FHFA STAFF WORKING PAPER SERIES

Land Valuation using Public Records and Kriging: Implications for Land versus Property Taxation in Cities

William D. Larson
and
Jessica Shui

This Version: April, 2021

Original: April, 2020

Working Paper

FEDERAL HOUSING FINANCE AGENCY
Division of Research and Statistics
400 7th Street SW
Washington, DC 20219, USA

Please address correspondence to William Larson (william.larson@fhfa.gov).

Land Valuation using Public Records and Kriging: Implications for Land versus Property Taxation in Cities

William D. Larson, Jessica Shui
FHFA Staff Working Paper
April, 2021

Abstract

We construct land values for each parcel in Maricopa County (Phoenix), Arizona, from 2000 through 2018 using a novel administrative dataset containing the universe of land sales and parcel records in the county. We then compare residential land values constructed using two classes of source data, vacant land sales and land under existing structures. Between 2012 and 2018, estimated land values for developed parcels are, on average, 14% higher when estimated using vacant land due to plattage effects and other unobserved factors. Growth rates are similar, facilitating the use of vacant land price indices to trace valuations over time from an accurate base year valuation. Dynamics between prices of Maricopa County land and housing suggest hypothetical land value tax revenues are more pro-cyclical than property tax revenues, with βs with respect to national house prices of 3.3 and 2.3, respectively. By 2018, houses had recovered 96% of pre-crisis (2007) values, but land had only recovered 66%. These findings demonstrate a source of risk of dependence on public revenues from land value taxes versus a base-period revenue-neutral property tax.

Keywords: land prices · price gradient · land value taxation · price dynamics

JEL Classification: R14, R21, R32

William D. Larson
Federal Housing Finance Agency
Division of Research and Statistics
400 7th Street SW
Washington, DC 20219, USA
william.larson@fhfa.gov

Jessica Shui
Federal Housing Finance Agency
Division of Research and Statistics
400 7th Street SW
Washington, DC 20219, USA
jessica.shui@fhfa.gov

1 Introduction

Land value taxation has long been held as a form of revenue generation superior to a general property tax.[1] A property tax, which implicitly taxes structures and land at the same rate, is thought to cause an inefficiently low provision of structures relative to a revenue-neutral tax on just the underlying land. With structure density affecting housing affordability, traffic congestion, labor supply, and other primary economic and social concerns, it is not difficult to understand why generations of urban economists and policymakers have held ambitions for cities to implement land value taxation regimes.

If land value taxation is almost universally held to be so much better for society compared to a general property tax, then why has it been so rarely implemented?[2] One major reason is the difficulty measuring land's value. Despite the assurances of George (1879),[3] the value of land underneath structures is incredibly difficult to estimate in a systematic manner, and until recently, has been unavailable over a broad cross section of cities in different time periods (Davis et al. (2021), "DLOS").[4]

Instead of estimating the value of land under structures, researchers and appraisers have sought to exploit variation in the price of vacant lots or sales of existing homes where the structure is worthless (e.g. teardowns or sales following severe damage to the structure) to

[1]See details in George (1879), Anas (2003), Anas (2015), Cho et al. (2011), Duke and Gao (2017), Dye and England (2010), Fischel (2013), Gallagher et al. (2013), Groves (2011), Lichfield and Connellan (1997), Maxwell and Vigor (2005), McCluskey and Franzsen (2001), Oates and Schwab (1997), Plummer (2010a), Plummer (2010b), Wallin and Zabel (2011), and Yang (2014).

[2]There are some notable examples of land taxation. For example, the United Kingdom has had four failed attempts implementing forms of land taxation in the last century; New Zealand replaced its land value taxation with property tax in the 90s due to political pressure and practical concerns arising from the complexity of land value capture (Almy (2013), Ingram et al. (2012), Beale et al. (2016), Biitir (2019)). Land value taxation also has a long history in the United States. However, as of today, only a few places (mostly cities in Pennsylvania) implemented a two-rate or split-rate tax scheme on land and structure while the vast majority of local governments use a combination of specific policies (e.g. BID tax) and a single rate property tax system where the tax burden falls evenly on the land and the structure on it. Several countries in Latin America and Africa (e.g. Ghana, Namibia and the rural land tax in Brazil) have adopted land taxation. In Asia, Singapore and Hong Kong serve as examples.

[3]"But, as a matter of fact, the value of land can always be readily distinguishable from the value of improvements." (George, 1879, p.421)

[4]There is a long literature that attempts to estimate the value of the land that is underneath structures. There are two main strands. The first is the hedonic literature, including notable contributions from Diewert et al. (2015), Kuminoff and Pope (2013), and Bostic et al. (2007a). These studies are useful experimental approaches that have been able address topics such as the option value of (re)development and fluctuations in the implicit price of land versus structures. The second is the residual method literature, including Davis et al. (2017). The work of Davis et al. (2021) falls in this latter line of inquiry.

estimate the value of land. The major issue with this approach is that these relatively direct land sales are typically infrequent, non-uniform across space, and the lots are often different than nearby lots where structures already exist. These factors complicate attempts to infer the value of land underneath structures using vacant land values.[5]

The first of two questions we address in this paper concerns this measurement issue: is vacant land a suitable proxy for land underneath structures, and therefore potentially usable for land assessment purposes? Using a new administrative dataset on vacant land transactions in Maricopa County, Arizona (a portion of the Phoenix-Mesa-Scottsdale MSA), we estimate the value of each 2007-vintage parcel in the county for every year between 2000 and 2018. To estimate these land values, we implement a number of spatial methods and determine kriging to be the most accurate. We then compare these land values estimated using vacant land transactions to those estimated in 2012 through 2018 by DLOS (2021), who estimate the value of land underneath structures.

There are three key findings related to Maricopa County. First, vacant land sells for a premium on a per-acre basis relative to land underneath structures, holding constant observable lot attributes, suggesting unobservable differences between vacant and built-up lots. Second, while tract-level estimates of land value are different when measured using vacant land versus land under structures, they are highly correlated in levels. However, due to noisiness in the series, growth rate associations are weak. Third, aggregation rectifies the tract-level noise issue, and the county-level growth rate is similar, facilitating the use of one measure as an instrument or proxy for the other. To construct an accurate time series, it may be possible to use an accurate valuation of land underneath structures for a base year, then project the time path using vacant land transactions.

The second question we address is the cyclicality of a revenue-neutral land value taxation regime. We believe there may be another reason land value taxation has not caught on among local governments. Because housing structures are infrequently renovated and construction costs change relatively slowly from year to year, rapid change in the value of housing typically occurs when the underlying land is appreciating or depreciating (see Bostic et al. (2007b), Davis and Heathcote (2007), Davis and Palumbo (2008), and Davis et al. (2017),

[5]This line of research includes Haughwout et al. (2008), Combes et al. (2012), Turner et al. (2014), Nichols et al. (2013), Albouy and Ehrlich (2015), and Albouy et al. (2018).

for example). If home values are more stable than land values over a housing cycle, then as a revenue source, land taxes may be more volatile and unpredictable than property taxes.[6]

Using the value of vacant land as a proxy for the value of land under structures, we estimate land and property tax revenue indices between 2000 and 2018. This allows us to trace the rise and fall of the revenue generated by two hypothetical tax regimes across a full housing cycle. The indices are constructed by fixing the index values and the attributes of the housing stock to a base year. Fluctuations in the estimated prices for these constant-quality parcels, when aggregated to the county level, gives the change in the revenue generated by the particular tax regime.[7] We find that land values (and implied land value taxes) in Maricopa county face a much higher degree of pro-cyclicality than property prices (and taxes). Land value tax revenues would have fallen 17% further than property tax revenues in the Great Recession, and would still be less than 70% of property tax revenues in 2018.

The remainder of this paper is structured as follows. First, we provide some background on land valuation in Section 2, along with a description of our setting, Maricopa County between 2000 and 2018. In Section 3, we describe and implement a variety of spatial imputation methods, including kriging, to estimate the value of each parcel in Maricopa County for each year in the sample. Armed with these estimates, Section 4 focuses on validation by comparing these land values to alternative land value and house price measures and both the tract and county level. The cyclicality of a stylized land value tax regime is explored in Section 5. Section 6 concludes.

2 Background and Setting

There are two major difficulties with estimating the value of land. The first is that land is typically purchased in a bundle along with a structure. The second is that not all parcels are transacted each period. Both of these issues have efficiency implications for estimation, with error rising due to either a fall in the number of observed transactions or measurement

[6] There is a line of research that suggests the "land residual" framework may overstate land appreciation, for instance, (Clapp et al., 2020). However, we show a strong correlation between appreciation of land using the residual method and the vacant land transaction method, suggesting any bias is likely small.

[7] Lutz et al. (2011) estimate that negative house price appreciation did not have a major effect on tax revenue because localities did not reduce their assessments of properties in a timely manner. Doerner (2012) also pointed out that assessment ratio, assessed value over market price, seldom reaches unity. We abstract from this sort of assessment behavior and instead assume assessors immediately tax landowners based on their best current estimate of the value.

error in separating the value of the land from the value of structure. The main issue in terms of bias, however, is the complex selection processes correlated with price: vacant land may be higher or lower in quality than built-up land, and parcels that are transacted may be of higher or lower quality than those that are not. Failure to account for these factors could easily lead to misleading estimates of parcel-level land values.

Databases available to property tax assessors typically do not include good estimates of the value of land underneath structures, with average assessed (and estimated "market") land values substantially below market values (DLOS, 2021).[8] But records of vacant land transactions are readily available, and these are accurate insofar as they reflect actual market transaction values. The first question posed in this research is whether vacant land transactions produce similar estimates of the value of land underneath structures.

To answer this question, we use a novel dataset on the universe of Maricopa County (Phoenix, Arizona) parcels and land transactions from 2000 through 2018, acquired via the Lincoln Institute of Land Policy. It is within this setting that our inquiry takes place. Using the vacant land transaction data, we estimate the value of each base-year parcel in each sample year. We fix the parcel universe to a base year in order to abstract from changes to composition and focus on the value of the lot, all else equal. The first parcel universe in the dataset is for 2007, so this is the base for the majority of our investigation. We also present the parcel universe for 2018 to illustrate changes in composition over the period and to establish the robustness of results to the choice of base year.

Maps of Maricopa County are depicted in Figure 1. The western and southwestern portions of the county are mostly desert and rural. The urban core is the area around Phoenix. Scottsdale and areas north are traditionally the wealthier areas of the county, and as we see from the land sale data in Figure 2, is also the location with the most expensive vacant land. Mesa and areas southeast were previously a low-cost areas, but land values in this area have increased substantially post-recession reflecting increasing demand. The Sun Cities are large planned communities and are a mid-price option in terms of land.

Summary statistics for the Maricopa County parcel database in 2007 are shown in Table 1.

[8] DLOS (2021) merged tax assessment land values and appraisal land values and showed that median appraised to assessed land value ratio is 1.21 (Figure 2(b)).

In this table, parcel counts and attributes are compared to the 2018 parcel database and all land-only transactions less than 10 acres conducted between 2000 and 2018. There are 96,154 vacant land transactions over the period in our sample. This is 5,652,698,457 square feet (129,768 acres), which is about 8% of the cumulative parcel area in the county. The final columns consist of adjustment factors used to standardize land values which will be discussed later. The table is divided into two parts, the first including counts, lot size, and other attributes, and the other including zoning.

