Valuing Housing Services in the Era of Big Data: A User Cost Approach Leveraging Zillow Microdata

Marina Gindelsky[a+]
U.S. Bureau of Economic Analysis

Jeremy G. Moulton[b]
University of North Carolina – Chapel Hill

Scott A. Wentland[c+]
U.S. Bureau of Economic Analysis

January 17, 2019

Abstract

Historically, residential housing services or "space rent" for owner-occupied housing has made up a substantial portion (approximately 10%) of U.S. GDP final expenditures. The current methods and imputations for this estimate employed by the Bureau of Economic Analysis (BEA) rely primarily on designed survey data from the Census Bureau. In this study, we develop new, proof-of-concept estimates valuing housing services based on a user cost approach, utilizing detailed microdata from Zillow (ZTRAX), a "big data" set that contains detailed information on hundreds of millions of market transactions. Methodologically, this kind of data allows us to incorporate actual market prices into the estimates more directly for property-level hedonic imputations, providing an example for statistical agencies to consider as they improve the national accounts by incorporating additional big data sources. Further, we are able to include other property-level information into the estimates, reducing potential measurement error associated with aggregation of markets that vary extensively by region and locality. Finally, we compare our estimates to the corresponding series of BEA statistics, which are based on a rental-equivalence method. Because the user-cost approach depends more on the market prices of homes, we find that since 2001 our initial results track aggregate home price indices more closely than the current estimates.

PRELIMINARY DRAFT – Please contact the authors for the most recent draft before citing.

Keywords: residential housing, Big Data, housing services, owner-occupied, space rent, home prices
JEL Classifications: E01, C80, R00

Acknowledgements: We would like to thank the organizing committee and participants of the 2018 NBER-CRIW Pre-Conference, Erwin Diewert, Dennis Fixler, Kurt Kunze, Brent Moulton, Brian Smith, and Randal Verbrugge for their helpful comments. All errors are our own.

+Disclaimer: Any views expressed here are those of the authors and not necessarily those of the Bureau of Economic Analysis or the U.S. Department of Commerce. Data provided by Zillow through the Zillow Transaction and Assessment Dataset (ZTRAX). More information on accessing the data can be found at http://www.zillow.com/ztrax. The results and opinions are those of the author(s) and do not reflect the position of Zillow Group.

[a] Office of the Chief Economist 4600 Silver Hill Rd, Suitland, MD 20746; marina.gindelsky@bea.gov
[b] Department of Public Policy, Abernethy Hall, CB 3435, Chapel Hill, NC 27599; moulton@email.unc.edu
[c] *Contact Author*. Office of the Chief Economist 4600 Silver Hill Rd, Suitland, MD 20746; scott.wentland@bea.gov

1. Introduction

Housing is an important part of the economy and the national economic accounts. As part of its tabulation of Personal Consumption Expenditures (PCE) within Gross Domestic Product (GDP), the Bureau of Economic Analysis (BEA) estimates aggregate expenditure on housing, measuring what households in the United States spend on housing services. Because a house is generally a long-lasting asset and the flow of its services is not consumed it its entirety in a single year, housing is not measured like many other consumption expenditures as simply the aggregate of home prices and quantities. The flow of housing services in GDP are, as a result, measured as conceptually most similar to rent for these services in a given period. For renters (tenant-occupied housing), this tabulation is straightforward, both intuitively and from an economic measurement standpoint, as it amounts to the aggregate sum of rents paid for all residential units over a given period. But, for conceptual consistency due to the fact that homeowners do not pay rent explicitly, the analogous calculation imputes market rents (also called "space rent") for the owner-occupied housing stock as if owners "rent" to themselves. The 2008 System of National Accounts (SNA) recommends this imputation for owner-occupied housing so the estimate of housing services is not arbitrarily distorted based on the decision to rent versus own a home.[1] Historically, these aggregate housing estimates for both tenant and owner-occupied housing have accounted for a substantial proportion of overall consumer expenditures and the economy more generally (approximately 16% of PCE, or about 10% of GDP final expenditures), and have been relatively stable over recent decades as shown in Figure 1 below.

[1] Specifically, the 2008 SNA states: "The production of housing services for their own final consumption by owner occupiers has always been included within the production boundary in national accounts, although it constitutes an exception to the general exclusion of own-account service production. The ratio of owner-occupied to rented dwellings can vary significantly between countries, between regions of a country and even over short periods of time within a single country or region, so that both international and inter-temporal comparisons of the production and consumption of housing services could be distorted if no imputation were made for the value of own-account housing services." (SNA 2008, 6.35, p. 99).

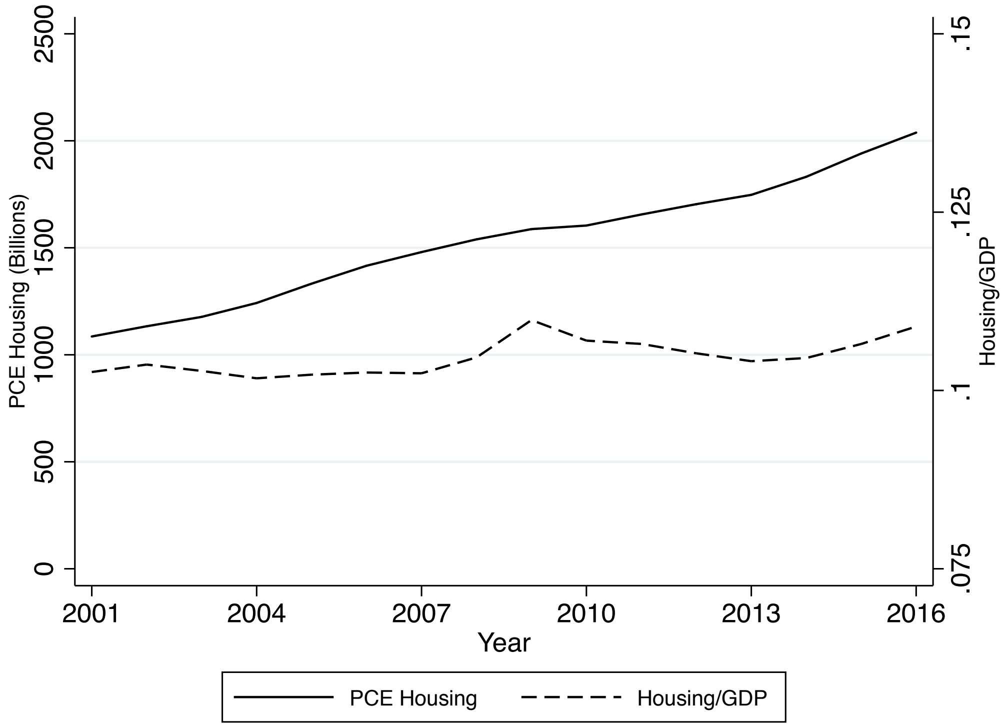

Figure 1: PCE Housing and PCE Housing/GDP
Source: U.S. Bureau of Economic Analysis, "Table 2.5.5: Personal Consumption Expenditures (PCE) by Function," bea.gov.

Yet, price indices of the national housing market such as the FHFA or Case-Shiller Price Index, while they do not exactly measure the same construct, show considerably more variation over time compared to housing services in PCE. To illustrate this contrast over the same time period, both the FHFA All Transaction Price Index and the Case-Shiller U.S. National Home Price Index are depicted below in Figure 2. A critical part of this difference is how housing services are measured and the corresponding underlying data. While indices like Case-Shiller are based on home *prices*, the BEA's current imputations of owner-occupied housing services primarily rely on designed survey data from the Census Bureau and a rental-equivalence method that bases its imputations on market *rents* of tenant occupied-homes. Hence, the purpose of this paper is to explore a method that relies more directly on market prices of the homes themselves, a user-cost approach, which utilizes "big data" from Zillow to provide a proof-of-concept alternative to the current rental-equivalence method used by BEA. However, we should state at the outset that this is not a paper about constructing an official account or arguing explicitly for a particular method; rather, we simply take the necessary first step of exploring its feasibility with a new "big data" source and provide initial estimates. Further, this also allows us to evaluate the extent to which the user cost method reflects broader price trends as compared to other data series.

