Parcel Size and Land Value: A Comparison of Approaches

Karl L. Guntermann
College of Business
Arizona State University
P. O. Box 874706
Tempe, AZ 85287
Phone: (480) 965-7206
Fax: (480) 965-8629
E-Mail: Karl.Guntermann@asu.edu

Alex R. Horenstein
Department of Economics
School of Business Administration
University of Miami
P.O. Box 248027
Coral Gables, FL 33124
Phone: (305) 284-1627
Fax: (305) 284-2985
E-Mail: horenstein@bus.miami.edu

Federico Nardari
Department of Finance
C.T. Bauer College of Business
University of Houston
334 Melcher Hall
Houston, TX 77204
Phone: (713) 743-4780
E-Mail: fnardari@uh.edu

Gareth Thomas
IHS
4521 Campus Drive, #336
Irvine, CA 92612-2621
Phone: 949-856-3368
Fax: 949-856-2044
E-mail: gareth.thomas@ihs.com

June 2014

The authors wish to thank the editor, three anonymous referees, Seung Ahn, and Crocker Liu whose comments substantially improved the content of this paper. The authors also thank Mark Ecker for useful discussions about the Bayesian estimation methodology and Mark Ecker and Hans Isakson for making their data available. The views expressed in the paper are those of the authors and are not representative of the views of IHS

Abstract

The analysis presented here uses simulated and real data sets to investigate the relative merits of parametric, semi- parametric and Bayesian methods in testing for the co-existence of plottage and plattage and in identifying the inflection point. Using artificial datasets generated with spatial correlation, inflection points at alternative locations over the range of sample sizes, different sample sizes and varying the relative quantities of small vs. large parcels, we find that the Bayesian approach method generally dominates the semi-parametric one. In turn, these two methods strictly dominate the parametric one.

1. Introduction

The appraisal literature has long included the concept of plottage value, which is defined in the Appraisal Institute's text as "The increment of value created when two or more sites are combined to produce greater utility". Because land assembly is not costless, it would occur only if the value of an assembled parcel were at least equal to the sum of the individual parcel values plus the costs associated with assembling them. These costs would include legal and other transaction related expenses, possible changes to infrastructure and higher prices extracted by holdout owners, etc. Land assembly costs would make the plottage portion of the parcel size-value curve convex, reflecting that land values are increasing at an increasing rate with size.

The term "plattage" was first used by Colwell and Sirmans (1978, 1980). It refers to the portion of the size-value curve where subdivision rather than land assembly is possible. Value can be created through the subdivision of land but that process also is not costless. For subdivision to occur, the total value of the smaller parcels would have to exceed their value in the original parcel by at least their subdivision costs, making the size-value curve concave over that range of parcel sizes. It is possible that both plottage and plattage could coexist in a market at the same time but they would be associated with different size parcels. If this is the case, then the relationship between price and parcel size would be "S" shaped, such as one of the options in Figure 1, rather than linear.[1] Point δ is the inflection point where land values stop increasing at an increasing rate and begin increasing at a decreasing rate. To the left of point δ, plottage would occur, while to the right of point δ, plattage would dominate.

2. The Literature

Rigorous analysis of the value-size relationship essentially began with the Colwell and Sirmans (1978, 1980) articles. These papers provided the initial explanation for the existence of an S shaped value-size curve, which ran counter to the prevailing view of a linear relationship between value and size. The models developed by Colwell and Sirmans found evidence for the existence of both plottage and plattage in residential lot data from Champaign, IL and Edinburgh, Scotland. More recent articles by Colwell and Munneke (1999), Thorsnes and McMillen (1998) and Ecker and Isakson (2005) specifically address the relationship between value and size. Tabuchi (1996) and Lin and Evans (2000) find evidence of plottage with data that includes very small parcels.

There is considerable evidence supporting the concave nature of the land value curve for large parcels but the evidence for convexity in the size range where plottage should occur is much weaker. Possible explanations for the lack of evidence include inappropriate model specification and data sets that do not contain enough sufficiently small parcels. It is possible that plottage exists but that the inflection point is beyond the range of the smallest parcels in the dataset and, hence, plottage cannot be detected even if the appropriate functional form is used in the tests. Colwell and Munneke (1999) attempt to detect plottage by comparing the degree of concavity in the land price function between the Chicago Central Business District (CBD) and the rest of Cook County. They focus on the CBD because of the greater likelihood that land assembly would be associated with redevelopment projects. The coefficient on the size variable in the models is greater than one for the residential tests, which supports a convex relationship, while the commercial and industrial tests are consistent with a linear relationship. The authors note that the plottage/plattage tests are indirect in the sense that the transactions used have not been identified as being associated with either assembly or subdivision activity. Therefore, the results may reflect a large number of non-assembly sales, making it impossible to detect plottage even though it might be present.

Thorsnes and McMillen (1998) re-examine the relationship between land values and parcel size using a parametric estimator for distance and a semi-parametric estimator for size. The flexibility of the semi-parametric specification permits the value-size function to reflect any shape (linear, convex or concave) in any portion of the size spectrum. Their model is estimated using data on 158 undeveloped residential parcels in the Portland, Oregon area that range in size from less than one-half acre to over 26 acres. Their results are consistent with studies that find the value-size relationship to be concave over the entire range of their data. However, they find no evidence that small parcels sell at a discount or that the value-size relationship is ever convex.

Ecker and Isakson (2005) develop a general model that reduces to the models previously discussed in special situations. It accommodates convexity for small parcels, concavity for large parcels and a non-deterministic shift point while accounting for spatial correlation. In their model, all parameters are fit simultaneously from a Bayesian perspective by using Markov Chain Monte Carlo Techniques. Their model is estimated using data on 646 residential, arms-length sales of vacant land in two medium sized, Midwestern urban areas (Cedar Falls and Waterloo, Iowa). Their data contains parcels that range from 640 to 4.6 million square feet.

Location also would be expected to have a significant effect on value and is typically measured using distance variables to proxy for accessibility. Transit cost savings associated with better accessibility play a critical role in location decisions and are reflected in the land value-distance relationship. Models testing for plottage and plattage include relatively few variables, typically size, distance and time, if the data might reflect changing market conditions. For example, Thornes and McMillen (1998) follow this format while using several distance variables (CBD, freeway interchange and arterial streets) including county dummy variables to proxy for transit cost savings. It is generally assumed that land value is a negative exponential function of distance and most models specify distance in this way. Value would be expected to decline at a decreasing rate from a central point, which traditionally has been the CBD but a similar value-distance relationship might be associated with other distinct sub-peaks of value such as freeway interchanges, or commercial or industrial centers.

There is evidence that the decline of value with distance may not be as large as previously believed. Colwell and Munneke (1997) show that using price per acre as the dependent variable biases the distance coefficient upward since parcel size tends to increase with distance, while price per acre declines as parcel size increases. Models that allow for nonlinearity between price and size result in a price-distance relationship that declines at a slower rate but the relationship still is negative and significant. The irregularity of the land price gradient with respect to distance has been demonstrated by Colwell (1998) and Colwell and Munneke (2003) through the use of piecewise parabolic multiple regression. The complexity of the price surface supports the use of multiple accessibility measures (Thorsnes and McMillen, 1998) and the expectation that price should decline with distance.

This paper has two objectives. The first is to test the accuracy of various models and estimation methodologies at detecting plottage and plattage when it is present in the data and at identifying the shift point from the convex to the concave portion of the relation.[2] It is possible that some studies produced inconclusive results not because of data limitations but because the models and methods used were incapable of consistently accepting or rejecting correct hypotheses about plottage and plattage. Many published studies used data sets that were relatively small. For example, the original Colwell and Sirmans (1978 and 1980) articles had twenty-six and 102 observations respectively while the Brownstone and DeVany (1991) study that "rediscovered" plottage and plattage had 85 observations. Equally important is the size range covered by the data used in empirical tests. While data to test the plattage portion of the curve are readily available, it is more difficult to gather data on small parcels to test for plottage. To compare and contrast alternative approaches we rely on an extensive Monte Carlo analysis calibrated to real data, which allows us to control the Data Generating Process and, thus, incorporate relevant features of the actual data. A second objective of this paper is to use two large data sets containing transactions covering a wide range of parcel sizes from different metropolitan areas in the empirical tests. The results will provide evidence on the consistency, similarities and differences among the parametric, semi-parametric and Bayesian approaches in estimating the inflection point between plottage and plattage as well as the possible co-existence of plottage and plattage

3. Models and Estimation Procedures

The three procedures that form the basis of our analysis are the parametric specification used in Colwell and Sirmans (1978), the semi-parametric model given in Thorsnes and McMillen (1998) and the Bayesian approach by Ecker and Isakson (2005).

Colwell and Sirmans' parametric model gives an exact form for the relationship between land value and area. This function is as follows:

$$\ln\left[y_i\right] = \ln\left[\beta_1^p\right] + \beta_2^p D_i + \beta_3^p \left(A_i - \delta\right)^{1/3} + \varepsilon_i \tag{1}$$

Where:

y_i = the price of the property

D_i = the distance of the property from a central urban point

A_i = the area of the property

δ = a shift parameter.

ε_i = an i.i.d normal error term with mean zero and variance σ^2

and the superscript p denotes a parameter in the parametric model.

This function (1) allows a relationship between land value and area that can exhibit both plottage and plattage. Land value is essentially an inverse cubic function of land area shifted by the shift parameter, δ. This shift parameter allows for the inflection point in the curve and represents the point at which the first derivative of land value with respect to area is at its maximum. If δ has a value less than the minimum value of A_i then the model would only exhibit plattage but if δ has a value greater than the maximum value of A_i then the model would only exhibit plottage. For values of *s* that lie within the range of values of A_i both plottage and plattage exist. Model 1 is estimated with respect to $\left\{\beta_1^p, \beta_2^p, \beta_3^p\right\}$, for a given δ. This estimation procedure is undertaken for a variety of different δ, and the value of δ that maximizes R^2 is the one used.

Colwell and Sirmans suggest that a further test of plottage/plattage can be undertaken by comparing this model to a model capable of detecting only plottage or plattage but not both simultaneously. This model is given by:

$$\ln\left[y_i\right] = \ln\left[\beta_1\right] + \beta_2 D_i + \beta_3 \ln\left[A_i\right] + \varepsilon_i \tag{2}$$

Here if β_3 takes a value less than one but greater than zero then the model includes only plattage, if β_3 takes a value greater than one then the model includes plottage and if β_3 is equal to one then land value is proportional to area. This model can be estimated through simple ordinary least squares. A J-test (see Davidson and Mackinnon (1981)) can then be undertaken to compare the two models, using the first regression as the null hypothesized true model. The fitted values from the second regression are included as an additional regressor in performing the J-test so that the estimated equation becomes:

$$\ln\left[y_i\right] = \ln\left[\beta_1\right] + \beta_2 D_i + \beta_3 \left(A_i - \delta\right)^{1/3} + \gamma Y^{est} + \varepsilon_i \tag{3}$$

where Y^{est} represents the fitted values of ln(*Price*) from the second regression. If the first model is correct, then the fitted values from the second regression should have no explanatory power on ln(*Price*) when they are included in model 1 and γ should be statistically insignificant meaning that a hypothesis that both plottage and plattage coexist can be rejected.[3]

The semi-parametric model of Thorsnes and McMillen estimates the following relationship:

$$y_i = g\left(A_i\right) + \beta_1^{sp} D_i + \varepsilon_i \tag{4}$$

Where the superscript sp stands for semi-parametric, $g\left(A_i\right)$ is some unknown function. This model is estimated using a kernel based non-parametric estimator for $g\left(A_i\right)$. Thorsnes and McMillen recommend a number of different kernels and in our paper we consider all of them. The Gaussian Kernel has the advantage of having continuous derivatives and is more easily implemented. One characteristic of the Gaussian kernel is that it puts weight on every observation when estimating the kernel density function at every point. This characteristic might be a disadvantage when the sample data has very sparse observations, as is typically case in the tails when using real estate data. For this reason, we also add to the estimations the Quartic and the Triangular kernels, which put weight only on observations that are within certain distance from the estimated point. More specifically, the kernels we use are:[4]

[Table 1 goes here]

where $u = \frac{a_i - a_j}{h}$, $a_i = \ln(A_i)$, h is the bandwidth of the kernel and $I(\bullet)$ is the indicator function. For the three cases we use as bandwidth the rule of thumb which is $h = n^{-0.2}$. For the Gaussian Kernel we also experiment with a variable bandwidth kernel density estimator as proposed in Silverman (1986) in order to mitigate the probable problems arising from sparse data in the tails.[5] This non-parametric estimation procedure attempts to estimate $g(A_i)$ on an observation-by-observation basis. Since there are no parameters involving A_i to be estimated, this procedure only provides point estimates and standard errors for the distance parameter, β^1_{SP}. However it is still possible for this procedure to provide inference on the existence of plottage/plattage.

