Online Appendix for
The Price of Residential Land for Counties, ZIP Codes, and Census Tracts in the United States

Morris A. Davis[a,*], William D. Larson[b], Stephen D. Oliner[c], Jessica Shui[b]

[a]*Rutgers Business School, 1 Washington Park Room 627, Newark, NJ 07102*
[b]*Federal Housing Finance Agency, 400 7th Street SW, Washington, DC 20219*
[c]*American Enterprise Institute, 1789 Massachusetts Ave, NW, Washington, DC 20036*

Appendix A. Spatial Interpolation Methods

This section describes various interpolation methods we compare in the paper. Our land price estimation approach begins by dividing the universe of parcels N within a geography into those with values that are observed (i.e. included in our working data) and those with values that are not observed. For convenience, below we denote the number of observed values in a geography as N^s. All land prices we report are calculated as the simple average of the price estimate (or actual value, when available) of each individual parcel within the county or county subaggregate (i.e. census tract). Accordingly, differences between land values within a given geography arise due the method used to estimate prices for parcels with unobserved values.

At the end of the day, we estimate the value of land when land value is not directly observed as a weighted average of land values of a subset of parcels within the county where land prices are directly observed. The estimated price for parcel i is calculated as the average of the n nearest (by proximity) neighbor parcels, indexed by j, with weights $\lambda_{i,j}$.[1] For a particular method, the estimated price for parcel i is:

$$\hat{p}_i = \sum_{j=1}^{n} \lambda_{i,j} p_{i,j} \tag{A.1}$$

The weights are assumed to sum to unity, or $\sum \lambda_{i,j} = 1$.

Appendix A.1. Spatial Statistics

Spatial statistical methods do not consider spatial relations in outcomes, only proximity. Here, we discuss three spatial statistics that are commonly used to interpolate prices spatially.

*Corresponding author
Email addresses: mdavis@business.rutgers.edu (Morris A. Davis), william.larson@fhfa.gov (William D. Larson), stephen.oliner@aei.org (Stephen D. Oliner), jessica.shui@fhfa.gov (Jessica Shui)

[1]So $p_{i,1}$ is the closest observed price to location i; $p_{i,2}$ is the second closest observed price to location i; and so on.

The general form of each of these statistics is below, where h is the distance between the location to be imputed, i, and another nearby sampled location, j, that is one of the n nearest locations. The exponent c gives the degree of decay in the weight that is due to distance between the parcels.

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

Null Estimator

The null estimator ("Null") sets n=N^s and $c = 0$, giving $\lambda_{i,j} = 1/n$. This gives the estimate of an individual parcel as the sample average.

Nearest Neighbor

The nearest neighbor estimator ("NN") also sets $c = 0$, giving $\lambda_{i,j} = 1/n$. But n is typically set to the 5 to 25 nearest observed prices. This gives the estimate of a parcel as the sample average of nearby parcels.

Inverse-distance weights

The inverse-distance weight estimator ("IDW") sets n generally within the range of 5 to 25 nearest observed prices as with the NN estimator. The calculation of λ then involves assuming an exponent c. This exponent is commonly set equal to $c = 2$, giving a relation between points that declines with the square of the distance. In this case, $\lambda_{i,j} = \frac{1}{h_{i,j}^2} \left(\sum_{j=1}^{n} \frac{1}{h_{i,j}^2} \right)^{-1}$

Appendix A.2. Geostatistics

In addition to spatial statistics, we include a single geostatistical estimator: ordinary Kriging. Ordinary Kriging is our preferred Kriging estimator because we eliminate known variation across parcels by way of standardization. In effect, we are performing a two-step regression-Kriging estimator, which can also be interpreted as estimating a rudimentary hedonic and then Kriging the residual.

As with the nearest neighbor estimator, n nearest neighbors are weighted and summed to generate predicted prices. The key difference with Kriging is λ is estimated based on the strength of the observed relationships between observations of different proximities within the sample such that it is the best linear unbiased predictor.