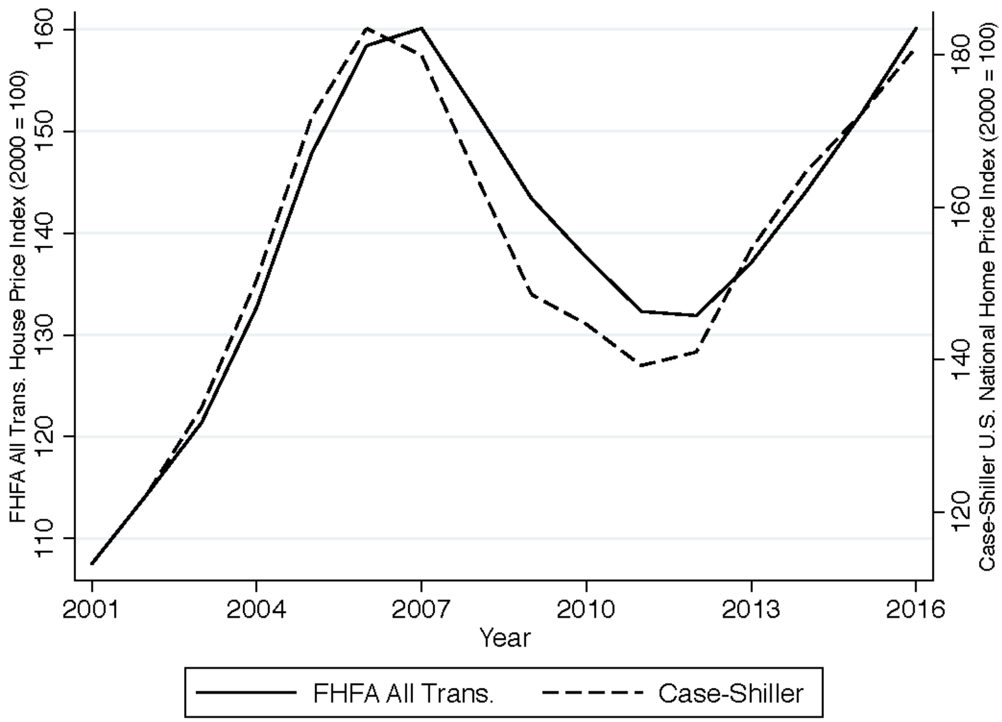

Figure 2: FHFA and Case-Shiller U.S. National Home Price Index
Source: https://fred.stlouisfed.org/series/CSUSHPINSA; /USSTHPI

The BEA's current approach based on rental data is the most common method used by national statistical agencies around the world (Katz 2017), in part due to the fact that countries collect high quality data on rents from nationally representative, specifically designed surveys of tenants and other sources. In contrast, data on home sales and corresponding home characteristics are primarily recorded by local municipalities, and thus often differ by locale, making a national effort to collect this data quite costly. Indeed, only in recent decades have most localities digitized these records, making rental survey data the most practical data source prior to the era of "big data." But, in the modern era companies like Zillow have privately undertaken a laudable effort to collect, compile, and organize a massive database of public data from local tax assessors' offices across the U.S. for the purposes of providing this information to users of their website. Zillow has recently provided much of their microdata to researchers free of charge, including those at BEA, which makes a user cost method based on fine-level price and home characteristic data more tractable, at least as a proof-of-concept effort to show how estimates built from national microdata stack up against current methods. This is important given that prior studies (for example, Verbrugge (2008), Garner and Verbrugge (2009), Aten (2018), and others) have found persistent and sizable differences between rental-equivalence and user cost methods using data from the Census Bureau and other aggregate data sources.

2. Background – Rental-equivalence vs. User Cost Approach

A central problem for statistical agencies is finding the right data; and, this is particularly true for imputing owner-occupied housing statistics, where the challenge is calculating transactions that are not *directly* measurable or observable. Hence, statistical agencies like the BEA measure the value of housing services indirectly using data that should closely approximate the market rent that homeowners expend. The two approaches briefly discussed above are those recommended by the 2008 SNA statistical framework: rental-equivalence and user cost. Conceptually, absent transactions costs and other market frictions, basic economic principles predict that market rents should approximately equal average cost (in the long run) if markets are competitive. The underpinning theory of this (approximate) equality can be derived from capital theory, which is based on Jorgenson's (1963, 1967) theory of capital and investment, where the rental cost of capital will equal its *ex ante* user cost (Katz 2009).[2] For example, if rent for an identical home was much higher than its user cost incurred by a homeowner, then more people would buy this preferred capital asset and fewer would rent, bidding down rents and bidding up home prices to the point where rents and costs are approximately equal.[3]

BEA's current method follows a rental-equivalence approach that uses data from the Census's Residential Finance Survey (RFS) to benchmark rent-to-value ratios for different value classes of properties, which is then used to impute average contract rent for owner-occupied properties across similar dimensions.[4] This weighted rental imputation constitutes what is often referred to as "space rent," which is then is multiplied by corresponding aggregate housing unit counts from the Census's American Housing Survey (AHS) to obtain the aggregate estimate of the total imputed rent of owner-occupied housing. For a more detailed discussion of the BEA's current method, refer to Mayerhauser and McBride (2007) and Katz (2017).

The rental-equivalence method is often cited as a preferred method for this imputation because most countries have relatively thick rental markets with substantial data on market rents. In fact, more than one-third of all housing units in the U.S. are rented to tenants. However, while the U.S. has a large number of tenant-occupied housing, the distribution of rental units is not the same as owner-occupied units (Glaeser and Gyourko 2009), as owner-occupied units have

[2] As a thought experiment, one can think of user cost in this context as measuring the net expenditure associated with purchasing a home at the beginning of a period, incurring cost during the period, and selling the home at the end of the period, abstracting away from transaction costs and other market frictions. According to Jorgensonian capital theory, the rental rate for this home set at the beginning of the period would equal this expected cost, ex ante. See also McFadyen and Hobart (1978) for an instructive cross-walk from Jorgenson (1967) to a user cost for housing in particular.

[3] Of course, this abstracts from risk, market imperfections, and transactions costs, which is particularly significant in housing (Bian, Waller, Wentland 2016). Thus, some gap might persist, but generally rents and user costs should move together over longer periods of time. In fact, recent empirical work by Goeyvaerts and Buyst (2019) has found a "strong correspondence" between rents and user costs using detailed micro-data.

[4] The BEA had last benchmarked these rent-to-value ratios using the 2001 RFS, the last time the data was available. Since then, the BEA has made quality and price adjustments primarily based on data from the BLS, which also relies on a rental-equivalence method for the CPI.

disproportionally more detached single-family residences (SFRs) and the distribution is skewed toward higher value homes. For additional discussion of this point and recent Census data illustrating these differences, see Aten (2018).

When rental markets are thin, the SNA recommends "other means of estimating the value of housing services," (SNA 2008, p. 109) which has led researchers and statistical agencies to explore alternative methods like a user cost approach, which utilizes data on the cost to the user of owning a home (e.g., interest, taxes, maintenance/depreciation, etc., which varies directly with the price of a home) rather than rents of different tenant-occupied homes. For an instructive review of this voluminous literature and novel examples of developing user cost estimates, see Diewert (2008a, 2008b), Katz (2004), Verbrugge (2008), Davis, Lehnert, and Martin (2008), Haffner and Heylen (2011), Hill and Syed (2016), Aten (2018) and numerous other papers on this topic.

A key advantage of the user cost approach is coverage of directly observable data. While tenant rents exist only for a subset of homes, a transaction price and corresponding costs associated with owning a home exist for nearly the universe of homes. While Gillingham (1983), Verbrugge (2008) and Diewert, Nakamura, and Nakamura (2009) and others have noted that the user cost approach has a number of weaknesses (e.g., greater volatility, sensitivity to interest rates, and conceptual issues with ex ante and ex post measurement), these would need to be weighed against weaknesses with the rental-equivalence approach (or any other approach, for that matter) to make the ultimate determination of which method to pursue. Nonetheless, weighing in on this debate falls outside the scope of this paper, as two necessary prerequisites for even considering a new approach are assessing whether it is feasible and conducting an initial evaluation of how the new estimates compare to the current approach, which is the aim of this paper.

3. Data

The novelty of this paper primarily resides with usage of new data. As we alluded to in the introduction, we use residential housing microdata from Zillow's ZTRAX data set. It contains transaction data as well as a large set of individual property characteristics for sales recorded from local tax assessor's data.[5] The data coverage is generally representative of the United States' national housing market, initially comprising 374 million detailed records of transactions across more than 2,750 counties.[6] This includes information regarding each home's sale price, sale date, mortgage information, foreclosure status, and other information commonly disclosed by a local tax assessor's office. We link each transaction to each home's property characteristics into a single dataset. The assessment data typically includes an array of characteristics one would find on Zillow's website or a local tax assessor's office describing the home, namely the size of the home (in square feet), number of bedrooms and bathrooms, year built, and a variety of other characteristics of the home.[7] We received all of this data in a somewhat raw form, requiring additional cleaning for research purposes.

We carefully scrutinized missing data and extreme values as part of our initial culling of outliers and general cleaning. The initial data set from Zillow contains sales of empty plots of land, some commercial property transactions, agricultural sales, and a host of types of properties that are outside the scope of the housing services estimates we aim to measure. Therefore, we limit the

[5] Data are provided by Zillow through the Zillow Transaction and Assessment Dataset (ZTRAX). More information on accessing the data can be found at http://www.zillow.com/ztrax. The results and opinions are those of the authors and do not reflect the position of Zillow Group. Nonproprietary code used to generate the results for this paper is available upon request of the authors.

[6] Because some states do not require mandatory disclosure of the sale price, we currently do not have price data for the following states: Idaho, Indiana, Kansas, Mississippi, Missouri, Montana, New Mexico, North Dakota, South Dakota, Texas, Utah, and Wyoming. Our method aggregates to the Census Division level by using housing unit counts from the ACS at the regional level. As a result, we assume that the states with data within a Census Division are reasonably representative of a state left out, which is an assumption we hope to explore in further research with supplemental data.

