

# The Value of Residential Land and Structures during the Great Housing Boom and Bust\*

## **Nicolai V. Kuminoff**

Department of Economics  
Arizona State University  
PO Box 879801  
Tempe, AZ 85287-9801  
Phone: 480-727-9802  
Email: [kuminoff@asu.edu](mailto:kuminoff@asu.edu)

## **Jaren C. Pope**

Department of Economics  
Brigham Young University  
180 Faculty Office Building  
Provo, UT 84602-2363  
Phone: 801-422-2037  
Email: [jaren\\_pope@byu.edu](mailto:jaren_pope@byu.edu)

Draft with revisions for *Land Economics*: April, 2012

\* We thank two anonymous referees, Dan Bromley, Leah Brooks, Morris Davis, Paul Ellickson, Gerald Korngold, Byron Lutz, David C. Lincoln, Elizabeth Plummer, V. Kerry Smith, David Wildasin, Kent Zhao, and seminar participants at the David C. Lincoln Fellowship Symposium and the Annual Conference of the National Tax Association for valuable comments and suggestions on this research. We gratefully acknowledge research support by the Lincoln Institute of Land Policy.

# The Value of Residential Land and Structures during the Great Housing Boom and Bust

## Abstract

This study examines how the value of residential land and structures evolved during the great housing boom and bust using data on more than a million residential properties that were sold in ten metropolitan areas between 1998 and 2009. We use a hedonic estimator to disentangle the market value of land and structures at a local (census tract) level. Our estimates reveal substantial heterogeneity in the evolution of the market value of land and structures *within* metropolitan areas. Surprisingly, lower value land at the urban fringes of metropolitan areas was the most volatile during the boom-bust.

*JEL Codes:* R14, R21, R31

*Keywords:* Land Value, Hedonic Method, Property Values

## 1. Introduction

Housing is a major source of wealth in the United States. The Federal Reserve's Flow of Funds Report documents that the asset value of owner-occupied housing for the entire U.S. was approximately 23 trillion dollars in 2006—more than the capitalized value of the NYSE, Amex, and Nasdaq exchanges combined. A house's value can be decomposed into two components: the value of the land on which the house is built, and the value of the structures that comprise the house itself. Decomposing property value into the value of land and structures is important for several reasons. First, some cities and counties tax land and structures at different rates (Chapman et al. 2009; Banzhaf and Lavery 2010; Cho, Lambert, and Roberts 2010). Successful implementation of a split-rate tax requires accurate estimates for each component of value. Second, structures depreciate differently from land. Documenting this difference is necessary for calculating tax code allowances for depreciation and for insurance companies to reimburse homeowners for damaged structures. Third, understanding how the value of land has evolved relative to the value of structures may help households, banks, and local governments to manage risk within their financial portfolios. Finally, tracking the evolution of land and structural values within and across metro areas may provide insights into the forces that drive boom-bust cycles in real estate.

There are three primary techniques for decomposing property values into the value of land and structures. The “teardown” approach derives land value from the sales prices (plus demolition costs) of properties that were purchased with the intention to tear down the existing structures (e.g. Rosenthal and Helsley 1994, Dye and McMillen 2007). The “replacement cost” approach infers land value from the difference between property values and the depreciated costs of replacing the structures on the property (e.g. Davis and Heathcoat 2007, Davis and Palumbo 2008). Finally, the “hedonic” approach estimates land value by regressing the sales prices of properties on the characteristics of the land and structures. In this case, the land value is commonly defined as the marginal implicit price per square foot of a lot (e.g. Clapp 1980; Glaeser, Gyourko, and Saks 2005). Bell et al. (2009) provide an overview of the three methods, describing their strengths and weaknesses. The difficulty with the teardown approach is the sparseness

of the data. There are too few teardowns to apply the methodology at a high level of spatial resolution in large geographic areas. Bell et al. describe a more nuanced set of tradeoffs between the replacement cost and hedonic approaches, but conclude that “the contribution [hedonic] principle of value seems more consistent with the notion of market value.”<sup>1</sup> Despite this conceptual advantage, the replacement cost approach is more commonly used. This is partly due to its apparent simplicity and relatively low data requirements. However, it also reflects concerns that omitted variables may confound hedonic estimates of land value.<sup>2</sup> If the houses located in neighborhoods with higher land values also tend to have nicer physical characteristics that are not readily observed by the analyst (e.g. granite countertops, hardwood floors) then the value of land will be confounded with the value of unobserved physical characteristics.

The objective of this paper is twofold: (i) to refine the hedonic approach to estimating land value in a way that mitigates omitted variable bias, and (2) to use our refined methodology to analyze how the market value of land and structures evolved within and across major metropolitan areas during the 2000’s. Our data describe the sale prices, physical attributes, and geographic locations of more than a million houses that were sold in ten metropolitan areas between 1998 and 2009: Boston, Cincinnati, Detroit, Los Angeles, Oakland, Philadelphia, Pittsburgh, San Francisco, San Jose, and Tampa. According to the S&P / Case-Shiller repeat sales index, residential property values in major metro areas more than doubled between 1998 and 2006 and then declined by approximately 40% between 2006 and the end of 2009 (figure 1).<sup>3</sup> Our data span this remarkable boom-bust cycle.

The location of each house conveys access to a specific bundle of local public goods and also defines the commuting opportunities that would be faced by a working household. These localized amenities may be in limited supply due to zoning regulations and other forms of development restrictions (Glaeser, Gyourko, and Saks 2005). As a result, it is important to recognize that land values may vary across neighborhoods within a metro area. Equally important is the need to recognize that the market values of land and structures may evolve differently over time. The relative price of land may increase over time as developable land becomes scarcer. Like-

wise, changes in credit constraints or wealth may alter the relative demands for the public and private attributes of housing in ways that differ across time and space.

To characterize the spatiotemporal variation in the value of land and structures, we estimate housing price functions for each metro area, while allowing the shape of the price function to change from year to year. Our estimator extends the conventional hedonic approach in two ways. First, we use fixed effects for Census tracts to capture spatial variation in localized amenities that contribute to land value through a parcel's location, rather than its size. Second, we add interactions between the fixed effects and square footage of living space to capture spatial variation in latent attributes of structures. We then generate estimates for annual average values of land and structures at the level of an individual Census tract. Our spatially explicit estimates are typically an order of magnitude larger than estimates based on the conventional hedonic approach.

When we aggregate our tract level estimates up to the level of a metro area they are generally consistent with the metro area averages reported prior to the boom by past studies using the replacement cost method (e.g. Davis and Palumbo 2008). The same is true after the bust. However, the two sets of estimates diverge during the boom-bust period. Our estimates for land values do not rise as fast during the boom or fall as quickly during the bust. Our estimates imply that the market value of structures exceeded their replacement cost during the height of the boom. The differences can be large—up to 100% for San Francisco. One potential explanation is that local housing markets are less than perfectly competitive. With a small share of houses on the market at any one time, the unique bundle of amenities provided by a desirable neighborhood may allow home sellers to command a markup on the structural characteristics of their houses, as Taylor and Smith (2000) first observed. Indeed, we find that neighborhoods with higher pre-boom land values (presumably the higher-amenity neighborhoods) had larger markups on structures during the boom. Over time, we would expect these markups to stimulate new construction, following the general logic of Tobin's q-theory.

Consistent with Davis and Palumbo's (2008) analysis of the variation *between* metro

areas we find that land values are more volatile in metro areas where the supply of housing is less elastic. Interestingly, we find the opposite pattern *within* metro areas. Neighborhoods at the urban fringe, where we would expect the supply of housing to be most elastic, were the neighborhoods that experienced the most volatility in housing prices and land values during the boom and bust. This general pattern can be seen in the Case-Shiller index. Figure 2 displays indices for the lowest, middle, and highest tier of houses (ranked by 2010 value) for Miami, San Francisco, Boston, and Atlanta. Within each metro area, the lowest value houses were the most volatile and the highest value houses were the least volatile.<sup>4</sup> We find that the higher value houses tend to be located closer to the city where the supply of land is least elastic and the lower value houses tend to be located at the outskirts of the surrounding suburbs where most of the new housing is built. This suggests that factors other than supply elasticity of housing are playing an important role in the evolution of land and structural values.

Overall, this research makes three contributions to the literature on land valuation. First, we use high-resolution spatial fixed effects to refine the conventional hedonic approach to decomposing housing prices into the implicit value of land and structures. Second, we use a consistent estimation strategy to provide new estimates for how land values evolved within and across several metro areas during the remarkable boom-bust cycle of the 2000's. Finally, while prior studies have reported that land values vary across space and time, we document two novel features of this variation that deserve more attention in future research: (i) the least valuable land at the urban fringes of metro areas was the most volatile during the recent boom-bust cycle; and (ii) the market value of structures exceeded construction costs during the boom, with the largest markups occurring in the most affluent neighborhoods.

The remainder of the paper proceeds as follows. Section 2 presents a simple model of the market for housing and uses it to define “land value” and “structure value” in the context of a hedonic price function. Section 3 explains our econometric approach. Section 4 summarizes the data we have assembled. Section 5 presents results. Section 6 discusses the implications of our findings and directions for future research, and section 7 concludes.

## 2. The Market Value of Land and Structures in a Metropolitan Area

We begin from a standard description of residential sorting. Heterogeneous households are assumed to choose from a stock of houses with different lot sizes and structural characteristics (e.g. bedrooms, bathrooms, sqft). Their collective location choices will in turn influence the supply of neighborhood amenities (e.g. public school quality, commute time to the city center, preservation of open space) through a combination of voting, social interactions, and feedback effects.<sup>5</sup> Formally, an individual household's utility maximization problem is

$$\max_{j,k} U_i(g_{kt}, l_{jk}, x_{jkt}, b; \alpha_{it}) \quad \text{subject to} \quad y_{it} = b + P_{jkt}. \quad (1)$$

In period  $t$ , household  $i$  selects one of  $j = 1, \dots, J_k$  houses located in one of  $k = 1, \dots, K$  neighborhoods. Their utility depends on the lot size of their parcel ( $l$ ), the structural characteristics of their house ( $x$ ), the amenities provided by their neighborhood ( $g$ ), and on the income they have left over to spend on the numeraire good ( $b$ ) after they pay the annualized after-tax price of housing ( $P_{jkt}$ ). The household's idiosyncratic preferences are represented by  $\alpha_{it}$ .

Sellers in this market may include a mix of developers and individuals selling their houses. There is no need to be more specific about the supply side of the market. Under a pair of weak restrictions on consumer preferences, any market outcome consistent with utility maximizing behavior can be described by a hedonic price function.

ASSUMPTION 1.

- a.  $U_i(g_{kt}, l_{jk}, x_{jkt}, b; \alpha_{it})$  is strictly increasing in  $b$  for all  $b \in (0, y_{it})$ .
- b. Let  $\succeq_i$  represent household  $i$ 's preference ordering over all potential location choices that satisfy their budget constraint.  $\succeq_i$  is invariant to  $i$ 's actual location choice.

The first condition is self explanatory. The second condition simply limits the scope for any one

household to influence prices or the supply of neighborhood amenities. For example, suppose household  $i$  has exceptionally bright children. If  $i$  were to move from their current house in school district R to a new house in school district S, then school quality may increase marginally in S due to peer effects, and decrease marginally in R. These adjustments may be followed by changes in housing prices. Condition  $b$  implies that these changes must be sufficiently small to leave  $i$ 's preference ordering over the two houses unchanged.<sup>6</sup> The need for this restriction becomes apparent in the proof of theorem 1, which is simply a variation on results derived by Bajari and Benkard (2005).<sup>7</sup>

**THEOREM 1.** Suppose that assumption 1 holds for every household. Then for any two houses,  $j, k$  and  $r, s$ , it must be true that  $P_{jkt} = P_{rst}$  if  $g_{kt} = g_{st}$ ,  $l_{jk} = l_{rs}$ , and  $x_{jkt} = x_{rst}$ .

*Proof.* Suppose  $i$  chooses  $j, k$  given  $P_{jkt} > P_{rst}$ . Then  $U_i(g_{kt}, l_{jk}, x_{jkt}, y_{it} - P_{jkt}; \alpha_{it}) < U_i(g_{st}, l_{rs}, x_{rst}, y_{it} - P_{rst}; \alpha_{it})$  because utility is strictly increasing in the numeraire. This preference ordering is invariant to whether  $i$  locates in  $j, k$  or  $r, s$ . Therefore,  $j, k$  cannot be a utility-maximizing location for  $i$  in period  $t$ , which is a contradiction. QED

Theorem 1 states that property values will be functionally related to neighborhood amenities, lot sizes, and structural housing characteristics during a single period. Relative to the majority of the empirical literature that maintains the assumptions of Rosen's (1974) hedonic model as a basis for measuring the willingness to pay for urban amenities, theorem 1 is notable for what it does not assume. We do not require the market to be perfectly competitive. Nor do we require households to be free to choose continuous quantities of every housing characteristic in every neighborhood. Some characteristics may be approximately continuous; others may be discrete. For example, households may be free to choose square footage continuously over some range, whereas public school quality may change discretely as one crosses the border between two



adjacent school attendance zones (Black 1999).

Relaxing continuity and perfect competition means that we lose the ability to interpret equilibrium marginal implicit prices for characteristics as measures that identify consumers' marginal willingness to pay for those characteristics or firms' marginal costs (Feenstra 1995, Taylor and Smith 2000, Kuminoff, Smith, and Timmins 2010). If discreteness in the choice set prevents buyers from setting their marginal rates of substitution between characteristics equal to the corresponding implicit price ratios, then the market clearing prices for individual characteristics may understate or overstate individual marginal willingness to pay. Likewise, if sellers can exert some degree of market power, then implicit prices for specific characteristics may not equal their marginal costs.

