

Pinto, Edward J.; Oliner, Stephen D.; Davis, Morris A.

Working Paper

House prices and land prices under the microscope: a property-level analysis for the Washington, DC area

AEI Economics Working Paper, No. 2014-05

Provided in Cooperation with:

American Enterprise Institute (AEI), Washington, DC

Suggested Citation: Pinto, Edward J.; Oliner, Stephen D.; Davis, Morris A. (2014) : House prices and land prices under the microscope: a property-level analysis for the Washington, DC area, AEI Economics Working Paper, No. 2014-05, American Enterprise Institute (AEI), Washington, DC

This Version is available at:

<https://hdl.handle.net/10419/280517>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



AMERICAN ENTERPRISE INSTITUTE

House Prices and Land Prices under the Microscope: A Property-Level Analysis for the Washington, DC Area

Morris A. Davis

University of Wisconsin-Madison

Stephen D. Oliner

American Enterprise Institute and UCLA

Edward J. Pinto

American Enterprise Institute

House Prices and Land Prices under the Microscope: A Property-Level Analysis for the Washington, DC Area

Morris A. Davis
University of Wisconsin-Madison

Stephen D. Oliner
American Enterprise Institute and UCLA

Edward J. Pinto
American Enterprise Institute

May 2014

Abstract

In this paper, we use a new property-level dataset to estimate the price of land since 2000 for nearly 600,000 detached single-family homes in the Washington, DC metro area. We employ these estimates to characterize the recent boom/bust cycle in land prices and house prices in the Washington area at a fine level of geography. The data show that land prices were more volatile than house prices everywhere, but especially so in the areas where land was inexpensive in 2000. In those areas, the land share of property value jumped during the boom, and this rise in the land share was a useful predictor of the subsequent crash in house prices. These results highlight the value of focusing on land for assessing house-price risk. In ongoing work, we are exploring how much of the spatial variation in land-price changes can be linked to high-risk mortgage lending and fundamentals such as housing rents, income growth, and employment.

The views expressed are ours alone and do not represent those of the institutions with which we are affiliated. We received expert research assistance from Rahee Jung, Jessica Li, Lilla Lukacs, Eleanor Qian, and Shuyi Yu. In addition, we received helpful comments from Sankar Bokka, Norrine Brydon, Robert Dorsey, Stuart Gabriel, Stuart Rosenthal, Joe Tracy, Stijn Van Nieuwerburgh, Mike Viehweg, Susan Wachter, and participants at the March 2014 International Conference on Collateral Risk in Berlin. We thank FNC, Inc. and Marshall & Swift, a CoreLogic company, for providing the data for the paper.

1. Introduction

One of the key lessons from the housing boom and bust, the foreclosure surge, and the financial crisis is that participants in the housing and mortgage markets, along with policymakers, need better information to evaluate the risks in these markets. Our goal in this paper and related work is to help fill this gap. With the right data and tools in place, we believe that the turbulence in housing and mortgage markets can be tempered.

In other work, we have constructed indexes of the risk in newly originated residential mortgages based on a rigorous stress test.¹ These indexes measure risk using key characteristics of the mortgage at origination: the loan-to-value ratio, the borrower's debt payment-to-income ratio and FICO score, and the purpose of the mortgage (home purchase or refinance), loan term, loan type (fixed or adjustable rate), and tenure type (primary owner-occupied, second home, or investor). These variables are powerful predictors of default risk, but an enhanced measure — and one that we intend to implement in the future — would account as well for the risk of a decline in the value of the underlying collateral, the house itself.

The research that we present here is our first attempt to quantify collateral risk at a fine level of geography. Many other studies (Davis and Palumbo, 2008; Houghwout, Orr, and Bedoll, 2008; Kok, Monkkonen, and Quigley, 2014; and Nichols, Oliner, and Mulhull, 2013, among others) have shown that most of the risk in the price of a house reflects the value of the underlying land. In this paper, we estimate the price of residential land from 2000 to 2013 for nearly 600,000 detached single-family homes in the Washington, DC metropolitan area. These properties represent 90 percent all such homes in the area under study. To our knowledge, no

¹ These indexes are produced by the International Center on Housing Risk at the American Enterprise Institute and are available at www.HousingRisk.org. The indexes currently include about three-quarters of all home purchase loans originated monthly. The broadest index covers the country as a whole (the National Mortgage Risk Index, or NMRI); separate indexes are published for each state and some metropolitan areas.

previous study has estimated land values for individual parcels on such a large scale. Our goal is to apply the same methods described in this paper to estimate the price of residential land for numerous other metropolitan areas in the United States.

As emphasized by Davis and Heathcote (2007), Davis and Palumbo (2008), and Nichols, Oliner, and Mulhall (2013), housing can be viewed as a bundle consisting of a structure that provides shelter and land that provides utility because of its particular location. Housing structures are relatively easy to reproduce using a known technology whereas the supply of land ready for development is inelastic. If the supply curve for structures is flat, and the supply curve for land is steep, then shocks to the demand for housing should result primarily in an increase in the quantity of structures built and an increase in the price of land.

Although housing can be affected by both supply shocks and demand shocks, the available evidence suggests that demand shocks account for most of the time-series variation in house prices and housing market activity. Well-calibrated models with only supply shocks (Davis and Heathcote, 2005; Iacoviello and Neri, 2010; Favilukis, Ludvigson, and Van Nieuwerburgh, 2011; and Garriga, Manuelli, and Peralta-Alva, 2012) cannot explain the volatility of house prices. In addition, a casual look at aggregate data strongly supports the notion that demand shocks drive housing markets. Figure 1 graphs the four-quarter percent change in real house prices against single-family housing starts, both for the U.S. as a whole. The two series are very highly correlated. Figure 1 is consistent with the idea that housing markets are best characterized as subject to large demand shocks that simultaneously boost prices and quantities. In this situation, land will account for the bulk of the price risk in the housing bundle.

We estimate land prices with a method that takes advantage of the rich information in our dataset. Briefly, we begin by measuring the implied land value for newly-built homes as the difference between the observed sale price and the estimated construction cost for the new structure. This method does not require any estimate of the depreciation of housing structures, a significant benefit given the absence of such information for local housing markets. The resulting land prices for newly-built homes provide a market-based "stake in the ground" for a specific time period, which we apply (with some adjustments) to all existing homes in the same zip code. Finally, we create property-level time series for home value, land value, and structure value that incorporate the stake in the ground and that reflect the changes over time in indices of house prices and construction costs for the property's local area.

Our results for land prices at a given time largely conform to the predictions of a standard urban model (Muth, 1969; Mills, 1972). In the Washington, DC metro area, the price of land tends to be less expensive and lot sizes tend to increase the farther parcels are located from the city center. At the same time, we find that the Washington area cannot be characterized by a uniform downward land-price gradient from the city center, as the price of land is relatively low in close-in areas populated by less affluent residents.

What we find with regard to land price over time is quite striking. In the places where land was cheap in 2000 (the outer suburbs and closer-in areas with limited locational amenities), the price of land jumped more than 500 percent during the boom from 2000 to 2006. In contrast, in the affluent inner suburbs where land was expensive in 2000, its price increased about 200 percent over the same period. We see the same pattern in reverse during the bust. The price of land declined about 20 percent from its peak in affluent areas but lost a staggering 75 percent of its peak value in initially cheap areas. Although land prices were everywhere more volatile than

house prices, the price of land was most volatile in areas where land was inexpensive and represented a small fraction of home value in 2000.

Unlike the pattern for land prices, house prices in the Washington area rose by roughly the same amount on average from 2000 to 2006 in places with different initial land prices. However, after the peak, the pattern for house prices mimicked that for land prices: steep price drops where land was initially cheap and much milder declines in places where land was expensive. This implies that an explosive rise in land prices, even when not accompanied by a relatively large increase in house prices, provided a valuable signal about the risk of a severe house-price drop. This result should inform the risk assessments done by market participants and policymakers. In ongoing work, we are exploring how much of the spatial variation can be linked to differences in the amount of high-risk mortgage lending and fundamentals such as housing rents, income growth, and employment.

Our final result relates to the "land leverage" hypothesis of Bostic, Longhofer, and Redfearn (2007). This hypothesis predicts that house prices will be most volatile in areas with the highest land share of property value. The prediction will hold if land prices are everywhere more volatile than house prices and if the ups and downs in land prices are similar in the various areas under study. Our results are consistent with the first condition but not the second, as we find that the land-price cycle was much sharper in the areas with low land shares. Consequently, in the Washington, DC area, the drop in house prices after 2006 was the most severe in areas with low land shares — the opposite of the prediction of the land leverage hypothesis. We believe our paper, which permits the measurement of land prices and land shares with fine geographic detail, is the first study to provide evidence that contradicts the hypothesis.

The rest of the paper is organized as follows. The next section discusses the unique, property-level dataset for the Washington, DC area that we use for the analysis. Section 3 describes in detail how we estimate land prices for the individual properties in the dataset. Section 4 presents a snapshot of the cross-sectional variation in the Washington metropolitan area, while section 5 lays out the time-series results for the recent boom/bust cycle. Section 6 concludes.

2. Data

a. Geographic coverage

Our dataset covers the city of Washington, DC and four large counties in the Washington metropolitan area: Montgomery County and Prince George's County in Maryland and Fairfax County and Prince William County in Virginia. The two Maryland counties, taken together, surround all but the southwestern part of Washington, DC. Fairfax County sits to the southwest of Washington, separated from the city by two relatively small jurisdictions in Virginia, Arlington County and the City of Alexandria.² Prince William County is a more distant suburban area that borders Fairfax County. We included Prince William in the dataset to provide additional coverage of the exurban part of the Washington metropolitan area.

Table 1 presents basic information about the five jurisdictions in the dataset. Fairfax County and Montgomery County are the most populous areas, with each having slightly more than a million residents as of 2012. The four counties plus Washington, DC have a total population of about 4 million people. Compared with the United States as a whole, the five

² At present, we do not have data for these two jurisdictions. We are attempting to obtain data for both to fill the small geographic hole in our dataset.

jurisdictions on average are quite affluent, have a well educated population, and are ethnically diverse.

However, there are some significant differences across the jurisdictions. Fairfax County and Montgomery County have high median income, high educational attainment, and a relatively small share of the population that is Black or Hispanic. In contrast, Prince George's County has considerably lower median income, a much smaller of adults with a college education, and a population that is almost 80 percent Black or Hispanic. Prince William County and Washington, DC have a blend of the characteristics at these extremes.

b. Source data

The core of our dataset consists of property-level information for detached single-family homes in the four counties and Washington, DC. The primary data source is the National Collateral Database (NCD) produced by the mortgage technology company, FNC, Inc. As described in Dorsey et al. (2010), the NCD covers virtually all residential properties throughout the United States, blending data from public records and home appraisals. The property-level information includes physical characteristics, location, the history of sale prices, appraised value, and the latest tax assessment. Importantly, the NCD file shows the year in which the home was built, its sale price when new, and its lot size. This information enables us to estimate land value with the methodology described in section 3.

We merge the NCD file with two other pieces of information for each home. The first is FNC's proprietary estimate of the home's market value as of the third quarter of 2013. FNC maintains a set of automated valuation models (AVMs) and selects the model that performs best for a given location.³ For example, in Montgomery County, the preferred AVM keys off the

³ For further information, see "FNC Automated Valuation Models," April 2014 (<http://www.fncinc.com/>). This document is available from FNC on request.

home's tax assessment, while in Prince William County, the preferred model is a hedonic regression estimate.

The other piece of information merged with the NCD record is the estimate of the home's reconstruction cost as of 2013:Q3 from Marshall & Swift, a CoreLogic company. Marshall & Swift's construction cost estimates are used extensively by property insurance companies, building contractors, appraisers, and government agencies. Their estimate of reconstruction cost represents the full cost of rebuilding the home from the ground up as a new structure. This includes the cost of materials and labor, equipment rentals, builder and subcontractor profit margins, permits and fees, and all applicable taxes associated with residential construction in the local area. The estimated reconstruction cost differs across homes based on a number of factors, but the main drivers are location, total square footage of living space, age of the existing structure (which affects the materials used), number of stories, and whether the home is attached or detached from other housing units.

As noted above, each home's estimated market value and reconstruction cost pertain to single time period, 2013:Q3. We create property-level time series for both variables with local house price indices and construction cost indices. The house-price indices are those produced by FNC using a spatial hedonic model estimated with the NCD data (see Dorsey et al., 2010, for details). FNC constructs the indices monthly back to January 2000 at various levels of geography, including the five-digit zip-code level. For each home, we use the index for the five-digit zip code in which it is located.⁴ The construction cost indices come from Marshall & Swift. These indices measure the cost of constructing several different models of homes in a given area, incorporating the various costs described above. Marshall & Swift produces the indices monthly

⁴ The indexes for individual metropolitan areas and higher-level aggregations are posted at <http://www.fncrpi.com/>. The zip-level indexes are available from FNC on request.

for three-digit zip codes throughout the country. The level of aggregation is higher than for the house price indices because construction costs do not vary within narrowly defined areas. We use the construction cost index for the three-digit zip code in which each home is located.

c. Property-level dataset

To create the dataset for our empirical work, we merged the three sources of property-level data and then removed properties in the NCD file that could not be address-matched to the Marshall & Swift file or that did not have an AVM. Among the remaining properties, we applied a series of data-quality screens.

The first of these screens removed properties in the NCD file that likely are either townhouses or condominiums. We eliminated properties that Marshall & Swift classified as condominiums or end-unit townhouses.⁵ Then, to screen out other attached units, we dropped the properties with very small lots.⁶

We applied three additional screens to take out properties with evident data errors or properties that were outliers. Specifically, we removed properties that were listed as having the same address as another property, properties with extremely large lots (in the top ½ percent for their respective county or Washington, DC), and properties for which the difference between the AVM and the estimated construction cost was in the bottom 1 percent or top 1 percent for their zip code.⁷ We did the latter trim at the zip level to allow for systematic differences in land values.

⁵ The Marshall & Swift file only identifies townhouses that are end units. It does not distinguish other townhouses from detached homes.

⁶ Specifically, for the Maryland and Virginia counties, we removed properties with lots in the bottom ½ percent of the county's lot size distribution. This threshold implied a minimum lot size of roughly 3500 square feet for properties retained in the dataset. Manual data checking indicated that relatively few townhouses in the four counties had lots larger than this threshold. In Washington, DC, lots tend to be quite small, even for detached homes. Accordingly, we set the minimum lot size at 2500 square feet.

⁷ For zip codes with relatively few properties, we did this trim at the "zip group" level. We discuss the distinction between zip codes and zip groups below.

Table 2 shows the size of the resulting dataset, together with the 2012 Census estimate of the number of detached single-family homes in each jurisdiction. Our dataset contains a total of about 593,000 properties, compared with the Census estimate of about 660,000 — a coverage rate of 90 percent. For the four counties and Washington, DC individually, the coverage rates range from about 85 percent to 95 percent. Hence, our dataset contains nearly the universe of detached single-family homes in the areas that we analyze.

d. Zip codes and zip groups

The number of properties in each five-digit zip code varies widely. Of the 162 five-digit zips in the dataset, several have more than 10,000 properties, while others have just a handful. These differences reflect the underlying variation in the total number of residential properties by zip code and the split of these properties between detached single-family homes and other types of housing.

To deal with this variation, we attached the zip codes with sparse data to an adjacent zip code (or in some cases, more than one adjacent zip) to create geographic units with at least 1000 properties. This aggregation condensed the original 162 zip codes into 119 geographic units that we refer to as "zip groups". About three-quarters of the zip groups consist of a single zip code, and most of the others combine two zips. Appendix table A1 lists all the individual zip codes in the dataset and the created zip groups, showing the number of properties in each zip group, the jurisdiction in which it is located, the cities and towns that it contains, and information related to the estimation of land value.