[7] Zillow's ZTRAX data contains separate transaction files by state, where all transactions need to be linked to corresponding assessment records. With guidance from Zillow, we were able to merge the bulk of the data, but not without some data loss (which figures into the size of our final sample).

sample to single family homes, townhouses, rowhomes, apartments, condos, and properties that are most closely associated with the current estimates. We winsorize acreage at five acres (limiting the influence of large farms) and outlier homes that are on the upper tail of the distribution (i.e. are larger than 10,000 square feet or have more than five bedrooms, more than three bathrooms).[8] We also drop homes from the prediction regression sample that sold for less than 20,000 dollars, the bulk of which are not arms-length transactions. We further cull homes that were built prior to 1865 or report a negative age of home (i.e. sale year – year built). While the Zillow data set contains a vast number of property characteristics, in our initial analysis we primarily rely on the variables described above that have the most coverage nationally to limit how much data we effectively discard.[9] We limit the sample years to 2001 through 2016, as the data is most complete for the vast majority of the states in our sample.

To ensure the quality of the final sample, we compared our cleaned Zillow sample to the U.S. Census American Community Survey (ACS) to ensure that this administrative data aligned with carefully collected (albeit more limited) survey data provided by the Census. Generally, there are only a limited set of characteristics of homes in both the ZTRAX data and the ACS (e.g., number of bedrooms, year built, number of rooms, tax amount, and an indicator for whether the property has more than 10 acres). When we compare them in aggregate, we find that they are quite similar in terms of their summary statistics. In untabulated results, we found that these shared variables across data sets had median and mean values that fell within a few percentage points of one another.

[8] We also create indicator variables equal to one if the property had missing characteristic values or reported a lot size of zero or there are missing bedrooms or bathrooms.
[9] In untabulated regressions, we conducted a sensitivity analysis for subsets of the sample that employed more property characteristics to determine whether the results are sensitive to omitted variables for which we can control. Our results were generally robust to omitting variables that have more limited coverage.

4. Methodology – An Idiosyncratic User Cost Approach

4.A. Overview

Generally, our approach using this microdata is motivated by constructing estimates from the bottom-up, as we estimate a user cost for *each individual property* in our dataset and then aggregate upward to produce a weighted national-level estimate. We begin by estimating a simplified user cost of housing services for each home in the data set using the formula:

$$U_{it} = P_{it}(i_t + \gamma_i + \tau_{it} - \mathrm{E}[\pi_i])$$

where for a given property (i) in quarter (t) P is the price of an individual home, i is the individual's discount rate (as a proxy we use the average nominal interest rate on a 30-year mortgage in quarter t),[10] γ is a constant representing "housekeeping expenses" of depreciation and maintenance cost of 3.5%,[11] τ is the individual property's effective tax rate, and $\mathrm{E}[\pi]$ is expected appreciation (revaluation) for a given year as 2%, which assumes homeowners have a very long-term view of home prices appreciating approximately the same as overall inflation in the economy.[12] We vary the latter assumption in a second user cost calculation discussed later in the paper, where price expectations are based on recent home price appreciation/depreciation in one's local area. Our

[10] While the dataset includes individual interest rates for transacted properties, the coverage is not as universal as other variables. However, it is customary for user cost estimates to use a single market interest rate to reflect the financial opportunity cost of the long-term asset (e.g., see Aten 2018). Conceptually, if a homeowner purchased a home with a 4% mortgage, but rates have since risen to 7%, the latter rate more closely represents the opportunity cost in that time period, as the homeowner could alternatively be earning a return on that equity of a similar long-term asset. The results and time series dynamics are similar if we use 10-Year Treasury or 30-Year Treasury rates.

[11] A depreciation rate of 1.5% is common to the literature (e.g., Aten (2018) and Verbrugge (2008)), and Gill and Haurin (1991) use a constant of (1.5% + 2% =) 3.5% for the combined maintenance and depreciation term. Conceptually, there is wear and tear on a home that would be similar to what a renter would incur in the analogous tenant-occupied counterfactual. Because these costs (on average) would be priced into a tenant's rent, it is logical to factor this into the imputation for owner-occupied properties.

[12] Verbrugge (2008) rigorously considered a variety of measures of $\mathrm{E}[\pi]$ using different forecast techniques, concluding that, "a very long horizon appreciation forecast (such as a long moving average), or an inflation forecast, should be used in the user cost formula" (p. 694). During the period we study, the Federal Reserve had maintained either an explicit or implicit target of 2% inflation over the long run (see, for example, their policy statements on their website regarding 2%: https://www.federalreserve.gov/faqs/money_12848.htm). Ex post, inflation, particularly in the housing market, departed from this target; but, use as an *ex ante* measure may not be unreasonable. For robustness, we consider a method where $\mathrm{E}[\pi]$ is determined by recent experience with price inflation in one's local area.

primary contribution to the literature is estimating national property-level user costs using idiosyncratic price and property tax data, which we describe in more detail below.

4.B Idiosyncratic P – Actual and Predicted

Because we have fine, transaction-level price data, we are able to use actual market prices for *P* (when available). While turnover varies considerably by state and locality, approximately one-third of properties in our dataset sold at least once within the window we study (from 2001-2016). If property *i* was purchased in the first quarter of 2010, for example, then for that quarter the *actual* price was used for the transacted property (P in the formula above). For the value of the home in the following quarter we posit that the price is simply the transacted price plus the average price appreciation/depreciation of the housing stock of the county (which we estimate using the same hedonic model we use for our price imputations discussed below). We use the same logic for the quarters proceeding that sale until there is a new sale of that property. We also apply this logic backward in time for a given property's first sale in this sample period.[13] This conforms most closely to the principles of valuation laid out by the System of National Accounts (SNA), where market prices are "the basic reference for valuation in the SNA" (SNA 2008, p. 22),[14] and thus much of our aggregate calculation flows directly from millions of observed market prices underlying the housing stock.

For homes that did not sell during our sample period, we impute their prices based on transactions of similar homes that sold in each quarter using a hedonic model.[15] Conceptually,

---

[13] This method would likely be altered if it were implemented in national accounts over a longer time-series, as a single transaction price adjusted for inflation may be less predictive of the actual price in other years as the time series becomes much longer. For example, we may limit forward and backward interpolations to a single five or ten year window; but, because our time series here only covers fifteen years we take the simplified approach.
[14] More specifically, the SNA recommends that statistical agencies use market prices when market prices are available, but "in the absence of market transactions, valuation is made according to costs incurred (for example, non-market services produced by government) or by reference to market prices for analogous goods or services (for example, services of owner-occupied dwellings)" (SNA 2008, p. 22).
[15] Within-quarter hedonic regressions avoid issues of controlling for macro-level relevant time-varying factors that could bias predictions if not properly accounted for in the model.

most of a home's value can be explained by its physical characteristics, location, and time; hence, our hedonic model uses sale prices of similar homes along these dimensions to estimate an imputed market valuation for each home in our data set. Therefore, we impute $\hat{P}$ based on the following hedonic model for each quarter separately:

$$Sale\ Price_{ij} = \alpha + \sum \beta X_i + \gamma LOCATION_j + \sum \delta\, sq.ft._i * LOCATION_j + \sum \varphi\, acreage_i * LOCATION_j + \varepsilon$$

where X is a set of physical characteristics (bedrooms, bathrooms, age of the structure, living area measured by square feet, lot size measured by acreage, whether the home was a single story, whether the home had a basement, and whether the home was new construction), location fixed effects, and interaction of location fixed effects with square footage and acreage, respectively.[16] For practicality in estimation, we initially use zip code fixed effects, although we obtain similar estimates (albeit, more precise model fit with higher R^2) using finer-level geographic fixed effects like Census block groups and Census tracts.[17] To avoid making predictions with thin cells, we specify that a given zip code have at least ten sales in the quarter of estimation. If not, we estimate the same model only for observations that do not meet this threshold using county (FIPS) level fixed effects. While intensive for processing, allowing square footage and acreage to vary by location encapsulates the idea that valuation of these attributes varies widely across areas. For

[16] While the Zillow ZTRAX data contains a lot more information about individual properties that would help with valuation, we chose the variables with extensive coverage across all states in the data set. When compared to a fuller model that includes many more home characteristics, the marginal gain in precision was small compared to the potential loss in observations due to missing data in states/localities that do not regularly report certain variables. When one of the key characteristics (e.g. bedrooms) was missing, we bottom coded it and included a missing indicator in the regression rather than drop the sale entirely. We also included an indicator in the regression for whether the home had extreme values for any of these characteristics to account for non-linearities, as opposed to just dropping these observations as well.

[17] We have also explored a semi-log specification, where sale price is logged, which produces similar results given how we treat outliers in the model. Indeed, the model fit is improved with the semi-log form in other specifications.

example, an additional 500 square feet in a home in New York City, will be valued much differently than the same addition upstate in Syracuse.[18]

4.C. Property Taxes

Property taxes vary widely across states and municipalities. As of 2017, the highest property tax state was New Jersey with an average effective tax rate of 2.31%, whereas Hawaii and Alabama have average rates of 0.32% and 0.48%, respectively.[19] Even within states there is considerable variation. Hence, for accurate estimates of user cost we attempt to account for the idiosyncratic nature of a property's taxes. Because the Zillow data is collected primarily from local tax assessor office databases, the coverage of property taxes is quite good. We use individual tax data to determine a property's effective tax rate based on a denominator of P (actual or predicted price) rather than the corresponding assessment value associated with each property in the data.[20]

We made this choice for a couple reasons. First, regarding the denominator, the assessment value is often much lower than the market value, so if we apply the rate based on the assessed value to the market value of P in the user cost calculation we would overestimate the amount homeowners pay in our calculation. The degree of mis-assessment of value varies considerably by locale, and in some cases it is by design of local policies for states like California to have assessments tied to historical values for longer tenured homeowners. Second, this approach better reflects the average effective tax rate, because like other elements of the tax code, homeowners do not all pay the same posted rate due to local property tax relief exemptions and relief for special groups (Moulton, Waller, and Wentland 2018). Finally, in the present study we are unable to

[18] This approach is used commonly in the hedonic valuation literature for housing and land. See, for example, Kuminoff and Pope (2013).

