The Value of Clean Dairy Air: Accounting for Endogeneity and Spatially Correlated Errors in a Hedonic Analyses of the Impact of Animal Operations on Local Property Values

Isaac Bayoh, Elena Irwin, and Brian Roe[1]

Contact:
Elena G. Irwin, Assistant Professor
Dept. of Agricultural, Environmental, and Developmental Economics
Ohio State University
2120 Fyffe Rd, Columbus, OH 43210
Telephone: 614-292-6449
Fax: 614-292-0078
Email: irwin.78@osu.edu

Paper prepared for the American Agricultural Economics Association Annual Meeting, Denver, CO, August 1-4, 2004.

Short Abstract: We study the effect of livestock operations on property values using a hedonic analysis in five Ohio townships. Unlike previous studies, we account for endogenous livestock location variables and spatially correlated errors. Results suggest failure to correct for these problems results in biased estimates of livestock impacts on property values.

[1] The authors are Graduate Student, Assistant Professor and Associate Professor, Department of Agricultural, Environmental and Development Economics, Ohio State University. The authors would like to thank Joe Beiler and Jon Rausch for their considerable contributions to the conception and data gathering portions of this research. All remaining errors are those of the authors. The authors also recognize support from the Center for Urban and Regional Analysis at Ohio State University and from the Ohio Agricultural Research and Development Center.

Introduction

The costs and benefits of livestock operations for local communities is often a hotly debated issue among local residents, farmers, and public officials. While such operations have been demonstrated to create some positive impacts through jobs creation and increased tax revenues (Hamed et al, 1999), animal facilities can also generate a range of negative externalities for proximate residences. Potential costs include odors, increased traffic, and increased water treatment costs that not only can affect rural residents' quality of life, but also property values. Evaluating the impact of livestock operations on the tax base is critical as property taxation provides the bulk of funds for local schools and services in many communities. Development strategies that encourage or discourage development of livestock operations will impact the property tax base both for the parcel of land that might house the livestock operation and, potentially, through spatial externalities that may influence the value of neighboring parcels.

Despite the importance property values play in rural communities and the growing interest in understanding the impacts of animal operations on property values, relatively few studies statistically analyze the impact of animal operations on property values (e.g., Ables-Allison and Connor 1990; Herrings, Secchi, and Babcock, 2003; Palmquist, Roka and Vukina, 1997; Park, Seidl and Davies 2002; Ready and Abdalla, 2003; Seipel *et al*. 1998; Taff, Tiffany and Weisberg 1997). These studies, all of which employ a hedonic pricing model to estimate the marginal effects of nearby livestock on housing property values, have demonstrated a range of effects. For example, an Iowa study of hog operations estimated that the location of a mid-sized hog operation within one-fourth of a mile to the northwest of a residence reduces its value by 26% (Herrings, Secchi, and Babcock, 2003), whereas other studies from Colorado (Park, Seidl and Davies, 2003) and Minnesota (Taff, Tiffany and Weisberg, 1997) have actually found a positive effect associated with nearby animal operations. While it is expected that effects will vary across regional markets and geographical areas, such wide-ranging results make it difficult for researchers and policymakers to generalize these results beyond the specific study area.

Hedonic models are subject to several methodological concerns that, if not addressed, can generate biased or inefficient estimates Two issues that are particularly relevant to the use of hedonic models in estimating spatial externality effects are the potential endogeneity of the location of the spatial externality source (in this case, livestock locations) and spatial error autocorrelation (Irwin and Bockstael, 2001; Irwin, 2002). Of the studies listed above, only Park, Seidl and Davies (2003) address spatial autocorrelation and none address the potential problem of endogeniety associated with livestock location (although Ready and Abdalla (2003) address this issue for other neighboring land uses). Thus, the divergence in empirical results may be in part due to these methodological problems.

Endogeneity problems associated with the measurement of proximate livestock effects will arise if the location of livestock facilities is influenced by residential property values in an area. For example, rather than animal operations causing neighboring parcels to have a low transaction price, it may be that low transaction prices attract large animal operations to the area.[2] Indeed, anecdotal evidence suggests that many large animal operations choose a site because it is far from other people and employment centers and, hence, potentially unattractive for development due to transportation costs. Problems of spatial error autocorrelation can arise, which, if the livestock variables are exogenous, will merely cause estimates to be inefficient (and thus can invalidate hypothesis testing). However, if the livestock variables are endogenous, the spatial error autocorrelation will generate a second source of bias in the estimated impact of livestock on property values (Irwin 2002).

We study the distant-dependent effects of livestock operations and their related characteristics (e.g., size and species) on residential property values in five Ohio townships using housing transactions data from 1999-2001. We begin by estimating a standard hedonic model using ordinary least squares. We then test for residual spatial error correlation, using a Lagrange multiplier test, and for the endogeneity of the livestock variables using a Hausman test. Both test

[2] See Irwin and Bockstael (2001) and Gayer (2000) for similar arguments in the context of neighboring open space and hazardous waste sites respectively.

statistics are found to be highly significant. We then correct for spatial autocorrelation using a spatial error model (Anselin, 1988) and for both spatial error and endogenous regressors using a three-stage GMM estimation approach developed by Kelejian and Prucha. In comparing the results across these three models, we find the significance and sign of the livestock variables changes across models. The results imply the impact of livestock operation on residential property values can be misrepresented when these methodological concerns are not addressed. We conclude that, without careful study of both the issues of spatial correlation of errors and the possible endogeneity of livestock locations, the existing body of empirical estimates may offer little unbiased guidance during community discussions, community level policy making and court decisions.

