

The Distribution of Urban Land Values: Evidence from Market Transactions

David Albouy
University of Illinois and NBER

Gabriel Ehrlich
Congressional Budget Office

© 2014 Lincoln Institute of Land Policy

Lincoln Institute of Land Policy Working Paper

The findings and conclusions of this Working Paper reflect the views of the author(s) and have not been subject to a detailed review by the staff of the Lincoln Institute of Land Policy.

The views expressed in this paper are the authors' and should not be interpreted as the views of the Congressional Budget Office.

Contact the Lincoln Institute with questions or requests for permission to reprint this paper. help@lincolnist.edu

Lincoln Institute Product Code: WP14xxx [Provided by Lincoln staff]

Abstract

We analyze land values from market transactions across the United States. Across space, land values are distributed widely and almost log-normally. The strongest predictor of value per acre is lot size, followed by location, and then time. Urban and agricultural land markets appear unified after accounting for fixed conversion costs. The relationship between peak metropolitan land values and populations, agricultural values, and arcs of urban expansion is roughly consistent with monocentric-city predictions and within-metro estimates of the rent gradient. Our estimates suggest land receives 7 percent of income and the cost elasticity of urban population is 3 percent of income.

About the Authors

David Albouy is associate professor of Economics at the University of Illinois at Urbana-Champaign. His research interests include public, urban, and labor economics, with particular interests in federal tax and transfer policy, housing, and local quality of life. He can be contacted at albouy@illinois.edu.

Gabriel Ehrlich is principal analyst at the Congressional Budget Office. His research interests include macroeconomics and urban economics. He can be contacted at Gabriel.Ehrlich@cbo.gov.

Acknowledgements

We would like to thank seminar participants at the Lincoln Institute of Land Policy and Tulane University for helpful comments. Any errors are our own.

Table of Contents

1. Introduction.....	1
2. Literature Review	2
3. Data Description.....	4
3.1. Source of Land Data	4
3.2. Non-Price Characteristics of the Land Parcels	5
3.3. The Statistical Distribution of Land Prices	6
4. Intra-Metropolitan Variation in Urban Land Values	10
4.1. Lot Size and Distance	10
4.2. Using the Land-Rent Gradient the Estimate Land's Share of Income.....	12
4.3. Intended Use and Land-Use Regulations.....	14
4.4. Overall Co-variation of Land Values with Space, Time, and other Observables.....	14
5. Inter-Metropolitan Variation in Urban Land Values	16
5.1. The Urban Fringe and Agricultural Land	18
5.2. Central Land Values, Population, Size, and Amenities	20
6. Conclusion	23
References.....	23

The Distribution of Urban Land Values: Evidence from Market Transactions

1: Introduction

Land is a fundamental factor of the economy, and its value is one of the primary topics of urban economics. The price of unimproved land measures the value of property rights over space itself. Much of the theoretical literature in urban economics seeks to explain the determinants of land prices through local amenities, access to employment, and regulations affecting land use. Land values are central to understanding property prices and assessments, the impact of land-use policies, and taxes levied on property or land. In some contexts, land values may also be used to determine the costs of urban agglomeration, the optimal level of public-good provision, the welfare effects of local policies, and even the optimal size of a city.

Despite the importance of land values, data on urban land values have remained rare. Data used for cross-sectional purposes typically focus on a single city. In this paper, we provide a novel look at land values using observed market transactions in the CoStar/COMPS database. To our knowledge, our paper is the first of its kind to make a cross-sectional examination using data that span a wide range of U.S. cities.

We describe both the wide variation in land values within and across metropolitan areas. When possible, we do so through the lens of the stand urban model of the mono-centric city, but also consider other important determinants. We begin by considering the statistical distribution of land of cities across the United States and within metropolitan areas. Next, we consider how the value of individual lots depend on their acreage, quarter of purchase, intended use, local regulations, metro area, and proximity to the center of that metro area. At the county level, we consider how urban land values are related to agricultural values. Across metro areas, we examine how the highest land values in a metropolitan area relate to its population, the opportunity cost of land, the arc at which the city may expand from its center and intra-metropolitan variation in amenities.

Our statistical analysis, described in section 3, reveals several striking features. First, the statistical distribution of land values per acre is decidedly convex, skewed to the right, and roughly log-normal, both nationally and within metropolitan areas. Second, the standard deviation of log prices within metro areas is greater than one and roughly similar across different cities. In levels, the distribution of land values fans out considerably with the metropolitan average, while in logarithms the distribution hardly fans out at all.

When we consider the determinants of individual parcels in section 4, we find that lot acreage is the strongest single predictor of price per acre: it explains over 40 percent of the variation in a simple regression. Indicators for metropolitan area increase the explanatory power of the regression to 57 percent. When we use various measures of distance from the city center to predict prices, distance measured in driving time appears to predominate versus straight-line distance and driving distance. Using the slope of the rent gradient, simple mono-centric theory, and a value for the costs of commuting, we impute an income share of residential land of roughly 7 percent.

We also consider how lot values depend on stated “intended-use” variables and regulations affecting land use. Retail uses are associated with higher values and industrial use, lower values. Regulations are associated with lower land values, especially when held for development, but not for land intended for residential development. Indicators for quarter of sale explain some of the variation, but much less than the spatial variables. Interactions between time of sale and other variables do not appear particularly important.

In section 5, we consider variation in land values at the county and metropolitan level. We find that land values for urban uses at the 10th percentile in each county are strongly and positively related to values for agricultural uses. Non-linearity in the relationship is likely to reflect costs of converting agricultural land to urban use.

We finish by examining how the central land value of a city depend ultimately on the population size of a metro area, the agricultural values at its fringe, and the available arc of land available for expansion away from the center. The analysis is embedded within a circular mono-centric city model with an expandable border. Central land values appear to be largely consistent with this theory, and predict that the elasticity of central land rents with respect to population is in the vicinity of 0.4. Multiplying this by the income share of land results in a predicted elasticity of urban costs with respect to population of 3 percent of income.

2: The Statistical Distribution of Urban Land Values

In this section we examine the statistical distribution of land values. Such an exercise was much harder to execute with previous datasets. This paper stands out from the rest of the literature for being one of the only cross-sectional analyses of land values across the United States using transaction data. Cross-sectional studies on land values have typically focused on a single city. Most have not used data from market transactions.

2.1: Previous Analyses of Land Values

The value of land has been a central component of land values since the Physiocrats of Quesnay (1758), who argued that land was the source of all value. Ricardo (1817) introduced the concept of marginal analysis to explain the origin of land value, and advocated for the taxation of land values, an idea later popularized by George (1879). Von Thunen (1826) developed a sophisticated monocentric model of agricultural values, which was echoed in the standard urban model developed by Alonso (1964), Mills (1967), and Muth (1969). Arnott and Stiglitz (1979) later argued that taxing land values would lead to optimal public good provision, coining the “Henry George Theorem.” These ideas were expanded by Hochman (1981) and others, reinforcing the centrality of land values in capturing the value of local amenities and economies of scale in production and consumption.

Two articles have used transactions data from the CoStar COMPS data to examine the variation in transaction prices in several cities over time. The first, by Sirmans and Slade (2012), aggregates land values over 20 metro areas to develop quarterly indices of land values from 1991

to 2009, which are separated by residential, commercial, and industrial uses. The second, by Nichols, Oliner and Mulhall (2103), pools uses and develops semi-annual indices from the mid-1990s to mid-2011 for 23 metro areas. These analyses make seek to control for spatial variation in land variation to focus on variation over time.

A few other articles have used the CoStar-COMPS database to examine land-value variation within metro areas. These cover Chicago (Colwell and Munneke 1997), New York (Haughwout, Orr, Bedall 2008), the San Francisco Bay Area (Quigley, Kok, Monkokken 2010).¹ Prior to these works, most land value analyses in the United States used long but fairly thin data sets from Chicago provided by Hoyt (1936), Olcott (1910-1970) and later sources.

The only article that examines cross-metro variation in housing prices across the United States is by Davis and Palumbo (2008). They construct a land-value index for 46 metro areas from 1984 to 2006 by measuring “the value of land residually as home value less the replacement cost of the structure.” In general, they estimate that land’s share of home values --- 32 percent in 1984, and 51 percent in 2004 --- that are very high relative to most previous estimates. The methodology is subject to several caveats. First, the residual method only applies to limited cross section of homes. Second, the residual measure of land value also captures any profits developers make, added costs due to land-use regulations, and additions to the property not observed in the data. The data have no defined location within metro areas, making much of the analysis we do below impossible.

Davis and Heathcote (2007). Case (2007) uses a similar approach to assess the combined value of land in residential and non-residential uses. He notes that the residual method can easily produce negative values for land underneath commercial properties.²

Measuring land values using market transactions avoids many of the difficulties inherent in the residual approach, but faces problems of sample selection. Market transactions are not determined randomly, but require someone to buy and to sell. In discussing agricultural land, Ricardo (1817, 2.11) remarked that the most favorably situated” would be developed first. This would suggest that our land sample is negatively selected. At the same time, many transactions, particularly central ones result from teardowns. Teardowns may be positively selected since developers generally seek to replace an existing structure with a larger one.

An added complication is that the observed parcel prices differ in size, and larger lots represent a larger quantity of land, although spatial correlation undermines weighting all acres equally.

Traditional methods of correcting for sample selection bias have not been used, as no one has yet created a data set with unsold lots. Previous studies have considered how land use may be selected across observed lots. McMillen et al. (1992) argue that land converted from agricultural

¹ Other studies have examined Vancouver’s tear-down market (Rosenthal and Helsley 1994), and assessments from Pittsburgh (Epple, Gordon, and Seig, 2010) selected Florida counties (Ihlandfeldt 2007).