3. Estimating land value

a. Overview

We estimate land value for individual homes in the Washington, DC area with an approach that has its roots in traditional land valuation practices (see Babcock, 1932, for an early reference) and that blends the methods used in two strands of the modern literature. One strand has relied on observed land sales (for recent examples, see Albouy and Ehrlich, 2013a and 2013b; Kok, Monkkonen, and Quigley, 2014; and Nichols, Oliner, and Mulhall, 2013). This approach has the virtue of incorporating observed market prices. However, the volume of land transactions in the older, established parts of a metropolitan area is generally too sparse to estimate changes in land values over time for specific localities. Reflecting this limitation, Nichols, Oliner, and Mulhall (2013) estimated an aggregate land price index for each of 23 metropolitan areas, but did not attempt to calculate indexes for different parts of each metro area. An alternative approach (see, for example, Davis and Palumbo, 2008) measures land value indirectly as the difference between the value of a representative house and the depreciated reconstruction cost of the structure on the lot. The data requirements to implement this method are lower than for the transaction-based approach. But calculating land value as a residual means that the estimates inherit the measurement errors elsewhere in the accounting framework.

Our method takes advantage of the strengths of each approach, while circumventing their weaknesses. Specifically, we use market prices to estimate land value for individual homes at a given time and then compute internally consistent measures of home value and structure value. This provides a market-based "stake in the ground." We then create property-level time series for home value, land value, and structure value that incorporate the stake in the ground and that reflect the changes over time in indices of house prices and construction costs for the property's

zip group. Importantly, this method does not require any estimate of depreciation for housing structures — a significant benefit given the absence of information on depreciation for local housing markets.

b. Estimating land value for new homes

We take advantage of two key pieces of property-level information: the sales price for a house when it was first built and the estimated cost of rebuilding the structure as new. Let $PHNEW_{i,z,t}$ denote the sale price of new home i located in zip group z and built in period t , and let $RCNEW_{i,z,s}$ denote the reconstruction cost for the as-new structure in period s . The estimated reconstruction cost is as of 2013:Q3 for all properties, while the sale prices are almost entirely from earlier periods. We move the estimated reconstruction cost back to the period of the initial home sale using the Marshall & Swift index of construction costs. With this re-dating, the estimated value of land for a specific new home in period t is

$$(1) \quad PL_{i,z,t} = PHNEW_{i,z,t} - RCNEW_{i,z,t}.$$

For this method to work, the estimated reconstruction cost in 2013:Q3 must reflect the same structure as when the house was built. Homes that have had a major renovation or expansion fail this requirement and thus cannot be used in equation 1. Because our dataset does not provide the history of improvements for a property, we assume that homes built in 2000 or after have not had major structural changes by 2013. Thus, the earliest period for which we use equation 1 is 2000:Q1.

Two other considerations define the sales used in equation 1. First, we employ only arms-length sales of non-distressed property to help ensure that the implied land values will be applicable to the full population of properties. And second, we include sales that occur not only

in the year the home was built, but also in the year after, as some new homes may sit on the market for a while before being sold.

An implicit assumption behind equation 1 is that reconstruction cost serves as a good proxy for the unobserved market value of the new housing structure. We believe this is a reasonable assumption. The equality of reconstruction cost and structure value is a basic equilibrium condition in a housing market with ongoing construction activity by profit-maximizing builders (see Schulz and Werwatz, 2011, and Rosenthal, 1999, for example). Although this condition will not hold exactly at all times, builders have an incentive to adjust the level of construction activity to close gaps that open up. The limited available evidence suggests that the equilibration occurs fairly rapidly (Rosenthal, 1999). Accordingly, using reconstruction cost as a proxy for the market value of a new structure should not generate significant measurement error.

The transactions in our dataset include some sales of vacant land in addition to sales of finished homes, but the two types of sales are not identified explicitly. It is important to distinguish new home sales from land sales because equation 1 is only applicable to the former; for land transactions, the sale price itself measures the property's land value. We classify the sales in our dataset as follows. Let PL^* denote the result from using equation 1 without knowledge of whether the sale involves a finished home or land. If PL^* is positive and greater than 10 percent of the sale price, we assume this is a new home sale and use equation 1 to value the land. Conversely, if PL^* is negative and greater than 10 percent of the sale price in absolute value, we assume this is a land sale (since the sale price is significantly below the cost of the structure that ultimately appears on the site); in such cases, we set the land value equal to the observed sale price. Finally, when PL^* is within 10 percent of the sale price in absolute value,

we do not classify the transaction as either a land or new home sale and set it aside.⁸ Using this market-based procedure, about 9 percent of the properties in the full dataset have an estimated land value.

The number of resulting land values differs widely by zip group. Not surprisingly, there are many observations in outlying areas where the housing stock is relatively new and far fewer observations in closer-in areas that were almost completely developed before 2000. As shown in appendix table A1, about half of the zip groups in Washington, DC have fewer than 25 estimates of land value over 2000-13, while most zip groups in Prince William County have more than 1,000. Although more observations clearly would be preferred to fewer, even a small number of land values — if reasonably representative — can provide the required "stake in the ground" for our methodology.

c. Estimating land value for all homes in a single period

The next step is to use the market-based land values for new homes in zip group z to impute land values for all homes in that zip group. One approach would be to regress the market-based land values on lot size, time, and possibly other variables, using the regression coefficients to compute fitted land values for other houses in the zip group. The difficulty, however, is that some zip groups have too few observations to produce reliable coefficients from such a regression.

In light of this issue, we use an alternative approach. To begin, we designate a "reference quarter", denoted by t_R , for each zip group. The reference quarter is the period in which we set the stake in the ground. If the initial period in our analysis, 2000:Q1, contains at least ten

⁸ We also excluded sales with apparent data errors. Specifically, we dropped transactions for which the sale price was more than five times greater than the property's AVM or less than 10 percent of the AVM. In addition, we excluded properties with multiple sales records on a single date when the lowest and highest sale prices differed by more than \$5000; when the prices differed by less than \$5000, we used the first sales record.

market-based land values, it serves as the reference quarter. If there are fewer than ten such land values in 2000:Q1, we combine sales from 2000:Q1 through the quarter that contains the tenth sale. The reference quarter in this case is set to be the average quarter of the included sales.

For each zip group, we calculate the median land value per square foot and median lot size for sales in the reference quarter (or the multiple quarters used to define the reference quarter, as the case may be). We use medians to reduce the influence of outliers, and we separate total land value into the value per square foot and lot size in order to account for the established finding that land value per square foot depends on lot size.⁹ For notation, let $PLSF_{med,z,t_R}$ and LOT_{med,z,t_R} denote, respectively, the median land value per square foot and the median lot size among the sale properties in zip group z in the reference period t_R (or the quarters that contribute to the reference period).¹⁰ In addition, let $PLSF_{j,z,t_R}$ and $LOT_{j,z}$ denote, respectively, the estimated land value per square foot and the lot size of the j th property in zip group z , where $j = 1, \dots, J$ indexes the full set of properties, not simply those built since 2000. No time subscript is needed for LOT because lot size for a specific property is constant. With this notation, the estimate of land value for property j in the reference period is:

$$(2) \quad PL_{j,z,t_R} = PLSF_{j,z,t_R} LOT_{j,z} = \left[PLSF_{med,z,t_R} f(LOT_{j,z}, LOT_{med,z,t_R}) \right] LOT_{j,z},$$

⁹ This is the so-called "plattage effect." Nichols, Oliner, and Mulhall (2013) estimated that a 10 percent increase in lot size is associated with a decline of about 5 percent in land price per square foot. Larger residential lots sell for less per square foot than smaller lots subject to the same zoning because the extra square footage cannot be used to build an additional house. Discussions of the lack of proportionality between square footage and lot value date back as far as Hurd (1903), Bernard (1913), National Association of Real Estate Boards (1927), and Babcock (1932).

¹⁰ In 19 of the zip groups, we calculated the median land value per square foot and the median lot size based solely on the presumed home sales, omitting the presumed land sales. We did this because the land values from these latter sales tended to be substantially higher than the land values from the presumed home sales. In all likelihood, these presumed land sales were actually home sales that we misclassified when using the allocation rule described above. The column in appendix table A1 labeled "Type of sales used" indicates the zip groups for which we used only the presumed home sales.

where $f(LOT_{j,z}, LOT_{med,z,t_R})$ is a factor (described below) that adjusts the median land price per square foot computed from the sales properties for the difference in lot size between those properties and property j .

To calculate the adjustment factor f , we pool the sale properties across all the zip groups and regress the natural log of the land price per square foot for each property on a set of zip group dummy variables, a third-order polynomial in time, and a third-order polynomial in lot size. The dummy variables control for differences in the average level of land prices across zip groups, while the polynomial in time controls for the cycle in land prices. The estimated polynomial in lot size $\left(\sum_{k=1}^3 \hat{\alpha}_k LOT^k\right)$ determines the adjustment factor for each property:

$$(3) \quad f(LOT_{j,z}, LOT_{med,z,t_R}) = \exp\left(\sum_{k=1}^3 \hat{\alpha}_k LOT_{j,z}^k - \sum_{k=1}^3 \hat{\alpha}_k LOT_{med,z,t_R}^k\right).$$

Figure 2 plots the estimated adjustment factor for lots ranging in size from 5000 square feet to 90,000 square feet (roughly two acres), relative to a baseline quarter-acre lot (10,890 square feet). As shown, large lots have a much lower price per square foot than small lots, consistent with the findings in the literature. For example, the price per square foot for a half-acre lot (21,780 square feet) is only about 60 percent of that for the baseline quarter-acre lot.¹¹

d. Estimating home value and structure value for all homes in a single period

Given the property-level estimate of land value in the reference quarter from equation 2, we complete the picture for each property by estimating its total market value (land plus structure) and the structure value alone in the reference quarter. Recall that the dataset includes

¹¹ Figure 2 scales the adjustment factor to be in terms of a quarter-acre lot. However, the median lot size for the sale properties in a given zip group will not be exactly a quarter acre except by chance. Therefore, in terms of figure 2, the adjustment factor for a given property is the value shown for the property's lot size divided by the value for the median lot size for the sale properties in its zip group. For example, if a given property has a lot size of 15,000 square feet and the median lot for the sale properties in its zip group is 30,000 square feet, the adjustment factor is about $0.85/0.45$, or slightly less than 2.

an AVM estimate of each home's market value in 2013:Q3. We move that estimate back to the reference quarter with the FNC house price index for the home's zip group. That is,

$$(4) \quad PH_{j,z,t_R} = \left(PFNC_{z,t_R} / PFNC_{z,2013:Q3} \right) AVM_{j,z,2013:Q3}.$$

Using PH from this equation, the estimate of the (depreciated) structure value for each home in the reference period is then

$$(5) \quad RC_{j,z,t_R} = PH_{j,z,t_R} - PL_{j,z,t_R},$$

where PL is calculated in equation 2.

e. Creating time series

The property-level estimates of total home value, land value, and structure value in the reference period provide the stake in the ground from which we can compute property-level time series from 2000:Q1 to 2013:Q3.

For zip groups with a reference period of 2000:Q1, we roll $PH_{j,z,2000:Q1}$ forward quarter by quarter with the price ratio $PFNC_{z,t+1} / PFNC_{z,t}$ for the home's zip group. Similarly, we roll $RC_{j,z,2000:Q1}$ forward quarter by quarter with the ratio of the Marshall & Swift construction cost index, $PMS_{z,t+1} / PMS_{z,t}$, for the home's zip group. This process generates a property-level home value and structure value in each quarter t from 2000:Q1 through 2013:Q3. The value of land in quarter t is then calculated as $PH_{j,z,t} - RC_{j,z,t}$.

For zip groups with a reference period after 2000:Q1, the only difference is that we roll PH_{j,z,t_R} and RC_{j,z,t_R} backward to 2000:Q1 as well as forward to 2013:Q3 to generate the full quarterly time series for each variable. With these time series in hand, we calculate the value of land in quarter t as $PH_{j,z,t} - RC_{j,z,t}$, just as we did for the zip groups with a reference period of 2000:Q1.

In both cases, we have created property-level time series for home value, land value, and structure value from 2000:Q1 to 2013:Q3. These property-level series are the building blocks for the analysis that follows. We aggregate across the properties in a zip group to obtain time series for average home, land, and structure values, along with the average land share of home value.

f. Adjustments

As a check on the resulting series, we examined the implied land shares for the zip groups. For the large majority of zip groups, the land shares presented no obvious issues. However, in some zip groups, the starting value in 2000:Q1 was close to zero or even negative, before rising over time, while in other zip groups, the land share exceeded 90 percent in some periods. Anomalous values such as these could arise if the sale properties — from which the market-based land values were determined — are not representative of the full set of properties in the zip group. This could occur, for example, if the homes built since 2000 were located in parts of the zip group for which land prices were generally higher (or lower) than average.

For these groups, we adjusted the starting value for the land shares in 2000:Q1. Specifically, the shares below five percent were adjusted to be exactly five percent, while those greater than 60 percent were changed to that value.¹² These adjustments were intended to nudge the 2000:Q1 land shares toward reasonable values. We multiplied the adjusted land share by the original estimate of home value in 2000:Q1 to obtain an adjusted estimate of land value (and an adjusted estimate of structure value as home value minus land value). With this revised stake in

¹² There were also two zip groups with no estimated land share because of an absence of new home sales. For these zip groups, we set the starting value of the land share to be 40 percent, about the average for the other zip groups.

the ground, we then re-created the time series for each property with the procedure described above.¹³

4. Cross-sectional Patterns

This section presents a snapshot of the variation across the zip groups for house prices, land values, and other property characteristics. We use a series of heat maps that divide the zip groups into quintiles, using data for 2013:Q3, the latest period in our dataset. In each map, the darkest shade represents the quintile with the highest values for each variable and the lightest shade represents the lowest quintile.

Figure 3 shows the pattern of house prices in 2013:Q3 across the zip groups, where the house price for each zip group is the average AVM for the homes in that group. As shown, house prices are the highest in Fairfax County, the southwestern part of Montgomery County (Bethesda, Chevy Chase, and Potomac), and the part of Washington, DC that borders Montgomery County. These areas include the most sought-after addresses in the Washington area. House prices are lower in outlying areas (Prince William County and the northern part of Montgomery County) and in locales with relatively low income and generally high proportions of Black and Hispanic residents (Prince George's County and the adjoining parts of Montgomery County and Washington, DC). The spread of average AVM values across the zip groups is wide, ranging from less than \$295,000 in the lowest quintile to between \$620,000 and \$1.6 million in the highest quintile.

Figure 4 presents the parallel distribution of average lot values, calculated as described in section 3. As can be seen by comparing figures 3 and 4, lot values mirror the pattern for home

¹³ Appendix table A1 identifies the 21 zip groups (of 119 in total) for which these adjustments were made. The time-series results presented in section 5 are essentially the same whether we include or exclude the 21 zip groups from the analysis.

prices — highest in Fairfax County and the affluent parts of Montgomery County and Washington, DC, and lowest in outlying areas and Prince George's County. Moreover, the range of average lot values is extremely wide. Most zip groups in Prince George's County have average lot values below \$100,000, while the zip groups in Fairfax County, Montgomery County, and Washington, DC in the highest quintile have average lot values that range from about \$430,000 to more than \$1 million. Clearly, house prices are high in some places and lower in others largely because of differences in the value of the underlying land.

As indicated in figure 5, lots tend to be relatively small in close-in areas and larger further out. None of the zip groups in Washington, DC proper has an average lot size greater than 9,000 square feet, while zip groups in the exurbs have lots that average as much as nearly four acres. This pattern is consistent with the canonical urban model of Muth (1969) and Mills (1972), which predicts that households respond to the higher price of land closer to the city center by economizing on land use.¹⁴

Given the wide differences in average lot size across the zip groups, measured land prices per square foot will be influenced by the plattage effect described in section 3. That is, large lots will have a lower value per square foot, all else equal, simply because much of the square footage is above and beyond that needed to build a house at that location. To control for this effect, we calculated the average land price per square foot for a quarter-acre lot in each zip group, using the adjustment function in figure 2. This standardized land price does a better job than the raw price per square foot of capturing the amenity value of land in various locations.

¹⁴ Technically, the prediction of the model relates to the size of the housing structure relative to the lot, rather than to lot size per se (see McMillen, 2006). The heat map for this ratio (not shown) displays qualitatively the same pattern as figure 5.

As shown in figure 6, the spatial pattern for the quarter-acre land price is very similar to that for house prices from figure 3. This high correlation confirms that differences in location-specific amenity value are an important driver of the variation in house prices across zip groups.

