Frequently Asked Questions

Experimental Dataset for the Price of Residential Land

DATASET:

https://www.fhfa.gov/PolicyProgramsResearch/Research/PaperDocuments/Land-Prices_DLOS_2019_9Oct.xlsx

FHFA WORKING PAPER 19-01:

https://www.fhfa.gov/papers/wp1901.aspx

Q1: What are the "land prices" that we are measuring?

A: There are two land price series in the dataset. Both of the land price series represent land underneath single-family structures (as opposed to vacant land or land under multifamily buildings or for some other land use type). The first land price series is called "land price (standardized)" which represents the price based on a ¼ acre lot in the specified geography. The second series is called "land price" and represents the market value of an average single-family lot within the specified geography.

Q2: What other land price measures exist?

Previous research using vacant land sales (Haughwout, Orr, and Bedoll (2008), Nichols, Oliner, and Mulhall (2013) and Albouy, Ehrlich, and Shin (2018)) have only been able to either estimate land prices on a small scale within a metro area, or on a large scale at metro-level due to data scarcity. Another thread of literature measures the price of land as the difference between house value and the replacement cost of the structure on the land. These studies have also produced metro-level land indices and the most recent study, Davis, Oliner, Pinto, and Bokka (2017), have generated the level of land prices and changes in land prices at the ZIP code group level for the Washington, DC metropolitan area. To our knowledge, our paper is the first study to produce these estimates at a fine geography (ZIP code and census tract level) for nearly the entirety of the United States.

Q3: Are these official FHFA indices?

A: No, they should be considered as experimental or developmental indices. We have released indices for various geographies and time periods. They are subject to changes in methods, geographic and time period coverage, and component data.

Q4: What data are used to construct the indices?

A: We use two main data sources. The first is the Uniform Appraisal Dataset ("UAD", see https://www.fanniemae.com/singlefamily/uniform-appraisal-dataset and http://www.freddiemac.com/singlefamily/sell/uniform_appraisal.html), from which we recover appraised land price per acre. We then employ CoreLogic's county assessor dataset as our "universe" of all single-

family units in an area. We use these two sources, along with the methods described in the paper, to calculate our land price series.

Q5: Why are certain geographies/time periods missing?

A: We require observations of newly constructed or renovated homes in the UAD. Thus, observed appraisals must be of a single-family unit, for homes with a low effective age (less than or equal to 15), and must not be nearly equal to the tax assessor value from the CoreLogic assessor database. Areas with few numbers of these sorts of transactions will have missing index values. This tends to include places with dense (i.e. multifamily) housing or extremely rural locations.

Q6: How do you create indices if there are not enough observations?

A: In this case, we do not report an index value. We have strict filters for admissible raw data points from the UAD, including limits to interior square feet, lot size, home value, and effective age.

Q7: Are these series nominal or real?

A: The annual indices are nominal. The pooled indices are based on 2013 dollars. To convert these values to real terms for particular base years, one might consider the Consumer Price Index produced by the Bureau of Labor Statistics (http://data.bls.gov/cgi-bin/srgate). The most common series used for these adjustments are the "all items" (CUUR0000SA0) and "all items less shelter" (CUUR0000SA0L2).

Q8: How might one connect demographic data (or other data sources) with this data file?

A: Each of the datasets has a numerical code that can usually be linked to Census or other demographic data. For ZIP codes we suggest using ZCTAs (ZIP Code Tabulation Areas) that are defined by the Census Bureau because a variety of statistics are collected and reported. More information can be found at https://www.census.gov/geo/reference/zctas.html. To our knowledge, the Census Bureau does not offer a crosswalk mapping but third-party providers have already created such databases. One such example is on the UDS Mapper website at http://www.udsmapper.org/zcta-crosswalk.cfm.

Q9: My area is not listed or is missing a value for a year. Can you provide it?

A: Unfortunately, data limitations (e.g. too few sales) typically prevent us from providing indices in certain locations or years. We only release information when it has been aggregated to an appropriate level. If you believe your region has been overlooked, you are welcome to let us know and we can investigate further.

Q10: Why have the land values changed across working paper versions?

As we refine the methods and data used in the construction of the measures, changes will result. The most impactful of these changes include the definition of the standardized lot area (now ¼ acre; previously the ZIP code average) and the effective age filter (now 15 years; previously 10 years). In addition, to remove the potential bias due to the land component of appraisals anchoring to the tax-assessed value of land, we exclude from our working sample any property within 2% of the assessed value for either the land component or the total value. As a result, we have a final sample of 7.9 million valid cost-approach appraisals from which we estimate land values.

Q11: How do land prices appear across geographic areas?

Prices tend to be higher in coastal areas and near large cities. Prices in rural areas tend to track agricultural land values. In general, land prices fall with distance to the Central Business District (CBD) because households are willing to pay less per unit of housing as commuting costs rise. This negative relationship is called downward slowing land price gradient. As shown in Figure 5 in the paper and also below, this pattern is clear in two Core Based Statistical Areas (CBSAs), Washington-Arlington-Alexandria, DC-VA-MD-WV and San Francisco-Oakland-Hayward, CA, and it is not unique to these two areas.