Empirical Evidence of the Impact of Livestock Operation on Housing Values

The existing empirical evidence on the impact of livestock operations on residential property values reveals a range of estimated effects. Several studies find that proximity to various animal facilities depresses housing values. Palmquist, Roka, and Vukina (1997), in the only published study to-date, evaluated the influence of hog feedlot operations on 237 residential properties in North Carolina and find a 4.75% decrease in property values for houses located within a half-mile of the animal facilities. Herriges, Secchi and Babcock (2003) use 1,145 rural residential homes sales transacted between 1992 and 2000 in Iowa and find that the magnitude of the effect of hog facilities on the value of neighboring residential property sales value depends both on distance and the direction of the facility from the house. For example, they estimate that a 250,000 pound capacity operation located ¼ of a mile to the northwest of a house is associated with a 26% lower housing sales price while the same operation located ¼ mile to the east of the house is associated with a 13% (though statistically insignificant) lower sales price. Ables-Allison and Connor (1990) use property sales from 1986-89 from areas located near eight Michigan hog operations and found modest negative effects that decayed with distance. For example, for properties within 1.6 miles of the operations, an additional hog reduced the sales value of the average property by $1.74 per property; by $0.53 per property for houses located within 1.6 to 2.3 miles; and by $0.13 per property for houses located within 2.3 to 3.5 miles of the hog operations. Hamed, Johnson and Miller (1999) estimate a loss of value of land located within three miles of a contained animal feeding operation (CAFO) to be about $112 per acre using data on 99 rural arm-length transactions from Missouri. Lastly, Ready and Abdalla (2003) estimate the effects of poultry, cattle, and hogs using 8,090 sales of single family homes sold between 1996 and 2002 in Pennsylvania and find that only facilities located within 1.6 kilometers (about one mile) of houses had a significant impact on sales values. Poultry facilities had the largest, negative effect on property values while cattle (beef and dairy combined) had the smallest, though the differences between species was only marginally different in a statistical sense. Averaged across all species, a new cluster of animal production buildings of average size (200 animal equivalents or larger) located within 500 meters (1/3 of a mile) of a house is estimated to lower its resale value by about 6.4%; by about 4.1% within 800 meters (1/2 mile); and by 1.6% within 1200 meters (3/4 mile). Interestingly, this is one of the few studies to compare the effect of animal agriculture equally to other land uses such as landfills and airports. The effect of animal agriculture was on house price was about half to two-thirds as large as the effect of landfills and did not extend as far.

In contrast, other studies have found positive effects associated with proximate livestock operations. For example, in the only study that accounts for spatial error autocorrelation, Park, Siedl, and Davies (2003) used 3,354 residential housing sale data from Colorado to test for the varying impacts of livestock operations by species. Their results indicate that the effect of livestock on property values differed by species, size and distance between the animal facility and the house. For example, the location of averaged-sized beef or dairy operation (about 330 animal units) or of a poultry operation (about 430 animal units) were found to generate positive impacts on the values of houses located within a mile of the operations. Taff, Tiffany and Weisberg

(1996), in a study of 292 rural residential properties transacted in 1993 and 1994 in Minnesota, found that the addition of a feedlot of an existing home could improve property values by as much as 9.4 percent. The authors suggest that increased demand from feedlot owners or workers may be causing the positive effect of livestock proximity on house prices.

While the precise estimates of livestock impacts on housing values varies considerably across these studies, other findings appear somewhat consistent. For example, there is evidence from several studies that impacts vary by size and species, with hogs often found to be the most offensive. For example, both Herriges, Secchi and Babcock (2003) and Park, Seidl and Davies (2003) find that larger hog operations are associated with lesser impacts on property values. The authors suggest that the larger facilities tend to be newer and incorporate either better management or improved technology, each which can ameliorate the negative effects of the facilities on neighboring property values. Lastly and not surprisingly, most studies find that the impacts diminish relatively quickly over distance. Ables-Allison and Connor (1990) appear to find the most persistent effects over distance (a decrease of $0.13 per property per hog for houses located within 2.3 to 3.5 miles of the livestock operation), which extends only to 3.5 miles. Most of the other studies do not report significant effects beyond one mile.

The Hedonic Pricing Model and Spatial Externality Effects

From valuing quality change in the automobile industry (Griliches, 1971) to assessing the cost of crime rates on housing values (Gibbons, 2002), the hedonic pricing model has been widely applied in economics. This approach is used to estimate the implicit prices of attributes associated with any differentiated market good that is described by a bundle of characteristics. The hedonic price function relates the price of the differentiated market good to the quantities associated with its attributes and the partial derivative of the hedonic price function with respect to any given attributes is its implicit market price (Freeman, 1993). In case of housing, houses are seen as a bundle of attributes that include the structural characteristics of the house, neighborhood features, and the relative location of the house to surrounding amenities or disamenities. The general formulation of a hedonic pricing model is:

$$P = \alpha + \beta X + \gamma N + \varepsilon \quad (1)$$

where P represents the value of the house, X is a vector of house characteristics and A a vector of location characteristics. In the standard model, the vector of residuals ε is assumed to be iid.