Finally, figure 7 portrays the average land share of property value across the zip groups in 2013:Q3. Comparing figures 6 and 7 shows that the land share tends to be high in the zip groups with high land prices and vice versa. Notably, the zip groups in the affluent part of Washington, DC, the adjacent areas of Montgomery County, and the close-in parts of Fairfax County all have land shares in the highest quintile or the next quintile. In contrast, almost all the zip groups in Prince George's county are in the lowest or second-lowest quintile. This difference in land shares will play a central role in the analysis of the housing cycle in the next section.¹⁵

5. The Price Cycle Since 2000

a. House prices

We use the zip-level hedonic price indexes published by FNC, Inc. to examine the appreciation in house prices from 2000 and 2006 and the subsequent decline from 2006 to 2012. The heat maps in figures 8 and 9 portray the respective periods.

Both figures show substantial variation across the zip groups, consistent with the heterogeneity found in other recent studies of house prices using data at a fine level of geography (Guerrieri, Hartley, and Hurst, 2013; Ferreira and Gyourko, 2011; Mian and Sufi, 2009). From 2000 to 2006, prices in the lowest quintile of zip groups rose 89 to 127 percent, compared with a range that tops 200 percent in the highest quintile. Then, from 2006 to 2012, prices in the

¹⁵ Despite this general pattern, some zip groups have land shares that differ considerably from those in neighboring zip groups. Generally, these zip groups are ones for which we adjusted the land shares from the raw values produced by the estimation procedure. These zip groups include 20721 in Prince George's County, the combination of 22134 and 22172 in Prince William County, and 20120, 20151, and 20171 in Fairfax County, all of which have land shares (even after adjustment) above those in adjacent zips. The land shares for these zip groups should be regarded as less reliable than for those for other zip groups.

quintile with the largest declines fell 41 to 51 percent, while prices dropped less than 15 percent in the quintile with the smallest declines. Four zip groups in this latter quintile actually saw prices rise on net over 2006-12. Moreover, one of those zips (20815, in Montgomery County) is located less than five miles from another zip in the same county (20902) where prices dropped 34 percent. Clearly, the Washington, DC area cannot be characterized as a single, homogeneous housing market.

Interestingly, the house price declines shown in figure 9 have a pronounced geographic pattern. Prices fell the least in the most affluent areas covered by our dataset — Fairfax County, the northwest part of Washington, DC, and the adjoining part of Montgomery County. In contrast, the steepest declines were concentrated in the outlying areas and in Prince George's County. That is, the exurbs and places with large Black and Hispanic populations were hit the hardest during the bust.

These two seemingly distinct areas share an important characteristic: they both have relatively low land prices (as we saw in figure 6). Thus, the places in the metropolitan area with cheap land experienced the most severe house-price crash.¹⁶ This pattern suggests that further analysis of the role of land could be fruitful.

b. Land prices, house prices, and construction costs over the cycle

We begin by examining the movement in land prices over our full sample period. To focus on the connection between the level of land prices and the magnitude of the price cycle, we aggregate the 119 zip groups into quintiles based on the quarter-acre land price in 2000:Q1.

¹⁶ Figure 6 displayed land prices in 2013:Q3, the latest period in our dataset. The use of 2013:Q3 could raise concerns that the magnitude of the post-2006 house price decline mechanically influenced the zip-group distribution of land prices shown in figure 6, making the correlation uninteresting. However, as we will show, the same correlation emerges between the level of land prices in 2000:Q1 (the initial period in the dataset) and the post-2006 house price decline. Hence, the connection between the level of land prices and the severity of the house price decline is not an artifact of using land prices as of 2013:Q3.

Figure 10 presents the resulting time series for land prices. The figure demonstrates that the amplitude of the price cycle was systematically related to the initial level of land prices. The price swing — both the rise and then the decline — was the greatest for the zip groups where land was initially the cheapest. As one moves from the lowest quintile of land prices to each higher quintile, the price cycle becomes progressively milder.¹⁷

To provide perspective on the size of the swing in land prices, figures 11 and 12 compare the percent changes in land prices, house prices, and construction costs for the zip groups aggregated into the same land-price quintiles as in figure 10. Figure 11 presents the comparison for the boom phase of the cycle (2000-06), while figure 12 covers the bust (2006-12).¹⁸ In both figures, house prices are measured by the FNC zip-level indices, while the construction cost indices are those produced by Marshall & Swift.

Three important results are immediately evident from figure 11. First, during the boom phase of the cycle, land prices rose much more than house prices in every quintile, while construction costs increased only modestly. Hence, the rise in house prices over 2000-06 largely reflected the appreciation in land value. Second, house prices increased by similar amounts in each of the quintiles. This implies that monitoring the rise in house prices quintile by quintile would have provided little information about the magnitude of the post-2006 price drop, which we know from figure 10 was especially large in the quintiles with low land prices. Third, unlike the nearly uniform rise in house prices, the increase in land prices was exceptionally large in the lowest land-price quintile, at more than 500 percent, and then moderated from quintile to quintile

¹⁷ This result is the land-price counterpart to the pattern shown for house prices in the Case-Shiller indexes for separate price tiers (<http://us.spindices.com/index-family/real-estate/sp-case-shiller>). For each of the 20 included metropolitan areas, the magnitude of the post-2000 house price increase and the subsequent decline was greater for low-price homes than for high-price homes. Guerrieri, Hartley, and Hurst (2013) document the same pattern over 2000-06 for a broader set of metropolitan areas.

¹⁸ We end in 2012 rather than 2013 because, as shown in figure 10, land prices were at or near their bottom in 2012. Thus, ending in 2012 allows us to portray the extent of the price drop more accurately than if we include the recovery in 2013.

(though it remained large even in the highest quintile). Consequently, land prices contain information about potential overvaluation beyond that in house prices themselves.

Turning to figure 12, what we observe after the 2006 peak mirrors in some respects what happened during the run-up through 2006. Land prices fell more than house prices in every quintile, with staggering losses in the lowest two price quintiles, where land prices plunged 75 percent and 60 percent, respectively. Because construction costs continued to rise over 2006-12, the drop in land prices more than accounted for the decline in house prices, emphasizing the central role of land in house-price swings. One difference, however, from figure 11 is that the magnitude of the house-price drop varies across the quintiles, and markedly so. Whereas house prices retreated only a bit more than 10 percent on net in the zip groups in the highest land-price quintile, they fell nearly 40 percent in the lowest quintile. Differences of this magnitude likely would be associated with sharp differences in the performance of the underlying mortgage loans, highlighting the potential value of information from land price movements in assessing risk.

c. Land share of property value

Another key difference across the quintiles relates to the land share of property value. Figure 13 plots the time series of the average land share in each land-price quintile. As would be expected, the land share is low in the locations where land is cheap and high where land is expensive. For example, in 2000, the land share ranged from a shade less than 25 percent in the lowest land-price quintile to more than 50 percent in the highest quintile. At the end of the sample period in 2013, the gap was even wider.

The difference in land shares across the quintiles explains arithmetically why land prices rose so much more than house prices over 2000-06 in the lowest quintiles. With land representing a small share of house value, the accounting relationship that connects house prices

to land and structure prices could only hold if land prices had risen much more than house prices. In the quintiles with higher land shares, the same accounting relationship would hold with a smaller rise in land prices relative to house prices.

Our results concerning land shares have important implications for the land-leverage hypothesis of Bostic, Longhofer, and Redfearn (2007). This hypothesis states that house prices will be more volatile, all else equal, in places where land represents a relatively large share of property value. Bostic, Longhofer, and Redfearn found that the hypothesis was supported in their detailed study of house prices in Wichita, Kansas. A necessary, but not sufficient, condition for the hypothesis to be valid is that land prices have wider swings than structures prices (so that land prices are more volatile than house prices). The findings from numerous previous studies have satisfied this condition.¹⁹ Our results for the Washington, DC area do as well, as we found that land prices rose more and then fell more than house prices in every land-price quintile.

Nonetheless, our results do not support the central prediction of the hypothesis. That is, we do not find that house-price volatility is the greatest in the zip groups with the highest land shares. Rather, we see the opposite pattern, as shown in table 3. The zip-group quintile with the lowest land prices and the lowest land shares experienced the most severe drop in house prices after 2006. Conversely, the two quintiles with the highest land shares had the mildest post-2006 decline in house prices. During the 2000-06 boom, there were only small differences in house price appreciation across the quintiles, but the slight differences go against the land-leverage hypothesis, with house prices rising the least in the quintile with the highest land share in 2000.

¹⁹ See Davis and Heathcote (2007), Davis and Palumbo (2008), Haughwout, Orr, and Bedoll (2008), Nichols, Oliner, and Mulhall (2013), and Sirmans and Slade (2011) for supporting evidence in U.S. metropolitan areas. For international evidence, see Bourassa et al. (2009, 2010), Schulz and Werwatz (2011), and Wu, Gyourko, and Deng (2012).

Our results do not support the land-leverage hypothesis because land prices were so much more volatile in the zip groups with low land shares than in the zip groups with high shares. Evidently, land prices in the areas with low land shares had a larger speculative component than land prices elsewhere. We will be analyzing the determinants of these land-price patterns in future research.

d. Predictive power of changes in land shares

Reflecting these differential movements in land prices, the land share has been more volatile since 2000 in the zip groups with initially cheap land than in the zip groups with more expensive land (recall figure 13). An important question is whether the amount by which the land share rose during the boom helps predict the magnitude of the house price decline during the bust. If so, real-time monitoring of land prices and land shares would be useful for market participants and policymakers.

To address this question, each panel of figure 14 plots the percent change in house prices from 2006 to 2012 against the rise in the land share in percentage points from 2000 to its maximum value in any year through 2006.²⁰ Each diamond represents a single zip group. Panel A includes all 119 zip groups, while panel B excludes the 21 zip groups for which we adjusted the land shares.

The figure shows that the rise in the land share during the boom did, in fact, have predictive power for the subsequent drop in house prices. In panel A, the regression line fit through the scatter plot has a highly significant slope coefficient of -0.95, implying that an additional 10 percentage point rise in the land share was associated with a 9.5 percentage point

²⁰ We use this definition of the rise in the land share because for some zip groups the land share peaked in 2005 rather than 2006.

drop in house prices after the peak. In panel B, which removes some of the outliers in panel A, the regression fit is somewhat tighter (the R^2 is 0.37 rather than 0.31), and the slope coefficient, -1.22, is about 25 percent larger in absolute value.²¹

To interpret this result, note that a rise in the land share indicates that the gap between house prices and construction costs has widened. Large increases in this gap proved to be unsustainable over the recent housing cycle. Eventually, house prices mean-reverted to fundamentals, bringing down the land share of property value.

6. Conclusions and Next Steps

We have provided a detailed picture of the recent boom/bust cycle in house prices and land prices for the Washington, DC metropolitan area, using an unprecedented dataset that covers close to the universe of detached single-family homes in the area under study. The rich property-level information allows us to estimate land value for the nearly 600,000 homes in the dataset. To our knowledge, no previous research has estimated residential land values at the property level on such a large scale.

The paper yields important new facts about the recent housing cycle in the Washington area. First, the swing in house prices and land prices varied widely across locations. The cycle was mildest in the affluent, close-in parts of the metro area. It was considerably greater in the more distant suburbs and in areas with a large Black or Hispanic population — places where land is relatively cheap. Second, land prices were more volatile than house prices everywhere, but especially so in the areas with initially inexpensive land. This held true even for different

²¹ Note that the rise in the land share retains its predictive power even after controlling for the rise in house prices from 2000 to 2006. In a regression of the change in house prices over 2006-12 on the change in house prices over 2000-06 and the maximum rise in the land share over 2000-06, the coefficient on the land share is significant at the 1 percent level. Moreover, the R^2 of this regression is 0.40, compared with only 0.25 when the land share is excluded.

locations in relatively close proximity. Third, changes in the land share of property value were a useful predictor of subsequent change in house prices: the areas with the largest increases in the land share tended to suffer the sharpest drop in house prices during the bust. The predictive power of this single variable is especially notable because it works across a large, diverse metropolitan area subject to a variety of shocks. These results highlight the value of focusing on land for assessing house-price risk and suggest that market participants and policymakers should be particularly attentive to situations that involve a rapid appreciation of land prices from initially low levels.

Our results also cut against the land-leverage hypothesis, which holds that house prices will be more volatile in areas where land represents a large share of property value. Previous research has supported the hypothesis when assessed across cities, but our study is the first to examine it within a large, diverse metropolitan area. We find exactly the opposite of what the hypothesis predicts. House prices were most volatile in the parts of the metropolitan area with low land shares because the land price swing in those areas was so much wider than elsewhere.

Thus far, we have documented the cycle in land and house prices for the Washington area since 2000 at a fine level of geography. The next step will be to analyze the factors that produced the patterns that we observe. This follow-on research will focus on changes in employment, income, and loan quality at the zip-code level. An important goal will be to assess the contribution of high-risk mortgage lending during the boom to the explosion in land prices in certain areas.

The analysis for the Washington, DC area is part of a larger project that will encompass a number of metropolitan areas across the United States. Data collection has already begun for an initial set of other cities: Boston and Miami on the East Coast; Los Angeles, Phoenix, and

Seattle in the West; and Chicago, Detroit, Memphis, and Oklahoma City in the interior of the country. These cities were chosen to represent the range of housing markets seen across the United States. Eventually, we intend to expand the coverage to include all of the largest metropolitan areas in the country and a broader sampling of mid-size cities.

References

- Albouy, David, and Gabriel Ehrlich. 2013a. "The Distribution of Urban Land Values: Evidence from Market Transactions." Working paper. (<https://sites.google.com/site/gabrielehrlich/>)
- Albouy, David, and Gabriel Ehrlich. 2013b. "Metropolitan Land Values and Housing Productivity." Working paper. (<https://sites.google.com/site/gabrielehrlich/>)
- Babcock, Frederick M. 1932. *The Valuation of Real Estate*. McGraw-Hill.
- Bernard, Alfred Duncan. 1913. *Some Principles and Problems of Real Estate Valuation*. United States Fidelity and Guaranty Company.
- Bostic, Raphael W., Stanley D. Longhofer, and Christian L. Redfearn. 2007. "Land Leverage: Decomposing Home Price Dynamics." *Real Estate Economics* 35(2): 183-208.
- Bourassa, Steven C. et al. 2010. "Land Leverage and House Prices." Swiss Finance Institute Research Paper Series no. 10-48.
- Bourassa, Steven C. et al. 2009. "House Price Changes and Idiosyncratic Risk: The Impact of Property Characteristics." *Real Estate Economics* 37(2): 259-78.
- Davis, Morris A., and Jonathan Heathcote. 2007. "The Price and Quantity of Residential Land in the United States." *Journal of Monetary Economics* 54(8): 2595-2620.
- Davis, Morris A., and Jonathan Heathcote. 2005. "Housing and the Business Cycle." *International Economic Review* 46(3): 751-84.
- Davis, Morris A., and Michael G. Palumbo. 2008. "The Price of Residential Land in Large US Cities." *Journal of Urban Economics* 63(1): 352-84.
- Dorsey, Robert E., Haixin Hu, Walter J. Mayer, and Hui-chen Wang. 2010. "Hedonic Versus Repeat-Sales Housing Price Indexes for Measuring the Recent Boom-Bust Cycle." *Journal of Housing Economics* 19: 75-93.
- Favilukis, Jack, Sydney Ludvigson, and Stijn Van Nieuwerburgh. 2011. "The Macroeconomic Effects of Housing Wealth, Housing Finance and Limited Risk Sharing in General Equilibrium." Unpublished paper, New York University.
- Ferreira, Fernando, and Joseph Gyourko. 2011. "Anatomy of the Beginning of the Housing Boom: U.S. Neighborhoods and Metropolitan Areas, 1993-2009." NBER Working Paper 17374.
- Garriga, Carlos, Rodolfo Manuelli, and Adrian Peralta-Alva. 2012. "A Model of Price Swings in the Housing Market." Federal Reserve Bank of St. Louis Working Paper 2012-022A.