There are several noteworthy facts in this table. The first is the changing composition of the parcel database between 2007 and 2018. The number of parcels increased by 6.15% and the average lot size increased by 2.91%, indicating that the land area covered by the parcel database increased substantially, with the newest additions being larger-than-average parcels. The number of cul-de-sac and gated parcels increased by 11.07% and 16.84% and the number of greenbelt parcels nearly doubled (+99.87%) and preserve land increased by 153.24%, suggesting a greater focus on conservation efforts. Planned communities such as Sun City were substantial contributors to the increase in parcels, increasing by 79.02%.

In terms of zoning, the major increase in parcels between 2007 and 2018 was the increase of planned communities and manufactured housing. Land transactions tended to have a greater representation of low-density residential zoning compared to existing parcels (500% greater share of 1-acre lots), agricultural, commercial, and industrial. High-density residential lots were the least likely to be transacted vacant relative to the share in the parcel database. Overall, these databases have rich variation in observed attributes, alongside a broad distribution of transactions across both space and time.

3 Parcel-Level Land Value Estimation

Because values are not observed for every parcel in every time period, it is necessary to estimate them. Spatial imputation models can be used to estimate an unobserved value at some location using a weighted average of observed values at nearby locations. Tobler's (1970) first law of geography is assumed when constructing the weights: "everything is related to everything else, but near things are more related than distant things." In practice, this means assuming nearer observations have at least as much weight as observations that are far away, all else equal. Beyond this basic setup, there are myriad methods to spatially impute unknown values. Which should be used?

3.1 Methods

To frame the discussion, suppose N observations of P exist across space, where P represents the price of land. We wish to estimate $\hat{P}$ at some location i. Then, ordinally index each of the P observations using i, j, which is the jth of the nearest observations to i. Let $\lambda_{i,j}$ be the weight given to observation $P_{i,j}$ in the construction of $\hat{P}$. Assume λ is weakly decreasing in h, the distance between the points irrespective of location or direction. Finally, assume a linear functional form such that the estimated value is an additive sum of the weights and the n nearest observed values, with $\sum_{j=1}^{n} \lambda_{i,j} = 1$ for each i.

$$\hat{P}_i = \sum_{j=1}^{n} \lambda_{i,j} P_{i,j} \tag{1}$$

Within this framework, there are two ways to proceed: make assumptions regarding λ, or estimate λ using optimization methods.

3.1.1 Assumption Methods

The first class of estimators we consider involves assumptions regarding the observation weights λ. The simplest prediction method is called the **null estimator**. In this estimator, the imputed value is the simple average of all of the observations in the dataset, with $n = N$ and $\lambda = 1/N$. This is the baseline to which spatial imputation accuracy is typically evaluated. The **nearest neighbor** estimator is the same as the null estimator, except it is the simple average of the n nearest values (say, 10 or 20) and not the entire dataset.

Another common method is called **inverse-distance weighting**, which is a constrained case where

$$\lambda_{i,j} = \frac{1}{h_{i,j}^{c}} \left(\sum_{j=1}^{n} \frac{1}{h_{i,j}^{c}} \right)^{-1} \tag{2}$$

The exponent c gives the degree of decay in the weight that is due to distance between the parcels. This is often set to 2, giving a rate of decay equal to the squared distance between the locations. These assumption methods are often favored due to their ease of use but are weakly inferior to predictions from models that optimize λs. As we will see, there are cases where an assumption-based weights result in estimates that are statistically indistinguishable from optimization approaches.

3.1.2 Optimization Methods

There are various ways to optimize linear weights, but each is called "kriging" after D.G. Krige, an engineer who pioneered the approach. In each version, rather than make assumptions about λ, the weights are estimated based on the strength of the observed relationship between observations of different proximities within the sample.

The most straightforward of the kriging methods is called "ordinary kriging," and involves solving a constrained minimization problem under *intrinsic stationarity*, which requires two key assumptions (see Sherman (2011) for derivation). Let us slightly alter our notation for a moment, with $P(s)$ indicating the price at location s, and $P(s+h)$ being another observed price h distance away from s. The two assumptions are: that the estimates are unbiased around an unknown mean, $E[P(s)] - \mu_p = 0$, and the variance between any two points is only related to the distance between them, $Var[P(s+h) - P(s)]/2 = \gamma(h)$. This function γ is called a "semivariogram" with 2γ defined as the "variogram". The function γ plays a crucial role in the calculation of the kriging weights. Note that γ is invariant to direction or location; only proximity matters.

The problem consists of minimizing the sum of the squared prediction errors subject to the constraint that the weights sum to unity for each i.

$$\min_{\lambda} E\left[\sum_{i=1}^{n}\left(\sum_{j=1}^{n}\lambda_{i,j}P_{i,j} - P_i\right)^2\right] \; s.t. \; \sum_{j=1}^{n}\lambda_{i,j} = 1 \; \forall i \tag{3}$$

This problem involves a number of steps to solve (see the solution in Sherman (2011)), but the key takeaway is that it hinges on the calculation of γ. This function is given an assumed form (e.g. exponential or spherical) and the relevant parameters are estimated using the data. In practice, there are two main decisions when solving an ordinary kriging problem: (1) the number of nearest neighbors to consider n, and (2) the functional form of

γ.[9] While these could presumably be optimized along with the kriging weights, this is not typically done. However, we will undertake some sensitivity analysis to establish robustness to alternative nearest-neighbor and functional form assumptions.

3.2 Spatial Imputation using Kriging

3.2.1 Standardization

Before moving on to the estimation of the semivariogram, it is useful to standardize the values of transacted parcels to eliminate predictable sources of variation across lots. By eliminating this class of variation, the spatial surface is more efficiently estimated. For our standard lot, we choose 1/4 acre with R10 zoning (about 1/4 acre maximum) to maintain some intuition around the results, but this decision is arbitrary and does not affect the estimation of the variogram. We use a linear regression specification for standardization under the assumption that attributes and zoning have constant, multiplicative effects throughout the county. The simple standardization equation we estimate using OLS is as follows:

$$\ln p_{it} = b_1 \ln sqft_{it} + b_2 R10_{it} + A' year_{it} + C' X_{it} + e_{it} \tag{4}$$

In the equation above, the vector X includes the set of covariates found in Table 1, which consists of variables describing the amenity level of the lot (e.g. corner, adjacent to golf course, adjacent to railroad) and the zoning.[10] The variables $year_{it}$ are dummy variables set to 1 for each parcel in the particular year. This model is estimated over the 96,154 land-only transactions in Maricopa County between 2000 and 2018. Estimates of $\hat{b}_1$, $\hat{b}_2$, $\hat{A}$, and $\hat{e}_{it}$ are used to calculate standardized land values.

[9] Two other methods that deserve special attention are "regression kriging" and "spatio-temporal kriging". In regression kriging, deterministic covariates are estimated simultaneously with a spatial surface of residuals. This estimator is more efficient than ordinary kriging with unmodeled, known covariates because the variation is captured by the spatial relation. Spatio-temporal kriging incorporates information from different time periods into the spatial surface. The computational burden of this approach is substantial because the γ is modeled using two arguments, spatial proximity h and temporal proximity τ. Due to the large number of observations in Maricopa County, it is not feasible to undertake spatio-temporal kriging. Covariates are eliminated through standardization in a first-stage procedure, described later in this section, obviating the need for regression-kriging.

[10] Therefore, the standardized lot should be interpreted as a 1/4 acre lot with R10 zoning and zero amenities observed and reported in the assessor dataset.

$$\ln \tilde{p}_{it} = \hat{b}_1 \ln(1/4\ Acre) + \hat{b}_2 + \hat{A}_{year} + \hat{e}_{it} \quad (5)$$

These parameter estimates are found in the final columns of Table 1, with a_t[11] omitted for brevity but available upon request. Of particular note is the plattage parameter of 0.58 on $\hat{b}_1$, indicating that lot size increases value at a decreasing rate; a 10% increase in the lot size only increases the lot value by 5.8%. Accordingly, larger lots have a lower price on a per-acre basis, all else equal.

Variables indicating amenity or accessibility are positively associated with values. Corner lots have an 20% premium and lots with the freeway access indicator have a 33% premium. Lots in gated communities, planned developments, and those adjacent to golf courses all sell at a premium. Conversely, being near a rail or other transit line is negatively associated with value. Zoning estimates are relative to agricultural zones and are thus mostly positive. Commercial zones have the largest premia and residential have premia that generally decrease with density.

Overall, estimates tend to conform with priors concerning the effects of various attributes on land values. While there is almost certainly simultaneous causality of lot size, lot value, and zoning, and we caution against using these estimates for structural applications, these partial effects allow us to eliminate systematic variation in land values across parcels in a reduced-form manner.

3.2.2 Variogram

The standardized values in the dataset are then grouped into $N-1 \times N-1$ pairwise combinations for each year, where N is the number of observations. As an intermediate step, we collapse transactions to a fine grid at latitude/longitude increments of 0.001, or about 365 feet. Standardization facilitates simple averaging where multiple transactions exist within a grid cell. For each pair of grid points, the squared difference between the two standardized (log) land values is calculated and divided by two, and plotted as a function of the distance between the two points. A functional form with parameters, θ, for this relation

[11] Coefficients for year fixed effects.

is assumed $\gamma(h, \theta)$.[12] The empirical semivariogram $\gamma(h, \hat{\theta})$ is then estimated. This function is used to estimate the covariance between land values h distance apart, and the set of these covariances for the nearest n observed locations to i are used to solve for the kriging weights λ.

Figure 3 shows several of the estimated semivariograms. Each panel depicts squared deviations (divided by two) that are binned by h as dots with the fitted empirical semivariograms as a line. The functional form is the same in each (spherical) but the parameter estimates are different. This suggests that the spatial correlation between land prices is changing over time and it is important to estimate new variogram parameters each year.

3.2.3 Imputation Fit

To evaluate the fit of rival spatial prediction methods, spatial weights are calculated using an 80% training sample that is held fixed across each method. We then take predictions for a 20% hold-out sample and calculate the root mean squared error (RMSE). This rudimentary cross-validation serves as a check on the accuracy of the spatial imputation procedure.

Table 2 contains 8 columns of RMSEs, each corresponding to a spatial prediction method, across 19 years (2000 through 2018). Model 1 is the null estimator, calculated as the simple average across Maricopa County within the year. Model 2 is the nearest neighbor, which is the simple average of the 50 nearest observed values. Model 3 considers inverse-distance weighting. Models 4 through 8 show the results for various kriging estimators, with 4-6 considering different numbers of nearest neighbors, 7 an alternative functional form, and 8 a less granular spatial grid.