Calculation of the weights proceeds in four steps. The first step involves calculating pairwise differences in values between each pair in the sample within a certain distance range. The next step collapses and bins the semivariances (half of the squared differences) into averages by the distance between the points. The third step fits a curve, referred to as a "variogram," to this set of binned averages. The fourth step uses this curve to estimate semivariances between any two points and then construct the weights.[2]. Once the weights are known, predictions and prediction error variances can be readily calculated for any location.

As an illustration of this procedure, we present kriging steps for land prices in Washington, DC, pooled between 2012 and 2019. There are about 110,000 parcels and 16,000

[2]For a more in-depth overview of kriging, see Hengl (2007) or Sherman (2011).

sampled standardized land prices. To start, differences and semivariances γ are calculated for each pair of points in N.[3]

The results from the first two kriging steps are shown in Figure A.1. The hollow circles represent semivariance averages within each of the 15 distance bins. Distances are reported as miles but are in terms of latitude/longitude degrees in the programs.[4]

In step three, a functional form for the relationship between the semivariances and the distances (the hollow circles) is assumed and fit to the data. The fitted curve, as shown by the blue line, is typically upward sloping, indicating the greater the distance the higher the variance. The spherical functional form has three parameters, a_0, a_1, and r that we estimate

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

The three parameters combine to give the "sill" which is the value to which the variogram asymptotically approaches as the distance between points approaches infinity, or a_0+a_1; the "nugget" which is the value of the variogram when distance approaches zero, or a_0, and the "range," r, which is the value of h when the variogram reaches the sill. Figure A.1 shows a fitted spherical functional form for Washington, DC, with $\hat{a}_0=0.06$, $\hat{a}_1=0.55$ and $\hat{r}=9.18$ miles. This function is used to estimate the semivariance between any two points given a distance between them.[5]

In the fourth step, define a linear system that gives the best linear predictor of the weights λ for an unsampled location i. γ is an $n\times 1$ vector of estimated semivariances between i and its n nearest points, indexed by j (Sherman, 2011). $\mathbf{\Gamma}$ is an $n\times n$ matrix of semivariances between these n nearest points. These matrices are augmented in the standard fashion with a Lagrange multiplier and column/row vectors of ones and a zero to normalize the weights to sum to one. These give the vector of weights λ_i, an $n\times 1$ vector.

$$\begin{bmatrix} \lambda_\mathbf{i} \\ \mathcal{L} \end{bmatrix} = \begin{bmatrix} \mathbf{\Gamma_i} & \mathbf{1}_n \\ \mathbf{1}'_n & 0 \end{bmatrix}^{-1} \begin{bmatrix} \gamma_\mathbf{i} \\ 1 \end{bmatrix} \tag{A.4}$$

Then, the predicted value for location i is $\hat{p}_i=\lambda_\mathbf{i}'p_{i,j}$ and the prediction error variance is $\hat{V}_{p_i}=\lambda_\mathbf{i}'\gamma_i$. The prediction error variance is relevant because it is used in the Jensen's inequality correction for log prices, $\hat{P}_i=\exp(\hat{p}_i+\frac{1}{2}\hat{V}_{p_i})$.

Appendix A.3. Comparison

We compare the fit of each spatial interpolation method by comparing actual (log) standardized land prices for a 20% hold-out sample to predicted (log) land prices estimated using

[3] The semivariance for prices at two points i and j is half of the squared difference, $\frac{1}{2}(p_i-p_j)^2$. Isotropy (i.e. the direction between the points does not affect the strength of the relationship) is a standard assumption, which we make here in order to express proximity using a single variable, h.

[4] In other words, we take the square root of the sum of the squared differences between two sets of coordinates. Since our distances are generally small, we use this simplified distance measure as a proxy for the actual distance, which varies due to changes in latitude.

[5] Other functional forms are common, especially the exponential function, $\gamma(h;a_0,a_1,r)=a_0+a_1(1-\exp(-\frac{h}{r}))$. We consider the exponential function as well but the spherical functional form has a slightly better fit, on average.

an 80% training sample. We consider a number of values for the number of nearest neighbors and the overall distance boundary considered.