The reason for relaxing continuity and perfect competition in our conceptual model is that in reality some neighborhoods are constrained by geographic features such as water bodies, steep terrain, and public land that can produce discreteness in the supply of amenities, whereas other neighborhoods use zoning regulations to explicitly constrain further urban development. If a constrained neighborhood provides access to a unique bundle of amenities, then the amenity bundle may convey market power to property owners in the neighborhood (e.g. see Taylor and Smith, 2000). We return to this point in section 6, where we consider local market power as a potential explanation for some of our results.

Our specification for the hedonic price function,  $P_{jkt} = P(g_{kt}, l_{jk}, x_{jkt})$ , describes a spatial landscape at a single point in time where prices, amenities, and location choices are all defined such that no household would prefer to move, given its income and preferences. This is a single-period snapshot of market outcomes; it may or may not be a long-run steady state. Current period incomes and preferences may reflect temporary macroeconomic factors. Credit may be unusually easy (or difficult) to obtain relative to a long run equilibrium. The average household may be unusually optimistic (or pessimistic) about the future asset value of housing. Budget constraints may reflect other temporary macroeconomic shocks. As all of these factors change

over time, households may adjust their behavior in ways that alter the shape of the price function and generate boom-bust cycles.

During a boom-bust cycle, the evolution of the price function can be decomposed into changes in the market value of land and structures. To illustrate this, we first define the market value of a property at a single point in time as its current annualized price.

DEFINITION 1.  $P_{jkt} \equiv P(g_{kt}, l_{jk}, x_{jkt})$  is the *market value* of property  $j,k$  in period  $t$ .

The value of the underlying land is then defined by the thought experiment where we remove all of the structural characteristics from the property.

DEFINITION 2.  $LV_{jkt} \equiv P(g_{kt}, l_{jk}, 0)$  is the *land value* of property  $j,k$  in period  $t$ .

$LV_{jkt}$  measures what a vacant (but otherwise identical) parcel to  $j$  would sell for in the same neighborhood.<sup>8</sup> This definition of land value captures the spatial tradeoff between commuting costs and accessibility to the city center (Alonso 1964; Muth 1969; Mills 1967) as well as the value of local public goods and urban amenities conveyed by the neighborhood (Tiebout 1956).<sup>9</sup>

Finally, subtracting land value from total market value yields the value associated with a property's structural characteristics,  $x_{jkt}$ .

DEFINITION 3.  $SV_{jkt} = P_{jkt} - LV_{jkt}$  is the *structural value* of property  $j,k$  in period  $t$ .

While it is conceptually straightforward to decompose property value into the value of land and structures, empirical implementation presents several challenges.

### 3. Estimating the Market Value of Land and Structures

#### 3.1 Background

If life were more like a laboratory experiment, there would be no need to *estimate* land values.

Sales of vacant parcels would be randomly distributed throughout metropolitan areas and we would simply measure their transaction prices. The problem, of course, is that vacant land sales typically occur at the fringes of urban areas. We rarely observe such transactions occurring in built-up neighborhoods. In an established neighborhood, the closest substitute for a vacant land sale is likely to be a “teardown.”

When an existing structure is purchased with a plan to demolish it and build new housing, the value of the underlying land should equal the sale price of the developed parcel less demolition costs. Rosenthal and Helsley (1994) were the first to apply this idea to infer land values from teardown properties in Vancouver, B.C. In subsequent work, Dye and McMillen (2007) and McMillen (2008) refined the econometrics to control for the non-random selection of which parcels are torn down and provided new evidence on land values in Chicago. While teardowns can support a convincing quasi-experimental approach to measuring land value, the active markets are too few and too thin to apply the method broadly across the United States or at a high level of spatial resolution throughout a single metro area.

Since the lack of data makes it difficult to measure the market value of land directly, analysts have sought to estimate it indirectly from hedonic regressions or replacement cost equations. Both strategies begin by rearranging the decomposition in definition 3,

$$P_{jkt} = L V_{jkt} + S V_{jkt}. \quad (2)$$

Given data on the structural characteristics of houses and their transaction prices, equation (2) can be used to estimate land values. In the replacement cost framework, two maintained assumptions are sufficient to guarantee the estimator will be consistent. First, the market for housing is assumed to be sufficiently competitive that the market value of a structure will equal the cost of rebuilding that structure in its current condition:  $replacement\ cost_{jkt} \equiv RC_t(x_{jkt}) = S V_{jkt}$ . Second, the replacement cost function is assumed to be known. Under these assumptions, one can obtain a consistent estimate for land value as the residual obtained by subtracting replace-

ment cost from the price of housing,

$$\overline{LV}_{jkt} = P_{jkt} - RC_t(x_{jkt}). \quad (3)$$

Davis and Heathcoat (2007) applied this logic at the national level to develop the first macroeconomic index of residential land value in the United States. Davis and Palumbo (2008) refined their methodology to control for variation in property values and construction costs across major metropolitan areas. They developed a database describing the value of land in 46 major metropolitan areas between 1984 and present.<sup>10</sup>

During boom-bust cycles, the replacement cost framework tends to attribute most of the changes in property values to speculation on land. This follows from the mechanics of (2)-(3). If residential construction costs are relatively stable during a period when property values are rising rapidly, then observed changes in property values will be interpreted as changes in land value. This was exactly what happened during the recent boom. The replacement cost model indicates that the ratio of land value to property value on the West Coast increased from 61% in 1998 to 74% in 2004, for example (Davis and Palumbo, 2008). We have no doubt that the market value of land did increase during the boom. However, the replacement cost estimates for the magnitude of the change may be too high if housing markets are less than perfectly competitive or if zoning restrictions and permitting requirements drive a wedge between construction costs and effective replacement costs in the short run.

The hedonic approach to estimating land values avoids the need to specify replacement costs or assume that markets are perfectly competitive. Instead, the key maintained assumption is a parametric specification for the relationship between the sale price of a house and its characteristics. Equation (4) presents a linear example reflecting the spatiotemporal structure of past hedonic land value estimators.

$$P_{jkt} = \xi + \delta \cdot l_{jk} + x_{jkt}\beta + \varepsilon_{jkt}. \quad (4)$$

In this case,  $\hat{\delta}$  provides an estimate of the implicit marginal price of land and  $\hat{\delta} \cdot l_{jk}$  provides an

estimate for the property's land value. Efforts to estimate  $\delta$  from data on individual housing sales date back at least to Clapp's (1980) study of land values in Chicago.<sup>11</sup> Over the years, the methodology has been refined to allow more flexible parametric specifications for the hedonic price surface (Cheshire and Sheppard, 1995) and extended to compare estimates across 21 metropolitan areas (Glaeser, Gyourko, and Saks, 2005).

There are two key challenges to developing credible hedonic estimates for land values. The first challenge—omitted variable bias—is widely recognized. For example, one might expect that houses built on larger lots will also tend to be built using higher quality materials. Because data on building materials are typically unavailable, their effect on sale prices will be confounded with the value of land (McMillen, 2008). Another concern is that an estimate for the depreciation of structures (from the coefficient on age) may be confounded with unobserved neighborhood amenities because all of the houses in a subdivision tend to be built at about the same time (Davis and Palumbo, 2008). More generally, there is always likely to be some degree of spatial correlation between observed parcel characteristics and unobserved neighborhood amenities that will ultimately bias the estimator (Kuminoff, Parmeter, and Pope, 2010).

The second challenge is to choose a specification for the price function that is sufficiently flexible to capture the key features of spatial and temporal variation in land values. Past studies have focused on allowing the per unit price of land ( $\delta$ ) to vary flexibly within a metropolitan area (for example, see Cheshire and Sheppard, 1995). While this is an important dimension of heterogeneity, we hypothesize that it is equally important to distinguish between the variable (i.e. quantity-based) and fixed (i.e. access-based) components of land value, while recognizing that marginal implicit prices for both may change over time due to changing market conditions.<sup>12</sup>

Access matters. This is a central theme of public, urban, and environmental economics. Commuters value access to the central business district (CBD). Homeowners value access to local public goods and amenities that contribute to their quality of life. These values are fundamental to the models of urban spatial structure and neighborhood formation that build on the work of Tiebout (1956), Alonso (1964), Mills (1967), and Muth (1969). Within a neighborhood,

the value of access will be approximately fixed, independent of parcel size. As one moves to a different neighborhood with higher crime rates, lower quality schools, and/or a longer commute to the CBD, the value of access may drop sharply. To identify spatial variation in access value separately from spatial variation in the per/unit price of land, the analyst must observe several housing transactions within each neighborhood during an interval over which land values are relatively stable.<sup>13</sup> To identify temporal variation in each of these components, the analyst must observe a large number of observations in each period. Our econometric model is specially designed to accomplish these tasks using data on the universe of housing sales within a metropolitan area together with controls for omitted variables.

### **3.2 Refining the Hedonic Approach to Estimation**

Our approach to estimating land values relies on micro data that are sufficiently rich to allow us to estimate annual price functions for metro areas, while simultaneously using spatial fixed effects to capture the market value of latent attributes of land and structures. In the case of land, the issue is that no existing database provides comprehensive coverage of spatial variation in access-based amenities below the level of a county. This is important because amenities often vary significantly within a county. To measure this variation we use spatial fixed effects for neighborhoods, which we define to be Census tracts.<sup>14</sup> Within a tract, access to amenities will be approximately fixed. Children will be assigned to public schools in the same school district, their parents will face the same commuting opportunities, and there will be little or no variation in crime rates or air quality. Thus, we would expect tract fixed effects to absorb the composite value of access to these and other neighborhood amenities.

In the case of structures, micro-level data are typically limited to the attributes recorded by the county assessor. Some houses have hardwood floors, granite countertops, skylights, solar panels, and spas. Unfortunately, these improvements are rarely noted in the county records. If the quality of building materials varies systematically across neighborhoods in ways we do not observe, then their average effect on property values may be confounded with our estimates for

the fixed component of neighborhood land value. To mitigate this potential source of confounding, we add a set of interactions between the fixed effects for neighborhoods and the square footage of the house. The resulting terms are intended to capture systematic variation across neighborhoods in the average value of a square foot of structural improvements.

We adopt a semi-log form for the estimation, regressing the log of transaction prices for all of the single-family residential properties sold in a metro area during year  $t$  on their lot sizes, their structural characteristics, and two sets of fixed effects,

$$\ln(P_{jkt}) = \underbrace{\xi_{kt} + \delta_{kt}l_{jkt}}_{LV_{jkt}} + \underbrace{\gamma_{kt}sqft_{jkt} + x_{jkt}\beta_t}_{SV_{jkt}} + \varepsilon_{jkt}. \quad (5)$$

The first two terms after the equality correspond to the property's land value.  $\xi_{kt}$  denotes the neighborhood fixed effects. They will measure the component of land value that is constant across all the houses sold within tract  $k$  during year  $t$ , regardless of lot size. The neighborhood amenities that enter  $\xi_{kt}$  may also interact with the size of the lot to influence the variable component of land value. For example, the marginal value of yard size may be larger in quieter neighborhoods with lower crime rates. Therefore, we allow the coefficient on lot size,  $\delta_{kt}$ , to vary over neighborhoods as well.

The third and fourth terms after the equality correspond to the value of structural improvements.  $x_{jkt}\beta_t$  measures the component of property value that can be explained by the housing characteristics that are observed. While we allow the implicit prices of characteristics to change over time, we restrict them to be fixed within a metropolitan area during the course of a year.  $\gamma_{kt}sqft_{jkt}$  measures systematic variation in the average value of a square foot of living space that varies across neighborhoods due to unobserved structural improvements.

Finally, we interpret the error term  $\varepsilon_{jkt}$  as the composite of three effects. It will reflect: (i) unobserved idiosyncratic structural improvements that differ from the tract average; (ii) idio-

syncratic access to amenities within a neighborhood;<sup>15</sup> and (iii) misspecification in the shape of the price function. To mitigate the first two effects, we aggregate our micro-level estimates for the value of land and structures to report averages for Census tracts, counties, and metropolitan areas. This also allows us to compare our results to estimates from the prior literature. While our resulting point estimates surely contain some error, and their standard errors may be affected by spatial autocorrelation, we expect the magnitude of the bias in our point estimates to be smaller than in previous hedonic studies because of the ways in which our model enhances spatial and temporal resolution and controls for omitted variables.<sup>16</sup> To evaluate the impact of these refinements, we use the fact that our model nests the conventional hedonic specification as a special case. Equation (5) reduces to (4) if we omit spatial fixed effects ( $\gamma_{kt} = \xi_{kt} = 0$ ) and restrict the implicit price per acre of land to be constant within a metropolitan area ( $\delta_{kt} = \delta$ ).

#### **4. Data and Summary Statistics**

Our analysis is based on more than one million observations on the sales of single-family residential properties across the United States. We purchased the data from DataQuick. This widely used commercial vendor of real estate data assembled the data from assessor's offices in individual towns and counties. The data include the transaction price of each house, the sale date, and a consistent set of structural characteristics, including square feet of living area, number of bathrooms, number of bedrooms, year built, and lot size. Using these characteristics, we performed some standard cleaning of the data, removing outlying observations, removing houses built prior to 1900, and removing houses built on lots larger than 5 acres.

The data also include the physical address of each house, which we translated into latitude and longitude coordinates using GIS street maps and a geocoding routine. The lat-long coordinates were then used to assign each house to its corresponding census tract. The tract-level assignment provides the needed spatial resolution to analyze trends in land values within and across metro areas during the boom-bust cycle. Furthermore, it allows us to use spatial fixed



effects to control for the average effect of latent variables within each tract.