² On some properties, the value assigned to land can also be very high, which may reflect physical or regulatory burdens involved with building on acquired land, rather than the costs of acquisition itself.

to urban use tends to be below average quality for agricultural purposes, but above average quality for urban use. Colwell and Munneke (1997) finds fairly little evidence of bias on values due to selection between residential, commercial, and industrial uses. Munneke and Slade (2000, 2001) also find little evidence of selection bias in the Phoenix office market. Finally, Fisher et al. (2007) find no evidence of selection bias in their study of commercial real estate properties. Although in theory the selection bias problem appears to be rather difficult, the evidence available suggests that it may be less problematic in practice.³

They regress land prices on various characteristics, and find, among other results that land prices decline in lot size, distance from the city center. Nichols et al. (2013) use the data to construct time series for commercial and residential land prices using market transactions across 23 metropolitan areas. However, the indices in their paper are not comparable cross-sectionally. This paper complements that work by considering a broader group of metropolitan areas, although over a shorter time period, and focusing on research questions related more to variation in land values over space rather than time. Sirmans and Slade (2011) construct a time series index of land prices using observed market transactions, but do not examine variation across or within metropolitan areas. Combes, Duranton, and Gobillon (2012) use land transaction from France to estimate the costs of urban agglomeration in the context of a monocentric city model.

3: Data Description

3.1: Source of Land Data

We collect data on land transactions recorded in the CoStar COMPS database.⁴ The database includes transaction details for all types of commercial real estate, including what CoStar terms “land.” Here, we take every land sale in the COMPS database provided by CoStar University, which is provided for free to academic researchers. The use of data from CoStar University, rather than through CoStar’s paid service, limits the data we are able to collect for each parcel.⁵ Appendix figure A1 illustrates a typical brochure, which lists information concerning address, sale date, sale price, lot size, and proposed use.⁶

We restrict the sample to transactions that occurred between 2005 and 2010 in a Metropolitan Statistical Area (MSA).⁷ Dropping sales prior to 2005 excludes over 80,000 observations, substantially reducing the sample size. We study the period since 2005 because most of the

³ It should be noted that these papers do not make strong efforts at imposing restrictions on the selection equation to make the selection effects non-parametrically identified.

⁴ This is the same source data as in Albouy and Ehrlich (2012).

⁵ We collect the data from summary brochures produced by the COMPs database. The brochures are produced in pdf format, which we read into a format suitable for analysis using commercially available optimal character recognition software.

⁶ Unfortunately, the brochures do not contain data regarding greenfield versus brownfield status or grading and paving status.

⁷ We use the June 30, 1999 definitions provided by the Office of Management and Budget. The data are organized by Primary Metropolitan Statistical Areas (PMSAs) within larger Consolidated Metropolitan Statistical Areas (CMSAs).

MSAs in the COMPs database are missing or have extremely sparse coverage until 2005. Thus, we study a large group of MSAs over a relatively short time span, rather than a smaller group of MSAs over a longer time span, as in Nicols et al. (2013), consistent with our desire to compare land values within and across a wide range of metropolitan areas. The dataset further excludes all transactions CoStar has marked as non-arm's length or without complete information for lot size, sales price, county, and date, or that appear to feature a structure. It also excludes observations we could not geocode successfully, and those that were recorded to have zero distance or more than 90 miles or 100 minutes of driving time away from the center of their respective MSA,⁸ leaving us with 68,757 observed land sales. Finally, in order to improve the comparability of land sales across MSAs, we exclude sales with a lot size less than an eighth of an acre or more than 320 acres (one half a square mile), for a final sample of 57,157 land sales.

TABLE 1: DESCRIPTIVE STATISTICS OF NON-PRICE CHARACTERISTICS OF LAND MARKET TRANSACTIONS, UNWEIGHTED, 2005-2010

Total Observations	57,157	<i>Intended Use</i>	
Average lot size in acres (std. deviation)	15.96 (35.47)	Hold for Development	19.2%
Median lot size in acres	3.42	None	16.2%
		Single Family	10.8%
		Retail	8.3%
		Industrial	7.6%
		Multi-family	6.5%
<i>Average Distance from Center</i>		Office	6.5%
Driving Time in Minutes (std. deviation)	30.59 (16.36)	Hold for Investment	3.6%
Driving Distance in Miles (std. deviation)	22.85 (15.71)	Food/Restaurant	1.6%
Euclidean Distance in Miles (std. deviation)	15.47 (11.48)	Hotel	1.0%
		Parking	0.9%
Year of sale	2005 2006 2007 2008 2009 2010	<i>Metro Area population</i>	
		< 500,000	8.8%
		500,000 to 1,000,000	11.0%
		1,000,000 to 2,000,000	19.9%
		2,000,000 to 4,000,000	24.9%
		> 4,000,000	35.5%

3.2: Non-Price Characteristics of the Land Parcels

Un-weighted non-price statistics of the market transactions are shown in Table 1. The average lot is 16 acres, but the lot size distribution is highly right-skewed, with a median of only 3.4 acres. The average property is 23 miles in driving distance, or half an hour in driving time, from the city center. The average straight-line distance is only 15 miles, roughly the distance from central Los Angeles to Santa Monica or from central Washington DC to Fairfax City, Virginia. While

⁸ Approximately 4,000 otherwise valid observations are excluded by this requirement.

the plots appear to be somewhat peripheral, there is still substantial coverage in the central city. Appendix figure A2 provides maps of the sales distribution for selected cities, which show a mix of central and peripheral lots.

There is a strong cyclical pattern in the timing of land sales in the sample. There are roughly twice as many sales per year during the housing-boom years of 2005 to 2007 than during the housing slump period of 2008 to 2010. We present evidence regarding the extent to which this pattern might raise selection concerns later in the paper. The observations cover metro areas across a range of sizes, although the coverage appears to be slightly more representative of cities with over 1 million inhabitants.

The dataset includes a field for the property's 'intended use', which while far from exhaustive or complete, contains some potentially useful indicators. 11 percent of parcels are intended for single family development, 8 percent each for retail and industrial, and 6.5 percent each for multifamily and for office, with smaller percentages for restaurants, hotels, and parking. These categories are not mutually exclusive, and 16 percent of parcels have no stated intended use.

3.3: The Statistical Distribution of Land Prices

This section considers the statistical distribution of land values, keeping in mind that there may be different ways of weighting the sample in order to produce a representative sample. While various studies have considered the spatial distribution of urban land values, to our knowledge none has considered the statistical distribution of land values explicitly.

Ideally, a random sample of all plots of urban land would be used to obtain a consistent estimate of the distribution. With randomly selected lots, weighting parcels by their size would allow an approximation of all urban land. However, there are reasons to believe that the plots in the dataset are not randomly selected, which may provide an argument for considering alternative weighting schemes. More generally, different weighting schemes may help to answer different questions.

Table 2 contains descriptive statistics for the data sample under different weighting schemes, corresponding to different concepts of a representative sample. The first column displays unweighted statistics. The arithmetic mean price per acre is \$468,000, versus a geometric average of \$181,000, consistent with the presence of right-skewness in the distribution. However, the skewness of the distribution of log prices is negative, with some excess kurtosis. Column 2 displays the summary statistics weighted by lot size, so the results may be conceived as representing the typical acre of land within the sample. The most notable effect is to reduce the average value per acre substantially, to \$98,000 for the arithmetic mean and \$26,000 for the geometric mean. This reduction stems from the so-called "plattage effect," discussed in detail in section 4. This pattern may stem from larger plots being of lower quality than smaller plots.⁹ Even absent such selection issues, the strong spatial correlation of land prices suggests that weighting by land area will provide less representative statistics than an un-weighted average.

⁹ It is worth noting that Combes et al. (2012) find a similar pattern in a data set of French land values that is less likely to suffer from selection bias than the CoStar COMPs dataset.

An alternative weighting scheme is to weight each plot according to the area of a predefined geography. Column 3 displays statistics weighted by county area, so that sales in larger counties will have more weight relative to sales in smaller counties. Therefore, these statistics are more representative of the typical acre of land in an urban area. However, there may still be concerns that the land values are somewhat peripheral within each county.

Next we consider weights that depend on economic values. Column 4 displays statistics weighted by parcel value, giving more weight to larger plots, or plots with higher value per acre. These statistics may be considered to represent the typical dollar of land value in the sample. The arithmetic mean property value rises to \$2.5 million, while the geometric average rises to \$410,000. Notably, the skewness and excess kurtosis of the distribution in logs both fall in magnitude relative to the un-weighted case.

Finally, column 5 shows summary statistics weighted by population density of the Census PUMA where the land sale is located. This measure may therefore be conceived as representing the distribution of land values where people currently live. It is representative of residential urban land and perhaps most useful when considering land as input into housing production. The arithmetic mean price per acre is \$1.3 million in this scheme, with a geometric mean of \$333,000. The skewness of the log distribution is near zero, although there is non-trivial excess kurtosis.