Table A.1 reports, by year, median RMSEs across counties for Kriging and the other spatial interpolation methods we consider for the hold-out sample. The table shows that kriging provides the greatest interpolation accuracy in every sample year. Table A.2 provides a bit more color for these results in terms of the mean, median, and standard deviation of RMSEs across the 2,378 counties in the pooled sample. The table shows that each of the kriging estimates at the various parameterizations shown yields similar average fit, with means RMSEs ranging from 0.393 to 0.395. We interpret this finding as indicating that nether the boundary nor the number of nearest neighbors considered tends to affect the estimates at the ranges considered. Overall, we interpret these findings as lending support to our decision to use the 20 nearest neighbors and a 6.9 mile boundary in our county-specific kriging procedure.

We also evaluate the accuracy of the Kriging procedure for our application by seeing if it can recover land values that endogenously arise from a simple rendition of the standard monocentric city model that we can compute analytically. We simulate two data sets from this model, one in which land values are measured perfectly and one in which land values are measured with error and then check the extent to which Kriging can replicate the analytic gradient of land prices. In the data set in which land values are perfectly measured, Kriging nearly exactly replicates the analytic gradient. In the data set in which land values are measured with error, Kriging produces relatively small average errors. This is described in the next section.

Appendix B. Monte Carlo Simulation of Standard Urban Model

Assume that a city lies on a featureless plane with a region called the central business district (CBD) at its center. This district provides all employment and because commuting is costly in a way we specify precisely, households wish to live near the CBD. Spatial equilibrium requires identical households to have identical utility in all locations in the city. As we show, this implies that households consume less housing at a higher price near the CBD. Additionally, the housing production function implies housing is produced with high density and structure intensity near the CBD where land prices are high, and with low density near the edge of the city.

To be precise, assume a person consuming c units of consumption and h units of housing receives utility of

$$(1-\alpha)\ln c + \alpha \ln h \tag{B.1}$$

If a person lives distance d from the city center, their wage after commuting is $w\,(1-td)$ where t is the percentage of income that must be paid to commute for each unit of distance d. Denote the rental cost per unit of housing d units from the city center as q_d^h. A person living d units from the city center faces the budget constraint of:

$$w\,(1-td) \;\; = \;\; c + q_d^h h \tag{B.2}$$

A person choosing to live in location d units away from the CBD maximizes utility (B.1) subject to the budget constraint (B.2) by choosing optimal consumption c_d and housing h_d

of

$$\begin{aligned} c_d &= (1-\alpha)\, w\, (1-td) && \text{(B.3)} \\ q_d^h h_d &= \alpha w\, (1-td) && \text{(B.4)} \end{aligned}$$

This means maximized utility at distance d from the center can be written as $\mathcal{U}_d$

$$\begin{aligned} \mathcal{U}_d &= (1-\alpha)\ln\left[(1-\alpha)\, w\, (1-td)\right] + \alpha \ln\left[\alpha w\, (1-td)\, / q_d^h\right] \\ &= \kappa_u + \ln w + \ln(1-td) - \alpha \ln q_d^h && \text{(B.5)} \end{aligned}$$

where κ_u is a constant equal to $\alpha \ln \alpha + (1-\alpha)\ln(1-\alpha)$. In equilibrium, we assume all locations have to provide the same utility, for example location d and d' must satisfy

$$\mathcal{U}_d = \mathcal{U}_{d'}$$

Then from equation (B.5) this implies

$$\begin{aligned} \ln(1-td) - \alpha \ln q_d &= \ln(1-td') - \alpha \ln q_{d'}^h \\ \frac{q_{d'}^h}{q_d^h} &= \left(\frac{1-td'}{1-td}\right)^{\frac{1}{\alpha}} && \text{(B.6)} \end{aligned}$$

Equation (B.6) governs the rate at which housing rental prices per unit change with distance from the CBD, roughly t/α percent per unit of d.

