

A New Approach to Computing Hedonic Equilibria and Investigating the Properties of Locational Sorting Models

Nicolai V. Kuminoff *
(*Corresponding Author*)

Department of Economics
Arizona State University
Main Campus, PO BOX 873806
Tempe, AZ 85287-3806
Phone: 480.727.9802
Fax: 480.965.0748
kuminoff@asu.edu

and

Abdul Salam Jarrah
Department of Mathematics and Statistics
American University of Sharjah
P.O.Box. 26666
Sharjah, U.A.E.
Phone: 971.6.515.2397
Fax: 971.6.515.2950
ajarrah@vbi.vt.edu

October, 2009

* We have benefited from several conversations with H. Allen Klaiber and V. Kerry Smith, and from excellent suggestions by two anonymous referees that led us to investigate the multiplicity of equilibria. We also thank Paul Fackler, Ray Palmquist, Chris Parmeter, Dan Phaneuf, Jaren Pope, and William Strange for helpful comments and suggestions on this research.

A New Approach to Computing Hedonic Equilibria and Investigating the Properties of Locational Sorting Models

ABSTRACT:

This paper outlines a new way to solve the traditional housing market assignment problem and uses it to investigate the properties of hedonic equilibria. Our approach to computing equilibria is based on Rosen's (1974) bid function. It has four desirable features: (i) convergence implies a hedonic equilibrium; (ii) convergence is guaranteed if a hedonic equilibrium exists; (iii) it can solve for a new equilibrium following a shock to the market; and (iv) if multiple equilibria exist, it can identify them. The algorithm is applied to micro data from San Joaquin County, California, where the choice of a home provides access to public schools in particular school districts. First we calibrate the algorithm to approximately reproduce actual housing prices in San Joaquin County as a hedonic equilibrium. Then we introduce a policy that improves school quality in selected school districts. We find that there are several possibilities for the new equilibrium. For each of these potential equilibria, we compare the marginal willingness to pay for school quality with the rate at which the improvement is capitalized into property values. The resulting capitalization rates differ substantially from marginal willingness to pay.

KEY WORDS: Capitalization, hedonic, locational equilibrium, multiplicity, sorting

JEL CODES: R21, C15, C52, H41, Q51.

1. Introduction

The hedonic property value model is one of the most trusted methods for estimating what consumers are willing to pay for non-market goods and services. Over the past 40 years, economists have used the model to measure the willingness to pay for a wide range of urban amenities including air quality (Ridker, 1967), water quality (Dornbusch and Barrager, 1973), earthquake risk (Brookshire et al., 1985), school quality (Black, 1999), and airport noise (Pope, 2008a). This line of research matters for public policy. Hedonic estimates have served as a basis for litigation and have helped to inform the regulatory process (Palmquist and Smith, 2002). The willingness of judges, juries, and regulators to trust the results from hedonic models makes it especially important for us to understand their capabilities and limitations.

Recent advances in hedonic modeling have sought to improve our estimates for willingness to pay by tracking the way housing markets adjust to unexpected shocks. Some studies have used quasi-experimental designs to identify the rates at which shocks to amenities are capitalized into property values (Chay and Greenstone, 2005; Davis, 2004; Linden and Rockoff, 2008; Pope, 2008a; Pope, 2008b). Others have used structural models to predict how future shocks will affect migration patterns, capitalization rates, and consumer welfare (Sieg et al., 2004; Smith et al., 2004; Walsh, 2007). Both of these approaches have raised new questions about the empirical properties of hedonic equilibria.

A key issue for the structural models is the potential multiplicity of equilibria. If there are several equilibria that may follow a shock to an amenity, how does one choose between them? Will a researcher's choice drive their predictions for welfare measures? For the quasi-experimental models, a key issue is the distinction between the capitalization rate for an amenity and the marginal willingness to pay (MWTP). Hedonic theory predicts that the two measures

will generally differ, but it does not predict the size of the difference (Starrett, 1981; Bartik, 1988). Will the difference be small enough to treat the capitalization rate as an approximation to MWTP? It is important to answer this question because there is considerable interest in using quasi-experimental methods for nonmarket valuation (Greenstone and Gayer, 2009).

The difficulty with investigating the properties of hedonic equilibria is that, in general, there is no closed-form expression for the equilibrium price function. We address this difficulty in the first half of the paper by outlining a simple way to compute equilibria numerically. The second half of the paper applies the new algorithm to a particular set of micro housing data in order to investigate whether the hedonic equilibrium will be unique and whether the capitalization rate for an unexpected shock to an amenity will approximate MWTP.

Our approach to computing equilibria is rooted in hedonic theory. We use Rosen's (1974) bid function to develop an iterative bidding algorithm that solves for a vector of prices and an assignment of people to homes that jointly clear the market for housing. The algorithm has four desirable properties: (i) convergence implies a hedonic equilibrium; (ii) convergence is guaranteed if a hedonic equilibrium exists; (iii) it can solve for a new equilibrium following a shock to the market; and (iv) if multiple equilibria exist, it can identify them. In previous work on simulating hedonic equilibria, Cropper et al. (1988, 1993) and Banzhaf (2003) solved the housing market assignment problem using a linear programming algorithm developed by Wheaton (1974). We compare both algorithms on a typical hedonic data set and find that the iterative bidding approach is faster, converges more reliably, and reduces the curse of dimensionality in the need for computer memory. These features allow us to compute equilibria for data sets that are an order of magnitude larger than those used in previous studies.

To investigate the potential multiplicity of equilibria, we apply the new algorithm to

housing data from San Joaquin County, California. Households with heterogeneous preferences and incomes are assumed to derive utility from housing characteristics, public school quality, and other amenities. We find that the data support several equilibria. Equilibria associated with high utility have low prices, and those associated with low utility have high prices. This result complements Bayer and Timmins (2005), who find that multiple equilibria can arise from social interactions in residential location choice. In contrast, our results demonstrate that multiple equilibria can arise in a conventional hedonic model without social interactions.

To investigate whether capitalization rates approximate MWTP, we first calibrate the model to approximately reproduce actual housing prices in San Joaquin County as a hedonic equilibrium. Then we introduce a policy that improves school quality in selected school districts. While there are several possibilities for the new equilibrium, they imply similar capitalization rates. Regardless of which equilibrium we choose, capitalization rates do not approximate MWTP. They systematically overstate MWTP in the lowest quality school district and they systematically understate MWTP in the highest quality school district. These differences arise from the way heterogeneous households sort themselves across the market in a hedonic equilibrium.

The remainder of the paper is organized as follows. Section two defines the urban landscape and formalizes the equilibrium conditions. Section three outlines the iterative bidding algorithm and provides convergence proofs. Section four summarizes the data and uses it to provide benchmarks on computational performance. Sections five and six report our results on multiplicity and capitalization. Finally, section seven concludes with suggestions for future research. Matlab code for the algorithm and data to replicate our results are provided in a supplemental appendix.

2. Characterizing a Locational Equilibrium

Hedonic property value models begin from the following problem: the availability of housing and public goods varies across an urban landscape and each household chooses to occupy the location in that landscape that provides its preferred bundle of goods, given its preferences, income, and the relative prices involved. Every household pays for its location choice through the price of housing. The problem can be formalized using the characteristics approach to consumer theory (Lancaster, 1966). That is, the utility a household obtains from each location can be written as a function of the characteristics of that location.

Let the urban landscape consist of J homes, each of which is defined by a vector of characteristics, x_j , where $j = 1, \dots, J$. This includes structural characteristics of the home, such as the number of bedrooms, the number of bathrooms, square feet, and lot size, as well as local public goods and amenities, such as crime rates, school quality, air quality, and access to open space. A household's utility depends on the characteristics of housing and public goods at its location and on its consumption of a numeraire composite commodity, c . Households are heterogeneous. They differ in their income, y , and in their preferences, α . Let the population of households be indexed from $i = 1, \dots, I$, where $I = J$. Then the utility obtained by household i at location j can be represented as: $U(x_j, c, \alpha_i)$. Each household, i , is assumed to choose a specific house and a quantity of c that maximize its utility subject to a budget constraint:

$$\max_{j,c} U(x_j, c; \alpha_i) \text{ subject to } y_i = c + p_j. \quad (1)$$

In the budget constraint, the price of the numeraire is normalized to one, and p_j represents annualized expenditures on house j .

A locational equilibrium is achieved when every household occupies its utility-maximizing location and nobody wants to move, given housing prices, housing characteristics, and the exogenous provision of local public goods.¹ In order to define this concept more formally, let b_{ij} denote household i 's bid for the j^{th} home, and let A_{ij} be an assignment indicator where $A_{ij} = 1$ if and only if household i occupies home j . Then a locational equilibrium must satisfy:

$$b_{ij} = \max_i \{ b_{ij} \} \text{ iff } A_{ij} = 1, \quad (2)$$

$$\sum_i A_{ij} = \sum_j A_{ij} = 1, \text{ for all } i, j. \quad (3)$$

In other words, each household occupies exactly one home, for which it has the maximum bid.

In Rosen's (1974) model of a hedonic equilibrium, bids are expressed as a function of housing characteristics and preferences. To see this, let \tilde{u} be some reference level of utility, and consider an indifference surface over which x and c vary, while \tilde{u} stays the same:

$\tilde{u} = U(x, c; \alpha)$. Assuming utility is monotonically increasing in c , the function can be inverted to solve for c .

$$c = U^{-1}(x, \tilde{u}; \alpha). \quad (4)$$

Inserting (4) into the budget constraint and rearranging terms allows a household's maximum willingness-to-pay for a home to be expressed as a function of its characteristics and the household's income, preferences, and reference utility:

$$b = y - U^{-1}(x, \tilde{u}; \alpha). \quad (5)$$

This is Rosen's (1974) bid function. We can use it to solve for a locational equilibrium, given a

¹ See Bayer and Timmins (2005, 2007) for a discussion of locational equilibria and the estimation of neighborhood choice models with endogenously determined public goods.

parametric specification for the utility function, information on preferences and income, and data on housing characteristics.

3. Solving for a Locational Equilibrium

We use Rosen's bid function to solve for a vector of prices and a unique assignment of households to homes that jointly clear the market for housing. This requires conducting a series of hypothetical second-price auctions for individual homes until subsequent bidding has no further effect on prices. While the equilibrium conditions in (2)-(3) are not imposed on any individual step of this process, we demonstrate that convergence of the algorithm implies a locational equilibrium.

