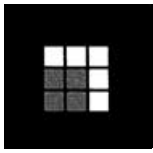
2002 V30 4: pp. 505–532
REAL ESTATE
ECONOMICS



Predicting Spatial Patterns of House Prices Using LPR and Bayesian Smoothing

John M. Clapp,* Hyon-Jung Kim** and Alan E. Gelfand***

This article is motivated by the limited ability of standard hedonic price equations to deal with spatial variation in house prices. Spatial patterns of house prices can be viewed as the sum of many causal factors: Access to the central business district is associated with a house price gradient; access to decentralized employment subcenters causes more localized changes in house prices; in addition, neighborhood amenities (and disamenities) can cause house prices to change rapidly over relatively short distances. Spatial prediction (e.g., for an automated valuation system) requires models that can deal with all of these sources of spatial variation. We propose to accommodate these factors using a standard hedonic framework but incoporating a semiparametric model with structure in the residuals modeled with a partially Bayesian approach. The Bayesian framework enables us to provide complete inference in the form of a posterior distribution for each model parameter. Our model allows prediction at sampled or unsampled locations as well as prediction interval estimates. The nonparametric part of our model allows sufficient flexibility to find substantial spatial variation in house values. The parameters of the kriging model provide further insights into spatial patterns. Out-of-sample mean squared error and related statistics validate the proposed methods and justify their use for spatial prediction of house values.

Hedonic models of house prices are plagued by inadequate specification of spatial price processes. It is known that house prices vary dramatically over space, but this is difficult to model. Dubin (1992) defines the problem: "Most hedonic estimations show few significant coefficients on the neighborhood and accessibility variables" (p. 433). It follows that there are many unmeasured spatial variables omitted from the hedonic equation. Dubin summarizes literature attempting to deal with the problem through the early 1990s.[1]

*University of Connecticut, Storrs, CT 06269 or johnc@sba.uconn.edu.
**University of Oulu, FIN-90014 Finland or Hyon-Jung.Kim@oulu.fi.
***University of Connecticut, Storrs, CT 06269 or alan@stat.uconn.edu.

[1] The motivation for this literature usually emphasizes mortgage-lending decisions, which require accurate predictions (*e.g.*, correct valuations) of house prices (see, *e.g.*, Can and Anselin 1998).

We view spatial variation in house prices as the sum of many causal factors. Two categories of causal factors may be modeled in the mean function:[2] (1) House prices change relatively slowly over long distances due to access to major points of interest such as the downtown (CBD), an airport, or an ocean beach; (2) Employment subcenters, school districts, higher elevations, a water view and the like require more proximity to influence house prices. These variables cause spatial patterns that work over shorter distances than the first category.

Mills and Hamilton (1984, especially Chapter 6) summarize "new urban economics" theory for category 1 spatial patterns; these are the factors giving shape to the skyline and presumed to cause major changes in house prices.[3] According to this theory, we can expect distances measured to the CBD and other major points of interest to influence house values over the entire metropolitan area.

Anas, Arnott and Small (1998) emphasize the growing complexity of urban areas in their summary of literature dealing with both categories of spatial variables. Technological change and deregulated global competition drive subcenter development. Many subcenters containing substantial portions of metropolitan employment give rise to cross-commuting patterns. It is reasonable to expect that such a complex pattern will reduce the distance over which any subcenter can influence the mean function of house prices.

A third category of causal variables (3) includes physical and political variables that cause neighborhood house prices to be spatially autocorrelated. For example, subdivision approvals lead to neighborhoods being developed at about the same time and with similar building characteristics. Can and Anselin (1998) and Gillen, Thibodeau and Wachter (2001) discuss a number of variables omitted from categories 2 and 3 that can be modeled with spatial autocorrelation.

Our approach is to model omitted spatial variables in the mean function as well as in the disturbances by combining spatial autocorrelation with a nonparametric technique. Thus, we attempt to reduce the degree of spatial autocorrelation, and to shorten its range, by using the nonparametric technique to model omitted variables in the mean of the hedonic function. By doing so, insight is gained into spatial structure without the need to measure all variables causing spatial variation in house prices.

---

[2] The mean function is the $f(x)$ in a typical regression equation: $y = f(x) + e$.

[3] Mills and Hamilton (1984) have a brief discussion of the effects of employment subcenters on land value gradients.

Our approach is based on Host (1999), who models spatial variation in air quality with local polynomial regressions (LPR) in the mean function and spatial autocorrelation in the disturbances. Unlike Host, we use a semiparametric framework: the local regression model (LRM). The LRM uses a linear function to estimate the subset of coefficients on building characteristics and measurable spatial variables, X. Unmeasured spatial variables produce house prices that are a highly nonlinear function of latitude and longitude, W:

$$\ln SP = X\beta_1 + f(W) + \varepsilon,$$

where ln SP is the natural logarithm of sales price. The nonlinear function, $f(W)$, is estimated with LPR methods. The parametric and nonparametric parts of the mean function are jointly estimated. Our paper is based on the idea that LPR gives sufficient flexibility to find substantial spatial variation in the value of a standard house, and that the smoothness assumption required by LPR makes it particularly suitable for finding the effects of the subcenters documented by Anas, Arnott and Small (1998).

Parametric methods have been used to estimate house value surfaces. Jackson (1979) and Clapp, Rodriguez and Pace (2000) used a polynomial expansion in latitude and longitude for the $f(W)$ term.[4] Jackson used terms up to the fourth power, dropping terms that were not statistically significant. In theory, a high degree polynomial could fit an arbitrarily flexible value surface. However, collinearity and loss of degrees of freedom may prevent use of a polynomial of sufficiently high degree. Empirical tests presented below are designed to compare this approach with the LRM.

A Bayesian spatial approach is used to further model unmeasured spatial variables in the error term.[5] Bayesian spatial modeling allows more flexibility for inference than classical approaches (*e.g.*, Gelfand *et al.* 2002 and Agarwal *et al.* 2001). The classical asymptotic standard errors associated with the estimated parameters of the spatial process are often inappropriate, especially for a small number of observations. Instead, the Bayesian framework enables us to provide complete inference in the form of a posterior distribution of each model parameter as well as the predictive distribution for out-of-sample observations.

---

[4] This "trend surface analysis" was originally developed in the 1950s. See Agterberg, Gaile and Willmott (1984) for a literature review.

[5] We will follow Host's terminology. *High frequency* spatial patterns refer to spatial relationships in the error term. Spatial autocorrelation is thought to be particularly suitable for spatial patterns that vary over short distances. We will use the terms *measured spatial variables* and *unmeasured spatial variables* to refer to the first two terms of the LRM.

The approach used here is partially Bayesian, as we attempt to quantify unmeasured spatial component of residuals after using ordinary least squares (OLS) and LPR. We anticipate that properties close to each other in geographic distance will show more similar selling prices than those apart. Hence, we expect stationary spatially correlated house prices after adjusting for structure in the mean function.

Stationary spatial processes are usually characterized with three parameters called range, sill, and nugget (Cressie 1993). The range is defined as the distance over which spatial correlation between observations dies off. The sill parameter describes the variability of the process; the partial sill describes spatial variability. The nugget for our application explains a variety of micromarket variability, which leads to different selling prices for two identical houses at essentially the same location. The results section shows how these parameters aid in evaluating omitted spatial variables.

We employ a standard geostatistical modeling specification for LRM residuals since each transaction in our data has an associated geocoded location. Prediction at new locations naturally follows from the predictive distribution of the specified model with the samples of posterior distributions of model parameters. Then we are equipped to provide predictive point estimation for cross validation as well as the interval estimates. The Bayesian fitting is implemented using Gibbs sampling (Gelfand and Smith 1990) with a Metropolis–Hastings sampler.

We will develop the following applications of our model.

It is useful to model unmeasured spatial variables in the mean function because this provides for consistent estimators of the parameters on the OLS part of the model. Moreover, the LRM allows measurement of spatial patterns of house prices: for example, a map of iso-value lines can be constructed in order to find unobserved subcenters.