While we conducted the econometric analysis for ten metro areas, we focus on four of them in greater detail in order to illustrate our main results: Miami, FL; San Francisco, CA; Boston, MA; and Charlotte, NC.<sup>17</sup> We selected these four because each has complete data between 1998 and 2008, they provide geographic variation on populous areas in the United States, they provide variation in the supply elasticity of land, and they differ in the intensity of their boom-bust cycles. Figure 2 illustrates the differences in the sizes of their booms and busts using the Case-Shiller Home Price Index. Each panel also reports Saiz's (2010) estimates for the supply elasticity of housing.<sup>18</sup>

Table 1 provides summary statistics for the housing transactions that we observe in Miami, San Francisco, Boston, and Charlotte. The first two rows of each panel illustrate that the average sale price rose in all four areas between 1998 and 2006. The size of the increase was most striking in Miami (\$162k to \$410k) and San Francisco (\$343k to \$809k) where prices more than doubled in nominal terms. These increases do not reflect any obvious changes in the composition of houses on the market. The structural characteristics of the average sale property are essentially constant over the study period. In each area, the median transaction was a single-family house with 3 bedrooms, 2 baths, and between 1600 and 1900 square feet of living area. Naturally, Charlotte and Miami have newer housing stocks than San Francisco and Boston. Lot sizes also tended to be larger in Charlotte and Boston than in Miami and San Francisco, reflecting variation in the balance between sales from the cities and suburbs.<sup>19</sup>

## **5. Results**

### **5.1 Comparison to Pre-Boom Estimates from the Existing Literature (1998-1999)**

We begin by comparing our estimates for land values to previous figures generated by the conventional hedonic estimator in Glaeser, Gyourko and Saks (2005) [henceforth GGS] and the replacement cost estimator developed by Davis and Palumbo (2008) [henceforth DP]. Neither study had the benefit of our spatially delineated micro data on actual housing sales. Instead, they

combined data from the American Housing Survey with other sources to generate estimates for average land values within several metropolitan areas. Fortunately, some of their estimates overlap with the spatial dimensions of our data prior to the onset of the housing market boom, providing an opportunity for comparison. The purpose of the comparison is to investigate how our refinements to the hedonic land value estimator influence the accuracy of our results. We would expect DP's replacement cost calculations to generate reasonable estimates for land values prior to the boom. The relative stability of the market in the mid to late 1990s would have allowed developers considerable time to meet demand, mitigating any wedge between construction costs and effective replacement costs. Thus, we view DP's estimates as the most reliable baseline for comparison.

The task of estimating land values is a relatively small component of the overall analysis in GGS. Their main objective is to test the hypothesis that land use regulations impose an effective tax that explains the rise in housing prices in major metropolitan areas. To illustrate their point and to compare housing prices to construction costs, GGS estimate the "free-market cost of land" using a conventional hedonic model (similar to equation 4 above) for 21 metro areas based on data from the 1998 and 1999 installments of the American Housing Survey (AHS).<sup>20</sup> We have the requisite information to develop comparable estimates for 10 of their 21 metro areas. Conveniently, DP also report estimates for all 10 areas.

To provide the best possible comparison, we focus on the subset of our data that overlap with the information used by GGS. Specifically, we limit our data to the year that matches the year in which each metro area was covered by the AHS (either 1998 or 1999). Then we subdivide metropolitan areas to match the disaggregate definitions used in the AHS. This means subdividing the San Francisco Consolidated metropolitan statistical area into the San Francisco, Oakland, and San Jose primary metro areas, for example. While our micro data still differ from the AHS in terms of the number of observations and the richness of information on structural characteristics, their spatial and temporal dimensions are the same.

The first column in Table 2 simply reproduces the estimates of land value (on a per-acre

basis) from table 4 of GGS. In column [2], we report the results from our attempt to come as close as possible to replicating their estimating equation, given the differences between the variables in our data and the AHS micro data.<sup>21</sup> A quick comparison between columns [1] and [2] confirms that the two sets of estimates are quite similar (with Tampa as the exception). Overall, the estimates line up with our general intuition for which metro areas ought to have more expensive land. San Francisco, San Jose, Oakland, and Los Angeles have the highest measures of land value whereas Detroit and Tampa have the lowest. However, all of the estimates seem implausibly low for the late 1990s. Could you really buy an acre of land in San Francisco for under \$200,000 or in Boston for under \$30,000? A likely explanation is that the conventional hedonic estimator does not fully capture the fixed component of land value associated with access to the local public goods and amenities ( $\xi_{kt}$ ).

Column [3] reports the corresponding replacement cost estimates for land value from DP. They used information published by R.S. Means Company (2004) to develop metro-level estimates for replacement cost. Their measures for housing prices were developed by combining data on price levels in each metro area during AHS survey years with time-series data on the percentage change in housing prices from Freddie Mac's Conventional Mortgage Housing Price Index (CMHPI). The rank order in column [3] is similar to the first two columns, but the replacement cost estimates are typically an order of magnitude larger! While there are some slight variations in the data sets used to develop the estimates in columns [1] and [3], none seem capable of generating order of magnitude differences.<sup>22</sup> It seems more likely that the differences are due to estimation procedures. In particular, the access-based component of land value associated with local public goods would be included in the replacement cost estimates and excluded from the estimates generated by conventional hedonic regressions.

Column [4] reports the estimates from our refinement to the hedonic estimator, using the specification in equation (5).<sup>23</sup> Generalizing the conventional hedonic model to allow for access-based amenities and latent housing characteristics increases our estimates by an order of magni-

tude (moving from column [2] to column [4]). The resulting estimates align much more closely with the estimates from DP's replacement-cost model.

Finally, it is important to reiterate that the similarity between our estimates and DP's occurs during a two-year period prior to the boom. As we track the two sets of estimates over the course of the boom-bust cycle, we see some interesting differences.

## **5.2 The Evolution of Average Land Values during the Boom-Bust Cycle (1998-2009)**

We estimated equation (5) for each (metro area, year) combination from 1998 and 2009. Table 3 summarizes results for the four metro areas where we have a complete set of data: Miami, San Francisco, Boston, and Charlotte. It reports our measures for the evolution of land values and the share of property value attributed to land ("land share"), alongside the replacement cost estimates from DP.<sup>24</sup> The hedonic measures were generated by averaging our parcel-specific estimates for land values and improved values over all of the housing transactions in each metro area. There are some obvious differences between the two sets of estimates at the market's peak.

Figure 3 illustrates the differences graphically. Focusing on the first column in the figure, it is clear that land values estimated by both methods rise and fall during the boom and bust. Prior to the boom, the two sets of estimates are similar. The same is true following the bust. However, the peak amplitude is much larger in the replacement cost estimates.

The second column of Figure 3 illustrates how estimates for the value of structures evolved over the same period. The hedonic model suggests that the market value of structures rose and fell in tandem with the market value of land during the boom-bust cycle. The replacement cost measures rose steadily, following a similar trend in every metro area. Once again, the differences between the hedonic estimates and the replacement cost measures are largest at the height of the boom.

Does the difference between the two sets of estimates reveal something interesting about the behavior of housing markets during the boom-bust cycle? Or does it merely reflect differences in the underlying data? While our comparisons were made along a consistent set of spa-

tial and temporal dimensions, the underlying micro data are not the same. DP's replacement cost estimates are based on integrating the AHS data with Freddie Mac's Conventional Mortgage Housing Price Index (CMHPI), whereas our hedonic estimates come from assessor data. In principle, the assessor data describe the universe of housing transactions, whereas the CMHPI is limited to transactions where: (a) the transaction was a repeated sale; and (b) the buyer took out a conventional mortgage that was purchased or insured by Freddie Mac or Fannie May.

Figure 4 compares the evolution of average property values in the two datasets. The assessor data suggest slightly smaller increases in property values during the boom. One possible explanation is that less expensive transactions were more likely to be associated with unconventional mortgages. Another explanation is that new houses built over this period tended to be located near the urban fringe where land values (and property values) were lower. In any case, the differences between the two measures of average property value in Figure 4 are dwarfed by the differences in estimated land values in Figure 3. Thus, the differences between the hedonic and replacement cost estimates for land value appear to be tied to methodology, not the underlying data.

Data differences aside, the main economic implication of our comparison between the hedonic and replacement cost estimates is that, during the boom, the market value of structures may have exceeded their replacement costs. To further investigate this possibility, we examine the spatial variation in the evolution of land values within each metro area.

### **5.3 Within-Metro Heterogeneity in the Evolution of Land Values (1998-2009)**

Figure 5 illustrates the spatial heterogeneity in land values across counties in the greater San Francisco and Boston metropolitan areas.<sup>25</sup> The left-most maps display the land value of the average residential property sold in 1998, the change in average land value during the boom (1998-2006), and the change in average land value during the bust (2006-2009). In the San Francisco metro area the counties with the highest land values in 1998 are San Francisco and San Mateo followed by Marin and Santa Clara. These same counties experienced the largest increas-

es in land value during the boom and the smallest decreases during the bust. Looking at the left-most maps in Figure 5 for the Boston metro area reveals a similar pattern.

The right-most maps in Figure 5 focus on the ratio of land value to total property value. The three maps display the average land share in 1998 and the subsequent changes during the boom and bust periods. Focusing first on the greater San Francisco metro area, we see that areas with higher land shares in 1998 (e.g. San Francisco and San Mateo) tended to see drops in land share during the boom and increases during the bust. Again, a similar pattern emerges in the Boston metropolitan area.

Overall, the spatial heterogeneity in the evolution of land values within the Boston and San Francisco areas seems somewhat counterintuitive. The counties that experienced the least volatility in land values during the boom-bust cycle are the same counties that we would expect to have the most inelastic supply of housing. This pattern is the opposite of what the prior literature has observed about the variation in land values between metro areas (e.g. Davis and Palumbo 2008). Figure 2 provides an example of the stylized fact that housing prices (and land values) tend to be more volatile in metro areas where the supply of land is relatively inelastic.<sup>26</sup> Why would the volatility of land values be so different within a metro area?

To further investigate the relationship between land value and housing supply, Table 4 summarizes trends in land values and permits issued for the construction of new housing units in the San Francisco Bay Area. Column [1] reports the baseline number of owner occupied housing units by county from the 2000 Census and column [2] reports the number of new permits for construction of single-family residential (SRF) housing units. The counties are ranked by column [3], which reports the ratio of column [2] to column [1]. The ratios are smallest for San Francisco and its adjacent coastal counties (Marin and San Mateo). The same is true if we look at the ratio of all new permits to all housing units in column [4]. This ratio is much higher for San Francisco because it includes permits to build apartment units. It also includes all housing units in the denominator, regardless of occupancy status. In the absence of county-level estimates for the supply elasticity of housing, columns [3]-[4] provide a crude proxy for the respon-

siveness of housing supply during the boom.

Comparing the ratios in columns [3]-[4] with the values of land and property in columns [5]-[12] highlights five interesting trends.<sup>27</sup> First, at the start of the boom period, property values and land values were higher in counties where the supply of housing was less responsive. This is true whether we look at the median self-reported property values in column [5], the mean of actual transaction prices in column [6], or our estimates for mean land values in column [7]. Second, land tends to represent a smaller share of total property value in counties where the housing supply is more responsive (comparing columns [6] and [7]). Third, while the counties with the least responsive housing supply experienced the largest nominal increases in land values during the boom (column [8]), these increases were relatively small in percentage terms (column [9]). Fourth, during the bust, the counties with the least responsive housing supply experienced the smallest decreases in land values in both nominal and percentage terms (columns [10]-[11]). Finally, and perhaps most strikingly, the counties with the least responsive housing supply had large net gains in land value between 1998 and 2009 whereas the fastest growing counties (Contra Costa, Napa, and Solano) lost most of the land value that had accumulated during the boom. Overall, these trends support our initial hypothesis that the Bay Area counties with the most volatile property values and land values during the boom-bust cycle also had the most elastic housing supplies.

Finally, we report an intriguing pattern in our estimates for the ratio of land value to total property value. We further disaggregate our results to the level of a Census tract and regress the change in the land value share of each tract between 1998 and 2006 on its baseline land value share in 1998. We find that census tracts with high initial land shares in 1998 tended to see smaller increases in land shares during the boom. Table 5 summarizes the results, by metro area. For example, the coefficient for the San Francisco metro area indicates that a 1 percentage point increase in a census tract's 1998 land value share was associated with a 0.574 percentage point decrease in the size of the change in the tract's land share between 1998 and 2006. The net effect is an increase (decrease) in the land share for tracts with initial land shares below (above)

62.5%.<sup>28</sup> This negative correlation holds for all of the metro areas and ranges from -0.393 in Pittsburg to -0.754 in San Jose. Furthermore, the R-squared values between 0.24 and 0.59 suggest that the initial land share in 1998 explains much of the variation in land shares during the boom.

One explanation for this pattern is that areas with high land shares in 1998 (presumably high amenity areas) saw both land values and structural values increase during the boom, but structural values went up relatively more due to markups arising from spatial market power associated with the inelastic supply of access to amenities. Another explanation is that areas with low land shares in 1998 (presumably low amenity areas) saw large increases in land values relative to structural values because (a) the relatively elastic supply of land in low amenity areas kept the implicit price of structures pinned to construction costs; and/or (b) the general relaxation of credit constraints during the boom had the largest impact on demand in these areas. Without imposing additional structure on the data, we cannot disentangle the relative importance of these explanations. We discuss them briefly in the hope of motivating future research.

## 6. Discussion

Conventional wisdom suggests that variation in the volatility of housing prices *across* metro areas is primarily due to heterogeneity in the supply elasticity of land. In areas with physical and legislated constraints to urban development, the market price of housing will be relatively sensitive to demand shocks fueled by speculation and relaxation of credit constraints. These demand shocks will be translated into higher land values. Our results do not contradict this hypothesis. However, the land supply hypothesis is not sufficient to explain the variation we observe in the amplitude of the boom and bust *within* metro areas.