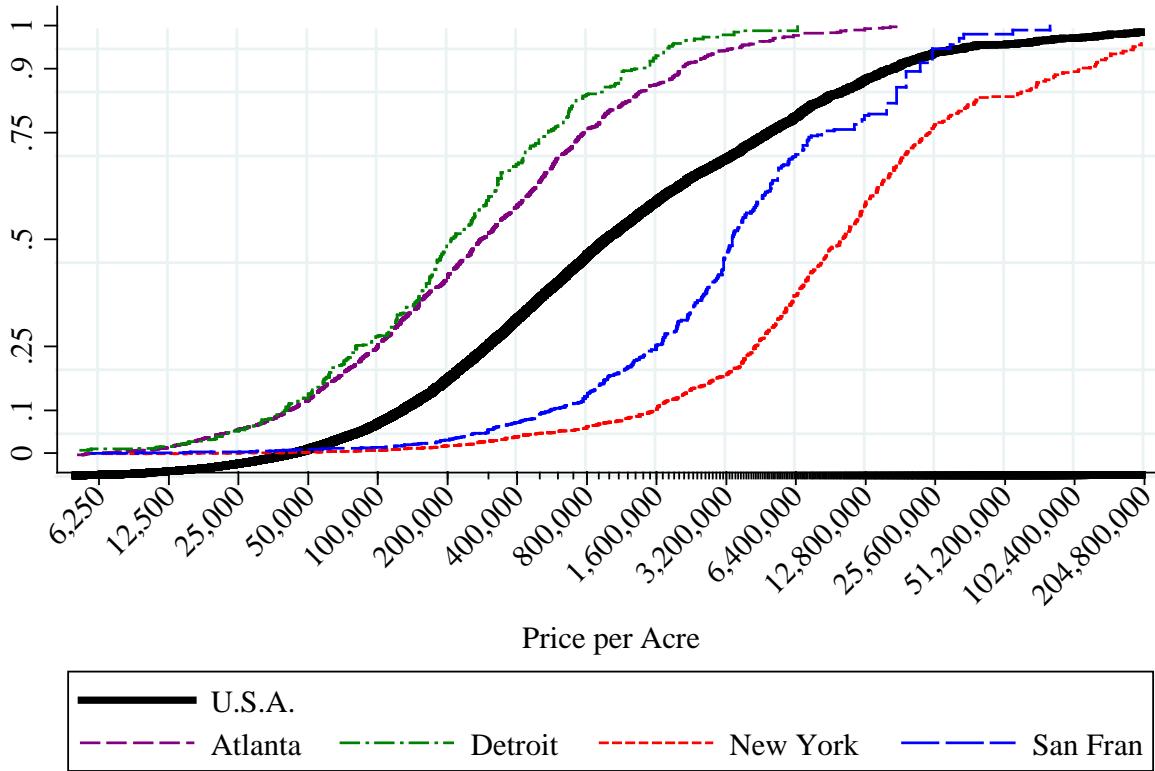
**TABLE 2: BASIC DESCRIPTIVE STATISTICS ON PRICE PER ACRE OF LAND:
NATIONAL SAMPLE**

Weights Used	Area of Property		Cty Area/ Observs	Value of Property	Population Density
	(1)	(2)	(3)	(4)	(5)
Arithmetic Mean (thousands)	\$1,280	\$165	\$518	\$13,587	\$10,177
Standard Deviation (thousands)	\$9,516	\$1,504	\$1,994	\$59,957	\$41,884
Geometric Mean (thousands)	\$272	\$48	\$150	\$801	\$1,077
Mean of Logarithm	12.5	10.8	11.9	13.6	13.9
Standard Deviation of Logarithm	1.61	1.44	1.63	2.13	2.06
Skewness of Logarithm	0.04	0.22	-0.29	0.69	0.24
Kurtosis of Logarithm	3.62	3.13	3.25	3.43	2.97

Regardless of which weighting scheme is used, it is striking that the skewness is generally close to one and the kurtosis is to three. While formal tests of normality will strongly reject that log land values are normally distributed, the lognormal distribution may still be useful as a first approximation. Furthermore, the roughly symmetrical distribution of log values implies that the distribution of level values is largely convex, particularly at the high end, as most urban models would suggest.

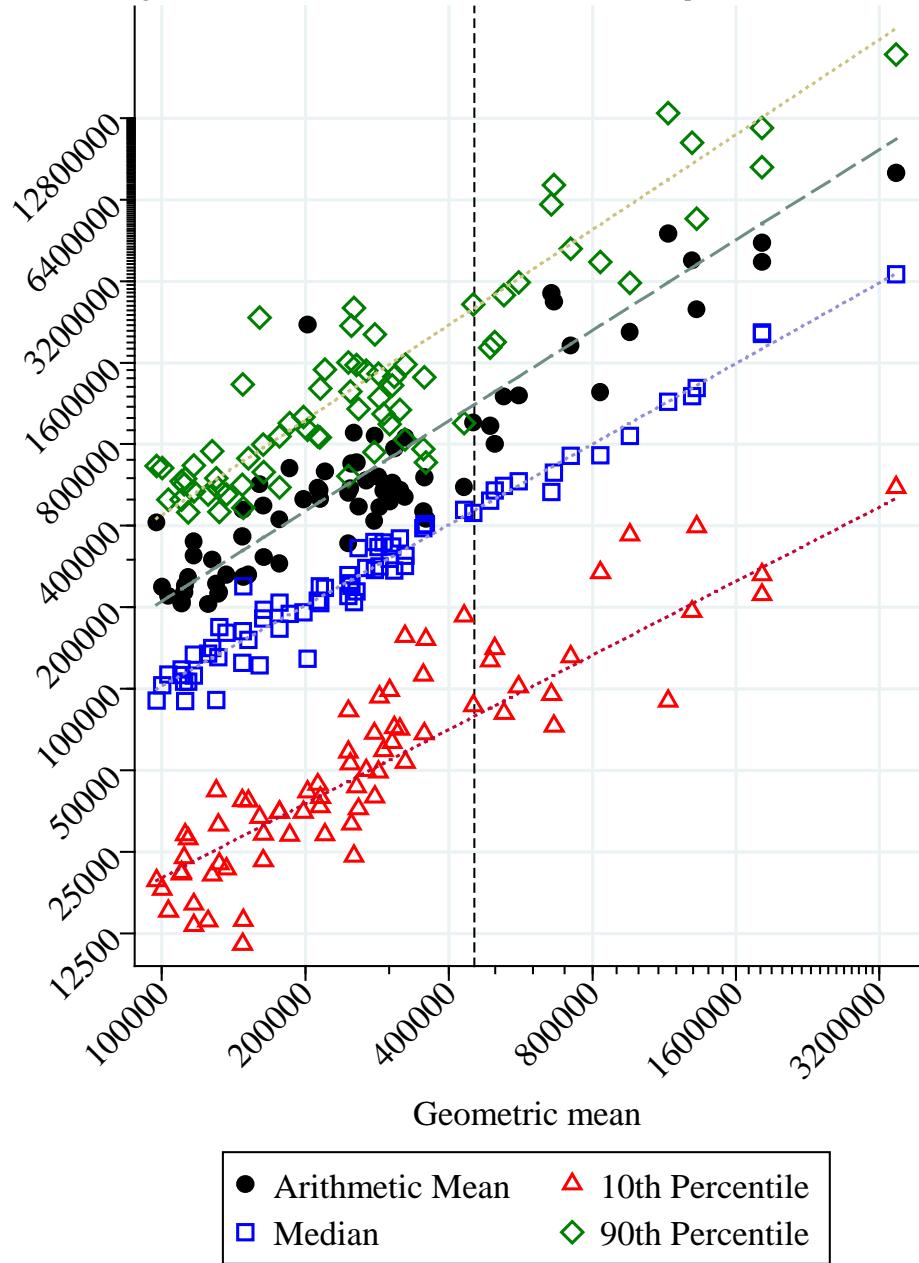
Figure 1 displays the cumulative distribution of land values both nationally and for four metropolitan areas, Detroit, Atlanta, San Francisco, and New York, for the period 2006 to 2008. The x-axis of the figure is in logs, so the relatively smooth, ‘s’-shaped distributions suggest lognormally distributed land values. Detroit and Atlanta have lower than national average land values at nearly every quantile of the distribution. The two cities similar low end land values, with 25 percent of lots in each city valued at less than \$100,000 per acre. Atlanta’s upper end is considerably higher, with a share of land valued at more than \$400,000 per acre nearly 10 percentage points higher. As expected, San Francisco and New York have much higher land values than the national average at almost all quantiles of the distribution. Visually, the figure suggests there may be some evidence of “fanning out” in land values across cities, in the sense that the distributions are not equally far apart at each quantile.

Figure 1: Cumulative Distribution of Land Values by Metro Area: 2006-2008



To examine this possibility in more detail, figure 2 plots the mean, median, 10th percentile, and 90th percentiles of land values per acre by MSAs against the geometric mean. We include only data from 2006 to 2008 from cities with at least 50 land sales in that time. The slopes of the lines of best fit are surprisingly close to one another. Table 3 shows that one cannot reject the hypothesis that the slopes of the median, 10th, and 90th percentiles are equal to one. The slope of the line for the arithmetic average is statistically larger than one, but it is not statistically distinguishable from the slopes of the individual quantiles. We interpret these results as consistent with the hypothesis that land values are distributed lognormally within metropolitan areas.

Figure 2: Distribution of Land Values across U.S. Metropolitan Areas: 2006-2008



**TABLE 3: REGRESSIONS OF MSA-LEVEL MEASURES OF LAND VALUES WITH
MEAN LOG PRICE PER ACRE**

Dependent Variable	Standard				
	Log Mean Price per Acre - Arithmetic Avg.	Deviation of Log Price per Acre	10th Percentile Log Price per Acre	50th Percentile Log Price per Acre	90th Percentile Log Price per Acre
	(1)	(2)	(3)	(4)	(5)
Mean Log Price per Acre	1.108	0.095	0.910	0.989	1.160
:Geometric Avg.	(0.038)	(0.033)	(0.054)	(0.011)	(0.053)
Constant	13.919	1.380	11.290	13.033	14.740
	(0.054)	(0.046)	(0.070)	(0.014)	(0.072)
Number of Observations	231	219	231	231	231
Adjusted R-squared	0.933	0.123	0.845	0.985	0.895

Robust standard errors, clustered by CMSA, reported in parentheses.

4: Intra-Metropolitan Variation in Urban Land Values

In this section, we consider the within-metropolitan area determinants of land values, focusing on the influence of lot size, distance from the city center, time of sale, proposed property use, and regulatory environment. We examine alternative specifications for lot size and find that the elasticity of land values with respect to lot size is not constant. We also examine alternative specifications for distance from downtown, and find that driving time appears to predominate versus other distance measures in determining land values.