Even though the hedonic pricing model is a method that can generate useful results, there are several methodological challenges associated with estimating unbiased and efficient estimates. Specification of the model's functional form is difficult as theory suggests little in the way of specific forms or restrictions and thus the selection of functional form is largely a strategic one (Haab and McConnell, 2003). Cropper et al (1987) use Monte Carlo simulations to examine the influence of functional form on deviations of the estimated marginal price from their true values. When all attributes are observed, the linear and quadratic box-cox functions perform better than the semi-log and double-log functions. However, in the presence of omitted attributes, the double-log and semi-log perform best. Because of the nature of housing markets, many structural and neighborhood variables are highly collinear and thus multicollinearity is another frequent issue in hedonic estimation. Other issues include defining the extent of the housing market and the sample selection problem that arises due to the fact that the only houses for which we observe a price are those that were actually sold.

Two additional methodological concerns arise when using hedonic pricing models of housing to estimate the effects of spatial externalities: the potential endogeneity of the variable generating the spatial externality and spatial error autocorrelation (Irwin and Bockstael, 2001). The endogeneity issue arises from the possibility that the neighboring variable (e.g., land use) that is generating the spatial externality of interest is itself a function of the residential land market. For example, the amount of open space around a house is likely to be endogenous in a hedonic model of housing values since open space is often privately held and can be developed as residential land (Irwin and Bockstael, 2001). In our case, it is certainly possible that the location of livestock facilities is influenced by an area's residential land market since operators may purposefully choose lower valued land that is relatively far from more populated areas. If so, then the estimates of the livestock effect will be biased. In particular, a model in which this endogeneity was not corrected could overstate the negative spatial effect of livestock on housing values since it would ignore the fact that livestock operations and lower valued land are spatially correlated due to the operator's location decision. The traditional way that researchers have dealt with an endogeneous explanatory variable is by using one or more instrument variables (IV) that serve as a proxy for the endogenous variable. The most common IV method is the two-stage least squares (2SLS) model in which IV are used to get an estimate of the endogeneous variables. The results of the first stage are then included in the model as variable instead of the actual measure.

The second methodological challenge that is particularly challenging when using a hedonic model to estimate the influence of spatial effects is spatial error autocorrelation. Spatial effects in a model are typically associated either with a substantive spatial interaction process or with spatially correlated errors, which are the result of either omitted variables that are themselves spatially correlated or mismeasurement of one or more spatial explanatory variables. While some researchers have focused on substantive spatial interactions among neighboring housing prices (Can, 1992; Ionnides, 2003), we focus here on the problem of spatial error correlation, which is more likely to be an issue in the context of spatial externalities. In mathematical terms, spatial error autocorrelation can be expressed by a condition on the moment matrix, where ε is a vector of spatially correlated error terms and *i* and *j* are neighbors:

$$Cov(\varepsilon_i,\varepsilon_j)=E(\varepsilon_i,\varepsilon_j)-E(\varepsilon_i).E(\varepsilon_j) \neq 0 \quad \text{for } i \neq j \tag{2}$$

The statement above states that autocorrelation should be suspected when the covariance matrix is non-zero, in which case the OLS estimates are still unbiased, but are inefficient. However, as illustrated by Irwin and Bockstael (2001), if both spatial error correlation and endogenous spatial externalities are present in a hedonic model, then the spatial error correlation will create an additional source of bias associated with the estimated spatial externality coefficient. This is because, in the case of endogeniety, the land use on neighboring parcel *j* that generates the spatial externality that impacts the price of house *i* is itself a function of house *i*'s price. Thus, if the error terms associated with neighbors *i* and *j* are correlated, then the spatial externality variable, which is a function of the residential value of parcel *j* and thus the error term associated with *j*, will be biased due to this correlation.

There are several techniques that researchers have employed to control for spatial error autocorrelation. By far the most common is to assign structure to the underlying spatial error process by defining the underlying spatial relationship among all locations. Given this assumed structure, which is a maintained assumption of the model, the number of parameters associated with the spatial error process that need to be estimated is reduced to just one. While a number of different processes are possible, the most commonly used are the spatial moving average (SMA) and the spatial autoregressive (SAR) processes. Given the specification of the general model in (1), the SMA and SRA error processes are defined respectively as:

$$\varepsilon = \rho W \varepsilon + \mu \quad \text{and} \tag{3}$$

$$\varepsilon = \rho W \mu + \mu, \tag{4}$$

where ε is now an Nx1 vector of spatially correlated error terms; μ is an Nx1 vector of iid errors; W is an NxN spatial weights matrix comprised of elements w_{ij} that represent the researcher's maintained assumption regarding the spatial relationship between observations *i* and *j*; and ρ is the spatial correlation parameter to be estimated. For each *i*, the element w_{ij} of the matrix W will be non-zero only if *j* belongs to the neighborhood set (that is, if *j* and *i* are spatially correlated). By convention, the diagonal elements of W, w_{ii}, are set to zero. The specification of which process to use is based on the researcher's conception of the spatial error process. The SMA process is used when the researchers believe that the stochastic correlation exist only between neighbors. Alternatively, when the researcher believes that the error process propagates throughout the system, such that all the observations are related to all others, then the SAR process is proven to be the one to use. Likewise, the specification of W is also at the researcher's discretion. Many options are available. The most common are binary contiguity weights, in which $w_{ij} = 1$ if i and j are neighbors and 0 otherwise and inverse distance weights, in which $w_{ij} = \frac{1}{d_{ij}}$ with d_{ij} being the distance between i and j. Given the specification of the error process, estimation of this model proceeds by transforming the model such that the resulting error, μ, is iid. For example, in the case of the SAR error process, the model is transformed by using matrix notation to rewrite (3) as $\varepsilon = (I\text{-}\rho W)^{-1}\mu$ and substituting this into (1) and then rewriting the model so that the error is iid:

$$P = \alpha + \beta X + \gamma N + (I - \rho W)^{-1} u \qquad (5)$$

Because the results of the model are conditional on the specification of *W* and the spatial error process, it is standard to estimate a spatial error model with alternative specifications of *W* to gauge the extent to which the results are robust to different specifications.