Within the metro areas that we studied, housing prices were relatively volatile in neighborhoods at the urban fringe, where the supply of land for housing is relatively elastic. Decomposing this price volatility into the market value of land and structures revealed two other interesting trends. First, we saw that the average market value of land and structures tended to



rise and fall in tandem. Since construction costs rose steadily during 1998-2009, our results suggest the presence of a wedge between construction costs and the market value of structures. Second, we saw that the size of the wedge tended to be larger in neighborhoods with higher land shares prior to the boom.

One possible explanation for the relatively high volatility at the urban fringe is that the relaxation of credit constraints was particularly important for lower income households, allowing them to purchase houses at the fringe of the suburbs. Recent work by Landvoigt, Piazzesi, and Schneider (2011) supports this hypothesis.<sup>29</sup> Using a calibrated assignment model, they suggest that cheaper credit for low income households effectively drove up housing prices for the “low end of the market” in San Diego during the boom. Likewise, Mian and Sufi (2011) use micro data on homeowners to document that during the boom there was positive correlation between borrowing and house price appreciation, and Mian and Sufi (2009) use zip-code level data to document that during the boom there was negative correlation between income growth and subprime mortgage credit. However, these correlations do not offer a specific explanation for the wedge we observe between construction costs and the market value of structures. In the remainder of this section, we discuss two market forces that may help to explain the wedge: imperfect competition and q-theory. Both present interesting opportunities for further research.

## **6.1 Imperfect Competition**

If markets were perfectly competitive with no barriers to entry, we would expect the market value of structures to be pinned to construction costs. Our results indicate this is not the case. One explanation is that barriers to entry convey some degree of market power to homeowners in exclusive neighborhoods. The construction industry may be close to perfectly competitive. However, builders cannot simply build more houses in established neighborhoods. Furthermore, development restrictions and zoning regulations often limit the ability of homeowners to expand their houses. With a small proportion of houses on the market at any one time, the unique bundle of amenities provided by a desirable neighborhood may allow home sellers to charge a markup

on the structural characteristics of their houses. If there are only a few large houses on the market in the best school district, for example, the implicit market value of square footage may be bid far above construction costs due, in part, to the demand for access to high quality schools. This hypothesis is consistent with our observation that the neighborhoods experiencing the largest increase in market values of structural characteristics were the neighborhoods with the largest pre-boom land values (presumably the highest amenity neighborhoods).

To illustrate the comparative statics of the market power hypothesis, we use Kuminoff and Jarrah's (2010) iterative bidding algorithm (IBA) to simulate hedonic equilibria with heterogeneous households and houses. The IBA uses a numerical procedure to solve for an assignment of people to houses and a vector of prices that jointly support a hedonic equilibrium, given an initial stock of housing and a set of draws from the joint distribution of income and preferences.

We use the IBA to simulate market outcomes in a stylized metropolitan area containing two built-up neighborhoods, A and B. Each neighborhood is defined to have 100 lots of identical size ( $l_i = 7000$  sqft). In neighborhood A, half of the lots contain "small" houses, uniformly drawn from  $x_i \sim [1000, 2000]$  sqft. The remaining lots contain "large" houses, uniformly drawn from  $x_i \sim [2000, 3000]$  sqft. In neighborhood B, the large and small houses are drawn from the same uniform distributions, but only 20% of the houses are large. The only other difference between the two neighborhoods is that B has more desirable amenities:  $g_B = 75 > 50 = g_A$ . Utility is specified as a Cobb-Douglas function of housing and neighborhood attributes:

$$V_{ijk} = \ln(y_i - p_{jk}) + \alpha_{i1} \ln l_{jk} + \alpha_{i2} \ln x_{jk} + \alpha_{i3} \ln g_{jk} + \alpha_{i4} \ln x_{jk} g_{jk}.$$

Finally, the joint distribution of preferences is drawn from a gamma distribution, and income is drawn from the same empirical distribution used by Kuminoff and Jarrah (2010).<sup>30</sup>

In the resulting equilibrium, large houses and small houses are both more expensive in neighborhood B, because it provides access to higher quality amenities. Our main point is that the large houses in B also command larger price premiums because they are in limited supply.

To isolate the price premium, we begin by calculating the difference between the average price of large and small houses in each neighborhood. For example,  $(\bar{p}_j | j \in A \ \& \ 2000 < x_j < 3000) - (\bar{p}_j | j \in A \ \& \ 1000 < x_j < 2000)$  measures the difference between the average prices of large and small houses in neighborhood A. Differencing removes the market value of land. This follows because all houses in A have identical lots and they provide access to same amenity. The same is true for all houses in B. Therefore, the percentage markup on square footage in neighborhood B can be defined as,

$$100 * \left[ \frac{(\bar{p}_j | j \in B \ \& \ 2000 < x_j < 3000) - (\bar{p}_j | j \in B \ \& \ 1000 < x_j < 2000)}{(\bar{p}_j | j \in A \ \& \ 2000 < x_j < 3000) - (\bar{p}_j | j \in A \ \& \ 1000 < x_j < 2000)} - 1 \right]. \quad (6)$$

Figure 6A graphs the relationship between the markup and the difference in amenities provided by the two neighborhoods. In the baseline equilibrium (i.e.  $g_B - g_A = 75 - 50 = 25$ ) there is a 63% premium on the market value of structures in the high amenity neighborhood. As  $(g_A - g_B) \rightarrow 0$ , spatial market power diminishes and the equilibrium markup approaches zero.

It is important to reiterate that our hedonic framework in this paper is consistent with the possibility of market power. While Rosen's (1974) welfare interpretation of the hedonic gradient relies on the maintained assumption of perfect competition, we did not maintain that assumption in order to prove that market outcomes can be described by a hedonic price function in theorem 1. As Feenstra (1995) demonstrated, introducing imperfect competition into a hedonic equilibrium simply changes the interpretation of the price function coefficients. They describe the implicit market prices of product characteristics, which reflect marginal costs plus markups or discounts.

Taylor and Smith (2000) provided the first hedonic evidence of market power in the market for beach rental properties in North Carolina. In particular, they found access to the beach gave owners the ability to charge markups on structural features of the house that were difficult to modify, such as the number of bedrooms. In our model, markups can enter through variation

in the tract-specific coefficients on square feet of living space. It would be interesting to investigate the extent to which this variation can be explained by spatial variation in the quality of local public goods and urban amenities, perhaps using a higher-resolution version of the quality-of-life indices that have been constructed at the county level (e.g. Blomquist, Berger, and Hoehn 1988). One could also consider generalizing our model to allow for more spatial variation in the implicit prices (and markups) for other structural characteristics.

## 6.2 Q-theory

A second explanation for the wedge between construction costs and the market value of structures is that new houses take time to build. In a long-run equilibrium we would expect the market value of an additional unit of each structural characteristic to equal its marginal construction cost. In the short run, however, a positive demand shock may lead market values to exceed construction costs. As the ratio of market value to construction cost increases, so does the incentive for new development. This is the basic idea behind Tobin's "q-theory" of capital investment. Figure 6B illustrates his logic in the context of our hedonic simulation. It graphs the difference ratio that enters the markup formula in (6) against the share of "large" houses in the high amenity neighborhood. Holding the amenity differential fixed, we simulate the transition path to a long-run equilibrium by incrementally "remodeling" the small houses in the high amenity neighborhood as large houses. Each time we remodel an additional 10% of houses, we solve for a new set of equilibrium prices and location choices. While the total number of houses is constant throughout this exercise, there is an increase in the total square footage of living space. As living space increases, the market-clearing difference ratio moves closer to 1, consistent with Tobin's description of the transition to a long-run equilibrium.<sup>31</sup>

While Tobin (1978) noted the potential for his model to explain the evolution of housing prices, there have been few applications. Part of the difficulty is that the logic of q-theory applies to structures, not to land, since we usually think of the supply of land as being fixed.<sup>32</sup> Thus, to evaluate the testable implications of q-theory we need credible estimates for the market

value of structures.<sup>33</sup> Given credible estimates for the value of structures and construction costs, the interpretation of q-ratios is still complicated by several factors, including the presence of taxes, search costs, lender fees, and the aggregation over individual housing characteristics, each of which would be expected to have a separate marginal q.<sup>34</sup> These factors may lead to q-ratios that differ from 1 in a long-run equilibrium.

As a first pass at assessing the predictions of q-theory in the housing market, we approximate q by dividing our annual estimates for the market value of structures by the replacement cost estimates for the cost of rebuilding the average home in Miami, San Francisco, Boston, and Charlotte. The results appear to be broadly consistent with the implications of q-theory.<sup>35</sup> Figure 7 illustrates that our pre-boom and post-bust estimates for structural value would imply q-ratios close to 1 in some areas. During the boom we generally see q-ratios increase and then decrease, consistent with lagged construction. In Miami, for example, our estimate for the q-ratio increases from 0.92 in 1998 to 1.5 in 2007, and then decreases back to approximately 0.92 by 2009. The decrease in the q-ratio is likely due to decreased demand as well as increased supply. Boston and Charlotte show similar patterns although their q-ratio is more elevated at the beginning of the time frame. San Francisco however begins with a q ratio of 1.68, peaks at over 3 in 2005 and then drops back to 1.69 in 2009. Modeling the dynamics of housing supply and demand as they relate to q-theory would be an interesting direction for future research.

## **7. Conclusion**

The boom-bust cycle in real estate during the 2000's was staggering in its size and impact on the world economy. It is extremely important to characterize the dimensions of this volatility and to understand why they arise. Hedonic pricing models provide a useful conceptual framework for this task, but recent studies have questioned their ability to deal with unobserved attributes of houses and neighborhoods. We answered this question by refining the standard hedonic pricing model to address omitted variables and by using our refined estimator to characterize how the

market values of land and structures evolved within and between several major metropolitan areas between 1998 and 2009.

Concerns about omitted variables led us to make three refinements to the conventional hedonic strategy for inferring the market value of residential land. We used spatial fixed effects to recognize that land values may be highly localized due to the spatial distribution of nonmarket amenities; we controlled for unobserved housing attributes by allowing the implicit per/unit prices of structures to vary from neighborhood to neighborhood; and we exploited data on the universe of housing sales to recognize that shocks to wealth, credit, tastes, and amenities may cause the shape of the housing price function to change from year to year. These refinements increased our estimates by an order of magnitude. Our estimates for the average market values of land in major metropolitan areas were very similar to the pre-boom estimates from Davis and Palumbo's (2008) replacement cost model, but an order of magnitude larger than estimates from Glaeser, Gyourko & Saks's (2005) conventional hedonic model.

Consistent with prior studies of the boom-bust period, we found that land values were generally more volatile in metropolitan areas where the supply of developable land was less elastic. Focusing on the spatiotemporal variation within metro areas yielded two novel findings. The least valuable land at the urban fringe was the most volatile, and the market value of structures exceeded construction costs during the boom, with the largest markups occurring in the most affluent neighborhoods. Interestingly, the volatility in the market value of structures was far greater than what has been assumed by prior replacement cost studies that pin the value of structures to construction costs. Finally, we suggested two potential explanations for the wedge between construction costs and the market value of structures: imperfect competition and  $q$ -theory. Formal tests of these hypotheses await future research.

Finally, our results also have some implications for the related literature on land value taxation. Over the years there has been considerable interest in the possible efficiency gains from replacing the property tax with a tax on the market value of land or a split tax with separate rates on land and structures (e.g. Banzhaf and Lavery 2008; Cho, Lambert and Roberts 2010).

One of the stylized facts about land value taxation is that, if implemented, it would lead to more variable revenue streams than the current property tax because land values are more susceptible to speculation (Bourassa 2009). At a practical level, part of the challenge with implementing a tax on land is determining its market value. Our findings have three implications for this literature. First, the replacement cost approach may overstate the value of land during a boom-bust cycle. Second, the bias may not be neutral. Our results suggest it would be largest in the highest-amenity neighborhoods. To the extent that homeowners in these neighborhoods collect markups on structures, they would have a disincentive to invest in structural improvements if they were effectively taxed on these improvements by a replacement cost scheme for determining land value. Finally, our estimates suggest that moving from a property tax to a land tax may actually help to stabilize revenue streams for some municipalities.

## References

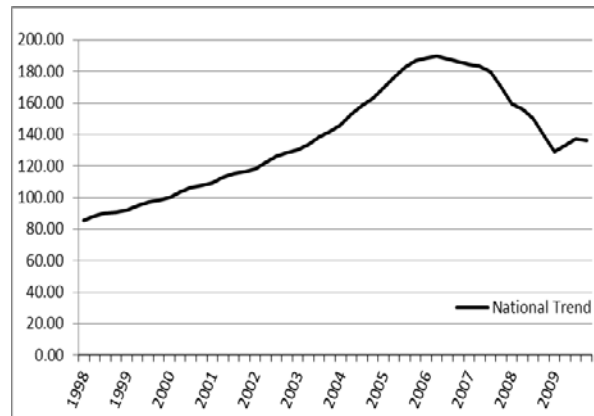
- Abbott, Joshua H. and H. Allen Klaiber. 2010. "An Embarrassment of Riches: Confronting Omitted Variable Bias and Multi-Scale Capitalization in Hedonic Price Models." *Review of Economics and Statistics*, forthcoming.
- Alonso, William. 1964. *Location and Land Use: Towards a General Theory of Land Rent*. Cambridge: Cambridge University Press.
- Bajari, Patrick and C. Lanier Benkard. 2005. "Demand Estimation with Heterogeneous Consumers and Unobserved Product Characteristics: A Hedonic Approach." *Journal of Political Economy*, 113(6): 1239-76.
- Banzhaf, H. Spencer, and Nathan Lavery. 2010. "Can the Land Tax Help Curb Urban Sprawl? Evidence from Growth Patterns in Pennsylvania." *Journal of Urban Economics*. 67: 169-179.
- Bell, Michael E., John H. Bowman, and Jerome C. German. 2009. "Assessment Requirements for a Separate Tax on Land." In *Land Value Taxation*, ed. Richard F. Dye and Richard W. England. Cambridge, MA: Lincoln Institute of Land Policy.
- Black, Sandra E. 1999. "Do Better Schools Matter? Parental Valuation of Elementary Education." *Quarterly Journal of Economics*, 114(2): 577-99.
- Blomquist, Glenn C., Mark C. Berger, and John P. Hoehn. 1988. "New Estimates of Quality of Life in Urban Areas." *American Economic Review*, 78(1): 89-107.
- Bourassa, Steven C. 2009. "The Political Economy of Land Value Taxation," In *Land Value Taxation*, ed. Richard F. Dye and Richard W. England. Cambridge, MA: Lincoln Institute of Land Policy.
- Chapman, Jeffrey I., Robert J. Johnston, and Timothy J. Tyrrell. 2009. "Implications of a Land Value Tax with Error in Assessed Values." *Land Economics*, 85(4): 576-586.
- Cheshire, Paul, and Stephen Sheppard. 1995. "On the Price of Land and the Value of Amenities." *Economica*. 62: 247-267.
- Cho, Seong-Hoon, Dayton M. Lambert, and Roland K. Roberts. 2010. "Forecasting Open Space with a Two-Rate Property Tax." *Land Economics*. 86(2): 263-280.
- Clapp, John M. 1980. "The Elasticity of Substitution for Land: The Effects of Measurement Errors." *Journal of Urban Economics*. 8 (2): 255-263.
- Davis, Morris. A. and Michael. G. Palumbo. 2008. "The Price of Residential Land in Large U.S. Cities," *Journal of Urban Economics*, 63(1), 352-384.