We show that spatial autocorrelation in house price equation residuals can be reduced or possibly eliminated by applying the LRM.

The model improves out-of-sample prediction of house prices, particularly in neighborhoods with high curvature in the value gradient. These predictions have obvious applications to automated valuation systems.

The next two sections develop methodology; the exposition is aimed at readers who are not familiar with local polynomial regressions or Bayesian kriging. The fourth section presents the data and the fifth discusses empirical results. The final section contains conclusions.

The Local Regression Model

The purpose of this section is to present the details of the LRM method and to compare it to existing methods for house value analysis.

The parametric part of our models starts with a standard hedonic function: the regression of log of sales price ($\ln SP_i$) on a vector of housing characteristics (X_i) and a physical location vector (W_i) for a given house, i, at a given point in time.

$$Y_i = \ln SP_i = X_i\beta_1 + W_i\beta_2 + \varepsilon_i, \tag{1}$$

where ε_i, an i.i.d. noise term, is assumed to be normally distributed for the purposes of OLS hypothesis testing. The extensions of OLS proposed here allow relaxation of this assumption. Implicit prices of building characteristics are contained in part of the vector β_1.

We put standard spatial variables into the vector X_i. Thus, in our model, X_i includes distances to points of interest and town dummy variables. A price surface based solely on these variables would exhibit regular or infrequent change, as in monocentric city or Tiebout capitalization models.

Unmeasured spatial price variation is modeled with the W_i vector. Models include trend surface analysis (*i.e.*, a polynomial of degree p in latitude and longitude), dummies for ZIP codes or Census tracts and local polynomial smoothing regressions conditioned on latitude and longitude.

High frequency spatial price variation is modeled with the ε_i vector using the Bayesian spatial process specification discussed in the introduction. The three categories of spatial price variation are added together to predict house prices at a given out-of-sample location for cross-validation. We will show that this approach improves understanding of the variables influencing house prices and that spatial prediction can be improved.

The LRM

The LRM is a semiparametric approach using a linear model to estimate some of the parameters in Equation (1). For example, the parameters on all the housing characteristics in the matrix X can be estimated initially by OLS. Then the residuals from this regression can be fit with Bayesian and/or LPR models.

The LRM begins by estimating Equation (1) with OLS. Then, partial residuals are calculated using Equation (2):

$$\hat{\eta}^0 = Y - \mathrm{X}\hat{\beta}_1^0, \tag{2}$$

where $Y = \ln SP$, $\hat{\beta}_1^0$ is an estimated vector of implicit prices of property characteristics, and $\hat{\eta}^0$ is the partial residual from Equation (1); the subscript i has been suppressed on all variables. The superscript will index the iteration number. The nonparametric (LPR) part of the model is:

$$\tilde{\eta}^0 \equiv smooth(\hat{\eta}^0 \mid \mathrm{W}), \tag{3}$$

where W has two columns, latitude and longitude. The function that produces the smooth will be explained in the next section. The estimated disturbance for iteration 0 is

$$\hat{\xi}^0 = \hat{\eta}^0 - \tilde{\eta}^0. \tag{4}$$

The mean zero error term, ξ, results from negotiation between heterogeneous buyers and sellers and from spatial autocorrelation.

To fully implement this method, one can iterate back and forth between the parametric and the nonparametric parts of the model until parameters converge, assuring orthogonal residuals in the two parts of the model. Conceptually, this works by dividing Y into two parts. $\hat{Y}_1$, the estimated OLS part of the model plus disturbance, is defined as follows at the beginning of iteration 1:

$$\hat{Y}_1^1 \equiv X\hat{\beta}_1^0 + \hat{\xi}^0. \tag{5}$$

Estimate the following OLS model to find new estimates for β_1:

$$\hat{Y}_1^1 = X\beta_1^1 + \xi. \tag{6}$$

The second part of the original Y values is given by η, defined as follows for iteration 1:

$$\hat{\eta}^1 \equiv Y - X\hat{\beta}_1^1. \tag{7}$$

Thus, Y is decomposed as: $Y = Y_1 + \eta = X\beta_1 + \tilde{\eta} + \xi$.

The iterative process continues until none of the $\hat{\beta}$s change by more than 5% in absolute value. By construction, we have achieved orthogonality between the OLS estimates and the LPR smoother. Thus, Robinson's (1988) concern has been addressed. The $\hat{\beta}$ vector is a consistent estimator of β, and the LPR part of the model is not statistically related to the OLS estimators.[6]

[6] Moyeed and Diggle (1994) prove consistency for the parameters of a semiparametric model, fitted with the backfitting method. Their Theorem I shows that bias and variance of the estimator tend to zero. Simulation studies show reasonable small sample-properties

The LPR Smoother

The LPR smoother can be introduced by imagining that a number, q, of houses trade at or near a given point in space, denoted by the fixed two-element vector w_0, the latitude and longitude at that point. Assume that the true values of β_1 are known exogenously. We calculate location value plus a disturbance term as the vector $\tilde{\eta} + \xi = Y - X\beta_1$.

Nonparametric smoothing produces a local average by down-weighting observations that are more distant from the fixed point. Thus, the nonparametric smooth of $\tilde{\eta} + \xi$ at that point is defined as

$$\tilde{\eta}(w_0) = \sum_{i=1}^{q} \frac{K_h\,(W_i - w_0)\,(\hat{\eta}_i)}{\sum_{i=1}^{q} K_h\,(W_i - w_0)}, \tag{8}$$

where the weighting function, $K_h(\cdot)$, is an inverse function of distance and h is the bandwidth, discussed below. Here, $\hat{\eta}_i$ is estimated after all iterations in Equations (2) through (7) and the parameters converge.

The average in Equation (8) with $\hat{\hat{\xi}}_i$ replacing $\hat{\eta}_i$, that is, the average error term, will tend to zero as the sample size gets large. Thus, we will have found a point on the value surface. When these estimates are replicated at every point in $w = \{w_i\}$, a rectangular grid covering the study area, then we would have estimated the entire value surface.

Equation (8) gives a weighted average of η in the local neighborhood around (w_0). This Nadaraya–Watson estimator gives less weight to more distant points. We adopt a product kernel:[7]

$$K_h(W_i - w_0) = K_{hw_1}(W_{i1} - w_{01})K_{hw_2}\,(W_{i2} - w_{02})\,. \tag{9}$$

The subscripts 1 and 2 on the W variables indicate latitude and longitude. The kernels $K_{hw_1}(\cdot)$ and $K_{hw_2}(\cdot)$ can be thought of as probability density functions

---

for these estimators (see Zeger and Diggle 1994). The objective function in this literature is AMISE (asymptotic mean integrated squared error). We leave technical details to the highly accessible books by Fan and Gijbels (1996), Hastie, Tibshirani and Friedman (2001) and Wand and Jones (1995).

[7] Product kernels are standard in the literature; a different bandwidth, h, is used for each dimension. Bandwidth selection is a trade-off between high variance (bandwidth is too small) and high bias (bandwidth is too large). This article uses a cross-validation method for bandwidth selection (see Wand and Jones 1995, Chapter 4). Our locally adaptive bandwidths increase, if necessary, until at least 20 observations are within one bandwidth at each point; experiments with 10 and 30 observations did not change the results substantially.

with standard deviations ("bandwidths") equal to hw_1 and hw_2; in fact, a normal pdf with range truncated at one standard deviation is used as the kernel function.[8]

Equation (8) is a special case of local polynomial regression (LPR). We have a specific point in space, $w_0 = (w_{01}, w_{02})$, and at locations $W_i = (W_{1i}, W_{2i})$, we have unobserved location values, η_i. Local polynomial regression now takes the form of Equation (10):

$$\eta_i = \gamma_0 + (\mathrm{W}_i - \mathrm{w}_0)\gamma_1 + (\mathrm{W}_i - \mathrm{w}_0)^2\gamma_2 + \cdots + (\mathrm{W}_i - \mathrm{w}_0)^p\gamma_p + \xi_i. \qquad (10)$$

Here and in Equation (11) the exponents are taken element by element: $(W_i - w_0)^p\gamma_p = (W_{1i} - w_{10})^p\gamma_{1p} + (W_{2i} - w_{20})^p\gamma_{2p}$. Here, the γ_j ($j = 1, \ldots, p$) are column vectors with number of elements equal to the columns of W_i; γ_0 is a scalar. Note that when W_i equal w_0 then Equation (10) reduces to γ_0, the parameter of interest. Thus, LPR fits a surface to the η-values conditional on the values of w given by w_0. For example, w is a rectangular grid of equally spaced latitude and longitude points that span the data; the level of η is estimated conditional on each knot of the grid.