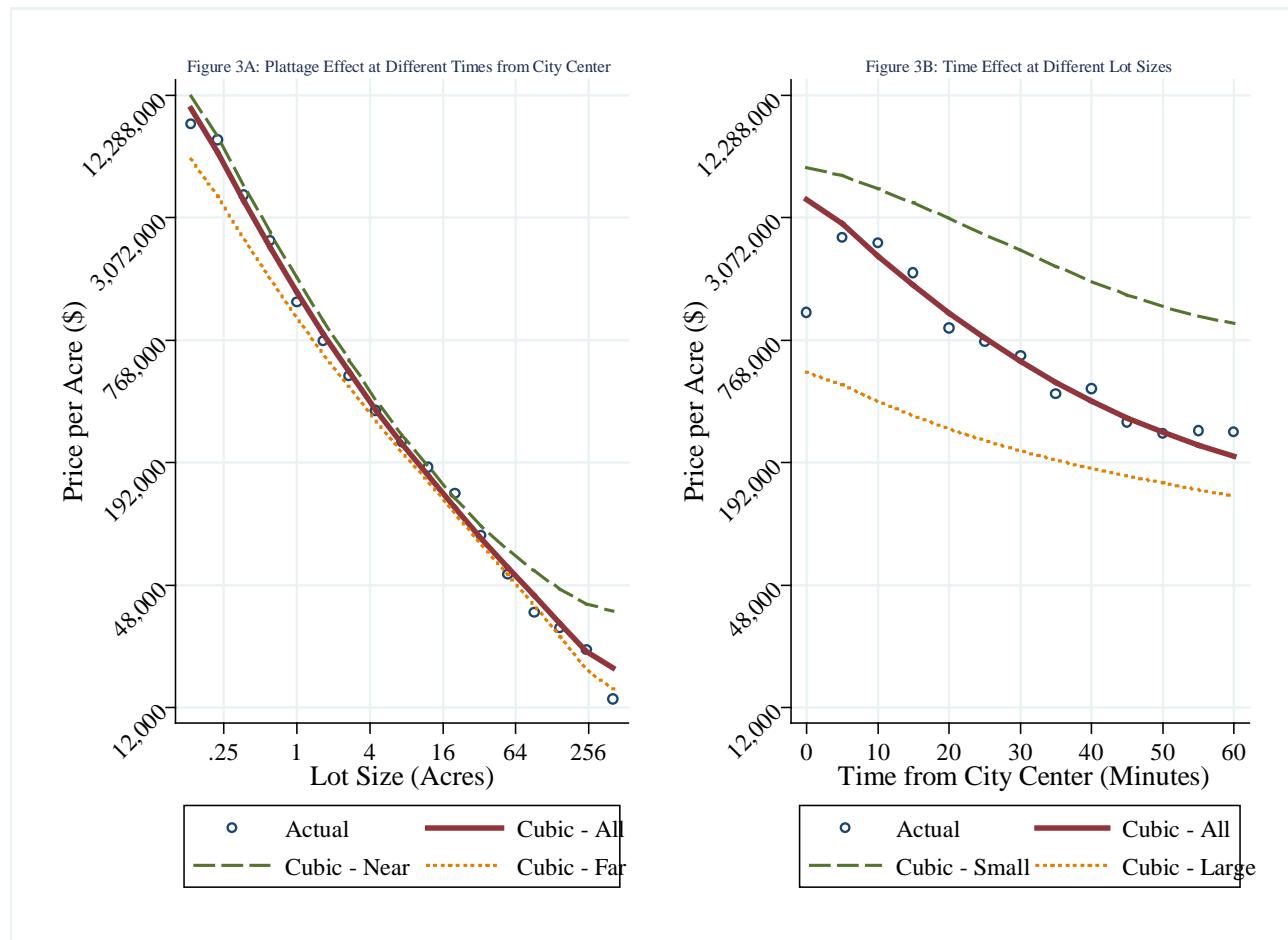
4.1: Lot Size and Distance

Two of the major determinants of land values are plot size and distance from city center. The relationship with plot size is very strong empirically, although it does not have a strong theoretical basis. The relationship with distance has always had a strong basis theoretically, although the empirical basis has not always been as strong.

The finding that lot size plays a key role in determining land values is quite robust to the inclusion of the other controls we consider. This finding is consistent with the so-called “plattage effect” that is well-documented in the literature, for instance by Colwell and Sirmans (1978, 1980, 1993, and references therein). They attribute the tendency for parcel value to increase less than proportionately with parcel size to holdout effects and the costs of subdividing and developing land. Nichols et al. (2013), who find similar plattage effects to those we estimate, argue that the effects are likely to stem from the existence of an optimal scale for buildings of a certain type, which reduces the value of land beyond the necessary scale. Another possibility is

unobserved heterogeneity in parcel quality: if smaller parcels tend to be of higher quality, the observed plattage effect could arise from omitted variable bias.¹⁰

To provide a basic picture of the data, Figure 3 illustrates the pattern of land values across lot sizes and distances from downtown. The solid line in Figure 3A shows predicted value per acre as a function of lot size estimated over the entire sample, while the blue circles show corresponding cell means. Here we see the elasticity of prices with respect to lot size is not quite constant, as previous research has indicated. The downward slope is not quite constant and appears to grow weaker with lot sizes greater than four acres. This suggests that the plattage effect is weaker for very large lot sizes, although this could be due to omitted variables rather than plot size itself. The left panel also shows the estimated functions for the sub-sample that is closer than average to the city center, as the green dashed line, and the sub-sample that is farther than average from downtown, as the dotted yellow line. The nearness of these lines to the full sample line illustrates how important lot size is relative to distance in determining prices. It also suggests that plattage effects are weaker for large lots closer to city centers than away from them.



¹⁰ Using data from Champaign, IL, Colwell ad Sirmans (1978) also documented so-called “plottage” effects, where the price gradient is convex, as opposed to concave, for lot sizes less than 3/8 of an acre. As may be seen in the graph, we did not find such a pattern in the national data.

The right panel, Figure 3B, shows the estimated distance function, measured as the Google Maps-reported driving time from downtown. The solid line and blue circles show the estimated function and cell means, respectively, for the whole sample. The estimated function is steep at first, flattens out slightly after about 20 minutes from downtown, and flattens out more after 40 minutes. The dotted yellow and dashed green lines show the estimated distance functions for the sub-samples with larger and smaller than average lot sizes. These lines suggest that the distant gradient may be exaggerated by the plattage effect, as larger lots tend to be further away. The 'small' and 'large' curves are similarly shaped, and suggest that land at the city center is worth four times than land an hour away. The gradient also appears to be slightly convex, weakening with distance, as standard monocentric city models would suggest.

Table 4 shows regression results confirming the visual evidence in figure 3. Column 1 shows results from a regression of land values on a set of MSA fixed effects; the R^2 of the regression is 0.46. Column 2 shows results from the same regression, adding single variables for log lot size and linear distance from the city center. The elasticity of price per acre with respect to lot size is estimated to be -0.6, implying that a 10 percent increase in plot size is associated with only a 4 percent increase in the price of the plot. This estimate is consistent with other estimates in the literature.

4.2: Using the land rent gradient to estimate land's share of income

The semi-elasticity of price per acre with respect to hours' drive to downtown, β_d , is -1.5, which corresponds to the 75 percent reduction in value from the city center to the urban fringe seen in Figure 3B. Standard urban theories, pioneered by Alonso (1964), Mills (1979) and Muth (1969), imply that the value of this coefficient should equal the ratio of the cost of an hour of commuting to the share of income derived from land. We use a particular formulation related to that in Combes et al. 2012).

Let the indirect utility function the product of income, w , metropolitan quality of life, Q , and a function, $\tilde{v}(d)$, of distance, d , from the central business district (CBD). It is inversely proportional to the index of price of housing $p(d)$, according to the power η . $V(w, p; Q, d) = wQ\tilde{v}(d)[p(d)]^\eta$. In logarithms, that is

$$\ln V = \ln w + \ln Q + \ln \tilde{v}(d) - \eta \ln p(d) \quad (1)$$

Via Roy's identity, it is immediate that η is the household expenditure share on housing. Housing is produced with a constant-returns-to-scale technology. Its cost is determined by the price of land $r(d)$, the price of labor w , and housing productivity A_Y , according to the unit cost function $c_Y(r, w; A_Y) = r^\phi w^{1-\phi} (A_Y)^{-1}$. We assume markets are competitive, and firms make zero profits, so that the price of housing in logarithms is

$$\ln p(d) = \phi \ln r(d) + (1 - \phi) \ln w - \ln A_Y \quad (2)$$

Note the parameter ϕ is the cost share of land in housing. Using these formulae, it is straightforward to show that the land rent gradient follows a negative exponential distribution:

$$\ln r(d) = \ln r(0) - \frac{1}{\phi\eta} [\ln \tilde{v}(0) - \ln \tilde{v}(d)] \quad (3)$$

The term $\ln \tilde{v}(0) - \ln \tilde{v}(d)$ captures how much the quality of location declines with distance to the CBD. Meanwhile, the term $\phi\eta$ captures the share of income that accrues to residential land.

As commuting is solely responsible for changes in location quality within a metro area, and d is measured by hours, then, then the formula can be approximated linearly, to correspond to the regression equation

$$\ln r(d) = \ln r(0) - \frac{1}{\phi\eta} \frac{\tilde{v}'(d)}{\tilde{v}(\bar{d})} d \quad (4)$$

Here $\tilde{v}'(d)/(\tilde{v}(\bar{d}))$ is equal to the marginal cost from an hour of commuting, as expressed as a fraction of total income.