Alternative techniques that attempt to correct for spatial error autocorrelation include the spatial sampling and the spatial differencing techniques. The spatial sampling technique consists of selecting a subset of observations from the full sample such that the spatial correlation among the selected observations is small or non-existent. Because spatial dependency is generally believed to be greater the closer two locations are to each other, this method relies on selecting observations by random and then dropping out their nearest neighbors so that the remaining observations are a minimum distance apart (Carrion-Flores and Irwin, in press; Irwin, 2002). The spatial differencing method is analogous to the first differencing method used in panel datasets and relies on creating new observations by differencing two correlated observations. Observations within a certain distance or neighborhood are transformed by subtracting a selected observation's attributes from the neighboring attributes. In this way, new observations are created that are o longer spatially dependent and thus the error term is presumed to be iid (Gibbons, 2002).

An additional challenge arises when both endogenous spatial externality variables and spatial error autocorrelation are present. Because the standard correction for spatial error autocorrelation relies on transforming the model such that the transformed errors are no longer spatially correlated (e.g., equation 5), estimation of this model requires non-linear estimation methods. This prevents the use of standard IV models that are used to correct for endogenous regressors IV, such as 2SLS. An extension of this approach that allows for the spatial error correction to be incorporated into a 2SLS approach is the generalized spatial 2SLS model developed by Kelejian and Prucha (1998). This model was developed for the estimation of endogenous spatial lags with spatial errors; we adapt it to estimate the hedonic model with spatial error and the endogenous livestock variables. The estimation of this model is broken down in three stages that allow estimation of an initial vector of coefficients using 2SLS, β_{2SLS}, which are used to generate an estimate of the spatial error coefficient, ρ. The parameter vector β_{2SLS} is then reestimated in a model that controls for the spatial error autocorrelation using the estimated value of ρ. More specifically, the first stage of the estimation yields two stage least squares estimates:

$$\beta_{2SLS} = \left(x'Px\right)^{-1}\left(x'Py\right) \qquad (6)$$

where β_{2SLS} is the vector of parameter estimates, x is a matrix of independent variables, y is the vector of dependent variables, and P is the projection matrix. The matrix P is built from a matrix of instruments (H): $P = H(H'H)^{-1}H'$. The estimated parameter vector β_{2sls} is consistent. However, it does not incorporate any information about the spatial process and thus is inefficient. The second step consists of using Generalized Method of Moments (GMM) estimation to obtain estimates for ρ and σ_ε^2 which are the spatial error coefficient and the variance of the iid error term respectively. Kelejian and Prucha (1995) suggest the following three-system equation:

$$\begin{aligned} \tilde{u} &= y - x\beta_{2sls} \\ \tilde{\tilde{u}} &= W\tilde{u} \\ \tilde{\tilde{\tilde{u}}} &= W^2\tilde{u} \end{aligned} \qquad (7)$$

Let

$$\Gamma = \frac{1}{n}\begin{pmatrix} 2(u'\tilde{\bar{u}}) & -(\tilde{\bar{u}}'\tilde{\bar{u}}) & 1 \\ 2(\tilde{\bar{\bar{u}}}'\tilde{\bar{u}}) & -(\tilde{\bar{\bar{u}}}'\tilde{\bar{\bar{u}}}) & tr(W'W) \\ E(\tilde{u}'\tilde{\bar{\bar{u}}} + \tilde{\bar{u}}'\tilde{\bar{u}}) & -(\tilde{\bar{u}}'\tilde{\bar{\bar{u}}}) & 0 \end{pmatrix}$$

$$\gamma = \frac{1}{n}\begin{pmatrix} \tilde{u}'\tilde{u} \\ \tilde{\bar{u}}'\tilde{\bar{u}} \\ \tilde{u}'\tilde{\bar{u}} \end{pmatrix}$$

The relationship between Γ and γ is given by

$$\gamma = \Gamma\alpha + \nu$$

ν being a set of residuals

$$\text{where } \alpha = \begin{pmatrix} \tilde{\rho} \\ \tilde{\rho}^2 \\ \tilde{\sigma}_\varepsilon^2 \end{pmatrix}$$

The first and the third elements of the vector α are estimates of ρ and $\tilde{\sigma}_\varepsilon^2$, but a more efficient set of estimates is given by the nonlinear least squares estimator that is based on the following relationship $\gamma = \Gamma\alpha + \nu$. These consistent and more efficient estimators are $\tilde{\tilde{\rho}}$ and $\tilde{\tilde{\sigma}}_\varepsilon^2$.

