The impact of federal housing policy on housing demand and homeownership: Evidence from a quasi-experiment

Morris A. Davis
Rutgers University
Stephen D. Oliner
American Enterprise Institute, UCLA
Tobias J. Peter
American Enterprise Institute
Edward J. Pinto
American Enterprise Institute

AEI Economics Working Paper 2018-01
Updated November 2019

© 2019 by Morris A. Davis, Stephen D. Oliner, Tobias J. Peter, and Edward J. Pinto. All rights reserved.

The American Enterprise Institute (AEI) is a nonpartisan, nonprofit, 501(c)(3) educational organization and does not take institutional positions on any issues. The views expressed here are those of the author(s).

The Impact of Federal Housing Policy on Housing Demand and Homeownership: Evidence from a Quasi-Experiment

Morris A. Davis
Rutgers University

Stephen D. Oliner
American Enterprise Institute and UCLA

Tobias J. Peter
American Enterprise Institute

Edward J. Pinto
American Enterprise Institute

November 2019
Previous version: June 2019

Abstract

Federal housing policy promotes homeownership by subsidizing mortgage debt for many households with few assets and low credit scores. In this paper, we exploit the Federal Housing Administration's (FHA's) surprise 50 basis point cut to its annual mortgage insurance premium in January 2015 to study the impact of federal housing policy and interest rates on housing demand for a population of households likely to be influenced by changes to policy. The premium cut, which reduced monthly payments the same amount as a three-quarter percentage point drop in the mortgage rate, increased the purchasing power of the typical FHA borrower by 6 percent. Our analysis shows FHA borrowers increased the value of the housing they purchased by 2.5 percentage points relative to a control group of borrowers in areas with minimal FHA presence. The rise in spending reflected an increase in constant-quality home prices, with no significant change in the quality of housing purchased by FHA buyers. We also estimate that the premium cut induced approximately 17,000 households to become first-time homebuyers in the initial year after the cut, an increase that fell far short of the FHA's projection.

We thank ATTOM Data Solutions for providing the data used in this paper. We received helpful comments from Darryl Getter, Chris Herbert, Theresa Kuchler, Kevin Park, Thomas Philippon, Daniel Ringo, Johannes Stroebel, Stijn Van Nieuwerburgh and other participants at conferences and seminars. The views expressed herein are ours alone and do not represent those of any other individuals or the institutions with which we are affiliated. Corresponding author: Morris A. Davis, mdavis@business.rutgers.edu, 608-358-0717.

1. Introduction

A primary goal of federal housing policy is to encourage homeownership and investment in housing by subsidizing homeowners' borrowing costs and by making mortgages available to households with limited resources for a down payment and low credit scores. Broadly speaking, economists have few reliable, data-driven estimates of the impact of these policies on homeownership, house prices, or housing affordability. Understanding how federal policy affects the homeownership rate and the value of housing purchased is difficult because large changes to federal policy are infrequent, tend to be universally applied, and are often correlated with significant macroeconomic shocks.

In this paper, we use a recent, quasi-natural experiment to investigate how exogenous changes in federal policy affect housing demand for borrowers that tend to be on the edge of homeownership. On January 7, 2015, the Federal Housing Administration (FHA) reduced the annual mortgage premium it charges to guarantee new loans processed on or after January 26th by 50 basis points, from 1.35 percent of the loan balance to 0.85 percent. The effect of this cut was to increase the purchasing power of the typical FHA borrower by about 6 percent.[1] This announcement was a surprise to the marketplace (Bhutta and Ringo 2017)[2], and it was not accompanied by pricing changes at the other major federal agencies that guarantee mortgages – the government-sponsored enterprises Fannie Mae and Freddie Mac (the GSEs) and the Department of Veteran's Affairs (VA). Because the premium cut was unanticipated, it solves the usual problem of finding exogenous variation in borrowing costs with which to estimate the effects on housing demand.

We use a differences-in-differences (diff-in-diff) approach to analyze the impact of the FHA premium cut. Even though the cut was a surprise that affected a subset of market participants, establishing a compelling diff-in-diff analysis poses a challenge for three reasons. The first challenge is that the premium cut applied to all FHA borrowers. For identification, we compare the change in behavior for FHA borrowers to another group of homebuyers. Our analysis uses GSE borrowers as the control group. A potential concern is that the borrowers who select GSE loans could be different than those who choose FHA loans in *unobservable* ways that cannot be controlled for in a diff-in-diff

[1] This estimate holds the homeowner's monthly payments fixed after the premium cut, where the monthly payments include repayment of loan principal and interest, homeowner's insurance (assumed to be 0.35 percent of property value), property taxes (assumed to be 1.2 percent of property value), and the FHA premium. The assumed loan is a 30-year fixed rate mortgage with an interest rate of 4 percent, the average interest rate on FHA mortgages in 2015.