- Davis, Morris. A. and Jonathan Heathcote. 2007. "The Price and Quantity of Residential Land in the United States," *Journal of Monetary Economics*, 54(8), 2595-2620.
- Dye, Richard F. and Daniel P. McMillen. 2007. "Teardowns and Land Values in the Chicago Metropolitan Area." *Journal of Urban Economics*, 61 (1), 45-63.
- Feenstra, Robert C. 1995. "Exact Hedonic Price Indexes." *The Review of Economics and Statistics*. 77 (4): 634-653.
- Glaeser, Edward L., Joseph Gyourko, and Raven Saks. 2005. "Why is Manhattan So Expensive? Regulation and the Rise in Housing Prices." *Journal of Law and Economics*, 48 (2): 331-369.
- Jackson, Jerry R. 1979. "Intraurban Variation in the Price of Housing." *Journal of Urban Economics*, 6(4): 464-479.
- Jud, G. Donald and Daniel T. Winkler. 2003. "The Q Theory of Housing Investment." *Journal of Real Estate Finance and Economics*, 27(3): 379-392.
- Kuminoff, Nicolai V. and Abdul S. Jarrah. 2010. "A New Approach to Computing Hedonic Equilibria and Investigating the Properties of Locational Sorting Models." *Journal of Urban Economics*, 67(3): 322-335.
- Kuminoff, Nicolai V., Christopher F. Parmeter, and Jaren C. Pope. 2010. "Which Hedonic Models Can We Trust to Estimate the Marginal Willingness to Pay for Environmental Amenities?" *Journal of Environmental Economics and Management*, 60 (3): 145-160.
- Kuminoff, Nicolai V., V. Kerry Smith, and Christopher Timmins. 2010. "The New Economics of Equilibrium Sorting and Its Transformational Role for Policy Evaluation." *NBER Working Paper #16349*.
- Landvoigt, Tim, Monika Piazzesi, and Martin Schneider. 2011. "Housing Market(s) in San Diego." *Mimeo*.
- McKinnish, Terra, Randall Walsh, and T. Kirk White. 2010. "Who Gentrifies Low-Income Neighborhoods?" *Journal of Urban Economics*, 67 (2): 180-193.
- McMillen, Daniel P. 2008. "Teardowns and Hedonic Land Value Function Estimation using Non-Sample Information." *Mimeo*.
- Mian, Atif, and Amir Sufi. 2009. "The Consequences of Mortgage Credit Expansion: Evidence from the U.S. Mortgage Default Crisis." *Quarterly Journal of Economics*, 124(4): 1449-1496.

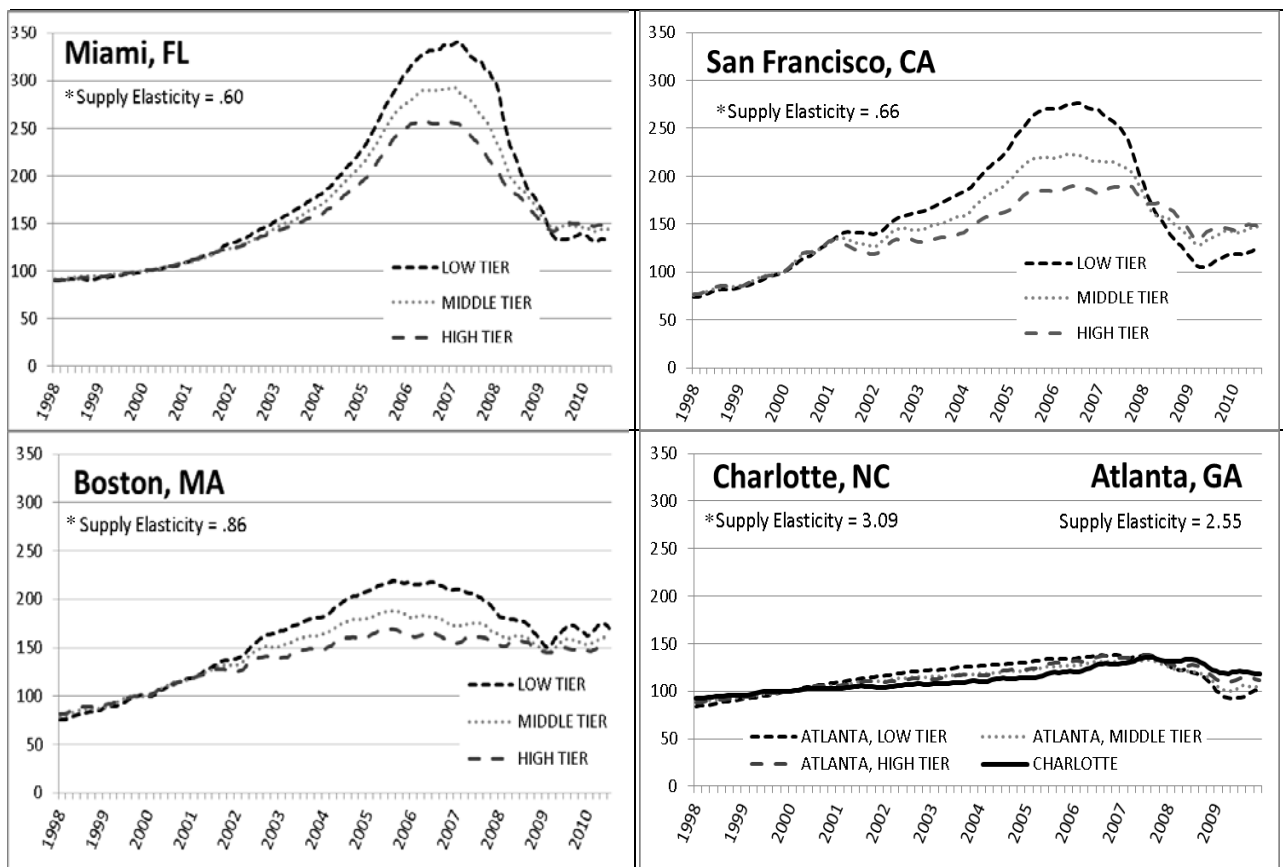
- Mian, Atif, and Amir Sufi. 2011. "House Prices, Home Equity-Based Borrowing, and the U.S. Household Leverage Crisis." *American Economic Review*, 101(5): 2132-2156.
- Mills, Edwin S. 1967. "An Aggregative Model of Resource Allocation in a Metropolitan Area." *American Economic Review*, 57(2):197-210
- Muth, Richard F. 1969. *Cities and Housing*. University of Chicago Press, Chicago.
- Rosen, Sherwin. 1974. "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition." *Journal of Political Economy*, 82(1): 34-55.
- R.S. Means. 2004. "Square Foot Costs." 25<sup>th</sup> Edition, Reed Construction Data, Kingston, MA.
- Rosenthal, Stuart S. and Robert W. Helsley. 1994. "Redevelopment and the Urban Land Price Gradient". *Journal of Urban Economics*, 35(2): 182-200.
- Saiz, Albert. 2010. "The Geographic Determinants of Housing Supply." *Quarterly Journal of Economics*, forthcoming.
- Taylor, Laura O. and V. Kerry Smith. 2000. "Environmental Amenities as a Source of Market Power." *Land Economics*, 76(4): 550-568.
- Tiebout, Charles. 1956. "A Pure Theory of Local Expenditures." *Journal of Political Economy*, 64(5): 416-24.
- Tobin, James. 1978. "Monetary Policies and the Economy: the Transmission Mechanism." *Southern Economic Journal*, 44(3): 421-431.
- Wildasin, David E. 1984. "The q Theory of Investment with Many Capital Goods." *American Economic Review*, 74(1): 203-210.

**Figure 1: Standard & Poor’s Case-Shiller National Housing Price Index**



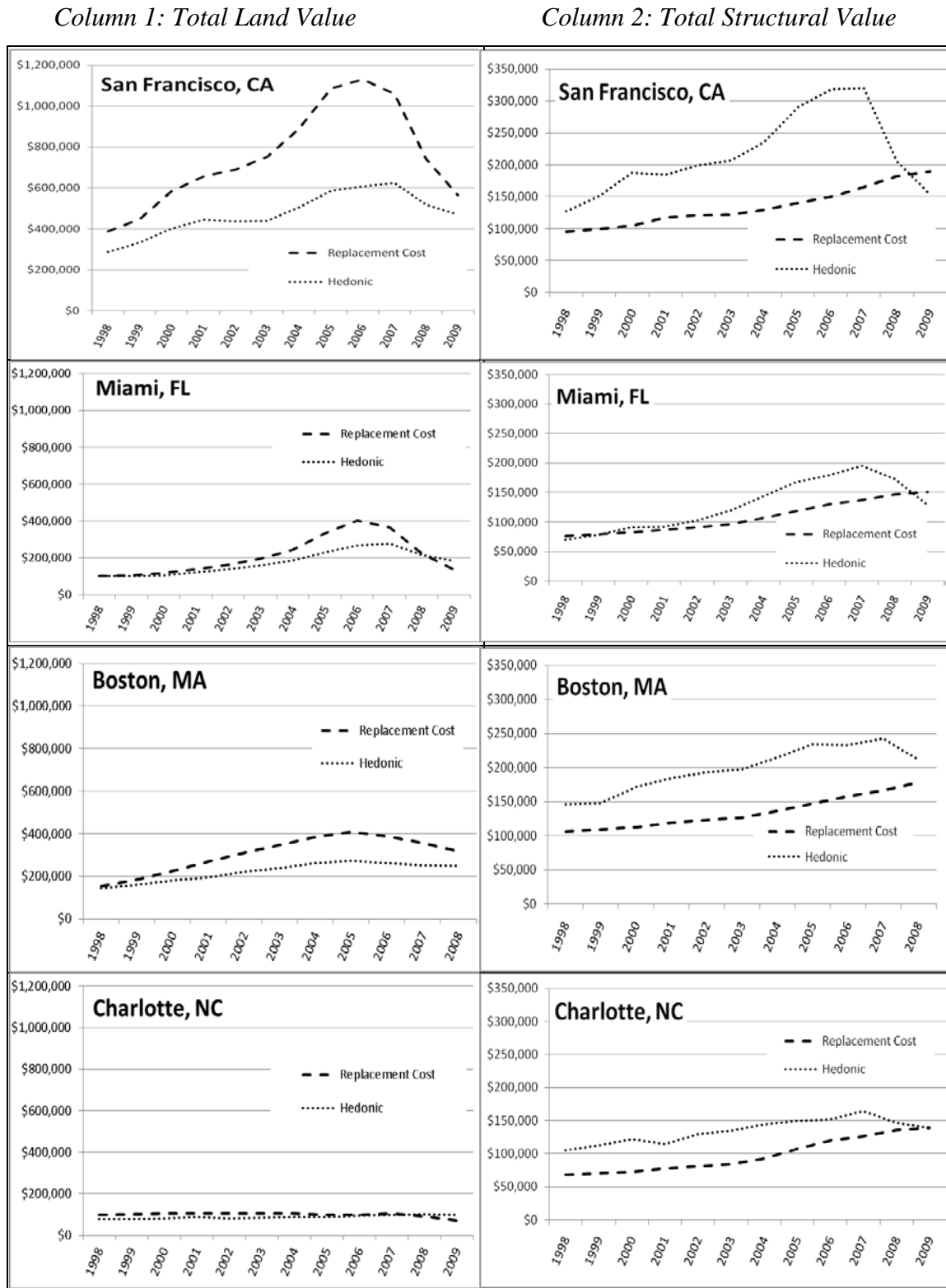
Note: This figure shows that housing prices more than doubled from 1998 to 2006, but then declined substantially from 2006 to 2009 for 20 metro areas included in the index. The data for this figure comes from the the S&P / Case-Shiller U.S. National Values Home Price Index. For documentation, see: <http://www.standardandpoors.com>.

