Bayesians in Space: Using Bayesian Methods to Inform Choice of Spatial Weights Matrix in Hedonic Property Analyses[*]

Julie M. Mueller and *John B. Loomis*[+]

ABSTRACT. The choice of weights is a non-nested problem in most applied spatial econometric models. Despite numerous recent advances in spatial econometrics, the choice of spatial weights remains exogenously determined by the researcher in empirical applications. Bayesian techniques provide statistical evidence regarding the simultaneous choice of model specification and spatial weights matrices by using posterior probabilities. This paper demonstrates the Bayesian estimation approach in a spatial hedonic property model estimating the impacts of repeated wildfires on house prices in Southern California. We find that improper choice of spatial model and weights can result in up to 5 percent difference in estimated coefficients and in our case study up to a $15 Million difference in total benefits of reducing wildfires in Los Angeles County.

Key Words: *Spatial Hedonic Models, Wildfires, Bayesian Estimation*

JEL Classification Codes: *Q17, R18, C3*

1. INTRODUCTION

Evaluating the effects of residential area environmental amenities, land uses, and hazards on society has frequently involved using differences in house prices to reveal the marginal benefits or costs to households. The results of these hedonic property estimations are often incorporated into benefit-cost analyses with policy implications. Thus, it is essential for policy-makers to have accurate measurements of estimated implicit prices from hedonic property models. Since hedonic property models are based on the selling prices of homes within the same geographical area, the models are spatial in nature. Proper specification of the spatial models is essential for unbiased and efficient parameter estimates.

Testing for and modeling spatial dependence involves use of a spatial weights matrix. At present, the spatial weights matrix is exogenously chosen by the researcher and implemented prior to estimation of the models. Despite increased use of spatially corrected models and importance of the spatial weights in model selection, little statistical guidance is available to aid the researcher in choosing the spatial weights matrix when estimating spatial models (Anselin, 2002). In addition, testing for the presence of spatial dependence and determining the appropriate

spatial model specification is conditioned on the researcher's exogenous specification of a spatial weights matrix.

When comparing results from Maximum Likelihood estimations to those from General Method of Moments, Bell and Bockstael (2000) found that coefficient estimates from a spatial hedonic property model were more sensitive to the choice of weighting matrix than to the method of estimation. They compared 10 coefficients estimated with 6 different weighting matrices and calculated a percent difference in coefficients equal to the difference in the estimates divided by the mean of the two estimates. The average percent difference in coefficient estimates was 201 percent, with some estimations differing by as much as 1,000 percent. Given that the estimated coefficients in hedonic property models are used to inform policies, the large variation in estimated coefficients due to the researcher's choice of spatial weights matrix could result in major differences in the policy implications of estimated implicit prices. Bell and Bockstael note that with Maximum Likelihood and General Method of Moments estimation techniques "nested tests of alternative spatial weight matrices are computationally infeasible for large datasets, and non-nested tests have not proved particularly helpful." In contrast, LeSage and Pace (2010) find that spatial weights matrices matter little when effect estimates are calculated correctly. We use posterior probabilities to choose both weights and specification, and our data supports the claim of LeSage and Pace—we find only small differences in effects due to spatial weights.

The spatial weights matrix is an $n \times n$ array representing neighbor relationships between all observations in a given dataset. Thus, when choosing the weights, the researcher is essentially making $(n - 1) \times n$ choices about neighbor relationships (one decision for each observation in the sample relative to all other observations) and if using traditional ML techniques, all neighbor decisions must be made before testing for spatial dependence and choosing model specification. Traditionally, researchers will choose the spatial weights matrix that maximizes the value of Moran's I statistic as a best practice approach. In this paper, we offer an alternative method for choosing spatial weights. Bayesian estimation techniques allow comparison of posterior probabilities for the non-nested problem of simultaneously choosing spatial weights and spatial model specification. Testing whether an environmental feature has a statistically significant effect on house prices requires precise coefficient estimates. The magnitude of the coefficients determines the magnitude of implicit prices, thus obtaining efficient and unbiased estimates is of direct importance to policy applications. A subset of possible spatial weights is chosen for demonstrating our approach, and these weights are typical for hedonic property analyses. Our approach could be used to choose among any number of potential weights that might be appropriate for the particular spatial model being estimated. Therefore, the use of Bayesian techniques is worth investigating. To the authors' knowledge, this is the first application of Bayesian methods to simultaneously inform model choice and spatial weights matrix choice in a hedonic property model.