Thorsnes and McMillen suggest one possible process to test for plottage/plattage would be to look at the first and second derivatives of $g(A_i)$. They first take the average of the second derivatives, $\frac{1}{n}\sum_{i=1}^{n} g''(A_i)$, and test whether it is significantly different from zero. If this average is not significantly different from zero, then the hypothesis of a linear relationship cannot be rejected. Thorsnes and McMillen also take the average of the first derivatives, $\frac{1}{n}\sum_{i=1}^{n} g'(A_i)$ and observe whether this value is less than one (indicating a concave price-area relationship), greater than one (indicating a convex relationship), or equal to one (indicating a linear relationship).
By observing only the average of the first and second derivatives of $g(A_i)$ Thorsnes and McMillen are only able to make statements on the average shape of the function and are unable to test whether both plottage and plattage co-exist. They recommend use of averages since “the non-parametric estimator does not provide accurate point estimates of the second derivatives”[6].

Ecker and Isakson (EI (2005) henceforth) use a Bayesian procedure to estimate at the same time 1) the regression parameters and the spatial correlation structure for the convex part (small lots) as well as for the concave part (large lots) of the land-value relation, 2) the change point from convexity to concavity. With the Bayesian approach all unknown parameters are regarded as random variables. This is one of the main differences with the classical econometric procedures where the random variable is the error between the real value of the unknown parameter and the estimated one, not the parameter itself. Bayesian analysis requires three elements: the data, a likelihood function (i.e., sampling distribution) dictated by the model specification, and prior beliefs about the parameters. Following Bayes rule, the joint posterior density of the parameters is proportional to the product of the likelihood function and the prior density on the parameters. Bayesian inference is accomplished by examining the joint posterior density of the parameters of interest.

EI (2005) suggest that the land-value relation can be tested through the following specifications for, respectively, large and small parcels:

$$\ln\left[y_i^l\right] = \beta_0 + \beta_1 t_i + \beta_2 h_l(A_i) + \sum_{k=3}^{p} \beta_k x_{ik} + \varepsilon_i^l \tag{5}$$

$$\ln\left[y_i^s\right] = \alpha_0 + \alpha_1 t_i + \alpha_2 h_s(A_i) + \sum_{k=3}^{p} \alpha_k x_{ik} + \varepsilon_i^s \tag{6}$$

Where:

the subscript *s* refers to small parcels and the subscript *l* refers to large parcels.

$h_r(A_i), r = l, s$ is a function of the parcel areas, like $h_l(A_i) = \log(A_i)$ in the standard plattage model.

p is the number of covariates

t_i refer to time of sale

x_i is the rest of covariates like zoning status and relative distances.

$\varepsilon_i^r \sim N(0, \tau^2 + \sigma^2), r = l, s$

Spatial correlation is allowed for as follows. First, the covariance among error terms is specified as $Cov(\varepsilon_i^r, \varepsilon_j^r) = \sigma^2 \rho((d_{ij}), \phi)$ for both small and large parcels, where d_{ij} is the Euclidean distance between the sites s_i and s_j. Next, the complete covariance matrix is modeled as:

$\Sigma = \begin{pmatrix} \Sigma_s \; \Sigma_{ls} \\ \Sigma_{sl} \Sigma_l \end{pmatrix}$, where the individual elements of each sub-matrix are:

$$\Sigma_{s_{ij}} = Cov(\log(Y_s(s_i)), \log(Y_s(s_j))) = \tau^2 + \sigma^2 \rho(d_{ij}, \phi) \tag{7a}$$

$$\Sigma_{l_{ij}} = Cov(\log(Y_l(s_i)), \log(Y_l(s_j))) = \tau^2 + \sigma^2 \rho(d_{ij}, \phi) \tag{7b}$$

$$\Sigma_{sl_{ij}} = Cov(\log(Y_s(s_i)), \log(Y_l(s_j))) = \sigma^2 \rho(d_{ij}, \phi) \tag{7c}$$

With these assumptions, all the parameters in equations (5) and (6) can be estimated within a unified modeling framework.

The function h provides the plottage or plattage relationship between the log of parcel size and the log of price. Given the value of the shift point represented by parameter δ, the data is divided as follows:

- If $A_i \leq \delta$ then $h(A_i) = h_s(A_i)$ where h_s is a convex function that generates plottage.
- If $A_i > \delta$ then $h(A_i) = h_l(A_i)$ where h_l is a concave function that generates plattage.

This rather general model allows for the specification of different parameters for the plottage (s) and plattage (l) sections of the data. For the function $h_r(A_i)$ EI (2005) suggest using a Box-Tidwell function of the following form:

$$h_l(A_i, \lambda_l) = \begin{cases} \dfrac{A_i^{\lambda_l} - 1}{\lambda_l} & \text{if } \lambda_l \neq 0 \\ \log(A_i) & \text{if } \lambda_l = 0 \end{cases} \quad \text{for the large areas and,}$$

$$h_s(A_i, \lambda_s) = \begin{cases} \dfrac{A_i^{\lambda_s} - 1}{\lambda_s} & \text{if } \lambda_s \neq 0 \\ \log(A_i) & \text{if } \lambda_s = 0 \end{cases} \quad \text{for the small areas.} \qquad (8)$$

The parameter λ is estimated from the data. The second derivative of $h(A_i, \lambda)$ with respect to A reveals convexity when $\lambda > 1$ and concavity when $\lambda < 1$. Note that when $\lambda = 0$ the functional form of the model is identical to the "classical" logarithmic form of equation (2). Similarly, the shift parameter δ is estimated from the data and its entire posterior distribution can be assessed rather than relying on the more ad-hoc devices typically used by the parametric (OLS-based) and semi-parametric methods. EI (2005) state that "only the Bayesian technique allows one to assess and test for plottage (through a posterior probability) in the presence of both spatial correlation and a non-deterministic change point"[7]. Whether this translates into accurate empirical estimates is an issue we intend to tackle with our simulation study. We apply a Bayesian Markov Chain Monte Carlo (MCMC) scheme following EI (2005). Details and specification of the (rather uninformative) priors are contained in the Appendix.

4. Monte Carlo Simulations

In order to compare the ability of the three approaches to detect the presence of both plottage and plattage and to identify the shift point, a number of Monte Carlo experiments were undertaken. In selecting the data generating process used in the simulations we aim not to penalize any given modeling structure and estimation method. In other words, we generate the data so that, a priori, a given approach is not expected to outperform the alternatives. This is especially relevant for the function that generates convexity and concavity in the land-value relation, as discussed below. At the same time, we want the artificial data to reflect the more relevant empirical regularities uncovered by previous studies.

Specifically, the price of the i^{th} parcel can be written as $Y(A_i, s_i, CBD1_i, CBD2_i)$ where A_i is the size of the parcel, s_i denotes the vector of absolute or exact spatial coordinates and $CBD1_i$ and $CBD2_i$ measure the hypothetical distance to two different central business districts. Then, the models are:

$$\log(Y_s(A_i, s_i, CBD1_i, CBD2_i)) = \alpha_o + \alpha_1 h_s(A_i) + \alpha_2 CBD1_i + \alpha_3 CBD2_i + e_i^s \text{ for } A_i \leq \delta$$

$$\log(Y_l(A_i, s_i, CBD1_i, CBD2_i)) = \theta_o + \theta_1 h_l(A_i) + \theta_2 CBD1_i + \theta_3 CBD2_i + \varepsilon_i^l \text{ for } A_i > \delta$$

where $\varepsilon_i^r \sim N(0, \tau^2 + \sigma^2), r = l, s$; $Cov(\varepsilon_i, \varepsilon_j) = \sigma^2 \rho((d_{ij}), \phi)$ where d_{ij} is the Euclidean distance between the sites s_i and s_j, and where the correlation structure is the same as in (7a), (7b) and (7c).

For the function $h(.)$ a slightly different specification is used from the one in EI (2005), which makes it easier to compare the simulation results from the various approaches[8]. While EI (2005) used $h(A_i) = \frac{A_i^\lambda - 1}{\lambda}$, in this paper we use $h(A_i) = (A_i - \delta)^\lambda$. In order to generate plottage and plattage a shift point δ is specified such that $\min(A_i) < \delta < \max(A_i)$. Then,

$$h_s(A_i) = -(\delta - A_i)^{1/\lambda_s} \text{ for every } A_i \leq \delta \tag{9}$$

$$h_l(A_i) = (A_i - \delta)^{\lambda_l} \text{ for every } A_i > \delta \tag{10}$$

For the plottage part of the distribution to be convex λ_s should be >1 (since $1/\lambda_s$ should be <1) and for the plattage part to be concave λ_l should be <1. Since one of the main purposes of the simulations is to determine the reliability of the estimation techniques to detect plottage and plattage, it is important to generate distributions for a wide range of possible land value curves, which is accomplished by changing the values of λ_s and λ_l. As their values approach one, the relationship between parcel log size and parcel log price becomes linear, while curvature increases as λ_s and λ_l^{-1} take on larger values. For example, if $\lambda_s = 3$ and $\lambda_l = 1/3$ plottage and plattage would be generated with the same curvature as proposed by Colwell and Sirmans (1978) but in this paper their methodology will now be tested in the presence of spatial correlation. As an illustration, Figure 2 shows the land value curves for three specifications of λ_s and λ_l. Overall, nine different sets of assumptions about the shape of the land value curve were used to generate the simulated data.

Table 2 displays the values of each model's parameter across the nine assumed scenarios. The table is divided into two panels, the main parameters are in Panel a) and what we called secondary parameters are in Panel b). The distinction is because the main parameters are those who directly affect the curvature of the price-size relationship and, thus, modifying them will modify the difficulty in estimating the inflection point and the coexistence of plottage and plattage.

The first scenario is the baseline case in which the plottage and plattage areas are perfectly symmetric and the inflection point (δ) is located in the middle of the sample. In Cases 2 to 4 we break the symmetry and generate a less pronounced curvature for the plottage and plattage areas to assess the performance of the estimation procedures under more difficult circumstances. In these cases the inflection point is still in the middle of the sample and the quantity of observations at both sides of the inflection point are the same. Case 5 is like Case 3 except that we increase tenfold the level of spatial correlation through parameter τ^2. Case 6 is designed to study the performance of the estimation methods when one side of the sample has fewer observations than the other. It resembles Case 4 but the plottage area contains only 10% of the observations. Case 7 has the same objective as Case 6 but adds more spatial correlation. Thus, Case 7 is like Case 5 but with fewer observations in the plottage side of the sample. Finally, Case 8 and 9 are designed to study the performance of the estimation methods when the inflection point is toward one tail of the sample, as it appears to be found in several previous empirical studies. Therefore, Cases 8 and 9 resemble Cases 5 and 7 respectively but with an inflection point towards the lower tail of the sample.