3.1. The Iterative Bidding Algorithm (IBA)

As in Wheaton (1974), the process of solving for a locational equilibrium begins by assigning each household i a reference level of utility, \tilde{u}_i . This can be used together with data on the distribution of housing characteristics and data on the joint distribution of income and preferences to solve for each household's bid for each home. The bids are then used to conduct a second-price auction for each property.

Each auction requires the highest bidder to pay the second highest bid plus a marginal increment, $\varepsilon > 0$. Consider an auction for the j^{th} home. It can be decomposed into three steps:

- i. Given $\tilde{u}_i, \alpha_i, y_i, x_j$, solve for $b_{ij} = y_i - U^{-1}(x_j, \tilde{u}_i; \alpha_i)$ for all i .
- ii. Rank the bids and set p_j equal to the second highest bid plus ε .
- iii. Update utility for k , the highest bidder: $\tilde{u}_k = U(x_j, y_k - p_j; \alpha_k)$.

Every household submits a bid and the highest bidder, k , is assigned to home j . Assuming ties

for the highest bid occur with zero probability, household k will pay a price that is less than or equal to its bid, $p_j \leq b_{kj}$. Since utility is assumed to be monotonic in the numeraire, the new assignment must either increase k 's reference utility or leave it unchanged. Monotonicity also implies that an increase in k 's reference utility will decrease its bid for every other home. After solving for p_j and updating \tilde{u}_k , we move on to auction the next home.

The IBA continues running second-price auctions until the occupant of every home is paying an ε above the second highest bid for that home. The complete algorithm consists of four steps:

Iterative Bidding Algorithm (6)

- (6.a) Order all the houses in the market from 1 to J .
- (6.b) Define $\alpha_i, y_i, \tilde{u}_i$ for each i .
- (6.c) Conduct an auction for each house and update utility for the highest bidder.
 - 1. Solve for p_1 and k , and update $\tilde{u}_k = U(x_1, y_k - p_1; \alpha_k)$.
 - 2. Solve for p_2 and k , and update $\tilde{u}_k = U(x_2, y_k - p_2; \alpha_k)$.
 - .
 - .
 - .
 - J. Solve for p_J and k , and update $\tilde{u}_k = U(x_J, y_k - p_J; \alpha_k)$.
- (6.d) If (6.c) did not change the price of any home, stop. Otherwise repeat (6.c).

The subscripts on p_1, \dots, p_J in (6.c) refer to specific homes, whereas the k subscript on utility identifies the individual bidder who wins each auction. In general, different individuals will win different auctions. However, unlike the adaptation of Wheaton's (1974) linear programming algorithm in Cropper et al. (1988, 1993) and Banzhaf (2003), the IBA does not constrain households to be uniquely assigned to homes on intermediate iterations of the algorithm. More precisely, the equilibrium condition in (3) may be violated on an intermediate iteration of (6.c).

For example, the household who won the auction for home 1 in step 6.c.1 may have subsequently won the auction for home J on step 6.c.J. In this case the decrease in p_J that motivated the household to move from 1 to J will violate the convergence criterion in (6.d). Therefore, the algorithm will begin a new iteration of (6.c). The IBA continues to iterate over (6.c) until the vector of housing prices converges, signaling the market has cleared. Market clearing implies that, at the current vector of prices, no household can increase its utility by moving to a different home.

Notice that the IBA systematically decreases housing prices. Each auction either decreases the price of a home or leaves it unchanged. The systematic nature of this process guides our choice for the initial reference level of utility. $\tilde{u}_1, \dots, \tilde{u}_I$ must be defined such that the vector of prices on the first iteration of the algorithm lies above the equilibrium. Since prices always decrease, we must start the IBA at a point that lies above the equilibrium price vector if we hope to converge to it.

If the IBA converges to a vector of prices, that vector must satisfy the conditions for a locational equilibrium. Moreover, if a locational equilibrium exists, the algorithm is guaranteed to converge.² Before proving these propositions in section 3.3, we first present numerical and graphical examples of the algorithm to illustrate its mechanics and provide intuition for the subsequent proofs.

3.2. *Graphical and Numerical Illustrations of the IBA: A Three-Home Example*

Suppose utility can be represented by the Cobb-Douglas function in (7), so that household i 's bid

² If no equilibrium exists, the algorithm will not converge. The specification for utility in (1) is too general to prove that a locational equilibrium will exist at positive prices, or at all. It is possible to prove existence if one is willing to impose additional restrictions on the structure of preferences and the stock of housing. For examples, see Bayer and Timmins (2005), Epple and Romer (1991), or Ekeland (2008).

for the j^{th} home can be expressed as (8).

$$U_{ij} = \ln(c) + \alpha_i \ln(x_j). \quad (7)$$

$$b_{ij} = y_i - \exp[\tilde{u}_i - \alpha_i \ln(x_j)]. \quad (8)$$

Now consider three households, A , B , and C , who bid on homes 1, 2, and 3. Table 1A provides each household's income and its utility from the characteristics of each home, $\alpha_i \ln(x_j)$. The households differ in how they rank homes 2 and 3, but they would all prefer to live in home 1. With this in mind, we define the initial reference level of utility as $\tilde{u}_i = \alpha_i \ln(x_1)$. This is the point at which each household spends all but one dollar of its income on its favorite home. Since the IBA decreases prices systematically, starting the algorithm at this point guarantees it will not converge to a nonsensical solution where housing expenditures exceed income.

Table 1B tracks the adjustment process as the IBA searches for a locational equilibrium. The algorithm begins by conducting a second-price auction for home #1. Each household bids its income less one dollar, reflecting our choice for \tilde{u}_i . The household with the highest bid, A , is assigned to live there. It pays \$64,500—the second highest bid plus ε , which is set to \$1 in this example. After updating A 's utility, household B is assigned to live in home #2 at a price of \$56,118. Then household B wins the auction for home #3 and pays \$55,444. This completes the first iteration of the algorithm. Notice that the assignment of households to homes is not unique.

When we return to home #1 on the second iteration, household A remains the highest bidder. However, the second highest bid (\$64,353) has decreased because B 's utility has increased through its assignment on the first iteration. A decrease in the second highest bid for home #1 causes its price to decrease. The same is true for homes #2 and #3. While the second iteration concludes with a unique assignment of households to homes, the vector of prices has not converged. All three prices are lower than on the previous iteration. Therefore, the auctions

continue until the price vector finally converges on the 14th iteration of the algorithm. At this point, each household has the highest bid for the home they occupy, paying $\varepsilon = \$1$ above the next highest bid. This is a numerical example of a locational equilibrium.

We can use the numerical example in table 1B to depict the IBA graphically. Since the utility function in (7) is monotonic in the numeraire, the bid for a particular home in (8) will be monotonically decreasing in the level of utility. Tracing out how an individual household's bids for each home vary with its level of utility allows us to define a locus of b_1, b_2, b_3 combinations in price space. This locus of points forms a string. Figure 1A graphs these "bid indifference strings" for each of the three households. A single point on a string identifies a b_1, b_2, b_3 combination at which the corresponding household's utility is fixed. In other words, the household would be exactly indifferent between the three homes if they were sold at these prices. The highest point on each string is defined by the bid triple that corresponds to the household's initial reference level of utility. For example, the string for household C ends at the point $(b_1 = 56999, b_2 = 56117, b_3 = 55443)$ which are its bids for the three homes on the first iteration of the IBA (table 1B). As we move down a string toward the origin, bids decrease and utility increases. Thus, the bid strings simply provide a graphical representation of household preferences in price space.

The IBA systematically searches along the bid strings, from top to bottom, until it finds a locational equilibrium. To follow the movement of the IBA, consider the hyperrectangle in figure 1B. It is defined by a lower vertex at the origin and an upper vertex at the highest bid in each dimension. We have left it transparent to illustrate that it contains the three bid strings. Every time the IBA decreases the price of a home, the upper vertex moves toward the origin. Graphically, one face of the hyperrectangle is pushed inward, toward the origin, on each step of

the IBA. Figures 1C, 1D, and 1E illustrate the first iteration of this process. First, figure 1C shrinks the hyperrectangle in the p_1 dimension until its face is only an ε away from touching a second bid string. The bid string that passes through this face belongs to household A, the highest bidder for home #1. Figure 1D repeats this process, shrinking the hyperrectangle in the p_2 dimension and assigning household B to home #2. Household B is subsequently assigned to home #3 when figure 1E shrinks the hyperrectangle in the p_3 dimension. This process continues until it is no longer possible to push any face of the hyperrectangle toward the origin without passing through a second household's bid string.³

Figure 1F illustrates convergence of the algorithm. In this locational equilibrium, households A, B, and C are assigned to homes 1, 3, and 2, and the equilibrium price vector is defined by the upper vertex of the hyperrectangle. The price each household pays for the home it occupies is an ε above what the next highest bidder is willing to pay. For example, household C occupies home #2. The price it pays (\$42,289) is one dollar above household A's bid (\$42,288) since $\varepsilon = \$1$. Graphically, C's bid string is the only one to pass through the face of the hyperrectangle in the p_2 dimension. A's bid string lies just inside the hyperrectangle, an ε away from the face. A visual inspection of figure 1F reveals that we cannot shrink the hyperrectangle by more than ε in any dimension without forcing two strings to pass through a single face. This is the graphical representation of a locational equilibrium.

3.3. *Key Properties of the IBA*

Convergence of the IBA is guaranteed to satisfy the conditions for a locational equilibrium under our maintained assumptions that utility is increasing monotonically in the numeraire and that ties

³ The supplemental appendix includes Matlab code that creates a "movie" of the shrinking hyperrectangle.

for the highest bid do not occur. We state this formally as proposition 1.