To approximate the marginal cost of an hour of commuting, we parameterize value based on the average cost. About 10 percent of the working day (25 minutes each way) and 5 percent of labor income is spent commuting according to the American Community Survey and Survey of Income and Program Participation. Netting out federal taxes from the time cost and accounting for non-labor income suggests that commuting costs are equal to roughly 9 percent of total income. If the typical two-way commute is indeed 50 minutes and the monetary cost of commuting is proportional to commute time, then an hour of commuting is worth 10.8 percent of total income. Then our coefficient estimate implies that the share of income accruing to land should be equal to $0.108/1.49$, or 7.3 percent. While this is a rough calculation based off of a precariously simplified model, the resulting income share to land is very close to the one produced by Case (2007).

Column 3 adds dummy variables for several common categories of controlled use (see below), namely, industrial, retail, single family, office, and ‘hold for development’, in addition to a dummy for no proposed use. These categories do not partition the data, as a parcel can have multiple proposed uses or no proposed use, and we do not include several proposed uses, such as parking or medical, that are rare in the data. Controlling for proposed use does not meaningfully change the coefficients on lot size or distance from downtown, and increases the explanatory power of the regression only marginally. Therefore, column 3 may support the hypothesis of a unified market for vacant land. Column 4 includes quadratic terms in lot size and distance from downtown, as well as an interaction between the two. The addition of the quadratic terms reduces the magnitude of the coefficient on lot size slightly, while the squared terms yields a concave shape, consistent with some findings Kau and Sirmans (1979). The estimates for the distance from downtown coefficients show that the distance effect is initially very negative, but that it falls off with distance. This could be due to a variety of causes: declining transportation costs; increasing shares of land in housing costs or housing in expenditures, either from substitution or the sorting of high housing demanders to the urban fringe. The interaction term is insignificant.

**TABLE 4: LAND VALUES AND OBSERVABLE SITE CHARACTERISTICS:
METROS ARES, ACREAGE AND DISTANCE WITH INTENDED USE CONTROLS**

	Dependent Variable: Log Price per Acre			
	(1)	(2)	(3)	(4)
Log lot size (acres)	-0.607 (0.003)	-0.593 (0.003)	-0.569 (0.006)	
Log lot size (acres) squared			-0.008 (0.002)	
Driving time from city center (hours)	-1.487 (0.024)	-1.416 (0.023)	-2.837 (0.063)	
Driving time from city center (hours) squared			1.317 (0.058)	
Log lot size (acres) x Driving time from city center (hours)			-0.008 (0.013)	
Number of Observations	57,157	57,157	57,157	57,157
Adjusted R-squared	0.462	0.709	0.726	0.729
CMSA Fixed Effects?	Yes	Yes	Yes	Yes
Controls for Intended Use?	No	No	Yes	Yes

4.3: Intended Use and Land-Use Regulations

In table 5, we examine the roles of property type, intended use, and the regulatory environment in determining land values. The table displays two columns, in both of which we have controlled for cubic polynomials in lot size and driving distance from downtown, interactions between the two up to a quadratic term, and MSA fixed effects. In column 1, we also include a set of dummies for common intended uses in the sample: industrial, retail, single family development, office, and ‘hold for development’. Additionally, we include a dummy for no proposed use. As noted in section 3, a property can have more than one intended use, and the categories considered here are not exhaustive of all intended uses in the sample.

To examine the importance of land-use restrictions we use the Wharton Residential Land Use Regulatory Index of Gyourko et al. (2008) for the county where the parcel is located.

Having no intended use lowers a parcel's predicted value per acre by 21 log points, while an industrial intended use lowers the predicted value by 41 log points. An intended use of single family development is roughly neutral as a predictor of price, while intended uses of office and hold for development predict roughly 5 log points higher prices. An intended use of retail predicts a price 24 log points higher, the strongest positive association of the considered intended uses. A one standard deviation increase in the Wharton Index, a measure of regulatory stringency related to local land use, predicts a nearly 7 log point decrease in price per acre.

In column 2, we examine whether the relationship between price and intended use depends on the regulatory environment by interacting the intended use dummies with the Wharton Index. Including the interaction terms does not substantially change the coefficients on the level terms in the regression. Most of the interaction terms are not statistically significant, but the interaction between an intended use of single family and the Wharton Index is estimated to be positive 5%, so that the net effect of land-use restrictions appears to be zero for residential purposes. The interaction term on an intended use of hold for development is negative 9%, implying that a more stringent regulatory environment is associated with lower prices for such properties. This interaction may reflect greater costs of navigating the zoning process for properties that are not yet developed. The estimated value effects should be treated with caution, as more valuable land may be more prone to being regulating, as McDonald and McMillen (1993) observe in Chicago.

**TABLE 5: LAND VALUES AND OBSERABLE SITE CHARACTERISTICS:
INTENDED USE AND LAND-USE REGULATION**

	Dummy Only (1)	Dummy Level (2a)	Interaction w/WRLURI (2b)
No intended use	-0.206 (0.013)	-0.203 (0.013)	-0.029 (0.020)
Intended use: industrial	-0.411 (0.017)	-0.420 (0.017)	0.042 (0.024)
Intended use: retail	0.243 (0.016)	0.247 (0.016)	-0.026 (0.024)
Intended use: single family	-0.011 (0.015)	-0.039 (0.019)	0.050 (0.022)
Intended use: office	0.042 (0.018)	0.037 (0.018)	0.032 (0.026)
Intended use: hold for development	0.051 (0.013)	0.070 (0.013)	-0.091 (0.020)
WRLURI	-0.067 (0.012)	-0.048 (0.015)	
Number of Observations	51,209	51,209	
Adjusted R-squared	0.649654	0.649896	
CMSA Fixed Effects?	Yes	Yes	

Includes a third-order acreage effect. WRLURI is the Wharton Residential Land Use Regulatory Index

4.3: Overall Co-variation of Land Values with Space, Time, and other Observables

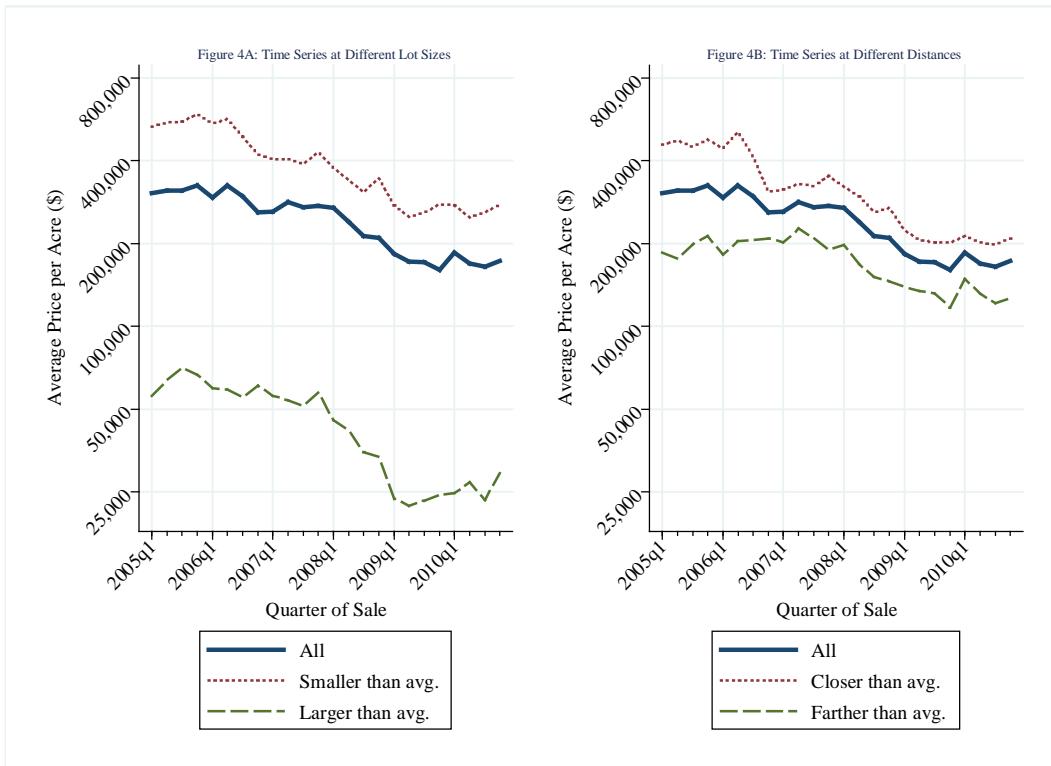
Before examining the determinants of land values within and across metropolitan areas in detail, we present a simple variance decomposition of log price per acre. The results are displayed in table 6. In the first column, we regress prices on a cubic polynomial in lot size and a set of intended use dummies. The adjusted R^2 of the regression is 0.57, implying that these variables alone predict 57% of the variation in log prices in our sample. In column 2 we add a cubic polynomial in driving distance from downtown and a set of MSA fixed effects. Controlling for space in this manner increases the adjusted R^2 of the regression noticeably to 0.73. In column 3, we control for time rather than space, by adding a set of quarter of sale dummies to the controls in column 1. The adjusted R^2 , 0.60, is only modestly higher than in column 1, and substantially lower than in column 2. We take this as evidence that space is more important than time in predicting urban land values. In column 4, we include both the space and time controls from columns 2 and 3. The explanatory power of the regression is slightly higher than in column 2, with an adjusted R^2 of 0.743. Finally, in column 5, we examine whether a flexibly estimated model with several interactions can improve substantially on the predictive power of the model in column 4. Accordingly, we add controls for MSA-lot size interactions, MSA-driving time interactions, quarter of sale-lot size interactions, and quarter of sale-driving time interactions. Despite the flexibility of this model in predicting land values, the predictive power of this specification is only modestly higher than in column 4, with an adjusted R^2 of 0.76. Therefore, it appears that a relatively small number of variables can predict most of the variation in land values in our sample. The relatively parsimonious model of column 4 accounts for nearly three-quarters of this variation for instance. Furthermore, lot size and spatial location have a stronger association with land values than does time of sale.