There is a clear rank-ordering of the accuracy of the methods at the parcel level. The least accurate is the null estimator, with a hold-out RMSE of .97, meaning the standard deviation of an average residual is nearly an order of magnitude, relative to the county

[12] The spherical functional form requires estimating three parameters.

$$\gamma(h; a_0, a_1, r) = \begin{cases} a_0 + a_1(\frac{3}{2}(\frac{h}{r}) - \frac{1}{2}(\frac{h}{r})^3), & 0 < h < r \\ a_0 + a_1, & h \geq r \end{cases}$$

The three parameters combine to give the "sill" which is the value to which the semivariogram asymptotically approaches as the distance between points approaches infinity, or $a_0 + a_1$; the "nugget" which is the value of the semivariogram when distance approaches zero, or a_0; and the "range," r, which is the value of h when the semivariogram reaches the sill.

mean. The nearest neighbor method is the next worst, with an RMSE of 0.72. Inverse-distance weighting and all but one of the kriging estimators are indistinguishable and tied for the best, on average, with RMSEs of 0.66. Assumptions for the kriging methods considered include variation in the number of nearest neighbor values and spherical versus exponential functional forms.[13] Grid size does matter, with an order-of-magnitude expansion of the grid increasing the RMSE to 0.68. Based on these results, we proceed with kriging due to its theoretical properties, with 50 nearest observations under a spherical functional form with a grid of 0.001 latitude/longitude degrees.

3.3 Results

Using the chosen kriging estimator and the full sample of vacant land transactions, we estimate a full 0.001 degree latitude/longitude grid of standardized land values throughout Maricopa County for each year between 2000 and 2018. For each year, the gridded standardized land value surface is projected onto each land parcel in the master Maricopa County parcel record for 2007. We maintain this base-year parcel database to hold constant any composition effects in land values.

Figure 4 depicts a smoothed surface of the spatial imputation grid. Because some areas have no transactions in some years, we limit the results depicted in the grid to the concave hull represented by the yearly transactions. The panels show standardized land value surfaces for 2000, 2006, 2009, 2012, and 2018, respectively. Between 2000 and 2006, rapid appreciation occurred in most areas of the county. By 2009, some areas had cooled, but around Scottsdale (center-right) remained strong. Prices had collapsed in most areas by 2012, but have slowly recovered since, especially Mesa (bottom-right). A fan graph of percentiles of parcel-level land values for each year is depicted in Figure 5. This figure has a log scale, reflecting the right-skewness of the land value estimates for any particular year. Standardized land values rise and fall as reflected in the spatial figures, from a median of about \$22,000 per 1/4 acre in 2000 to \$63,000 in 2007, falling to \$19,000 in 2011, and finally rising to about \$38,000 in 2018.

To restore the value of each lot to reflect its characteristics, the adjustment parameters from the standardization procedure are re-added back into the predicted standardized land value,

[13] For some comparison, 20% hold-out RMSE estimates in repeat-sales house price indices are approximately 0.1 to 0.2 (Bogin et al., 2019) depending on the holding period, and 20% hold-out estimates of values calculated using kriging and land-under-structures is about 0.5 (DLOS, 2021). Vacant land is roughly comparable but somewhat larger than the values found in this prior work using this other source data.

creating a predicted "as-is" land value for each parcel. These as-is values are depicted in Figure 6. Note the scale in this figure is much wider and more skewed than in Figure 4, implying that standardization served to mute much of the variation over space to facilitate spatial modeling, as intended. Phoenix and Scottsdale are the two highest-value areas, with land values per 1/4 acre ranging between $0.5 and $5 million in many tracts. Values tend to decline radially from this high-priced band of tracts.

4 Valuations Based on Vacant Land versus Land Underneath Structures

The prior section introduced a panel of parcel-level land value estimates for each year from 2000 to 2018. We have affirmed the internal consistency of the estimates based on the theoretical properties of the estimator, and efficiency based on a 20% hold-out sample. Further validation is both possible and necessary. The validation exercises in this section focus on single-family residential real estate.

The first exercise compares aggregated county-level land value estimates to an alternative method of calculation in DLOS (2021), which also uses kriging to spatially impute values to unpriced parcels, but uses different source data–direct appraisals of value of land underneath structures. Both of these values track each other closely in the growth rate, but are different in the level due to differences in both observable and unobservable characteristics, which we will demonstrate.

The second validation exercise examines the link between land prices and house prices both in terms of tract-level levels and growth rates. Here is where parcel-level volatility translates to tract-level volatility and causes problems. The correlation between levels is relatively strong, but the correlation in growth rates becomes statistically indistinguishable from zero.

Our conclusion is that land valuation using vacant transactions is reliable in the growth rate at the county level, but not in the level. However, at the tract or lower level, neither is particularly reliable. In cases where tract or parcel level values are necessary, it may be prudent to implement some sort of intertemporal aggregation method or smoothing.

4.1 County-Level Values and Growth Rates

For comparison with land values in Section 3, we use estimates of land underneath single-family structures produced by DLOS (2021). This paper uses cost-approach appraisals of newly constructed or renovated homes to infer the land value as the difference between the transaction price of a property and the replacement cost of the structure. Valid appraisals are imputed to all single-family parcels using ordinary kriging using methods similar to those in the present paper. These estimated values are averaged to construct tract- and county-level levels and growth rates between 2012 and 2018. These estimates have proven to be reliable in terms of both internal and external validity, with external validity established by comparing calculated values to appreciation rates of house prices from Bogin et al. (2019) and cross-sectional gradients across a large number of cities.

Table 3 shows summary statistics comparing the single-family residential subset of the Maricopa County parcel database, land transactions, and the parcels covered in DLOS (2021). The average residential parcel has a similar lot size, at about 10,800 square feet in the Maricopa dataset versus 10,300 in the DLOS dataset. The average land value is \$181,000 in the transactions database. After valuing each parcel in the entire county for 2018, the average estimated land value is about \$106,000 using vacant transactions and \$102,000 using land underneath structures from the DLOS database, though these are based on slightly different samples. The average difference between county land value levels is shown by year in Figure 7. Between 2012 and 2018, vacant land transactions generate land values that are, on average, 14% higher than land underneath structures.

This level difference deserves additional discussion so we can understand the nature of this 14% average gap. Let us begin by comparing tract-level average characteristics of residential land under structures with the value of land for raw vacant land transactions zoned for single-family residential. The average lot size for vacant land is typically much larger than existing built-up lots, as shown in Figure 9, panel (a). Raw value differentials are small for tracts where lot sizes are similar (panel b), whereas value differentials are large for tracts where lot sizes are different (panel c). In the latter case, for higher vacant values, vacant land is more expensive than land under structures; for lower values, it is cheaper.

These simple figures underscore the need to standardize values before imputing to all parcels. We find in our standardization equations that larger lots sell for a discount on a per-acre basis compared to smaller lots, with a value-size elasticity of 0.58. Therefore, when the value of vacant land is imputed to smaller parcels, the value is larger than the raw per-acre average. This is how it is possible to end up with the case where raw values are lower on a per-acre basis, but average imputed values are higher, compared to land values estimated directly from lots with structures.

Figure 9 panel (d) shows land value differences plotted alongside lot size differences. There is also a line on the plot representing the difference expected from the plattage effect. The quadratic line of best fit is above this line for most of the range, indicating that large vacant lots are not sold for a large enough discount as would have been predicted by the vacant land plattage effect estimated over vacant land transactions. It is this differential that is the source of the difference in the level between the value of land calculated using vacant land versus land underneath structures.

This difference in the levels also highlights issues related to unobservable factors related to vacant land sales. Because the values in panel (b) are so similar for similar lot sizes, we infer that the large level differences are not due to differences in the option value of redevelopment nor teardown costs, and that unobservable differences are not everywhere in the city. Rather, we infer that there are large, positive, unobserved factors that are positively correlated with the size of residential lots. Accordingly, direct use of vacant land transactions may result in biased estimates of land underneath structures in Maricopa County.

Growth rates, on the other hand, are quite aligned. Both follow a similar time-path of directional changes while maintaining similar magnitudes. The growth rate from 2012 to 2013 is 21% using the vacant land sales data versus 15% when using land under structures. From 2013 to 2014 the growth rates are 16% and 10% respectively. As the growth rates slow, the precise rates depart but the directional signs and signs of changes are maintained. The only major difference in the series is 2018, the last year in the series, where the vacant land growth rate falls to zero while the growth rate for land under structures is 13%. Overall, given that these series are constructed with different source data, the strong association between both directional changes and magnitudes is encouraging.

4.2 Census Tract Levels and Growth Rates

When disaggregating to the tract level, the source of the differences in the aggregate values comes into focus. The first panel of Figure 8 shows a scatterplot of the 900 tracts for which land values are available in both datasets for 2018. There is a strong correlation of land value levels with a slope parameter of the value-value relation of 0.53 with a standard error of 0.04. High-valued tracts appear to be the source of the difference between the two county-level aggregate series, as the line of best fit moves further away from the 45-degree line.

On the other hand, the clear and relatively stable growth rates found at the county level are seemingly masked by noise in the disaggregate. Tract-level growth rates, measured as the log-difference between 2012 and 2018, have a very low correlation. The slope is -0.02, with a standard error of 0.05, implying a slightly positive but not statistically significant correlation. As we saw from Figure 7, there is a strong relation at the county level, implying that the low slope coefficient in the tract-level scatterplot is indeed due to noise and not to a lack of association.

Confirming the notion that the growth rate of the value of land estimated using vacant land sales is confounded by noise at the tract level are plots of changes in land value versus change in house prices. As shown by Davis and Heathcote (2007), Davis and Palumbo (2008), and others, under certain conditions, the growth rate of house prices can be expressed as a weighted average of the growth rate of structure costs and the growth rate of land prices. Because construction costs are roughly constant within a metro area, house price appreciation should respond on a less-than one-for-one basis to a change in land values. The higher the initial land value share, the closer this relation should be to unity. For land values estimated using land-under structures found in DLOS (2021), this slope parameter is 0.33 with a standard error of 0.05, indicating a 10% change in land values is associated with a marginal 3.3% increase in the price of housing, a plausible estimate. On the other hand, this slope parameter is 0.10 with a standard error of 0.02, suggesting about 1/3th of the magnitude of the marginal effect. While this may be accurate for extremely low land shares, this estimate is unlikely to be the population parameter.

In sum, we conclude that annual county-level land values and growth rates are relatively accurate and can be relied upon, though the land value levels are only usable to price vacant land. Tract levels are also reasonably accurate, as evidenced by the strong correlation with an alternative and already-validated estimate of the value of land. Tract-level growth rates, on the other hand, are masked by substantial noise and are suspect on an individual basis. In Maricopa County and elsewhere, it may be possible to measure level values for a base year using a thorough study measuring the value of land underneath structures directly, then extrapolate to other years using vacant land sales.

5 Implications for Land Value Taxation

Thus far, we have done two things in this paper. First, we estimated the value of land underneath structures using vacant land sales, and second, we have shown that these estimates to be reasonably accurate at the county level in Maricopa County. Because counties have available a record of land transactions, their location, and price, it is reasonable to suggest that the methods and the findings in this paper, if applied by tax assessors, could open up new possibilities for large-scale land valuation in cities. While our study focuses only on Maricopa County, the procedures we follow for estimating and assessing land valuation methods are likely reproducible in a variety of settings.

Our focus now turns to the implications of our findings as they relate to a hypothetical land value tax regime. It is crucial for tax assessors to understand a particular aspect of land value taxation that has not been the subject of much research that we are able to address using the land values estimated in this paper: the behavior of land prices over a housing cycle. Because land is typically the more volatile component of a property, a tax exclusively on the value of land would therefore be more volatile in terms of revenue generated.