2. SPATIAL HEDONIC PROPERTY MODELS

The hedonic property method is well established as a revealed preference method for obtaining non-market values of environmental amenities or hazards such as air quality improvements (Kim et al., 2003), water quality (Leggett and Bockstael, 2000) and nuclear waste transport (Gawande and Jenkins-Smith, 2001). Rosen (1974) first proposed the detailed theoretical construct for the hedonic property model based on the proposition that identical

houses in otherwise similar neighborhoods will have different prices if the neighborhoods have different levels of an environmental amenity. In order to determine the marginal implicit price of an environmental amenity using a hedonic property model, it is thus necessary to control for other characteristics that determine house price, such as structural characteristics, neighborhood demographics, and housing market trends. See Taylor (2002) and Palmquist (1991) for a comprehensive discussion of the theoretical aspects of hedonic property models.

Although past hedonic property models have incorporated several distance measures to amenities or hazards as explanatory variables, failure to properly model and account for spatial dependence causes econometric problems that may result in biased and inefficient coefficient estimates (Brasington and Hite, 2005). Following Kim, Phipps, and Anselin (2003) if data exhibit a spatial lag process, a Spatial Autoregressive (SAR) model is appropriate. A SAR model uses Maximum Likelihood (ML) techniques to estimate

$$(1) \; \mathbf{y} \; = \; \rho \mathbf{W} \mathbf{y} + \mathbf{X} \boldsymbol{\beta} \; + \boldsymbol{\varepsilon}$$

with $\mathbf{y}$ as the dependent variable where $\mathbf{X}$ is a matrix of independent variables including a constant term, $\mathbf{W}$ is the n by n weighting matrix, $\boldsymbol{\beta}$ is a vector of estimated coefficients, $\boldsymbol{\varepsilon} \sim$ MVN($0, \sigma^2 \mathbf{I}_n$), ρ is the spatial autoregressive parameter and $\mathbf{I}_n$ is an $n \times n$ identity matrix.

Spatially correlated errors are likely to occur when measurement error is related to location (Anselin and Bera, 1998). In a hedonic property model, if some neighborhood effects are not captured in the demographic characteristics, OLS will result in spatially correlated errors and hence be inefficient. The Spatial Error Model (SEM) is as follows:

$$(2a) \; \mathbf{y} \; = \; X \boldsymbol{\beta} + \mathbf{u}$$

$$(2b) \; \mathbf{u} \; = \; \lambda W \mathbf{u} + \boldsymbol{\varepsilon}$$

Here, λ is a coefficient on the spatially correlated errors.

A more general spatial specification is that of the Spatial Durbin Model (LeSage and Pace, 2009). The SDM allows for a spatially lagged dependent variable as well as spatially lagged explanatory variables. A Spatial Durbin Model is applicable when uncertainty about the nature of the spatial dependence exists. Because hedonic property models provide the possibility of having spatial dependence in both the explanatory variables and the dependent variable, a SDM is a reasonable model to explore. The SDM is specified as:

$$(3) \; \mathbf{y} \; = \; \rho \mathbf{W} \mathbf{y} + \mathbf{X} \boldsymbol{\beta} + \mathbf{W} \mathbf{X} \boldsymbol{\gamma} + \boldsymbol{\varepsilon}$$

where $\boldsymbol{\gamma}$ represents the estimated coefficients on the spatially weighted explanatory variables.

3. SPATIAL WEIGHTS

As seen in Equations (1)–(3), correcting for spatial dependence involves the use of a spatial weights matrix. The weights matrix models the "neighbor" relationship within the observations of the dependent variables. Weights can be based on contiguity or distance. The nature of most hedonic property data dictates spatial weights based on distance because house sale data do not provide information on parcel borders and it would be unlikely to have sales of adjacent houses in the same time period. Intuitively, each of our houses is considered a point in space. $\mathbf{W}$ is an $n \times n$ weights matrix with zeros on the diagonal. Wherever there is a nonzero element in the weights matrix, we consider two parcels "neighbors." If observations i and j are

not considered neighbors, w_{ij}=0. MATLAB[1] was used to generate four weighting matrices to be used in our estimations—three are based on nearest neighbor rule and one on Euclidean distances between observations.