[19] Variation in property taxes across states gained attention during the national coverage of the Tax Cuts and Jobs Act of 2017. For example, USA Today ran a story comparing effective property tax rates across the U.S.: https://www.usatoday.com/story/money/personalfinance/2017/04/16/comparing-average-property-taxes-all-50-states-and-dc/100314754/

[20] We currently have one year of tax amount data from Zillow, but updating this data more often (preferably annually) may be required if this method is to be used for national accounts measurement.

accurately determine the *net* tax bill for each homeowner or precisely consider the full range of offsetting tax benefits that come with homeownership (namely, mortgage interest deductions and state and local tax deductions on federal taxes); but, if we are able to successfully link this data with administrative data, then we will be able to construct a credible estimate of these benefits in future work.[21]

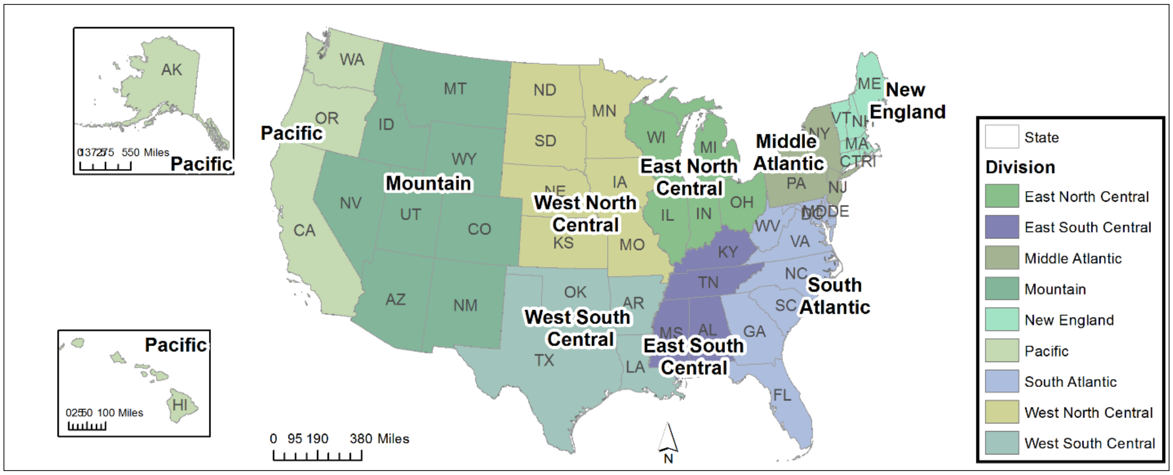

Figure 3: Census Divisions
Source: https://www.census.gov/geo/reference/webatlas/divisions.html

4.D. Quantity, Housing Counts, and Aggregation

Once we obtain user-cost estimates for millions of individual properties across the United States, we then aggregate to a weighted national estimate of housing services based on the corresponding quantities of the housing stock by location/region, type of home (single family residence (SFR) vs. non-SFR), and number of bedrooms. We use the weighted unit counts of the housing stock from Census's American Community Survey (ACS), which provide a yearly count of the aggregate number of residential housing units by Census Division, depicted in Figure 3. For illustrative purposes, refer to Table 1, where we show the calculation of our national estimate

[21] Linkages to Census administrative data records, for example, would also allow us to better estimate maintenance and other costs for households (or, at least regionally – where wear and tear from climate and other factors may contribute to households reporting systematically different levels of maintenance expenditures) and to better understand housing market dynamics of populations of homeowners vs. renters.

for Q4 of 2016. For each Census Division or region of the U.S., we multiply the average user cost for each type of home (SFR vs. non-SFR) for each bedroom category.[22]

This method of aggregation assumes that the non-missing data is reasonably representative of the missing data. For example, Indiana's sale prices are missing from the ZTRAX data set, as it is among the non-disclosure states that does not ordinarily record sale prices in public use tax assessor data. Hence, our final aggregate estimates must assume that the average user costs imputed from sales in its region (Illinois, Michigan, Ohio, and Wisconsin) reflect the Indiana market.[23] Missing data itself is not a prohibitive limitation for constructing national accounts, as statistical agencies always have limited data; but, the issue is more a matter of the extent of the representativeness of the data we do have.

While many of these states are reasonably represented by their neighboring states' housing markets, as the Indiana case may be, one exception might be Texas (the largest state for which we have missing price data) where the current method may be the most problematic, simply because of the variability of the housing markets within the state. If this method, or some variation of it using similar data, were to be adopted by the BEA, supplemental data would be required to verify these assumptions or to re-weight the estimates to better represent the missing states' housing markets. The scope of this study, however, is to explore how far this particular "big data" set can go toward this end.[24]

[22] We use bedrooms as a proxy for size of the home to create categorical differences that more accurately reflect the weighted total. The bins are numbered 1 through 5+ in Table 1. However, for states that did not have good coverage of the number of bedrooms, we assumed that the distribution of user cost approximately aligned with the distribution of bedrooms and assigned homes to corresponding bins of bedrooms. In future work, we will explore using county-level quantity counts, as finer location averages could be more relevant that averages by physical characteristics.

[23] Recall that one of the limitations of this data set is that there is no price data from the following states: Idaho, Indiana, Kansas, Mississippi, Missouri, Montana, New Mexico, North Dakota, South Dakota, Texas, Utah, and Wyoming. Maine is also excluded due to limited data in a number of quarters of our sample period.

[24] The American Housing Survey (AHS) also has high quality data on the unit counts of the housing stock, but the survey is only available every other year. While the counts are not always identical across surveys, the differences are relatively small. In future work, we plan to use linked Census data or Zillow's assessment data to construct our own unit weights with our data.

Table 1: User Cost Aggregation – Summary Calculation for 2016q4

		SFR			Non-SFR		
Division	**Bedrooms**	**Ave. User Cost ($)**	**Q**	**P*Q ($B)**	**Ave. User Cost ($)**	**Q**	**P*Q ($B)**
New England	1	12,427	78,841	1	22,559	767,907	17
	2	14,681	490,837	7	18,242	1,008,062	18
	3	19,305	1,605,454	31	25,675	418,328	11
	4	30,048	838,200	25	20,360	83,239	2
	5	49,335	206,494	10	22,261	29,104	1
Middle Atlantic	1	6,401	141,935	1	48,426	2,591,742	126
	2	10,158	1,031,546	10	12,662	2,622,026	33
	3	15,710	3,609,221	57	9,099	1,756,793	16
	4	27,799	2,237,124	62	11,116	303,307	3
	5	58,282	583,540	34	12,212	115,127	1
East North Central	1	4,319	218,903	1	16,533	1,755,527	29
	2	6,429	1,938,344	12	11,818	2,484,126	29
	3	11,151	6,567,881	73	15,881	796,746	13
	4	22,073	2,990,114	66	14,092	107,214	2
	5	33,863	670,174	23	15,910	36,214	1
West North Central	1	8,025	146,868	1	19,499	767,992	15
	2	9,914	1,043,513	10	12,400	955,087	12
	3	14,119	2,680,432	38	15,830	284,014	4
	4	20,643	1,526,842	32	15,111	53,393	1
	5	25,653	474,931	12	16,925	12,170	0
South Atlantic	1	7,156	196,669	1	16,131	2,046,952	33
	2	9,174	1,919,499	18	15,593	3,258,395	51
	3	14,448	7,543,817	109	22,214	1,598,763	36
	4	27,072	3,747,649	101	29,760	243,581	7
	5	65,141	1,105,643	72	40,482	34,846	1
East South Central	1	3,227	93,315	0	13,776	446,856	6
	2	4,207	734,721	3	11,173	693,547	8
	3	7,853	2,895,815	23	12,793	214,781	3
	4	16,650	1,058,912	18	15,058	27,240	0
	5	35,857	246,716	9	16,955	5,556	0
West South Central	1	1,832	192,651	0	12,973	1,382,770	18
	2	5,083	1,171,105	6	4,291	1,346,042	6
	3	9,532	4,647,022	44	12,406	384,662	5
	4	24,412	2,158,298	53	0	55,459	0
	5	19,780	415,247	8	0	10,433	0
Mountain	1	10,802	129,086	1	14,315	786,290	11
	2	14,008	762,322	11	14,279	1,061,530	15
	3	15,726	2,602,678	41	19,680	368,171	7
	4	26,479	1,598,170	42	30,294	54,078	2
	5	46,510	626,676	29	39,849	10,203	0
Pacific	1	20,746	316,702	7	38,308	2,534,134	97
	2	24,809	1,578,474	39	28,077	2,880,140	81
	3	28,878	5,078,692	147	34,520	944,312	33
	4	40,681	2,935,190	119	36,507	152,479	6
	5	59,695	758,066	45	40,828	35,263	1
			Subtotal (SFR)	1,454		*Subtotal (non-SFR)*	761
						Total User Cost: 1,454 + 761 =	**2,215**