- Guerrieri, Veronica, Daniel Hartley, and Erik Hurst. 2013. "Endogenous Gentrification and House Price Dynamics." *Journal of Public Economics* 100: 45-60.
- Haughwout, Andrew, James Orr, and David Bedoll. 2008. "The Price of Land in the New York Metropolitan Area." Federal Reserve Bank of New York, *Current Issues in Economics and Finance* 14(3): 1-7.
- Hurd, Richard M. 1903. *Principles of City Land Values*. Real Estate Record Association.
- Iacoviello, Matteo, and Stefano Neri. 2010. "Housing Market Spillovers: Evidence from an Estimated DSGE Model." *American Economic Journal: Macroeconomics* 2(2): 125-64.
- Kok, Nils, Paavo Monkkonen, and John M. Quigley. 2014. "Land Use Regulations and the Value of Land and Housing: An Intra-Metropolitan Analysis." *Journal of Urban Economics* 81: 136-48.
- McMillen, Daniel P. 2006. "Testing for Monocentricity." In *A Companion to Urban Economics*, edited by Richard J. Arnott and Daniel P. McMillen, chapter 8.
- Mian, Atif, and Amir Sufi. 2009. "The Consequences of Mortgage Credit Expansion: Evidence from the U.S. Mortgage Default Crisis." *Quarterly Journal of Economics* 124(4): 1449-96.
- Mills, Edwin S. 1972. *Studies in the Structure of the Urban Economy*. Resources for the Future.
- Muth, Richard F. 1969. *Cities and Housing*. The University of Chicago Press.
- National Association of Real Estate Boards. 1927. *Real Estate Appraising*.
- Nichols, Joseph B., Stephen D. Oliner, and Michael R. Mulhall. 2013. "Swings in Commercial and Residential Land Prices in the United States." *Journal of Urban Economics* 73(1): 57-76.
- Rosenthal, Stuart S. 1999. "Residential Buildings and the Cost of Construction: New Evidence on the Efficiency of the Housing Market." *The Review of Economics and Statistics* 81(2): 288-302.
- Schulz, Rainer, and Axel Werwatz. 2011. "Is There an Equilibrating Relationship between House Prices and Replacement Cost? Empirical Evidence from Berlin." *Journal of Urban Economics* 69(3): 288-302.
- Sirmans, C.F., and Barrett A. Slade. 2011. "National Transaction-based Land Price Indices." *The Journal of Real Estate Finance and Economics*. Published online. Available online at www.springerlink.com/content/
- Wu, Jing, Joseph Gyourko, and Yongheng Deng. 2012. "Evaluating Conditions in Major Chinese Housing Markets." *Regional Science and Urban Economics* 42(3): 531-43.

Table 1. Population and Demographic Characteristics

County/city	Population	Median Household Income	Percent with Bachelors Degree	Percent Black or Hispanic
Fairfax Co., VA	1,118,602	\$107,923	57.8	24.8
Montgomery Co., MD	1,004,709	94,767	56.8	34.6
Prince George's Co., MD	881,138	72,254	29.8	78.4
Prince William Co., VA	430,287	95,427	37.2	40.4
Washington, DC	632,323	64,610	51.7	58.7
Memo: U.S. total	313,914,040	51,771	28.6	29.1

Note. Population and percent Black or Hispanic are as of 2012; the definition of Black or Hispanic is the Census line item for Hispanic or Latino of any race plus the line item for Black or African American alone and not Hispanic or Latino. Median household income and percent of adults with a bachelors degree are averages over 2010-12. Median household income is in 2012 dollars.

Sources: U.S. Census Bureau, 2012 American Community Survey, Table DP05, and 2010-12 American Community Surveys, Tables DP02 and DP03.

Table 2. Count of Detached Single-Family Homes

County/city	Dataset	Census	Percent coverage
Fairfax Co., VA	177,648	193,491	91.8
Montgomery Co., MD	154,609	182,854	84.6
Prince George's Co., MD	156,708	169,466	92.5
Prince William Co., VA	68,554	76,858	89.2
Washington, DC	35,241	37,044	95.1
Total	592,760	659,713	89.9

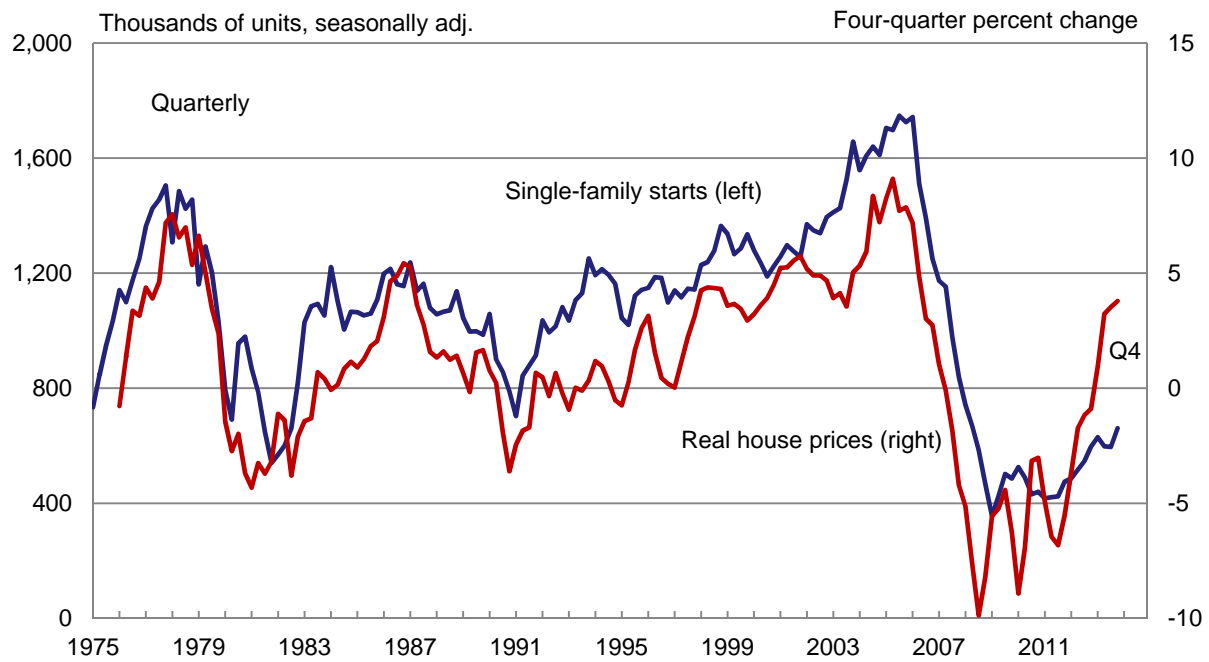
Sources: Authors' calculations based on dataset created from data provided by FNC, Inc. and Marshall & Swift, a CoreLogic company; U.S. Census Bureau, 2012 American Community Survey, Table DP04.

Table 3. Land Shares and Changes in House Prices by Zip-Group Quintiles

Quintiles, quarter-acre land price, 2000:Q1	Land share (pct., annual average)			House-price change (pct.)	
	2000	2006	2012	2000-06	2006-12
Lowest	22.9	59.8	24.1	145.7	-38.8
Second	33.1	66.2	40.0	156.0	-35.9
Third	38.7	67.4	48.8	142.5	-27.6
Fourth	51.8	74.8	63.1	147.2	-21.7
Highest	52.8	72.2	64.5	119.4	-11.4

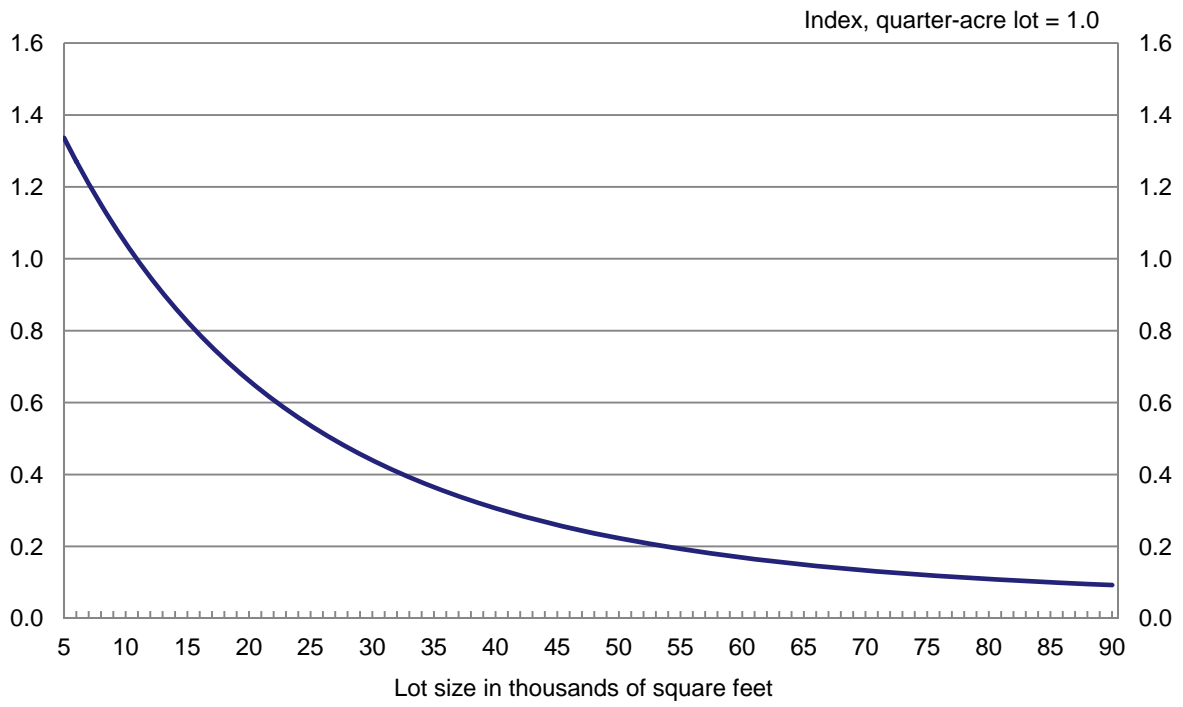
Sources: Authors' calculations based on dataset created from data provided by FNC, Inc. and Marshall & Swift, a CoreLogic company.

Figure 1
Single-family Housing Starts and Real House Prices



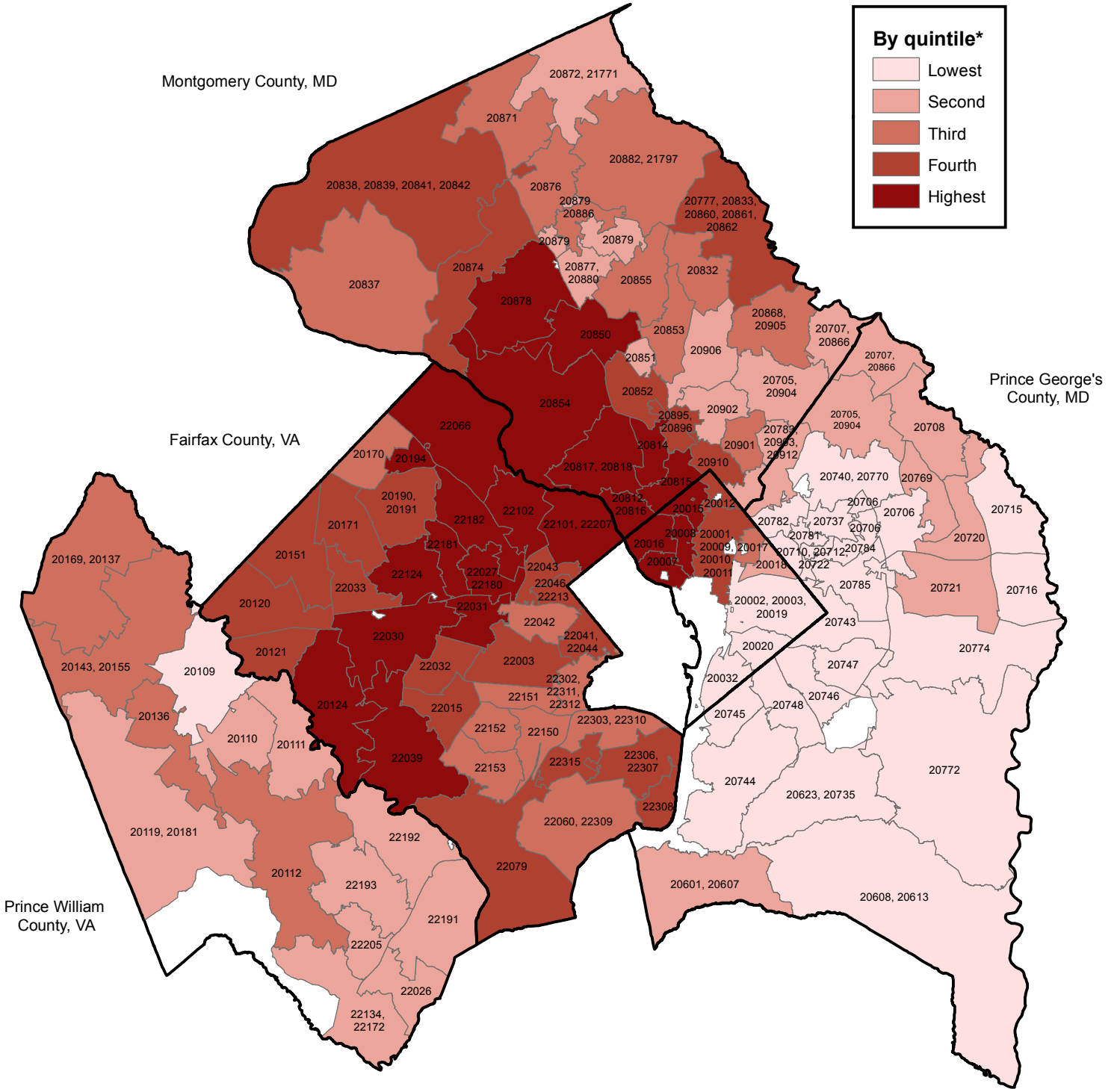
Note: Real house prices are measured as the FHFA all-transactions house price index divided by the price index for personal consumption expenditures.
 Source: Single-family housing starts, Census Bureau; real house prices, FHFA and Bureau of Economic Analysis.

Figure 2
Estimated Relationship Between Lot Size and Land Price per Square Foot



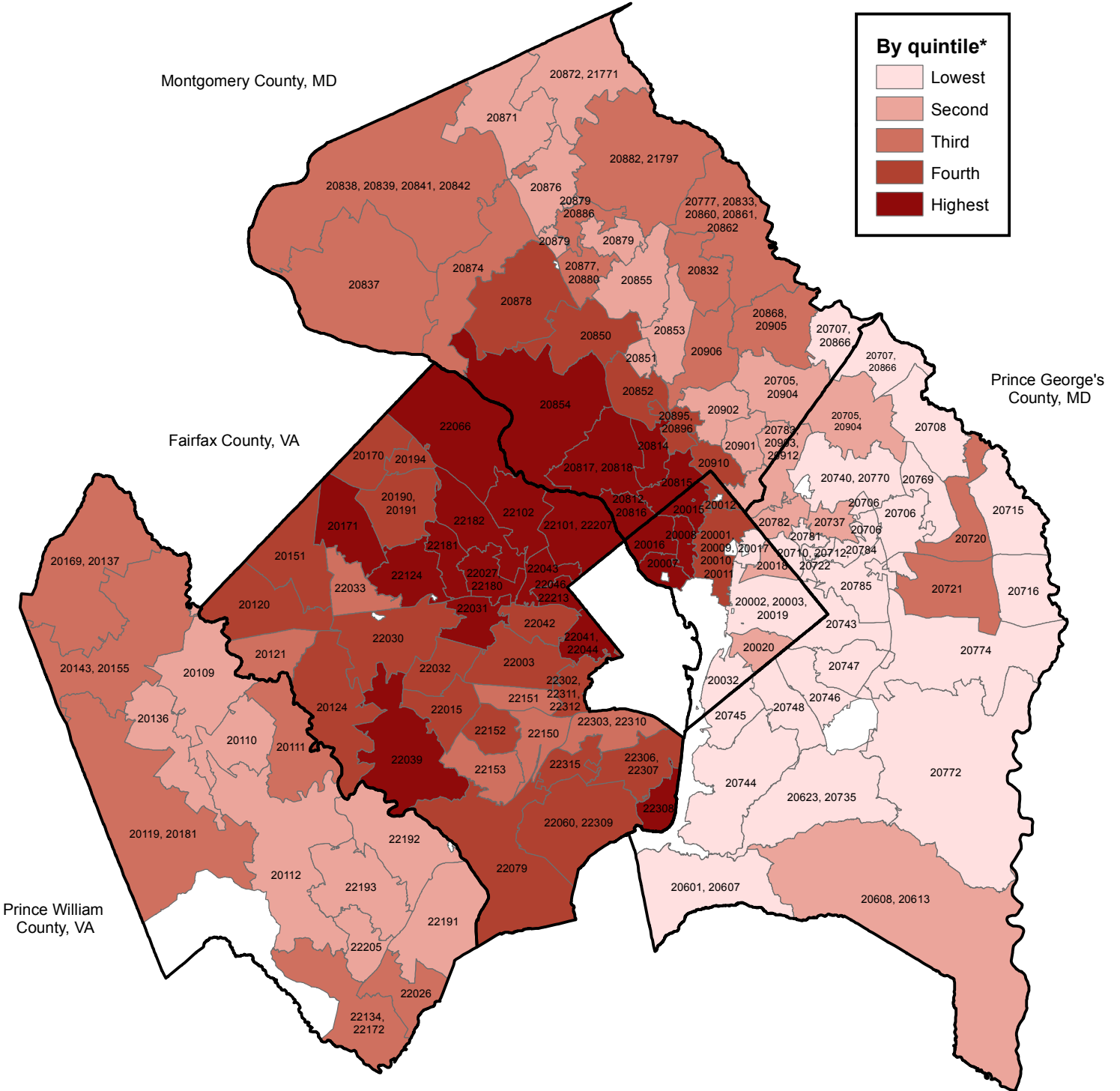
Source: Authors' calculations using data from FNC, Inc. and Marshall & Swift, a CoreLogic company.