Proposition 1. *If $U'(c) > 0$ and ties for the highest bid occur with zero probability, there must be some sufficiently small $\varepsilon > 0$ such that convergence of the IBA to a vector of prices and an assignment of households to homes satisfies the conditions for a locational equilibrium.*

Proof. A locational equilibrium occurs when equations (2) and (3) are satisfied simultaneously. Convergence of the IBA always implies (2). Suppose the algorithm converges at a point that violates (3). Then at least one household must be assigned to at least two homes. Without loss of generality, let household i be assigned to homes j and $j+1$, with equilibrium prices p_j^* and p_{j+1}^* . Now consider the final auction for $j+1$. Having been previously assigned to j , household i 's bid is $b_{i,j+1} = y_i - U^{-1}[x_{j+1}, \tilde{u}(x_j, y_i - p_j^*, \alpha_i), \alpha_i]$. Since ties are assumed to occur with zero probability, the next highest bid must be lower and household i will be assigned to $j+1$ and pay the price $p_{j+1}^* \leq b_{i,j+1}$. If this condition holds with equality, replace ε with $\varepsilon/2$ so that $p_{j+1}^* < b_{i,j+1}$. Since utility is monotonic in the numeraire, $\tilde{u}(x_{j+1}, y_i - p_{j+1}^*, \alpha_i) > \tilde{u}(x_j, y_i - p_j^*, \alpha_i)$, which means household i is no longer willing to pay p_j^* for home j . Therefore, the algorithm cannot converge at the point p_j^*, p_{j+1}^* . \square

The same restrictions we use to guarantee that convergence of the IBA implies a locational equilibrium are also sufficient to guarantee that existence of a locational equilibrium implies convergence of the IBA. Notice that proposition 1 implies the IBA cannot stop at any

point that lies above a unique equilibrium. Proposition 2 shows that if the IBA starts at a point that lies above a unique equilibrium, it cannot bypass that equilibrium.

Proposition 2. *If $U'(c) > 0$ and a unique equilibrium price vector exists inside the hyperrectangle defined by a lower vertex at the origin and an upper vertex at the starting vector of housing prices, the IBA cannot bypass it.*

Proof. Let P^* denote the unique vector of prices that defines a locational equilibrium. In order to bypass P^* , the IBA would have to move from a point \bar{P} to a point \underline{P} such that two conditions hold:

- (i) $\underline{P}_j, P_j^* \leq \bar{P}_j$ for all $j = 1, \dots, J$.
- (ii) $\underline{P}_j < P_j^* \leq \bar{P}_j$ for exactly one j .

Suppose the bypass occurs on the auction for the j^{th} home. Let $b_j^2(\bar{P})$ denote the second highest bid for j , which depends on the current vector of prices. Using this notation, we can rewrite the first inequality in (ii) as $b_j^2(\bar{P}) + \varepsilon < b_j^2(P^*) + \varepsilon$, which implies (iii):

$$(iii) \quad b_j^2(\bar{P}) < b_j^2(P^*).$$

From the definition of a locational equilibrium, $b_{ij}(P^*) \leq P_j^*$ for all i . Since $P_j^* \leq \bar{P}_j$ for all j , the monotonicity restriction implies $b_{ij}(P^*) \leq b_{ij}(\bar{P})$ for all i . This contradicts (iii). \square

Figure 1 can help to provide some intuition. From a graphical perspective, proposition 2 simply recognizes that in order to bypass an equilibrium, the IBA would have to push one face of

the hyperrectangle through at least two bid strings. This would violate the algorithm's adjustment criteria. The same logic can be applied to a situation with multiple equilibria. Put simply, the IBA cannot bypass *any* equilibrium. However, with multiple equilibria the order in which the IBA iterates over homes may determine which of the equilibria it converges to. To see this, partition the set of all possible equilibria into two sets, S and T , such that any price vector contained in S is strictly higher than every price vector contained in T and the price vectors contained in S are unordered.⁴ Now consider the special case where S is a singleton. In this case, it is fairly obvious that the algorithm will converge to the unique equilibrium price vector in S regardless of the order in which we iterate over homes. We state this as a corollary to proposition 2.

Corollary 1. *Let $S = \{p^{**}\}$, where p^{**} is the highest equilibrium price vector. The IBA will converge to p^{**} .*

Proof. Let p^0 denote the IBA starting value, where $p_j^0 \geq p_j^{**}$ for all j . Moving from p^0 to a point below p^{**} requires passing through the equilibrium point p^{**} . This possibility is ruled out by proposition 2. \square

In the more general case where S is not a singleton, the algorithm will converge to one of the equilibrium points in S . Which of these equilibria it converges to may depend on the order in which we iterate over homes. While it is difficult to develop a general proof for this situation, it

⁴ This refers to the usual vector ordering where two price vectors p^1 and p^2 are *unordered* if $p_j^1 > p_j^2$ and $p_k^1 < p_k^2$ for at least one j, k pair with $j \neq k$.

is certainly possible to experiment in the context of a simulation. Alternatively, if one is willing to impose specific restrictions on the structure of preferences and the stock of housing, it is possible to guarantee uniqueness or to restrict the set of possible equilibria (Ekeland, 2008).

3.4. Discussion

We have demonstrated that the iterative bidding algorithm allows us to solve for a vector of hedonic prices and an assignment of people to homes that jointly define a locational equilibrium. Our approach avoids the need to impose the equilibrium assignment condition on each step of the algorithm. Nevertheless, under the twin assumptions that utility is monotonic in the numeraire and that ties for the highest bid do not occur, we can guarantee that convergence of the IBA implies a locational equilibrium and that existence of a locational equilibrium guarantees convergence of the IBA. These assumptions deserve some additional discussion.

The assumption that ties for the highest bid occur with zero probability implies that there must be a minimal degree of heterogeneity among houses and households. No two households can be identical in their incomes, preferences, and initial reference utilities. Since identical households would always have identical bids, they would eventually share the highest bid for a home. Likewise, no two homes can provide identical bundles of characteristics. In a locational equilibrium, identical homes must have identical prices. A household assigned to one of these homes would obtain the same utility from its twin and the occupants of both homes would have identical bids for each.

A model in which no two homes are identical would be consistent with an urban landscape in which space matters. If spatially delineated amenities are conveyed through the location of a home, it seems reasonable to assume that households will perceive each home as

being unique. This is consistent with the empirical hedonic literature. Homes with identical structural characteristics tend to be priced differently according to their proximity to urban and environmental amenities such as air quality (Chay and Greenstone, 2005), school quality (Black, 1999), crime (Linden and Rockoff, 2008; Pope, 2008b), and airport noise (Pope, 2008a). Furthermore, the assumption that houses and households are unique is consistent with the continuity assumptions that underlie Rosen's (1974) theoretical model of a hedonic equilibrium. While we conjecture that the IBA could be extended to include decision rules that would allow us to solve for a locational equilibrium in which a subset of the houses and households are identical, there is little to gain from doing so since our current framework only requires that each pair of houses and households differ marginally.

The assumption that utility is monotonically increasing in the numeraire is more fundamental to the mechanics of the IBA. Monotonicity is what guarantees prices decrease as the algorithm iterates over auctions for each home, and it also underlies our proofs of convergence. These results come at little cost. Since monotonicity is one of the maintained assumptions of Rosen's model, the mechanics of the IBA are consistent with hedonic theory.

While it is reassuring to know that the IBA will converge to a hedonic equilibrium, if one exists, our proofs do not provide any insight into the computational burden of the algorithm or its convergence speed. The next section provides some evidence on these practical considerations.

4. Calibrating the IBA and Evaluating its Performance

4.1. The Urban Landscape: San Joaquin County, California

Located in the middle of California's central valley, San Joaquin is one of the largest agricultural counties in the nation, with annual production value well over a billion dollars. Three quarters of

its 564,000 residents live in seven cities: Escalon, Lathrop, Lodi, Manteca, Ripon, Stockton, and Tracy. Their locations are traced out by the density of census tracts and the locations of recent housing sales shown on the map in figure 2. Nearly all of the white space on the map is farmland. Production agriculture accounts for over 80% of the land use in the county, surrounding and dividing the seven cities. This landscape evokes the “city surrounded by farmland” metaphor Starrett (1981) used to motivate his conceptual model of how public goods are capitalized into housing prices.

To provide a representation of San Joaquin’s urban housing market that is consistent with the empirical hedonic literature, we assume that homebuyers care about the structural characteristics of their homes, the demographic composition of their neighborhoods, and their proximity to spatially delineated public goods and amenities. Table 2 reports summary statistics for the variables we use to create this landscape. Data on the price and structural characteristics of 9,634 homes sold in the county between 1995 and 1998 were purchased from a commercial vendor.⁵ Each of these homes is differentiated by the following structural characteristics: *number of bedrooms, age, building size, and lot size*. We define neighborhoods as Census tracts and attach data on three attributes from each tract to the homes they contain: *mean time-to-work, median household income, and share of the population under 18*.⁶ GIS data from the California Farmland Mapping and Monitoring Program were used to develop two proxy measures of land use near each home: *distance to the nearest grazing land* and *distance to the nearest water body*.

While these measures are admittedly crude, both are statistically significant and economically

⁵ The summary statistics in table 2 are based on the original transactions data. After preparing table 2, we added a random shock to the price and structural characteristics of each home in order to create a “fake” data set that we can distribute without violating contractual obligations to the vendor. See the appendix for details.

⁶ For simplicity, our simulations treat the demographic characteristics of communities as fixed. This approximates a situation in which only a small fraction of households are mobile. Recent research has illustrated how neighborhood demographics may be determined as outcomes of a sorting process across communities (see, for example, Bayer and Timmins, 2005). Integrating this form of endogeneity into the IBA would be an interesting topic for further research.

important in a simple linear regression of housing prices on housing characteristics and urban amenities. The last two columns of table 2 report these results. All else constant, homes closer to grazing land tend to be more expensive, while those closer to water bodies tend to be less expensive. We hypothesize that because grazing land is located in the foothills of the county, it provides relatively scenic views and closer access to opportunities for outdoor recreation. Meanwhile, the main water body in the county is a series of tributaries from the Sacramento Delta that pose a flood risk for nearby residents in wet years.

With few exceptions, students in San Joaquin County are required to attend schools located within the geographic boundaries of the school district in which they live. This assignment creates a link between the choice of a home and public school quality.⁷ Keeping this in mind, we develop a proxy measure for school quality that we assign uniformly to homes within each district. Specifically, we use *10th grade math score* in 1998, reported by the Standardized Testing and Reporting (STAR) program.⁸ Then we convert the average score for each district into its corresponding percentile in the distribution of all California school districts. San Joaquin's seven school districts range from the 54th percentile to the 80th percentile in the statewide distribution of school districts, ranked by STAR math score.