Armed with the land value series estimated in this paper, it is simple to estimate the revenue generated by a hypothetical and highly stylized residential land value tax regime over the housing cycle for Maricopa County. This is contrasted with a full property tax regime, constructed in the same manner, over the same period. The "method" is trivially simple under three assumptions. If one assumes a constant housing and land composition for the county, a constant tax millage rate for each regime, and that assessments are equal to market valuations, a tax revenue index is simply a value index weighted to some base year. Accordingly, we construct land and property tax revenue indices for Maricopa County using only the land value series estimated in this paper alongside the repeat-sales house price index found in Bogin et al. (2019), set to a common base year.

To quantify the variability of Maricopa County house prices and land prices, we estimate bivariate models relating each to the annual log-difference of national house prices and then to each other. The slope (β) parameter (and standard error) on the national change is 2.3

(0.28) over 2000 through 2018 (18 observations), indicating that for every percentage point change in the national house price index, house prices in Maricopa County change by 2.3%. When Maricopa County land prices are used, this parameter is estimated to be 3.3 (0.59). When Maricopa land prices are the dependent variable and Maricopa house prices are the right-hand side variable, $\hat{\beta}$ is 1.46 (0.16). Overall, these simple statistics indicate land prices in Maricopa County are more than 3 times as volatile as national house prices, and about 50% more volatile than local house prices.

When indexed as a revenue series in 2007, both land and property tax values are set to 100 in 2007. This represents the period where both are set to be equal and changes from this point represent differences in revenue. In this period, actual property tax revenue for Maricopa County was about $397 million. By 2011, in our stylized property tax regime, revenues would have fallen to 51% of the 2007 value, or $202 million.[14] This is 50% more than the land tax, which would have generated just 34% of the 2007 value, or $135 million. By 2018, revenue in the property tax regime would have almost fully recovered to the 2007 value at 96%, but the land tax revenue would have only recovered to 66%.

This base-2007 comparison is the minimum-revenue scenario for a land tax. Were a revenue-neutral land value tax implemented in 2000, revenues would have skyrocketed to 297% in 2007, fallen to 102% by 2011, and recovered to 195% in 2018. This compares to property taxes which are only above land taxes for two years, 2011 and 2018.

Overall, a hypothetical land tax regime in Maricopa County would have been highly sensitive to the period in which it was implemented due to rapid fluctuations in the price of land relative to property. Accordingly, practitioners and tax assessors may wish to find some way to mute the effect of base-year choice on revenue. Then, based on the likelihood of excess volatility in land versus property tax revenues, some sort of reserve fund might be necessary, which can spread the volatility risk over a decade or more.

6 Conclusion

Economists have long argued that taxing structures leads to an inefficiently low provision of housing and other real estate capital investments. Though a property tax is incentive

[14] In reality, primary property tax revenue did not decrease much in 2011, it was until 2014 ($387 million) that it almost fell to the 2007 level ($397 million), indicating changes to some combination of composition, the ratio of average assessed to market value, and changes to tax millage rates.

compatible with amenity provision and its base is long-lasting (Glaeser, 1996), a land tax would lead to less distortion in building activities and more sustainable growth patterns by encouraging efficient development. Land values are not typically easy to estimate over a county, making it difficult to implement land taxation regimes, forcing areas to adopt property tax regimes instead. Our paper aims to help rectify this problem by constructing estimates of the value of land using public land records and an optimizing model of spatial imputation, kriging.

In this paper, for Maricopa County, Arizona, we have shown kriging to be error-minimizing at the parcel level compared to alternative methods. We have shown the level estimates of land values produce using vacant land sales to be inflated versus the value of land under structures. However, given a proper base-year valuation, growth rates at the county level appear to be applicable to developed lots. We believe our results to be an encouraging start that points to the potential value of undertaking similar studies in other counties.

Before moving forward with any land tax regime, governments should be aware of the volatility of land values relative to property values experienced in Maricopa County, and the associated risk to public finances were a land tax regime to have been implemented there. Because of volatility in land prices prior to the implementation of the hypothetical land tax, there is substantial sensitivity to base-year valuations. During booms, coffers become flush as land prices rise due to rapid increases in the option value of development. On the other hand, busts may be particularly painful, as land prices tend to decline more than structures and therefore property prices. Margins for adjustment assumed away within the paper may be able to smooth revenues but be politically infeasible, including quick adjustments to land tax rates in response to valuation changes. The slowness of changes to assessments, however, would serve to mute revenue effects of market valuation changes, though tax inequities may arise as a result (see Lutz, Molloy, and Shan, 2011; McMillen, 2008).

References

Albouy, D. and Ehrlich, G. (2015). The distribution of urban land values: Evidence from market transactions. In *Proceedings. Annual Conference on Taxation and Minutes of the Annual Meeting of the National Tax Association*, volume 108, pages 1–32.

Albouy, D., Ehrlich, G., and Shin, M. (2018). Metropolitan land values. *Review of Economics and Statistics*, 100(3):101–120.

Almy, R. (2013). *Global compendium and Meta-Analysis of property tax systems.* Lincoln Institute of Land Policy Working Papers.

Anas, A. (2003). *Taxes on buildings and land in a dynamic model of real estate markets.* Edward Elgar Publishing Limited.

Anas, A. (2015). *The Effects of Land Value Taxation in a Computable General Equilibrium Model.* Lincoln Institute of Land Policy Working Papers.

Beale, T. M., Miller, R. A. C., Ishemo, A., and Murray-Stuart, C. A. (2016). *Towards Property Tax Compliance: A Case Study of Attitudes Toward Paying Property Taxes in Jamaica.* Lincoln Institute of Land Policy Working Papers.

Biitir, S. B. (2019). *Designing Land Value Capture Tools in the Context of Complex Tenurial and Deficient Land Use Regulatory Regimes in Accra, Ghana.* Lincoln Institute of Land Policy Working Papers.

Bogin, A., Doerner, W., and Larson, W. (2019). Local house price dynamics: New indices and stylized facts. *Real Estate Economics*, 47(2):365–398.

Bostic, R. W., Longhofer, S. D., and Redfearn, C. L. (2007a). Land leverage: decomposing home price dynamics. *Real Estate Economics*, 35(2):183–208.

Bostic, R. W., Longhofer, S. D., and Redfearn, C. L. (2007b). Land leverage: Decomposing home price dynamics. *Real Estate Economics*, 35(2):183–208.

Cho, S.-H., Kim, S. G., and Roberts, R. K. (2011). Measuring the effects of a land value tax on land development. *Applied Spatial Analysis and Policy*, 4(1):45–64.

Clapp, J. M., Cohen, J., and Lindenthal, T. (2020). Valuing urban land with land residual and option value methods: Applications to house price dynamics and volatility. *Available at SSRN 3618293.*

Combes, P.-P., Duranton, G., and Gobillon, L. (2012). *The costs of agglomeration: Land prices in French cities.* CEPR discussion paper No. DP9240.

Davis, M. A. and Heathcote, J. (2007). The price and quantity of residential land in the united states. *Journal of Monetary Economics*, 54(8):2595–2620.

Davis, M. A., Larson, W. D., Oliner, S. D., and Shui, J. (2021). The price of residential land for counties, zip codes, and census tracts in the united states. *Journal of Monetary Economics*, 118:413–431.

Davis, M. A., Oliner, S. D., Pinto, E. J., and Bokka, S. (2017). Residential land values in the washington, dc metro area: New insights from big data. *Regional Science and Urban Economics*, 66:224–246.

Davis, M. A. and Palumbo, M. G. (2008). The price of residential land in large U.S. cities. *Journal of Urban Economics*, 63(1):352–384.

Diewert, W. E., Haan, J. d., and Hendriks, R. (2015). Hedonic regressions and the decomposition of a house price index into land and structure components. *Econometric Reviews*, 34(1-2):106–126.

Doerner, W. M. (2012). Dynamic and disparate distributions of property assessments. Working Paper.

Duke, J. M. and Gao, T. (2017). *Public Acceptability and Land Value Taxation: An Experimental Economics Investigation*. Lincoln Institute of Land Policy Working Papers.

Dye, R. F. and England, R. W. (2010). *Assessing the theory and practice of land value taxation*. Lincoln Institute of Land Policy Working Papers.

Fischel, W. A. (2013). Fiscal zoning and economists' views of the property tax. *Available at SSRN 2281955*.

Gallagher, R. M., Kurban, H., and Persky, J. J. (2013). Small homes, public schools, and property tax capitalization. *Regional Science and Urban Economics*, 43(2):422–428.

George, H. (1879). *Progress and poverty: An enquiry into the cause of industrial depressions, and of increase of want with increase of wealth. The Remedy*. K. Paul, Trench & Company.

Glaeser, E. L. (1996). The incentive effects of property taxes on local governments. *Public choice*, 89(1-2):93–111.

Groves, J. R. (2011). *Estimating the Responsiveness of Residential Capital Investment to Property Tax Differentials*. Lincoln Institute of Land Policy Working Papers.

Haughwout, A., Orr, J., and Bedoll, D. (2008). The price of land in the New York metropolitan area. *Federal Reserve Bank of New York: Current Issues in Economics and Finance*, 14(3).

Ingram, G. K. et al. (2012). *Value capture and land policies*. Lincoln Institute of Land Policy Working Papers.

Kuminoff, N. V. and Pope, J. C. (2013). The value of residential land and structures during the great housing boom and bust. *Land Economics*, 89(1):1–29.

Lichfield, N. and Connellan, O. (1997). *Land value taxation in Britain for the benefit of the community: History, achievements and prospects*. Citeseer.

Lutz, B., Molloy, R., and Shan, H. (2011). The housing crisis and state and local government tax revenue: Five channels. *Regional Science and Urban Economics*, 41(4):306–319.

Maxwell, D. and Vigor, A. (2005). *Time for Land Value Tax?* Institute for Public Policy Research.

McCluskey, W. J. and Franzsen, R. C. (2001). *Land value taxation: a case study approach*. Lincoln Institute of Land Policy Working Papers.

McMillen, D. P. and Weber, R. N. (2008). Thin markets and property tax inequities: A multinomial logit approach. *National Tax Journal*, pages 653–671.

Nichols, J. B., Oliner, S. D., and Mulhall, M. R. (2013). Swings in commercial and residential land prices in the united states. *Journal of Urban Economics*, 73(1):57–76.

Oates, W. E. and Schwab, R. M. (1997). The impact of urban land taxation: The pittsburgh experience. *National Tax Journal*, pages 1–21.

Plummer, E. (2010a). *The effect of land value ratio on property tax protests and the effects of protests on assessment uniformity*. Lincoln Institute of Land Policy Working Papers.

Plummer, E. (2010b). Evidence on the distributional effects of a land value tax on residential households. *National Tax Journal*, 63(1):63–92.

Sherman, M. (2011). *Spatial statistics and spatio-temporal data: covariance functions and directional properties*. John Wiley & Sons.

Tobler, W. R. (1970). A computer movie simulating urban growth in the detroit region. *Economic Geography*, 46:234–240.