[2] Bhutta and Ringo marshal several pieces of evidence to support the surprise nature of the premium cut. In particular, they searched through mortgage and housing industry publications and found no articles or blog posts that anticipated the cut. They also found no evidence from Google searches that households had expected it.

regression. Persistent unobservable differences are not a problem because they are removed by the diff-in-diff. The potential threat to identification comes from changes in unobservables for FHA borrowers over the estimation window that differ from any such changes for GSE borrowers. To reduce this risk, our estimation strategy uses a relatively short window so that any changes to unobservables are likely to be small. In addition, we show that the sorting of borrowers into FHA and GSE loans can be predicted very accurately from pricing differences between these loans that depend solely on the borrower's credit score and loan-to-value ratio. This suggests that the observable differences between FHA and GSE borrowers – which we control for in the diff-in-diff – are far more important than unobservable differences.

The second challenge arises from the fluid boundary between FHA and GSE borrowers. We show that the premium cut induced a substantial number of GSE borrowers to switch to FHA loans. As a result, the observable characteristics of both the FHA treatment group and the GSE control group could have changed after the premium cut. We account for this potential change in mix by including controls for a broad set of borrower characteristics in our baseline regressions and our robustness checks. The controls in our baseline regressions include the borrower's county of residence, income, credit score, gender, race, ethnicity, and whether the borrower sought pre-approval for the loan; robustness tests add the borrower's loan-to-value ratio and debt-to-income ratio (which is not available for some loans) to the baseline set of controls.

The final challenge is that the FHA policy change indirectly affected other homebuyers in the same geographic market. Indeed, we document that the premium cut boosted constant-quality house prices in census tracts with a relatively high share of FHA mortgages and this price increase spilled over to non-FHA borrowers in the tract, which reduced their purchasing power. To minimize the influence of these spillover effects, our control group consists of borrowers receiving a GSE mortgage in census tracts where FHA mortgages are infrequent. This control group is locationally separate from our treatment group, which consists of FHA borrowers living in census tracts where FHA mortgages account for a relatively high proportion of all home purchase mortgages.

To conduct the analysis, we merge data from a number of sources. Our combined dataset contains information on a large share of the FHA and GSE mortgages originated in 23 counties across the country during 2013-2015. We show that in the initial year after the premium cut, the prices of FHA-financed homes accelerated about 2.5 percentage points vis-à-vis GSE-financed homes, implying a semi-elasticity of housing demand to the mortgage rate (the intensive margin of demand) of roughly 3.4

among FHA homebuyers.[3] This result provides a calibration target for macroeconomists looking to anchor a model-predicted interest-rate elasticity of housing demand. We also show that rising constant-quality prices due to stronger demand fully accounted for the 2.5 percent increase; we find no evidence that FHA buyers opted for homes with more amenities or a better location.

In addition to this analysis of market impacts, we evaluate the effectiveness of the policy change in boosting homeownership. When the premium cut was announced, the FHA anticipated that it would spur 250,000 new homebuyers to purchase their first home over the next three years – an average of roughly 83,000 per year over 2015-2017.[4] Although the number of FHA first-time buyers increased about 176,000 in the first year after the premium cut, we estimate that the premium cut itself was responsible for only about 17,000 of these buyers. The 95 percent confidence band around this point estimate is relatively wide, but even the top end of the band is less than half of the FHA's projection. The rest of the 176,000 increase in FHA first-time buyers in 2015 consisted of borrowers who likely would have received mortgages backed by other government agencies or borrowers who would have chosen to buy homes because of trend growth in the economy, unrelated to the premium cut. For first-time buyers using FHA mortgages, our point estimate implies a one-year semi-elasticity of homeownership to interest rates of 4.2.

The premium cut created losers as well as winners and implied a redistribution of wealth. In addition to the direct benefit for FHA borrowers, existing homeowners gained from the induced increase in constant-quality home prices in areas with a sizable FHA presence. The benefits for these two groups came at the expense of home buyers who did not use FHA financing, as the rise in constant-quality home prices eroded their purchasing power. As discussed in Section 4, we estimate that in 2015 there were about 500,000 non-FHA first-time buyers who lived in Census tracts with a significant presence of FHA borrowers. We also estimate that each of these non-FHA first-time buyers paid an average of roughly $6,200 extra for their home as a result of the premium cut. Thus, non-FHA first-time buyers as a group paid about $3.1 billion extra due to the premium cut (500,000 such buyers x $6,200 per home).