The secondary parameters are chosen as follows: A_i, $CBD1_i$ and $CBD2_i$ are generated from a uniform distribution. For the parameters α and θ we picked fixed values across the nine scenarios. The only exception is in Case 1, where we impose the constraint that the α's and θs are identical in order for the plattage and plottage parts to be perfectly symmetric. Finally, the parameters σ^2, τ^2 and ϕ were picked in line with the results obtained by EI.[9]

The data generating process corresponding to each set of assumptions was used to produce two sample sizes: for Cases 1 to 7, 100 observations and 500 observations; for Cases 8 and 9 sample sizes are 100 and 250, respectively. The smaller sample size approximates the number of observations in most of the empirical studies of plottage and plattage while the larger sample permits a test of the sensitivity of the simulation results to sample size. In Cases 1 through 5 and in Case 8 we generate an equal number of small and large parcels, whereas in Cases 6, 7 and 9 we confine the percentage of small parcels to only 10% of the overall sample size: this is also a feature that appears to be commonly found in previous studies. For each set of assumptions and sample sizes 1,000 datasets are generated and each of the three models is estimated.

In the parametric approach we estimate Model (1). The semi-parametric method estimates Model (4). For both models we estimate the shift point δ and report the standard deviation and root mean squared error (RMSE) of δ across 1,000 simulated datasets along with their accuracy in identifying an inflection point that is within the range of the simulated data. The Bayesian MCMC methodology estimates the model made up of eq. (5), (6), (7), (9) and (10). The main objects of interest we report are posterior summaries (namely, mean and standard deviation) for the shift parameter along with its RMSE across datasets and with the probabilities of plottage and plattage. Like for the other methods, the posterior quantities are averaged across simulated datasets.

When using the parametric approach we infer the inflection point by finding the δ that maximizes the R^2 of the imposed functional form. When using the semi-parametric approach, the inflection point is selected by finding the first point in which the average second derivative of the kernel function moves from positive to negative. Therefore, in both cases we only have a point estimate for the inflection point but we do not have a known distribution for the point estimators. More importantly, these two procedures are silent on whether the proposed model does better than a baseline model assuming no inflection point: Thus, using a specification test for the models is necessary. For the parametric approach, the J-test is a natural way to test whether the assumed model fits the data better than a baseline model that does not allow for the coexistence of plottage and plattage. We report the percentage of times in which, across simulations, the J-test rejects the coexistence of plottage and plattage at a 5% significance level.[10] For the non-parametric method, we reject the coexistence of plottage and plattage if the average second derivatives of $g(A)$ does not change sign from positive to negative at any given point in the data range. We report the percentage of times across simulations that the coexistence of plottage and plattage is rejected. On the other hand, when using a Bayesian approach, the inflection point is itself viewed as a random variable (as every other parameter in the model), for which the estimation naturally generates confidence intervals and, hence, posterior probabilities that plottage and plattage coexist. At the broader level, the differences in the inferential tools (J-test vs. Bayesian posterior probabilities), hinges upon the fundamentally different foundations between Bayesian and frequentist approaches. In the latter, a given hypothesis is tested for acceptance or rejection at a chosen significance level, while in the former a subjective probability statement about the hypothesis is made.

5. Simulation Results

The results of the simulations are reported in Tables 3, 4 and 5 for the parametric, semi-parametric and Bayesian estimation respectively. The ability to estimate the coexistence of plottage and plattage across methodologies is evaluated in two dimensions. First, we report how close the estimated δ is to the true one. The column "Mean Shift" report the average value of the estimated δ across simulations, which should be close to 15 for Cases 1-7 and close to 7 for Cases 8-9. Columns "Stdev Shift" and "RMSE Shift" report two measures of dispersion for the estimated δ: the standard deviation and the RMSE of the estimated shift points across simulations respectively. The standard deviation measures the dispersion of the estimated δ with respect to its average estimated value across simulations while the RMSE measures the dispersion of δ with respect to its true value. As these two measures approach zero, the accuracy of the method improves.

The parametric model has been the technique traditionally used by researchers and its discriminatory power is marginal at best. The approach is most accurate in Case 1, which is the original Colwell-Sirmans specification, but even in such case it rejects the true hypothesis that plottage and plattage were present in the data in 12.3 percent of the simulations. The method does, however, do a good job on average in identifying the inflection or shift point (15) in the data in Case 1. The results were somewhat worse in Case 2 while the technique essentially failed in all other cases where the land value curve was closer to being linear. The semi-parametric model was considerably more accurate in all cases compared to the parametric model. However, the semi-parametric model accuracy declined in Cases 6,7,8 and 9, where small parcels constituted only ten percent of the sample or when the shift point occurs close to the lower tail of the size distribution, which more closely reflects the size distribution of parcels in most empirical studies. Among the different kernels used in the semi-parametric estimation, the Quartic kernel is the one that best predicts the coexistence of plottage and plattage while the estimation with the triangular Kernel is the one that has the lowest RMSE for the shift point. The Gaussian kernel does not behave very well as it rejects the co-existence of plottage and plattage a much larger percentage of times relatively to the other kernels. In terms of identifying the location of the shift point, there seems to be no consistent pattern pointing to a better performance of a specific kernel. One exception is the triangular kernel, which tends to underestimate especially in the smaller samples. The Gaussian kernel, though, displays an inferior performance when compared to both the Quartic and the Triangular kernels in term of dispersion of the estimates (higher standard deviation across datasets and higher RMSE). In Cases 8 and 9, in which the shift point is not in the middle of the sample (asymmetric shift point), the parametric and semi-parametric techniques are unable to correctly estimate the size at which the shift point occurs. In these cases the parametric model is unable to infer the coexistence of plottage and plattage while the semi-parametric is able to do so, especially when using the Quartic Kernel.

The simulation results from the Bayesian approach appear to support the superiority of this technique over the parametric and semi-parametric methods in several dimensions. The shift point was determined quite accurately regardless of the curvature of the land value function, although the accuracy somewhat deteriorates in Cases 6 and 7, the two simulations where large parcels dominated the samples. Moreover, when the shift point is asymmetric as in Cases 8 and 9, the Bayesian technique is the only one able to accurately estimate δ.The Bayesian approach correctly identified both plottage and plattage in the data in well over 90 percent of the simulations when the shift point occurs in the middle of the parcel size distribution, except once again in Cases 6 and 7, where both plottage and plattage are given a posterior probability around 86-88%.[11] The Bayesian estimation, though, encounters some problems in assessing the probability of plottage when the shift point occurs at the very low end of the size distribution when there are just a handful of observations (ten, to be exact) on the convex portion of the curve. Increasing the sample size from 100 to 500 (or, to 250 for Cases 8 and 9) did not systematically improve or change the performance of any method in all cases. The only exception appears to be an increased accuracy of the Bayesian approach in identifying the inflection points and, especially, the probability of plottage n the harder Cases, 6-9.

6. Applications to Real Data

In addition to testing the accuracy of the three techniques for detecting plottage and plattage and for estimating their shift point using simulated data, a second contribution of this paper lies in applying the methods and models using two different databases. Empirical estimates of the inflection point with data from both Arizona and Iowa will allow for a comparison of model properties that would not be possible with just one database. The Arizona database contains information provided by Costar, Inc. on 295 land parcels in the Phoenix metro area that were sold in 1998 and 1999. An additional 127 improved single-family lot transactions were provided by Marketron, Inc. and were added to the database to increase the number of small parcels making the total sample size 422.[12] The Costar transactions ranged from 18,865 to 41,327,948 sq. ft. while the improved lots ranged from 3,851 to 65,993 sq. ft.

Dummy variables were added to the Arizona model specifications to control for land improvements reflected in the lot prices. The Finished Lot dummy controls for the land development costs reflected in the price of the lots relative to the raw land parcels while the Platted/Eng (engineered) dummy controls for intermediate stages of development where a final plat map has been recorded but land improvement costs have yet to be incurred. The ESL (environmentally sensitive land) dummy variable controls for the added development cost associated with some lots with ESL zoning. Finally, the Density variable picks up differences in zoning that might be reflected in lot prices. The typography of the Phoenix metro area has led to development that can generally be classified into the Southeast or Northwest Valleys. For this reason distances from each parcel to a central point in either the Southeast (Dist SE) or Northwest (Dist NW) Valley are used as covariates to measure proximity to the urban fringe.

The Iowa database used by EI (2005) contains 646 residential, arms-length sales of vacant land in and around Cedar Falls and Waterloo Iowa and was obtained from a public source. The sales span the period from 1980 to 2000 during which land values appreciated at a slow but steady rate. The location of the parcels is known because the state plane (x-y) coordinates of the centroid of each parcel were collected and geo-coded. Distances to the two CBDs are used as covariates. As in EI (2005), we also include a term for time of sale (see eq. (5) and (6) since the data spans 20 years. Table 4 contains summary statistics for the data used in the estimation.

7. Empirical Results

7.1 Results with Arizona Data

The results from the estimation procedures for Arizona are presented in Table 7a and Table 7b. The Arizona regression includes three dummy variables: ESL, Finished Lots and Platted/Engineered. The other covariates are Size, Density and two distance variables. Results for the Bayesian estimation are reported in table 7a while results for the parametric and semi-parametric estimations are reported in table 7b. Some parcels of different sizes have the same or virtually the same x-y coordinates and this causes problems in the matrix of spatial correlations. Therefore, for robustness purposes the results for two Bayesian estimations are reported. Arizona Bayesian 1 keeps the smallest parcel among the duplicate x-y coordinates while Arizona Bayesian 2 keeps the largest parcel among the duplicates. For the parametric and semi-parametric procedures we report standard errors below each parameter estimate, with the exception of the shift parameter, for which standard errors are not available. For the Bayesian procedures we apply the same model specification used with simulated data (see eq. (5), (6), (7), (9), and (10)) and the MCMC algorithm detailed in the Appendix, with the same priors described herein. From the estimation output we report the posterior mean and the posterior 95% confidence interval (in parenthesis).[13] For the semi-parametric estimator we use four different kernels: Gaussian, Quartic and Triangular using a fixed bandwidth equal to the rule of thumb ($n^{-0.2}$) and a Gaussian Kernel using a variable bandwidth (smoothing parameter). The algorithm used for the variable bandwidth kernel is explained in detail in the appendix section.

The explanatory power of the parametric and semi-parametric models is quite good. As opposed to what we find in the simulations, the Gaussian Kernel seems to find the inflexion point at small parcel size values while the Quartic and Triangular Kernels find it at very large sizes. This is because the behavior of the later kernels is very erratic when data is sparse given that they only take into account neighboring observations for the calculations and there might be very few or none in the end tails. We will show later in our figures that this is the case in our data. It is important to notice that adding a variable bandwidth to the Gaussian Kernel does not seem to improve the estimation very much. The null hypothesis of coexistence of plottage and plattage is not rejected by any semi-parametric estimation method. Results with the parametric and semi-parametric methods with the Gaussian Kernel are consistent with recent studies on urban land prices such as Colwell and Munneke (1997, 1999) but slightly lower than the R^2 reported by Thornes and McMillen (1998). The coefficients on the distance parameters have the correct sign in every estimation, their magnitudes are plausible and they are also fairly consistent between the parametric, semi-parametric and Bayesian approaches. As expected, the coefficient for size in the parametric model is highly significant; the coefficients for all other variables are also significant except for Finished Lots. Similar inferences can be drawn from the estimates of the regression coefficients in the semi-parametric model.

The two Bayesian specifications yield very similar posterior quantities (Table 5, last four columns), indicating that keeping the smallest or the largest parcel within duplicate x-y coordinates does not alter inferences in any appreciable manner. The estimates for the parameter ϕ imply a range of effective spatial correlation[14] with a posterior mean around 65,000 ft and a 95% confidence interval between about 35,000 and about 120,000 ft. Based on the posterior mean, parcels within approximately 12 miles of each other are spatially correlated.