TABLE 6: ACCOUNTING FOR VARIANCE IN LAND VALUES

	Dependent Variable: Log Price per Acre				
	(1)	(2)	(3)	(4)	(5)
Lot size polynomial	Yes	Yes	Yes	Yes	Yes
Intended use dummies	Yes	Yes	Yes	Yes	Yes
Driving time polynomial		Yes		Yes	Yes
CMSA Fixed Effects		Yes		Yes	Yes
Quarter of sale dummies			Yes	Yes	Yes
CMSA-lot size interactions					Yes
CMSA-driving time interactions					Yes
Quarter of sale-lot size interactions					Yes
Quarter of sale-driving time interactions					Yes
Adjusted R-squared	0.569	0.730	0.596	0.743	0.761

These results are quite interesting because a number of studies have documented substantial time series variation in United States land values. Davis and Palumbo (2007) and Davis and Heathcote (2007) document a large increase in the price of residential land in the United States in the years prior to 2005 using residual methods described in the introduction. Using the same source data as in this paper, Nichols et al. (2013) document that land transaction prices peaked in 2006-2007 in the cities that they study, before falling an estimated 50% from their peak by mid-

2011. In figure 4, we document a similar decline in our dataset, which spans the period 2005 through 2010 and considers a larger sample of MSAs. The figure displays the fitted time trend for land values per acre for the second quarter of 2005 through the fourth quarter of 2010. The time trend for the entire sample, displayed as the solid blue line both in panels A and B, shows that the land values peaked in the second quarter of 2006 at a geometric average value of \$326,146, and fell to a low of \$161,168 in the fourth quarter of 2009. One natural question is whether this trend complicates inference using this sample. We interpret figure 4 as indicating that inference is unlikely to be meaningfully complicated by the time trend in land values over the sample period. Panel A shows the fitted time trend for larger than average parcels, as the green dashed line, and smaller than average parcels, as the dotted red line. Panel A suggests both that the time trends for small and large lots were quite similar over the sample period, and that the time series variation in that period was small relative to the variation associated with different lot sizes. Similarly, panel B displays the fitted time trend for parcels closer (red dotted line) and farther (dashed green line) than average to downtown. Again, the time trends for the two types of parcels are broadly parallel. Visually, the log price variation between close and far parcels is of roughly the same magnitude as the variation over time. These results are consistent with the decomposition of variance in table 6, in which including quarter of sale dummies increases the explanatory power of the regression only modestly. Overall, we interpret the evidence as indicating that time variation in land values over the sample period is unlikely to pose problems for the other results in the paper.

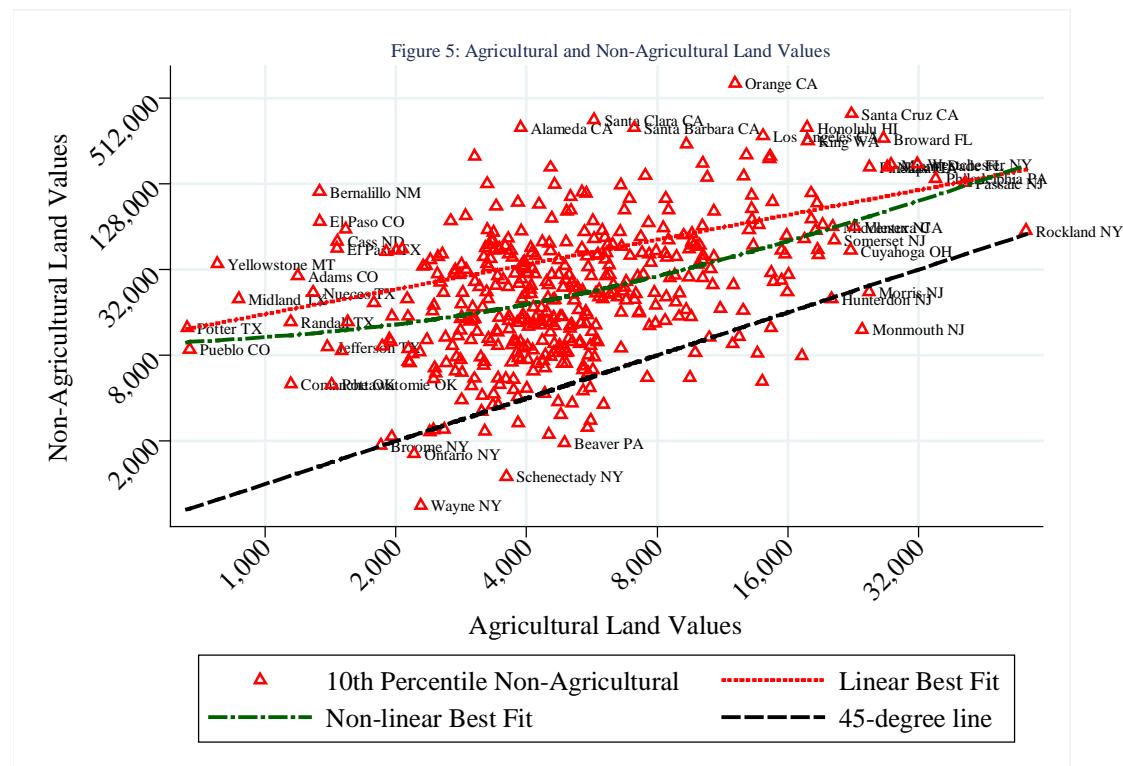


5: Inter-Metropolitan Variation in Urban Land Values

We now examine variation in land values across metropolitan areas. Two topics of particular interest are the degree of unification between the urban and agricultural land markets, and the role of metropolitan attributes in determining land values.

5.1: The Urban Fringe and Agricultural Land

Standard urban theory suggests that in the presence of a unified land market, the value of land on the urban fringe, denoted here as d , should equal the land's value in agricultural use. Because we cannot identify the exact location of the urban fringe, we instead take land values at the 10th percentile for each county as a measure of fringe land values, $r(d)$. Figure 5 plots these 10th-percentile land values on the y-axis against average agricultural land values on the x-axis.¹¹ The agricultural land values are from the USDA Economic Research Service for the year 2007. Table 7 presents regression results for the relationship between urban and agricultural land values. Column 1 reports that the elasticity of the 10th percentile of urban land values with respect to agricultural values is 0.66, significantly below one, and the regression constant is significantly above one. Assuming the 10th percentile of urban land values can be taken to represent the urban fringe, these values suggest a deviation from a perfectly unified land market. The variation in urban land values at the fringe is greater than the variation in agricultural values across counties, as visual inspection of figure 5 illustrates.



¹¹ The figure and lines of best fit omit counties with the lowest and highest one percent of agricultural land values.

However, deviations between urban and agricultural land values may arise if there are considerable costs to converting agricultural land to urban land. Most importantly, agricultural land may need to be cleared and surveyed before being used for urban purposes. These costs are likely to vary little across cities, and are likely to be large relative to agricultural land values where the latter are small. To examine whether conversion costs drive a wedge between agricultural and urban land values, we estimate the non-linear equation:

$$\ln r(\underline{d}_i) = \beta_a \ln (c + \text{ag-value}_i) + \epsilon_i \quad (5)$$

where c represents the cost of converting agricultural land to urban use. The line of best fit from the non-linear regression is displayed as the dashed green line in figure 5. The results in column 2 of Table 7 imply an estimate of c of \$3,972 per acre. This number is roughly equal to the median agricultural land value of \$3,979. Thus, for the typical city, an acre of land at the urban fringe (i.e. the 10th percentile) appears to derive roughly half of its value from improvements. This result is surprisingly consistent with Mills' (1998) "guess" that land at the urban fringe derives roughly 50 percent of its value from improvements. The slope coefficient β_a in the non-linear regression increases to 1.10, which is statistically larger than but much closer to one.

TABLE 7: AGRICULTURAL AND NON-AGRICULTURAL LAND VALUES

Dependent Variable: 10th Percentile Log Price per Acre			
	Linear Specification (1)	Non-linear Specification (2)	Non-linear Specification (3)
Log agricultural land value per acre	0.662 (0.087)		
Log of combined agricultural land and implied conversion costs per		1.095 (0.017)	1.430 (0.209)
Implied conversion costs per acre		3,792 (1,097)	7,216 (2,779)
Constant	4.335 (0.744)		-3.538 (2.211)
Number of Observations	442	439	439
Adjusted R-squared	0.173		

County-level observations. Standard errors clustered at MSA level.

Note that this equation does not have a constant term. A constant might be justified theoretically if there is a conversion cost proportional to the agricultural land value, which would produce a positive constant, or if there is a conversion cost to agriculture, which would produce a negative constant. Column 3 of table 7 displays the non-linear regression with a constant term. The

constant is estimated to be negative but statistically indistinguishable from zero, with a level value of only -\$35. The standard errors on the other coefficients also increase substantially with the addition of the constant term. Overall, we believe the specification in column 2, which does not allow for conversion costs to vary with agricultural land values, is most accurate.