Note that we can also work out how the quantity of housing changes as a function of distance to the CBD. We start by using the definition of utility and substituting in optimal consumption but keeping housing

$$\mathcal{U}_d = (1-\alpha)\ln\left[(1-\alpha)\, w\, (1-td)\right] + \alpha \ln h_d \qquad \text{(B.7)}$$

Once we impose $\mathcal{U}_d = \mathcal{U}_{d'}$, this gives us

$$\begin{aligned} (1-\alpha)\ln(1-td) + \alpha \ln h_d &= (1-\alpha)\ln(1-td') + \alpha \ln h_{d'} \\ \rightarrow \frac{h_{d'}}{h_d} &= \left(\frac{1-td'}{1-td}\right)^{-\frac{1-\alpha}{\alpha}} && \text{(B.8)} \end{aligned}$$

Now that we have worked out how housing quantities h and prices per unit q_d^h vary from the city center, we can also work out how the quantities and prices of land and structures change with distance. Temporarily suppressing the distance subscripts, assume competitive builders build housing using land l and structures s according to a CES production function

$$h = \left[(1-\theta)\, s^\rho + \theta l^\rho\right]^{\frac{1}{\rho}} \qquad \text{(B.9)}$$

with $\rho \in (-\infty, 1]$. Assume each unit of housing generates revenue of q^h; further, assume each unit of land costs q^l and each unit of structure costs 1. Builders maximize

$$q^h \left[(1-\theta)\, s^\rho + \theta l^\rho\right]^{\frac{1}{\rho}} - s - q^l l \qquad \text{(B.10)}$$

The first-order conditions for optimal structures are

$$1 = q^h h^{1-p} (1-\theta) s^{p-1} \tag{B.11}$$

$$\rightarrow s = \left[q^h (1-\theta)\right]^{\frac{1}{1-\rho}} h \tag{B.12}$$

This implies that once we know q^h and h, we also know s. Note that because we know s, we also know $q^l l = q^h h - s$. Now consider the first-order condition for optimal land:

$$q^l l = q^h h^{1-p} \theta l^p \tag{B.13}$$

and thus

$$l = \left[\frac{q^l l}{q^h h^{1-\rho} \theta}\right]^{\frac{1}{\rho}} \tag{B.14}$$

Given a set of parameters, we can compute how quantities and prices and expenditures on housing, structures and land change with distance from a CBD. For a rough calibration, we set $\alpha = 0.25$ based on the median housing budget shares of renters as documented by Davis and Ortalo-Magné (2011). The other parameters we set to match some approximate features of a city. We set $t = 0.02$ such that people 10 miles from the CBD consume about double the housing than people at the CBD but spend 20% less.[6] We jointly set $\theta = 0.90$ and $\rho = -2.0$ such that land's share of value rises from about 15% 10 miles from the CBD to about 55% at the CBD. We normalize the price per unit of housing to 1 at the CBD and normalize the quantity of housing consumed at the CBD such that the total expenditure at the CBD is for a \$1 million house. As noted earlier, we assume the price per unit of housing structure is 1.0 everywhere in the metro area. Table A.3 shows prices, quantities and expenditures on housing, structures and land as well as land's share of house value, the quantity of land once we normalize the size of a single-family plot at the CBD to 0.25 acres, and land price per acre.

We simulate two data sets based on the calibration of this model. In both data sets, we draw 100 observations for houses in neighborhoods uniformly between 0 and 3.5 miles from the CBD; 200 observations for houses in neighborhoods uniformly between 3.5 and 7.5 miles from the CBD; and 300 observations for houses in neighborhoods uniformly between 7.5 miles and 10 miles from the CBD. In the first data set we assume no quantities or prices are measured with error. This enables us to see the accuracy of the Kriging procedure with regards to this application in an ideal environment.

In the second data set, we allow for i.i.d. measurement error in both the value of housing and the value of structures.[7] This simulation gives us some intuition for how the Kriging procedure performs under conditions where land is imperfectly measured. Denote $\widetilde{q_d^h h_d}$ as observed housing value and $\widetilde{s}$ as observed structures costs. $\widetilde{q_d^h h_d}$ and $\widetilde{s}$ are determined as

$$\begin{aligned} \widetilde{q_d^h h_d} &= q_d^h h_d \left(1 + e_d^h\right) \qquad & e_d^h &\sim U[-0.10, 0.10] \\ \widetilde{s} &= s\left(1 + e_d^s\right) \qquad & e_d^s &\sim U[-0.10, 0.10] \end{aligned}$$

[6]This is referring to single-family homes.