Kernel weights are applied when estimating Equation (10):

$$\mathrm{Min}(\gamma) \sum_{i=1}^{n} \left\{\eta_i - \gamma_0 - \cdots - (W_i - w_0)^p\gamma_p\right\}^2 K_h(W_i - w_0), \qquad (11)$$

where the weights are applied to each of the variables, including the constant term (the vector of ones). The location value surface is the estimated γ_0 vector with one element for each element of w.

The LPR smoother has good theoretical properties provided the underlying function, $f(W)$, is smooth.[9] For example, if the value surface is twice differentiable, then the LPR will converge to that surface as the sample gets large. The LPR has superior ability to provide flexibility to the estimated function and it has good properties at the boundary points.

Wand and Jones (1995) summarize the theory of the LPR. Accessible discussions of LPR are given in Fan and Gijbels (1996), Yatchew (1998), and in Hastie, Tibshirani and Friedman (2001). These sources indicate that the LPR, and semiparametric models that have the same form as the LRM, are applied frequently in the medical field.

---

[8] The summations in the denominator, Equation (8), emphasize that the kernel density must sum to one.

[9] The theoretical properties of the LPR distinguish it from the nonparametric part of the McMillen and Thorsnes (2000) model.

The Nadaraya–Watson (NW) estimator given by Equation (8) is a special case of Equations (10) and (11): When the degree of the polynomial is equal to zero ($p = 0$), then the NW estimator follows immediately. The idea of a locally weighted average carries over to Equation (10): $\hat{\gamma}_0$ is the estimate of $\tilde{\eta}$ at the target point, w_0. The polynomial function in Equation (10) serves only to allow for curvature in the neighborhood of w_0.

An OLS regression, linear in W, is a special case of Equations (10) and (11) when the bandwidths become very large. In this case, the estimators are no longer local and all of the points in the data sample receive equal weight, just as with the OLS estimator.

The Partial Bayesian Spatial Model

The approach of the previous section provides a smoothing of the OLS residuals from a standard hedonic model. In particular, the iterative scheme presented alternately updates the hedonic coefficients and then uses LPR to smooth the resulting residuals. For instance, if the algorithm is stopped at iteration t, yielding coefficients $\hat{\beta}_1^t$ and residuals vector ξ^t, then, as proposed in the introduction, we claim to have extracted both measured and unmeasured spatial components present in the mean function.

The objective of this section is to propose an additional spatial smoothing to see if any further spatial structure remains in $\hat{\xi}^t$. To do so we assume a customary spatial process plus white noise process partitioning of $\hat{\xi}(w_i)$, that is, suppressing t,

$$\hat{\xi}(w_i) = \delta(w_i) + \varphi(w_i). \tag{12}$$

In Equation (12), the $\delta(w_i)$ are assumed to come from a stationary Gaussian spatial process with mean 0, spatial covariance function $\text{cov}(\delta(w),\delta(w')) = \sigma^2\rho(w - w';\phi)$. In particular we took ρ of the form $\exp(-\phi\|w - w'\|)$ where $\|d\|$ denotes the length of the vector d. In Equation (12), $\phi(w_i)$ is pure error, that is, the $\phi(w_i)$ are i.i.d. $N(0, \tau^2)$. Such modeling for residuals is standard (see, *e.g.*, Cressie 1993, p. 128).

In the present case, the two components in (12) are needed. We must allow for essentially pure error to be all that remains. We expect that there is still a need for the spatial component after using LPR to estimate unmeasured spatial variation.

We adopt a Bayesian approach in fitting (12) to $\hat{\xi}$ to enable full inference regardless of the model parameters in (12). Also, prediction at new locations is

routine and provides a full predictive distribution; by way of contrast, a point estimate and an approximate standard error are produced by ordinary kriging (Cressie 1993). Hence, the overall approach is only partially Bayesian since it is appended to the foregoing non-Bayesian algorithmic approach to obtain $\hat{\xi}$.

More precisely, we use priors for σ^2, ϕ, and τ^2 which are inverse gamma, gamma and inverse gamma, respectively. We make them partially databased but still rather noninformative, that is, we use the data to center the priors but make the prior variability quite large. Fitting of the Bayesian models arising from (12) is standard using simulation methods. We use a Markov Chain Monte Carlo algorithm, which is customarily referred to as a Metropolis within the Gibbs sampler. Such approaches create a Markov chain (each number in a sequence depends on previous numbers) whose stationary distribution is the posterior for the model unknowns. In the present case, these are σ^2, ϕ, and τ^2. The objective of the simulation is to obtain samples from this posterior distribution. We can draw as large a sample as we wish and, using this sample, we then can infer to arbitrary accuracy about any feature of the posterior distribution. More precisely, at least two independent trajectories of the chain are run for a large number of iterations until, using appropriate convergence diagnostics, a determination is made that the chain is at equilibrium.

The preconvergence iterations are referred to as the burn-in. After the burn-in, transitions of the chains are continued, and, to gather essentially independent samples, only the realizations at every kth iteration are retained where k is chosen so that the autocorrelation at lag k is sufficiently small. See, for example, Carlin and Louis (2000) or Gelman *et al*. (1995) for an introductory presentation of this methodology. In the present instance, two chains were burned in for 10,000 iterations. Then, an additional 25,000 iterations were run and thinned ($k = 25$) to obtain a posterior sample of size 1000. The posterior samples provide full inference about the (partial) sill, nugget, range and variograms at any distance. The predictive distribution for $\varepsilon(w_0)$**,** that is, at a new location, is

$$f(\xi(w_0) \mid \hat{\xi}) = \int f\big(\xi(w_0) \mid \hat{\xi}, \sigma^2, \phi, \tau^2\big) f(\sigma^2, \phi, \tau^2 \mid \hat{\xi}). \tag{13}$$

In (13) the second distribution under the integral sign is the posterior, which is not available analytically, but from which we obtain samples using the above simulation. The first distribution under the integral is a normal. Hence, given an observation from the posterior, we may immediately obtain a random observation from the resulting conditional normal distribution. Doing this for each posterior sample, we obtain a sample from the predictive distribution in (13). In fact we can do this for as many new locations as we wish. Also, adding $X(w_0)\hat{\beta}$ to each sample from (13) gives a realization that is approximately from the predictive distribution for $Y(w_0)$. Hence we obtain a sample from a predictive

distribution for $Y(w_0)$ given the observed data. This sample enables predictive point estimation for cross-validation using, for example, mean square error. It also enables simple $1 - \alpha$ prediction intervals for $Y(w_0)$ using, for example, the $\alpha/2$ quantile and the $1 - \alpha/2$ quantile of the sample.

The Data and Validation Sample

The data set covers 2,338 single-family transactions from January, 1996, through April, 1999. Figure 1 shows the geographic area which includes six small towns in Massachusetts with sales prices and dates, interior square footage, building age, bathrooms, lot size, latitude and longitude (see Table 1). This is an area with high socioeconomic status. The average house value was over \$445,000 (=exp(13.01)) in 1999 dollars, nearly two-thirds of adults had a B.A. degree in 1990, and per capita income was about \$37,000.

This study focuses on the town of Lincoln, Massachusetts, because the Lincoln tax assessor made the data available. Lincoln contains a historic central point known as the "Flower Pot." It also has a military air base that is close to the town dump. The distances from each of these two points are used as explanatory variables to identify measured spatial patterns of house value, that is, we expect that these distances will shape spatial structure throughout the study area.

Figure 1 ■ The six-town area.