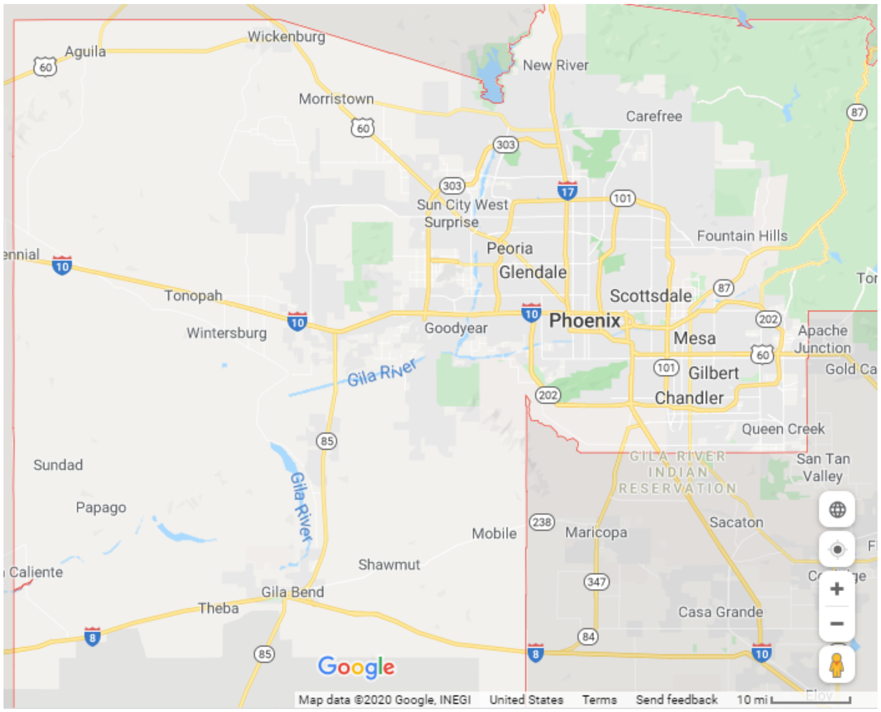
Turner, M. A., Haughwout, A., and Van Der Klaauw, W. (2014). Land use regulation and welfare. *Econometrica*, 82(4):1341–1403.

Wallin, B. and Zabel, J. (2011). Property tax limitations and local fiscal conditions: The impact of proposition 21/2 in massachusetts. *Regional Science and Urban Economics*, 41(4):382–393.

Yang, Z. (2014). *The Effects of the Two-Rate Property Tax: What Can We Learn from the Pennsylvania Experience?* Lincoln Institute of Land Policy Working Papers.

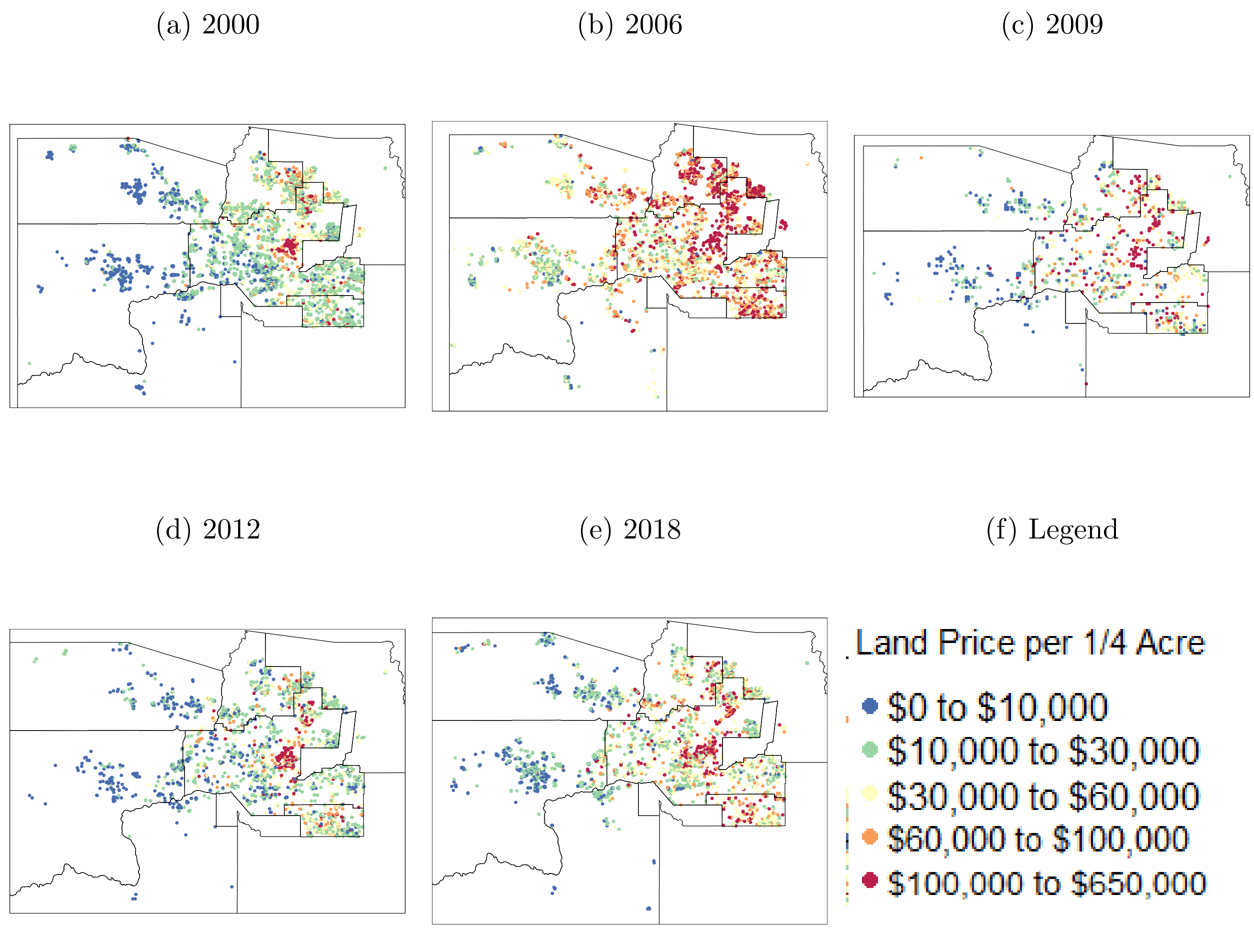
Figure 1: Maricopa County, Arizona

(a) Road Network



(b) Topography

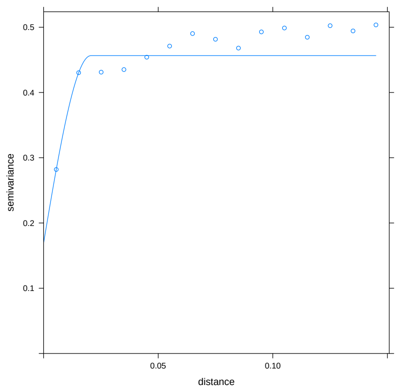
Figure 2: Land Sales with Standardized Values, Select Years



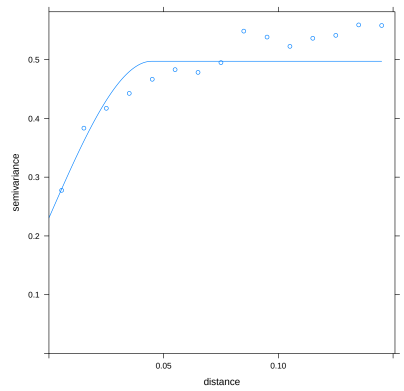
Notes: Land values are in terms of as-is price per 1/4 acre.