Figure 3
Average AVM, 2013:Q3



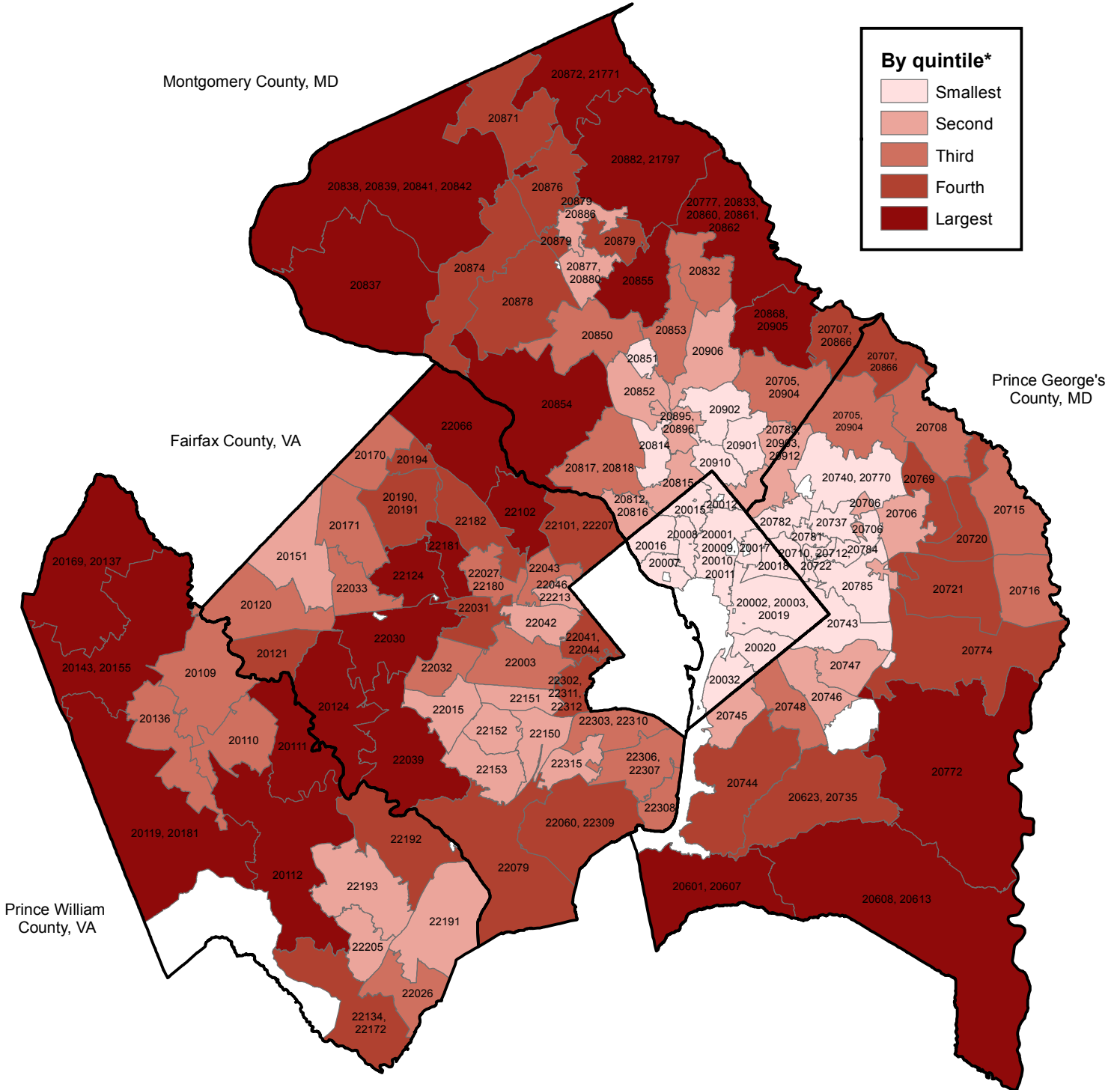
*Range of average AVM values by quintile (rounded to nearest \$1000) is \$128,000 to \$294,000, \$294,000 to \$416,000, \$416,000 to \$542,000, \$542,000 to \$621,000, and \$621,000 to \$1,593,000.
 Source: Authors' calculations using data from FNC, Inc.

Figure 4
Average Lot Value, 2013:Q3



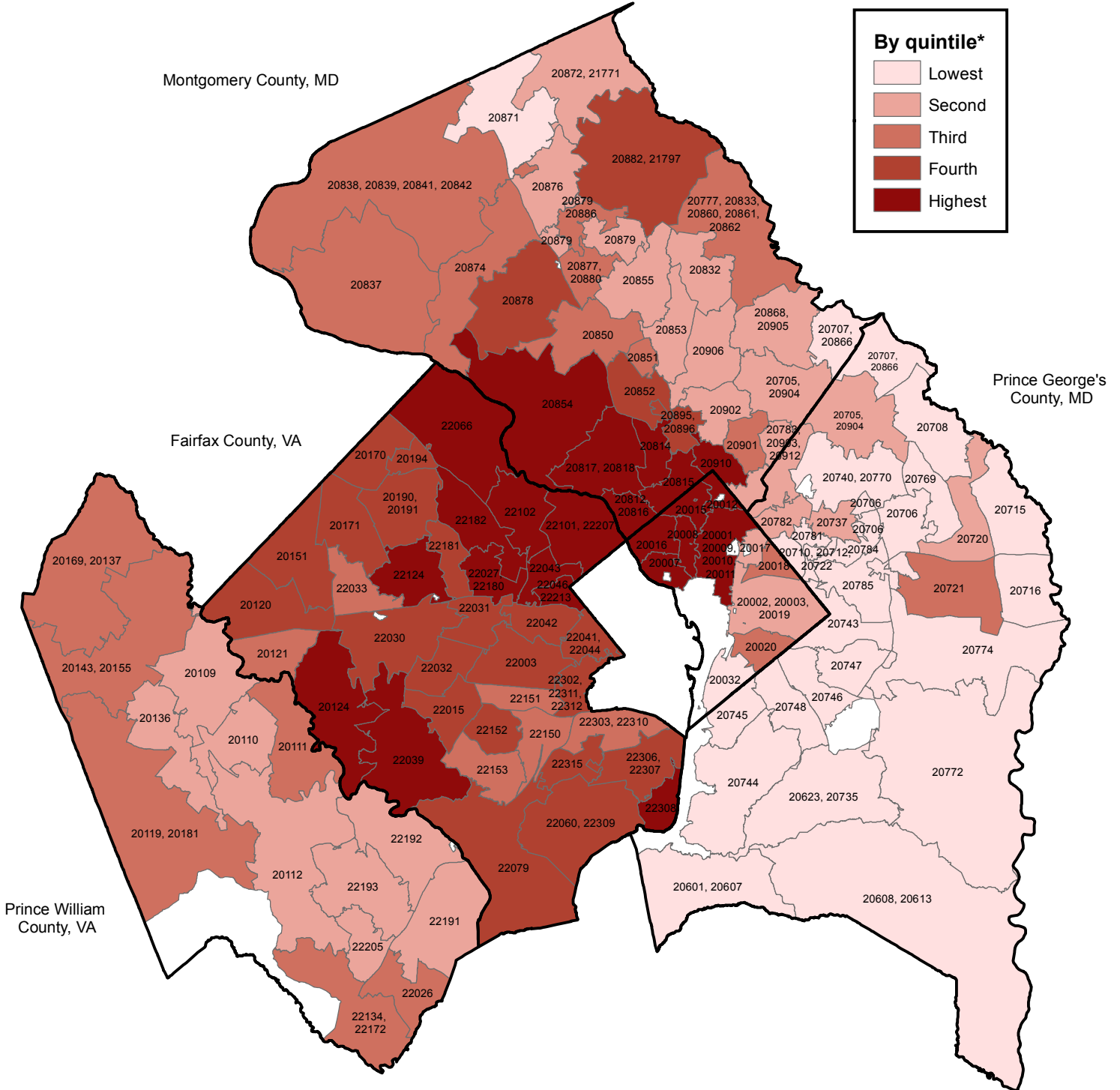
*Range of average lot value by quintile (rounded to the nearest \$1000) is \$6,000 to \$96,000, \$96,000 to \$193,000, \$193,000 to \$290,000, \$290,000 to \$432,000, and \$432,000 to \$1,057,000.
 Source: Authors' calculations using data from FNC, Inc. and Marshall & Swift, a CoreLogic company.

Figure 5 Average Lot Size, 2013:Q3



*Range of average lot size by quintile (rounded to the nearest 1000 square feet) is 4,000 to 9,000, 9,000 to 13,000, 13,000 to 17,000, 17,000 to 30,000, and 30,000 to 167,000.
Source: Authors' calculations using data from FNC, Inc.

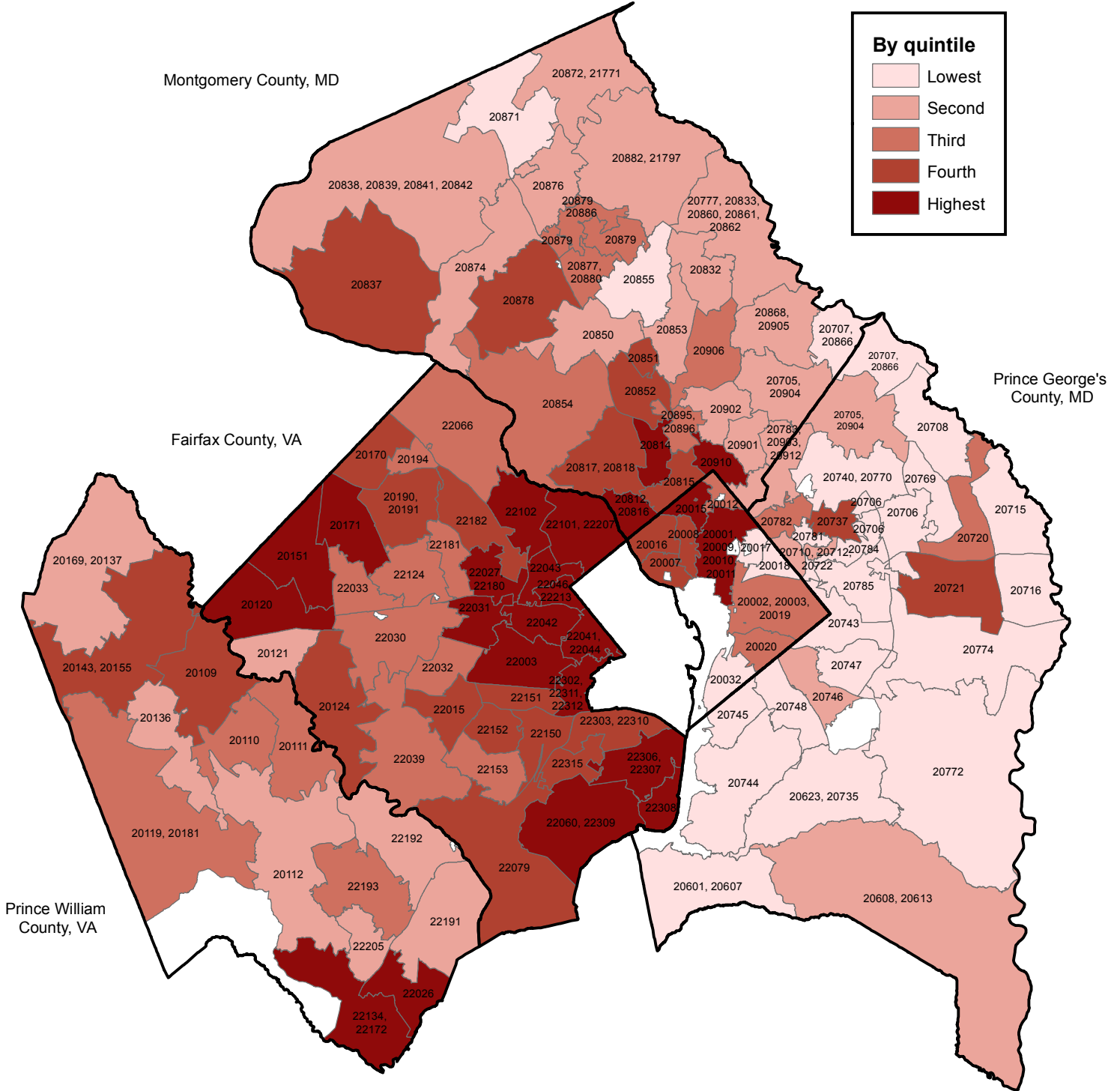
Figure 6
Average Standardized Land Price per Square Foot, 2013:Q3



*Range of average land price per square foot for a quarter-acre lot by quintile (rounded to the nearest dollar) is \$1 to \$10, \$10 to \$17, \$17 to \$24, \$24 to \$37, and \$37 to \$119.

Source: Authors' calculations using data from FNC, Inc. and Marshall & Swift, a CoreLogic company.