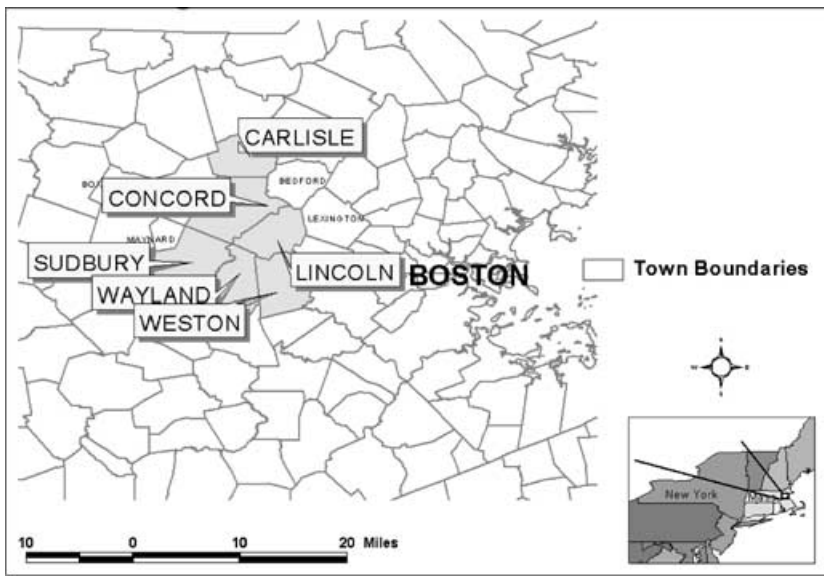


Table 1 ■ In-sample and validation sample statistics.

Variable	Mean	Std. Dev.	Variance	Minimum	Maximum
Panel A: In-Sample Statistics: 2,011 Observations					
lnSP99	13.01	0.49	0.24	10.60	15.36
lnlotsze	10.64	0.82	0.67	7.61	14.25
lninterSF	7.73	0.44	0.20	6.55	9.45
latitude	42.41	0.06	0.00	42.32	42.55
longitude	−71.37	0.05	0.00	−71.48	−71.26
building age	42.32	36.12	1304.42	0.00	283.00
age > 50	19.30	43.76	1915.19	0.00	283.00
age > 50-squared	3,094.82	7,416.15	54,999,238.71	0.00	80,089.00
baths	2.02	0.89	0.79	1.00	8.00
dist. Flower Pot, miles	5.65	2.04	4.15	0.24	9.97
dist. air base, miles	6.81	2.39	5.74	0.60	11.68
Carlisle	0.08	0.27	0.07	0	1
Lincoln	0.07	0.25	0.06	0	1
Concord	0.23	0.42	0.18	0	1
Sudbury	0.29	0.45	0.21	0	1
Wayland	0.20	0.40	0.16	0	1
Weston	0.13	0.33	0.11	0	1
YY96	0.35	0.48	0.23	0	1
YY97	0.32	0.47	0.22	0	1
YY98	0.27	0.44	0.19	0	1
YY99	0.07	0.25	0.06	0	1
Panel B: Validation Sample Statistics: 327 Observations					
lnSP99	13.07	0.49	0.24	11.57	14.73
Lnlotsze	10.62	0.81	0.66	8.47	14.63
lninterSF	7.77	0.42	0.18	6.66	8.86
Latitude	42.41	0.06	0.00	42.32	42.56
Longitude	−71.37	0.05	0.00	−71.47	−71.27
building age	43.00	39.49	1559.14	0.00	265.00
age > 50	20.63	47.10	2218.03	0.00	265.00
age > 50-squared	3,403.11	7,977.87	63,646,440.64	0.00	70,225.00
Baths	2.07	0.81	0.65	1.00	5.00
dist. Flower Pot, miles	5.57	1.96	3.83	0.22	9.86
dist. air base, miles	6.69	2.39	5.72	0.43	11.65

Notes: Variable definitions: *lnSP99* = natural log of verified sales price, updated to 1999; *lnlotsze* = log of lot area (square feet); *lninterSF* = log of living area of the house; *building age* is years from construction to sale; *age > 50* is building age if greater than 50 years, otherwise zero; *age > 50-squared* is the square of this variable. *Baths* count the number of full and half baths; half baths are rounded up. Distances (in miles) are along a cord, uncorrected for the earth's curvature, from the indicated location to each property. They are calculated from latitude and longitude with a cosine correction for the degrees of latitude. The Flower Pot is the historic center of the town of Lincoln. The military air base is also near the Lincoln town dump. YY_{yy} variables are dummies for the year of the sale. Statistics for time and town dummies, not used for out-of-sample estimates, are for 2,338 sales; their means add to one.

The validation sample was designed to randomly select about 15% of the observations, without replacement. After the duplicate observations were discarded, 327 out-of-sample observations were retained. This left 2,011 in-sample observations (see Table 1, Panels A and B). The descriptive statistics leave no doubt that the validation sample was randomly drawn from the same population as the in-sample data.

Turning to regressions in Table 2, the town dummy variables appear to do a good job of capturing the effects of local public goods and taxes. Each town has its

Table 2 ■ OLS used to update the sales to 1999 for Lincoln, Massachusetts.

Regression with Town and Time Dummies				
Valid cases:	2338	Dependent variable:		*lnSP*
Missing cases:	0	Deletion method:		None
Total SS:	599.185	Degrees of freedom:		2319
R-squared:	0.689	*R*bar-squared:		0.686
Residual SS:	186.539	Std. error of est.		0.284
F(18,2319):	284.995	Probability of *F*:		0
Variable	Estimate	Standard Error	*t*-value	Prob. > |*t*|
constant	94.170	25.984	3.624	0.000
Lnlotsze	0.100	0.010	10.107	0.000
LninterSF	0.628	0.023	26.926	0.000
Latitude	1.252	0.578	2.167	0.030
Longitude	1.968	0.472	4.168	0.000
building age	−0.003	0.001	−4.300	0.000
age > 50	0.001	0.000	3.445	0.001
age > 50-squared	0.000007	0.000002	2.925	0.003
Baths	0.071	0.010	6.876	0.000
dist. Flower Pot, miles	−0.056	0.019	−2.978	0.003
dist. Air base, miles	0.066	0.023	2.838	0.005
Carlisle	−0.080	0.075	−1.061	0.289
				Lincoln omitted
Concord	0.143	0.042	3.437	0.001
Sudbury	0.095	0.045	2.106	0.035
Wayland	0.033	0.035	0.932	0.351
Weston	0.246	0.039	6.355	0.000
YY96	−0.215	0.025	−8.528	0.000
YY97	−0.143	0.025	−5.610	0.000
YY98	−0.033	0.026	−1.293	0.196
				YY99 omitted

Notes: See Table 1 for variable definitions. The breakpoints on building age come from experimentation with alternative curvatures. The dependent variable, *lnSP*, is log of sales price at the time of the sale. *lnSP99* was produced from this regression by multiplying the town dummies and time dummies (last eight variables) by their coefficients and subtracting from *lnSP*.

own government, responsible for the school system, other public services and taxation. In particular, the strong positive coefficient for the town of Weston is consistent with casual empiricism indicating an excellent school system in that town. Also, the socioeconomic status of Weston is only slightly behind that of Lincoln. Whatever the reason, houses in Weston fetch a 24.6% premium when compared to Lincoln.

The time dummy coefficients have plausible magnitudes, indicating that house prices increased by 21.5%, with an estimated standard error of about 2.5%, over the four-year period (Table 2).[10] Previous research (Clapp 2001) showed that house price indices (*i.e.*, the time pattern of house prices) are not significantly different over the six towns, or for neighborhoods within towns. Therefore, sales prices were updated to 1999 for the town of Lincoln; town dummies and time dummies were multiplied by their coefficients, the last eight rows of Table 2. This inner product was subtracted from the dependent variable to produce updated values, *lnSP99*. This process did not affect the other explanatory variables. These updated sales prices are described in the first row of Table 1, Panel A, and, for the validation sample, Table 1, Panel B.