Figure 3: Variograms of Standardized Land Prices, Select Years

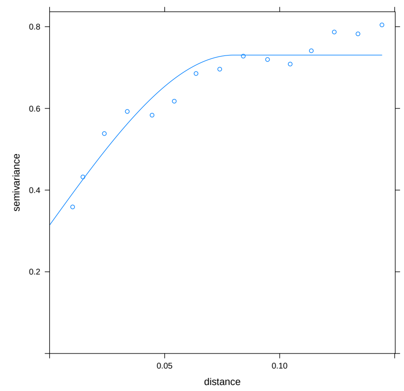
(a) 2000



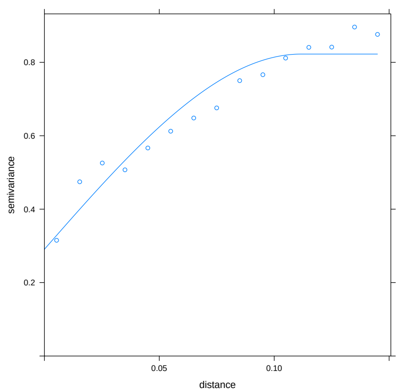
(b) 2006



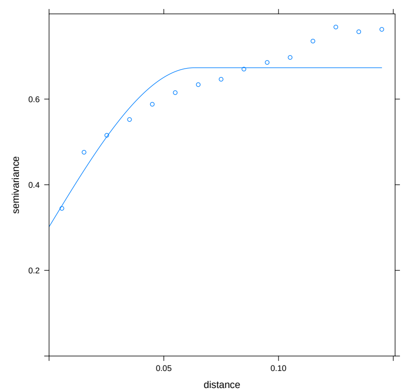
(c) 2009



(d) 2012

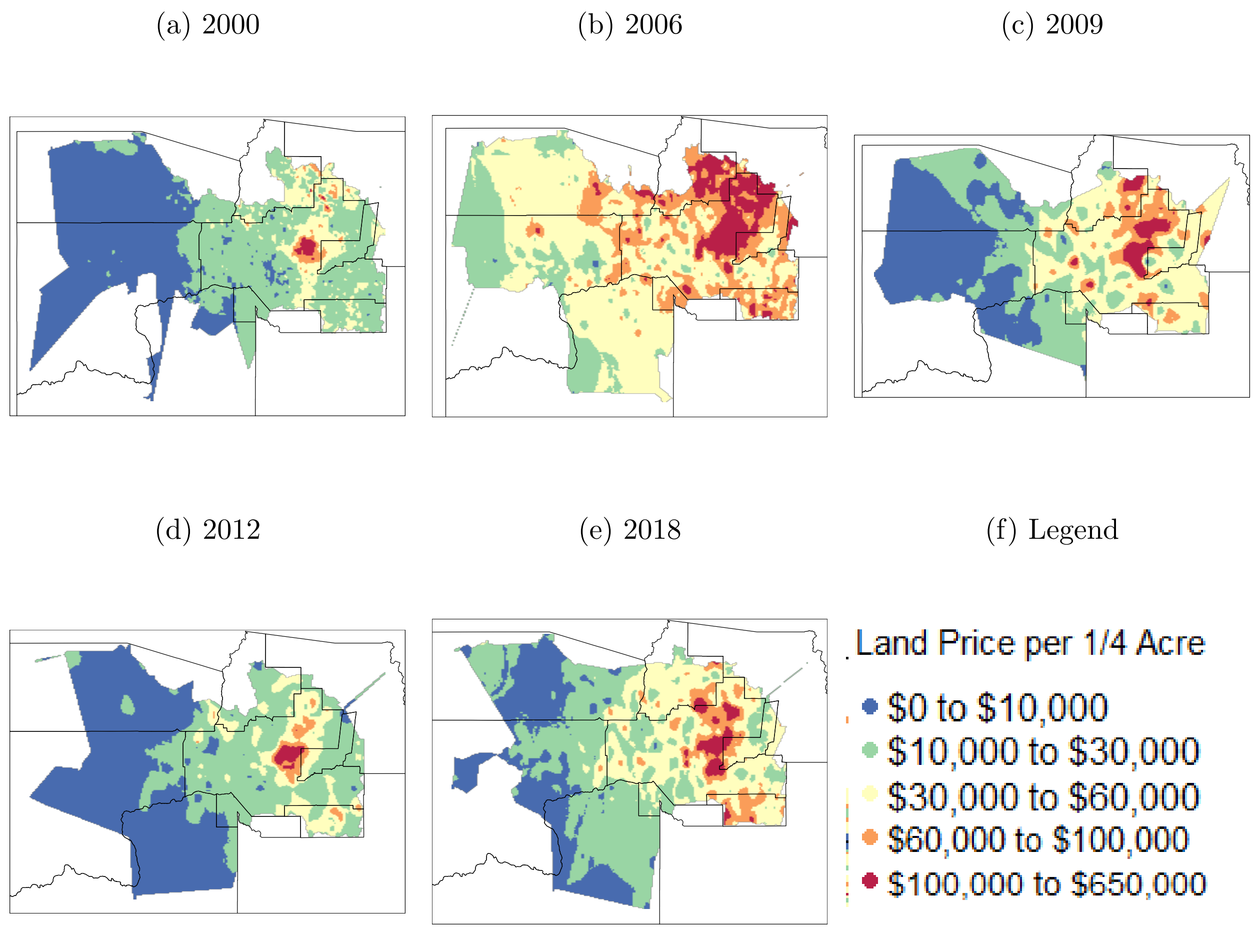


(e) 2018



Notes: These figures depict semivariograms γ estimated separately for each year. Dots represent distance bins which are collections of pairwise squared differences, divided by two. The fitted line is based on a spherical functional form.

Figure 4: Standardized Land Price Surfaces, Select Years



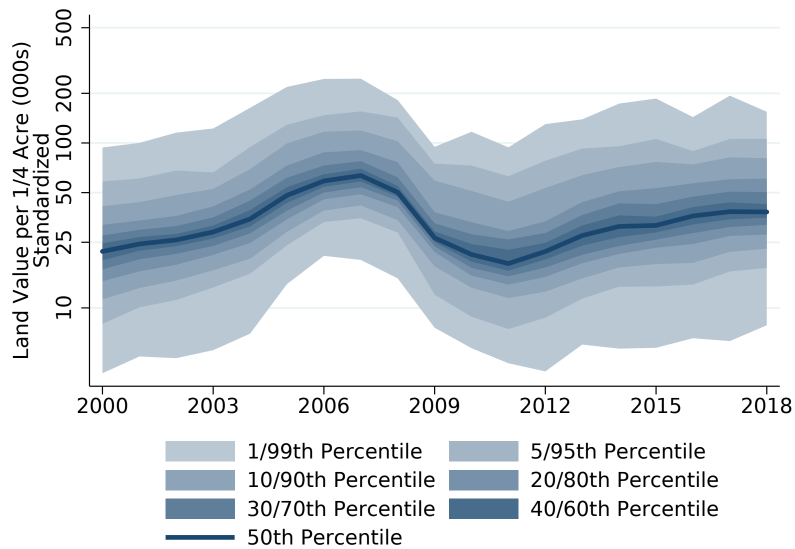
(a) 2000 (b) 2006 (c) 2009

(d) 2012 (e) 2018 (f) Legend

Notes: Land values are standardized to a 1/4 acre lot, zoned R-10 residential.

Figure 5: Distribution of Standardized Land Value per 1/4 Acre and Number of Land Transactions

(a) Distribution of Standardized Land Value per 1/4 Acre



Notes: This figure depicts percentile ranges of land values within a particular year.

(b) Number of Land Transactions

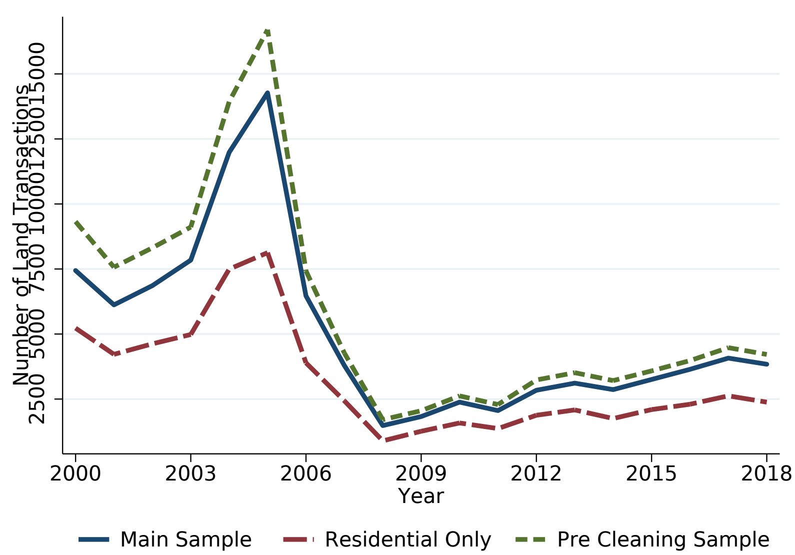


Figure 6: Land Price per 1/4 Acre, As-Is (2018), Census Tracts

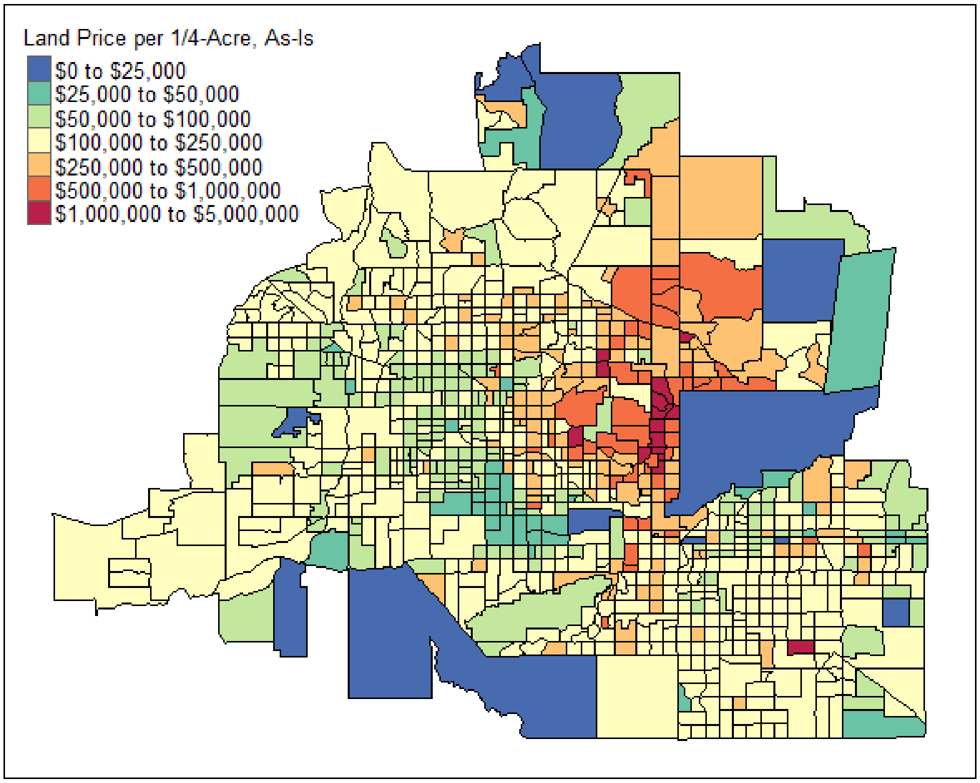
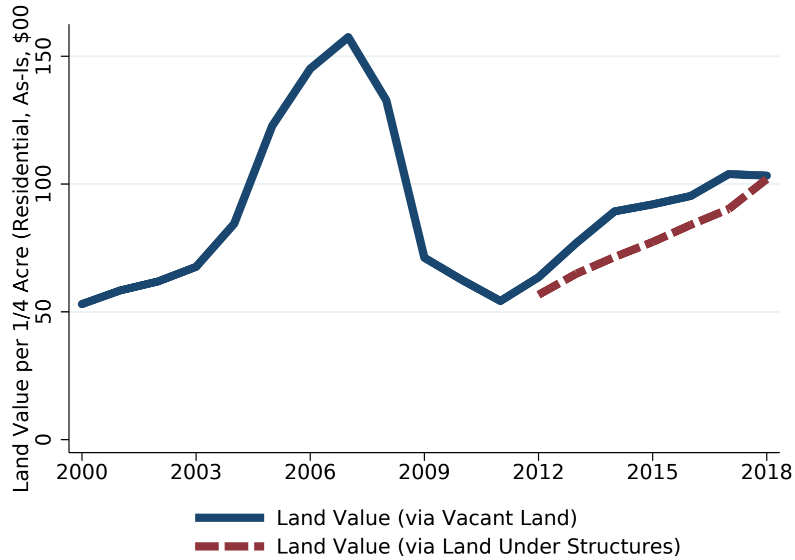
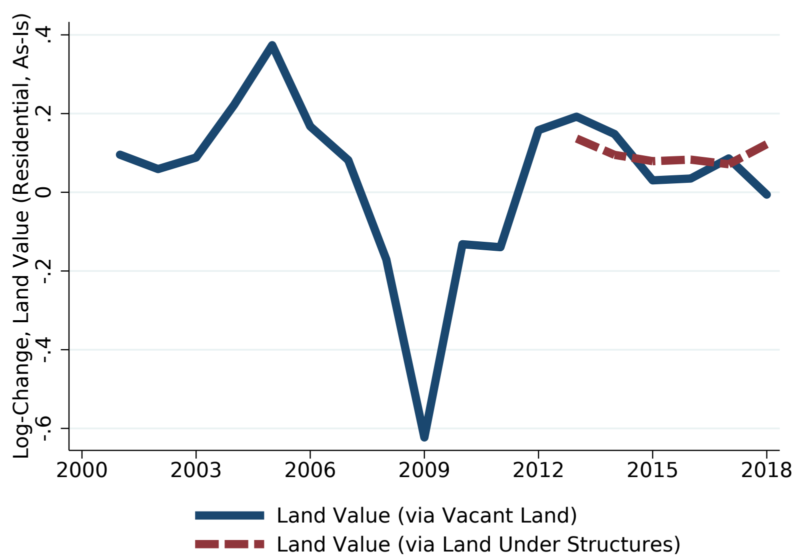


Figure 7: Land Values Created using Vacant Land vs Land under Structures

(a) Levels



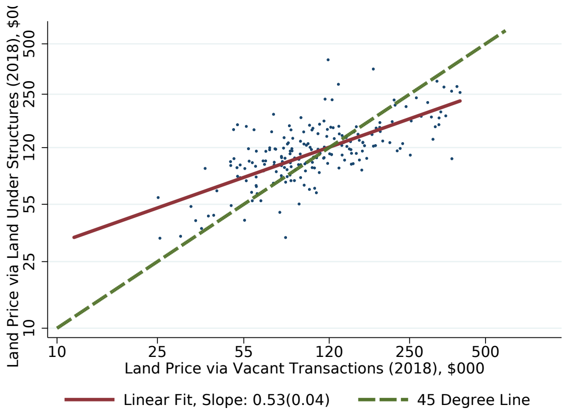
(b) Appreciation



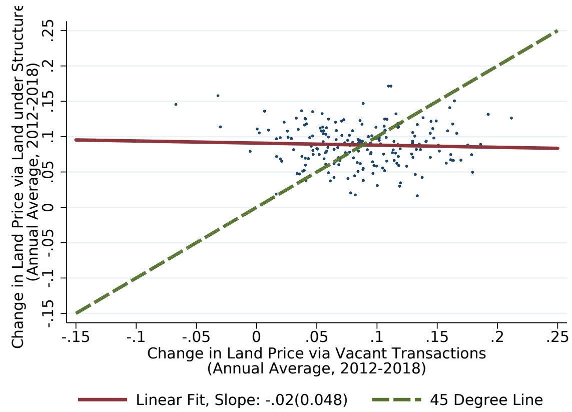
Notes: The source of the Land Value (via Vacant Land) is described in Section 3 of this paper. The source of the Land Value (via Land Under Structures) is found in Davis et al. (2021), which is calculated using a method similar to the one in this paper but uses different raw source data for land valuation.