**Figure 2: Heterogeneity in the Evolution of Housing Prices across & within Metro Areas**



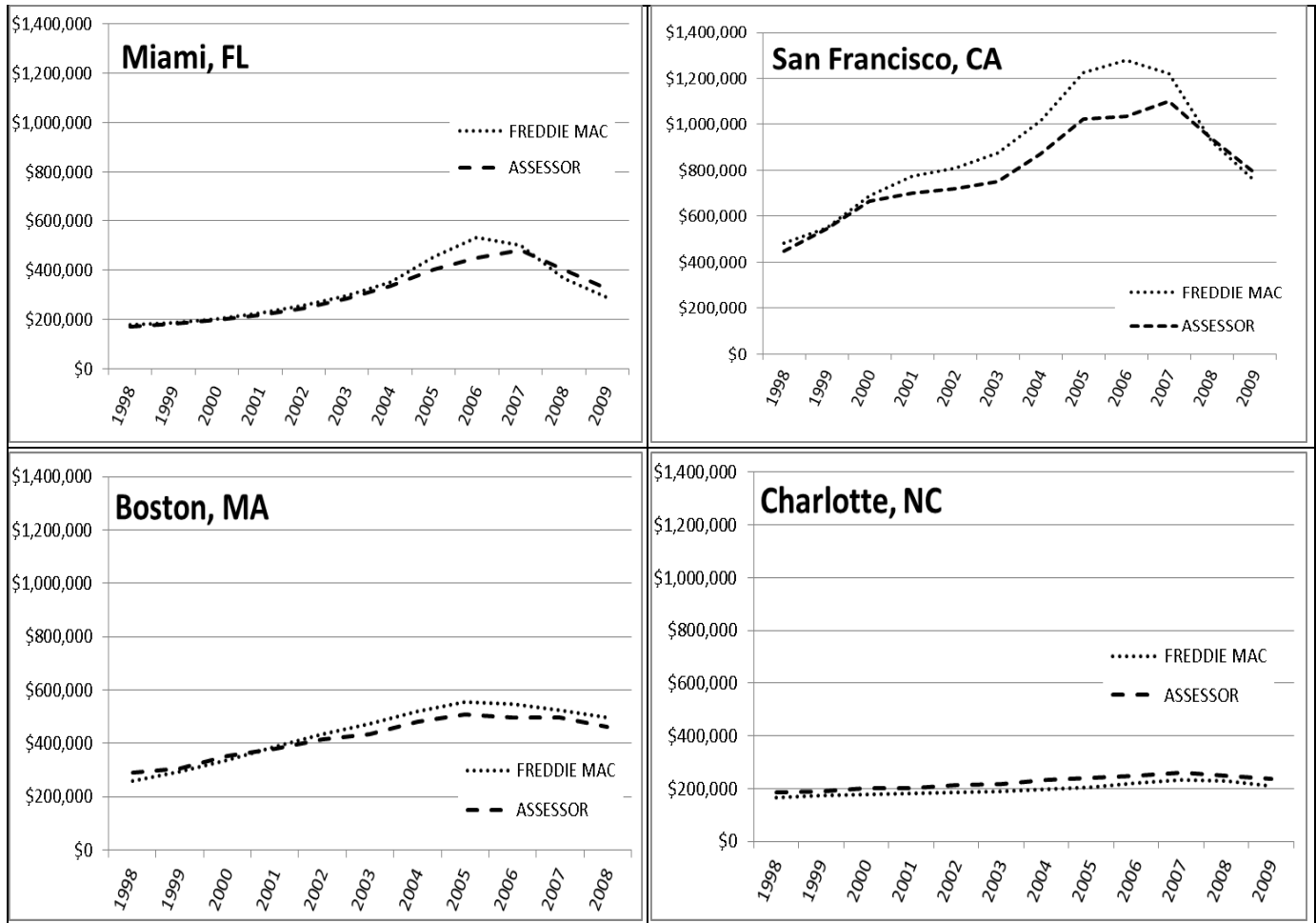
Note: This figure shows the substantial heterogeneity in price changes both across and within metro areas. The three lines show the evolution of prices for a metro’s bottom “tier”, middle “tier”, and top “tier” of the price distribution. Breakpoints are defined by metro area as of August 2010. The data for this figure also comes from the the S&P / Case-Shiller Home Price Index. \*Supply elasticities are based on Saiz (2010)

**Figure 3: Evolution of Total Land Value and Structural Value for Four Metro Areas**



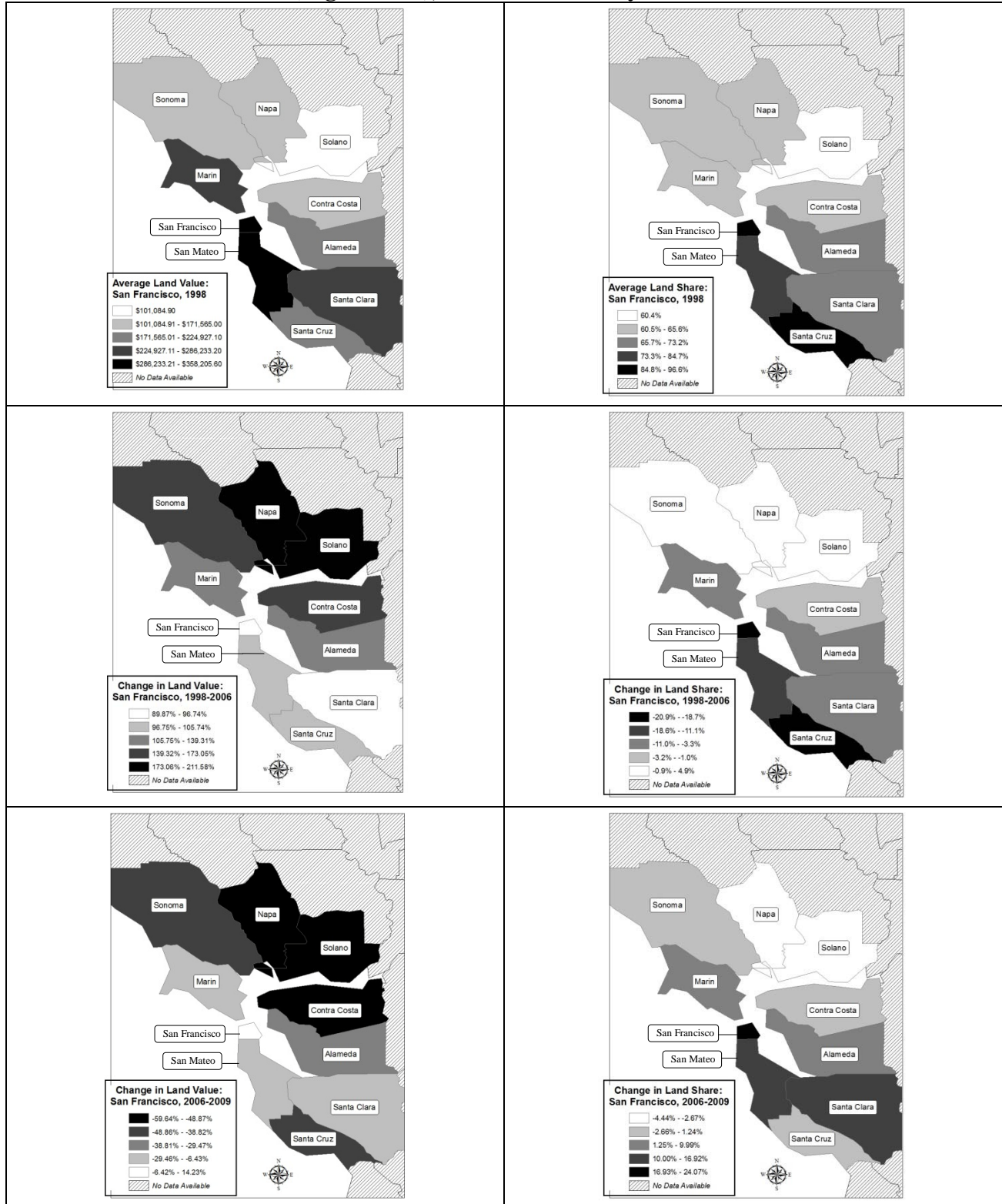
**Note:** Column 1 shows the evolution of total land value and Column 2 shows the evolution of total structural value by PMSA. Both columns show estimates derived using the hedonic method and the replacement cost method.

**Figure 4: Comparing Our Assessor Data to Freddie Mac’s Conventional Mortgage Home Price Index (CMHPI)**



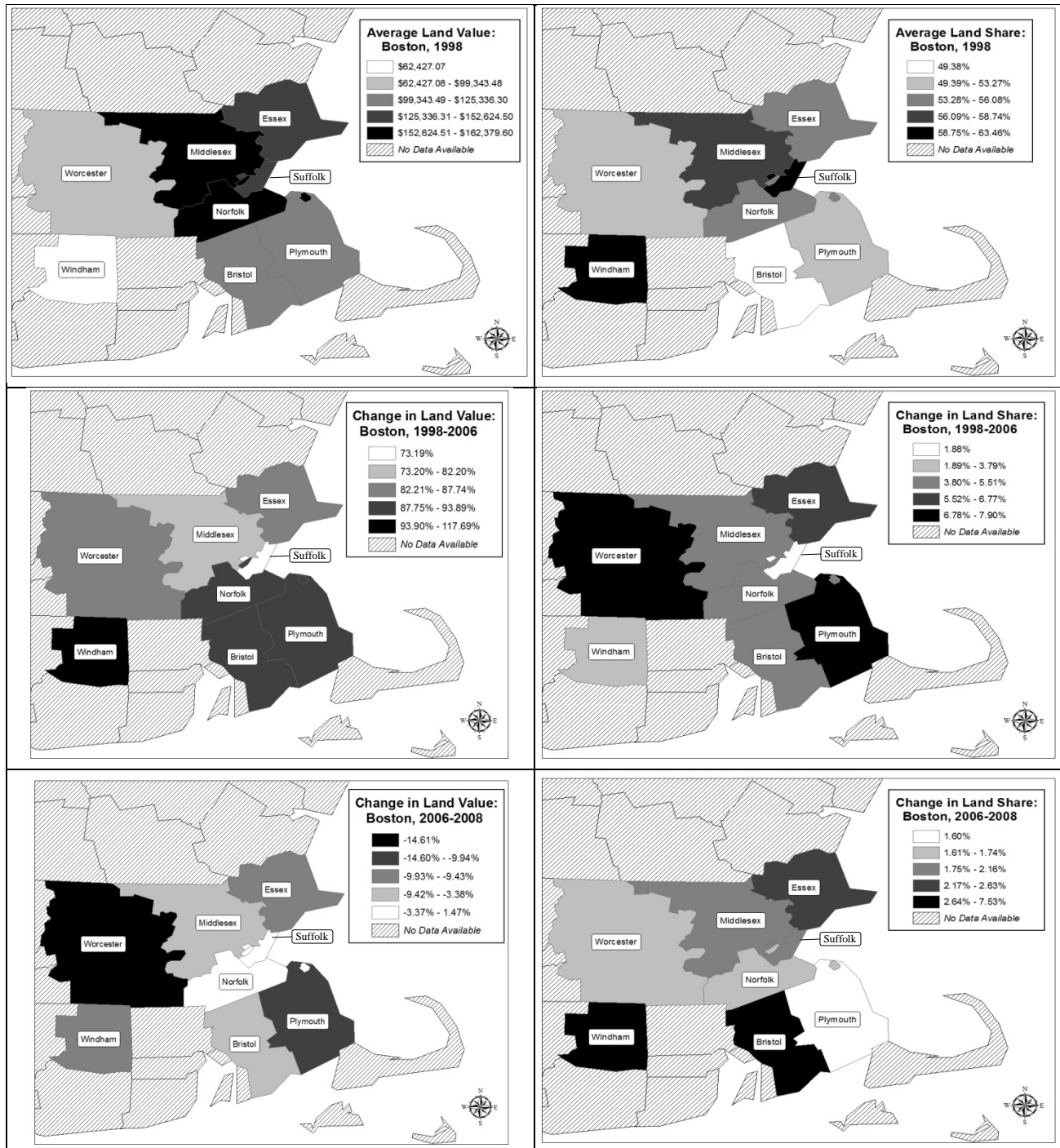
Note: Figures are produced using our assessor data and data from Freddie Mac’s “Conventional Mortgage Home Price Index” as documented in Davis and Palumbo (2008).

**Figure 5: Within Metro Land Value Heterogeneity (Left Panel) and Heterogeneity in the Evolution of Land Shares (Right Panel), San Francisco Bay Area**



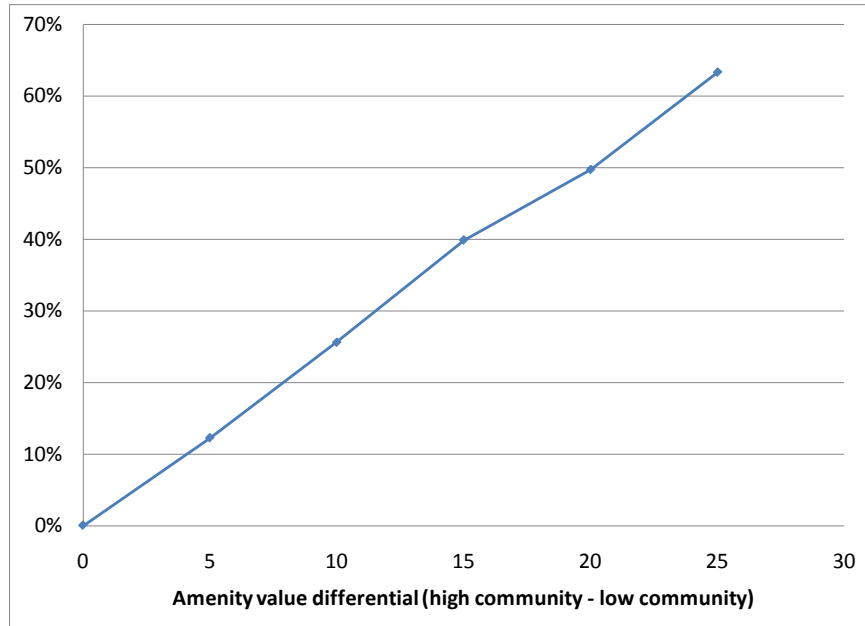
Note: These figures were produced using our hedonic approach to estimating land value. Black represents greatest absolute change, positive or negative.