Finally, because housing expenditures are defined as an annualized measure in the hedonic bid function, the price data were converted into rents using the formula suggested by Poterba (1992). Specifically, the relationship between the sale price of a home (P) and its annualized user cost (R) can be expressed as (9):

$$R = [(1 - \tau)(i - \tau_p) + r + m + \delta - \pi]P. \quad (9)$$

⁷ A series of empirical studies have exploited the spatial discontinuities associated with school district boundaries to measure the extent to which school quality is capitalized into housing prices. This work begins with Black (1999).

⁸ These data are available from the California Department of Education.

The marginal tax rate (τ) was based on the U.S. average (23%); California’s Proposition 13 fixes the annual property tax rate (τ_p) at 1%; and the 7% interest rate (i) represents an annual average of the 30-year fixed rate mortgage as reported by the Federal Home Loan Mortgage Corporation from 1995 to 2005. The annual risk premium ($r = 4\%$), maintenance rate ($m = 2\%$), and depreciation rate ($\delta = 2\%$) were all set to Poterba’s suggested values. Finally, the land appreciation rate ($\pi = 5\%$) was calculated as the annual average inflation rate for the consumer price index of housing in the San Francisco Consolidated Metropolitan Statistical Area from 1995 to 2005, as reported by the Bureau of Labor Statistics.⁹ The resulting formula implies the annual rental rate for housing equals 9.17% of its price.

4.2. *The Specification for Utility and Performance Benchmarks*

Given a definition for the stock of housing and the joint distribution of income and preferences, we can use the IBA to solve for a hedonic price vector that defines a locational equilibrium. In our simulations, the joint distribution of income and preferences is based on the Cobb-Douglas specification for utility in (10):

$$U_{ij} = \ln(c) + \alpha_i \ln(x_j), \quad \text{where } \alpha_i = \omega(\tilde{\alpha}_i/y_i), \quad \text{and } \tilde{\alpha} \sim \Gamma(\delta, \beta). \quad (10)$$

A vector of scaling parameters, ω , recognizes that relative preferences may vary systematically over the ten different housing characteristics.¹⁰ Dividing $\tilde{\alpha}$ by income implies negative correlation between income and the overall strength of preferences for housing relative to all other goods. While negative correlation is not required by the IBA, we find that it improves our

⁹ San Joaquin County is adjacent to the San Francisco Consolidated Metropolitan Statistical Area and approximately 16% of its working residents commute to work in the Bay Area.

¹⁰ This “horizontal” specification for preferences is consistent with the more general depictions of preference heterogeneity in the literature on structural estimation of locational sorting models (e.g. Bajari and Kahn, 2005; Bayer and Timmins, 2007).

ability to fit the model to the San Joaquin data. Previous studies have also found that negative correlation between income and preferences helps to explain why households are less than perfectly stratified across residential communities by income (Epple and Sieg, 1999; Sieg et al., 2004). Finally, to draw from the joint distribution of income and preferences, we use (10) to combine a random draw from $\Gamma(\delta, \beta)$ with an independent random draw on y from the distribution of household income provided in the 2000 Census of Population and Housing.¹¹

While the Cobb-Douglas specification for the utility function is somewhat rigid, the gamma distribution of preferences is sufficiently flexible to allow us to approximately reproduce the pattern of housing prices in San Joaquin County. This was done by using the Nelder-Mead algorithm to solve for the 12x1 vector $[\omega \ \delta \ \beta]$ that minimizes the distance between the predicted and observed distributions of housing prices.¹² Figure 3 contrasts the difference between these distributions for two different sample sizes: 200 and 2000. The solid line in panel A represents the empirical cumulative distribution function of actual prices for 200 homes in San Joaquin.¹³ The dashed line represents the equilibrium prices assigned to those homes in our simulation. While the predicted prices for some homes differ considerably from their actual values, the simulation reflects the general price trend in our data. This is reinforced by the close match between the corresponding simulated and empirical probability density functions in panel B. Panels C and D illustrate that these results do not change much when we increase the sample size to 2000. Overall, our simulated equilibria appear to provide a reasonable approximation to

¹¹ The Census distribution reports the number of households in each of 16 income bins. The lowest bin ($y < \$5,000$) was dropped under the assumption that households in this category are retired or purchasing housing out of savings. The highest bin ($y > \$200,000$) was truncated at \$300,000.

¹² An alternate version of the simulation used independent draws from 10 gamma distributions with different shape and scale parameters. This specification was abandoned because it greatly increased computational time without substantially improving model fit.

¹³ Recall that these are annualized housing prices. Converting them back to actual housing prices would require multiplying by 1/.0917.

the housing market in San Joaquin County.

Table 3 summarizes how the equilibrium distribution of hedonic prices varies with the number of homes in the simulation. Each row of the table presents means and standard deviations of the results from 30 Monte Carlo replications. Increasing the sample size has little impact on the distribution of equilibrium prices. The mean, standard deviation, and interquartile range are virtually unchanged as the simulated market size increases from 200 to 2000. The main difference occurs at the right tail of the distribution. This is because the prevalence of extreme draws from the joint distribution of income and preferences increases with the sample size. A small number of households with high incomes and strong preferences for housing, relative to the numeraire, bid up the prices of the largest homes. The impact of these extreme values is reflected in the difference between the actual and predicted skewness and kurtosis.

From a computational perspective, the IBA has three desirable features. First, the algorithm is quite simple. Looping over the iterative bidding process in (6.c) requires only 15 lines of code, given the specification for utility in (10). Second, the algorithm does not require working with large matrices. The iterative process in (6.c) requires calculating an $N \times 1$ vector of bids. In comparison, the linear programming approach used by Cropper et al. (1988, 1993) and Banzhaf (2003) requires iterating over a series of $N \times N$ matrices. With two thousand homes, these matrices exceed Matlab's memory capacity on a computer running a 32-bit version of Windows. Finally, the IBA is fast. With 200 homes, the average Monte Carlo replication requires 531 iterations of (6.c), which takes 5 seconds. Naturally, increasing the number of homes increases computational time. Increasing the size of the market by an order of magnitude (from 200 homes to 2000) increases the average computational time from 5 seconds to 17 minutes. While the number of iterations increases linearly, the number of calculations increases

exponentially due to the need to calculate the $N \times 1$ bid vector on each step of every iteration of the algorithm. In larger markets, where computational time may pose a constraint, one could modify the algorithm to avoid redundant calculations, such as the bids made by households with incomes below the current sale price of the home.

For comparison, the last column of table 3 reports the convergence speed of the linear programming algorithm (LPA) for sample sizes of 200, 500, and 1,000. We found that good starting values were needed to guarantee convergence. Therefore, starting values were defined by adding a small shock to the equilibrium utilities identified by the IBA.¹⁴ This ensured that both algorithms converged to virtually the same set of prices. We would expect that starting the LPA near an equilibrium would reduce computational time. Nevertheless, the IBA is still considerably faster. With 200 homes, the average Monte Carlo replication of the LPA takes 36 second to converge—seven times as long as the IBA. The relative performance of the IBA improves with the size of the simulation. In samples of 1000 homes, the LPA takes more than forty times as long to converge.

Finally, it is worth noting that we have found the IBA to perform consistently on other problems. We have calibrated the IBA to different data sets, increased the sample size to as many as 10,000 homes, and solved for equilibria using Translog and Diewert utility functions. Throughout this process, we have observed similar performance in terms of computational time and have not had any difficulty obtaining convergence.

5. Identifying a Multiplicity of Hedonic Equilibria

The IBA provides a simple way to investigate whether there are multiple hedonic equilibria.

¹⁴ The LPA systematically increases housing prices as it iterates over a series of assignment problems, each of which is formulated as a linear program. We used Matlab's linear programming solver and our code follows the formulation of the primal and dual problems outlined in Banzhaf (2003).

Recall from the corollary to proposition 2 that the IBA will converge to the highest equilibrium price vector located below the point at which the algorithm begins. Therefore, by systematically varying the algorithm's starting value, one can determine whether the data support multiple equilibria.

To test whether the San Joaquin data support multiple equilibria, we fixed the draws on households and homes and varied the algorithm's starting value. We began from the highest point in price space that satisfies each consumer's budget constraint. Then, after solving for an initial equilibrium, we increased each household's reference utility and solved for a new equilibrium. Increasing the reference utility decreases the IBA's starting value in price space, allowing us to test whether the data support equilibria at lower prices. Repeating this process over a wide range of starting values confirmed that the data do support multiple equilibria.

Figure 4 depicts the set of equilibria in a sample of 100 homes. It illustrates how the equilibrium prices that we recovered for two particular homes (#2 and #13) vary with the IBA's starting value. The horizontal axis measures the share of income spent on housing at the initial reference level of utility. The vertical axis measures the price of housing in the resulting equilibrium. We began by defining reference utility such that each household spends 99% of its income on its least favorite home. This produced an equilibrium at the prices ($p_{13} = \$11,011$, $p_2 = \$10,373$). Then we decreased the share of income spent on housing in the reference consumption bundle to 98%. This allowed us to identify a lower equilibrium at prices ($p_{13} = \$10,599$, $p_2 = \$9,910$). Equilibrium prices were unchanged when we continued to incrementally decrease the income share to 85%. However, decreasing it to 84% revealed a third equilibrium at prices ($p_{13} = \$10,328$, $p_2 = \$9,711$) where we remain until the share drops below 80%. Continuing this process until the income share reached 1% identified 23 different

equilibria.¹⁵

The pattern of results in figure 4 is representative of the entire sample. That is, all 100 homes had step functions similar to the ones shown for #2 and #13. While housing prices vary systematically across the equilibria, housing assignments typically remain the same. The average household was assigned to 1.8 different homes in the 23 different equilibria. Finally, while increasing the sample size did not affect the overall pattern of results, it did increase the number of equilibria. When we repeated the analysis for a sample of 1000 homes, we identified 85 equilibria. This seems intuitive. As the sample size increases linearly, the number of ways that we can order households over homes increases exponentially, giving us more potential assignments to satisfy the equilibrium conditions.