Figure 8: Tract-Level Comparisons

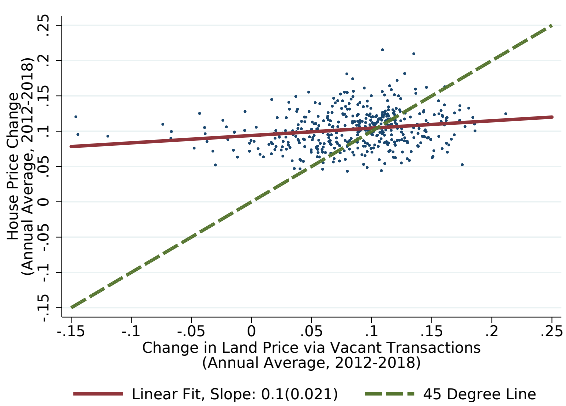
(a) Land Values (As-is), Vacant Transactions vs Land Under Structures, 2018



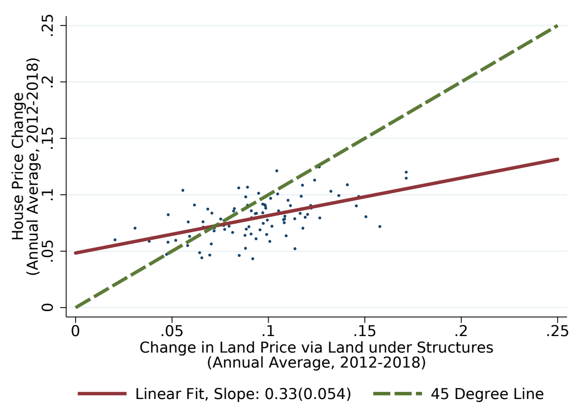
(b) Land Value Changes: Vacant Transactions vs Land Under Structures, 2012-2018



(c) Land Value (via Vacant Transactions) vs House Price Changes, 2012-2018

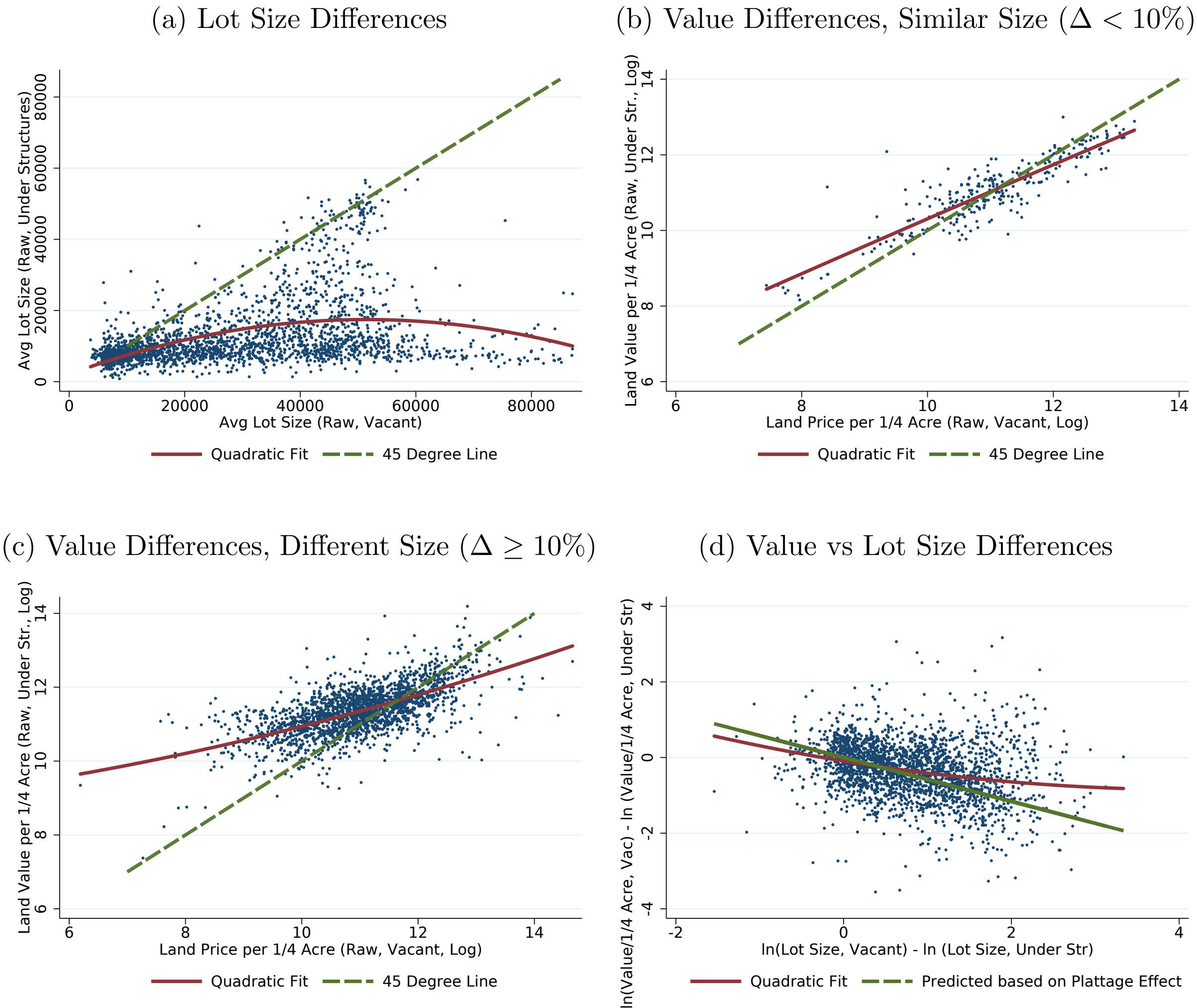


(d) Land Value (via Land under Structures) vs House Price Changes, 2012-2018



Notes: Land Value (via Vacant Transactions) is the land values estimated using vacant transactions in Section 3 of this paper. Land Value (via Land Under Structures) is taken from Davis et al. (2021), which uses similar methods to Section 3 applied to land underneath structures. House Price Changes are based on tract-level annual repeat-sales house price indices from Bogin et al. (2019).

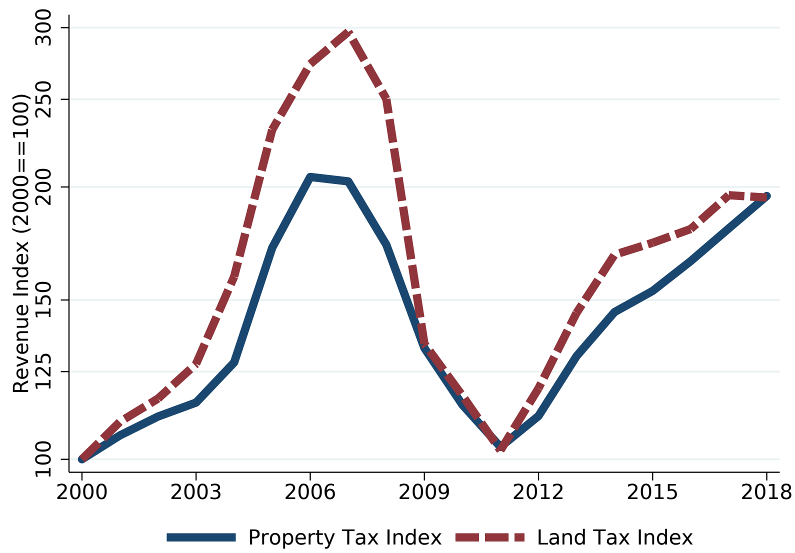
Figure 9: Raw Characteristics of Tract-Level Averages for Residential Land– Vacant Land versus Land Under Structures



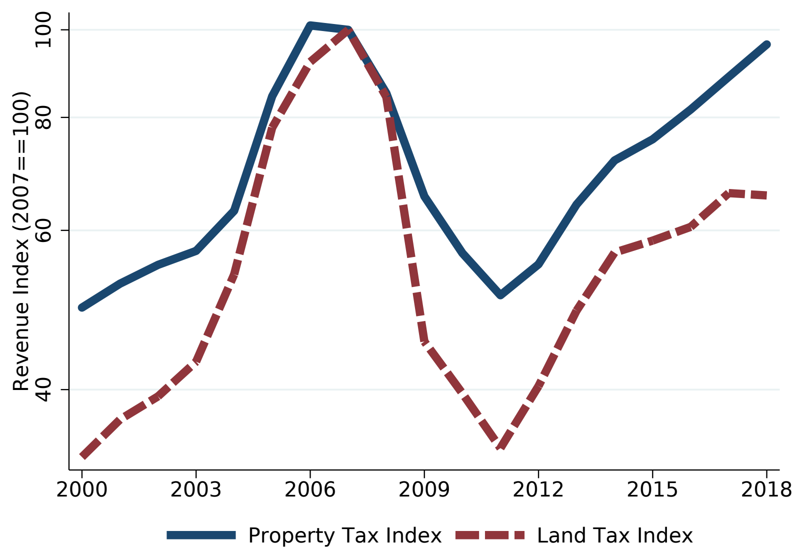
Notes: Scatterplots present annual tract-level values from the respective raw source datasets for each year between 2012 and 2018. The source of the raw land values and lot sizes for vacant land is the Maricopa County vacant transactions database. The source of the raw land values and lot sizes for land underneath structures is the source data described in Davis et al. (2021) as the subset of all mortgage appraisals with a low effective age where the value is not anchored to a tax assessed value.

Figure 10: Tax Revenue Indices

(a) Base 2000



(b) Base 2007



Notes: Both figures use identical data; the only difference is the base year (2000 vs 2007). The land tax revenue index is the land value index calculated in this paper under two assumptions: a constant composition of parcels based in 2007, and a tax rate. The property tax revenue index is an indexed version of the county-level annual repeat-sales house price index from Bogin et al. (2019).