For purposes of calculating a distance matrix, let the distance between two parcels be represented by $d_{i,j}$. With a distance matrix, we consider parcels "neighbors" if they are located within the minimum Euclidean distance such that every parcel has at least one neighbor. Suppose this distance equals *b*. For each parcel *i*, let *z* represent the number of neighbors within Euclidean distance *b* from parcel *i*. For our distance-based weights matrix, $w_{ij} = d_{ij}$, $w_{ij} = d_{ij}$ if $d_{i,j} < b$.

With Bayesian estimation techniques for spatial econometrics, the choice of spatial weights matrix and spatial model are often made jointly, and can be analyzed using posterior probabilities. The use of Bayesian estimation to test hypotheses is relatively common in statistics, although only relatively recently used in applied econometrics. Kass and Raftery (1995) exemplify several situations in which they use Bayesian estimation methods. In particular, they emphasize the idea that post-estimation model choice when using posterior probabilities does not require alternative methods to be nested. Hepple (2004) clearly outlines a brief application of Bayesian estimation methods as do LeSage and Pace (2009).

4. BAYESIAN ESTIMATION

Bayesian estimation is based on inference using Bayes' Theorem. Let **θ** represent a vector of *k* unobserved parameters, and suppose our **DATA** is a matrix of *n* observations dependent upon these unobserved parameters. Bayes's Theorem is given in Equation (4):

(4) $f(\theta|\textbf{DATA}) = \frac{f(\textbf{DATA}|\theta)f(\theta)}{f(\textbf{DATA})}$

Where $f(\theta|\textbf{DATA})$ is the posterior distribution of θ, $f(\textbf{DATA}|\theta)$ is the sampling distribution for **θ** (proportional to the likelihood), $f(\theta)$ is the prior density of the data, and $f(\textbf{DATA})$ is the "marginal likelihood" or In other words, Bayes' Theorem states:

$Posterior \propto Likelihood \times Prior$ where $\propto$ denotes "is proportional to." The posterior distribution contains all the information needed to make inferences about individual parameters $\theta_1 \dots \theta_k$. $f(\theta)$ represents our beliefs about the parameter **θ** before any data are observed. This "prior" belief about θ illuminates the subjective philosophy of Bayesian estimation. However, it also allows the researcher to incorporate valuable prior knowledge about the estimated parameters. $f(\textbf{DATA})$, or the "marginal likelihood" is essential in Bayesian model comparisons. The mathematically intensive part of Bayesian estimation is finding a solution to this open integral.

We use Bayesian estimation with Markov Chain Monte Carlo (MCMC) techniques (i.e. the Gibbs sampler and the Metropolis-Hasting Algorithm) to find the marginal posterior distributions for each parameter. Gibbs sampling is a relatively basic MCMC method that can be used when sampling from a multivariate posterior is not feasible, but sampling from the conditional distributions for each parameter is feasible (Lynch, 2007). When this is not feasible (as in the case of the spatial autocorrelation parameter) we utilize the Metropolis-Hastings

---

[1] Code for generating the distance based weights matrices was found at Donald Lacombe's website: http://www.rri.wvu.edu/lacombe/matlab.html.

Algorithm. A thorough explanation of the mathematics behind the Gibbs sampler is beyond the scope of this paper. Readers are referred to Lacombe (2008) and LeSage and Pace (2009) for a mathematical introduction to MCMC Gibbs sampling methodology in the context of spatial econometrics.

5. BAYESIAN MODEL CHOICE

Following the development of Bayesian model choice by Hepple (2004), suppose we want to compare two models, M_k and M_j using Bayes's Theorem. If so,

$$(5)\ p(\mathbf{M}_k|\mathbf{DATA}) = \frac{p(\mathbf{DATA}|\mathbf{M}_k)}{p(\mathbf{M}_k|\mathbf{DATA}) + p(\mathbf{DATA}|\mathbf{M}_j)} \frac{p(\mathbf{M}_k)}{p(\mathbf{M}_j)}$$

where $p(data|M_k)$ is the marginal likelihood of the data given M_k and $p(M_k)$ is the prior probability of the model M_k The extent to which the data favors one model over the other can be represented by the ratio of their posterior probabilities, or the posterior odds O_{kj}, where:

$$(6)\ O_{kj} = \frac{p(\mathbf{M}_k|\mathbf{DATA})}{p(\mathbf{M}_j|\mathbf{DATA})} = \frac{p(\mathbf{DATA}|\mathbf{M}_k)}{p(\mathbf{DATA}|\mathbf{M}_j)} \frac{p(\mathbf{M}_k)}{p(\mathbf{M}_j)}$$