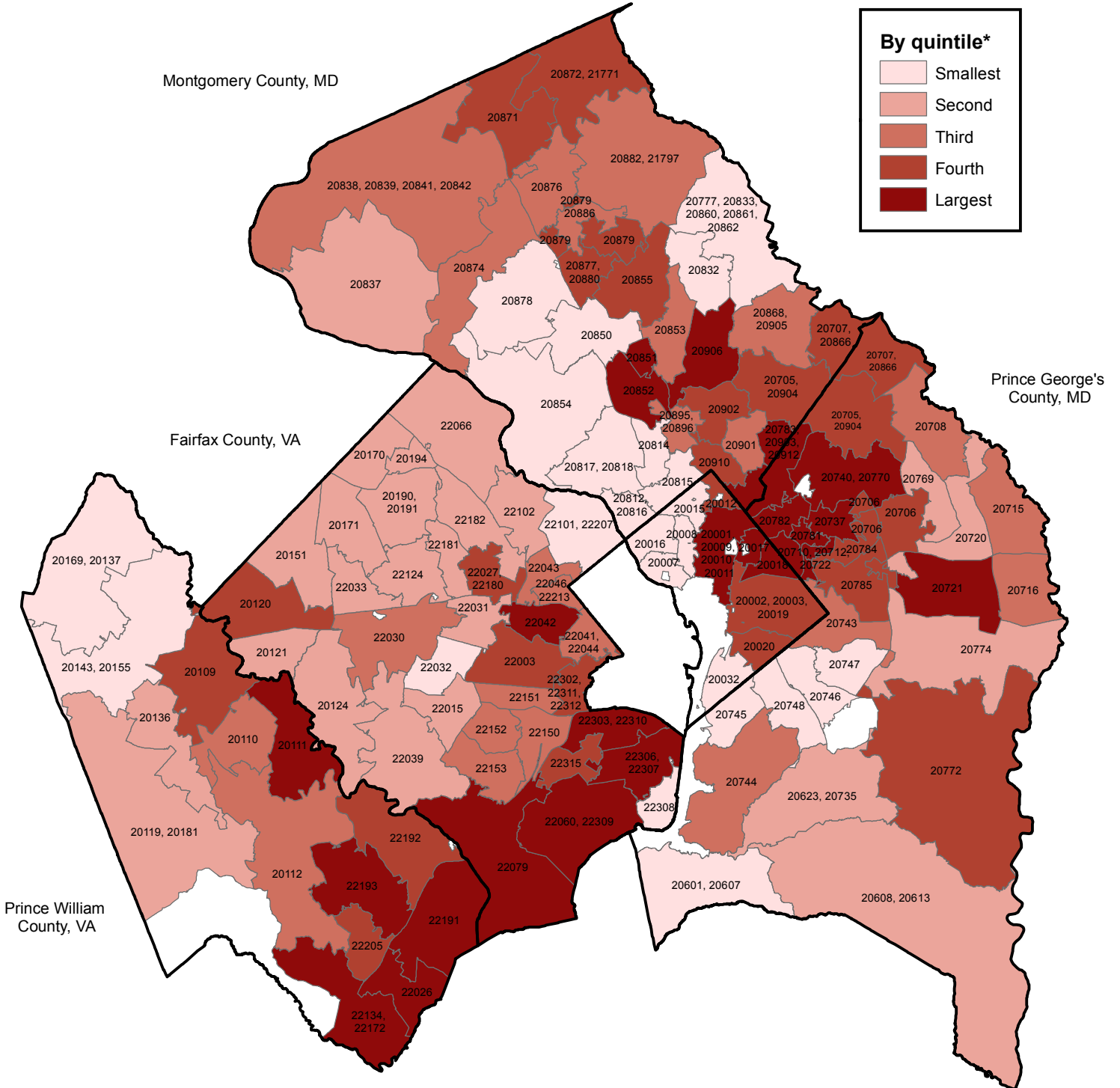
Figure 7
Average Land Share of Property Value, 2013:Q3



*Range of average land share of property value by quintile (rounded to the nearest 1%) is 3% to 35%, 35% to 47%, 47% to 58%, 58% to 70%, and 70% to 76%.

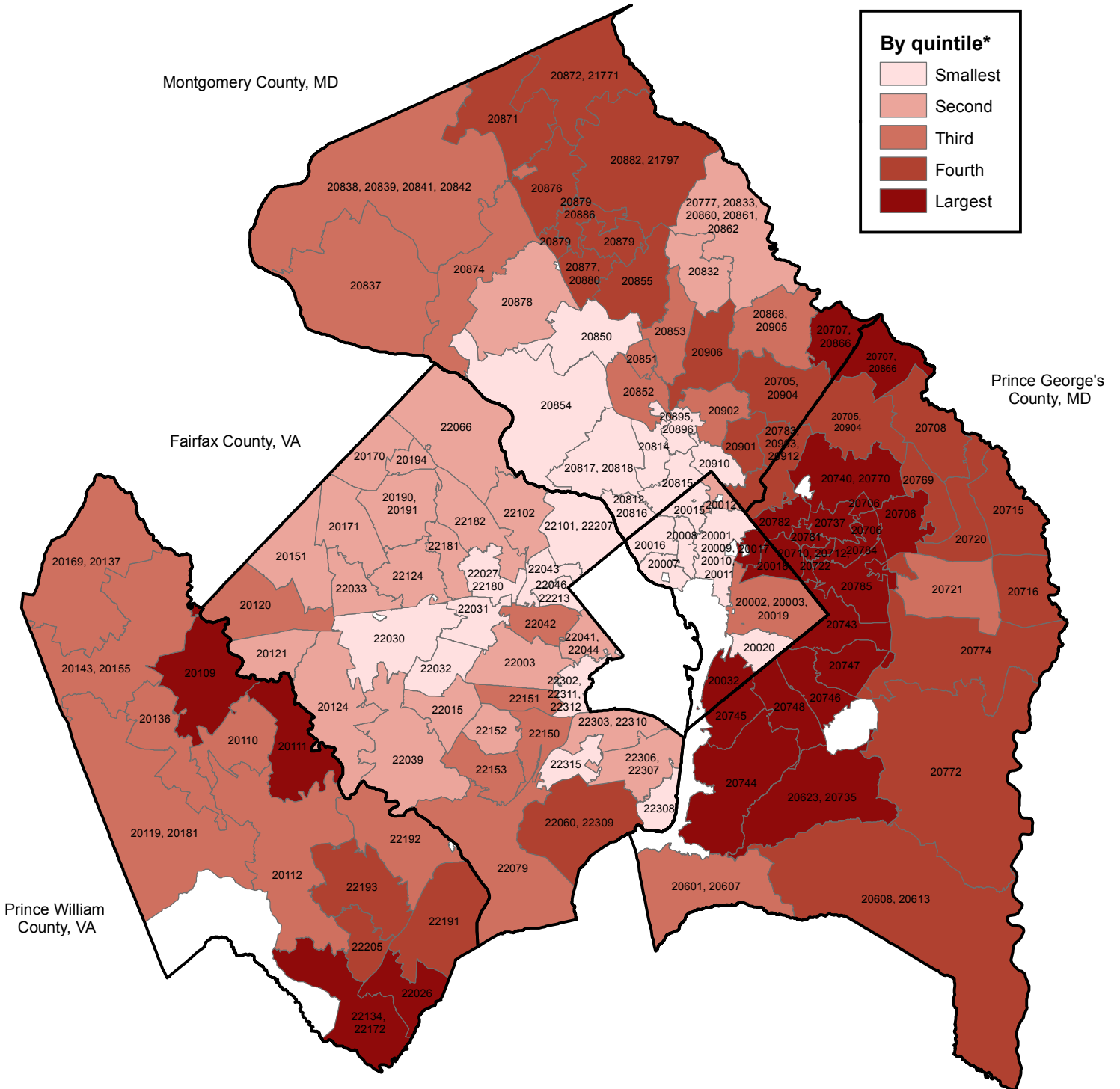
Source: Authors' calculations using data from FNC, Inc. and Marshall & Swift, a CoreLogic company.

Figure 8 House Price Increase, 2000-2006



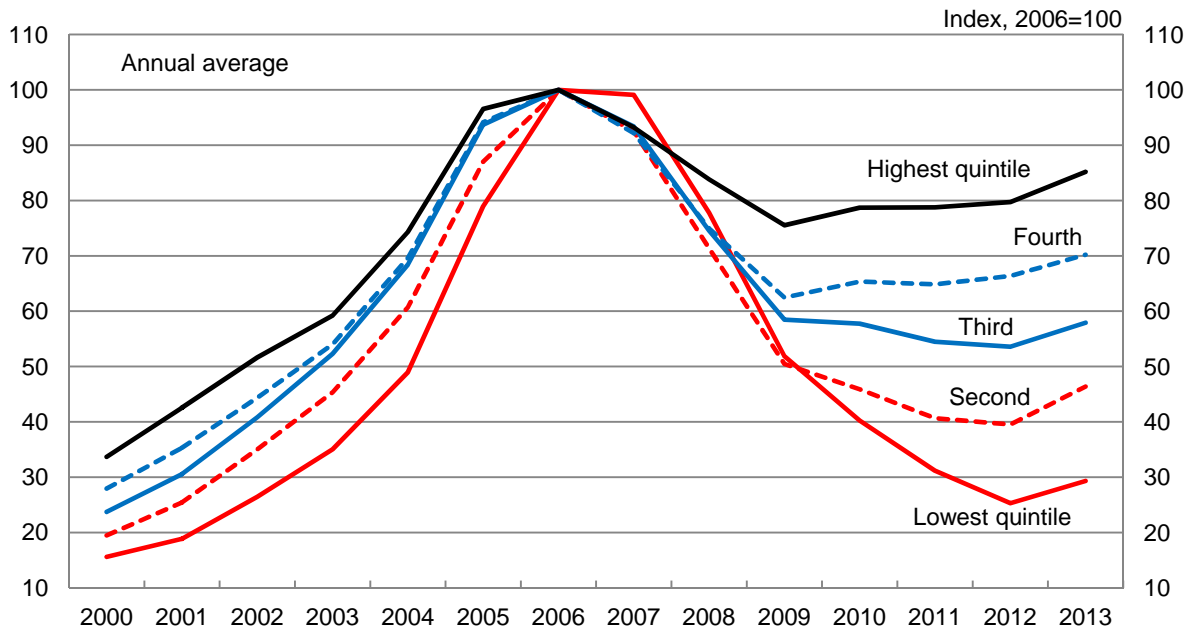
*Range of house price increase by quintile (rounded to the nearest 1%) is 89% to 127%, 127% to 130%, 130% to 146%, 146% to 165%, and 165% to 204%.
Source: Authors' calculations using data from FNC, Inc.

Figure 9 House Price Decline, 2006-2012



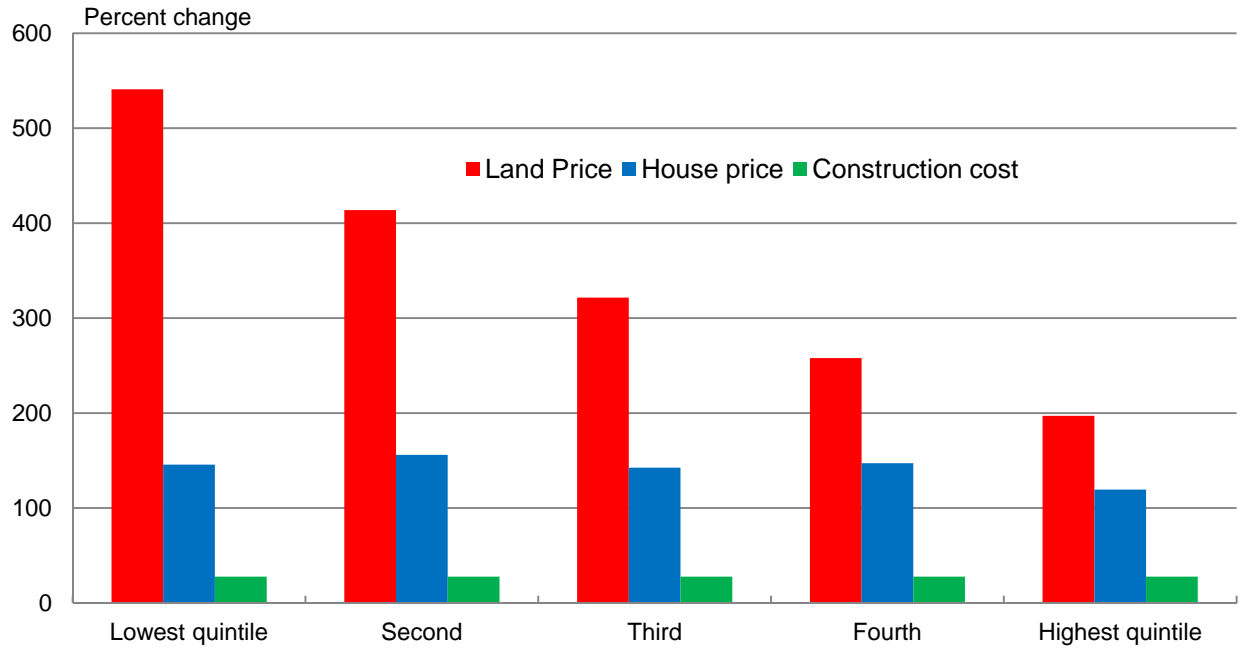
*Range of house price decline by quintile (rounded to the nearest 1%) is less than 15%, 15% to 22%, 22% to 34%, 34% to 41%, and 41% to 51%. Bottom quintile includes zips 20007, 20008, 20016, and 20815, where index rose.
Source: Authors' calculations using data from FNC, Inc.

Figure 10
Average Land Prices
 (for zips grouped into quintiles by quarter-acre land price in 2000:Q1)



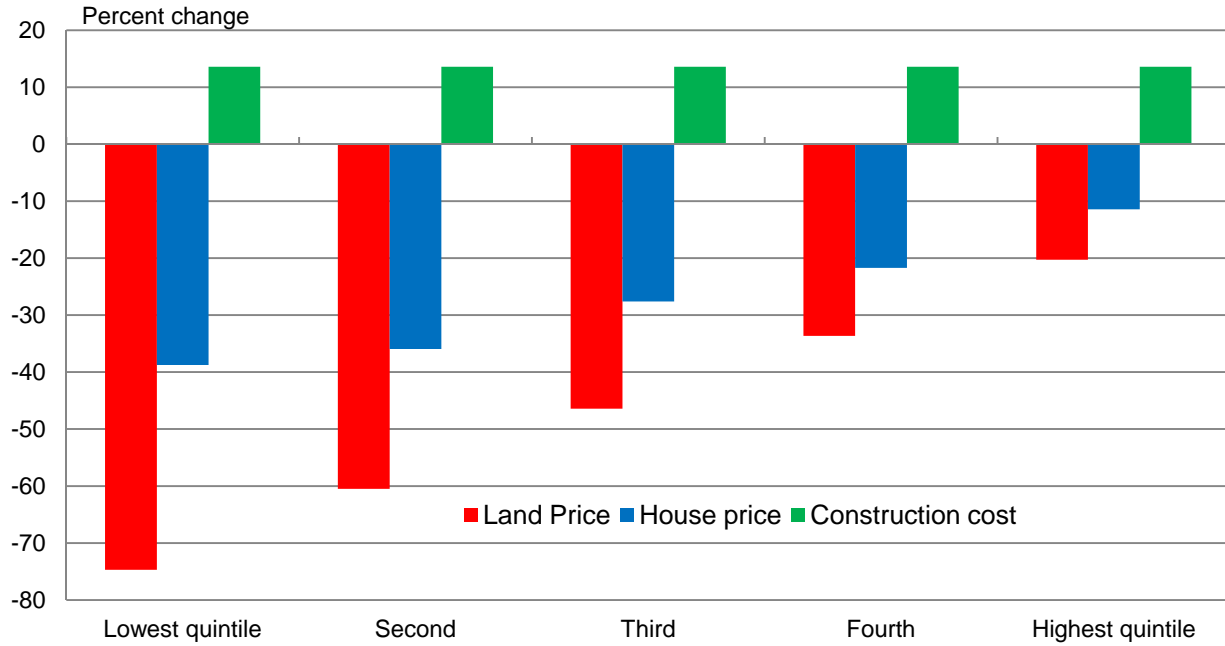
Note: Observations for 2013 represent average of Q1, Q2, and Q3.
 Source: Authors' calculations using data from FNC, Inc. and Marshall & Swift, a CoreLogic company.