Table 1: Summary Statistics (part 1 of 2)

	Parcel Database				Land Transactions			Adj. Parameters	
	2007	2018	Δ	%Δ	2000-2018	Δ vs 2007 parcels	%Δ vs 2007 parcels	Beta	SE
Count	1,521,892	1,615,476	93584	6.15%	96,154				
Aggregated land area covered (sqft)	72,659,433,007	79,374,342,858	6,714,909,851	9.24%	5,652,698,457	-67,006,734,550	-92.22%		
Median lot size(sqft)	7,444	7,427	-17	-0.23%	46,683	39239	527.12%		
Average lot size (sqft)	47,745	49,134	1389	2.91%	82,003	34259	71.75%	0.58	0.03
Parcel located on an arterial road	9.25%	9.35%	0.10%	1.09%	15.27%	6.02%	65.12%	0.02	0.01
Corner parcel	8.43%	9.24%	0.81%	9.59%	14.14%	5.71%	67.71%	0.20	0.01
Cul-de-sac parcel	3.01%	3.34%	0.33%	11.07%	8.94%	5.93%	196.87%	0.25	0.01
Parcel has freeway access	0.14%	0.15%	0.01%	8.65%	0.23%	0.09%	63.79%	0.33	0.06
Parcel located in a gated community	3.90%	4.56%	0.66%	16.84%	8.57%	4.67%	119.82%	0.31	0.01
Parcel located on a golf course	2.95%	3.03%	0.09%	3.05%	3.59%	0.64%	21.79%	0.63	0.02
Parcel located on a greenbelt	2.43%	4.86%	2.43%	99.87%	4.02%	1.59%	65.42%	0.22	0.02
Parcel located on a lake	0.67%	0.71%	0.04%	5.36%	0.37%	-0.30%	-44.30%	0.43	0.05
Parcel located on a major intersection	0.66%	0.53%	-0.13%	-19.78%	1.13%	0.48%	72.25%	0.49	0.03
Parcel located on a mountain	0.75%	0.68%	-0.06%	-8.67%	4.23%	3.49%	466.46%	0.18	0.02
Parcel is a pad site	0.14%	0.25%	0.11%	79.02%	0.95%	0.81%	577.59%	0.85	0.03
Parcel has a premium view	1.01%	0.90%	-0.11%	-10.55%	3.63%	2.62%	260.73%	0.03	0.02
Parcel located on a preserve	0.13%	0.34%	0.20%	153.24%	0.85%	0.71%	537.64%	0.12	0.03
Parcel located on a railroad line	0.11%	0.15%	0.03%	28.97%	0.12%	0.01%	9.44%	-0.10	0.09
Parcel has direct railroad access	0.04%	0.04%	0.00%	-1.93%	0.03%	-0.01%	-23.42%	-0.39	0.16
Parcel adjacent to an apartment/multi-family complex	0.06%	0.39%	0.34%	600.53%	0.08%	0.02%	39.16%	0.28	0.11
Parcel adjacent to commercial/industrual property	0.16%	0.98%	0.83%	523.21%	0.26%	0.10%	65.39%	0.10	0.06
Parcel adjacent to transmission line	0.23%	0.56%	0.33%	145.05%	0.56%	0.33%	144.69%	-0.45	0.04
Parcel adjacent to a waterway	0.25%	0.27%	0.02%	8.01%	0.41%	0.16%	63.66%	0.19	0.05
Parcel not accessible via a road	1.64%	2.10%	0.46%	27.95%	4.53%	2.89%	176.21%	-1.24	0.02
Parcel accessible via an unpaved road	2.94%	3.33%	0.39%	13.26%	23.11%	20.17%	685.80%	-0.77	0.01
Parcel accessible via a paved road	36.30%	35.97%	-0.33%	-0.91%	55.75%	19.46%	53.61%	-0.12	0.01
Parcel located in an air park	0.03%	0.04%	0.01%	19.90%	0.12%	0.09%	289.55%	0.01	0.09

Table 1: Summary Statistics (part 2 of 2)

	Parcel Database				Land Transactions			Adj. Parameters	
	2007	2018	Δ	%Δ	2000-2018	Δ vs 2007 parcels	%Δ vs 2007 parcels	Beta	SE
Agricultural Zoning District	0.37%	0.30%	-0.07%	-18.05%	0.86%	0.49%	132.87%	0.00	0.00
Convenience Commercial District	1.84%	1.42%	-0.42%	-22.90%	2.67%	0.83%	45.29%	-1.76	0.41
General Commercial District	0.59%	0.70%	0.11%	18.53%	1.35%	0.76%	128.83%	-1.92	0.41
Neighborhood Commercial Zoning District	0.25%	0.26%	0.01%	2.38%	0.53%	0.28%	110.08%	-1.82	0.41
Other Commercial Zoning District	0.94%	0.96%	0.02%	2.00%	1.14%	0.20%	21.32%	-1.37	0.41
Planned Industrial Zoning District	0.97%	0.89%	-0.08%	-7.91%	1.91%	0.94%	96.65%	-2.40	0.43
Light Industrial Zoning District	0.38%	0.30%	-0.08%	-21.11%	0.52%	0.14%	37.04%	-2.48	0.43
Other Industrial Zoning District	0.23%	0.23%	0.00%	-0.89%	0.58%	0.35%	148.16%	-1.94	0.43
Manufactured Housing Residential Zoning District	6.03%	7.92%	1.89%	31.34%	2.58%	-3.45%	-57.25%	-1.35	0.39
Multi-household Zoning Districts	6.17%	5.32%	-0.85%	-13.81%	2.83%	-3.34%	-54.15%	-1.72	0.39
Unknown or Other type of Zoning District	7.98%	6.52%	-1.46%	-18.28%	2.67%	-5.31%	-66.53%	-0.01	0.42
Planned Area Development Overlay Zoning District	14.47%	16.80%	2.33%	16.10%	10.08%	-4.39%	-30.36%	-0.98	0.39
SF Residential 6,000 sqft per Dwelling Unit	25.09%	24.35%	-0.74%	-2.95%	4.75%	-20.34%	-81.06%	-0.75	0.39
SF Residential 7,000 sqft per Dwelling Unit	6.80%	6.76%	-0.04%	-0.59%	1.34%	-5.46%	-80.31%	-1.06	0.40
SF Residential 8,000 sqft per Dwelling Unit	5.63%	5.61%	-0.02%	-0.28%	2.03%	-3.60%	-63.96%	-1.02	0.39
SF Residential 10,000 sqft per Dwelling Unit	4.45%	4.36%	-0.09%	-2.11%	4.49%	0.04%	0.84%	-2.60	0.40
SF Residential 15,000 sqft per Dwelling Unit	0.54%	0.62%	0.08%	14.80%	1.48%	0.94%	172.04%	-1.31	0.40
SF Residential 18,000 sqft per Dwelling Unit	1.34%	1.36%	0.02%	1.17%	3.94%	2.60%	193.67%	-1.89	0.40
SF Residential 35,000 sqft per Dwelling Unit	1.67%	1.64%	-0.04%	-2.22%	8.67%	7.00%	418.63%	-1.23	0.40
Rural Zoning District – One Acre per Dwelling Unit	5.61%	5.17%	-0.43%	-7.72%	35.75%	30.14%	537.58%	-1.34	0.39
Rural Zoning District – 70,000 sqft per Dwelling Unit	0.17%	0.16%	0.00%	-1.32%	0.89%	0.72%	435.31%	-0.84	0.40
Rural Zoning District – 190,000 sqft per Dwelling Unit	0.76%	0.69%	-0.07%	-9.80%	4.06%	3.30%	431.40%	-1.04	0.40
Rural Agricultural Zoning – One Acre Per Dwelling Unit	1.24%	0.82%	-0.42%	-33.89%	0.66%	-0.57%	-46.45%	-1.10	0.39
Other SF Residential Zoning	6.06%	6.38%	0.32%	5.35%	3.47%	-2.59%	-42.82%	-1.19	0.40
Residential Townhouse	0.40%	0.44%	0.04%	10.04%	0.75%	0.35%	86.00%	-0.33	0.40

Notes: This table describes the summary statistics of three different databases—all parcels in 2007, all parcels in 2018, and all land transactions between 2000 and 2018.

Table 2: Interpolation RMSE (20% hold-out sample)

		Method	Null	Null	IDW	Krig	Krig	Krig	Krig	Krig
		Neighbors	All	50	50	25	50	75	50	50
		Function	-	-	-	Sphere	Sphere	Sphere	Exponent	Sphere
		Grid	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.01
Year	Training	Hold-Out	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
2000	5928	1186	0.88	0.63	0.56	0.56	0.56	0.56	0.56	0.63
2001	5027	1006	0.89	0.64	0.56	0.57	0.57	0.57	0.56	0.65
2002	5576	1116	0.87	0.65	0.58	0.58	0.58	0.58	0.57	0.61
2003	6442	1289	0.89	0.65	0.59	0.59	0.59	0.59	0.59	0.54
2004	9237	1848	0.95	0.60	0.53	0.54	0.54	0.53	0.53	0.61
2005	10631	2127	0.98	0.69	0.62	0.64	0.63	0.64	0.63	0.64
2006	5300	1060	0.87	0.67	0.60	0.58	0.58	0.58	0.58	0.61
2007	3277	656	0.88	0.75	0.68	0.67	0.67	0.67	0.67	0.67
2008	1351	271	0.94	0.81	0.74	0.73	0.73	0.73	0.72	0.79
2009	1621	325	1.00	0.85	0.81	0.81	0.80	0.80	0.80	0.75
2010	2087	418	1.02	0.80	0.74	0.74	0.74	0.74	0.73	0.75
2011	1859	372	1.04	0.79	0.71	0.73	0.73	0.73	0.72	0.77
2012	2472	495	1.03	0.76	0.69	0.69	0.70	0.70	0.69	0.71
2013	2777	556	1.00	0.76	0.77	0.75	0.75	0.75	0.75	0.76
2014	2530	506	1.06	0.76	0.71	0.71	0.71	0.71	0.71	0.63
2015	2840	568	1.06	0.72	0.68	0.69	0.69	0.69	0.68	0.72
2016	3130	626	1.04	0.74	0.66	0.69	0.69	0.69	0.68	0.73
2017	3556	712	1.01	0.69	0.64	0.64	0.64	0.64	0.64	0.67
2018	3359	672	1.02	0.76	0.66	0.67	0.67	0.67	0.67	0.68
Mean	4158	832	0.97	0.72	0.66	0.66	0.66	0.66	0.66	0.68

Notes: Interpolation RMSE calculated as follows: 1) Estimate spatial grid using training sample. 2) Estimate an interpolated estimate for each hold-out grid point for each year. 3) Calculate the RMSE over all hold-out values.

Table 3: Single-Family Residential Land Source Comparison

	Single Family Residential Parcels			
	All Parcels, 2018	Land Trans., 2000-2018	Land Trans., 2018 only	Land under Str. (2012-2018)
Count	1,330,502	61,222	2,382	940,446
Average Lot Size (sqft)	10,834	36,872	35,570	10,311
Aggregated land area covered	14,414,130,547	1,576,913,575	81,309,800	9,697,122,385
Average Price per quarter-acre	105,838	175,041	180,652	101,787

Notes: This table presents the single-family residential subset of the estimates from Section in columns 1, 2, and 3. Column 4 is calculated from public estimates found in Davis et al. (2021).