Thus, the odds ratio for any pair of models, $O_{kj} = \frac{\pi_k}{\pi_j}$, where π_j is the posterior probability of model j. When comparing n possible models, if each of the models have equal prior probabilities, the posterior probability, π_j for model j is given in Equation (7)

$$(7)\ \pi_j = \frac{p(\mathbf{DATA}|\mathbf{M}_j)}{\sum_{i=1}^{n} p(\mathbf{DATA}|\mathbf{M}_i)}$$

and the researcher can compare several models using posterior probabilities. If the researcher believes several models to be equally probable, then post-Bayesian analysis, they should choose the model with the highest posterior probability. In the empirical analysis that follows, we illustrate how the Bayesian analysis allows us to simultaneously identify the appropriate spatial model and spatial weights matrix. We estimate six models using Bayesian analysis and choose the model with the highest posterior probability. We then compare differences in the implicit prices between the Bayesian identified "best" model and the other five model specifications.

6. DATA

We draw upon a rich dataset with repeated wildfires in Southern California. The dependent variable is the house sale price. A log-linear specification allows the marginal effect of each independent variable to vary with the level of the dependent variable, hence the marginal implicit price of an attribute changes as house price varies. This functional form has been shown to be a robust functional form in the face of limitations in the data (Cropper et al., 1988). The general hedonic property model follows the general form:

$$(8)\ P_{it} = f\ (E_{it}, S_i, N_i,)$$

Where P_{it}: Sale amount at decision date t, with sale amount deflated using the annual housing price index for Los Angeles, Orange, and Riverside Counties (1983 Base Year), E_{it}: Environmental variables of interest (e.g., wildfires) for house i at time t, S_i: Structural characteristics of house i, and N_i: Neighborhood demographics for house i. Our data is a cross section of single sales of single family homes. For each observation, we have information on the sale date of the home and its sale price. We used the sale date to determine the dummy variables representing whether a house sold after one or two wildfires.

The following are the independent variables included in our empirical specification:

E_{it}: Environmental and Location Variables:
- (i)   *After One Fire*: An indicator variable that equals one if a house sold after and is located within 1.75 miles of one wildfire[2]
- (ii) *After Two Fires:* An indicator variable that equals one if a house sold after and is located within 1.75 miles of two wildfires
- (iii) *Days Since Most Recent Fire:* Number of days since the most recent wildfire
- (iv) *Distance to USFS Land*: to the edge of the nearest National Forest owned land (meters)
- (v)   *Elevation:* Elevation of the house lot (meters) above sea level

The elevation of a house lot serves as a proxy for vegetation type (higher elevations tend to have more flammable vegetation in southern California). Houses located at higher elevations and nearer to forests have a higher risk of burning from a wildfire.

S_i: Housing Structure Variables:
- (i)   *Square feet*
- (ii) *Number of Bedrooms*
- (iii) *Number of Bathrooms*
- (iv) *Year Built*

Many of the housing characteristics are highly correlated. When all structural characteristics are included, the model shows signs of collinearity problems. *Square feet* is a commonly used explanatory variable in hedonic property models to control for size of housing structure, and therefore is the structural characteristic we decided to include. *Number of Bedrooms* is highly correlated with *Square Feet,* and is therefore dropped in the final specification. We also include *Year Built* as a measure of housing structure quality.

N_i: Neighborhood Demographics Factors:
- (i)   *Median Household Income*: Median household income in census tract (Year 2000 dollars)
- (ii) *Percent with No High School Degree*: Percent of residents in census tract above 18 years old with no high school degree

Other neighborhood characteristics commonly included in hedonic models are school district quality and household income (Taylor, 2002). A direct measure of school district quality is unavailable within our data, so a measure of the percent with no high school degree in a neighborhood is used as a proxy for the relative level of educational attainment in a particular community. Neighborhoods with high percentages of educated people generally have higher quality schools. Median household income is also included as a proxy for neighborhood desirability.