The maximum R^2 for the parametric model occurs at a shift point (*s*) of 47,500 SF and the J-test cannot reject a hypothesis about the coexistence of plottage and plattage. A different situation is observed for the semi-parametric estimation with Gaussian Kernel and the two Bayesian estimates where the data indicate a shift point at 11,069 SF (Gaussian kernel and fixed bandwidth), 10,406 SF (Gaussian kernel and variable bandwidth) and either 11,457 SF (Bayesian 1) or 12,137 SF (Bayesian 2), with fairly tight confidence bounds. The probability of plottage and plattage are quite high (>90%) for both Bayesian estimates.

Finally, the semi-parametric estimates of $g\left(A_i\right)$ are graphed in Figure 3.a and while the fitted values for the parametric estimation and the fitted values for the Bayesian estimation are graphed in Figure 3.b.[15] We plot $g(A_i)$ using the four different Kernels. The curves are quite irregular for the different kernels. The graph of $g(A_i)$ using a Gaussian Kernel appears to reflect two inflection points, one for very small parcels and the other for larger parcels. Using a variable bandwidth with the Gaussian Kernel does not change the conclusions in any appreciable way. The Quartic and Triangular Kernel do not seem to fit the data very well producing an almost linear relationship at the small end of the sample and an erratic behavior at the large end of the sample when the land size is large and data is very sparse. The fitted values from the parametric and Bayesian methods force an S-shaped price-size relation since a specific functional form is imposed.

7.2 Results with Iowa Data

The Iowa data includes Time, Size and two CBD distances (Cedar Falls and Waterloo) as covariates and the results from the estimation are reported in tables 8a and 8b for all three estimation procedures. For comparability between the original EI (2005) model and the slightly modified one used here in the simulation study, the Bayesian method is applied to three alternatives. Namely, the original EI (2005) model specification (see eq. (5) through (8) above) is estimated along with the modified model discussed earlier for the h(.) function (see eq. (9) and (10)) and two different assumptions about prior beliefs. The Bayesian estimation results are presented in table 8a while the parametric and semi-parametric using four different kernels as in the previous section are presented in table 8b. In Table 8a, Bayesian 1 refers to the modified model proposed in this paper with the relatively less informative priors about the shift parameter and the convexity/concavity parameters specified in the Appendix. Bayesian 2 is the original EI (2005) model but with the less informative priors used in Bayesian 1, and Bayesian 3 is the original EI (2005) model estimated with the same (relatively more informative) priors EI adopt for the inflection point and for the λs. In other words, Bayesian 3 wants to mimic EI (2005) as closely as possible: same data, same model specification, same priors. Bayesian 2 should allow to assess prior sensitivity when compared to Bayesian 3, while Bayesian 1 should allow to assess the sensitivity to the specification of the concave-convex (i.e., the h()) function.

All three approaches find evidence of the coexistence of plottage and plattage in the data. For the parametric model, the J-test cannot reject the coexistence of plottage and plattage while for all three Bayesian estimates the probabilities plattage are very high and similarly so across estimates. The posterior probability of plottage is estimated as close to 89% by the model we propose, while it falls to about 74% in the EI (2005) specification. The smallest inflection point is found by the semi-parametric model at 1,323 SF and 1,674 SF when using a Gaussian Kernel with fixed and variable bandwidth respectively. It is quite close to what suggested by the parametric approach and by two of the three Bayesian estimates, while a relatively larger shift point is found by the original EI model and priors at 2,753 SF. It is noteworthy that both the location of the shift point and, especially, the 95% CI around it reflect quite closely the posterior summaries presented by EI (2005) for the same Iowa dataset. The similarities in the results between Bayesian 1 and Bayesian 2 point to a fairly minor effect played by the specification of the h(.) function. In other words, either the one adopted in EI (2005) or the one we use in our simulation study appear to lead to very similar inferences about the existence and location of plottage and plattage. On the other hand, the, albeit not large, discrepancy in the inflection point estimates between Bayesian 2 and Bayesian 3 may indicate some relatively minor sensitivity to the prior specification for δ and/or for the λs. This is not surprising given that the sample of small parcels is typically quite small. When using the semi-parametric approach, the Quartic and Triangular kernel fail again when compared to the Gaussian by finding the inflection point at large, and quite unreasonable, values of parcel sizes.

Finally, the semi-parametric estimates of $g\left(A_i\right)$ and the fitted values of the parametric and the Bayesian estimations for the Iowa sample are graphed in Figure 4.a. An inflection point can be observed for very small parcels and after that the relationship between land area and value remains fairly constant until one reaches the largest parcels. In figure 4.b we show the fitted values for the parametric and Bayesian procedures which display a similar pattern to the one observed in 4.a.

While the range of inflection points estimated from the Iowa data across all procedures is fairly narrow, the same cannot be said for the Arizona data. The inflection point from the parametric model is an "outlier" compared to the semi-parametric and Bayesian estimates raising the question of whether the estimate differs so substantially because of problems with the Arizona data or the parametric technique itself. Both the Arizona and Iowa data bases contain far fewer sales of small parcels making them closer to the simulation data used for Cases 6 and 7 (Tables 2 and 3). The parametric technique was able to correctly predict the presence of plottage and plattage roughly 50 percent of the time or less in Cases 6 and 7, although in one case the estimated shift point, 15.27 was quite close to the value used to generate the data, δ=15. This would suggest that the empirical results with the Arizona dataset are an example of the unreliability of the parametric technique in some situations. Although in simulations the Quartic and Triangular Kernels seemed to work marginally better than the Gaussian kernel, real data analysis shows that the Gaussian kernel may be superior in some instances. This is likely because the other kernels are negatively affected by the presence of very sparse observations as it is the case in real data. Finally, using a variable bandwidth with the Gaussian kernel leads to slightly smaller inflection points than using a fixed kernel, but the difference is not substantial and results are very similar.

8. Conclusions

There is an extensive literature on the relationship between parcel size and land price with most researchers finding evidence that the relationship is non-linear, especially with respect to larger parcels. What has been lacking until now is an assessment of whether the different modeling and econometric techniques that have been used in empirical tests can accurately and reliably detect convexity (plottage) and concavity (plattage) if it is present in data over the full range of parcel sizes. The primary contribution of this paper is to test and compare the parametric, semi-parametric and Bayesian techniques on simulated data over a wide range of possible shapes for the land value function. The parametric model was quite accurate for the original Colwell-Sirmans specification of the parameters but failed to accurately or reliably detect plottage and plattage for all other sets of parameter specifications. The semi-parametric model was considerably more accurate than the parametric one in identifying the location of the shift point, except for the simulations where small parcels constituted only ten percent of the sample, which more closely approximates the size distribution of the data used in most empirical studies. Within the semi-parametric technique, the Gaussian kernel outperformed other types of kernels when using real data but underperforms with simulated data. The Bayesian technique was the most accurate in estimating the correct inflection point and could also detect both plottage and plattage over 85 percent of the time in seven of the nine simulation settings. It, too, was less accurate with the datasets containing only a limited number of small parcels but it still performed generally better than either the parametric or semi-parametric models. Overall, the simulation results demonstrate the higher reliability of the Bayesian technique for detecting plottage and plattage and for accurately estimating the inflection point, especially in the cases in which the shift point was in the tail of the sample size, which is expected to be the case in real data. The only instances where the Bayesian approach displayed some difficulties were the assessments of plottage probabilities in cases characterized by both an asymmetric shift point and a relatively small sample size (one hundred observations in our experiments).

The empirical results for the three different techniques were consistent using the Iowa data but the same cannot be said using the Arizona data. The original Ecker-Isakson results for Iowa were supplemented with estimates from the parametric and semi-parametric models. Based on the simulation results, it was interesting to see if those two models produced inflection points consistent with the Bayesian estimates, which they did for the Iowa data. However, the Arizona results were mixed with a much higher estimated inflection point from the parametric model. Given the reliability of the Bayesian approach in the simulations, the discrepancy between the Bayesian and the parametric estimates of the inflection point in the Arizona data suggests further that the latter approach may experience problems in certain situations.

Appendix 1: Bayesian Estimation Method

In this Appendix we detail the implementation of the Bayesian MCMC procedure adopted in the simulation exercise and in most of the analysis on real data. We also specify the, albeit slight, differences between the implementation we adopt and the one originally proposed by Ecker and Isakson (EI 2005). As reported in the main text, the differences reside in the less informative specifications we adopt for the prior distributions on some of the parameters relatively to EI (2005).

Combining the models for small and large parcels in eq. (5) and (6) and combining the observations in vector form one obtains (see eq (6) in EI (2005))

$$\log(Y) \equiv (\log(Y_s), \log(Y_l))` = X\gamma + e \qquad \text{(A1)}$$

with $e \sim \mathrm{N}(0, \Sigma)$

and where

Y_s and Y_l are vectors containing all observations on, respectively, small and large parcels,

$X = \begin{pmatrix} X_s & 0 \\ 0 & X_l \end{pmatrix}$, X_s and X_l are matrices whose first column is a vector of ones and containing in the remaining columns observations on a set of covariates, such as time and distance and including the h(.) functions, for small and large parcels and with the *0*s being conformant matrices,

$\gamma = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}$, with α and β being vectors that collect the regression parameters, and Σ is defined is by eq. (7a), (7b) and (7c) in the main text.

As suggested by EI (2005), the above model can be seen as a random effects spatial model. Namely,

$$\log(Y) = X\gamma + W + u \tag{A2}$$

with $u \sim N(0, \tau^2 I_N)$ and independent of W,

where $W \sim N(0, \Sigma)$ is the spatial random effect and N is the combined sample size (i.e., the total number of parcels).

Marginalizing over W the model can be written as

$$\log(Y) \sim N(X\gamma, \Omega) \tag{A3}$$

where $\Omega \equiv \tau^2 I_N + \sigma^2 H(\phi)$

Prior Distributions

Following EI (2005) we use independent priors:

$$\pi(\gamma, \lambda_s, \lambda_l, \delta, \phi, \tau^2, \sigma^2) = \pi(\gamma)\pi(\lambda_s)\pi(\lambda_l)\pi(\delta)\pi(\phi)\pi(\tau^2)\pi(\sigma^2)$$

We make the following distributional choices and select rather uninformative hyper-parameters:

$$\pi(\gamma) \sim N(0, \Gamma_0)$$
$$\pi(\lambda_s) \sim U(0,4)$$
$$\pi(\lambda_l) \sim U(0,1.5)$$
$$\pi(\delta) \sim U(\min_A, \max_A)$$
$$\pi(\phi) \sim IG(\frac{l_0}{2}, \frac{m_0}{2})$$
$$\pi(\tau^2) \sim IG(\frac{c_0}{2}, \frac{d_0}{2})$$
$$\pi(\sigma^2) \sim IG(\frac{a_0}{2}, \frac{b_0}{2})$$

where N denotes the normal distribution, U denote the uniform distribution and IG the inverse gamma. The prior on γ is essentially flat as Γ_0 is set to $10^6 I_r$, where I is the identity matrix and r is the dimension of γ. The parameters for the inverse gamma priors on σ^2 and τ^2 are selected so that the prior mean equals the OLS estimate of the error variance in a linear regression of log-size on all the covariates, while the prior variance is infinite (i.e., $d_0 = b_0 = 4$). All these choices are very close, if not identical, to those in EI (2005) for the corresponding parameters. For the spatial correlation parameter, ϕ, we also rely on an inverse gamma prior to ensure positivity. As EI (2005) suggest, we impose an infinite prior variance (i.e., $m_0 = 4$) and we set a prior mean so that the implied mean on the range of effective correlation (which is given by $3/\phi$, see EI (2005)) equals 1/10 of the maximum simulated distance across parcels. We use uniform priors on the concavity-convexity parameters, λ_s and λ_l: our choice is less informative than the beta priors adopted by EI (2005). Still, it assigns a 75% probability to the existence of both plottage and plattage. The flat prior for δ over the range of parcel sizes is, perhaps, the largest departure from EI (2005). They select an inverse gamma prior with infinite variance but with a mean around 5,000 ft^2 for the Iowa dataset. Given the crucial role played by the shift parameter, we feel more inclined towards our completely diffused prior.