Figure 11
Prices and Construction Costs: 2000-2006
(for zips grouped into quintiles by quarter-acre land price in 2000:Q1)



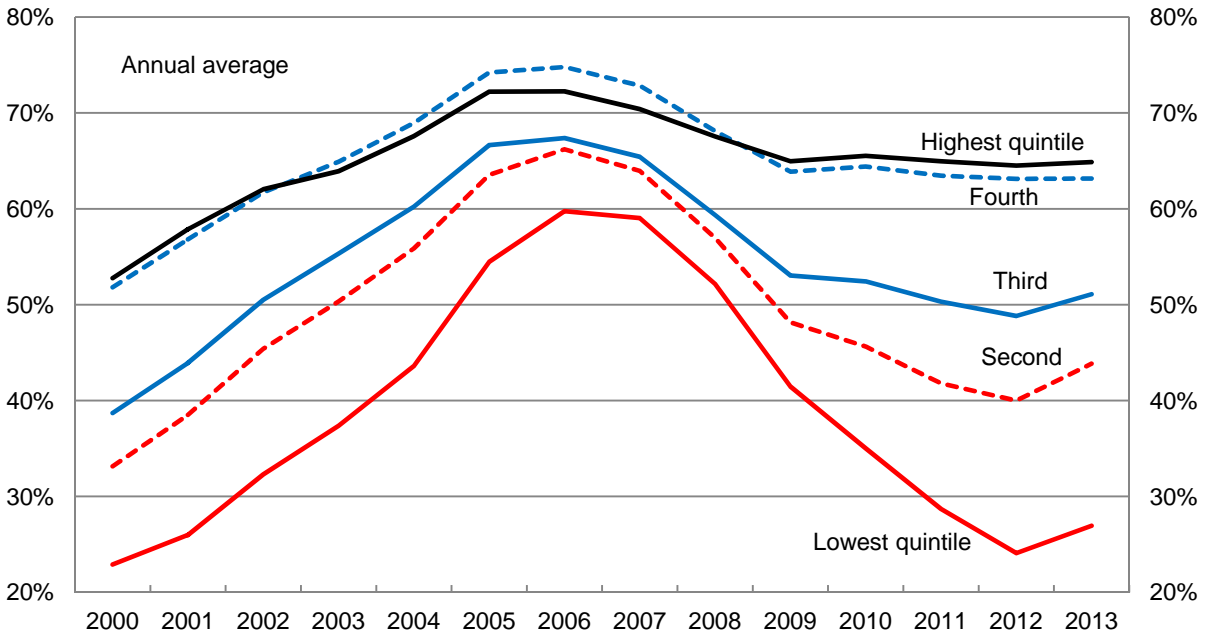
Source: Author's calculations based on data from FNC, Inc. and Marshall & Swift, a CoreLogic company.

Figure 12
Prices and Construction Costs: 2006-2012
 (for zips grouped into quintiles by quarter-acre land price in 2000:Q1)



Source: Author's calculations based on data from FNC, Inc. and Marshall & Swift, a CoreLogic company.

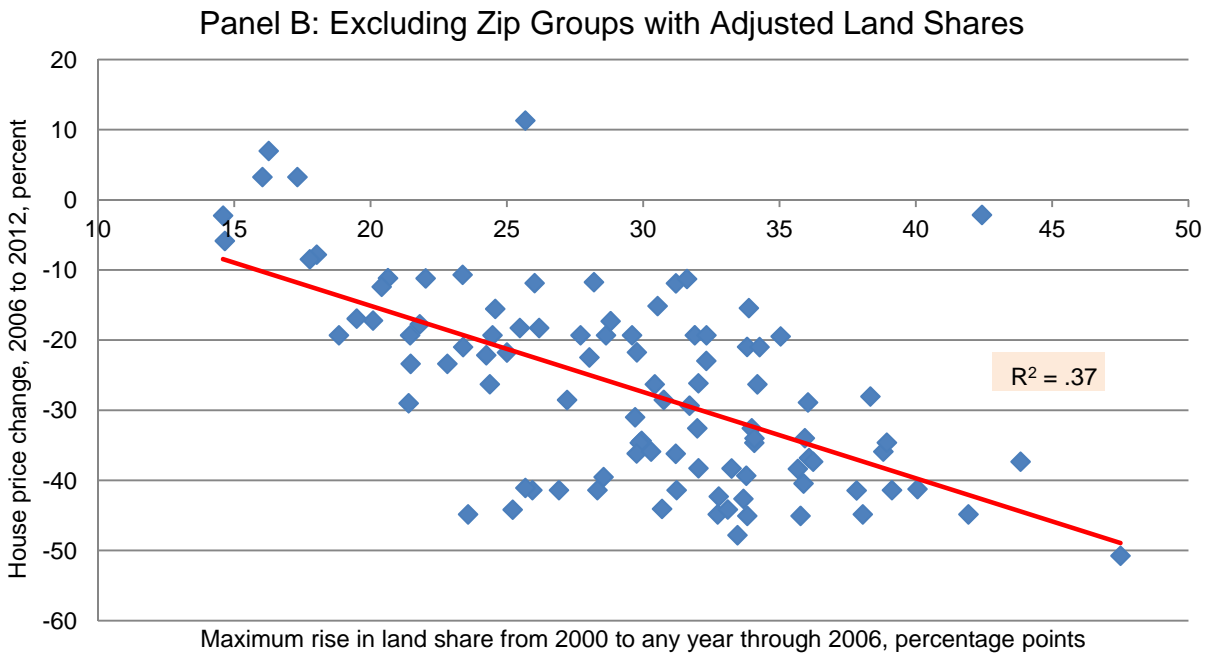
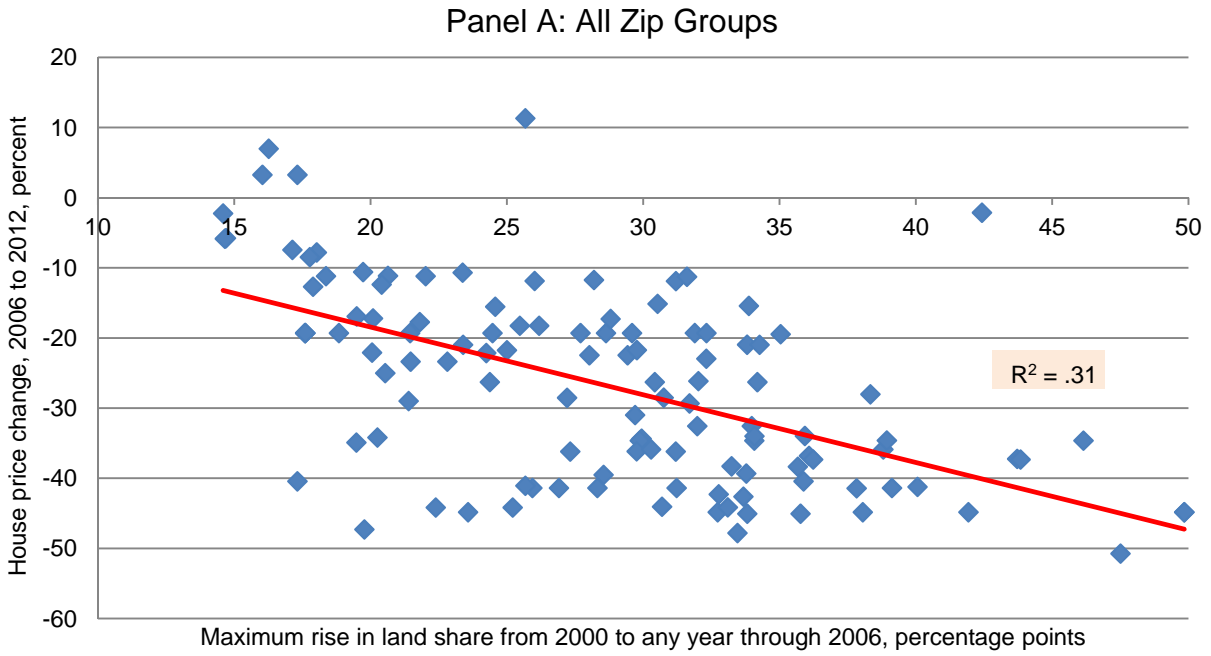
Figure 13
Average Land Share of Home Value
 (for zips grouped into quintiles by quarter-acre land price in 2000:Q1)



Note: Observations for 2013 represent average of Q1, Q2, and Q3.

Source: Authors' calculations using data from FNC, Inc. and Marshall & Swift, a CoreLogic company.

Figure 14
Predictive Power of Changes in Land Share



Source: Authors' calculations using data from FNC, Inc. and Marshall & Swift, a CoreLogic company.

Appendix

This appendix contains exhibits with supplementary information mentioned in the main part of the paper. Table A1 provides detailed information on every zip code and zip group in the dataset. Figures A1 through A4 are versions of figures 10 through 13, respectively, that exclude the zip groups with adjusted land shares.

Table A1. Detailed Information on Zip Groups

Group	Zip codes	City/Cities	County (or DC)	No. of homes	No. of new home sales	Type of sales used	Reference quarter	Land share	
								2000:Q1 (%)	Adjusted?
1	20007	Washington	DC	1473	29	H,L	2002:Q1	36.9	No
2	20008	Washington	DC	2263	21	H,L	2002:Q1	49.9	No
3	20012	Washington	DC	3176	3	H,L	2004:Q4	30.2	No
4	20015	Washington	DC	4101	19	H,L	2002:Q4	55.2	No
5	20016	Washington	DC	5493	90	H,L	2000:Q2	45.9	No
6	20017	Washington	DC	2312	14	H,L	2006:Q4	5.0	Yes
7	20018	Washington	DC	2876	21	H,L	2003:Q1	20.1	No
8	20020	Washington	DC	2855	72	H,L	2000:Q3	9.2	No
9	20032	Washington	DC	1238	39	H	2003:Q2	8.4	No
10	20001, 20009, 20010, 20011	Washington	DC	4410	10	H,L	2003:Q4	43.0	No
11	20002, 20003, 20019	Washington	DC	5044	202	H,L	2001:Q3	16.9	No
12	20120	Centreville	FF	5879	809	H,L	2000:Q1	60.0	Yes
13	20121	Centreville	FF	1587	117	H,L	2001:Q3	21.0	No
14	20124	Clifton	FF	3496	194	H,L	2000:Q1	46.0	No
15	20151	Chantilly, Fairfax	FF	4145	593	H,L	2000:Q1	60.0	Yes
16	20170	Herndon	FF	6685	331	H,L	2000:Q1	44.4	No
17	20171	Herndon, Oak Hill	FF	7864	644	H,L	2000:Q1	60.0	Yes
18	20194	Reston, Herndon	FF	1810	70	H,L	2000:Q2	27.5	No
19	22003	Annandale	FF	10,019	163	H,L	2000:Q3	56.6	No
20	22015	Burke, Springfield	FF	7205	123	H,L	2000:Q3	40.1	No
21	22030	Fairfax	FF	4024	647	H,L	2000:Q1	28.3	No
22	22031	Fairfax	FF	2871	173	H,L	2000:Q2	60.0	Yes
23	22032	Fairfax	FF	6931	175	H,L	2000:Q3	31.1	No
24	22033	Fairfax	FF	4386	501	H,L	2000:Q1	26.6	No
25	22039	Fairfax Station	FF	5308	153	H,L	2000:Q1	38.4	No
26	22042	Falls Church	FF	6329	72	H,L	2001:Q1	60.0	Yes
27	22043	Falls Church	FF	4009	232	H,L	2000:Q2	52.0	No
28	22066	Great Falls	FF	4932	369	H,L	2000:Q1	32.8	No
29	22079	Mason Neck, Lorton	FF	4118	2326	H,L	2000:Q1	53.0	No
30	22102	McLean	FF	3038	324	H,L	2000:Q1	60.0	Yes
31	22124	Oakton, Vienna	FF	3849	443	H,L	2000:Q1	34.9	No
32	22150	Springfield	FF	4584	265	H,L	2000:Q1	51.4	No
33	22151	Springfield, N. Springfield	FF	4511	36	H,L	2003:Q2	48.4	No
34	22152	Springfield, W. Springfield	FF	4586	124	H,L	2000:Q3	52.2	No
35	22153	Springfield	FF	6045	251	H,L	2000:Q1	33.8	No
36	22181	Vienna	FF	3062	127	H,L	2000:Q2	37.1	No
37	22182	Vienna	FF	6401	465	H,L	2000:Q1	51.3	No

Group	Zip codes	City/Cities	County (or DC)	No. of homes	No. of new home sales	Type of sales used	Reference quarter	Land share	
								2000:Q1 (%)	Adjusted?
38	22308	Alexandria	FF	4528	67	H,L	2000:Q4	60.0	Yes
39	22315	Alexandria	FF	2368	399	H,L	2000:Q3	36.4	No
40	22041, 22044	Falls Church	FF	3440	121	H,L	2001:Q2	58.8	No
41	20190, 20191	Reston	FF	3931	9	H,L	2001:Q1	57.2	No
42	22027, 22180	Vienna, Dunn Loring	FF	6218	436	H,L	2000:Q3	48.4	No
43	22046, 22213	Falls Church	FF	1255	50	H,L	2001:Q2	60.0	Yes
44	22060, 22309	Ft. Belvoir, Alexandria	FF	4821	355	H,L	2000:Q2	60.0	Yes
45	22101, 22207	McLean, Arlington	FF	8494	400	H,L	2000:Q2	60.0	Yes
46	22302, 22311, 22312	Alexandria	FF	3070	116	H,L	2000:Q2	60.0	Yes
47	22303, 22310	Alexandria	FF	6251	297	H	2000:Q2	32.3	No
48	22306, 22307	Alexandria	FF	5598	170	H,L	2000:Q3	51.5	No
49	20812, 20816	Bethesda, Glen Echo	MG	4417	119	H,L	2000:Q4	52.4	No
50	20815	Bethesda, Chevy Chase	MG	5759	123	H,L	2000:Q2	54.8	No
51	20814	Bethesda	MG	4664	223	H,L	2000:Q3	58.5	No
52	20817, 20818	Bethesda, Cabin John	MG	10,524	463	H,L	2000:Q1	49.6	No
53	20895, 20896	Kensington, Garrett Park	MG	5938	124	H,L	2000:Q2	27.1	No
54	20910	Silver Spring	MG	4986	59	H,L	2000:Q2	58.7	No
55	20901	Takoma Park, Silver Spring	MG	7638	8	H,L	2003:Q1	32.5	No
56	20902	Silver Spring, Wheaton	MG	9021	26	H,L	2000:Q4	27.6	No
57	20852	Rockville, N. Bethesda	MG	4291	23	H,L	2002:Q1	42.9	No
58	20854	Rockville, Potomac	MG	13,297	324	H	2000:Q1	39.4	No
59	20906	Silver Spring, Aspen Hill	MG	8599	158	H,L	2000:Q3	44.4	No
60	20853	Rockville	MG	8011	120	H	2000:Q2	25.1	No
61	20851	Rockville	MG	3227	1	None	None	40.0	Yes
62	20850	Rockville, Potomac	MG	6126	762	H,L	2000:Q1	26.9	No
63	20707, 20866	Laurel, Burtonsville	MG/PG	5692	1426	H,L	2000:Q1	23.2	No
64	20868, 20905	Silver Spring, Spencerville, Colesville	MG	5186	314	H	2001:Q1	36.7	No
65	20878	Gaithersburg, North Potomac, Darnestown	MG	9040	842	H,L	2000:Q1	43.7	No
66	20874	Germantown, Darnestown	MG	4017	624	H,L	2000:Q1	31.2	No
67	20837	Poolesville	MG	1342	49	H,L	2000:Q1	44.9	No
68	20777, 20833, 20860, 20861, 20862	Ashton, Highland, Brinklow, Sandy Spring, Brookeville	MG	3009	357	H	2000:Q1	30.0	No
69	20832	Olney	MG	4626	172	H	2000:Q2	26.1	No
70	20855	Derwood, Rockville	MG	3430	25	H	2002:Q2	5.0	Yes
71	20877, 20880	Washington Grove, Montgomery Village, Gaithersburg	MG	2888	74	H,L	2000:Q3	38.5	No
72	20879	Gaithersburg, Montgomery Village	MG	2751	169	H,L	2000:Q2	29.9	No
73	20886	Gaithersburg, Montgomery Village	MG	1797	1	None	None	40.0	Yes
74	20882, 21797	Gaithersburg, Woodbine, Laytonsville	MG	3844	370	H	2000:Q3	34.6	No
75	20876	Germantown	MG	3087	234	H,L	2000:Q1	31.5	No