Results

Ordinary Least Squares Models (Table 3)

The in-sample coefficients on lot size, interior square footage and bathrooms are plausible in sign and magnitude (Model 1, Table 3).[11] The effect of building age is significant with the expected negative sign for buildings up to 50 years old. For older buildings, it is important to model the historic character of these towns; the oldest building in the transactions data set is 283 years, and 118 sales are for structures more than 99 years old. Age has a highly nonlinear effect, with older buildings increasing in value.[12] Of course, better quality and better located older buildings tend to be preserved.

Distances from the Flower Pot and air base are significant in Model 1, with the expected signs. The result is surprising, because the two locations are only 2.4

[10] Quarterly time dummies were tried, but they did not change the explanatory power of the model.

[11] Note that the spatial dimensions of the OLS regressions have been specified as completely as possible in order to run a fair horse race between OLS, Bayesian kriging and the LRM. ZIP code and Census tract boundaries are coterminous with town boundaries in these small towns.

[12] The 50-year breakpoint and the squared term for age over 50 years come from experimentation with alternative specifications.

Table 3 ■ OLS models and backfitting results.

	Model 1 Base Model	Model 2 Backfitted Coefficients	Model 3 Polynom. in Distance		Model 4 Polynom. in Lat./Lon.	Model 5 Simulated Subcenter
Constant	56.664	0.021	24,358.394	constant	−53,512	−171,061
	(3.485)	(0.120)	(1.520)		(−1.246)	(−3.967)
	0.001	0.9	0.128		0.213	0.000
lnlotsze	0.075	0.081	0.098	*lnlotsze*	0.095	0.088
	(7.396)	(9.060)	(10.193)		(10.016)	(9.212)
	0.000	0.000	0.000		0.000	0.000
lninterSF	0.652	0.605	0.632	*lninterSF*	0.636	0.647
	(26.326)	(25.980)	(27.414)		(27.700)	(28.038)
	0.000	0.000	0.000		0.000	0.000
latitude	0.360		−991.169	*latitude*	1,369.5450	3,403.419
	(0.812)		(−1.512)		(1.450)	(3.587)
	0.417		0.130		0.147	0.000
longitude	0.909		−248.468	*longitude*	−688.109	−2,769.486
	(4.446)		(−1.502)		(−1.008)	(−4.040)
	0.000		0.133		0.313	0.000
building age	−0.002	−0.004	−0.002	*building age*	−0.002	−0.002
	(−3.086)	(−5.39)	(−4.406)		(−4.251)	(−4.337)
	0.002	0.000	0.000		0.000	0.000
age > 50	0.001	0.002	0.001	*age > 50*	0.001	0.001
	(2.655)	(3.680)	(3.499)		(3.495)	(3.670)
	0.008	0.001	0.000		0.000	0.000
age > 50-squared	0.000005	0.000008	0.000007	*age > 50-squared*	0.000006	0.000006
	(1.976)	(3.660)	(2.991)		(2.802)	(2.613)
	0.048	0.000	0.003		0.005	0.009

Table 3 ■ continued.

	Model 1 Base Model	Model 2 Backfitted Coefficients	Model 3 Polynom. in Distance		Model 4 Polynom. in Lat./Lon.	Model 5 Simulated Subcenter
baths	0.078 (7.026) 0.000	0.069 (6.670) 0.000	0.070 (6.930) 0.000	*baths*	0.071 (6.944) 0.000	0.068 (6.621) 0.000
dist. Flower Pot, miles	−0.05 (−2.896) 0.004	−0.063 (−10.24) 0.000	0.023 (0.874) 0.382	*dist. Flower Pot*, miles	0.045 2.050 0.040	0.071 (3.198) 0.001
dist. air base, miles	0.042 (2.071) 0.039	0.033 (6.360) 0.000	0.048 (1.788) 0.074	*dist. air base*, miles	−0.003 (−0.159) 0.873	−0.027 (−1.392) 0.164
dist. Flower^2			3.357 (1.495) 0.135	*latitude*^{2}	−27.305 (−3.152) 0.002	−70.92 (−8.151) 0.000
dist. air^2			−3.409 (−1.528) 0.127	*longitude*^{2}	−8.78 (−2.171) 0.030	−30.287 (−7.455) 0
distances interacted			0.044 (2.155) 0.031	*latitude; ts longitude*	−13.268 (−3.178) 0.002	−36.64 (−8.737) 0.000
*R*bar-squared	0.658	n/a	0.668	*R*bar-squared	0.667	0.724

Notes: t-values are in parentheses, coefficients are above, and *p*-values are below. Variables are defined in Table 1.

miles apart. One can conclude that the undesirable aspects of the air base (also close to a town dump) and the desirable aspects of locations near the Flower Pot are important to spatial patterns of house value in the six-town area.

A simple first-degree polynomial approximation to the house value surface is modeled with latitude and longitude where one degree is approximately 70 miles after an adjustment for latitude. The analysis shows that longitude is statistically significant but not latitude. Further exploration with higher degree polynomials in distances (Model 3) and coordinate position (Model 4) indicate that some of the added variables are statistically significant; also, the adjusted R^2 improves modestly. The quadratic polynomial in Model 4 can have at most one peak in the six-town area, so it has added limited flexibility to $f(W)$. Collinearity prevents inversion of the X; pr$X$ matrix when cubed terms are introduced, that is, before interesting house value patterns can be observed. This is a typical problem with trend surface analysis (Jackson 1979, and Agterberg, Gaile and Willmott 1984). Therefore, we turn to the LRM model as a way of measuring the effects of unmeasured variables on the house value surface.

LRM Model Results

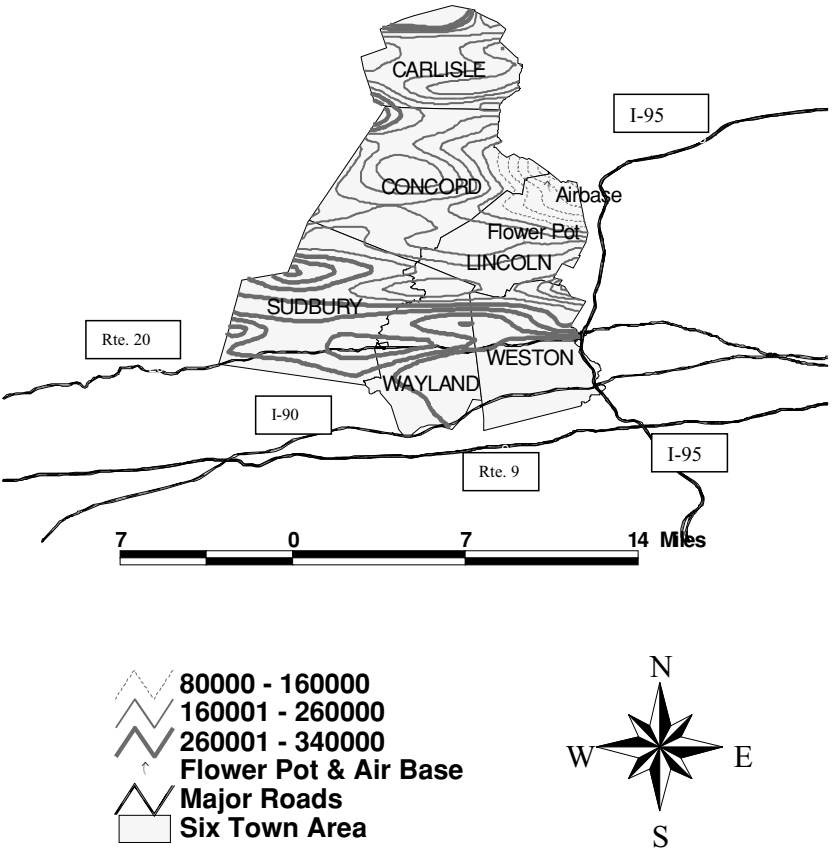
Model 2 (Table 3) shows the $\hat{\beta}$ parameters after the backfitting iterations given by Equations (2) through (7). These iterations are continued until the coefficients change by less than 5%, about seven iterations in this case.