General trends in real house prices may occur over time. To reflect these general market trends in the real house prices, we include a variable that indicates the time period of the sale of the house. We have single sale house price data with the sale date for each observation. We use

---

[2] The authors performed sensitivity analysis of different distance cutoffs and chose the 1.75-mile distance to keep a relatively large sample. Please see Mueller et al. (2009) for a detailed description of this analysis.

the sale date for each observation to generate a single value for the time trend variable and values for the *After Fire* indicator variables. The empirical model specification is as follows: [3]

(9)   *Log (Real Sale Amount) = β₀ + β₁*After One Fire+β₂*After Two Fires +*
       β₃*Days Since Most Recent Fire + β₄*Square Feet + β₅*Year Built +
       β₆*% with no High School Degree + β₇*Median Household Income+
       β₈*Distance to USFS Land + β₉*Elevation + β₁₀*Trend + ε

7. APPLICATION TO SPATIAL HEDONIC PROPERTY MODEL

We use Bayesian estimation methods to estimate several spatial hedonic property models. Equations (1), (2a), and (3) are estimated using the same dependent variable and independent variables, but with each of the four different types of weight matrices—4, 6 and 8 nearest neighbors, and distance. The Bayesian models are estimated using Gibbs sampling with 11,000 iterations, omitting the first 1,000 draws. With our Bayesian model, we use posterior probabilities to inform our choice of spatial models.[4] The prior value for β is 0 with a covariance of 1E+6 from a multivariate normal distribution. The prior distribution of σ is $\Gamma(0,0)$, and the prior distributions of ρ and λ are uniform(-1,1).

7.1. Model Comparison

We estimate and compare posterior probabilities for SEM, SAR, and SDM models using all four weights matrices.[5] Recall from Equation (7) that for Bayesian model comparison, post-estimation the researcher should choose the model with the highest posterior probability. As mentioned, the models need not follow the same specification of spatial dependence or spatial weights in order to use posterior probabilities to perform model comparisons. The posterior probabilities are reported in Table 1.[6] All models had plausible and consistent signs on the variables of interest. When estimating all twelve Bayesian spatial models, the posterior probability of the SEM model with distance-based weights is the highest. The sum of the posterior probabilities sums to one, with the probability on the SEM model with distance-based weights very close to one, and the other probabilities quite small. Therefore, we have strong evidence to conclude that the SEM with distance weights best represents the true data-generating process.

---

[3] An anonymous reviewer pointed out that all hedonic property equations such as Equation (9) suffer from endogenous explanatory variables. Besides econometric solutions, another solution to endogeneity is selecting a case study that involves a "natural experiment" in which a treatment variable is randomly assigned. Our study applies this natural experiment approach. In particular, whether a house was close to zero, one or two fires, and how many days passed between forest fires are exogenous events. We also note that the purpose of this paper is not to derive specific implicit prices to be used for benefit-cost analysis or policy purposes. Rather we wish to compare implicit prices among models that use exogenously chosen weights and weights chosen via a Bayesian approach. To the extent there is endogeneity it should equally affect all the implicit prices, but their comparative differences by spatial weights would likely not be changed.

[4] Bayesian model averaging can also be used to inform model choice with a large number of possible models and posterior probabilities to estimate. Detailed analysis of Bayesian averaging can be found in LeSage and Parent (2007). Because of the high posterior probability estimated on the SEM model with distance-based weights, and the computation time involved with Bayesian model averaging, the authors chose not to apply this methodology, and point to it as a valuable avenue for future research.

[5] We use the "model_probs" code from the Jim LeSage's spatial econometrics toolbox. The toolbox and all supporting documentation are available at www.spatial-econometrics.com.

[6] Full results from all models are available from the authors.

Table 1: Posterior Probabilities of Estimated Models

	4NN	6NN	8NN	Distance
SEM	5.68E-26	2.87E-23	1.94E-21	1.00E+00
SAR	5.23E-33	6.00E-33	5.37E-33	2.51E-16
SDM	6.55E-48	5.18E-45	2.36E-42	5.81E-24

7.2. Estimated Implicit Prices

The parameter estimates from our spatial models contain information about the impact of a wildfire on house prices for individual homes, yet they also contain information about the impact of a wildfire on the house prices in the neighborhood. In the SAR and SDM models, changes in the wildfire variables will result in both a direct impact on the individual house, and an indirect impact that is the average impact on the neighborhood as defined by the spatial weights $\mathbf{W}$. The estimated implicit price from a semi-log hedonic SEM is simply $\hat{\beta} \times \bar{y}$ because the SEM model is a special case of the Spatial Durbin model where the spatial autoregressive parameter, $\rho=0$. In contrast, the impact of a wildfire in our SAR and SDM models will have a spatial multiplier effect on y that can be expressed using a multiplier term $S_r(\mathbf{W})$, where $S_r(\mathbf{W})$ takes different forms depending upon the spatial model (LeSage and Pace, 2009). An average of the diagonal of the matrix $S_r(\mathbf{W})$ represents an average direct impact. The indirect impacts work through the spatial autoregressive parameter, ρ, and allow spillovers from a change in one explanatory variable to impact the value of an individual observation and its neighbors through the spatial weights. As discussed in Section 7.1, our model comparison leads us to the conclusion that the data-generating process is best represented by the SEM.[7] The direct impacts of variable r can be represented by