4.E Varying Ex Ante Expected Price Appreciation/Depreciation

Finally, we vary the E[π] term of ex ante expected price appreciation for robustness. Our default specification assumes a very long-run view of home price inflation of a constant 2% per year, despite the fact that homeowners during this period may very well have perceived price appreciation quite differently. To test how the results differ if homeowners had drastically different expectations than we assume in our default specification, establishing a lower bound of sorts, we assume the opposite end of the spectrum for our alternative specification. That is, if our default is that homeowners take a *constant long-run, national* view of price expectations, then the opposite might be a *variable short-run*, *local* view of price expectations. Thus, our alternative specification assumes that homeowners expect ex ante price appreciation to be their local (county-level) average price inflation from the prior quarter. This is calculated by taking the percent change of the median predicted price by county by quarter from our hedonic model estimates discussed above.[25] While this is somewhat simplistic, our goal is to provide a sense of a reasonable range of possible estimates, as a more moderate moving average as in Verbrugge (2008) may produce an estimate somewhere in between this range of results, albeit closer to the long-run default specification.[26]

5. Results

Our full set of results for all years and quarters in our sample appear in Table 2 and Figure 4, which shows both the total and average user cost estimates of housing services as well as the

---

[25] Note that this is not seasonally adjusted. Some of the volatility in prices will be from purely seasonal factors. This can be augmented by applying a standard seasonal adjustment. For now we report the raw, unadjusted results.

[26] Generally, countries that employ a user cost method for housing omit the E[π] term entirely, simplifying the calculation (Diewert and Nakamura 2009). One way of thinking about this simplification involves referring back to the reason why the E[π] term is factored in the calculation in the first place. As a thought experiment, the user cost method is often pitched as calculating the cost of an owner who purchases a home at the beginning of a period and sells it at the end (assuming away transactions costs). The E[π] term in that case would simply be the capital gain/loss during a given period; but, if the next period begins with repurchasing the same home at the price from the end of the last period, then the capital gain/loss is essentially erased immediately. For now, we remain somewhat agnostic to the different approaches by offering results for multiple ways of incorporating E[π] into user cost; our default specification comes at the suggestion of feedback we received from the NBER-CRIW Pre-Conference in 2018.

corresponding estimates by housing type (SFR vs. non-SFR). The first column in each panel provides estimates for our default specification, while the second provides the alternative specification that allows for price expectations to vary by quarter based on recent experience in the housing market. As expected, the latter specification shows greater volatility over time, generating some quarters with very small user cost values due to high expected price appreciation in those quarters, if expectations are based on very recent, very local price inflation. For simplicity in discussing the remaining results, we focus on the default specification as it is closer to more reasonable long-run expectations, ex ante.

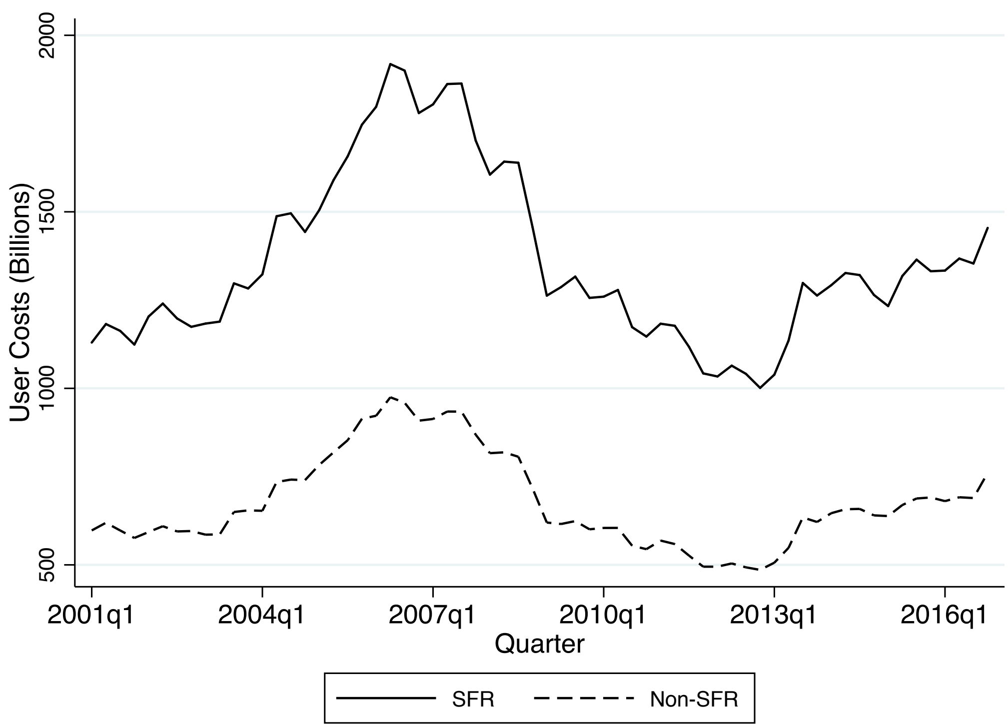

Figure 4: Total User Costs by SFR/Non-SFR (Default Specification)