$\tilde{\tilde{\rho}}$ and $\tilde{\tilde{\sigma}}_\varepsilon^2$ are defined as the minimizers of

$$\left[\gamma - \Gamma\begin{bmatrix} \tilde{\rho} \\ \tilde{\rho}^2 \\ \tilde{\sigma}_\varepsilon^2 \end{bmatrix}\right]' \left[\gamma - \Gamma\begin{bmatrix} \tilde{\rho} \\ \tilde{\rho}^2 \\ \tilde{\sigma}_\varepsilon^2 \end{bmatrix}\right] \tag{8}$$

Given the estimate of ρ from the generalized moment estimation outlined above, the third and final step is to transform the original regression equation via a Cochran-Orcutt type transformation to account for the spatial error autocorrelation and use this to reestimate the vector of regression parameters by 2SLS.

The Effects of Livestock Operations on Residential Property Values

We begin by restating the model in (1) to make explicit the hypothesized role of livestock effects in a hedonic model of residential property values:

$$P_i = \alpha + \beta X_i + \gamma N_i + \lambda A_i + \varepsilon_i \qquad (9)$$

where P_i is the transacted price of parcel i, X_i is a vector of parcel i's observable characteristics (e.g., lot size, age, size, amenities of house, road access), N_i a vector of locations attributes, A_i is a vector of variables that capture the intensity of animal production on the land around parcel i (e.g., distance to nearest animal operation, total distance-weighted and species-adjusted number of animals near the parcel, direction of parcel from the animal facility if prevailing winds exist), ε_i is the unobserved term which may be spatially correlated, and α, β, γ, and λ are parameter vectors to be estimated.

Data

The data are from a rural county in Ohio, Mercer County, a rural county located in the far western portion of the state adjacent to the Indiana border. With a total population of 40,000 people in 2000 and a single larger urban area of about 10,000 people, Mercer County's principal economic activity is agriculture. It is the top hog producer and second corn producer in the state of Ohio and approximately 88% of the county land is in agriculture. The data consist of 3,476 residential housing sales data from 1999 to 2001 from within a five township region of Mercer County. These data were made available to us in geocoded format by the Mercer County Auditor's Office and, in addition to the transacted price, include a variable of structural variables that describe the size and amenities associated with the house and lot. In order to include only arms-length transactions, we selected these observations from a larger set of observations by selecting only those with a sale price of at least 50% of the assessed value.[3] Additional

[3] This method of data selection has been used by Park, Seidl, and Davies (2003) in their study of the impact of livestock operation on residential property values in northern Iowa.

neighborhood and locational variables were generated for each observation using a Geographic Information System (GIS).

Data on the location, size, and type of livestock operations were collected by visual inspection of all agricultural land in the five township study area. Extension agents familiar with the area conducted a tour of the townships to verify the auditor's information on animal agricultural land uses and to record additional information on the livestock farms, including the species (dairy/beef, hogs, or poultry) and approximate size (number of head). This visual inspection revealed many properties that were recorded as animal operations in the auditor dataset that are no longer active and all of which were dropped from our dataset. Thus the visual inspection ensures us of having a dataset of premium quality containing all the pertinent information. Based on this, we used GIS to calculate a set of spatial livestock variables that captured the potential dimensions of the spatial effects associated with livestock. In particular, we hypothesize that the effects may vary by proximity to nearest facility as well as by the total number of animals within a given area of each house and that these effects are likely to vary by species. Accordingly we generate a set of distance variables that measures the distance from each house to the nearest livestock facility by species and density measures of the number of head of each species within 500, 1000, 2500, and 5000 feet of each house. Aggregate livestock measures of both the distance and density variables were created. In generating the density variables it was necessary to transform animals into equivalent units. We used the following equivalency factors developed by researchers at Iowa State University: 1.4 for each cow, 0.4 for each finishing hog and 0.0025 for each chicken. We use the two aggregate measures of livestock distance and density in the following estimation because of the limited number of instruments we use in the first stage estimation.

The other variables used to estimate the model include structural housing variables, available from the tax auditor database, and neighborhood attributes, generated using GIS. Inclusion of structural characteristics is necessary to control for variations across housing stock.

Several different structural attributes are included in order to control for this. These include the number of bathrooms, total square feet of the house, number of stories, approximate age, and lot size. In addition, we include several categorical variables indicating the type of heating system, the presence of a fireplace, the construction grade of the house, and the overall condition of the house. Lastly, we include the assessed value of any additional structures that are present on the property. In additional to the characteristics of the house itself, the value of a property is also a function of the attributes of a neighborhood. To control for neighborhood effects, several location-specific variables are included. These include a set of dummy variables delineated the local jurisdiction and school district to which each house belongs; distance to the local urban center, Celina, as well as to the nearest metropolitan area, Fort Wayne, Indiana; proximity to highway; and neighborhood population density, which is calculated at the block group level.

Empirical Results

We first estimate the hedonic model specified in (6) using ordinary least squares (OLS) estimation. Because it is likely that we have omitted variable in our dataset, we follow the recommendation from Cropper et al (1987) results and used a log-log formulation in which the left hand side variable and most of the right hand side were transformed into logarithm.[4] The results of this estimation are reported in the second column of Table 1. The OLS estimation yields a R-squared of 0.53. The results suggest that all the house characteristics variables are significant and of the expected sign, e.g., the number of bedrooms and building condition are positive and very significant. Other variables such as the age of the house are negative and significant revealing consumers' preference for newer houses. While the structural variables are significant and of the expected sign, we find that proximity to Fort Wayne, the nearest metropolitan area, is not significant. On the other hand, there is a premium for houses located in higher density areas, perhaps because high density areas provide more services such as banks and

[4] Exceptions include dummy and categorical variables as well as distance to highway, which enters as a quadratic.

hospital. The results also suggest that some school districts and some townships are more attractive than others. Lastly, the aggregate livestock variables results suggest that animal density around the residential property is insignificant and that proximity to an animal farm is *positive* and significant (since the distance variable is negative and significant).