[7]This measurement error can be thought of as a deviation from model-determined prices.

with e_d^h and e_d^s drawn independently. We then compute observed land value residually,

$$\begin{aligned}\widetilde{q_d^l l_d} &= \widetilde{q_d^h h_d} - \widetilde{s} \\ &= q_d^h h_d - s \quad + \quad \left(q_d^h h_d e_d^h - s e_d^s\right)\end{aligned}$$

Denote the term in parentheses as e_d^l, measurement error in land value. Even though the standard deviation of e_d^h and e_d^s are relatively small (5.7 percent each), the standard deviation of measurement error as a percent of true land value, measured as $e_d^l / \left(q_d^l l_d\right)$, in this second data set is much larger, 27 percent. The measurement error for land value is magnified because land value is residually measured and accounts for a relatively small fraction of home value (as we discuss earlier in the paper).

Table A.4 compares the estimates from Kriging for land price per acre by distance to CBD to the numbers we compute analytically in Table A.3 for the simulated data set measured without error (data set 1) and the simulated data set with measurement error (data set 2) for 0 to 9 miles to the CBD. When the data are measured with error, the Kriging results are less accurate, though the average error across the distance bins is relatively small at 4.2%.

Appendix C. Comparison of CBSA Results to Davis and Palumbo (2008)

Davis and Palumbo (2008) created the first dataset of publicly available land prices and land shares for residential land in a large number of large metropolitan areas. The data for the published article span the time period 1984-2004, but the authors have routinely updated the data and the publicly-available data are now available through 2018.[8] In this Appendix we compare land prices and land shares for the CBSAs and time period in which our data overlap with Davis and Palumbo's data. To facilitate this comparison, in the Davis and Palumbo data we merge Anaheim with Los Angeles, Oakland with San Francisco, and Dallas with Fort Worth to reduce the number of metro areas from 46 to 43 MSAs. Our data and the Davis and Palumbo data overlap for 2012-2018, so we compute average land prices and land shares over this period in each data set and compare the averages.

Broadly speaking, the data align very well. The correlation is 0.95 for the level of land prices and 0.87 for the land shares. However, variation in land-price levels and land shares at the metro area level is less extreme in our data than in the Davis and Palumbo data. To see this, in Figure A.2 we show scatter plots of log land prices (top panel) and land shares (bottom panel) along with regression lines. The graph comparing log land prices shows that, in the cross section of metro areas, our log land price increases by 0.73 percent for each one percent increase in the Davis and Palumbo data. Similarly, the graph comparing land shares shows that in the cross section of metro areas, our estimates of land share increases by 0.55 percentage point for each percentage point increase in the Davis and Palumbo data.

Ultimately, we believe the data in our paper are more accurate for three reasons: First, our sample sizes are orders of magnitude larger. For example, the maximum metro-area sample size in the Davis and Palumbo data is 2,513 in Salt Lake City (Table 2 of that paper). Second, Davis and Palumbo include all observations, including old homes. We drop old homes from the sample to avoid the problem of using construction cost as a proxy

[8]These data are currently available at *https://www.aei.org/historical-land-price-indicators/*.

for structure value for these homes. Finally, the geography of the specific counties that constitute a given metro area is unclear in Davis and Palumbo. They use whatever data appear in the Metropolitan-AHS survey files, and the documentation of those files is not clear. In our paper, we know exactly the counties comprising each metro area.[9]

References

Davis, M.A., Ortalo-Magné, F., 2011. Household expenditures, wages, rents. Review of Economic Dynamics 14, 248–261.