Table 2: Housing User Costs by Quarter from 2001 through 2016

	Full Sample				SFR				Non-SFR			
	Total User Cost ($B)	Total Alt. User Cost ($B)	Ave. User Cost ($)	Ave. Alt. User Cost ($)	Total User Cost ($B)	Total Alt. User Cost ($B)	Ave. User Cost ($)	Ave. Alt. User Cost ($)	Total User Cost ($B)	Total Alt. User Cost ($B)	Ave. User Cost ($)	Ave. Alt. User Cost ($)
2001q1	1,727	1,381	17,586	14,056	1,130	1,013	17,245	15,463	598	367	16,342	10,870
2001q2	1,801	1,182	18,341	12,030	1,182	805	18,108	13,027	619	377	16,986	10,927
2001q3	1,760	1,115	17,921	11,349	1,163	716	17,934	10,444	597	399	16,579	10,853
2001q4	1,700	1,676	17,311	17,068	1,124	1,069	17,390	15,627	576	607	16,172	15,744
2002q1	1,796	1,940	18,121	19,564	1,203	1,353	18,201	20,005	593	587	16,763	15,511
2002q2	1,850	1,423	18,657	14,356	1,240	939	18,771	14,735	609	484	17,221	14,576
2002q3	1,793	843	18,083	8,500	1,198	514	18,169	7,533	595	329	16,836	8,120
2002q4	1,770	1,158	17,854	11,683	1,174	825	17,824	12,249	596	333	16,796	11,012
2003q1	1,769	1,469	17,645	14,652	1,183	1,069	17,627	15,810	586	399	16,419	11,731
2003q2	1,775	1,407	17,706	14,035	1,189	923	17,698	14,666	586	484	16,346	14,209
2003q3	1,947	965	19,421	9,621	1,297	609	19,284	9,260	650	355	18,044	10,638
2003q4	1,937	1,070	19,323	10,669	1,283	765	19,039	11,616	654	305	18,121	10,343
2004q1	1,976	1,676	19,423	16,475	1,323	1,210	19,221	18,021	653	466	17,964	13,212
2004q2	2,222	1,239	21,843	12,179	1,487	866	21,579	13,274	735	373	20,018	11,849
2004q3	2,237	461	21,992	4,528	1,496	301	21,654	4,585	742	159	20,092	6,291
2004q4	2,183	1,342	21,459	13,193	1,443	904	20,941	13,724	740	438	20,196	13,312
2005q1	2,288	1,942	22,235	18,874	1,505	1,363	21,466	19,511	783	580	20,981	14,991
2005q2	2,408	1,151	23,399	11,186	1,590	833	22,634	13,433	819	318	21,773	10,572
2005q3	2,510	592	24,393	5,757	1,657	321	23,560	5,047	853	271	22,530	8,753
2005q4	2,661	1,573	25,854	15,282	1,747	1,025	24,779	14,939	914	548	24,093	16,468
2006q1	2,720	2,516	26,304	24,332	1,797	1,766	25,136	25,427	922	749	24,400	20,394
2006q2	2,893	2,579	27,982	24,945	1,919	1,773	26,839	25,575	975	806	25,750	21,355
2006q3	2,859	2,078	27,654	20,101	1,900	1,310	26,593	18,231	959	768	25,349	20,976
2006q4	2,688	2,942	25,998	28,457	1,780	1,935	24,946	26,868	908	1,008	23,974	26,461
2007q1	2,718	3,445	26,036	33,001	1,804	2,363	24,969	32,402	913	1,082	23,978	28,526
2007q2	2,796	2,729	26,790	26,143	1,862	1,868	25,820	26,342	934	861	24,511	23,063
2007q3	2,798	2,350	26,802	22,516	1,863	1,533	25,926	20,615	934	818	24,657	21,768
2007q4	2,571	3,145	24,629	30,132	1,702	2,055	23,723	27,830	869	1,090	22,916	27,408
2008q1	2,422	3,821	23,035	36,341	1,605	2,605	22,276	35,690	816	1,215	21,323	31,842
2008q2	2,461	3,313	23,408	31,513	1,642	2,266	22,814	31,039	819	1,047	21,432	27,724
2008q3	2,445	2,657	23,259	25,272	1,639	1,678	22,834	22,925	806	979	21,242	25,888
2008q4	2,173	2,967	20,665	28,219	1,456	1,930	20,399	26,408	716	1,037	18,948	25,126
2009q1	1,882	3,368	17,811	31,866	1,263	2,254	17,632	30,598	620	1,114	16,188	28,754
2009q2	1,902	3,094	18,001	29,278	1,287	2,047	18,016	28,523	616	1,047	16,082	27,772
2009q3	1,941	1,461	18,362	13,822	1,317	868	18,474	11,787	624	593	16,348	15,430
2009q4	1,857	1,731	17,570	16,376	1,256	1,139	17,611	15,381	601	592	15,785	14,793
2010q1	1,864	2,097	17,484	19,671	1,260	1,446	17,491	20,364	605	651	15,647	16,268
2010q2	1,883	2,248	17,662	21,086	1,279	1,526	17,806	21,387	605	723	15,671	19,568
2010q3	1,728	1,176	16,202	11,028	1,173	747	16,371	9,886	554	429	14,384	10,381
2010q4	1,691	2,225	15,859	20,871	1,147	1,518	16,003	20,938	544	707	14,128	18,567
2011q1	1,752	2,184	16,359	20,396	1,183	1,511	16,472	20,876	569	674	14,525	17,991
2011q2	1,736	2,337	16,211	21,823	1,177	1,618	16,382	22,895	559	719	14,286	18,467
2011q3	1,644	1,188	15,349	11,098	1,118	748	15,590	10,457	526	441	13,462	11,497
2011q4	1,537	1,632	14,353	15,235	1,042	1,083	14,506	14,504	495	548	12,782	13,112
2012q1	1,529	2,046	14,144	18,936	1,034	1,409	14,276	19,910	495	638	12,595	15,566
2012q2	1,568	1,609	14,512	14,890	1,064	1,139	14,685	15,804	504	470	12,765	11,546
2012q3	1,534	296	14,195	2,741	1,041	157	14,388	2,208	493	139	12,477	4,393
2012q4	1,487	1,039	13,760	9,613	1,001	701	13,814	9,520	486	337	12,290	8,812
2013q1	1,545	1,351	14,241	12,450	1,039	1,033	14,272	14,493	506	318	12,635	7,972
2013q2	1,684	1,335	15,520	12,305	1,135	956	15,586	14,109	548	379	13,677	10,502
2013q3	1,933	193	17,816	1,778	1,298	83	17,825	1,387	634	110	15,859	2,634
2013q4	1,884	1,246	17,370	11,485	1,263	875	17,315	11,926	622	371	15,496	9,080
2014q1	1,938	2,011	17,706	18,368	1,292	1,436	17,561	19,911	646	575	15,763	14,540
2014q2	1,984	1,817	18,126	16,601	1,327	1,273	18,040	17,744	658	545	16,107	14,663
2014q3	1,979	643	18,080	5,873	1,321	365	17,977	4,752	658	278	16,145	5,903
2014q4	1,905	1,506	17,399	13,759	1,265	1,073	17,193	14,493	640	433	15,645	10,934
2015q1	1,871	2,221	16,948	20,111	1,233	1,543	16,640	21,057	638	678	15,443	17,061
2015q2	1,988	1,526	18,003	13,825	1,318	1,132	17,788	15,828	669	394	16,103	8,755
2015q3	2,053	561	18,589	5,076	1,365	242	18,420	3,395	688	319	16,618	8,237
2015q4	2,023	1,609	18,318	14,572	1,332	1,030	17,938	13,611	691	579	16,579	12,961
2016q1	2,014	2,169	18,125	19,515	1,334	1,598	17,798	21,917	681	570	16,317	15,289
2016q2	2,059	1,675	18,530	15,072	1,368	1,072	18,246	15,220	691	603	16,595	13,752
2016q3	2,043	338	18,384	3,044	1,354	168	18,082	2,209	689	170	16,547	2,848
2016q4	2,215	1,785	19,933	16,061	1,454	1,189	19,358	15,829	761	595	18,132	14,297

The key figure of the paper is Figure 5, where we compare our average yearly user cost measure of housing services with the BEA's yearly estimate of housing services from PCE. Note that we compare the full estimates of aggregate housing services because we are estimating user cost for all residential homes in our sample, applying the same method to all homes whether they are owner-occupied or not.[27] Our aggregate measure of housing was initially much higher than the BEA's estimate in 2001, but this gap widened precisely when home prices throughout much of the U.S. appreciated considerably during the run up to the financial crisis and Great Recession.

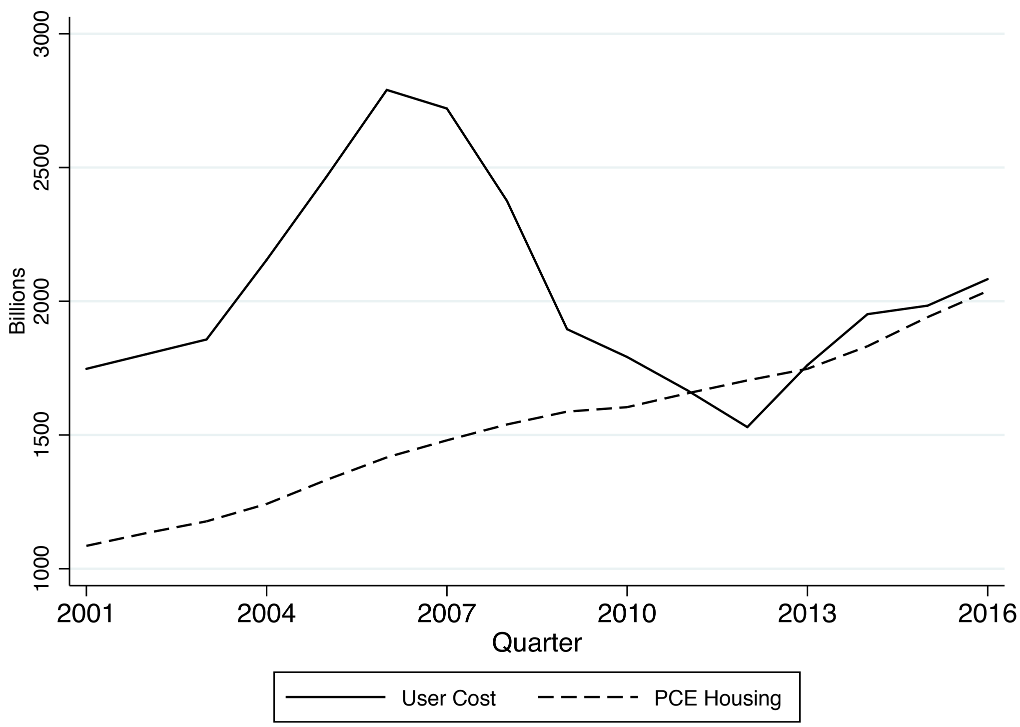

Figure 5: Total User Cost (Default Specification) Compared to PCE Housing Estimates

[27] Also note that aside from methodology, there are other small differences that remain. For example, we do not include the imputed rent for farm dwellings, as we cull properties zoned for agriculture and we do not have separate estimates for group homes, nor do we differentiate between vacant and occupied-dwellings. But, these estimates are small and relatively constant over time, so they would not account for much of the differences in price dynamics over time in Figure 5. With linked administrative data, future work could make vacancy rate adjustments to our user cost estimates.

The more pronounced path of the user cost-based estimate from 2001 through 2010, during the infamous bubble-bust years, bears a striking resemblance to national house price indices like Case-Shiller's or FHFA's, rising approximately $1 trillion from 2001 to the peak in 2007 (62%), with a similarly precipitous fall in the several years that followed. Broadly, this result is consistent with other recent work like Braga and Lerman (2019) who assess the divergence in consumer price index (CPI) measures using a user cost vs. rental-equivalence approach, also finding a stark contrast between the two approaches over these years. However, beginning around 2010, the user cost-based estimate of housing services using Zillow data has tracked much more closely to the housing estimate based on the BEA's current rental-equivalence method.

Our alternative specification of the user cost method, factoring in very recent, very local price expectations, depicts a more pronounced bubble and bust in its measurement of housing services of the same time period. Figure 6 shows price expectations producing a much sharper peak and trough with the alternative specification, with the level in recent years being considerably smaller than current BEA estimates of housing. But, given that this specification is much more aggressive in its price expectations assumptions, this result should be seen as one of the more volatile series this data can produce with this approach, and therefore interpreted with more than a grain of salt, so-to-speak. Indeed, this is one reason why most countries that actually employ the user cost method for housing in their national accounts or price indices often simplify this method further by omitting the price appreciation term in the user cost calculation (Diewert and Nakamura 2009).