However, as discussed earlier, these results are potentially biased due to the potential for spatial error correlation and endogenous regressors. Using the residuals from the OLS model, we test for both the presence of spatial error correlation and the endogeneity of the livestock variables. To test for spatial error, we used a Lagrange multiplier test, which provides the LM statistic with a critical value of 6.635. The calculated LM statistic is 15.58 and thus we reject the hypothesis of no spatial error. The Hausman test for endogenous livestock variables yields a statistic of 3.17×10^{15}, leading us to clearly reject the null hypothesis of exogenous livestock variables.

To examine the influence of spatial error and endogeneity separately on the coefficient estimates, we proceed by first estimating a model that corrects for spatial error only and then a model in which both the spatial error and endogeneity problems are corrected. We assume that the spatial error process follows a spatial autoregressive pattern and therefore, we estimate the model presented in equation (5). As discussed earlier, the specification of the weights matrix *W* is subject to the researcher's apriori beliefs about the underlying spatial structure of the data. We hypothesize that the error process is driven by omitted variables that are themselves spatially correlated, which are most likely to cause a declining gradient of spatial dependence over distance. For this reason, we specify a weight matrix based on the inverse distance between observations with a maximum cutoff point of 5000 feet (which is just less than one mile).

The results of the spatial model, reported in column 3 of Table 1, are not noticeable different with respect to housing characteristics; they all remain significant without any sign changes. However, the significance levels of some variables changes. Distance to Fort Wayne, which was insignificant by OLS, becomes positive and significant once the spatial error correction is in place. Overall there is an increase of the significance of the location variables. There are noticeable changes in the livestock variable estimates. The density variable estimate's sign switches to positive, but remains insignificant, while the distance variable loses significance when the spatial error correlation is controlled. Thus results from the spatial error model suggest that proximity to a livestock facility does not affect residential property value at all. Finally, the spatial error coefficient, ρ, is positive and significant, indicating a strong correlation between observations located nearby.

While these results are of some interest, they are unreliable due to the significant endogeniety of the livestock variables, which will bias the results with respect to the estimated livestock coefficients. To account for both the endogneity and spatial error autocorrelation problems, we employ the generalized spatial two-stage least squares (GS2SLS) model developed by Kelejian and Prucha (1998) presented in the previous section (equations 6-8). To identify appropriate instruments, we seek variables that explain the location of livestock (or more precisely, the density and location of livestock relative to houses in our sample), but that are not themselves correlated with the error term. Because livestock operators will seek to minimize their transportation costs of inputs and outputs, proximity to highway is a reasonable instrument. We include this term and the square of this term in the first stage hedonic estimation. The results of this model are reported in column four of Table 1. The results from this model are quite similar to the OLS results. All of the structural housing variables remain of the same sign and highly significant. Unlike the spatial error model estimate, the distance to Fort Wayne variable is once again insignificant. Neighborhood population density remains significant and positive and none of the other location variables switch signs or significance levels.

The most significant difference between the spatial error and generalized spatial 2SLS models resides in the livestock variables. They were all insignificant under the spatial error model. Upon implementing the generalized spatial 2SLS model, the two livestock variables are both positive and significant, indicating that, while distance from the nearest livestock facility is now positive (reflecting negative externalities), adjacency to livestock (i.e., density of livestock within 500 feet) is also positive. This suggests that two variable have competing effects. Using these results, consider how the construction of a standard size livestock facility (1,000 animal units) would affect the price of a house situated within 500 feet. If the house was previously surrounded by livestock, e.g., the new livestock facility lowers the distance from the house to its nearest livestock facility by only 1000 feet, the price of the property would actually appreciate according to the estimated model (-1000 feet * 0.0195 + 1000 animal units within 500 feet * 0.187 = $167.50). Alternatively, consider a case where the house had no livestock within three miles before the arrival of the new, 1,000 unit facility. In such a case the property value depreciates (-15,000*0.0195 + 1000*0.187 = -$105.50). This suggests that the qualitative effect of additional livestock is dependent upon the existing density of livestock in the immediate neighborhood of the house and that, in areas already heavily populated with livestock, additional livestock may possibly have a positive impact on housing prices.

Summary and Conclusions

We estimate a hedonic housing price model to estimate the impact of proximate livestock operations on residential property values. Preliminary results indicate that there are both positive and negative effects on residential property values associated with nearby animal operations. After controlling for both spatial error autocorrelation and the endogeneity of livestock locations, we find that distance from livestock facilities increases housing values, but that adjacent livestock operations *increase* the value of a house. While these results are quite preliminary, such a finding may capture that the property value effects caused by changes in the size and location of livestock facilities depends crucially on the existing levels of livestock. Specifically, we find preliminary evidence that additional livestock by parcels previously isolated from livestock facilities would drive property prices down while additional livestock by parcels already surrounded by livestock may actually cause prices to appreciate.

However, these results are very preliminary and may change once additional refinements to the empirical model have been made. In our next iteration of the model we plan to account for wind direction and to include additional instruments in the generalized 2SLS model, such as soil quality, parcel size, and water availability associated with surrounding livestock farms. In addition, we will consider combining this model with a control function model to account for the potential bias introduced by missing variables.