**Figure 5 (Continued): Within Metro Land Value Heterogeneity (Left Panel) and Heterogeneity in the Evolution of Land Shares (Right Panel), Boston Area**



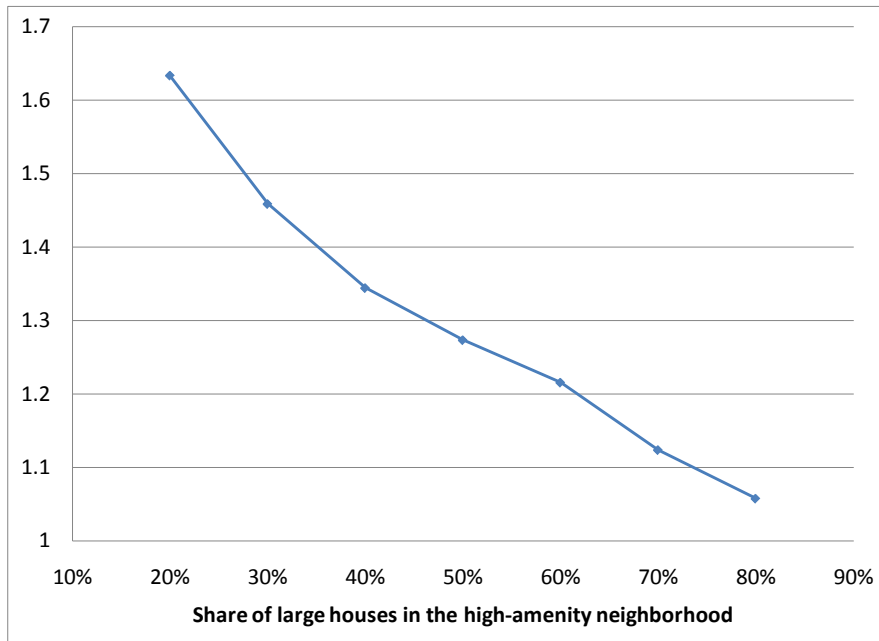
**Note:** These figures were produced using our hedonic approach to estimating land value. Black represents greatest absolute change, positive or negative.

**Figure 6A: Markup on \$/SQFT of Large Houses in the High Amenity Neighborhood**



Note: Markups are based on the hedonic simulation described in section 6. The percentage markups on the price of square footage are calculated using equation (6). In the low amenity neighborhood, 50% of houses are large and 50% are small. In the high amenity neighborhood, 20% of houses are large and 80% are small.

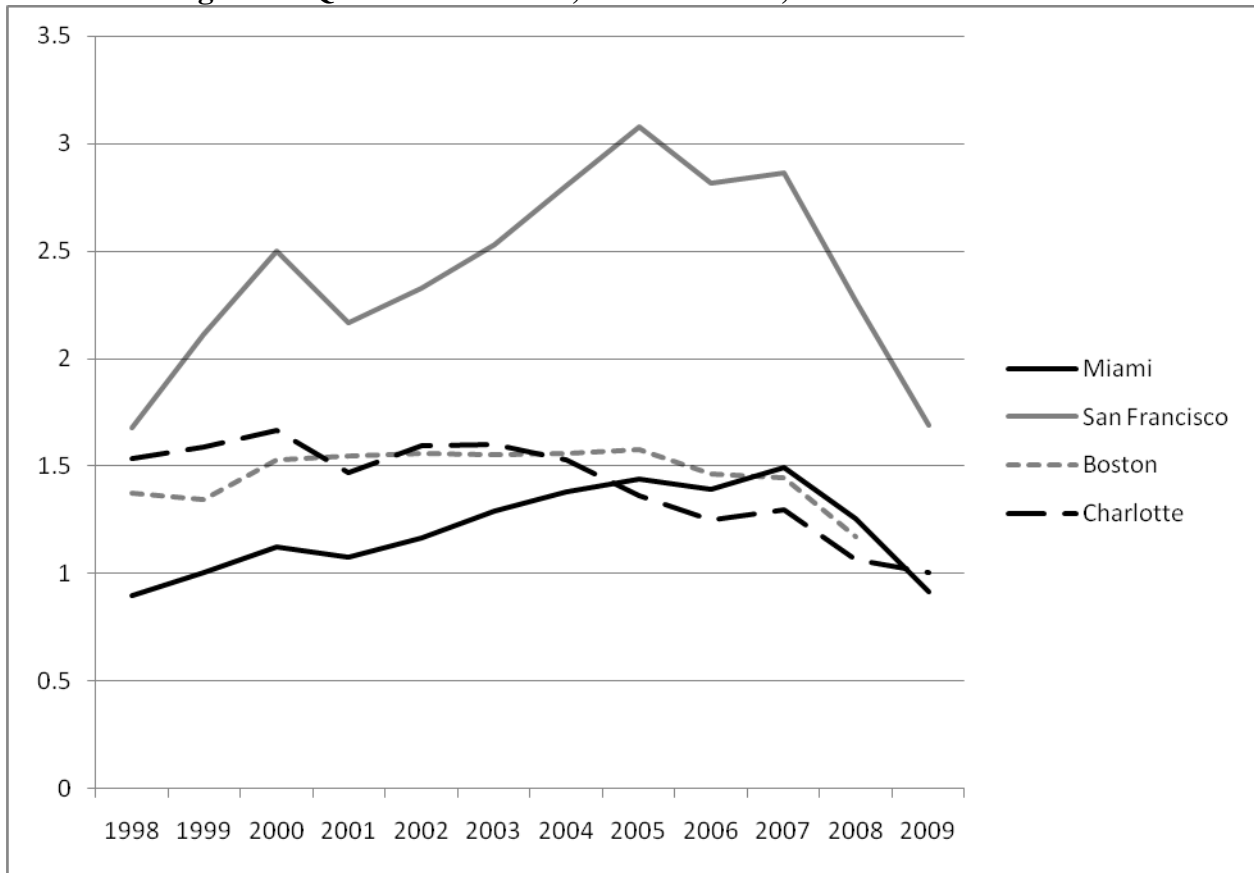
**Figure 6B: Q-Ratio and the Share of Large Houses in High Amenity Neighborhood**



Note: Q-ratios are based on the hedonic simulation described in section 6. Ratios are calculated using the difference ratio in the markup formula (6). The share of large houses in the low amenity neighborhood are held constant at 50%. The level of the amenity in the low (high) amenity community is held constant at 50 (75).



**Figure 6: Q-ratios for Miami, San Francisco, Boston and Charlotte**



Note: the q-ratios illustrated above are the ratio between the average market value of structures within the metro area and the structural costs of the average house within the metro area as given by the updated estimates of Davis and Palumbo (2008) that can be found at: <http://www.lincolinst.edu/subcenters/land-values/metro-area-land-prices.asp> .

**Table 1: Summary Statistics for Miami, San Francisco, Boston and Charlotte**

Variable		Mean	Median	St. Dev.	Min.	Max.	Obs.
<b>Miami</b>	Housing Price (1998)	161,679	129,000	144,276	17,500	3,500,000	9,799
	Housing Price (2006)	409,931	335,000	333,682	23,000	5,000,000	21,730
	Square Feet	1,948.49	1740	861.07	251	9997	194,242
	Bathrooms	2.11	2	0.80	0.5	10	194,242
	Bedrooms	3.18	3	0.83	1	10	194,242
	Year Built	1977	1983	20	1901	2008	194,242
	Lot Size (acres)	0.21	0.17	0.26	0.00	5	194,242
<b>San Francisco</b>	Housing Price (1998)	343,034	285,000	235,225	10,000	4,772,727	31,656
	Housing Price (2006)	809,485	710,000	430,730	15,000	4,900,000	42,193
	Square Feet	1,820.24	1640	769.58	260	9984	517,295
	Bathrooms	2.19	2	0.80	0.5	10	517,295
	Bedrooms	3.34	3	0.88	1	10	517,295
	Year Built	1968	1968	25	1900	2008	517,295
	Lot Size (acres)	0.21	0.14	0.31	0.00	5	517,295
<b>Boston</b>	Housing Price (1998)	275,770	234,000	177,842	16,321	4,150,000	14,399
	Housing Price (2006)	445,970	367,900	302,919	50,000	4,750,000	29,369
	Square Feet	1,875.02	1662	880.97	252	9989	281,920
	Bathrooms	1.93	2	0.83	0.5	9.5	281,920
	Bedrooms	3.25	3	0.82	1	10	281,920
	Year Built	1960	1960	30	1900	2008	281,920
	Lot Size (acres)	0.58	0.34	0.66	0.00	5	281,920
<b>Charlotte</b>	Housing Price (1998)	176,186	137,000	143,959	13,000	2,500,000	4,909
	Housing Price (2006)	231,841	175,000	205,418	7,900	3,700,000	22,552
	Square Feet	2090.49	1876	937.11	412	9968	129,596
	Bathrooms	2.44	2	0.86	0.5	10	129,596
	Bedrooms	3.31	3	0.70	1	10	129,596
	Year Built	1986	1994	21	1900	2009	129,596
	Lot Size (acres)	0.36	0.28	0.36	0.01	5	129,596

Note: Summary statistics for housing characteristics and lot size using the micro-level assessor data for single-family residential properties in Miami, San Francisco, Boston, and Charlotte.

**Table 2: Comparing Traditional Hedonic Estimates with Estimates Generated by the Replacement-Cost Method and our New Hedonic Method**

Metropolitan Area	Year	[1]	[2]	[3]	[4]
		GGS Hedonic Land Values (\$/Acre)	Our Approximation to GGS (\$/Acre)	DP Replacement Cost Land Values (\$/Acre)	Our New Hedonic Land Values (\$/Acre)
Boston	1998	29,621	20,038	237,063	212,523
Cincinnati	1999	17,424	25,700	131,220	217,927
Detroit	1999	16,117	5,227	96,927	238,939
Los Angeles	1999	112,820	67,954	804,555	857,309
Oakland	1998	101,930	94,525	976,995	908,507
Philadelphia	1999	35,284	16,988	104,087	198,530
Pittsburgh	1998	30,492	21,780	42,007	212,020
San Francisco	1998	178,596	192,971	2,421,461	1,716,395
San Jose	1998	170,755	125,017	1,533,329	1,337,703
Tampa	1998	16,117	871	122,822	176,039

Note: Column [1] reports selected land values from Table 4 of Glaeser, Gyourko & Saks (2005) or “GGS” converted to a per/acre basis. Col. [2] reports our replication of the GGS estimates, using a similar (but not identical) set of housing characteristics from assessor data. Col. [3] reports replacement-cost estimates from Davis and Palumbo (2008) or “DP.” Finally, Col. [4] reports results from our new hedonic fixed effects estimator.

**Table 3: Evolution of Land Values and Land Shares between 1998 and 2009**

	<b>Miami, FL</b>				<b>San Francisco, CA</b>			
	Hedonic Land Value	Replace. Cost Land Value	Hedonic Land Share	Replace. Cost Land Share	Hedonic Land Value	Replace. Cost Land Value	Hedonic Land Share	Replace. Cost Land Share
<b>1998</b>	100,833	100,192	0.67	0.57	286,638	384,895	0.73	0.80
<b>1999</b>	100,837	106,185	0.64	0.57	332,605	439,445	0.72	0.82
<b>2000</b>	105,339	116,585	0.62	0.59	401,393	577,058	0.72	0.85
<b>2001</b>	122,341	134,653	0.65	0.61	444,133	675,404	0.73	0.85
<b>2002</b>	138,410	159,482	0.64	0.64	438,264	685,531	0.69	0.85
<b>2003</b>	157,891	193,003	0.64	0.67	440,785	740,952	0.66	0.86
<b>2004</b>	187,726	234,739	0.64	0.69	505,022	868,277	0.67	0.87
<b>2005</b>	231,533	321,360	0.65	0.74	587,105	1,078,102	0.65	0.89
<b>2006</b>	267,107	407,144	0.68	0.76	607,558	1,143,481	0.66	0.88
<b>2007</b>	276,173	383,478	0.67	0.73	623,958	1,085,304	0.64	0.86
<b>2008</b>	218,320	230,180	0.67	0.60	517,277	781,603	0.65	0.80
<b>2009</b>	184,441	129,071	0.68	0.48	469,692	527,940	0.68	0.75

	<b>Boston, MA</b>				<b>Charlotte, NC</b>			
	Hedonic Land Value	Replace. Cost Land Value	Hedonic Land Share	Replace. Cost Land Share	Hedonic Land Value	Replace. Cost Land Value	Hedonic Land Share	Replace. Cost Land Share
<b>1998</b>	144,156	150,676	0.56	0.59	79,700	97,249	0.52	0.59
<b>1999</b>	158,979	178,134	0.59	0.63	77,395	102,112	0.50	0.59
<b>2000</b>	179,920	216,678	0.59	0.67	82,127	106,457	0.49	0.59
<b>2001</b>	196,329	264,326	0.60	0.69	88,791	105,246	0.53	0.58
<b>2002</b>	221,005	300,307	0.61	0.71	83,115	105,388	0.48	0.57
<b>2003</b>	237,622	344,810	0.62	0.73	84,361	106,634	0.47	0.56
<b>2004</b>	264,144	381,594	0.63	0.74	89,847	104,239	0.47	0.53
<b>2005</b>	272,830	413,243	0.62	0.74	89,869	98,610	0.46	0.48
<b>2006</b>	263,627	395,675	0.62	0.71	95,748	100,060	0.47	0.45
<b>2007</b>	253,552	363,197	0.60	0.68	97,479	106,024	0.46	0.46
<b>2008</b>	250,611	320,873	0.64	0.64	102,243	92,201	0.50	0.40
<b>2009</b>		277,680		0.61	97,778	70,959	0.50	0.34

Note: This table shows the evolution of total land value and the share of property values accounted for by land, by PMSA. For each metro-area, estimates are shown for both the hedonic method for deriving land values and the replacement cost method.