We conclude that although there is some evidence of frictions that impede a perfectly unified market for land, the land market appears to be much more unified across uses after accounting for conversion costs. Furthermore, it would appear that these conversion costs account for a substantial fraction of observed land values at the urban fringe. While it would be worthwhile to consider reasons to explain deviations around the regression line, we move ahead now to connect values at the urban fringe with those at the urban core.¹²

5.2: Central Land Values, Population, Size, and Amenities

This section considers the value of land at the urban center. According to the standard monocentric city model, the value of this land should be proportional to the per-capita costs of urbanization, typically measured through commuting. The elasticity of land values with respect to metropolitan population can then be taken as a measure of the diseconomies of scale experienced in cities. Recent work by Combes et al. (2012) has tried to estimate this elasticity using a rather different set of French land values. Here we use a modified form of their theory that incorporates fringe land values, and link our estimates to the estimated rent gradients estimated above, providing even stronger support for the theory.¹³

We follow their theory by allowing the elasticity of commuting costs with respect to total income $\tilde{v}'(d)/\tilde{v}(d) \approx \tau$ to be constant, deviating from the former linear specification. We allow for observable agricultural land prices $r(\underline{d})$, so that our method is much more direct. In this case, land values at the city center are given by

$$\ln r(0) = \text{constant} + \frac{\tau}{\phi\eta} \ln \underline{d} + \ln r(\underline{d}) \quad (6)$$

We do not typically observe a single well-defined radius for a metro area. It is not clear exactly where a city ends, and where agricultural uses begin. Nor is it guaranteed to be uniform and different directions. Thus, it is more practical and interesting to use population N as a regressor. Avoiding complications, assume that area of the city is given by $S = \pi\Theta\underline{d}^2$. The parameter Θ measures the angle in radians that land is provided around the CBD, or “arc of expansion.” In that case, it is possible to show that the downtown land values are related to the population size N through the following formula

$$\ln r(0) = \text{constant} + \frac{\tau}{2\phi\eta} \ln N + \ln r(\underline{d}) - \ln \Theta \quad (7)$$

¹² For more on land values and uses on the urban fringe, see Chicoine (1981) and McMillen (1989).

¹³ Our city model also allows all city residents to inhabit a city, while the Combes et al. model makes an ad hoc adjustment of city residents that live outside of the urban fringe whose “first principles is beyond the scope of [their] paper.”

If this model is correct, then $\tau/2$ measures the diseconomies of scale that accrue from having a larger city. Furthermore, according the theory the coefficient on log population should be related to that on estimated driving time, β_d . Namely, it should be minus half of that coefficient times the average commute time in hours, or $-(-1.4)/2(5/6) = 0.61$. In addition, the urban land value should increase proportionally with agricultural values and inversely with the number of radians available.

In addition, we consider the role of natural amenities. Typically, urban economists believe that land values should rise with the value of an urban amenity, for example, living by the coast. However, in a local market, the distribution of amenities should also matter. Because land values arise generally from differences in amenities, places with a greater variance of amenities may have greater land values overall. More concretely, if everyone can have a house on a hill, land values may be lower than if only some people can have a house on a hill. If the supply of hills is limited, then the price paid for the hill may be higher.

In the model above, uniform levels of amenities, seen in Q , are already accounted for in the population level, and thus should not appear in the above equation. Variation in amenities, however, changes the local value $\tilde{v}(d)$. Just as transportation costs raise land values by giving relative advantages in some areas relative to others, so do natural amenities.

We use two methods of estimating land values at the urban core. In the first, we simply take the 90th percentile of log price per acre within the MSA. In the second, we estimate the regression

$$\ln r_{ijt} = \alpha_j + \delta_t + \beta_A A_{ijt} + \beta_d d_{ijt} + \beta_X X_{ijt} + \epsilon_{ijt} \quad (8)$$

where r_{ijt} is the price per acre of parcel i in metro area j sold in quarter t , α_j is a set of MSA fixed effects, δ_t is a set of quarter of sale fixed effects, A_{ijt} is a cubic polynomial in log lot size, d_{ijt} is a cubic polynomial in driving time from the city center, X_{ijt} is a set of intended use dummies, and ϵ_{ijt} is a stochastic error. We take the average residual from this regression plus the estimated MSA fixed effect for each MSA as our second index of urban land values at the urban core.

The two indices serve as the dependent variables in the regressions on table 8. The first index requires less of a stand in us determining what the value of central land is. The second index is more appropriate for looking at the impact at variation in amenities, since it is more based on representative land values.

In column 1, we regress the 90th percentile of MSA land values on log MSA population. Land values are increasing in population, with an estimated elasticity of 0.65. In column 2, we also control for log agricultural land values, which determine a city's land area in the standard monocentric city model with endogenous city size. We adjust the USDA's estimated agricultural land values by adding the estimated conversion costs from column 2 of table 7. In turn, city land area helps to predict land values at the urban core. Higher agricultural land values are associated with higher land values at the core, with an elasticity of 0.83. This value is not statistically different from 1 at standard significance levels. Inclusion of agricultural land values reduces the

elasticity of core land values with respect to population to 0.50, which is rather close to the value predicted from the distance from CBD regression of 0.61. These regressions lend credence to the possibility that the elasticity of urban costs, $\frac{\tau}{2}$, with respect to population growth are somewhere around 3.7 to 4.5 percent of gross income.

In column 3, we add a measure for the fraction of the MSA's potential land area that cannot be developed because of a coastline or major body of water from Stephen Malpezzi. The elasticity of core land values with respect to the fraction of developable land is -0.58. This is consistent with the prediction that the fraction of developable land has a negative elasticity, although the elasticity is a bit shy of one.

In column 4 of table 8, we drop the fraction of area lost to coastline and water and add the average slope of land and its standard deviation at the PUMA level from Albouy et al. (2012), and average slope of land within a 50 kilometer radius of the city center.¹⁴ The estimates indicate that variation in this amenity within metro areas is more important in explaining land values as the average level of the amenity across metro areas, as predicted.

In column 5, we add variables for average commute times and their within-MSA standard deviations to see if they account for some of the association between MSA population and land values at the urban core. Ultimately these regressors do little to improve the explanatory power of the regression, and neither is statistically significant. In column 6, we estimate the regression from column 5 without the controls for MSA population. The explanatory power of the regression falls noticeably, but the standard deviation of land slope and the standard deviation of commuting times both predict significantly higher land values. We interpret this evidence as being consistent with the standard monocentric city model of land values, in which variation in amenities, not simply their levels, gives rise to land values.

Finally, in columns 7 and 8 we use the land value at the urban core index calculated as we describe above as the dependent variable. Column 7 mimics column 2; the results are broadly similar, although the coefficients both on MSA population and on adjusted agricultural land values are smaller when using the land value index rather than the 90th percentile land values. Column 8 mimics column 4. The results are again qualitatively similar, but the coefficient on the standard deviation of the average slope of land is now positive and statistically significant. Again, we interpret this result as reinforcing the classical insight from Ricardo (1817) that land values ultimately stem from differences in the quality of locations.

¹⁴The measure is slightly different than that measured by Saiz (2010), which is the fraction of land over water or with a slope greater than 15 percent.

TABLE 8: CENTRAL LAND VALUES, POPULATION, AREA, AND AMENITIES

Dependent Variable:	90th Percentile MSA Land Values						Land Values at MSA Center Index	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Log MSA population	0.653 (0.057)	0.496 (0.064)	0.456 (0.061)	0.483 (0.058)	0.468 (0.063)		0.379 (0.038)	0.359 (0.036)
Log agricultural land values plus conversion costs		0.829 (0.188)	0.838 (0.193)	0.742 (0.164)	0.777 (0.173)	1.264 (0.216)	0.777 (0.134)	0.678 (0.117)
Log (1 - Degrees lost to coast or major water/360)			-0.575 (0.180)					
Average slope of land				0.000 (0.039)	0.000 (0.038)	-0.043 (0.042)		0.018 (0.022)
Standard deviation of slope of land					0.084 (0.062)	0.070 (0.061)	0.167 (0.068)	0.121 (0.041)
Commuting time						-0.279 (0.368)	0.566 (0.441)	
Standard deviation of commuting time						0.824 (0.707)	2.961 (0.770)	
Number of Observations	239	237	219	235	235	235	237	235
Adjusted R-squared	0.431	0.480	0.548	0.487	0.491	0.389	0.497	0.519

All regressions include a constant term. Dependent variable is MSA average residual from regression of log price per acre on a cubic polynomial in lot size, a cubic polynomial in driving time from downtown, a set of quarter of sale dummies, and a set of intended uses. Degrees lost to coast or major water from Stephen Malpezzi.

6: Conclusion

Despite the many problems that are likely to arise when measuring land values, it appears that the market transactions are consistent with several predictions of neoclassical urban theory. The overall distribution of land values is convex, and may be roughly approximated by a log-normal distribution, with a wide standard deviation that widens slightly in higher value cities. The association of lot size with price per acre is strong and largely consistent with previous studies. Even though many commutes do not occur between suburbs and central cities, Land values still fall with travel time from downtown as the monocentric city would predict. Furthermore, under stricter assumptions, the estimate implies a share of income for residential land that is entirely plausible at around 7 percent.

We fully believe that variation in land values over time as examined by Oliner et al. (2013) and others is extremely important in understanding the urban as well as macro economy. Yet, even over the housing boom and bust cycle, we still found that space is a much greater determinant of land values than time is.

At the metropolitan level, we found compelling evidence that value of urban land at the fringe is close to that of agricultural land after accounting for conversion costs. As suggested by Mills decades ago, these costs appear to be about half of the value of land on the fringe, albeit a much smaller fraction of more central land values.

When considering urban land values at the city center, we find that they increase at roughly the square root of urban population, which is consistent with various forms of the monocentric city

model, including the one examined here. In the model we use, the elasticity of urban costs as a fraction of total income appears to be around 4 percent of income. We also find interesting evidence that within metro areas, land values rise not only with the average level of amenities but also with their variance. While there are certainly many imperfections with data and theory used in this analysis, the results of this paper suggest that the prices of observable land transactions do exhibit many regularities consistent with economic theory and intuition.