In the applications to real data, the only change we make to the prior hyper-parameters listed above is for ϕ, where we follow EI (2005) and set the prior so that the prior mean on the range of effective correlation is about 1,100 ft. We also experiment with a completely diffused prior for ϕ on the (0,1) range. No relevant changes in the posterior distributions were noted.

Posterior Sampling

Combining the chosen priors with the sampling distribution in (A2) or (A3) it follows that the posterior distributions for the parameters $\gamma, \lambda_s, \lambda_l$ and δ (the shift point) can all be sampled conditioning on the data and on Ω, while the random effects can be sampled from a multivariate normal distribution. Then, conditioning on W, one can sample the posteriors for σ^2 and ϕ. Finally, conditioning on the data, random effects and remaining parameters, one can sample the posterior for τ^2 from standard regression updates. The steps of the MCMC chain are summarized as follows:

1. Initialize $\lambda_s, \lambda_l, \delta, \phi, \sigma^2, \tau^3, \gamma$

2. Sample γ from $\gamma \,|\, Y, X, \Omega, \delta$

3. Sample λ_s, λ_l from $\lambda_s, \lambda_l \,|\, Y, X, \Omega, \gamma, \delta$

4. Sample δ from δ|Y, X, $\Omega, \gamma, \lambda_s, \lambda_l$

5. Draw W from $W \,|\, Y, X, \gamma, \phi, \sigma^2, \tau^2$

6. Draw σ^2 from $\sigma^2 \,|\, W, \phi, \gamma$

7. Draw ϕ from $\phi \,|\, W, \sigma^2$

8. Draw τ^2 from $\tau^2 \,|\, Y, X, W, \gamma$

9. Go to step 2 and repeat.

In step 2, the update for γ is of a standard form from the regression in eq. (A3) and it is, thus, a multivariate normal distribution. Similarly, the posterior in step 8 is a standard inverse gamma distribution from the regression in eq. (A2).

In step 5, the posterior for W is multivariate normal (see, e.g., Ecker and Gelfand (2003)) with covariance matrix $W = [\tau^{-2} I_N + \sigma^{-2} H(\phi)^{-1}]^{-1}$ and mean vector $W(Y - X\gamma)/\tau^2$. In step 6, the posterior is, in a standard way, inverse gamma with parameters $(a_0 + N)/2$ and $(b_0 + W`H(\phi)^{-1}W)/2$.

For the remaining steps (3, 4 and 7) the posterior are of non-standard form and, therefore, need to be sampled via the Metropolis-Hastings (MH) algorithm (see, among others, Chib and Greenberg (1995)). This involves generating a candidate value from a known density and, then, accept it with an easily computable probability, More specifically, for each of the three posteriors we use a random walk MH step where a candidate draw is generated using a normal distribution centered on the (log-transformed) current value and whose spread is tuned in order to achieve a good trade-off between acceptance rate and autocorrelation in the draws. After experimentation on both the simulated and the actual data, we find that the chain appears to mix well (low autocorrelation across draws) when the individual acceptance rates are between 40% and 50%.
In our applications on simulated data, we cycle through steps 2 through 8 for 25,000 iterations. We discard the initial 5,000 “burn-in” draws and retain the remaining draws for inferential purposes. As convergence diagnostics we use the inefficiency factors (see Chib, Nardari and Shephard (2006) for a description and application). We find that with the numbers of Gibbs iterations mentioned above, the chain mixes fairly well, with inefficiency factors ranging from below 10 for the α_0 through α_3 parameters, to around 20 for the τ^2 and β parameters, to about 55-60 for δ, α_4 and σ^2, to 85-90 for ϕ and the λs. We also experiment with retaining every 5 draws and obtain very similar posterior inferences. For the smaller sample sizes (100 obs), we check robustness by using 25,000 burn-in draws and collecting the subsequent 50,000 draws.

In our applications to real data, we experiment with simulation sizes between 25,000 (5,000 burn-in) and 200,000 (50,000 burn-in) draws. The inefficiency factors appear to be satisfactory even with the more conservative choices. We choose to report the results obtained with 50,000 collected draws (10,000 burn-in).

Appendix 2: Variable Bandwidth Kernel Algorithm

The algorithm for the variable bandwidth kernel implemented is the one developed in Silverman (1986) and Fox (1990). The implementation follows the description in Salgado-Ugarte and Perez-Hernendez (2003) and consists of three steps:

1- The first step consists in calculating the density estimate $\widehat{f}(x)$ using a fixed bandwidth. We followed the rule of thumb to pick the fixed bandwidth size, which is set equal to $n^{-0.2}$.

2- Then we calculate a local window factor for each observation X_k equal to $w_k = \left\{ \frac{\left[\prod_{i=1}^{n} \widehat{f}(X_i)\right]^{1/n}}{\widehat{f}(X_k)} \right\}^{0.5}$ where $\widehat{f}(X_k)$ is the density estimate in the first step.

3- Finally, using the weights estimated in the second step we estimate the final kernel estimator $\tilde{f}(x) = \frac{1}{nh}\sum_{i=1}^{n}\frac{1}{w_i}K\left(\frac{x-X_i}{w_i h}\right)$.

Table 1: kernels

Gaussian	$\frac{1}{2\pi}e^{-\frac{u^2}{2}}$
Quartic	$\frac{15}{16}(1-u^2)^2 I(\lvert u\rvert \le 1)$
Triangular	$\frac{15}{16}(1-\lvert u\rvert)I(\lvert u\rvert \le 1)$

Table 2: Simulation Setup

a) Main parameters

	Case 1	Case 2	Case 3	Case 4	Case 5	Case 6	Case 7	Case 8	Case 9
λ_s	1/3	2/5	5/7	5/6	5/7	5/6	5/7	5/7	5/7
λ_l	1/3	5/9	5/6	5/6	5/6	5/6	5/6	5/6	5/6
δ	15	15	15	15	15	15	15	7	7
Percentage of small parcels	50%	50%	50%	50%	50%	10%	10%	50%	10%

b) Secondary parameters

	Case 1	Case 2	Case 3	Case 4	Case 5	Case 6	Case 7	Case 8	Case 9
α_1	0.30	0.30	0.30	0.30	0.30	0.30	0.30	0.30	0.30
α_2	0.15	0.15	0.15	0.15	0.15	0.15	0.15	0.15	0.15
α_3	0.12	0.12	0.12	0.12	0.12	0.12	0.12	0.12	0.12
θ_1	0.30	0.30	0.30	0.30	0.30	0.30	0.30	0.30	0.30
θ_2	0.15	0.20	0.20	0.20	0.20	0.20	0.20	0.20	0.20
θ_3	0.12	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.17
τ^2	.08436	.08436	.08436	.08436	.8436	.08436	.8436	.8436	.8436
ϕ	.15	.15	.15	.15	.15	.15	.15	.15	.15
σ^2	.5512	.5512	.5512	.5512	.5512	.5512	.5512	.5512	.5512
Min($CBD1_i$)	0	0	0	0	0	0	0	0	0
Max($CBD1_i$)	10	10	10	10	10	10	10	10	10
Min($CBD2_i$)	0	0	0	0	0	0	0	0	0
Max($CBD2_i$)	15	15	15	15	15	15	15	15	15
Min(A_i)	5	5	5	5	5	5	5	5	5
Max(A_i)	25	25	25	25	25	25	25	25	25

Table 3

Parametric Model Simulations' Results

Case	Parametric									
	Small Sample					Large Sample				
	Mean Shift	**Stdv Shift**	**RMSE Shift**	**J-Test Rejections (%)**	**R2**	**Mean Shift**	**Stdv Shift**	**RMSE Shift**	**J-Test Rejections (%)**	**R2**
1	15.23	4.90	4.90	12.30	0.58	15.45	3.20	3.30	6.00	0.57
2	17.46	6.20	6.70	20.50	0.62	18.50	6.00	7.00	25.00	0.61
3	24.64	8.70	13.00	75.00	0.74	30.58	2.50	15.80	100.00	0.73
4	20.98	11.90	13.40	87.60	0.76	27.58	9.10	15.50	100.00	0.75
5	18.61	7.00	7.90	28.00	0.66	17.85	6.00	6.60	54.00	0.65
6	10.57	8.40	9.50	61.00	0.75	6.88	5.80	10.00	99.00	0.74
7	15.27	9.50	9.50	48.00	0.57	12.01	9.60	9.60	82.00	0.59
8	10.19	6.64	7.37	46.20	0.68	9.09	5.86	6.22	92.60	0.67
9	9.23	7.75	8.06	78.00	0.63	6.67	5.78	5.79	99.80	0.62

Table 4

Semi-Parametric Model Simulations' Results

a) Gaussian Kernel

Case	Small Sample					Large Sample				
	Mean Shift	Stdev Shift	RMSE Shift	Rejections (%)	R2	Mean Shift	Stdev Shift	RMSE Shift	Rejection (%)	R2
1	14.88	4.17	4.17	0.10	0.68	16.75	1.09	2.06	0.00	0.66
2	15.00	4.53	4.53	0.20	0.72	17.04	1.43	2.49	0.10	0.70
3	15.66	4.85	4.93	5.50	0.81	18.15	2.67	4.13	11.70	0.80
4	15.59	4.75	4.78	11.10	0.83	18.06	2.51	3.95	37.10	0.82
5	14.07	5.35	5.43	3.70	0.67	17.26	4.02	4.61	7.10	0.65
6	12.70	3.88	4.51	26.40	0.75	12.86	1.28	2.50	18.50	0.73
7	11.86	4.09	5.15	17.60	0.57	12.03	2.43	3.84	6.20	0.55
8	10.90	3.39	5.17	12.70	0.83	9.78	2.35	3.64	35.00	0.82
9	10.80	4.34	5.77	12.60	0.64	10.47	4.35	5.56	32.80	0.63

b) Quartic Kernel

	Small Sample					Large Sample				
	Mean Shift	Stdev Shift	RMSE Shift	Rejection (%)	R2	Mean Shift	Stdev Shift	RMSE Shift	Rejection (%)	R2
1	15.26	3.27	3.28	0.00	0.35	18.07	1.55	3.44	1.50	0.32
2	13.65	2.97	3.26	0.00	0.30	16.74	2.42	2.98	0.00	0.27
3	12.29	1.63	3.16	0.00	0.18	14.38	0.75	0.97	0.00	0.14
4	12.33	1.72	3.18	0.00	0.15	14.38	0.76	0.98	0.00	0.13
5	21.69	1.77	6.92	2.80	0.26	14.57	1.19	1.26	0.00	0.12
6	13.28	2.65	3.15	0.00	0.37	14.80	0.17	0.26	0.00	0.34
7	13.32	2.71	3.19	0.00	0.30	14.80	0.17	0.26	0.00	0.27
8	13.60	3.40	7.42	2.70	0.12	14.45	2.48	7.85	0.40	0.13
9	12.74	2.29	6.18	0.70	0.20	13.72	1.45	6.87	0.10	0.20

c) Triangular Kernel

	Small Sample					Large Sample				
	Mean Shift	**Stdev Shift**	**RMSE Shift**	**Rejection (%)**	**R2**	**Mean Shift**	**Stdev Shift**	**RMSE Shift**	**Rejection (%)**	**R2**
1	11.39	0.88	3.71	1.00	0.07	13.66	0.19	1.36	0.00	0.05
2	11.39	0.88	3.71	1.30	0.06	13.66	0.20	1.36	0.00	0.04
3	11.40	0.88	3.71	11.20	0.04	13.66	0.19	1.36	6.10	0.03
4	11.40	0.88	3.71	11.30	0.04	13.66	0.19	1.36	7.40	0.03
5	11.39	0.88	3.71	10.60	0.04	13.66	0.19	1.36	5.60	0.02
6	14.78	2.87	2.88	15.20	0.25	14.88	0.28	0.31	0.10	0.14
7	14.76	2.83	2.84	14.70	0.21	14.88	0.33	0.35	0.10	0.11
8	11.37	1.22	4.53	0.10	0.08	12.68	0.97	5.76	0.00	0.03
9	11.36	0.85	4.45	6.90	0.05	12.61	0.47	5.63	4.00	0.03