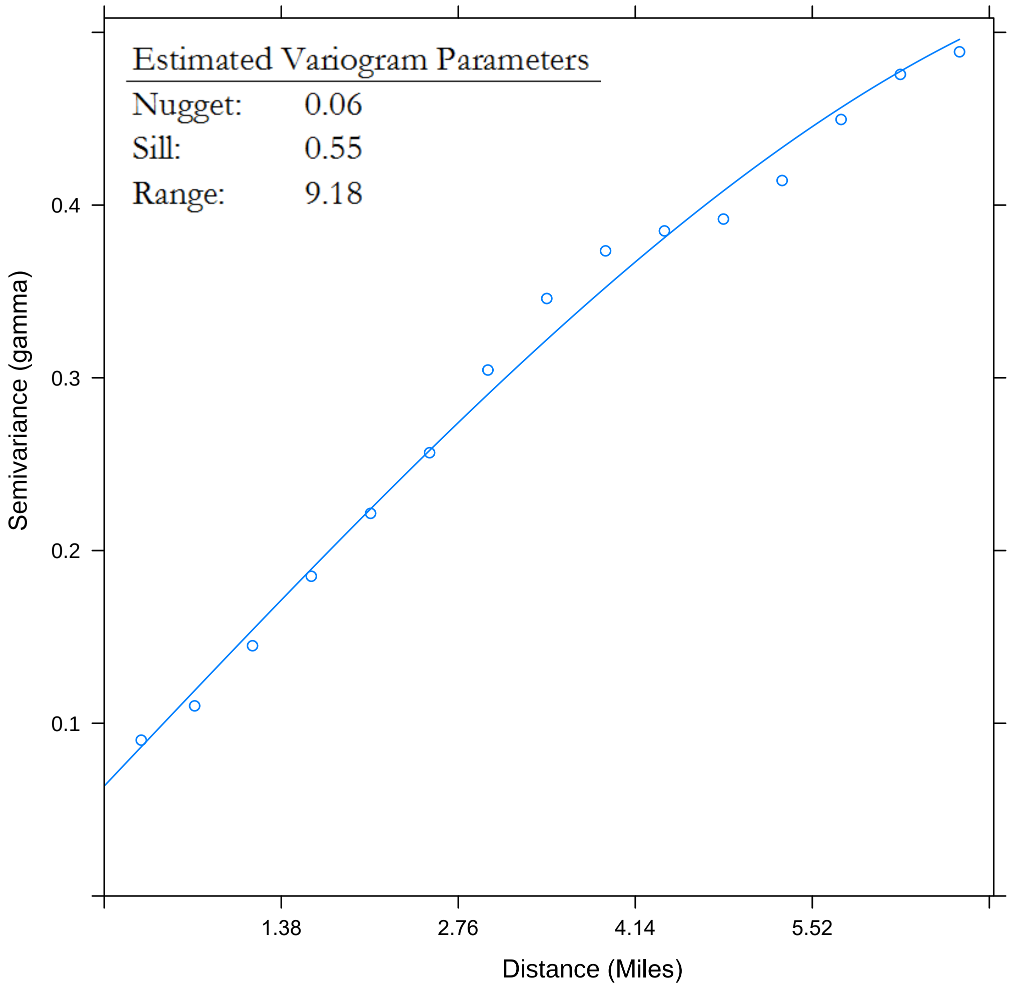
Davis, M.A., Palumbo, M.G., 2008. The price of residential land in large U.S. cities. Journal of Urban Economics 63, 352–384.

Hengl, T., 2007. A Practical Guide to Geostatistical Mapping of Environmental Variables. Luxembourg: Office for Official Publications of the European Communities.

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

[9]For these reasons, we suspect the statistically significant intercepts and slope coefficients less than 1.0 when we regress our data on the Davis and Palumbo data arise from attenuation bias due to measurement error in their data.