All the coefficients on building age are much larger in absolute value after the iterative fitting. This is the most obvious result of the fact that building age is highly correlated with location; houses are typically built in neighborhoods and each neighborhood is constructed at about the same time. The larger magnitudes of the age coefficients suggest that a better method of dealing with spatial variation, and making that variation orthogonal to the $X\hat{\beta}$ part of the model, has improved the estimators. The higher t-values after backfitting indicate more precisely measured coefficients.

The LRM provides interesting information about spatial patterns. Figure 2 shows that the LRM identifies spatial variation due to the air base and to the area near the Flower Pot. Also, substantial spatial variation in house value is observed in towns surrounding Lincoln. Most importantly, subcenters have clearly been identified in Concord, Sudbury and Wayland; other centers appear to border Carlisle and Concord.[13]

[13] The LPR is notable for unbiased estimation near boundary points. Cross-validation indicated that a polynomial of degree one should be used, degree two can be eliminated on theoretical grounds and degree three had high variance.

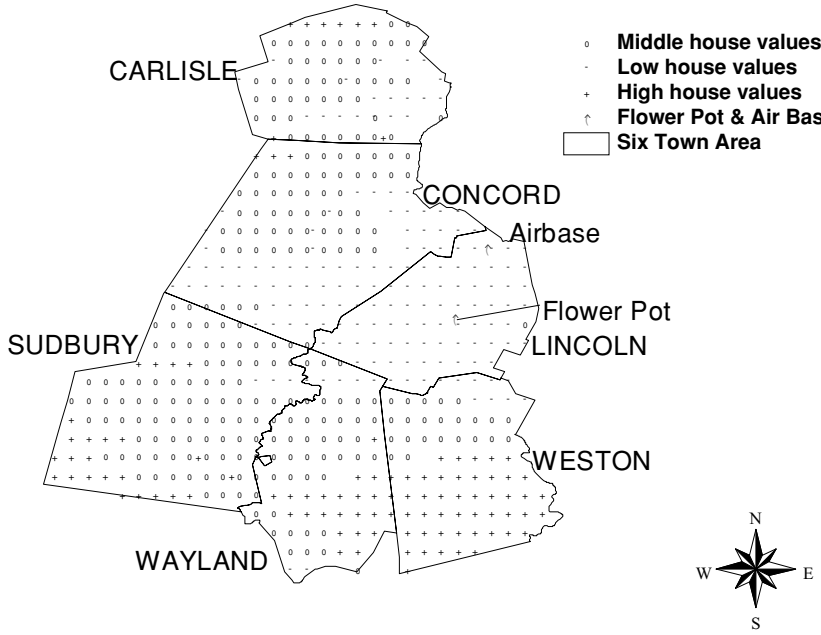
Figure 2 ■ Measured and unmeasured variables influence location value.



The influence of unmeasured spatial variables on house values is given by $\exp \tilde{\eta}$. This term makes a somewhat bigger contribution to the spatial patterns in Figure 2 than measured spatial variables (*i.e.*, the two distances). The variance of unmeasured divided by the variance of unmeasured plus measured is 56%.[14] Figure 3 shows the broad spatial patterns of house values due to unmeasured variables by dividing these values into three groups: low house values are more than 0.5σ below the mean, high are more than 0.5σ above the mean, and middle are all the rest. The σ is calculated from a cross section over the entire six-town area.

[14] By construction, the unmeasured component estimated by LPR has zero sample covariance with the measured spatial component estimated by OLS.

Figure 3 ■ Unmeasured spatial variables influence house values. The value of a standardized house is calculated as a function of latitude and longitude from the LPR part of the model. From the set of these location values, middle values refer to values within 0.5σ of the mean, low is less than 0.5σ below the mean, and high is more than 0.5σ above the mean. The figure is a top-down view of a value topographical map.



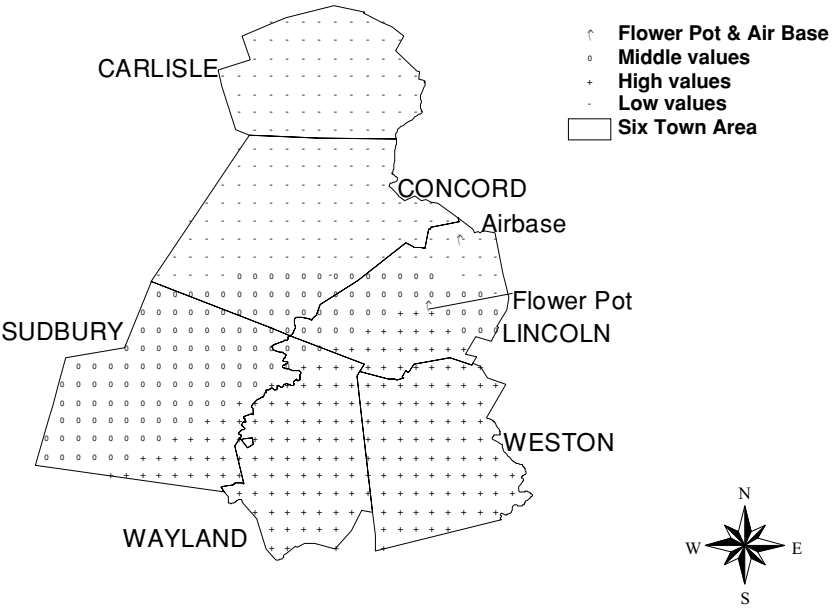
Compare Figure 3 to Figure 4, which contains the same calculations for measured spatial house values. Clearly, there is more spatial variation in Figure 3, where all towns except Lincoln have values in the high, middle and low ranges. Measured spatial patterns (Figure 4) dominate house value movements in the town of Lincoln. This is not surprising since the two landmarks used for these patterns are both in Lincoln.[15]

Bayesian Kriging Results

To examine any remaining spatial structure, the Bayesian spatial model (12) was fitted separately to the LRM residuals in (4) and also to the OLS residuals in (2). Figure 5 presents histograms of the posterior distributions for the model

[15] The model–low and medium patterns combined–produces considerable spatial variation in house values, from a minimum of about $80,000 to a maximum of over $580,000. The mean is $275,000 with a standard deviation of $80,000. The abilities of Figures 2 through 4 to show this are limited.

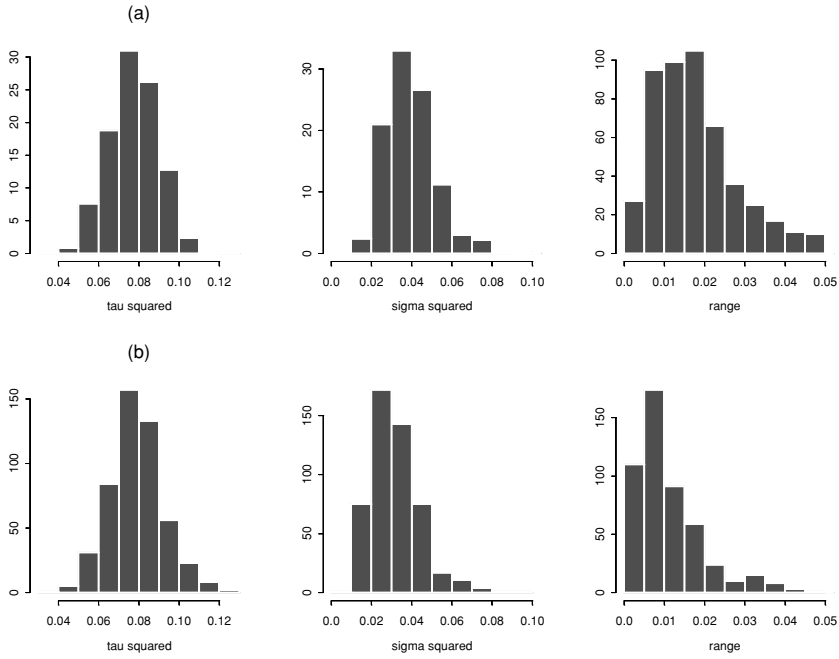
Figure 4 ■ Measured distances influence the spatial pattern of house values. The value of a standardized house is a function of distances to the Flower Pot and to the air base calculated from the OLS part of the model. From the set of these values, middle values refer to the values within 0.5σ of the mean, low is less than 0.5σ below the mean, and high is more than 0.5σ above the mean. The figure is a top-down view of a value topographical map.



parameters. Table 4 summarizes posterior means and 95% interval estimates for the three model parameters. As expected, it reveals slightly stronger evidence of spatial association for the OLS residuals than for the spatially refined LRM residuals. In either case, the pure noise variability is more than twice spatial variability; the residuals reflect primary microscale variability. The range estimate is roughly 1.035 miles for OLS residuals and 0.63 miles for LRM residuals, which indicates that LRM residuals are less spatially correlated. Thus, the LRM model better explains the measured spatial pattern of the data than the pure OLS approach. Overall variability ($\tau^2 + \sigma^2$) is larger for OLS residuals than for LRM residuals.