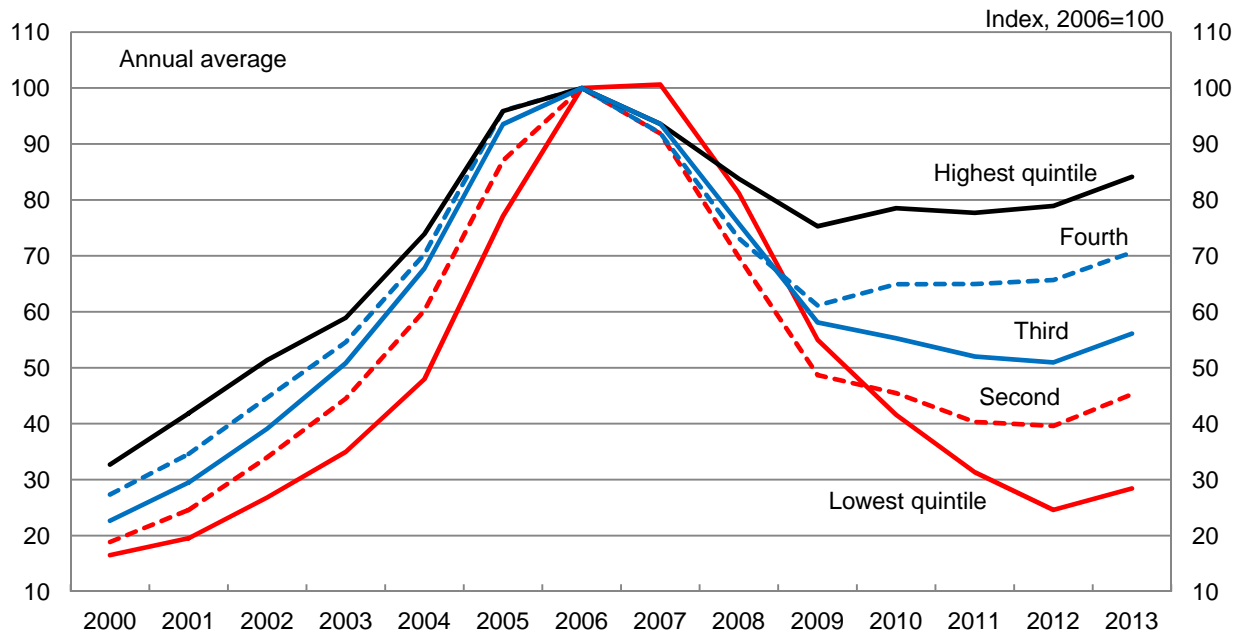
Group	Zip codes	City/Cities	County (or DC)	No. of homes	No. of new home sales	Type of sales used	Reference quarter	Land share	
								2000:Q1 (%)	Adjusted?
76	20838, 20839, 20841, 20842	Barnesville, Beallsville, Boyds, Dickerson	MG	2354	1343	H	2000:Q1	22.5	No
77	20872, 21771	Mount Airy, Damascus	MG	2816	164	H	2000:Q2	21.5	No
78	20871	Clarksburg, Hyattstown	MG	2232	1567	H	2000:Q4	5.0	Yes
79	20706	Lanham Seabrook, Glenarden	PG	7262	191	H,L	2000:Q2	31.2	No
80	20708	Laurel, Montpelier	PG	2596	244	H,L	2000:Q1	17.0	No
81	20715	Bowie	PG	8052	417	H	2002:Q2	6.3	No
82	20716	Bowie, Mitchellville	PG	3703	205	H	2000:Q2	22.5	No
83	20720	Bowie, Laurel	PG	4965	1180	H	2000:Q3	60.0	Yes
84	20721	Bowie, Mitchellville	PG	6156	1156	H,L	2000:Q2	60.0	Yes
85	20737	Riverdale	PG	2881	56	H,L	2004:Q2	55.1	No
86	20740, 20770	College Park, Berwyn Heights, Greenbelt	PG	5660	18	H,L	2004:Q1	11.5	No
87	20744	Fort Washington	PG	13,001	932	H,L	2000:Q1	25.1	No
88	20745	Oxon Hill, Forest Heights	PG	3642	3	H,L	2002:Q2	37.0	No
89	20746	Suitland, Morningside	PG	2906	20	H,L	2002:Q2	39.2	No
90	20747	District Heights, Forestville	PG	5215	192	H,L	2000:Q3	33.7	No
91	20748	Temple Hills, Hillcrest Heights, Camp Springs	PG	6088	141	H,L	2000:Q2	26.9	No
92	20769	Glenn Dale	PG	1933	389	H	2001:Q3	17.1	No
93	20772	Upper Marlboro	PG	9141	1923	H,L	2000:Q1	23.5	No
94	20774	Springdale, Upper Marlboro, Largo, Glenarden, Kettering	PG	8348	2153	H,L	2000:Q1	24.8	No
95	20781	Hyattsville	PG	2010	5	H,L	2003:Q3	5.0	Yes
96	20782	Chillum, University Park, Hyattsville	PG	4020	3	H,L	2001:Q3	37.6	No
97	20784	Hyattsville, Cheverly, Landover Hills, New Carrollton	PG	5226	1	H,L	2000:Q3	30.1	No
98	20785	Hyattsville, Landover, Cheverly	PG	3783	120	H,L	2000:Q1	26.0	No
99	20601, 20607	Accokeek, Waldorf	PG	3240	961	H,L	2000:Q1	22.6	No
100	20608, 20613	Brandywine, Aquasco	PG	3101	1028	H	2002:Q3	40.5	No
101	20623, 20735	Cheltenham, Clinton	PG	10,428	1123	H,L	2000:Q1	23.4	No
102	20705, 20904	Silver Spring, Beltsville, Colesville	MG/PG	12,163	1042	H,L	2000:Q1	30.3	No
103	20710, 20712, 20722	Brentwood, N. Brentwood, Bladensburg, Colmar Manor, Cottage City, Mt. Rainier Spring	PG	3064	24	H,L	2001:Q4	27.5	No
104	20783, 20903, 20912	Takoma Park, Hyattsville, Adelphi, Silver Spring	MG/PG	10,928	34	H,L	2000:Q4	28.7	No
105	20743	Capitol Heights, Seat Pleasant, Fairmount Heights	PG	7196	150	H,L	2000:Q2	31.8	No
106	20109	Manassas, Manassas Park, Sudley Springs	PW	2110	188	H,L	2002:Q1	60.0	Yes
107	20110	Manassas	PW	1354	133	H,L	2000:Q1	31.1	No
108	20111	Manassas, Manassas Park	PW	3533	608	H,L	2000:Q1	43.0	No
109	20112	Manassas	PW	7354	2301	H,L	2000:Q1	29.7	No
110	20119, 20181	Cattlett, Nokesville	PW	2037	403	H,L	2000:Q4	40.7	No

Group	Zip codes	City/Cities	County (or DC)	No. of homes	No. of new home sales	Type of sales used	Reference quarter	Land share	
								2000:Q1 (%)	Adjusted?
111	20136	Bristow	PW	5488	4339	H,L	2000:Q1	33.0	No
112	20169, 20137	Broad Run, Haymarket	PW	5283	3163	H	2000:Q2	22.4	No
113	20143, 20155	Catharpin, Gainesville	PW	7010	5537	H,L	2000:Q1	47.4	No
114	22205	Montclair, Dumfries	PW	3513	1117	H,L	2000:Q1	37.6	No
115	22026	Dumfries	PW	1835	1321	H,L	2000:Q1	55.0	No
116	22134, 22172	Triangle, Quantico	PW	1621	794	H,L	2002:Q1	60.0	Yes
117	22191	Woodbridge	PW	6892	1793	H,L	2000:Q2	27.7	No
118	22192	Woodbridge	PW	6995	1107	H	2000:Q2	27.9	No
119	22193	Woodbridge, Dale City	PW	13,529	2408	H,L	2000:Q1	40.6	No

Key

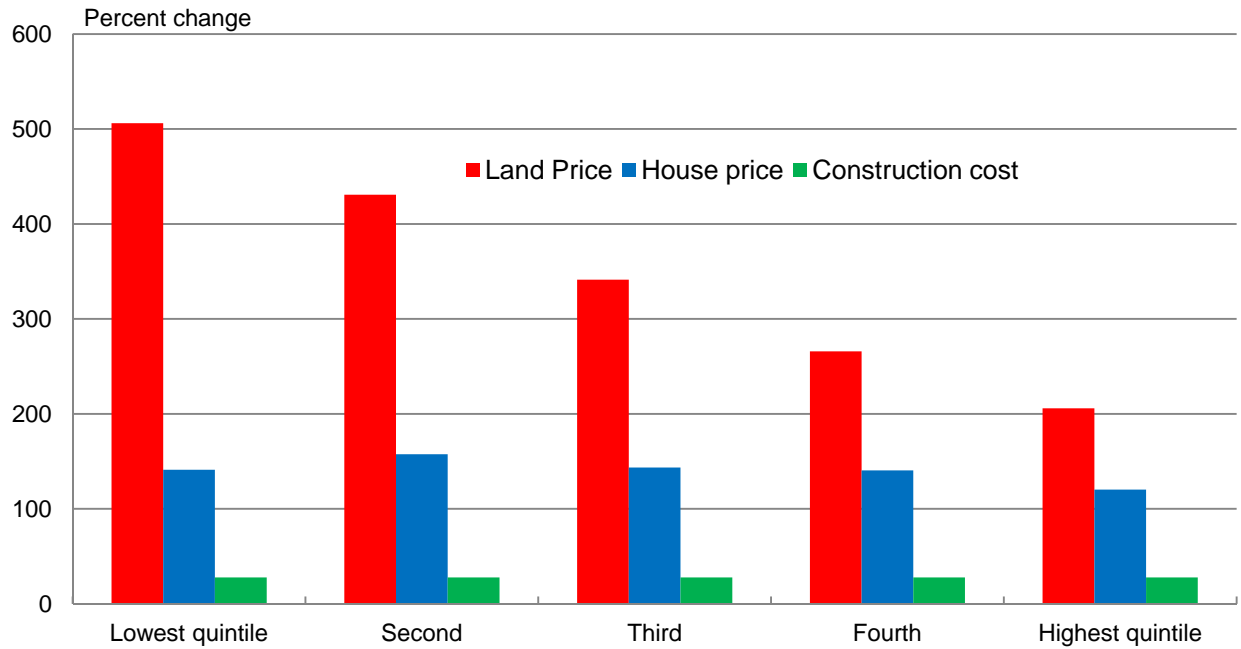
Counties: FF = Fairfax, MG = Montgomery, PG = Prince George's, PW = Prince William.
Type of sales: H,L = home and land sales, H = home sales only.

Figure A1
Average Land Prices
Excluding Zip Groups with Adjusted Land Shares
 (remaining zips grouped into quintiles by quarter-acre land price in 2000:Q1)



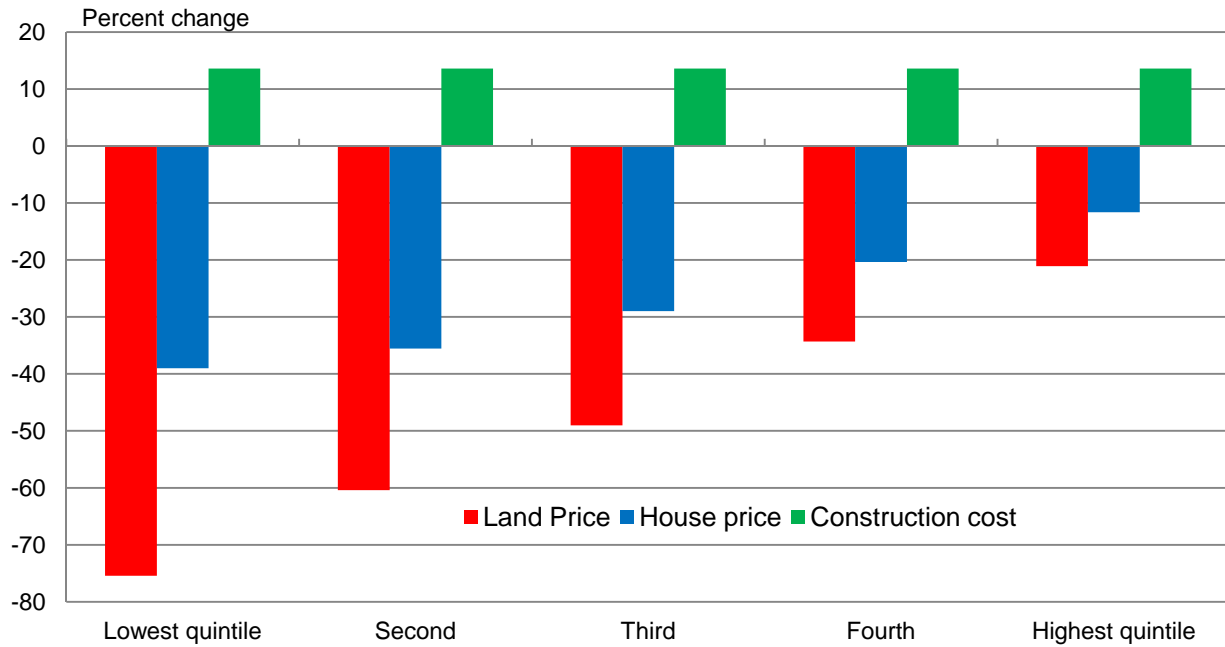
Note: Observations for 2013 represent average of Q1, Q2, and Q3.
 Source: Authors' calculations using data from FNC, Inc. and Marshall & Swift, a CoreLogic company.

Figure A2
Prices and Construction Costs: 2000-2006
Excluding Zip Groups with Adjusted Land Shares
 (remaining zips grouped into quintiles by quarter-acre land price in 2000:Q1)



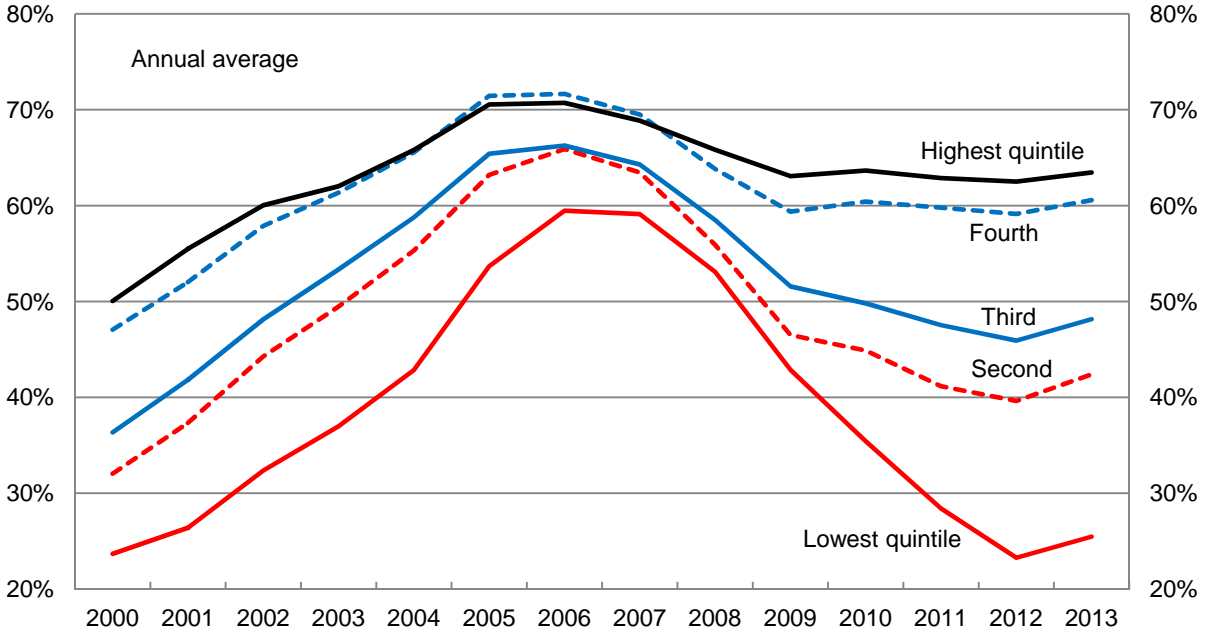
Source: Author's calculations based on data from FNC, Inc. and Marshall & Swift, a CoreLogic company.

Figure A3
Prices and Construction Costs: 2006-2012
Excluding Zip Groups with Adjusted Land Shares
 (remaining zips grouped into quintiles by quarter-acre land price in 2000:Q1)



Source: Author's calculations based on data from FNC, Inc. and Marshall & Swift, a CoreLogic company.

Figure A4
Average Land Share of Home Value
Excluding Zip Groups with Adjusted Land Shares
 (remaining zips grouped into quintiles by quarter-acre land price in 2000:Q1)



Note: Observations for 2013 represent average of Q1, Q2, and Q3.
 Source: Authors' calculations using data from FNC, Inc. and Marshall & Swift, a CoreLogic company.