Table 5

Bayesian Model Simulations Results

	Bayesian									
	Small Sample					Large Sample				
Case	**Mean Shift**	**Stdv Shift**	**RMSE Shift**	**Prob Plot**	**Prob Plat**	**Mean Shift**	**Stdv Shift**	**RMSE Shift**	**Prob Plot**	**Prob Plat**
1	14.98	2.30	2.30	0.96	0.97	14.99	0.20	0.10	0.98	0.98
2	15.05	2.40	2.40	0.97	0.96	14.87	0.30	0.10	0.98	0.98
3	14.69	2.60	2.60	0.96	0.97	14.92	0.70	0.50	0.97	0.91
4	14.85	2.90	2.90	0.96	0.96	14.98	0.70	0.50	0.92	0.91
5	15.35	4.20	4.20	0.95	0.92	14.78	1.20	1.60	0.98	0.96
6	13.40	4.90	5.10	0.87	0.89	15.30	4.20	17.90	0.82	0.87
7	13.87	5.20	5.30	0.92	0.87	14.92	5.50	30.70	0.86	0.88
8	7.25	2.40	2.42	0.68	0.99	7.26	1.67	1.69	0.88	0.99
9	8.20	4.38	5.81	0.69	0.99	7.89	3.05	5.33	0.76	0.92

Table 6 Descriptive Statistics for the Arizona and Iowa Data

Panel A

	Arizona							
	Price	Size (Square feet)	Distance 1	Distance 2	Density	*ESL	*Finished Lot	*Platted Engineered
N	422	422	422	422	422	422	422	422
Mean	$886,190	1,386,424	23.06	18.34	2.4727	0.0972	0.4028	0.0474
Standard Deviation	$1,605,307	2,759,258	13.23	7.56	2.0891	0.2965	0.4911	0.2127
Maximum	$15,809,589	20,803,385	91.01	69.24	10.62	1	1	1
Minimum	$17,000	4,100	0	0.67	0.20	0	0	0

*Dummy variables

Panel B

	Iowa				
	Price	Time (Years)	Size (Square feet)	Distance Cedar Falls (Feet)	Distance Waterloo (Feet)
N	646	646	646	646	646
Mean	$27,574	14.33	32,223	26,464	19,298
Standard Deviation	$34,011	5.65	190,000	14,434	11,201
Maximum	$500,000	20.56	4,623,023	58,870	43,983
Minimum	$1,000	0	640	298	1,474

Table 7a Empirical Results with the Arizona Data

	Arizona Bayesian 1		Arizona Bayesian 2	
	Plottage	Plattage	Plottage	Plattage
Intercept	5.8143 (-38.63, 49.82)	6.6857 (3.02, 9.22)	5.5982 (-38.36,49.61)	6.35 (3.66,9.44)
Distance SE	-0.0350 (-0.0827,0.0127)	-0.0123 (-0.0296,-0.0047)	-0.0363 (-0.0822,0.0087)	-0.0104 (-0.0177,-0.0030)
Distance NW	-0.0300 (-0.1307,0.0718)	-0.0313 (-0.0439,-0.0186)	-0.02114 (-0.0892,0.0484)	-0.0319 (-0.0441,-0.0195)
Size	0.4093 (-4.9413,7.5802)	2.072 (0.6344,4.6467)	0.8849 (-11.97,15.39)	2.2134 (0.6434,4.0205)
Density	0.0142 (-0.4348,0.4634)	0.1147 (0.0844,0.1450)	0.2649 (-1.1195,1.6519)	0.1279 (0.0982,0.1577)
ESL	-0.0701 (-62.14,62.42)	0.0939 (-0.1385,0.3267)	-0.2271 (-61.79,62.37)	0.1131 (-0.1102,0.3397)
Finished Lot	5.8919 (-38.01,49.98)	0.3465 (0.1305,0.5620)	5.4339 (-38.47,49.40)	0.3785 (0.1640,0.5905)
Platted Eng.	0.1580 (-61.92,61.95)	0.3761 (0.1177,0.6367)	-0.03497 (-62.21,62.09)	0.3326 (0.0810,0.5815)
Shift (s)	**11457** **(11073,13705)**		**12137** **(11111,13941)**	
R^2	___		___	
J-test	___		___	
Min/Max	3,851 SF/ 41,327,985 SF			
λ	1.5746 (1.0527,1.9288)	0.1051 (0.0613,0.1461)	1.5261 (1.0894,1.9499)	0.0947 (0.0667,0.1453)
ϕ	0.000053 (0.000026,0.000105)		0.000045 (0.000024,0.000083)	
τ^2	0.2175 (0.1724,0.2667)		0.2133 (0.1753,0.2562)	
σ^2	0.2589 (0.1313,0.4813)		0.2178 (0.1154,0.3892)	
Prob Plot/Plat	0.9192	0.9932	0.9017	1

Table 7b Empirical Results with the Arizona Data

	Arizona Parametric	Arizona Semi- parametric			
		Gaussian fixed bandwidth	Gaussian variable bandwidth	Quartic Kernel	Triangular Kernel
Intercept	12.96 (0.11)	___	___	___	___
Distance SE	-0.0267 (0.0022)	-0.0237 (0.0021)	-0.0219 (0.0024)	-0.0186 (0.0050)	-0.0269 (0.0045)
Distance NW	-0.0426 (0.0043)	-0.0356 (0.0041)	-0.0337 (0.0044)	-0.0172 (0.0091)	-0.0404 (0.0082)
Size	0.0173 (0.0007)	___	___	___	___
Density	0.0522 (0.014)	0.0873 (0.0155)	0.0776 (0.0175)	0.0775 (0.0338)	0.0961 (0.0307)
ESL	0.4491 (0.102)	0.4368 (0.0963)	0.4567 (0.1175)	0.2896 (0.2458)	0.3093 (0.2253)
Finished Lot	0.0232 (0.101)	0.1846 (0.1213)	0.2923 (0.1341)	-1.4273 (0.1571)	-0.8695 (0.1537)
Platted Eng.	0.4271 (0.142)	0.4130 (0.1334)	0.4514 (0.1490)	0.8308 (0.3087)	0.4256 (0.2862)
Shift (s)	**47500**	**11069****	**10406****	**13918988****	**9422028****
R^2	.79	.82	.78	.49	.58
J-test	0.45	___	___	___	___
Min/Max	3,851 SF/ 41,327,985 SF				
λ	___	___	___	___	___
ϕ	___	___	___	___	___
τ^2	___	___	___	___	___
σ^2	___	___	___	___	___
Prob Plot/Plat	___	___	___	___	___

** For the semi-parametric case the shift reported is the size at which the second derivative of $g(A)$ becomes negative for the first time.

Table 8a Empirical Results with the Iowa Data

	Iowa Bayesian 1		Iowa Bayesian 2		Iowa Bayesian 3	
	Plottage	Plattage	Plottage	Plattage	Plottage	Plattage
Intercept	14.39 (-50.59,57.11)	-5.2363 (-40.85,8.70)	-12.07 (-3163,34.70)	8.41 (6.20,10.60)	5.54 (-54.77,39.25)	8.38 (6.21,10.48)
Time	0.0437 (-1.0719,2.1992)	0.0671 (0.0521,0.0821)	-0.2250 (-0.8765,1.4795)	0.0682 (0.0533,0.0829)	0.0516 (-1.1315,2.3689)	0.0681 (0.0533,0.0830)
Distance Cedar Falls	-0.000061 (-0.0009,0.0013)	-0.000027 (-.00007,.00001)	-0.00024 (-.0009,.0008)	-0.000025 (-.000066,.000013)	-0.000075 (-.000963,.00143)	-0.000025 (-.000064,.000012)
Distance Waterloo	-0.000007 (-0.0002,0.00023)	0.000003 (-.00004,.00005)	-0.000034 (-.00029,.00022)	0.000006 (-.000043,.000058)	-0.000026 (-.000281,.000238)	0.000007 (-.00004,.000057)
Size	0.2695 (-0.1068,1.1924)	0.1306 (0.1206,0.3791)	0.0154 (-0.0013,0.1270)	0.0231 (0.0011,0.0983)	0.0085 (-0.0023,0.0802)	0.0224 (0.00126,0.0971)
Shift (s)	**1867 (1535,2445)**		**2012 (1552,2643)**		**2753 (2527,4280)**	
R^2	________________		________________		________________	
J-test	________________		________________		________________	
Min/Max	640 SF / 4,623,023 SF					
λ	2.0820 (1.2475,2.8619)	0.1042 (0.0046,0.2685)	1.0496 (0.5206,1.7089)	0.2934 (0.1056,0.6951)	1.1103 (0.5808,1.6184)	0.2916 (01055,0.6840)
ϕ	0.00098 (0.000039,0.000221)		0.000094 (0.000035,0.000230)		0.000100 (0.000038,0.00023)	
τ^2	0.8601 (0.7467,0.9846)		0.8554 (0.7419,0.9781)		0.8561 (0.7398,0.9808)	
σ^2	0.7719 (0.3548,1.5774)		0.7997 (0.3491,1.7126)		0.7726 (0.3434,1.5827)	
Prob Plot/Plat	0.8893	0.9979	0.7407	0.9995	0.7477	0.9952

Table 8b Empirical Results with the Iowa Data

	Iowa Parametric	Iowa Semi- parametric			
		Gaussian fixed bandwidth	Gaussian variable bandwidth	Quartic Kernel	Triangular Kernel
Intercept	8.610 (0.21)	___	___	___	___
Time	0.080 (0.007)	0.080 (0.007)	0.080 (0.007)	0.081 (0.016)	0.084 (0.0098)
Distance Cedar Falls	-0.0000219 (0.000003849)	-0.0000219 (0.0000038)	-0.0000215 (0.0000038)	-0.0000311 (0.00000792)	-0.0000268 (0.00000494)
Distance Waterloo	0.00000218 (0.000005096)	0.00000313 (0.00000532)	0.00000333 (0.00000531)	-0.0000182 (0.00000975)	-0.0000102 (0.00000635)
Size	0.016 (0.003471599)	___	___	___	___
Shift (s)	**2,618**	**1,674****	**1323****	**85813****	**25688****
R^2	.271	.277	.275	.228	.183
J-test	.30	___	___	___	___
Min/Max	640 SF / 4,623,023 SF				
λ	___	___	___	___	___
ϕ	___	___	___	___	___
τ^2	___	___	___	___	___
σ^2	___	___	___	___	___
Prob Plot/Plat	___	___	___	___	___

** For the semi-parametric case the shift reported is the size at which the second derivative of $g(A)$ becomes negative for the first time.