These results imply that there is little spatial autocorrelation in the LRM residuals and that changing from the OLS model to the LRM (*i.e.*, introducing additional spatial modeling into the mean function) reduces spatial autocorrelation. The downward shift in the posterior distribution of the range (Figure 5b compared to 5a) suggests that practitioners may be less concerned about omitted

Figure 5 ■ Histograms for posterior distributions of spatial model parameters showing OLS residuals (a) and LRM residuals (b). See text for details. τ^2 = nugget, the variance for standardized houses at essentially the same location. σ^2 = value of the spatial covariance function at zero distance (partial sill). $\sigma^2 + \tau^2$ = sill, which describes the variability of the process. Range is the distance in degrees before spatial autocorrelation dies out; multiply by 70 to get miles.



spatial variables once they have implemented the LRM. Of course, these particular results are due to the relatively homogeneous nature of these six towns in Massachusetts. Other areas may have considerable spatial autocorrelation even after implementation of the LRM.

The posterior distributions produced by the Bayesian model are useful to practitioners and to researchers. For example, the range indicates the distance within which comparable sales are particularly helpful to appraisers, assessors and lenders. The variability in this parameter (see Table 4 and Figure 5) indicates the confidence that can be placed in a given comparable sale. In this case, the highly skewed distribution of the range for LRM residuals indicates a range of no more than 0.02 degrees, only about 1.5 miles.[16] For the researcher, these

[16] Classical estimation of asymptotic standard errors would ignore this skewness.

Table 4 ■ Inference summary of Bayesian models for OLS and LRM residuals.

	2.5%	50%	97.5%
Panel A: OLS Residuals			
τ^2	0.053	0.078	0.100
σ^2	0.020	0.038	0.072
range	0.003	0.015	0.048
Panel B: LRM Residuals			
τ^2	0.054	0.078	0.107
σ^2	0.014	0.030	0.061
range	0.003	0.009	0.036

Notes: The percentages refer to percentile points on the cumulative frequency distribution calculated from the Gibbs sampler. τ^2 = nugget, the variance for standardized houses at essentially the same location; σ^2 = value of the spatial covariance function at zero distance (partial sill); $\sigma^2 + \tau^2$ = sill, which describes the variability of the process. Range is the distance in degrees before spatial autocorrelation dies out; multiply by 70 to get miles.

results indicate that the parameters of the OLS part of the model are efficiently estimated; little would be gained by simultaneously estimating spatial autocorrelation and the mean structure of the model.

Out-of-Sample Prediction Results

Table 5 shows that the benchmark OLS model has out-of-sample mean standard error (MSE) of 0.079, or about 7.9% of the predicted sales price. The estimated OLS bias (=mean error) is +0.028.

The LRM increases bias by over 14% and reduces MSE by 9%. Thus, as indicated by the mean absolute error, out-of-sample variance has been substantially reduced by the LRM.

Bayesian spatial modeling of OLS residuals has the opposite effect. It reduces bias by 9% but increases MSE by 22% due to a substantial increase in variance. Bayesian spatial modeling of LRM residuals increases both bias and variance by about 5% and 20%, respectively. The Bayesian modeling is expected to manifest increased variability since, unlike OLS and LRM, it treats model unknowns (σ^2, ϕ and τ^2) as random, endowing them with prior variability.

We conclude that both measured and unmeasured spatial variables are significant in our data set whereas high frequency spatial patterns are not. These high frequency patterns can add information when unmeasured spatial variables are left out, but a lot of noise is added also. This is likely an artifact of the six-town

subject area—as a small, fairly homogeneous suburb, high frequency patterns are relatively small. We note that Host (1999) used a similar model to find significant high frequency spatial patterns for air pollution.

Panels B and C of Table 5 compare two more flexible OLS models to the LRM, *with all distance variables in the LPR part of the model*. An important implication of these panels is that the LPR part can adequately measure the effects of the Flower Pot and air base. The MSE is reduced slightly by putting distance variables in the LPR part. This indicates that major centers do not have to be measured and that distance variables need not be calculated. This is an advantage for large-scale applications of the model, where consistent measurement of "the CBD" or other major points might be difficult.

Is a polynomial in latitude and longitude (see Jackson 1979) capable of capturing unmeasured spatial variation? Panels B and C of Table 5 show that these more flexible functional forms can capture some of the out-of-sample variation in house prices.[17] The advantage of the LRM is limited to a reduction of 3.6% for the polynomial in distance and 4.2% for the polynomial in coordinate positions.

As with the kriging results, the limited ability of the LRM to reduce MSE is likely due to the special characteristics of the study area. Moreover, it would be interesting to know if the LRM adds useful predictive ability in areas of high curvature in the underlying value surface. These issues are explored in the next subsection.

Simulating Additional Spatial Variability

We hypothesize that our model will work better on a data set that has more variation in the spatial pattern of house prices. For example, Anas, Arnott and Small (1998, p. 1440) and McMillen (2001, p. 457) point out that many cities have more than a dozen employment subcenters. The surprising information is that this includes "old" cities such as Chicago, which has 33 subcenters. In theory, our model could measure the impact of these subcenters on house values without the need to identify the location of each center. Since identification of subcenters is complex (McMillen 2001), this is an important advantage of the model.

To test this idea, we simulated the addition of a subcenter in the six-town area. The simulation was constructed as follows:

[17] The limits of this parametric approach are indicated by the fact that latitude and longitude cubed could not be added to the OLS model. Perfect collinearity prevented inversion of the $X'X$ matrix.

Table 5 ■ .Out-of-sample results for 327 observations.

				Bayesian			
	OLS	LRM	% Difference	OLS-resid.	% Difference	LRM-resid.	% Difference
Panel A: Base OLS Model (Model 1)							
Mean absolute error	0.188	0.178	−4.9%	0.220	17.4%	.207	10.6%
Mean error	0.028	0.032	14.4%	0.025	−8.9%	0.033	20.1%
MSE	0.079	0.072	−9.0%	0.096	21.8%	0.086	9.1%
Panel B: OLS with polynomial in distance (Model 3)							
Mean absolute error	0.183	0.178	−2.6%				
Mean error	0.025	0.035	41.0%				
MSE	0.074	0.071	−3.6%				
Panel C: OLS with polynomial for Lat./Lon. (Model 4)							
Mean absolute error	0.184	0.178	−3.0%				
Mean error	0.025	0.034	38.5%				
MSE	0.075	0.071	−4.2%				
Panel D: Simulated subcenter and polynomial for Lat./Lon. (Model 5)							
Mean absolute error	0.187	0.178	−5.2%				
Mean error	0.023	0.027	14.6%				
MSE	0.076	0.071	−6.8%				
Panel E: 152 observations within four miles of the Simulated Subcenter (same model as Panel D)							
Mean absolute error	0.156	0.147	−5.7%				
Mean error	0.046	0.043	−6.1%				
MSE	0.047	0.043	−9.6%				

Notes: In all cases, the model was estimated using the 2,011 in-sample observations; predictions were made for the 327 randomly selected out-of-sample observations. Specifications for the five models are defined in Table 3. Model 2 specifies the coefficients on the linear part of the LRM. Panels B through E use similar LRM specifications except that all location variables, including distances, are in the nonparametric (LPR) part of the model. The Bayesian model reveals little high frequency association (spatial autocorrelation), so it is not repeated for all models.