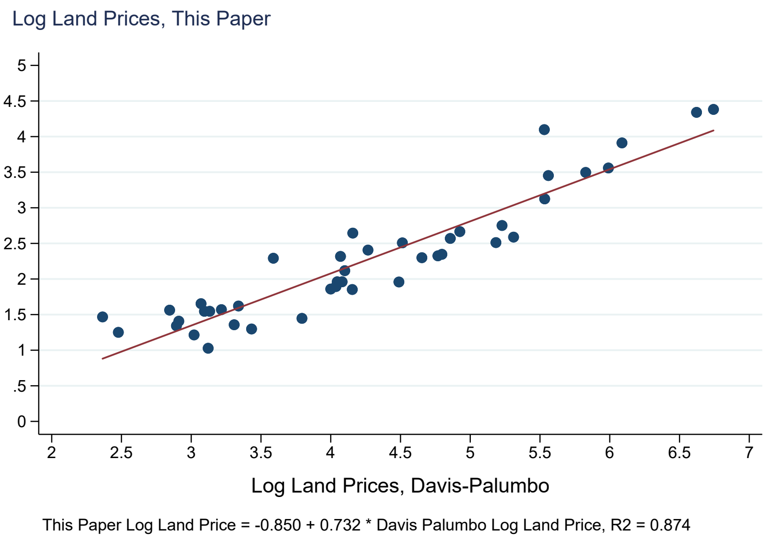
Figure A.1: Variogram for Washington, DC



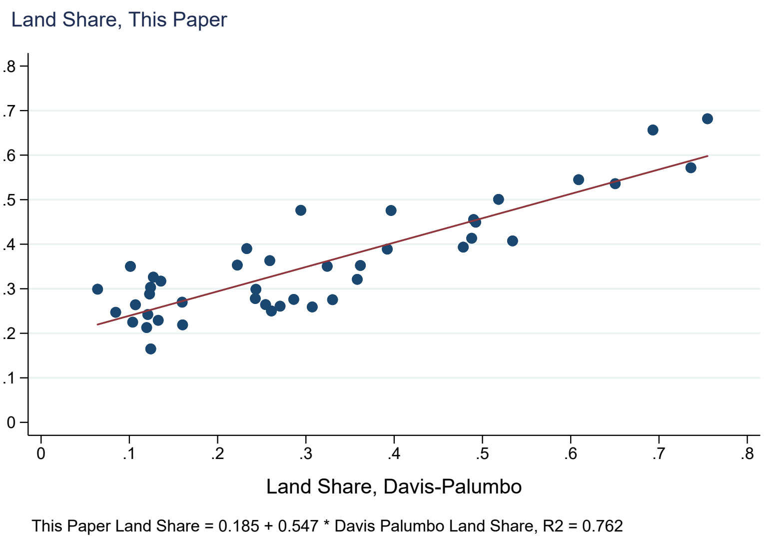
Notes: This figure presents the binned semivariances (hollow circles) and fitted variogram (blue line) for the District of Columbia. The variogram here is estimated for all parcel pairs extending to 6.9 miles.

Figure A.2: Comparison of CBSA-Level Land Prices and Shares to Davis and Palumbo (2008)

(a) Land Prices



(b) Land Shares



Notes: Panel (a) compares the log of land prices from Davis and Palumbo (2008) to the log of land prices from this paper for 43 CBSAs over the 2012-2018 period. For both data sets, the log of the average value of land prices over 2012-2018 is reported. Panel (b) compares average land shares over 2012-2018 for the same CBSAs.

Table A.1: Interpolation RMSE (20% hold-out sample)

	Kriging	IDW	NN	Null-Tract	Null-ZIP Code	Null-County	Hold-Out Obs
2012	0.394	0.420	0.430	0.440	0.444	0.513	99,395
2013	0.389	0.403	0.421	0.430	0.442	0.522	147,759
2014	0.387	0.408	0.418	0.424	0.438	0.513	126,714
2015	0.383	0.399	0.414	0.424	0.431	0.520	166,654
2016	0.379	0.395	0.410	0.419	0.430	0.506	201,222
2017	0.386	0.401	0.418	0.425	0.437	0.513	169,185
2018	0.385	0.403	0.421	0.424	0.434	0.519	142,231
2019	0.381	0.395	0.399	0.412	0.429	0.504	167,683

Notes: Interpolation RMSE calculated as follows. 1) Estimate an interpolated estimate for each hold-out parcel for each year. 2) Calculate an RMSE for each county for each year. 3) Calculate the median RMSE across counties (reported in table). IDW = inverse-distance weights, NN = nearest neighbor.