Figure 1: Plottage and Plattage

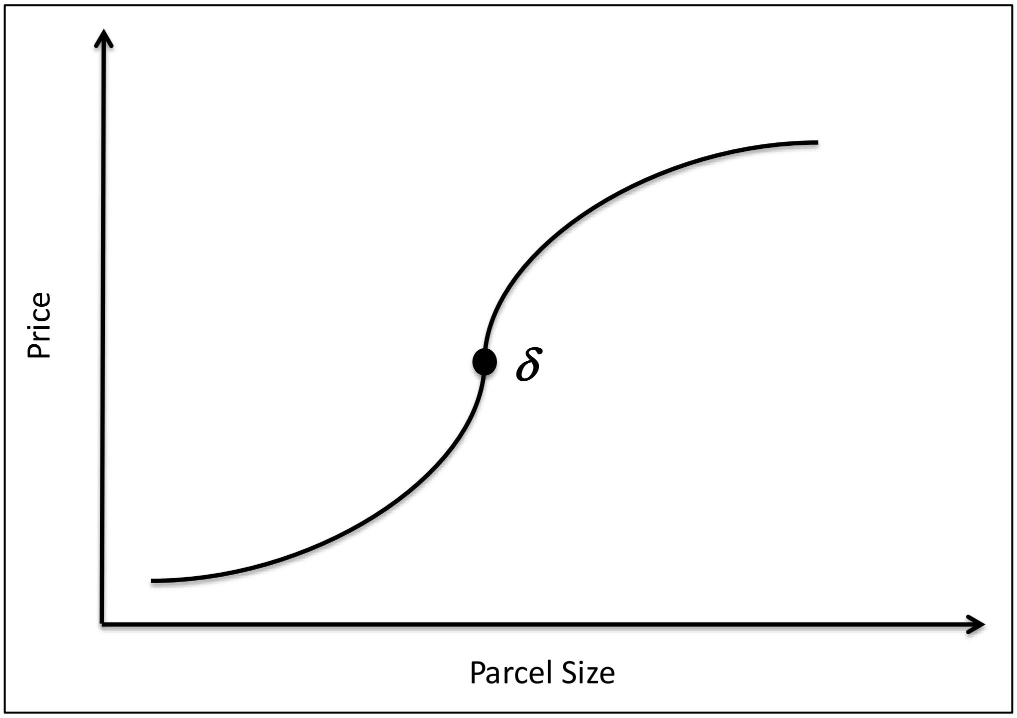


Figure 2 Shape of the Land Value Curve for Values of λ_s and λ_l

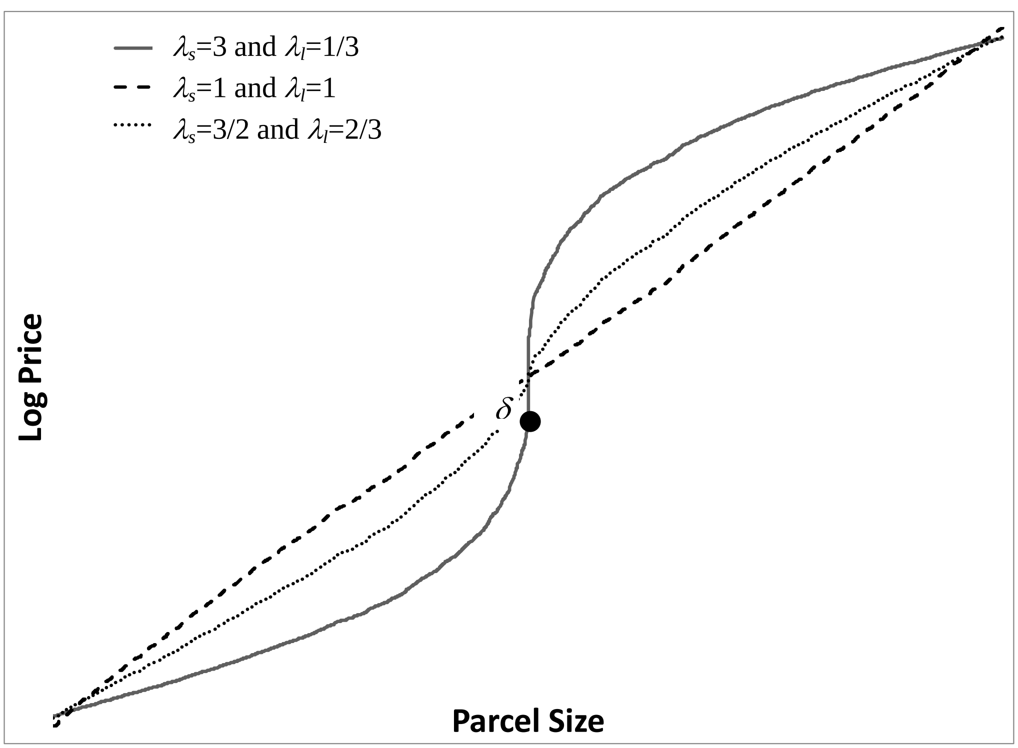
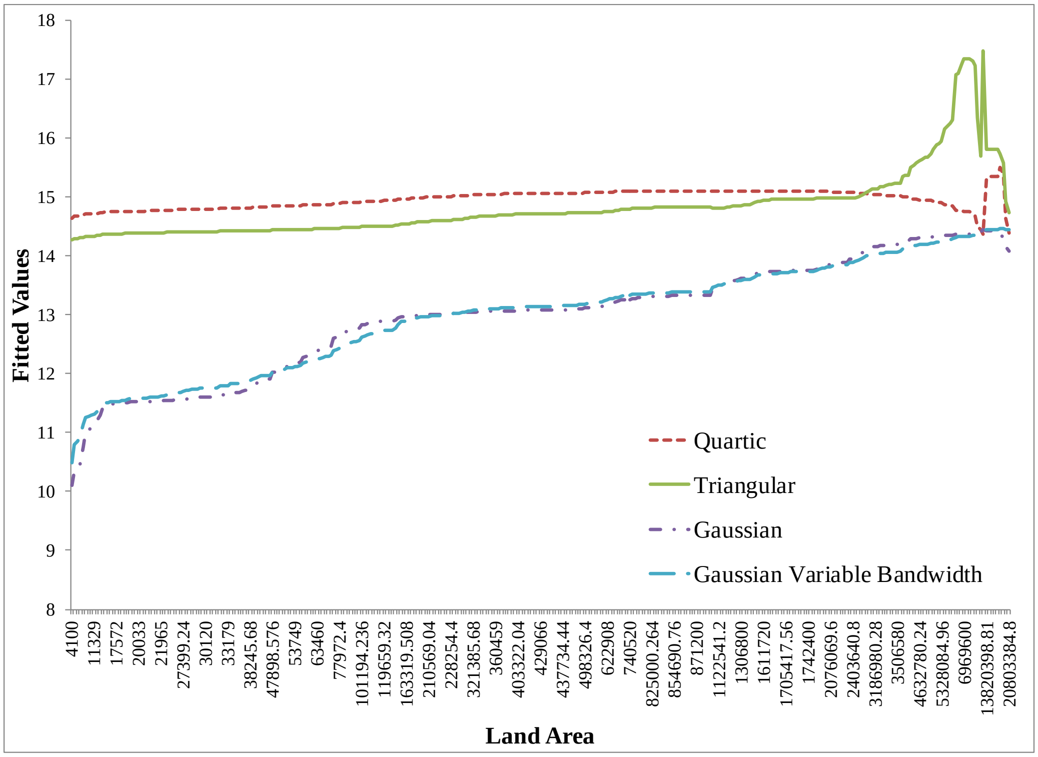
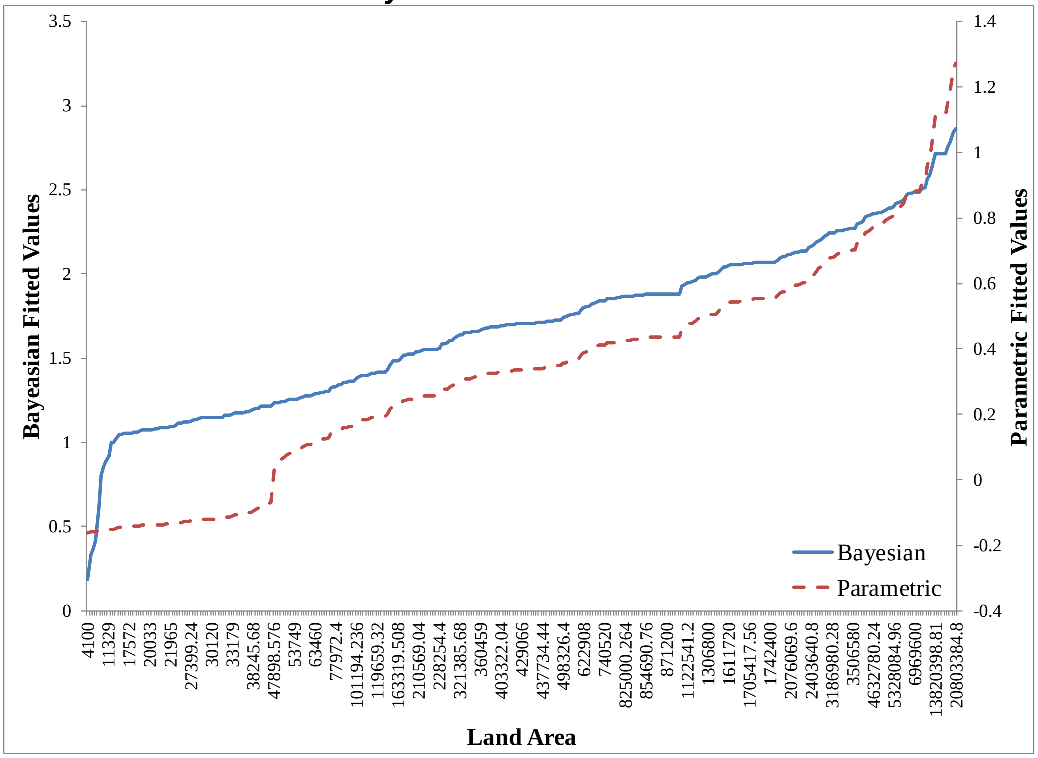


Figure 3 Estimation of the Size-Value relation - Arizona Data

Panel a: Semi-Parametric estimation



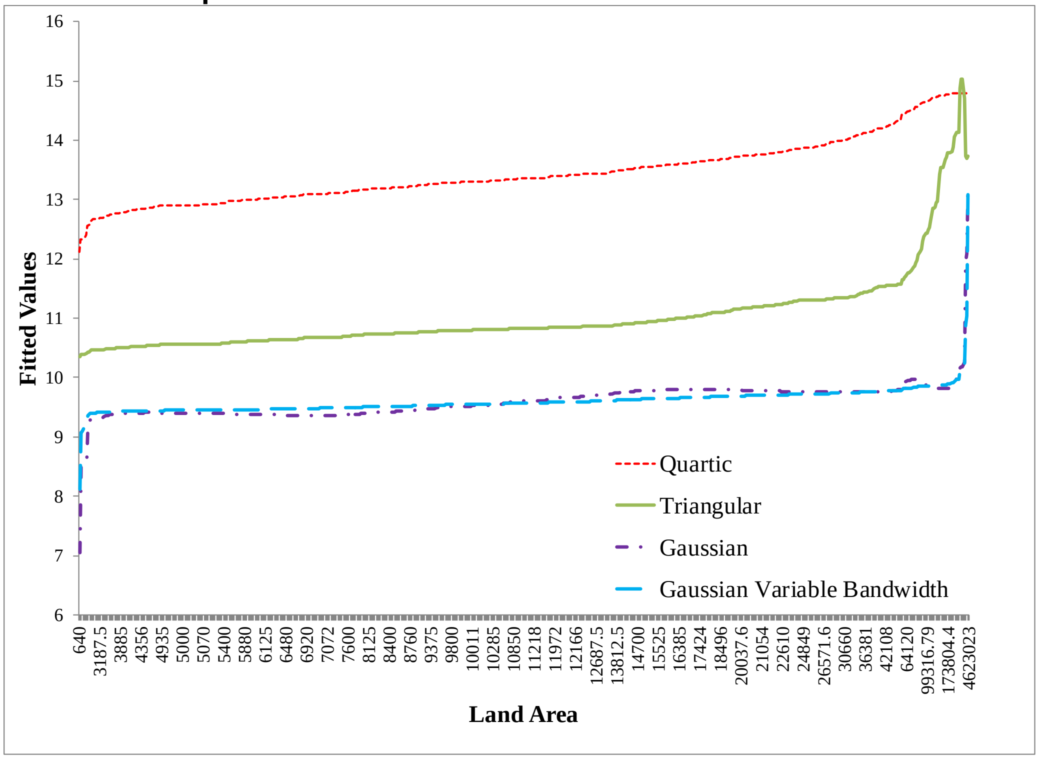
Panel b: Parametric and Bayesian* estimations