**Table 4: Housing Units, Property Values, and Land Values in the San Francisco Bay Area, 1998-2009**

	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]	[11]	[12]
county	Owner occupied units: 2000 (thousand)	Single- family residential permits: 2000-2007 (thousand)	New S.R.F. permits 2000-2007 / baseline owner occ units: 2000	All new permits 2000-2007 / baseline housing units: 2000	Median perceived housing value: 2000 (thousand)	Mean sale price: 1998 (thousand)	Mean land value: 1998 (thousand)	Mean land value change: 1998 to 2006 (thousand)	Mean land value change: 1998 to 2006	Mean land value change: 2006 to 2009 (thousand)	Mean land value change: 2006 to 2009	Mean land value change: 1998 to 2009 (thousand)
San Francisco Metro Area												
San Francisco	79.55	0.58	0.01	0.05	396	408	358	322	90%	97	-14%	419
San Mateo	135.61	5.67	0.04	0.04	469	430	328	332	101%	-42	6%	289
Marin	55.12	3.27	0.06	0.04	515	498	264	324	123%	-102	17%	222
Santa Clara	291.77	18.08	0.06	0.08	446	446	286	277	97%	-82	15%	195
Alameda	251.17	16.29	0.06	0.06	303	316	200	279	139%	-141	29%	138
Santa Cruz	43.43	4.17	0.10	0.06	378	273	225	238	106%	-200	43%	38
Sonoma	91.61	11.79	0.13	0.10	273	250	148	255	173%	-156	39%	99
Contra Costa	210.34	34.41	0.16	0.12	268	303	172	269	157%	-226	51%	43
Napa	23.49	4.30	0.18	0.12	251	231	137	291	212%	-209	49%	82
Solano	75.97	13.94	0.18	0.13	178	188	101	201	199%	-180	60%	21

Note: Col. [1] is based on the 2000 Census. Col. [2] is based on annual counts of permits for single-family residential construction reported by the SOCDS building permits database provided by huduser.org. Col. [3] is the ratio of col. [2] to col. [1]. Col. [4] divides the total number of permits for new housing units (all types) between 2000 and 2007 by the year 2000 stock of housing units (all types). The numerator is reported by the California Statistical Abstract. The denominator is reported by the 2000 Census. Col. [5] is the median self-reported housing value by county, from the 2000 Census. Col. [6] is the average transaction price from our county assessor data. Col. [7]-[12] are based on our estimates for land value of the average single-family residential property in each county.

**Table 5: Correlation between Increase in Land Share (1998-2006) and Baseline Land Share (1998)**

	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]
Dep. Variable = (land share in 2006 - land share in 1998)	<b>San Francisco</b>	<b>Boston</b>	<b>Miami</b>	<b>Charlotte</b>	<b>Los Angeles</b>	<b>Detroit</b>	<b>Oakland</b>	<b>Pittsburg</b>	<b>San Jose</b>	<b>Tampa</b>
<b>1998 Land Share</b>	-0.574*** (-0.062)	-0.448*** (-0.041)	-0.540*** (-0.060)	-0.520*** (-0.081)	-0.732*** (-0.026)	-0.646*** (-0.051)	-0.710*** (-0.038)	-0.393*** (-0.086)	-0.754*** (-0.068)	-0.639*** (-0.095)
<b>Constant</b>	0.359*** (-0.048)	0.306*** (-0.024)	0.383*** (-0.044)	0.235*** (-0.045)	0.528*** (-0.020)	0.375*** (-0.037)	0.483*** (-0.029)	0.237*** (-0.057)	0.519*** (-0.052)	0.386*** (-0.058)
<b># of Census Tracts</b>	103	359	141	70	650	180	247	48	142	143
<b># housing transactions</b>	47,687	181,617	132,498	82,291	312,059	88,570	172,735	24,592	75,863	100,526
<b>R-squared</b>	0.458	0.255	0.365	0.375	0.554	0.479	0.591	0.312	0.467	0.242

Note: Standard errors in parentheses; \*\*\* p<0.01, \*\* p<0.05, \* p<0.1

**SUPPLEMENTAL APPENDIX: NOT FOR PUBLICATION**

**Table A1: Housing Units, Property Values, and Land Values in the Boston Area, 1998-2008**

	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]	[11]	[12]
county	Owner occupied units: 2000 (thousand)	Single-family residential permits: 2000-2007 (thousand)	New S.R.F. permits 2000-2007 / baseline owner occ units: 2000	All new permits 2000-2007 / baseline housing units: 2000	Median perceived housing value: 2000 (thousand)	Mean sale price: 1998 (thousand)	Mean land value: 1998 (thousand)	Mean land value change: 1998 to 2006 (thousand)	Mean land value change: 1998 to 2006	Mean land value change: 2006 to 2008 (thousand)	Mean land value change: 2006 to 2008	Mean land value change: 1998 to 2008 (thousand)
<b>Boston Metro Area</b>												
Suffolk	39.36	0.81	0.02	N/A	187	249	153	112	73%	4	-1%	116
Norfolk	144.18	7.10	0.05	N/A	230	330	157	143	91%	-4	1%	138
Essex	139.93	8.38	0.06	N/A	220	296	145	127	88%	-26	10%	101
Middlesex	268.54	16.20	0.06	N/A	248	312	162	133	82%	-10	3%	123
Bristol	99.92	10.56	0.11	N/A	152	255	115	108	94%	-14	6%	93
Plymouth	110.22	12.14	0.11	N/A	179	267	125	116	92%	-24	10%	92
Worcester	149.39	21.78	0.15	N/A	146	211	99	85	86%	-27	15%	58

Note: Col. [1] is based on the 2000 Census. Col. [2] is based on annual counts of permits for single-family residential construction reported by the SOCDS building permits database provided by huduser.org. Col. [3] is the ratio of col. [2] to col. [1]. In col. [4] we were unable to obtain data on the total number of permits for new housing units in a format that would be comparable to the data on San Francisco in table 4. Col. [5] is the median self-reported housing value by county, from the 2000 Census. Col. [6] is the average transaction price from our county assessor data. Col. [7]-[12] are based on our estimates for land value of the average single-family residential property in each county.

## Grouped Footnotes

---

<sup>1</sup> In the literature on land valuation, the hedonic approach is also referred to as the “contribution” approach.

<sup>2</sup> Indeed, some prior hedonic estimates for land value seem strikingly low. For example, the estimates in Glaeser, Gyourko and Saks (2005) suggest that the value of a full acre of land in Boston in 1998 was less than 30k and in San Francisco it was less than 200k.

<sup>3</sup> The data comes from the the S&P / Case-Shiller U.S. National Values Home Price Index. For documentation, see: <http://www.standardandpoors.com>.

<sup>4</sup> One can find the same pattern in the other 16 major metropolitan areas tracked by the Case-Shiller index.

<sup>5</sup> The new empirical literature on Tiebout sorting stresses the need to recognize that neighborhood amenities are typically endogenous to the collective location choices made by the households in a metropolitan area (Kuminoff, Smith, and Timmins, 2010). For example, urban development may provide opportunities for dining and nightlife, while increasing traffic congestion and degrading air and water quality. Homeowners may be asked to vote on assessments to fund open space preservation and public schools. Academic performance among students in those schools may depend on the distribution of income and education among parents in the school district. While we do not model these mechanisms, our framework is consistent with their presence.

<sup>6</sup> Theorem 1 can also be proven under an alternative assumption that households ignore their own contributions to the supply of neighborhood amenities.

<sup>7</sup> Our theorem recognizes that neighborhood amenities may be determined endogenously through a Tiebout sorting process. In contrast, Bajari and Benkard (2005) characterize markets where product attributes (other than price) are determined exogenously. They also model unobserved product attributes and restrict utility to be Lipschitz continuous in order to guarantee Lipschitz continuity of the price function. While it is straightforward to add these elements to our model, they are unnecessary to guarantee the existence of a price function.

<sup>8</sup> This assumes the undeveloped parcel is also zoned for residential development.

<sup>9</sup> Cheshire and Sheppard (1995) distinguish between these two components of land value. While we could certainly do the same, it is not essential to our analysis.

<sup>10</sup> While the published version of Davis and Palumbo’s paper presents estimates for 1984 to 2004, the Lincoln Institute of Land Policy maintains a webpage where their estimates are updated as new data become available: <http://www.lincolnst.edu/subcenters/land-values/metro-area-land-prices.asp>

<sup>11</sup> In earlier work, Jackson (1979) used aggregate census tract data to estimate a coarse approximation to a hedonic price surface in Milwaukee. In principle, his results could also be used to develop an approximation to the value of land.

<sup>12</sup> In general, the shape of the equilibrium hedonic price function will vary with changes in tastes, wealth, regulations, and spatially delineated amenities. See Kuminoff, Smith, and Timmins (2010) and Kuminoff, Parmeter, and Pope (2010) for details.

<sup>13</sup> Abbott and Klaiber (forthcoming) make a similar point in the context of identifying what occupants are willing to pay for a particular amenity.

<sup>14</sup> The U.S. Census Bureau defines census tracts to be “as homogeneous as possible with respect to population characteristics, economic status, and living conditions.” See “Chapter 10: Census Tracts and Block Numbering Areas, U.S. Census Bureau”, Geographic Areas Reference Manual which can be found at <http://www.census.gov/geo/www/garm.html>.

<sup>15</sup> For example, all of the houses in a Census tract may be located near public open space but the handful of lots that are adjacent to the public lands may sell at an additional premium.

<sup>16</sup> A referee notes that we could have used an explicit spatial model to attempt to mitigate the bias from spatial correlation between housing attributes and omitted variables. Examples of explicit spatial models include the spatial error, spatial lag, general spatial, and spatial Durbin models. We chose not to use these models for three reasons. First, there is prior evidence that spatial fixed effects outperform explicit spatial models in terms of recovering accurate estimates for the implicit prices of housing attributes (Kuminoff, Parmeter and Pope 2010). Second, our fixed effects model produces estimates that are very similar to Davis and Palumbo’s (2008) replacement cost calculations prior to the boom, as we discuss in section 5.1. We interpret this evidence as validating our model since we expect the replacement cost methodology to provide good estimates for land values during periods of market stability. The third reason is that our sample sizes and covariate matrices are too large to implement the maximum likelihood routines for estimating flexible spatial autocorrelation models. As computing power continues to improve, an interesting avenue for future research would be to develop new models that combine spatial fixed effects with strategies for addressing more localized forms of spatial autocorrelation within neighborhoods.



- 
- <sup>17</sup> The other six are Cincinnati, Detroit, Los Angeles, Philadelphia, Pittsburgh, and Tampa.
- <sup>18</sup> His estimates are generated using information on geographic constraints, regulatory constraints, and pre-determined population levels in each metro area.
- <sup>19</sup> A related literature on gentrification investigates how the demographic composition of neighborhoods within a metropolitan area changes as people migrate from the suburbs to the cities, and vice versa. See McKinnish, Walsh and White (2010) for an interesting example of this line of research as well as citations to the broader literature.
- <sup>20</sup> The results (reported in their Table 4) support their hypothesis that the areas that we would expect to be more highly regulated have larger differences between construction costs and housing prices.
- <sup>21</sup> Our results are generated using a simple linear model estimated according to equation (4), using the combination of control variables that comes as close as possible to the specification from GGS. Complete results will be provided upon request.
- <sup>22</sup> For example, the CMHPI uses a slightly different definition for metro areas than the AHS. Also, since DP do not report average lot size, we use the average lot size in our data to convert the DP estimates to a \$/acre measure. So, there are certainly some differences in the estimates that are caused by differences in spatial-temporal components of the underlying datasets but we think that it is highly unlikely that these drive the large differences we document between GGS and DP.
- <sup>23</sup> We dropped tracts with fewer than 15 observations per year to better ensure accurate estimates of tract-specific land values. This drop (approximately 33.5% of our observations) was necessary given our census tract fixed effects identification strategy.
- <sup>24</sup> The replacement cost results are the Davis and Palumbo (2008) estimates that have been updated and provided at <http://www.lincolnst.edu/subcenters/land-values/>. Also note that we were unable to obtain assessor data for the year 2009 in Boston. Therefore all results reported for Boston are for the 1998 to 2008 time frame.
- <sup>25</sup> Charlotte and Miami have only 1 or 2 counties with available assessor data, so their maps are less interesting.
- <sup>26</sup> A related This is despite the fact that many have been concerned about gentrification in urban areas (i.e.) which is likely to dampen the volatility.
- <sup>27</sup> All five trends are also present in the Boston area. For brevity, we provide a table with results in the supplemental appendix.
- <sup>28</sup> 62.5% is the value for the 1998 land share that would correspond to a prediction of no change in the land share over the boom. It is calculated by dividing the regression intercept by the slope coefficient ( $=0.359/0.574$ ). The regression predicts decreases for larger baseline land shares and increases for smaller baseline land shares.
- <sup>29</sup> Also see their paper for a useful summary of the recent literature on the role of credit in the housing market boom and bust.
- <sup>30</sup> Matlab code to reproduce the simulation results is available as a supplemental online appendix.
- <sup>31</sup> Since the difference ratios are based on average prices, rather than marginal prices, we would not necessarily expect them to be exactly equal to 1 in a long-run equilibrium.
- <sup>32</sup> A rare exception is residential communities that are built on swamps that have been drained or wetlands that have been filled.
- <sup>33</sup> Jud and Winkler (2003) apply q-theory to housing values without decomposing them into the value of land and structures.
- <sup>34</sup> For a discussion of aggregation issues in a setting with multiple capital goods and citations to the broader literature on the empirical measurement of q-ratios, see Wildasin (1984).
- <sup>35</sup> We are grateful to David Wildasin for first bringing this to our attention.