatial Poisson ML Results

| | Poisson ML | | |
|---|---|---|---|
| Variable | Coefficients | S.E.[†] | Elasticities[††] |
| constant | –0.934*** | 0.281 | |
| msemp | 0.031*** | 0.004 | 0.464 |
| pelt10 | –0.002 | 0.002 | –0.116 |
| pemt100 | –0.029*** | 0.004 | –0.321 |
| tfdens | 0.006 | 0.010 | 0.0003 |
| mhhi | 0.020 | 0.009 | 0.017 |
| pop | 0.002*** | 0.0005 | 0.021 |
| cclass | 0.048*** | 0.013 | 0.852 |
| uer | 0.073*** | 0.022 | 0.310 |
| pedas | 0.130*** | 0.021 | 0.762 |
| awage | 0.019*** | 0.007 | 0.456 |
| netflow | 0.002 | 0.002 | 0.015 |
| proad | 0.103*** | 0.018 | 0.193 |
| interst | 0.007*** | 0.001 | 0.109 |
| avland | –0.009*** | 0.001 | –0.272 |
| bci | 0.080** | 0.037 | 0.480 |
| educpc | 0.004* | 0.002 | 0.051 |
| hwypc | –0.030 | 0.019 | –0.069 |
| metro | 1.265*** | 0.092 | 0.387 |
| micro | 0.572*** | 0.063 | 0.099 |
| Log Likelihood | –32,248 | | |
| Pseudo $R^2$ | 0.805 | | |

$N = 3078$; ***,**,* represent statistical significance at the 1%, 5%, 10% level, respectively. [†] Heteroskedasticity robust standard errors. [††] Elasticities represent the average effects across all observations.

Table 6. Single–Unit Manufacturing Establishment Births (2000–2004), Two-Step LIML SAR-Poisson Results

| | SAR-Poisson ad hoc constant, $c$ = 0.5 | | Elasticities†† | | SAR-Poisson IHS transformation | | Elasticities†† | |
|---|---|---|---|---|---|---|---|---|
| Variable | Coefficients | S.E.† | Direct | Indirect | Coefficients | S.E.† | Direct | Indirect |
| constant | –0.914*** | 0.288 | | | –1.020*** | 0.279 | | |
| $\rho$ | 0.181*** | 0.062 | | | 0.145*** | 0.049 | | |
| msemp | 0.026*** | 0.004 | 0.399 | 0.085 | 0.026*** | 0.004 | 0.400 | 0.066 |
| pelt10 | –0.003 | 0.002 | –0.132 | –0.028 | –0.003 | 0.002 | –0.132 | –0.022 |
| pemt100 | –0.027*** | 0.004 | –1.409 | –0.302 | –0.027*** | 0.004 | –1.401 | –0.231 |
| tfdens | 0.001 | 0.010 | 0.010 | 0.001 | 0.001 | 0.010 | 0.005 | 0.001 |
| mhhi | –0.012 | 0.008 | –0.425 | –0.091 | –0.012 | 0.008 | –0.424 | –0.070 |
| pop | 0.003*** | 0.0004 | 0.019 | 0.004 | 0.002*** | 0.0004 | 0.019 | 0.003 |
| cclass | 0.054*** | 0.013 | 0.930 | 0.199 | 0.054*** | 0.013 | 0.930 | 0.153 |
| uer | 0.048** | 0.024 | 0.207 | 0.044 | 0.047** | 0.024 | 0.205 | 0.034 |
| pedas | 0.117*** | 0.019 | 0.641 | 0.137 | 0.117*** | 0.018 | 0.670 | 0.111 |
| awage | 0.019** | 0.008 | 0.462 | 0.099 | 0.019** | 0.008 | 0.461 | 0.076 |
| netflow | 0.001 | 0.002 | 0.011 | 0.002 | 0.001 | 0.002 | 0.011 | 0.002 |
| proad | 0.082*** | 0.017 | 0.152 | 0.032 | 0.082*** | 0.017 | 0.151 | 0.025 |
| interst | 0.006*** | 0.001 | 0.089 | 0.019 | 0.006*** | 0.001 | 0.088 | 0.015 |
| avland | –0.006*** | 0.001 | –0.189 | –0.041 | –0.006*** | 0.001 | –0.189 | –0.031 |
| bci | 0.113*** | 0.032 | 0.079 | 0.017 | 0.114*** | 0.032 | 0.080 | 0.013 |
| educpc | 0.004** | 0.002 | 0.049 | 0.011 | 0.004** | 0.002 | 0.049 | 0.008 |
| hwypc | –0.034* | 0.020 | –0.061 | –0.013 | –0.034* | 0.020 | –0.061 | –0.010 |
| metro | 1.182*** | 0.096 | 0.410 | 0.088 | 1.184*** | 0.095 | 0.410 | 0.068 |
| micro | 0.545*** | 0.059 | 0.118 | 0.025 | 0.547*** | 0.059 | 0.119 | 0.020 |
| $c$ | 0.5 | fixed | | | | | | |
| $\theta$ | | | | | 0.781*** | 0.046 | | |
| Log Likelihood | –31,125 | | | | –31,108 | | | |
| Pseudo $R^2$ | 0.812 | | | | 0.812 | | | |

$N$ = 3078; ***,**,* represent statistical significance at the 1%, 5%, and 10% level, respectively. † Heteroskedasticity robust standard errors. †† Elasticities represent the average effects across all observations.