Figure 5: Land Prices, pooled cross-section (2013 prices)

(a) Washington
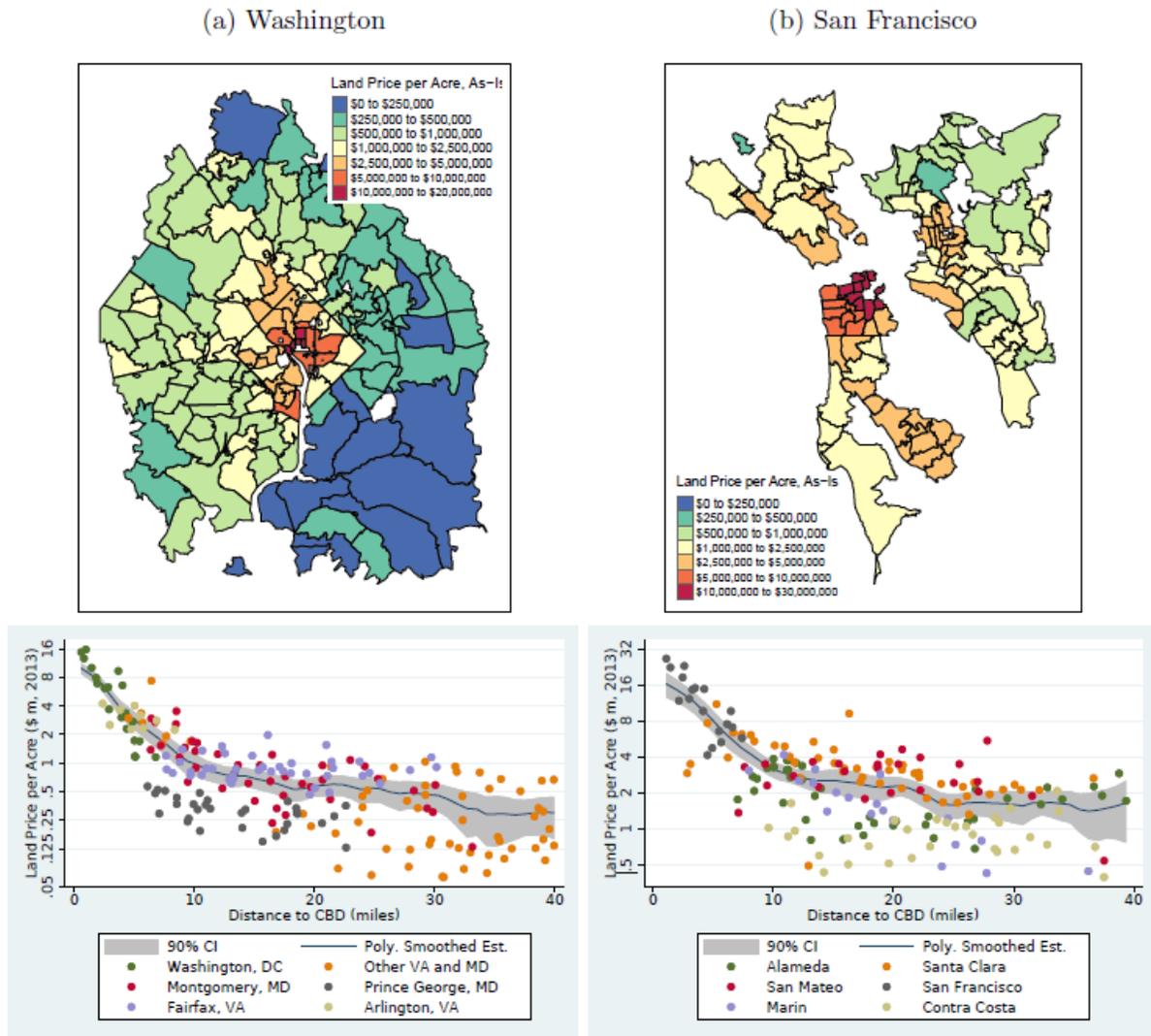
(b) San Francisco

Q12: Where do we see extremely high or low values?

Typically, urban cities with higher structure density will see the higher land values and leafy rural areas will see lower land values. The highest values can be found in New York, San Francisco, and Washington, DC. The lowest values are in rural areas of South Carolina, Georgia, and Mississippi. Specifically, for the

areas in which we compute data, land prices average about $15 million per acre in the most expensive ZIP codes in the Washington, DC metro area and about $20 million per acre in the San Francisco metro area.

Q13: Aren't appraisals anchored to other values?

Yes, Figure 3 in the paper shows the ratio of appraised value to tax-assessed value for all housing units in the U.S. that have an appraisal and a tax-assessment. From the bottom panel, it is clear that a substantial amount of appraisals anchor land value to the tax-assessed value of land. Relatedly, other papers have found that purchase-money mortgage appraisals anchor to contract price, and such phenomena is called contract price confirmation. This is not a concern in our context given the assumption that contract price reflects the market value.

Q14: Why is Manhattan so unique?

New York County, New York, has extremely high land values in the dataset and often little coverage for ZIP codes and Census tracts. The reason is that there are very few single-family homes in Manhattan. Those that remain are typically very large "brownstones" that are large in terms of interior square feet and expensive compared to other units in the dataset. They also occupy expensive land.

Q15: What do land shares tell us beyond the land price?

Land shares are defined as the ratio of the land value to the total value of the property. Land values and land shares would change in similar ways if lot sizes and home sizes were similar across space and over time. Differences between land shares and land prices therefore show (at least) three things: when land shares are extremely low, it is typically because land prices are low but the home value remains high because construction costs do not change as much across areas (e.g. materials, labor); when it is extremely high, it is typically because lot sizes land values are high and and/or home sizes cannot be optimized and changed; finally, it shows where house prices may be most volatile—abrupt changes in land shares can be a predictor of future house price volatility.

Q16: Land prices and shares are higher in the October 2020 release than earlier releases. Why?

For the October 2020 version, we implemented a needed correction when we transformed our log-prices to actual prices. See the text of the paper for details. Land prices tended to rise by 1-3%.