Our confirmation of multiplicity adds to previous findings by Bayer and Timmins (2005), who demonstrate that social interactions can lead to multiple equilibria when the household's location choice problem is defined in probabilistic terms. In contrast, our results demonstrate that multiplicity can arise in a deterministic setting without social interactions. With the evidence from both studies pointing to a multiplicity of equilibria, models of household location choice that aim to predict how markets will respond to future shocks have at least three options: (i) demonstrate the new equilibrium is unique; (ii) justify a particular equilibrium; or (iii) evaluate the sensitivity of policy implications to the set of potential equilibria. We return to this choice below in the context of our prediction for the rate that an unexpected shock to school quality will be capitalized into property values.

6. The Difference between Capitalization and MWTP in the San Joaquin Model

¹⁵ Randomly varying the order in which we iterate over homes had no effect on these results, suggesting that the equilibria follow a vector ordering.

Over the past few years, researchers have increasingly adapted the hedonic property value model to a quasi-experimental framework. Recent applications have used this strategy to investigate how housing prices react to unexpected shocks in urban amenities such as leukemia risk (Davis, 2004), air quality (Chay and Greenstone, 2005), information about airport noise (Pope, 2008a), and the presence of registered sex offenders (Linden and Rockoff, 2008; Pope, 2008b). These studies demonstrate that unexpected shocks to amenities *cause* housing prices to change. It is less clear what the resulting capitalization rates reveal about willingness to pay.

Some quasi-experimental studies make welfare calculations that effectively treat the capitalization rate as an approximation to MWTP (Davis, 2004; Chay and Greenstone, 2005; Linden and Rockoff, 2008). This interpretation is valid as long as households are identical and the shock to the amenity is marginal. However, if people are heterogeneous or the shock to the amenity is large then the theoretical equivalence between capitalization and welfare breaks down (Starrett, 1981; Bartik, 1988). In this case the sorting process that underlies hedonic equilibria can generate capitalization rates that understate or overstate MWTP. It is important to investigate the size of this difference given the current interest in using quasi-experimental methods for benefit-cost analysis (Greenstone and Gayer, 2009).

The remainder of this section uses the calibrated San Joaquin model to investigate the size of the difference between the MWTP for school quality and the rate at which an unexpected improvement is capitalized into property values. We shock the level of school quality in individual school districts, solve for new hedonic equilibria, and compare the resulting capitalization rates with the average ex-ante MWTP for the households who were “treated” by the shock.¹⁶

¹⁶ To calculate an exact welfare measure for the change in school quality, one would need to account for changes in housing prices and household location choices, as in Sieg et al. (2004) and Smith et al. (2004). The IBA provides

The multiplicity of potential post-shock equilibria is addressed by conducting a sensitivity analysis on the subset of equilibria that we judge to be plausible. First, we limit our analysis to equilibria in which the average price of housing in the improved school district *increases*. It seems reasonable to rule out a decrease in property values because the marginal utility from school quality is defined to be strictly positive. This restriction defines a lower bound on the new equilibrium prices. Then, following the logic in Bartik (1988), we define an upper bound at the point where the price of each home in the improved district increases by the maximum ex-ante willingness to pay of any household in the county. While these bounds restrict our analysis to a subset of equilibria that we believe to be plausible, we do not have a simple way to choose between them. Therefore, we calculate the capitalization rate for each of the plausible equilibria and report the minimum and maximum. These are interpreted as bounds on the range of consistent predictions for market capitalization.

Table 4 summarizes how households sort themselves across the urban landscape in the initial hedonic equilibrium that we calibrated to the San Joaquin data. For each of the seven school districts in figure 2, the table reports average values for the rental rate, household income, preferences for school quality, and the annualized willingness-to-pay for a marginal improvement in school quality. Not surprisingly, wealthier households tend to occupy more expensive homes. The two most expensive districts (Escalon and Tracy) have the highest average income, and the lowest income households tend to live in the least expensive district (Stockton). Yet the relationship between average income and the average rental rate is not strictly monotonic. Households in Lodi have lower average income and pay higher average rent than households in Manteca, for example. This reflects variation in the stock of housing within

the information needed to make these calculations. Comparing the total market capitalization of a shock with exact welfare measures would be an interesting topic for future research on using hedonic models for policy evaluation.

each school district together with negative correlation between income and preferences for housing characteristics.¹⁷ Finally, notice that households with the strongest preferences for school quality tend to locate in the districts with the highest quality schools (Manteca and Lincoln). These households also tend to have the highest annualized MWTP.

Table 5 compares the ex-ante MWTP for school quality with the market capitalization rate for two hypothetical policies that lead to 15-quantile increases for individual school districts. The first policy raises school quality in the district with the lowest initial test scores (Stockton) from the 54th quantile in the statewide distribution of test scores to the 69th quantile, holding school quality constant in all other districts. The second policy increases school quality in the district with the highest initial scores (Manteca), raising it from the 80th quantile to the 95th quantile while school quality is held constant everywhere else. After each policy experiment, the IBA is used to recover the set of plausible hedonic equilibria. For each of these equilibria, the capitalization rate is measured as (11), where D denotes the school district that experienced the improvement, q measures the school quality quantile, and the 0 and 1 superscripts distinguish between the pre-shock and post-shock equilibrium prices.

$$\left[\text{mean}_{j \in D}(p_j^1 - p_j^0) - \text{mean}_{j \notin D}(p_j^1 - p_j^0) \right] / (q_D^1 - q_D^0). \quad (11)$$

This is simply the difference between the average price changes that occur in the improved and unimproved areas, measured per unit of the improvement. Each row of table 5 reports the means and standard deviations of (11) from 30 Monte Carlo replications of each policy experiment, given a particular number of homes. As in earlier simulations, increasing the market size from

¹⁷ The average home in Lodi is larger (measured by sqft and lot size) and requires a shorter commute than the average home in Manteca. As a result, households who locate in Lodi tend to have stronger preferences for these characteristics than households in Manteca. At the same time, because of the negative correlation between preferences and income, the households in Manteca tend to have slightly higher incomes. These higher income households do not bid up the rental rates in Lodi because they prefer to live in Manteca.

200 homes to 2000 homes induces small changes in the mean capitalization rates and large decreases in their standard deviations.

While there are many plausible choices for the new equilibrium, they make similar predictions for the capitalization rate. For example, with a sample size of 200, our range of predictions for the rate at which an improvement to Stockton is capitalized into property values has an average lower bound of \$12.36 and an average upper bound of \$13.34. Figure 4 provides intuition for why this range is narrow. Large differences in absolute prices across the various equilibria correspond to small differences in relative prices because equilibrium prices tend to move together. Intuitively, when an improvement in Stockton's schools make it more attractive *relative* to other districts, equilibrium housing prices in Stockton will tend to increase *relative* to equilibrium prices in other districts, regardless of which equilibrium we choose.

The results from the improvement to Stockton illustrate the difficulty in equating capitalization rates with MWTP when the capitalization reflects a non-marginal shock and households have heterogeneous preferences and incomes. Stockton has the lowest initial test scores and the households who have chosen to live there have the lowest MWTP for an improvement (table 4). When Stockton's school quality improves, some residents from other districts would prefer to move there. Importantly, these are not the individuals with the strongest preferences for school quality. Stockton still has the lowest test scores of any district. The households who want to move there tend to have below average preferences for school quality. Nevertheless, because many of them are willing to pay more than Stockton's initial population, they bid up housing prices. In the largest version of the simulation (N=2000), the improvement is capitalized into housing prices at an annualized rate of \$11.40 to \$11.56 per 1-quantile increase in school quality. These capitalization rates are 22% to 24% higher than the annualized

MWTP for Stockton’s initial residents (\$9.35). Meanwhile, the range of values for the capitalization rate is 26% to 27% below the MWTP for the average resident in San Joaquin County (\$15.65). The bottom line is that the sorting behavior needed to clear the market following a shock causes the market capitalization rate to overstate average MWTP for the “treated” households and understate average MWTP for the population as a whole. These differences are substantial.

The pattern of results is reversed when there is an improvement to school quality in Manteca—the district with the highest initial test scores. While residents of Lincoln, Lodi, Stockton, Escalon, Ripon, and Tracy bid up housing prices in Manteca, the resulting capitalization rate reported in the last row of table 5 (\$16.87 to \$18.30) is still 13% to 20% below average MWTP for Manteca residents (\$21.01). Meanwhile, the capitalization rate is 8% to 17% higher than the average MWTP in the county. This is because the households who are interested in moving to Manteca after the improvement tend to value school quality more than San Joaquin’s average resident.

In their theoretical models of the capitalization process, Starrett (1981) and Bartik (1988) demonstrate that there are limits to what we can learn about how much households are willing to pay for an amenity from analyzing data on the rate that changes in that amenity are capitalized into housing prices. Our simulations demonstrate that this theoretical point can be empirically important for property value studies that use quasi-experiments to assess the value of urban amenities.

5. Conclusions and Future Research

For more than 40 years, economists have used hedonic models to measure the values that

homebuyers implicitly assign to urban amenities. Given the importance of this work for public policy, there has been surprisingly little effort to validate the resulting predictions. We have provided the means to address this gap in the literature by developing an iterative bidding algorithm for computing hedonic equilibria. The algorithm can be calibrated to approximately reproduce the actual prices in a large micro dataset as a hedonic equilibrium.

Our application of the model revealed three important points about the empirical properties of hedonic equilibria. First, there can be multiple equilibria in a conventional hedonic model without social interactions. Second, the rate at which an unexpected shock to an amenity is capitalized into property values can differ substantially from average MWTP. Finally, the existence of multiple equilibria need not limit the robustness of prediction. Our results on the difference between capitalization and MWTP are robust to the choice of a new equilibrium. That said, we must also stress that our results do not prove there will *always* be multiple equilibria or that capitalization rates will *never* approximate MWTP. Our results are conditioned by the features of our model. This includes the size of the shock being considered, the definition for the spatial landscape, the preference specification, and the standard assumptions of free mobility and perfect information. Future research on multiplicity and capitalization could use the IBA to investigate the properties of equilibria in a variety of settings.