(10) $\bar{M}(r)_{direct} = n^{-1}tr(S_r(\mathbf{W}))$.

It should be emphasized that there are no indirect impacts in the SEM model because there is no spatial autoregressive parameter.

For the SAR model, $S_r(\mathbf{W}) = (\mathbf{I}_n - \rho\mathbf{W})^{-1}(\mathbf{I}_b\beta_r)$ and for the SDM, $S_r(\mathbf{W}) = (\mathbf{I}_n - \rho\mathbf{W})^{-1}(\mathbf{I}_b\beta_r + \mathbf{W}\theta_r)$. We obtain estimates of direct impacts for the After Fire variables during the MCMC sampling process of our Bayesian estimation.

7.3 Policy Implications

As shown in Table 2, the estimated coefficient on *After One Fire* is -0.1767 in the SEM with distance weights. The mean house sales price (in 1983 dollars) is \$152,967. Using the Housing Price Index for California[8] this equates to approximately \$570,000 in today's dollars. In today's dollars, the SEM estimation using the distance matrix indicates that house prices will drop approximately \$100,537 after one fire. In today's dollars, the SEM estimation using the distance matrix indicates that house prices will drop approximately \$106,093 after one fire.

---

[7] Indirect and total impacts are available from the authors upon request. However, it is important to note that the total effects are not significant in the SDM models. We choose to use direct impacts to have a more meaningful comparison of estimated implicit prices from different model specifications.

[8] http://research.stlouisfed.org/fred2/series/CASTHPI

Table 2: SEM Estimation Results Using Distance-based Weights

Variable (*n*=1,762)	Coefficient	Std Deviation	*p*-level
Constant	4.877915	1.021533	<0.0001
After One Fire	-0.176369	0.019336	<0.0001
After Two Fires	-0.109692	0.024262	<0.0001
Days Since Most Recent Fire	0.000032	0.000011	0.0020
Square Feet	0.000310	0.000011	<0.0001
Year Built	0.003531	0.000536	<0.0001
No High School Degree	-0.006190	0.000640	<0.0001
Median Household Income	0.000000	0.000000	0.1348
Elevation (meters)	-0.000303	0.000048	<0.0001
Trend	0.000018	0.000007	0.0053
lambda	0.740751	0.045747	<0.0001

Because the SEM model does not suffer from bias, we also compare our results from the SEM model to simple OLS. OLS predicts a $11,896 larger price drop relative to the SEM with distance based weights. [9]

To calculate the total benefits of fire reduction, we need to consider the study area. We have approximately 54,000 single family residences in our study area. The *difference* in benefits from the two spatial models of reducing a first forest fire to 54,000 single-family residences amounts to a capital value or present value of over $300 million. Annualizing the $300 million difference in capital value at 5 percent to make it equivalent to annual fire-fighting budgets yields $15 million a year. When small differences in estimated coefficients may result in large differences in total benefits, it is vital to the researcher to properly specify the spatial weights and the spatial specification.

8. SUMMARY AND CONCLUSIONS

In this paper we address an important issue in spatial hedonic property models—how to simultaneously choose the spatial weights and corresponding spatial model specification. To provide one answer to this question we adapted Bayesian estimation techniques to inform choice of spatial weights matrix in a hedonic property model. Bayesian techniques can provide statistical evidence regarding the simultaneous choice of spatial model specification and spatial weights matrices in spatial econometrics, while more commonly applied ML techniques do not allow non-nested model comparison. Given our data, with improper model specification and weights matrix, estimated implicit prices were found to vary by 5 percent. While this is not a large difference in implicit prices with our small impact radius of 1.75 miles, with a larger impact area associated with wide reaching environmental effect, choice of spatial weights might have a larger impact on implicit prices. Nonetheless, from our empirical analysis it is evident that both the choice of the spatial model and spatial weights matrix does affect the magnitude of the estimated implicit prices. Having a statistical method that can endogenously and simultaneously determine both the spatial weights and spatial model is a useful advance in

---

[9] Detailed OLS results are available from the authors upon request.

conducting spatial econometric estimation of hedonic property models and their corresponding implicit prices used in benefit-cost analysis.