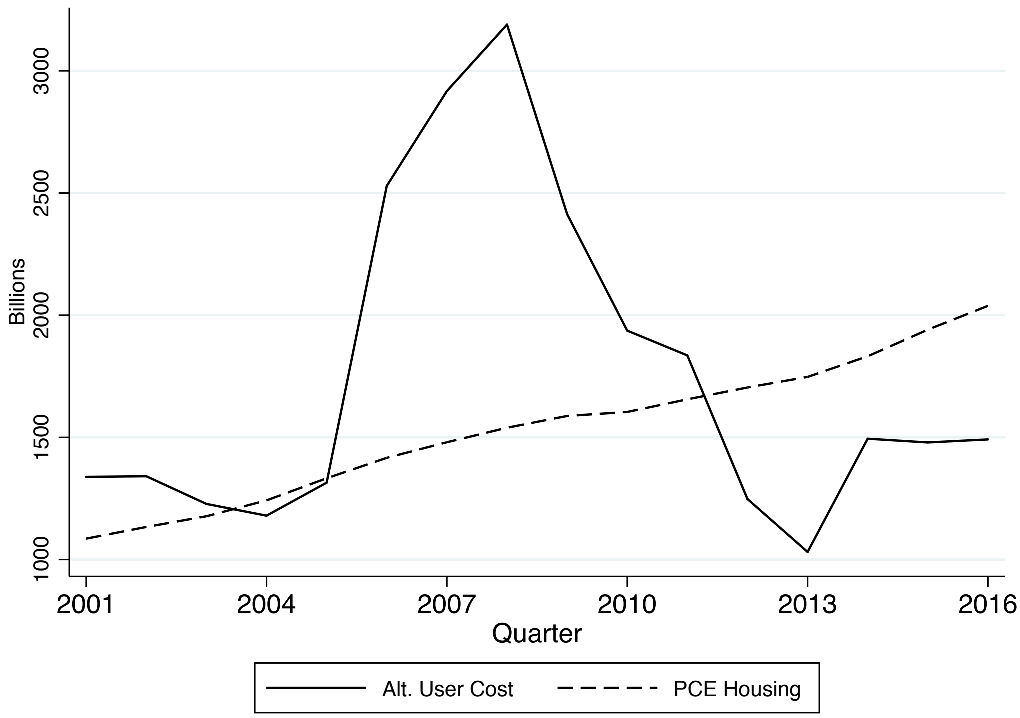

Figure 6: Total Alternative User Cost Compared to PCE Housing

An important benefit to calculating user cost estimates with microdata is that there is greater scope for separating estimates geographically or by housing type. More generally, national statistical offices face increasing demands by users for finer partitions of the national accounts, which is a key advantage of "big data" over traditional designed survey data that suffers to a greater extent from a thin cell problem. As an example, Figures 7 and 8 show average user cost by region (Census Division) for single family residences (SFR) and non-SFR's respectively, although the data easily allows us to provide measures at the county or zip code level (except, of course, for states with missing price data). As a reasonableness check, the estimates produce the expected results that the Pacific region has the highest average user costs of housing, followed by New England, with several regions at the bottom experiencing mild, if any, bubble-bust market dynamics. This is consistent with numerous other regional metrics of the housing market over this same period.

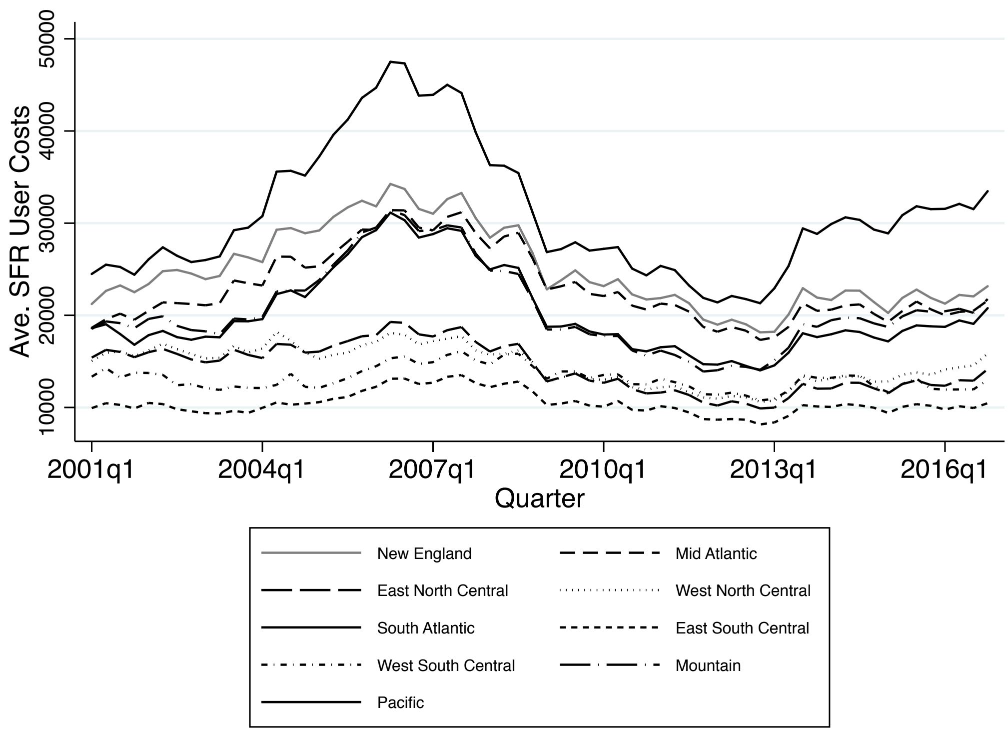

Figure 7: Average User Costs for SFR by Census Division

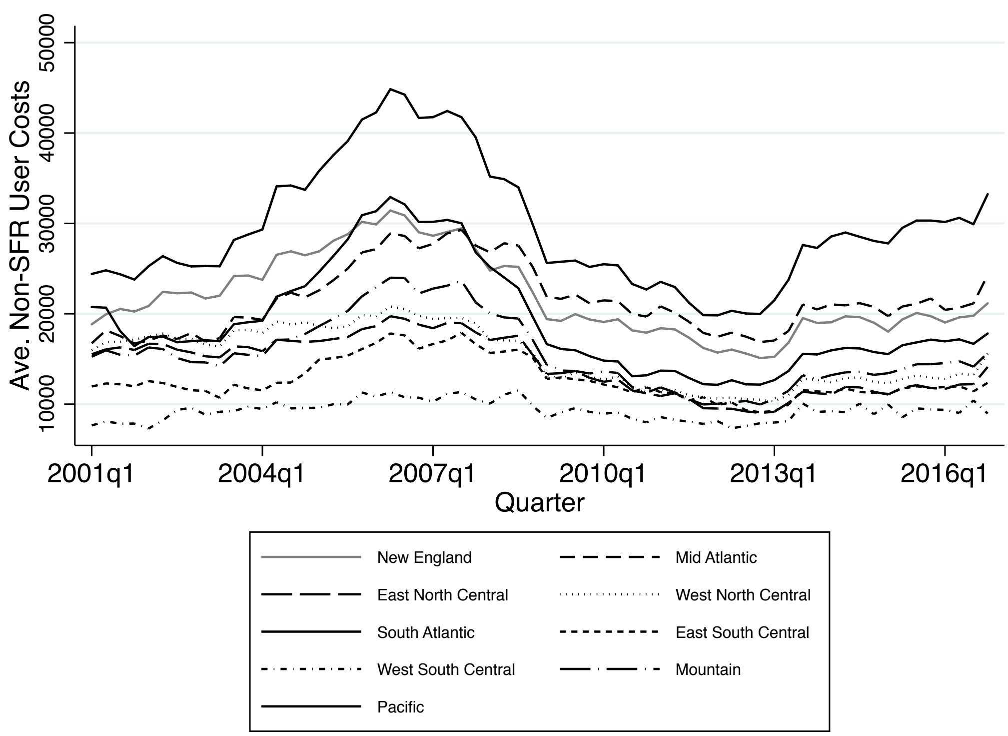

Figure 8: Average User Costs for Non-SFR by Census Division

Finally, while large aggregate estimates are often the focus of NIPA estimates, many users prefer per unit averages. Figure 9 depicts average user cost per residential unit and the corresponding BEA per unit space rent estimate. While the shape is nearly identical to Figure 5, the magnitudes may be helpful for assessing reasonability of the estimates.