- Randomly select a latitude and longitude coordinate within the six-town area;
- Add 50% to house values at this simulated subcenter;
- Decrease the add factor by eight percentage points per mile distance from this subcenter;
- Construct simulated ln*SP* values by adding the amounts calculated in steps 2 and 3 to the data;
- Rerun the model and out-of-sample statistics with the simulated ln*SP* vector.

The random choice in step 1 produced a subcenter near the southwest corner of Concord. Therefore, the maximum possible influence of the simulation was limited to a rough semicircle around the subcenter.

Simulation results are shown in Panels D and E of Table 5. With the simulated subcenter, the LRM can improve MSE by nearly 7% when compared to the quadratic in latitude and longitude. Moreover, the LRM improves MSE by 9.6% for 152 out-of-sample observations within four miles of the simulated peak in value. Thus, the LRM is capable of improving predictive performance over areas with substantial curvature in the underlying spatial pattern.

Software to Implement the LRM and Bayesian Models

We used a Gauss program developed by the authors for the LRM. Model estimation is made faster with some C++ programs driven by Gauss. However, speed appears to be secondary; the slowest program, cross-validation to find the optimal bandwidth, took less than three minutes with a 1.4 Ghz AMD processor. The Bayesian method was implemented with a Fortran program developed by the authors.

We have taken pains to use cutting-edge methods. However, the basic elements of the LRM are available in commercial software. Smoothing can be implemented in SAS, STATA and XploRe (available at http://www.xplore-stat.de). Yatchew (1998) stresses the simplicity of programming for smoothing algorithms. SpaceStat (available at http://www.spacestat.com) facilitates several forms of classical spatial autoregressive analysis. Dorfman (1997) discusses simple algorithms for applied Bayesian methods, and he guides the reader to web sites with useful programs. Software for Bayesian inference using Gibbs Sampling (BUGS software) is available at http://www.mrc-bsu.cam.ac.uk.

Conclusions

The influence of causal variables (measured or omitted) on the spatial pattern of house prices is categorized into three groups of variables. This article combines OLS, intended to capture measurable effects, local polynomial smoothing (LPR), intended to capture the effects on price of unmeasured spatial variables, and Bayesian spatial models designed to capture spatial autocorrelation introduced by unmeasured variables. Thus, our model extends the ability of hedonic models, including those estimated with a spatial autocorrelation function, to deal with unmeasured variables that cause spatial variation in house prices.

The LRM is capable of producing interesting and plausible iso-value lines (Figure 2). While it adds some bias to the mean out-of-sample error, it reduces variance enough to offset this and reduce MSE by 9%. Thus, the ability of the LRM to model unmeasured spatial processes produces improvements in predictive ability.

The Bayesian model produces estimates of parameters for the spatial process which models residual error from LRM or OLS. The model estimates a posterior distribution of these parameters. Thus, Bayesian modeling is capable of adding useful information to the OLS and LPR parts. The particular study area used here does not show substantial high frequency spatial patterns. Therefore, the data show that practitioners in this six-town area can focus on the other components of the spatial variation in house value.

Adding a random subcenter to the data simulated additional spatial variation in the mean function of house prices. The simulated data show that the model can improve out-of-sample predictive ability when compared to the most flexible parametric specifications. Thus, the model would appear to be most applicable in areas such as Chicago, where large numbers of subcenters add substantial spatial variation to house prices. It is likely that Bayesian spatial modeling would also add significantly to predictive power in such an area.

Our base model puts measurable spatial variables (two distance variables in the six-town area) into the linear part of the model. But out-of-sample tests show that this is unnecessary: The LPR part of the model can adequately capture the effects of distances to important points. This could be highly advantageous in a large-scale application to automated valuation.

The authors thank Kelley Pace for helpful suggestions and the Center for Real Estate and Urban Economic Studies at the University of Connecticut for research support. Tom Thibodeau and four anonymous referees made numerous constructive comments that have improved the article.

References

Agarwal, D., A.E. Gelfand, C.F. Sirmans and T.G. Thibodeau. 2001. *Nonstationary Spatial House Price Models.* Department of Statistics, University of Connecticut.

Agterberg, F.P., G.L. Gaile and C.J. Willmott. 1984. *Spatial Statistics and Models*. D. Reidel Publishing Company: Dordrecht, Holland.

Anas A., R. Arnott and K. Small. 1998. *Journal of Economic Literature* 36(3): 1426–1464.

Can, A. and L. Anselin. 1998. Spatial Effects in Models of Mortgage Origination. Draft paper presented at the midyear meetings of the American Real Estate and Urban Economics Association, Washington, DC.

Carlin, B.P. and T.A. Louis. 2000. *Bayes and Empirical Bayes Methods for Data Analysis*. Chapman and Hall/CRC: Boca Raton, FL.

Clapp, J.M. 2001. A Semiparametric Method for Valuing Residential Location. Working Paper. University of Connecticut.

Clapp, J.M., M. Rodriguez and R.K. Pace. 2000. Residential Land Values and the Decentralization of Jobs. *Journal of Real Estate Finance and Economics* 22(1): 43–62.

Cressie, N. 1993. *Statistics for Spatial Data*. Wiley-Interscience: New York.

Dorfman, J.H. 1997. *Bayesian Economics through Numerical Methods*. Springer: New York.

Dubin, R.A. 1992. Spatial Autocorrelation and Neighborhood Quality. *Regional Science and Urban Economics* 22(3): 433–452.

Fan, J. and I. Gijbels. 1996. *Local Polynomial Modeling and Its Applications*. Chapman and Hall: New York.

Gelfand, A.E., M. Ecker, J. Knight and C.F. Sirmans. 2002. The Dynamics of Location in House Price. Forthcoming.

Gelman, A., J. Carlin, H.S. Stern and D.B. Rubin. 1995. *Bayesian Data Analysis.* Chapman and Hall/CRC: Boca Raton, FL.

Gillen, K., T. Thibodeau and S. Wachter. 2001. Anisotropic Autocorrelation in House Prices. *Journal of Real Estate Finance and Economics* 23(1): 5–30.

Hastie, T., R. Tibshirani and J. Friedman. 2001. *Elements of Statistical Learning: Data Mining, Inference, and Prediction.* Springer: New York.

Host, G. 1999. Kriging by Local Polynomials. *Computational Statistics & Data Analysis* 29: 295–312.

Jackson, J.R. 1979. Intraurban Variation in the Price of Housing. *Journal of Urban Economics* 6: 464–479.

McMillen, D.P. 2001. Nonparametric Employment Subcenter Identification. *Journal of Urban Economics* 50(3): 448–473.

McMillen, D.P. and P. Thorsnes. 2000. The Reaction of Housing Prices to Information on Superfund Sites: A Semiparametric Analysis of the Tacoma, Washington Market. *Advances in Econometrics* 14: 201–228.

Mills, E.S. and B.W. Hamilton. 1984. *Urban Economics.* Scott Foresman and Company: Glenview, IL.

Moyeed, R.A. and P.J. Diggle. 1994. Rates of Convergence in Semi-Parametric Modeling of Longitudinal Data. *Australian Journal of Statistics* 36(1): 75–93.

Robinson, P.M. 1988. Root-*N*-Consistent Semiparametric Regression. *Econometrica* 56(4): 931–954.

Wand, M.P. and M.C. Jones. 1995. *Kernel Smoothing.* Chapman and Hall: New York.

Yatchew, A. 1998. Nonparametric Regression Techniques in Econometrics. *Journal of Economic Literature* XXXV: 669–721.

Zeger, S.L. and P.J. Diggle. 1994. Semiparametric Models for Longitudinal Data with Application to CD4Cell Numbers in HIV Seroconverters. *Biometrics* 50: 689–699.