The IBA could also be used to conduct Monte Carlo experiments on the performance of hedonic estimators. Previous studies have used Wheaton's (1974) linear programming algorithm for this task. Cropper et al. (1988) compare the accuracy of predictions for MWTP made by competing specifications for the hedonic price function, Cropper et al. (1993) compare hedonic and logit methods, Banzhaf (2003) investigates the impact of discreteness in the availability of an amenity, and Klaiber and Smith (2009) investigate quasi-experimental approaches to policy

evaluation. While all of these studies provide valuable insights, they only begin to address the issues associated with the current generation of hedonic estimators. Empirical hedonic studies increasingly use spatial fixed effects, spatial error and spatial lag models, semiparametric and nonparametric methods, quasi-experimental designs, and structural estimators. Which methods will provide the most accurate estimates for MWTP when some variables are omitted and others are measured with error? Which structural estimators will be least sensitive to misspecification of the preference function? The IBA provides the means to answer these and other important questions using sample sizes that are consistent with recent empirical studies.

Finally, we feel there is some potential to use the IBA as a tool for policy evaluation. The ability to calibrate preference functions to reproduce the actual distribution of prices in any given housing market offers a micro-level alternative to the calibrated equilibrium models of community location choice and school quality developed by Fernandez and Rogerson (1998), Nechevba (2000), and others. Our framework generalizes these models in the sense that it recognizes households do not just select a community. They select an individual home *within* a community. Likewise, households may differ in both their incomes and in their relative preferences for multiple housing characteristics and public goods. While our model could be characterized as the first calibrated simulation framework to simultaneously capture all of these features, it is more rigid than some existing studies in the sense that it treats the provision of local public goods as exogenous. A key feature of Fernandez and Rogerson (1998), Nechevba (2000), and also Walsh (2007), is the ability to model how public goods are determined endogenously as an equilibrium outcome of the sorting process. For example, the quality of local public schools may depend on the distribution of income and education among the households who live in the communities that comprise the school district. Generalizing the

iterative bidding algorithm to model this endogeneity would extend the literature on calibrated equilibrium models by providing a means to investigate the micro-level implications of public policy changes that affect markets for housing.

Appendix: Data and Code for Simulation

The data and Matlab code needed to reproduce the examples and simulation results reported in this paper are provided in a supplemental appendix. In order to provide micro data on the stock of housing in San Joaquin County without violating our contractual obligations to the commercial vendor, we have modified the price of each home, as well as its structural characteristics. Specifically, we added random shocks to each of the following variables: sale price, number of bedrooms, lot size, and age. The resulting fake data depict a housing market with a distribution of homes that is similar to San Joaquin's. Nevertheless, no individual observation is identical to any actual housing transaction in San Joaquin County. Therefore, the fake data we provide online have no commercial or research value beyond enabling others to reproduce our simulation results.

References

- Bajari, P., Kahn, M.E., 2005. Estimating Housing Demand with an Application to Explaining Racial Segregation in Cities. *Journal of Business and Economic Statistics* 23 (1), 20-33.
- Banzhaf, H.S., 2003. Hedonic Pricing in Realistic Urban Structures, or What if Tiebout Called and Nobody Sorted? Working Paper, Resources for the Future.
- Bartik, T.J., 1988. Measuring the Benefits of Amenity Improvements in Hedonic Price Models. *Land Economics* 64 (2), 172-183.
- Bayer, P., Timmins, C., 2005. On the Equilibrium Properties of Locational Sorting Models. *Journal of Urban Economics* 57, 462-477.
- Bayer, P., Timmins C., 2007. Estimating Equilibrium Models of Sorting across Locations. *The Economic Journal* 117, 353-374.
- Black, S.E., 1999. Do Better Schools Matter? Parental Valuation of Elementary Education. *Quarterly Journal of Economics* 114 (2), 577-599.
- Brookshire, D.S., Thayer, M.A., Tschirhart, J., Schulze, W.D., 1985. A Test of the Expected Utility Model: Evidence from Earthquake Risks. *Journal of Political Economy* 93 (2), 369-389.
- Chay, K.Y., Greenstone, M., 2005. Does Air Quality Matter? Evidence from the Housing Market. *Journal of Political Economy* 113 (2), 376-424.
- Cropper, M.L., Deck, L.B., McConnell, K.E., 1988. On the Choice of Functional Form for Hedonic Price Functions. *Review of Economics and Statistics* 70 (4), 668-675.
- Cropper, M.L., Deck, L., Kishor, N., McConnell, K.E., 1993. Valuing Product Attributes Using Single-Market Data: A Comparison of Hedonic and Discrete-Choice Approaches. *Review of Economics and Statistics* 75 (2), 225-232.
- Davis, L., 2004. The Effect of Health Risk on Housing Values: Evidence from a Cancer Cluster. *American Economic Review* 94 (5), 1693-1704.
- Dornbusch, D.M., Barrager, S.M. 1973. Benefits of Water Pollution Control on Property Values, Office of Research and Monitoring, U.S. Environmental Protection Agency, EPA-600/5-73-005.
- Epple, D., Sieg, H., 1999. Estimating Equilibrium Models of Local Jurisdiction. *Journal of Political Economy* 107 (4), 645-681.
- Epple, D., Romer, T., 1991. Mobility and Redistribution. *Journal of Political Economy* 99 (4), 828-858.

- Ekeland, I., 2008. Existence, Uniqueness and Efficiency of Equilibrium in Hedonic Markets with Multidimensional Types. Working Paper.
- Fernandez, R., Rogerson, R., 1998. Public Education and Income Distribution: A Dynamic Quantitative Evaluation of Education-Finance Reform. *American Economic Review* 88 (4), 813-833.
- Greenstone M., Gayer T., 2009. Quasi-Experimental and Experimental Approaches to Environmental Economics. *Journal of Environmental Economics and Management* 57 (1), 21-44.
- Klaiber, A.H., Smith, V.K. 2009. Evaluating Rubin's Causal Model for Measuring the Capitalization of Environmental Amenities. NBER Working Paper 14957.
- Lancaster, K.J., 1966. A New Approach to Consumer Theory. *Journal of Political Economy* 74 (2), 132-157.
- Linden, L., Rockoff, J.E., 2008. Estimates of the Impact of Crime Risk on Property Values from Megan's Laws. *American Economic Review* 98 (3), 1103-1127.
- Nechyba, T., 2000. Mobility, Targeting, and Private-School Vouchers. *American Economic Review* 90 (1), 130-146.
- Palmquist, R.B., Smith, V.K., 2002. The Use of Hedonic Property Value Techniques for Policy and Litigation. in: T. Tietenberg, and H. Folmer, (Eds.), *The International Yearbook of Environmental and Resource Economics 2002/2003*, Edward Elgar, Cheltenham, UK, pp. 115-164.
- Poterba, J.M., 1992. Housing and Taxation: Old Questions, New Answers. *American Economic Review* 82 (2), 237-242.
- Pope, J.C., 2008a. Buyer Information and the Hedonic: The Impact of a Seller Disclosure on the Implicit Price for Airport Noise. *Journal of Urban Economics* 63, 498-516.
- Pope, J.C., 2008b. Fear of Crime and Housing Prices: Household Reactions to Sex Offender Registries. *Journal of Urban Economics* 64, 601-614.
- Ridker, R., 1967. *Economic Costs of Air Pollution*, Frederick A. Praeger Publishers, New York.
- Rosen, S., 1974. Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition. *Journal of Political Economy* 82 (1), 34-55.
- Sieg, H., Smith, V.K., Banzhaf, H.S., Walsh, R., 2004. Estimating the General Equilibrium Benefits of Large Changes in Public Goods. *International Economic Review* 45 (4), 1047-1077.

Smith, V.K., Sieg, H., Banzhaf, H.S., Walsh, R., 2004. General Equilibrium Benefits for Environmental Improvements: Projected Ozone Reductions under EPA's Prospective Analysis for the Los Angeles Air Basin. *Journal of Environmental Economics and Management* 47 (3), 559-584.

Starrett, D.A., 1981. Land Value Capitalization in Local Public Finance. *Journal of Political Economy* 89 (2), 306-327.

Walsh, R.L., 2007. Endogenous Open Space Amenities in a Locational Equilibrium. *Journal of Urban Economics* 61 (2), 319-344.

Wheaton, W.C., 1974. Linear Programming and Locational Equilibrium: The Herbert-Stevens Model Revisited. *Journal of Urban Economics* 1, 278-287.

Table 1A

Utility from housing characteristics: 3-home example

	Household		
	A	B	C
income	68,910	64,500	57,000
<u>utility from the characteristics of:</u>			
home 1	12.1	25.7	17.0
home 2	10.3	20.3	10.3
home 3	10.8	21.6	9.7

Table 1BTracking the progress of the iterative bidding algorithm: 3-home example ^a

IBA Iteration	Home	Price	Bids by Household:			Corresponding Figure
			A	B	C	
1	1	64,500	68,909	64,499	56,999	1C
1	2	56,118	43,398	64,265	56,117	1D
1	3	55,444	53,276	62,298	55,443	1E
2	1	64,354	64,500	64,353	56,999	not shown
2	2	42,556	42,555	30,032	56,117	not shown
2	3	52,760	52,759	55,444	31,550	not shown
.	
.	
.	
13	1	64,308	64,308	64,307	56,983	1F
13	2	42,289	42,288	19,193	42,289	1F
13	3	52,597	52,596	52,597	31,079	1F
14	1	64,308	64,308	64,307	56,983	1F
14	2	42,289	42,288	19,193	42,289	1F
14	3	52,597	52,596	52,597	31,079	1F

^a Bold bids identify the household assigned to the home being auctioned.

Table 2Summary statistics and regression results for San Joaquin County, CA ^a

		Mean	Std. Dev.	Min	Max	Coef	t-stat
	saleprice (\$1000)	128	51	30	372	----	----
structural characteristics of home	bedrooms (#)	3.08	3.08	3.08	3.08	-4,627	-7.05
	building size (1000 sqft)	1.55	0.50	0.33	6.09	64,634	44.78
	lot size (1000 sqft)	7.72	6.78	1.32	214.32	661	5.36
	age (years)	25.92	20.96	1.00	98.00	-400	-19.46
Census tract demographics	mean time-to-work (minutes)	29.90	8.47	15.86	59.74	858	10.91
	median household income (\$1000)	47.48	15.35	11.19	85.00	401	11.29
	population under 18 (%)	31.15	5.54	12.10	43.96	-1,795	-20.88
Public goods and amenities	distance to grazing land (km)	8.28	4.40	0.03	16.01	-1,141	-10.09
	distance to water (km)	7.74	5.49	0.00	27.26	857	10.97
	STAR 10th grade math score (%)	70.53	10.50	54.00	80.00	118	3.07

^a Data contain 9,634 observations on homes sold between 1995 and 1998. Prices and income are measured in constant \$1999 dollars. Regression results are from a linear model with all variables measured in levels. T-statistics are based on heteroskedasticity-robust standard errors. $R^2=0.74$.