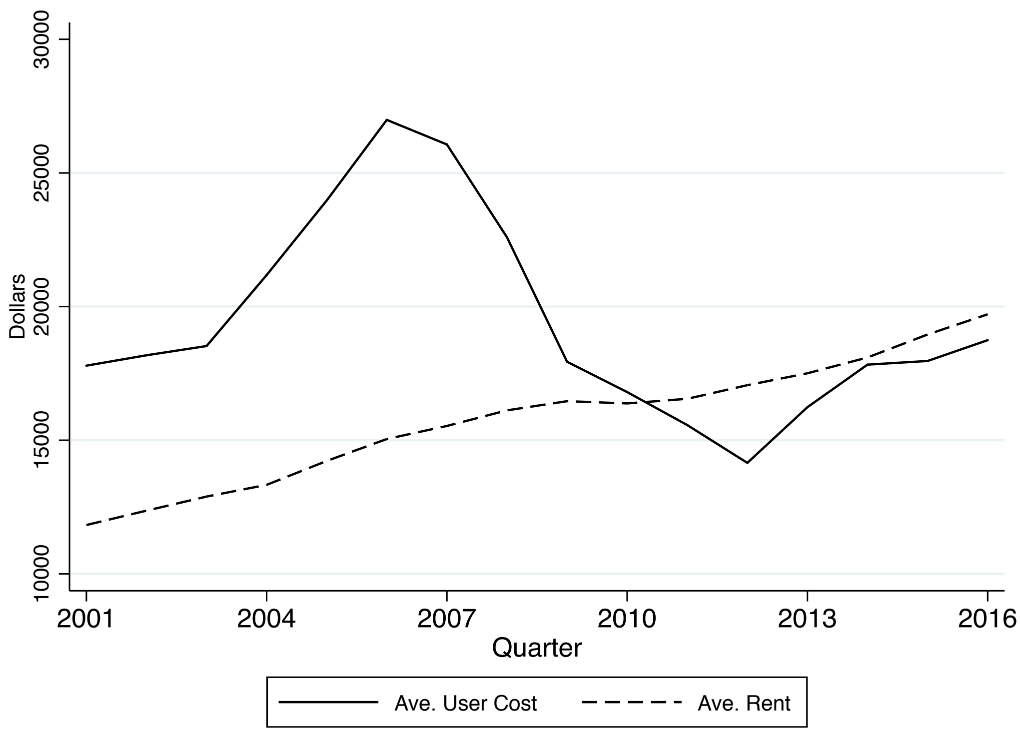

Figure 9: Average User Costs and PCE Average Rent

6. Discussion

We find that a user cost method using fine-microdata from Zillow can produce estimates of housing services comparable to the BEA's current method, at least for the most recent years we estimate. However, the departure from the rental-equivalence method during the first decade of this century (and, extended periods prior to that, based on other studies using different data) shows that convergence of these estimates is far from guaranteed. And, if there are systematic divergences, particularly when the housing sector is experiencing a pronounced boom-bust cycle, a central question for national statistical offices will be: to what extent should housing estimates reflect underlying asset appreciation (that does not appear in rental data), which may or may not be temporary? And, which conception of aggregate housing is more relevant to users of the data and policymakers?[28]

While these and other fundamental questions remain, there is a great deal of potential upside to incorporating new data and exploring new methods into the national accounts, which is a key motivator of this study. Statistical agencies are continuously seeking ways to lower response burden for survey respondents, which is of increasing concern in an era of falling response rates more generally, and to find more cost-effective means for delivering statistics to users. If "big data" sources can substantially improve precision for regional and type stratification, for example, or even supplement parts of the current method where data may be thin, then a wholesale replacement of the current method may be a false dichotomy, as a hybrid or supplemental approach could be a valid consideration as well. We leave this, however, to future research.

[28] There is evidence that the economic decisions of homeowners are, in fact, influenced by price appreciation/depreciation of their homes and housing wealth. See, for example, Mian and Sufi (2011), Mian, Rao, and Sufi (2013), Campbell and Cocco (2007), and Lowenstein (2018).

References

Aten, Bettina. 2018. Valuing Owner-Occupied Housing: an empirical exercise using the American Community Survey (ACS) Housing files, BEA Working Paper, March 2018.

Bian, X., Waller, B.D. and Wentland, S.A., 2016. The Role of Transaction Costs in Impeding Market Exchange in Real Estate. *Journal of Housing Research*, 25(2), pp.115-135.

Braga, B. and Lerman, R.I., 2019. Accounting for homeownership in estimating real income growth. *Economics Letters*, 174, pp.9-12.

Bureau of Economic Analysis. 2018. Table 2.5.5 Personal Consumption Expenditures by Function. Accessed from https://apps.bea.gov/iTable/iTable.cfm?reqid=19&step=2#reqid=19&step=2&isuri=1&1921=survey on December 17, 2018.

Campbell, J. Y. and Cocco, J. F. (2007). How do house prices affect consumption? evidence from micro data. *Journal of Monetary Economics*, 54 (3), 591–621.

Davis, M.A., Lehnert, A. and Martin, R.F., 2008. The Rent-price ratio for the aggregate stock of owner-occupied housing. *Review of Income and Wealth*, 54(2), pp.279-284.

Diewert, W. E. 2008a, The Treatment of Owner Occupied Housing and Other Durables in a Consumer Price Index, in W. E. Diewert, J. Greenlees and C. Hulten (eds.), Price Index Concepts and Measurement, NBER Studies in Income and Wealth, University of Chicago Press, 2008.

Diewert, W. E. 2008b, "Conclusions and Future Directions", The Paris OECD-IMF Workshop on Real Estate Price Indexes, Paris, November 6-7m 2006.

Diewert, W.E., Nakamura, A.O., 2009. Accounting for housing in a CPI. In: Diewert, W.E., Balk, B.M., Fixler, D., Fox, K.J., Nakamura, A.O., 2009. Price And Productivity Measurement, vol. 1 – Housing, Trafford Press, pp. 7–32 (Chapter 2). Available from: www.vancouvervolumes.com and <www.indexmeasures.com>

Diewert, W.E., Nakamura, A.O. and Nakamura, L.I., 2009. The housing bubble and a new approach to accounting for housing in a CPI. *Journal of Housing Economics*, 18(3), pp.156-171.

Garner, T.I. and Verbrugge, R., 2009. Reconciling user costs and rental equivalence: Evidence from the US consumer expenditure survey. *Journal of Housing Economics*, 18(3), pp.172-192.

Gill, H.L. and Haurin, D.R., 1991. User cost and the demand for housing attributes. *Real Estate Economics*, 19(3), pp.383-396.

Gillingham, Robert. 1983. Measuring the Cost of Shelter for Homeowners: Theoretical and Empirical Considerations, *Review of Economics and Statistics*, 65, 254–65.

Glaeser, Edward L. and Joseph Gyourko, 2009. "Arbitrage in Housing Markets," in E. L. Glaeser and John M. Quigley editors, Housing markets and the economy: risk, regulation, and policy: essays in honor of Karl E. Case, Lincoln Institute of Land Policy, Cambridge, MA.

Goeyvaerts, G. and Buyst, E., 2019. Do market rents reflect user costs? *Journal of Housing Economics*. Forthcoming.

Haffner, M. and Heylen, K., 2011. User costs and housing expenses. Towards a more comprehensive approach to affordability. *Housing Studies*, 26(04), pp.593-614.

Hill, R.J. and Syed, I.A., 2016. Hedonic price–rent ratios, user cost, and departures from equilibrium in the housing market. *Regional Science and Urban Economics*, 56, pp.60-72.

Jorgenson, D.W., 1963. Capital theory and investment behavior. *American Economic Review* 53 (2), 247–259.

Jorgenson, D., 1967. The theory of investment behavior. In *Determinants of Investment Behavior* (pp. 129-175). NBER.

Katz, Arnold J. (2004), "Estimating Dwelling Services in the Candidate Countries: Theoretical and Practical Considerations in Developing Methodologies Based on the User Cost of Capital Measure," Bureau of Economic Analysis #BE-54.

Katz, A.J., 2009. Estimating dwelling services in the candidate countries: theoretical and practical considerations in developing methodologies based on a user cost of capital measure. In: Diewert, W.E., Balk, B.M., Fixler, D., Fox, K.J., Nakamura, A.O., 2009. Price and Productivity Measurement, vol. 1 – Housing, Trafford Press, pp. 33–50 (Chapter 3). Available from: <www.vancouvervolumes.com> and <www.indexmeasures.com>.

Katz, Arnold J. (2017), "Imputing Rents to Owner-Occupied Housing by Directly Modelling Their Distribution", (WP2017-7), BEA Working Paper, August 2017.

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

Lowenstein, Lara. 2018. Consumption of Housing During the 2000s Boom: Evidence and Theory, Working Paper.

Mayerhauser, Nicole and Denise McBride. 2007. "Treatment of Housing in the National Income and Product Accounts," BEA Staff Study presented before the Society of Government Economists at the Annual Convention of the Allied Social Science Associations, December.

McFadyen, S. and Hobart, R., 1978. An alternative measurement of housing costs and the consumer price index. *The Canadian Journal of Economics*, *11*(1), pp.105-112.

Mian, Atif, and Amir Sufi. 2011. House Prices, Home Equity-Based Borrowing, and the US Household Leverage Crisis. *American Economic Review*, 101 (5): 2132-56.

Mian, A., Rao, K. and Sufi, A. 2013. Household balance sheets, consumption, and the economic slump. *Quarterly Journal of Economics*, pp. 1–40.

Moulton, J.G., Waller, B.D. and Wentland, S.A., 2018. Who Benefits from Targeted Property Tax Relief? Evidence from Virginia Elections. *Journal of Policy Analysis and Management*, 37(2), pp.240-264.

United Nations, Commission of the European Communities, International Monetary Fund, Organisation for Economic Co-operation and Development, and World Bank, 2008 System of National Accounts, United Nations, NY, 2008. https://unstats.un.org/unsd/nationalaccount/sna2008.asp

Verbrugge, R., 2008. The puzzling divergence of rents and user costs, 1980–2004. *Review of Income and Wealth*, 54(4), pp.671-699.