Although still preliminary, our results have implications both for methodology and policy. First, from a methodological perspective, we find that both spatial error autocorrelation and the endogeneity of the livestock variables are present and highly significant. Our results are consistent with Irwin and Bockstael (2001), who argue that, if the spatial externalities are endogenous, the estimates will be biased both because of the usual endogeneity bias as well as from the spatial error autocorrelation. In our case, the estimated coefficient associated with the livestock distance variable flipped from negative and significant in the OLS specification to positive but insignificant when just the spatial error was controlled to negative and significant when both the error and endogneity were controlled. On the other hand, the estimate of the livestock adjacency coefficient was insignificant in both the OLS and spatial error models, but became positive and significant when both the error and endogeneity were controlled. In contrast, most of the other estimated coefficients of the model remained stable across the different models (with the exception of proximity to Fort Wayne), suggesting that the spatial error, as expected, did not bias these estimates. The changes of sign and significance of location and livestock variables demonstrate the importance of accounting for endogeneity and spatial error when using a hedonic model to estimate spatial externality effects.

Our results have implications for local economic development strategies and policies regarding animal agriculture. As the public controversy surrounding the citing of large livestock enterprises continues to grow, many counties are considering altering property tax valuation methods to compensate for property value losses thought to be caused proximate livestock operations (e.g., Livingston v. Jefferson County Board of Equalization) and private land owners adjacent to livestock operations are using the civil court system to demand for compensation for property value losses (e.g., Blass, *et al.* v. Iowa Select Farms). We demonstrate that careful study of livestock effects necessitates consideration of both spatial error correlation and the possibility of the endogeneity of livestock citing. Given that most of the existing empirical studies have not accounted for these effects, the ability of these results to provide unbiased guidance in community level policy making and court decisions is called into question.

References

Abeles-Allison, M and Connor, L..J. 1990 "Analysis of Local Benefits and Costs of Michigan Hog Operations Experiencing Environment Conflicts," Agricultural Economics Report #536, Department of Agricultural Economics, Michigan State University.

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

Anselin, L. 2001. "Spatial Econometrics." In *A Companion to Theoretical Econometrics*: Blackwell Publishers.

Anselin, L 2002. " Under the Hood: Issues in the Specification and Interpretation of Spatial egression Models", *Agricultural Economics* 27, 247-267.

Anselin, L 2003. "Spatial Externalities, Spatial Multipliers and Spatial Econometrics," *International Regional Science Review*," 26 153-166.

Anselin, L and Kelejian, H. 1997. " Testing for Spatial Error Autocorrelation in the Presence of Endogeneous Regressors," *International Regional Science Review*, 20 153-183.

Bell, K and Bockstael, N. 2000. "Applying the Generalized Method of Moments Approach to Spatial Problems with Micro Level Data," *Review of Economic and Statistics*.

Brock, W., and S. Durlauf. 2001. "Interactions-Based Models." In The Handbook of Econometrics 5, ed. James J. Heckman and Edward E. Leamer. Amsterdam: Elsivier Science, North Holland.

Can, A. 1992, "Specification and Estimation of Hedonic Housing Price Models", *Regional Science and Urban Economics* 22, 453-474.

Cheshire, P., and S. Sheppard. 1995. "On the Price of Land and the Value of Amenities." *Economica* 62 (May): 247–67.

Cropper,M. L., L .B. Deck, andK. E.McConnell. 1988. "On the Choice of Functional Form for Hedonic Price Functions." *Review of Economics and Statistics* 70 (4): 668–75.

Freeman, M. 1993. The Measurement of Environmental and Resource Values. Washington D.C.: Resources for the Future.

Geoghegan, J., L. A. Wainger, and N. E. Bockstael.1997. ''Spatial Landscape Indices in a Hedonic Framework: An Ecological EconomicsAnalysis Using GIS.'' *Ecological Economics* 23 (3): 251–64.

Greene, W. H. 2000. Econometric Analysis, 4th ed. Upper Saddle River, N.J.: Prentice Hall.

Griliches, Z. 1971 *Price Index and Quality Change*. Havard University Press. Cambridge Massachussets.

Haab, Timothy C. and Kenneth E. McConnell, 2002. *Valuing Environmental and Natural Resources: The Econometrics of Non-Market Valuation*. Edward Elgar Publishing, LTD. Cheltenham, UK.

Hamed, M., Johnson, T. and Miller, K. 1999. "Impact of Animal Feeding Operations on Rural Land Values." Community Policy Analysis Center, University of Missouri-Columbia.

Herriges, J.A., Secchi, S., and Babcock, B.A. 2003. "Living with Hogs in Iowa: The Impact of Livestock Facilities on Rural Residential Property Values," Center for Agricultural and Rural Development Working Paper, 03-WP 342, Des Moines, IA, August.

Irwin, E .G., and N. E. Bockstael. 2001. "The Problem of Identifying Land Use Spillovers: Measuring the Effects of Open Space on Residential Property Values." *American Journal of Agricultural Economics* 83 (3): 698–704.

———. 2002. "Interacting Agents, Spatial Externalities, and the Endogenous Evolution of Residential Land Use Pattern." *Journal of Economic Geography* 2 (1): 31–54.

Kelejian, H and Prucha, I. 1995, "A Generalized Moment Estimator for Autoregressive Parameter in Spatial Model." Department of Economics, University of Maryland, Working Paper 95-03.