REFERENCES

Anselin, Luc. (2002) "Under the Hood: Issues in the Specification and Interpretation of Spatial Regression Models," *Agricultural Economics*, 27, 247–267.

Anselin, Luc and Anil Bera. (1998) "Spatial Dependence in Linear Regression Models," In Aman Ullah, David E. A. Giles (Eds.). *Handbook of Applied Economic Statistics*. Marcel Drekker: New York, 237–289.

Anselin, Luc, Anil Bera, Raymond Florax, and Mann J. Yoon. (1996) "Simple Diagnostic Tests for Spatial Dependence," *Regional Science and Urban Economics*, 26, 77–104.

Bell, Kathleen and Nancy Bockstael. (2000) "Applying the Generalized-Moments Estimation Approach to Spatial Problems Involving Micro-Level Data," *Review of Economics and Statistics*, 82, 72–82.

Brasington, David M. and Diane Hite. (2005) "Demand for Environmental Quality: A Spatial Hedonic Analysis," *Regional Science and Urban Economics,* 35, 57–82.

Cropper, Maureen L., Leland B. Deck, and Kenneth E. McConnell. (1988) "On the Choice of Functional Form for Hedonic Price Functions," *Review of Economics and Statistics,* 70, 668–675.

Gawande, Kishore and Hank Jenkins-Smith. (2001) "Nuclear Waste Transport and Residential Property Values: Estimating the Effects of Perceived Risks," *Journal of Environmental Economics and Management,* 42, 207–233.

Hepple, Leslie W. (2004) "Bayesian Model Choice in Spatial Econometrics," In James P. LeSage and R. Kelley Pace (eds), *Spatial and Spatiotemporal Econometrics* (Advances in Econometrics, Volume 18). Elsevier: Amsterdam, pp. 101–126.

Kass, Robert E. and Adrian E. Raftery. (1995) "Bayes Factors," *Journal of the American Statistical Association*, 90(430), 773–795.

Kim, Chong Won, Tim T. Phipps, and Luc Anselin. (2003) "Measuring the Benefits of Air Quality Improvement: A Spatial Hedonic Approach," *Journal of Environmental Economics and Management* 45, 24–39.

Lacombe, Donald J. (2008) "An Introduction to Bayesian Inference in Spatial Econometrics," Available at SSRN: http://ssrn.com/abstract=1244261, July 24.

Leggett, Christopher G. and Nancy E. Bockstael. (2000) "Evidence of the Effects of Water Quality on Residential Land Prices," *Journal of Environmental Economics and Management,* 39, 121–144.

LeSage, James P. and R. Kelley Pace. (2009*) Introduction to Spatial Econometrics*. CRC Press: Boca Raton, FL.

_____. (2010) The Biggest Myth in Spatial Econometrics, December 1, Available at SSRN: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1725503

LeSage, James P. and Olivier Parent. (2007) "Bayesian Model Averaging for Spatial Econometric Models," *Geographical Analysis* 39, 241–267.

Lynch, Scott M. (2007) *Introduction to Applied Bayesian Statistics and Estimation for Social Scientists*. USA: Springer Science+Business Media LLC.

Mueller, Julie M., John B. Loomis, and Armando González-Cabán. (2009) "Do Repeated Wildfires Change Homebuyers' Demand for Homes in High-Risk Areas? A Hedonic

Analysis of the Short and Long-Term Effects of Repeated Wildfires on House Prices in Southern California," *Journal of Real Estate Finance and Economics,* 38, 155–217

Palmquist, Raymond B. (1991) "Hedonic Methods," In John B. Braden and Charles D. Kolstad (eds), *Measuring the Demand for Environmental Quality*. Amsterdam: Elsevier, pp. 77–119.

Rosen, Sherwin. (1974) "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition," *Journal of Political Economy,* 82, 34–55.

Taylor, Laura O. (2002) "The Hedonic Method," In Patricia A. Champ, Kevin C. Boyle, and Thomas C. Brown (eds), *A Primer on Non-Market Valuation*. Kluwer: Amsterdam, pp. 331–394.