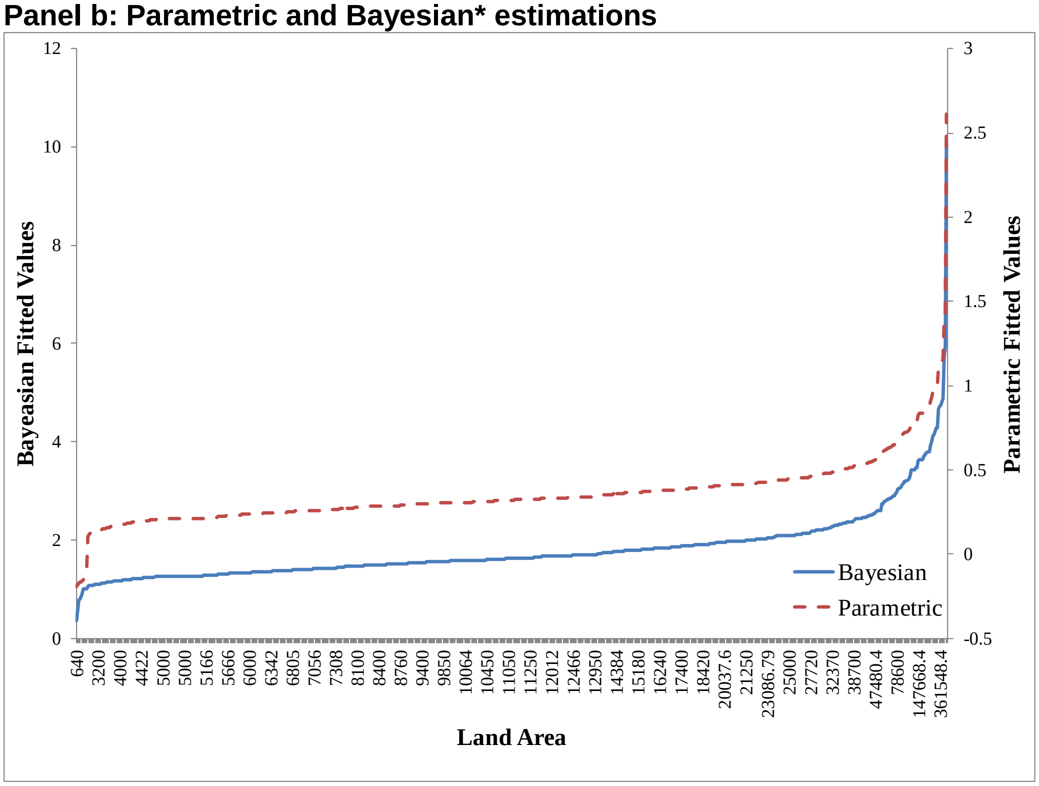
*In the Bayesian estimation line the inflection point has been indexed to the value of one in order to facilitate the graphical interpretation.

Figure 4 Estimation of the Size-Value relation - Iowa Data

Panel a: semi-parametric estimation



Panel b: Parametric and Bayesian* estimations



*In the Bayesian estimation line the inflection point has been indexed to the value of one in order to facilitate the graphical interpretation.

References

Appraisal Institute. *The Appraisal of Real Estate*, 12th edition, 2001.

Brownstone, D. and A. De Vany, Zoning, Returns to Scale, and the Value of Undeveloped Land, *Review of Economics and Statistics*, 1991 (Nov.), 73:4, 699-704.

Chib, S. and E. Greenberg, Understanding the Metropolis-Hastings Algorithm, *The American Statistician*, 49, 1995, 327–335.

Chib, S., F. Nardari, and N. Shephard, Analysis of High Dimensional Multivariate Stochastic Volatility Models, *Journal of Econometrics*, 134-2, 2006, 341-371.

Colwell, P., A Primer on Piecewise Parabolic Multiple Regression Analysis via Estimations of Chicago CBD Land Prices", Journal of Real Estate Finance and Economics, Special Issue on Spatial Econometrics, 1998, 17:1, 87-97.

Colwell, P. and H. Munneke, The Structure of Urban Land Prices, *Journal of Urban Economics*, 1997, 41 (May), 321-336.

Colwell, P. and H. Munneke, Land Prices and Land Assembly in the CBD, *Journal of Real Estate Finance and Economics*, 1999, 18:2, 163-180.

Colwell, P. and H. Munneke, Estimating a Price Surface for Vacant Land in an Urban Area, *Land Economics*, 1 March 2003, 79:1, 15-28(14).

Colwell, P. and C.F. Sirmans, Area, Time, Centrality and the Value of Urban Land, *Land Economics*, November 1978, 54:4, 514-519.

Colwell, P. and C.F. Sirmans, Nonlinear Urban Land Prices, *Urban Geography*, 1980, 1:2, 141-152.

Colwell, P. and C. F. Sirmans, A Comment on Zoning, Returns to Scale, and the Value of Undeveloped Land, *Review of Economics and Statistics*,1993, 75:4, 783-786.

Davidson and MacKinnon, *Estimation and Inference in Econometrics*. Oxford University Press, New York 1981.

Di Pasquale, D. and W. Wheaton, *Urban Economics and Real Estate Markets*, Prentice Hall,1996.

Ecker, M. and H. Isakson, A Unified Convex-Concave Model of Urban Land Values, *Regional Science and Urban Economics*, May 2005, 35:3, 265-277.

Fox, J. Describing univariate distributions in *Modern Methods of Data Analysis*, (J. Fox and J. S. Long, editors), Sage Publications, Newbury Park, CA, 1990.

Greene, William H., *Econometric Analysis*, 5th Ed., Pearson, 2003.

Lin, Tzu-Chin and A. W. Evans, The Relationship between the Price of Land and Size of Plot when Plots are Small, *Land Economics*, 2000, 76(3): 386-394.

Mills, E. S., The Value of Urban Land in *The Quality fo the Urban Environment*, (H. S. Perloff, editor), Johns Hopkins University Press, Baltimore, 1971.

Salgado-Ugarte, I. and M. Perez-Hernandez, Exploring the use of variable bandwidth kernel density estimators, *The Stata Journal*, 2003, 3(2): 133-147.

Silverman, B.W, Density Estimation for Statistics and Data Analysis, *Monographs on Statistics and Applied Probability*, London: Chapman and Hall, 1986.

Tabuchi, T., Quantity Premia in Real Property Markets, *Land Economics*, 1996, 72(2): 206-217.

Thornes, P., and D. P. McMillen, Land Value and Parcel Size: A Semi-parametric Analysis, *Journal of Real Estate Finance and Economics*, 1998, 17:3, 233-244.

[1] Specific functional forms that may generate the shape observed in Figure 1 will be described in Section III

[2] In principle, there is no reason for the existence of only one inflection point. If more than one inflection point exists in the data, the semi-parametric method studied in this paper has the advantage of not imposing any functional form for the size-price relationship, and has the potential to find more than one inflection point. However, most datasets contain very sparse data at larger parcel sizes, leading to possibly unreliable estimates of an inflexion point for these parcels. We leave this issue for future research.

[3] The J-test was introduced by Davidson and Mackinnon as a specification test for non-nested models in linear regressions. Suppose you have two competing non-nested models to explain variable Y. Model 1 is $Y = X\beta + \varepsilon_1$ and model 2 is $Y = Z\gamma + \varepsilon_2$. Estimating a compound model $Y = \alpha X\beta + (1-\alpha)Z\gamma + \varepsilon_3$ could be used to check which one of the competing models is correct. If $\alpha = 1$, then model 1 is correct (H_0) while if $\alpha = 0$ model 2 would be the correct one (H_1). Unfortunately the parameters of the compound model are not identified. However, Davidson and Mackinnon showed that we can use the linear regression $Y = X\beta + \eta\hat{Y}_2 + \varepsilon_4$ in which $\hat{Y}_2 = Z\hat{\gamma}$ to test the competing models. More specifically, they showed that if the t-statistic of the estimated parameter η is significant, then H_0 can be rejected. A similar logic can be used to test whether we reject or not H_1 (for a more detailed explanation see Greene (2005)).

[4] Thorsnes and McMillen also list the Epanechnikov and Uniform kernels. Results using Epanechnikov kernel are very similar to those using a Qartic one. The use of a Uniform kernel leads to very similar results to those obtained with the Triangular kernel. Results with the Epanechnikov and Uniform kernels are available upon request. We left them outside the paper for clarity and brevity in the exposition.

[5] In the simulations data is not sparse. Therefore, using a fixed bandwidth leads to essentially the same results as a variable bandwidth. In real data the two types of bandwidths might lead to different results

although we find in our datasets that using a fixed bandwidth leads to similar results as using a variable one.

[6] Thorsnes and McMillen (1998) pp 237. As it will be evident from our simulation study, in several instance the semi-parametric estimator is able to find the inflection point from plottage to plattage with relatively good accuracy at the first point estimate when the second derivative changes sign from positive to negative values.

[7] Ecker, M. and Isakson, H. (2005), pp. 267

[8] In the empirical section below we show that, when applied to the same real dataset (from Iowa), the two functional forms produce essentially identical inferences on all relevant dimensions.

[9] The only difference in the values of the parameters σ^2, τ^2 and ϕ between our paper and EI comes from ϕ. We increase the value of that parameter with respect to EI because the covariance matrix becomes non-invertible in some simulations for smaller values of ϕ. For generating the data applied in simulations we used Gauss 6.0. Finally, when we estimated the parameter ϕ using the Iowa data provided by EI we find with our Bayesian methodology similar results with respect to what they found.

[10] The J-Test is a common specification test used for model selection. For a detailed description see Greene (2003), p. 154-55.

[11] For each simulate dataset, the probability of plottage/(plattage) is estimated as the relative frequency with which the λ_s (λ_l) parameter is bigger (smaller) than one across the MCMC draws. We, then, report in Tables 2 and 3 the average of this relative frequency across the 1,000 simulated datasets.

[12] One limitation of the Arizona data is that if restricted to contain only rural parcels, then only a few small parcels are available. Previous work indicates that if an inflection point exists we should find it at relatively small parcel sizes with only a few thousand sqft at most. Therefore, we added improved parcels' data to the Arizona dataset together with dummy variables that control for the impact that those improvements might have on the final price of the raw land.

[13] Specifically, we provide 95% highest posterior density (HPD) intervals. A 100(1-α) HPD interval for a parameter θ has the property of being the smallest interval delimiting an area of 1-α under the posterior distribution of θ. Heuristically, a Bayesian HPD interval may be seen as similar to a frequentist confidence interval. In other words, when α = 0.05, the researcher is 95% confident that θ lies within the HPD.

[14] The range of spatial correlation is computed as 3/φ (see EI (2005)). The posterior quantities for the range are not reported in the tables in order to conserve space.

[15] Note that the scale of the point estimates with regard the size-price relation differ for each estimation method. While the Parametric and Bayesian estimates are relatively similar (generally smaller than 5), the ones corresponding to the Semi-parametric are bigger (generally larger than 10). The main reason for this difference to appear is that the function that relates size to price is scaled by the parameter β in the Bayesian and Parametric methods while it is not in the Semi-parametric method.