———. 1998. " A Generalized Spatial Two-Stage Least Squares Procedure for Estimating a Spatial Autoregressive Model with Autoregressive Disturbances." *Journal of Real Estate Finances and Economics*, 17, 99-121.

Manski, C. 1993. "Identification of Endogenous Social Effects: The Reflection Problem." *Review of Economic Studies* 60 (3): 531–42.

Park, D., Seidl A. F. and Davies, S.P. 2002. "Livestock Industry Location and Residential Property Values: A Spatial Hedonic Approach," Department of Agricultural and Resource Economics, Colorado State University.

Palmquist, R. B., Roka, F. M. and Vukina, T. 1997. "Hog Operation, Environment Effects and Residential Property Values" *Land Economics* 73(1): 114-124

Rapaport, Carol 1997 "Housing Demand and Community Choice: An Empirical Analysis" *Journal of Urban Economics*, 42, 243-260.

Ready, R and Abdalla, C. 2003, GIS Analysis of Land Use on the Rural-Urban Fringe: The Impact of Land Use and Potential Local Disamenities on Residential Property Values and on the Location of Residential Development in Berks County, Pennsylvania. Staff paper, Department of Agricultural Economics and Rural Sociology, The Pennsylvania State University.

Rosen, S. 1974 "Hedonic Price and Implicit Market: Product Differentiation in Pure Competition" *Journal of Political Economy*. 82, 34-55.

Seipel, M., Mubarak H., J. Rikoon, S. and Kleiner, A.M. "The Impact of Large-Scale Hog Confinement Facility Sitings on Rural Property Values," Department of Rural Sociology, University of Missouri, Columbia, MO.

Taff, S.J., Tiffany, D. and Weisberd, S 1996. "Measuring the Effects of Feedlots on Residential Property Value in Minnesota: A Report to the Legislature." Staff paper 96-12. Department of Applied Economics, University of Minnesota.

Table 1: Results from OLS, Spatial Error, and Generalized Spatial Two-Stage Least Square Models

Variable	OLS	Spatial Error	GS2SLS
Full Bath	0.1106720**	0.1051930**	0.1106720**
	0.0000000	*0.0000000*	*0.0000000*
Half Bath	0.0692510**	0.0665850**	0.0692510**
	0.0000030	*0.0000160*	*0.0000080*
Square Feet (log)	0.0944720**	0.0995580**	0.0944720**
	0.0034720	*0.0005770*	*0.0014840*
Heat (dummy)	0.0808320**	0.0819500**	0.0808320**
	0.0034150	*0.0025070*	*0.0029080*
Fireplace (dummy)	0.1389390**	0.1386950**	0.1389390**
	0.0000000	*0.0000000*	*0.0000000*
House Age	-0.0437700**	-0.0437400**	0.0437700
	0.0000000	*0.0000000*	*0.0000000*
Type of Story Building	0.0374830**	0.0360670	0.0374830
	0.0072310	*0.0276920*	*0.0219590*
Building Grade	0.0543100**	0.0513900	0.0543100
	0.0000180	*0.0004380*	*0.0001970*
Building Condition	0.0818230**	0.0799640	0.0818230
	0.0000000	*0.0000000*	*0.0000000*
Lot size (log)	0.0329920**	0.0288240	0.0329920
	0.0052890	*0.0110830*	*0.0020690*
Out Building Replacement value	0.4373840**	0.4303820	0.4373840
	0.0000000	*0.0000000*	*0.0000000*
Population Density (log)	0.017126**	0.018761**	0.015875**
	0.0124610	*0.0291240*	*0.0410710*
Dist to Fort Wayne (log)	0.1553130	0.138981**	0.0851600
	0.1665880	*0.0020950*	*0.4790990*
Coldwater school district	0.374595**	0.394541**	0.369651**
	0.0070850	*0.0153800*	*0.0182890*
Celina City school district	0.1113610	0.1260650	0.0926610
	0.3486440	*0.3612700*	*0.4866320*
Other school district	0.0755430	0.0453160	0.0278110
	0.4991520	*0.7278800*	*0.8240910*
Celina (dummy)	0.0125170	0.0230620	0.0162080
	0.6250060	*0.4661920*	*0.5896200*
Montezuma (dummy)	-0.0829900	-0.1074100	-0.0780900
	0.2069390	*0.2038120*	*0.3074350*
Marion Township	0.0642090	0.1283220	0.1791580
	0.6432340	*0.4240270*	*0.2139880*
Franklin Township	-0.0847600	-0.0464500	0.0133290
	0.2082330	*0.4892850*	*0.7901390*
Butler Township	0.32504**	-0.28229**	0.22839**
	0.0001450	*0.0072830*	*0.0094030*
Hopewell Township	-0.0402400	0.0104620	0.0209950
	0.5647690	*0.8949880*	*0.8029950*
Distance to the nearest farm	-0.00416**	-0.0026000	0.019489†
	0.0382340	*0.2854740*	*0.0555230*
Weighted animal density (500 ft)	-0.0000060	0.0000260	0.186981†
	0.9496910	*0.8057390*	*0.0537810*
ρ	N.A.	0.31103**	0.187817**
		0	*0*

Notes: Dependent variable is log(Price). N = 3,476. *p*-values are in italics. †, * and ** indicate significance at the 10, 5, and 1% levels, respectively.