Table 3

Monte Carlo analysis of the impact of market size on equilibria mean (standard deviation) from 30 replications

N	min	5th	25th	50th	75th	95th	max	mean	std	skew	kurt	iter ^a	time ^a	time ^a
<u>Estimated Distribution of Equilibrium Hedonic Prices (\$1,000)</u>												<u>IBA</u>	<u>IBA</u>	<u>LPA</u>
200	5.2 (0.5)	5.9 (0.6)	8.6 (0.6)	11.2 (0.6)	15.0 (0.7)	23.4 (1.7)	36.3 (5.0)	12.5 (0.5)	5.5 (0.5)	1.4 (0.3)	5.6 (1.5)	531 (276)	5 (2)	36 (8)
500	5.1 (0.2)	5.8 (0.2)	8.7 (0.3)	11.2 (0.3)	15.1 (0.4)	23.9 (0.9)	42.2 (4.3)	12.6 (0.3)	5.7 (0.3)	1.5 (0.1)	6.5 (1.2)	929 (425)	35 (16)	795 (227)
1000	5.0 (0.1)	5.8 (0.2)	8.7 (0.2)	11.3 (0.3)	15.1 (0.3)	23.9 (0.6)	48.0 (7.2)	12.6 (0.2)	5.8 (0.2)	1.6 (0.2)	7.1 (2.0)	1,852 (1,786)	238 (229)	10,175 (4,501)
2000	5.0 (0.0)	5.7 (0.1)	8.6 (0.2)	11.2 (0.2)	15.1 (0.3)	23.8 (0.5)	52.6 (10.4)	12.5 (0.2)	5.8 (0.1)	1.6 (0.2)	7.7 (2.3)	2,150 (663)	1,016 (328)	--- ---
<u>True Distribution of Prices (\$1,000)</u>														
9634	2.5	5.5	8.3	11.1	14.2	20.7	35.6	11.7	4.7	1.0	4.7	---	---	---

^a *Iter* is the number of iterations of equation (6.c) required for convergence to an equilibrium and *time* is the computational time (in seconds) on a standard desktop computer with a Pentium 4 processor running a 32-bit version of Windows for the iterative bidding algorithm (IBA) developed here and the linear programming algorithm (LPA) used in previous studies.

Table 4Equilibrium sorting across school districts (N=2000) ^a

School District	Math Score Quantile	rent	income	α_{school}	MWTP
Lincoln	78.3%	12,080	53,232	2.92	24
Lodi	74.3%	13,274	57,697	1.81	16
Stockton	53.6%	9,136	33,557	1.08	9
Escalon	74.3%	16,248	81,317	1.75	18
Manteca	80.1%	12,785	58,056	2.44	21
Ripon	73.5%	13,270	59,011	1.60	15
Tracy	78.7%	15,787	71,739	1.88	17

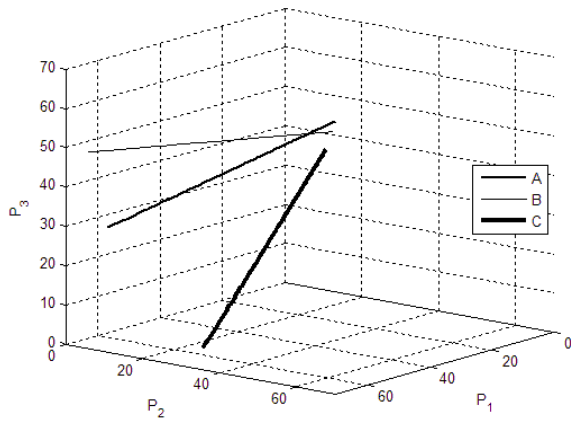
^a The figures in the table reflect an average over 30 Monte Carlo replications.

Table 5

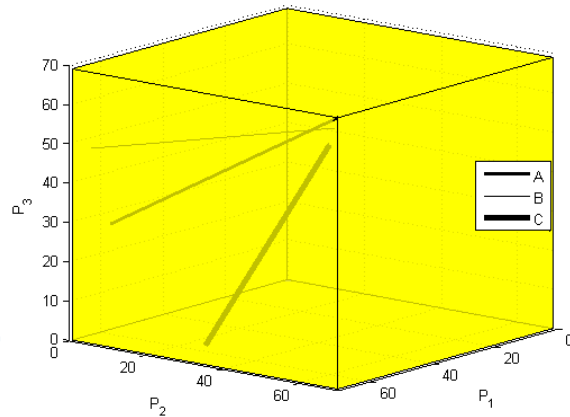
Comparison between capitalization rates and marginal willingness to pay
(means and standard errors from 30 Monte Carlo replications)

N	Average MWTP for households in:			Capitalization Rates for 15 Point Improvement to:			
	Stockton (score=53.6)	Manteca (score=80.1)	San Joaquin County	Stockton		Manteca	
				minimum ^a	maximum ^a	minimum ^a	maximum ^a
200	10.36 (1.65)	19.15 (3.70)	15.61 (1.42)	12.36 (2.70)	13.34 (3.47)	16.66 (2.99)	17.86 (3.45)
500	10.19 (1.38)	19.79 (2.20)	15.52 (0.72)	12.35 (1.88)	12.53 (2.03)	16.06 (2.00)	17.16 (2.53)
1000	9.44 (0.68)	21.01 (1.74)	15.76 (0.55)	11.25 (1.58)	11.54 (1.65)	17.13 (1.43)	18.43 (1.52)
2000	9.35 (0.47)	21.01 (1.06)	15.65 (0.40)	11.40 (0.75)	11.56 (0.79)	16.87 (1.11)	18.30 (1.15)

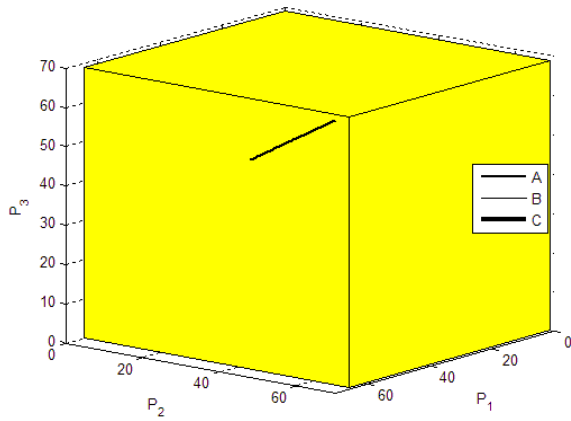
^a The minimum and maximum capitalization rates are calculated over the set of potential equilibria on each Monte Carlo replication.



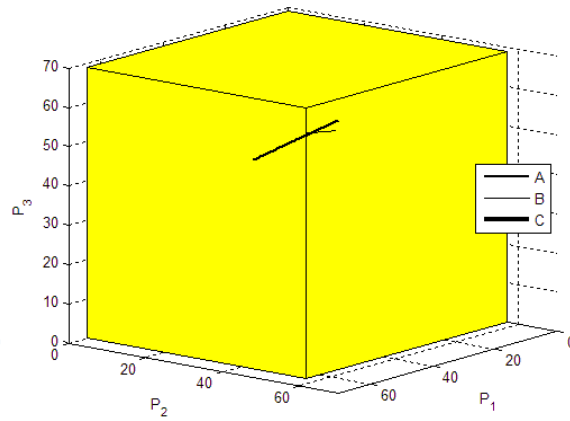
A. Three Bid Indifference Strings



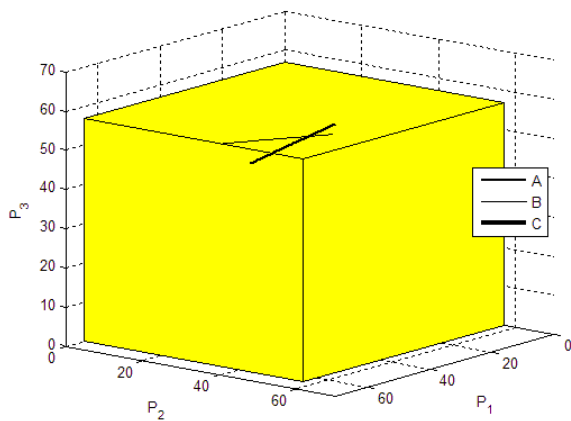
B. Initial Hyperrectangle



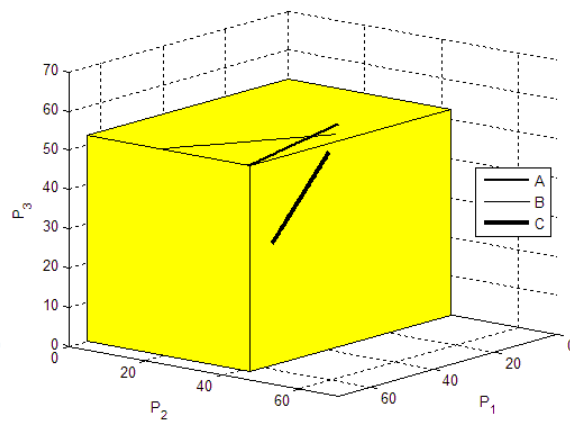
C. First Auction for Home 1



D. First Auction for Home 2



E. First Auction for Home 3



F. Convergence of the IBA

Fig. 1. A graphical depiction of the iterative bidding algorithm (axes measure \$1000)