Table A.2: Interpolation RMSE (20% hold-out sample), alternative parameterizations

Estimator	Mean	Median	SD
Null - County Average	0.514	0.576	0.228
NN - 20 NN	0.418	0.453	0.167
IDW - 20 NN	0.420	0.462	0.181
Kriging - 10 NN, 6.9 Mile Boundary	0.395	0.431	0.162
Kriging - 20 NN, 6.9 Mile Boundary	0.395	0.429	0.162
Kriging - 30 NN, 6.9 Mile Boundary	0.393	0.430	0.163
Kriging - 20 NN, 3.4 Mile Boundary	0.394	0.430	0.164
Kriging - 20 NN, 10.4 Mile Boundary	0.395	0.430	0.163

Notes: Sample is the pooled cross-section (2,378 counties). Interpolation RMSE calculated as follows. 1) Estimate an interpolated estimate for each hold-out parcel. 2) calculate an RMSE for each county for each year. 3) Calculate the median/mean/SD RMSE across counties (reported in table).

Table A.3: Predictions of Calibrated Urban Model

d	q_d^h	h_d	$q_d^h h_d$	s	q_d^l	l_d	$q_d^l l_d$	land share	l_d (acres)	q_d^l per acre
0	1.000	1,000,000	$1,000,000	$464,159	0.413	1,295,995	$535,841	54%	0.25	$2,143,364
1	0.922	1,062,482	$980,000	$480,054	0.354	1,411,219	$499,946	51%	0.27	$1,836,505
2	0.849	1,130,281	$960,000	$496,838	0.300	1,543,749	$463,162	48%	0.30	$1,555,320
3	0.781	1,203,972	$940,000	$514,581	0.251	1,697,826	$425,419	45%	0.33	$1,298,934
4	0.716	1,284,211	$920,000	$533,360	0.206	1,879,309	$386,640	42%	0.36	$1,066,527
5	0.656	1,371,742	$900,000	$553,260	0.165	2,096,587	$346,740	39%	0.40	$857,343
6	0.600	1,467,412	$880,000	$574,375	0.129	2,362,219	$305,625	35%	0.46	$670,706
7	0.547	1,572,189	$860,000	$596,810	0.098	2,696,126	$263,190	31%	0.52	$506,049
8	0.498	1,687,183	$840,000	$620,680	0.070	3,132,449	$219,320	26%	0.60	$362,959
9	0.452	1,813,671	$820,000	$646,116	0.047	3,736,426	$173,884	21%	0.72	$241,250
10	0.410	1,953,125	$800,000	$673,261	0.027	4,655,227	$126,739	16%	0.90	$141,134

This table reports model-generated values of the price-per-unit of housing q_d^h, housing quantities h_d, the value of housing $q_d^h h_d$, the value of structures s, the price-per-unit of land q_d^l, the quantity of land l_d, the value of land $q_d^l l_d$, land's share of house value, the quantity of land in acres, and the price per land per-acre as a function of the distance to CBD d.

Table A.4: Land price per acre, as predicted by the model and estimated via Kriging from two data sets

		From Kriging Procedure			
		Data Set 1		Data Set 2	
Distance	From Model	Predicted	% Error	Predicted	% Error
0	$2,143,364	$2,137,352	0.28%	$2,107,724	1.70%
1	$1,836,505	$1,836,186	0.02%	$1,762,402	4.00%
2	$1,555,320	$1,555,104	0.01%	$1,575,406	-1.30%
3	$1,298,934	$1,298,750	0.01%	$1,276,002	1.80%
4	$1,066,527	$1,066,499	0.00%	$998,657	6.40%
5	$857,343	$857,259	0.01%	$878,682	-2.50%
6	$670,706	$670,688	0.00%	$641,687	4.30%
7	$506,049	$506,015	0.01%	$450,198	11.00%
8	$362,959	$362,945	0.00%	$364,878	-0.50%
9	$241,250	$241,259	0.00%	$200,874	16.70%
Mean			0.03%		4.16%

This table reports model-generated price per-acre of land and the kriging-based estimates of that price when the model is simulated without measurement error (columns marked Data Set 1) and when the model is simulated with measurement error (columns marked Data Set 2).