References

- Albouy, David and Gabriel Ehrlich (2012) “Metropolitan Land Values and Housing Productivity.” NBER Working Paper No. 18110. Cambridge, MA.
- Albouy, David, Walter Graf, Hendrik Wolff, and Ryan Kellogg (2012) “Extreme Temperature, Climate Change, and American Quality of Life.” Unpublished Manuscript.
- Alonso, W (1964) *Location and Land Use*. Cambridge: Harvard University Press.
- Arnott, Richard and Joseph Stiglitz (1979), “Land Rents, Local Expenditures, and Optimal City Size”, *Quarterly Journal of Economics*, 93, pp. 471-500.
- Karl E. Case (2007) “The Value of Land in the United States: 1975-2006” in Gregory Ingram and Yu-Hong Hung, eds. *Land Policies and Their Outcomes*. Cambridge, MA: Lincoln Institute for Land Policy. pp. 127-147.
- Chicoine, David L. (1981) “Farmland Values at the Urban Fringe: An Analysis of Sale Price.” *Land Economics*. 57(3), pp.
- Colwell, Peter and Henry Munneke (1997) “The Structure of Urban Land Prices.” *Journal of Urban Economics*, 41, pp. 321-336.
- Colwell, Peter and C.F. Sirmans, (1978) “Area, Time, Centrality, and the Value of Urban Land.” *Land Economics* 54(4), pp. 514-519.
- Colwell, Peter and C.F. Sirmans (1980) “Non-Linear Urban Land Prices.” *Urban Geography* 1, pp. 353-362.
- Colwell, Peter and C.F. Sirmans (1993) “A Comment on Zoning, Returns to Scale, and the Value of Undeveloped Land.” *The Review of Economics and Statistics*, 75, pp. 783-786.
- Combes, Phillippe, Gilles Duranton and Laurent Gobillon (2012) “The Costs of Agglomeration: Land Prices in French Cities.” IZA Discussion Paper 7027.
- Davis, Morris A., and Jonathan Heathcote. “The price and quantity of residential land in the United States.” *Journal of Monetary Economics* 54.8 (2007): 2595-2620.

Davis, Morris and Michael Palumbo (2007) "The Price of Residential Land in Large U.S. Cities." *Journal of Urban Economics*, 63, pp. 352-384.

Fisher, Jeff, David Geltner and Henry Pollakowski (2007) "A Quarterly Transactions-based Index of Institutional Real Estate Investment Performance and Movements in Supply and Demand." *Journal of Real Estate Finance and Economics*, 34, pp. 5-33.

George, Henry (1879) *Progress and Poverty: An Inquiry into the Cause of Industrial Depressions and of Increase of Want and Increase of Wealth: The Remedy*. Robert Schalkenbach Foundation

Gyourko, Joseph, Albert Saiz, and Anita Summers (2008) "New Measure of the Local Regulatory Environment for Housing Markets: The Wharton Residential Land Use Regulatory Index." *Urban Studies*, 45, pp. 693-729.

Haughwout, Andrew, James Orr, and David Bedoll (2008) "The Price of Land in the New York Metropolitan Area." Federal Reserve Bank of New York Current Issues in Economics and Finance, April/May 2008.

Hochman, Oded (1981) "Land Rents, Optimal Taxation, and Local Fiscal Independence in an An Economy with Local Public Goods." *Journal of Public Economics* 15: 59-85.

Hoyt, H. (1933) *One Hundred Years of Land Value in Chicago*. Chicago: University of Chicago Press.

Ihlanfeldt, Keith R. (2007) "The Effect of Land Use Regulation on Housing and Land Prices." *Journal of Urban Economics*, 61, pp. 420-435.

Kau, James B. and C. F. Sirmasn (1979) "Urban Land Value Functions and the Price Elasticity of Demand for Houisng." *Journal of Urban Economics*, 6(1), pp. 112-121.

Kok, Nils, Paavo Monkkonen and John Quigley (2010) "Economic Geography, Jobs, and Regulations: The Value of Land and Housing." Working Paper No. W10-005. University of California.

Mayer, Christopher J. and C. Tsuriel Somerville (2000) "Land Use Regulation and New Construction." *Regional Science and Urban Economics* 30, pp. 639-662.

McMillen, Daniel P. (1989) "An Empirical Model of Urban Fringe Land Use." *Land Economics* 65(2), pp. 138-145.

McMillen, Daniel P. and John F. McDonald (1993) "Could Zoning Have Increased Land Values in Chicago?" *Journal of Urban Economics* 33, pp. 167-188.

Mills, Edwin S. (1967) "An Aggregative Model of Resource Allocation in a Metropolitan Area" *American Economic Review* 57, pp. 197-210

Mills, Edwin (1998) "The Economic Consequences of a Land Tax." in Dick Netzer, ed. *Land Value Taxation: Can It and Will It Work?* Cambridge, MA: Lincoln Institute for Land Policy. pp. 31-48.

Munneke, Henry and Barrett Slade (2000) "An Empirical Study of Sample-Selection Bias in Indices of Commercial Real Estate." *Journal of Real Estate Finance and Economics*, 21, pp. 45-64.

Munneke, Henry and Barrett Slade (2001) "A Metropolitan Transaction-Based Commercial Price Index: A Time-Varying Parameter Approach." *Real Estate Economics*, 29, pp. 55-84.

Muth, Richard F.. (1969) *Cities and Housing: The Spatial Pattern of Urban Residential Land Use*. Chicago: University of Chicago Press.

Nichols, Joseph, Stephen Oliner and Michael Mulhall (2013) "Commercial and Residential Land Prices Across the United States." *Journal of Urban Economics* 73, pp. 57-76

Olcott's Land Value Blue Book for Chicago. (1910-1970) Chicago: Olcott Co.

Quigley, John and Stephen Raphael (2005) "Regulation and the High Cost of Housing in California." *American Economic Review*. 95, pp.323-329.

Quigley, John and Larry Rosenthal (2005) "The Effects of Land Use Regulation on the Price of Housing: What Do We Know? What Can We Learn?" *Cityscape: A Journal of Policy Development and Research*, 8, pp. 69-137.

Ricardo, David (1817) *Principles of Political Economy and Taxation*.

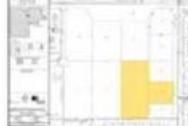
Saiz, Albert (2010) "The Geographic Determinants of Housing Supply." *Quarterly Journal of Economics*, 125, pp. 1253-1296.

Thorsnes, Paul (1997) "Consistent Estimates of the Elasticity of Substitution between Land and Non-Land Inputs in the Production of Housing." *Journal of Urban Economics*, 42, pp. 98-108.

Von Thunen, Johann. *The Isolated State in Relation to Agriculture and Political Economy*. New York. Palgrave.

Appendix A

Figure A1: Sample Brochure from CoStar COMPs Database

31	SW 75th Ave & Deer Valley Rd	SOLD
Glendale, AZ 85308	Maricopa County	
Sale Date: 12/06/2008 Sale Price: \$2,500,000 - Affidavit \$/AC Land Gross: \$381,679.39 (\$8.76/SF)	Land Area: 6.55 AC (285,318 SF) Lot Dimensions: - Proposed Use: -	
Parcel No: 200-19-004J, 200-19-014		
Comp ID: 1624780 Research Status: Research Complete	Sale Conditions: -	
32	SW 75th Ave & Roosevelt St	SOLD
Phoenix, AZ 85043	Maricopa County	
Sale Date: 06/11/2008 Sale Price: \$31,981,393 - Confirmed \$/AC Land Gross: \$428,623.41 (\$9.84/SF)	Land Area: 74.61 AC (3,250,195 SF) Lot Dimensions: - Proposed Use: -	
Parcel No: 102-43-002G		
Comp ID: 1543026 Research Status: Research Complete	Sale Conditions: -	
33	NW 75th Ave & Rose Garden Ln - 68 Lots, The Reserve at Eagle Heights	SOLD
Glendale, AZ 85308	Maricopa County	
Sale Date: 09/30/2008 Sale Price: \$6,650,000 - Affidavit \$/AC Land Gross: \$356,625.73 (\$8.19/SF)	Land Area: 18.65 AC (812,263 SF) Lot Dimensions: Irregular Proposed Use: Hold for Development [Partial List]	
Parcel No: 200-19-274, 200-19-275, 200-19-276, 200-19-277, 200-19-278, 200-19-279, 200-19-280, 200-19-281, 200-19-285, 200-19-286, 200-19-287, 200-19-288 [Partial List]		
Comp ID: 1595885 Research Status: Research Complete	Sale Conditions: -	
34	NE 82nd St & Princess Dr - I-1/pcd Zoned Acreage	SOLD
Scottsdale, AZ 85255	Maricopa County	
Sale Date: 08/08/2008 Sale Price: \$1,612,650 - Confirmed \$/AC Land Gross: \$1,490,443.62 (\$34.22/SF)	Land Area: 1.08 AC (47,132 SF) Lot Dimensions: Irregular Proposed Use: Office	
Parcel No: 215-07-219		
Comp ID: 1572119 Research Status: Research Complete	Sale Conditions: -	
35	SE 83rd Ave & Thomas Rd - Hispanic Conference Center	SOLD
Phoenix, AZ 85035	Maricopa County	
Sale Date: 11/26/2008 (153 days on mkt) Sale Price: \$2,103,816 - Confirmed \$/AC Land Gross: \$111,077.93 (\$2.55/SF)	Land Area: 18.94 AC (825,026 SF) Lot Dimensions: - Proposed Use: Commercial	
Parcel No: 102-38-006		
Comp ID: 1619982 Research Status: Research Complete	Sale Conditions: -	
36	10610 N 84th St	ESCROW
Scottsdale, AZ 85260	Maricopa County	
Asking Price: - \$/AC Land Gross: - Days on Market: 41 Sale Status: Escrow	Sale Type: Owner/User Land Area: 1.00 AC (43,429 SF) Lot Dimensions: - Proposed Use: -	
Parcel No: 202-22-009Q Sales Condition: -		