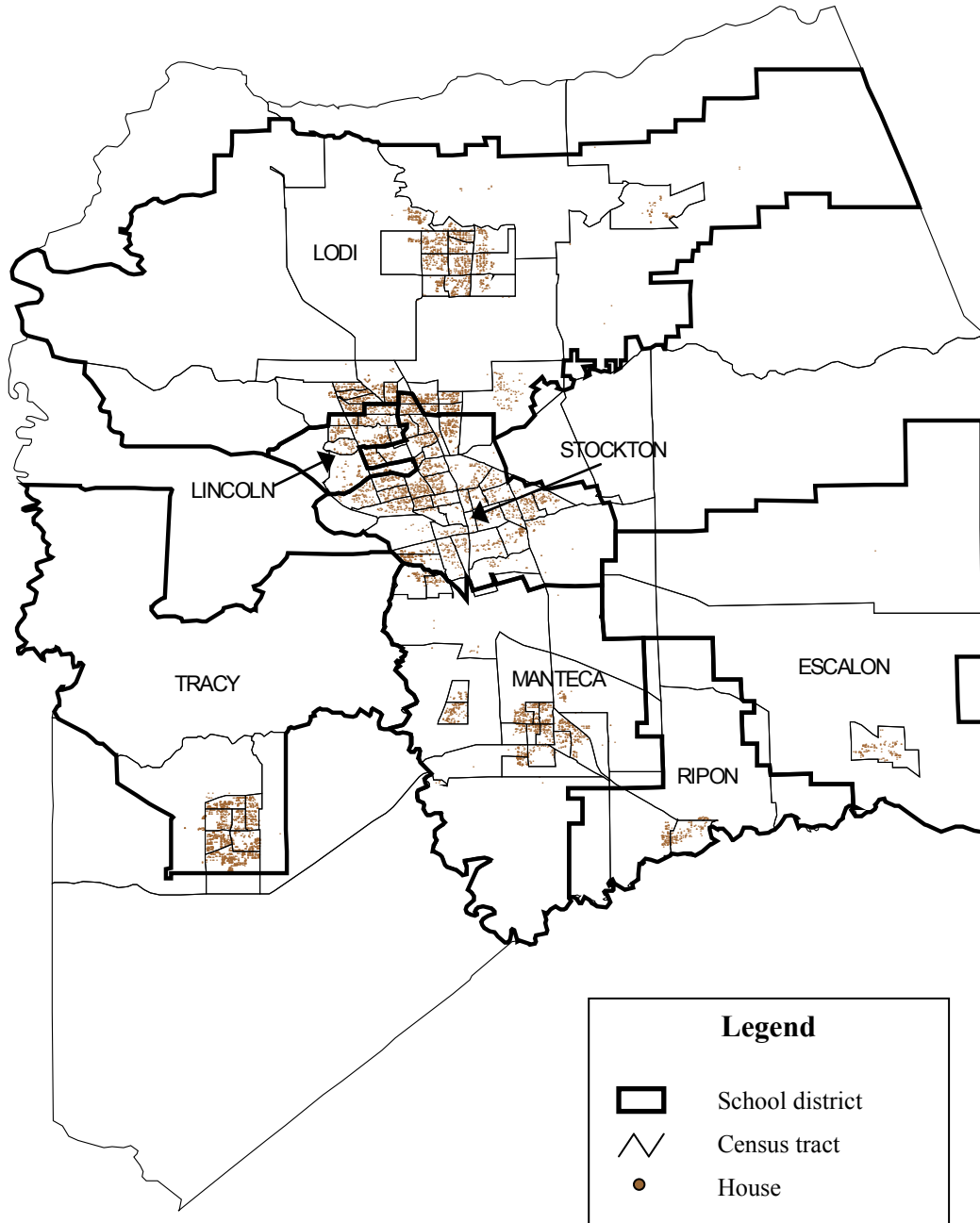
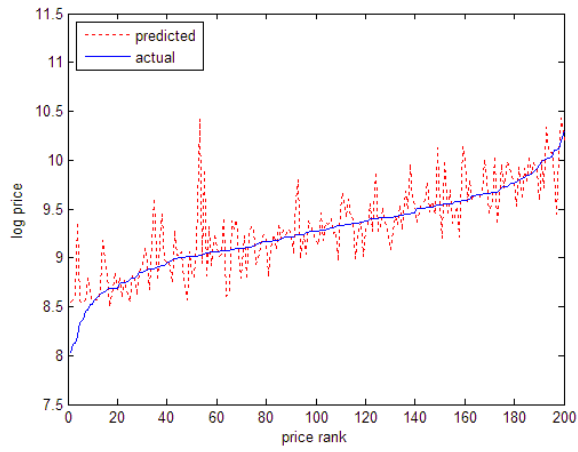
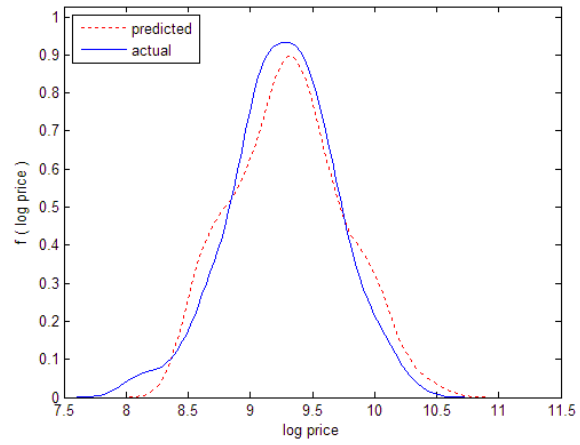


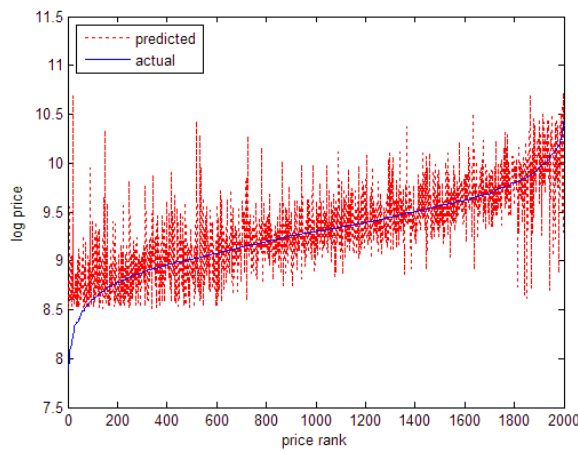
Fig. 2. San Joaquin County, California



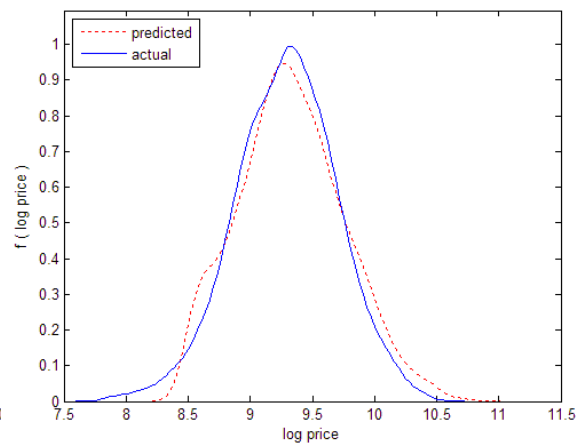
A. Deviations from CDF ($N = 200$)



B. Deviations from PDF ($N = 200$)



C. Deviations from CDF ($N = 2000$)



D. Deviations from PDF ($N = 2000$)

Fig. 3. Reproducing the distribution of housing prices in San Joaquin County, California

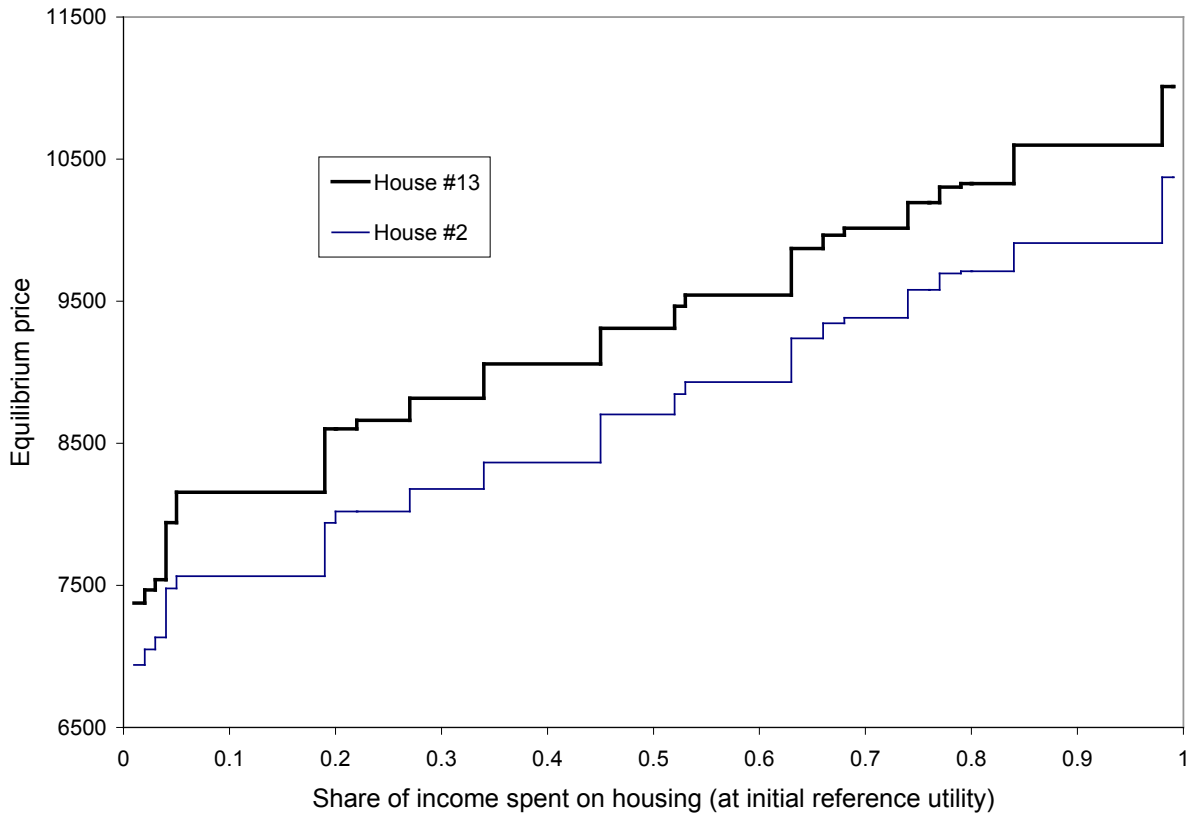


Fig. 4. Multiplicity of equilibria in a sample of 100 homes