As noted above, our findings should be of interest to macroeconomists studying the effect of federal policy on housing demand, house prices, and homeownership. Recent papers have analyzed the effect of tax policy (Sommer and Sullivan 2018), government mortgage guarantees (Jeske et. al. 2013 and Gete and Zecchetto 2017), financing constraints and interest rates (Favilukus et. al. 2017), limits on

[3] The 50 basis point reduction in the FHA premium has the same effect on monthly payments as a 73 basis point reduction in the mortgage interest rate. This imputed 73 basis point cut in the mortgage rate induced a 2.5 percent increase in the average value of homes purchased, for a semi-elasticity of 3.4 (2.5 divided by 0.73).
[4] https://archives.hud.gov/news/2015/pr15-001.cfm

debt-to-income ratios (Greenwald 2018), and alternative mortgage products (Chambers et al. 2009 and Corbae and Quintin 2015). The authors in this literature take care to match many features of the micro and macro data, as emphasized by Davis and van Nieuwerburgh (2015) and Piazzesi and Schneider (2016). Some of these macro models explicitly include an equation that links the mortgage rate to an endogenously determined risk-free rate and a gap between these two rates that is set outside of the model; others have a wedge between the risk-free rate and the mortgage interest rate that is endogenously determined by frictions or government institutions inside the model. Thus, all of these models have an embedded elasticity of housing prices and quantities with respect to the difference of the mortgage interest rate and the risk-free rate.[5] To our knowledge, however, authors in this literature have not checked if their models replicate the interest-rate elasticities of house prices and quantities because these elasticities *are not widely known for U.S. data*. Our paper provides macroeconomists with calibration targets for these elasticities with respect to exogenous shifts in the cost of mortgage borrowing.

There are only a few other papers besides ours that estimate the impact of interest rates on housing quantities and prices at the household level. Two key predecessors to our paper are Adelino et al. (2014) and DeFusco and Paciorek (2017). Adelino et al. (2014) use changes over time in the maximum size for GSE loans (the conforming loan limit) to estimate the semi-elasticity of housing demand among existing buyers (the intensive margin). Their estimates span a wide range – from 1.2 to 9.1 – with the estimates depending on the assumed rate spread between conforming and jumbo loans and the estimated change in house prices. DeFusco and Paciorek (2017) study bunching around the conforming loan limit to estimate the interest elasticity of mortgage debt at the household level. They find a one percentage point increase in the mortgage interest rate reduces the size of the first-lien mortgage taken out (not house prices or quantities) by 2 to 3 percent. This range is a bit below – but not significantly different than – our estimate along the intensive margin of roughly 3.4.[6]

[5] Examples of exogenous changes in this gap include an unexpected reduction in the tax deductibility of mortgage interest as in Sommer and Sullivan (2018) or a reduction in the subsidy to mortgage rates in government-guaranteed loans as in Jeske et. al. (2013) and Gete and Zecchetto (2017).

[6] When we re-run our analysis using mortgage debt instead of house values as the dependent variable, our results do not change. Evidence in Bhutta and Ringo (2017) suggests FHA borrowers are more likely to be constrained by payment-to-income constraints than the borrowers around the conforming loan limit that are the focus of the DeFusco and Paciorek study. This may explain why our estimates are higher than those of DeFusco and Paciorek, although the standard errors are large enough that we cannot reject that our estimates are the same as theirs. DeFusco and Paciorek's results are similar to those of Fuster and Zafar (2015), who use survey data to estimate that a 2 percentage point change in mortgage rates changes willingness to pay for a home by 5 percent.

Our analysis complements that of Bhutta and Ringo (2017), who also study the impact of the FHA premium cut. Bhutta and Ringo focus mainly on the increase in the number of home purchase loans induced by the premium cut (the extensive margin). Their results fall within the relatively wide confidence band for our estimate. Bhutta and Ringo also include a brief analysis of the impact of the premium cut on house prices and find no evidence of such price effects, contrary to our results. We believe our approach, which uses a large property-level dataset to create well-defined treatment and control groups, has more power to uncover price effects than the Bhutta-Ringo analysis, which relies for the most part on estimated median home prices at the zip-code level.

Finally, we would note that our results reflect the combined influence of household responses and market conditions. The FHA premium cut occurred during the 29th month of a national seller's market based on data published by the National Association of Realtors (NAR).[7] With tight inventories of homes for sale, we show that the boost to demand from the premium cut resulted in a substantial rise in home prices and an erosion of purchasing power for non-FHA buyers. Had the same policy change been implemented in a buyer's market with more plentiful supply, our estimated impact on prices might have been smaller and the number of new homebuyers greater.[8]

2. Data

2.1. Property-level data

Our primary dataset, provided by ATTOM Data Solutions (ATTOM), contains detailed property-level data based on public records drawn from tax assessments, sales transactions, and recorded mortgage loans. The mortgage information includes the loan purpose (allowing us to separate home purchase loans from refinance loans), the loan amount, the exact date of the transaction, and the loan type (FHA, VA, a construction loan, or a large residual category that includes conventional loans and loans guaranteed by the Rural Housing Service (RHS)).[9] Finally, the data also include an estimated home value for each property, based on ATTOM's Automated Valuation Model (AVM). The ATTOM data cover 23 counties across 12 states over the period 2013-2015. Although we selected these counties because of their large FHA loan counts, they also have some of the nation's largest conventional loan totals by

[7] The NAR defines a seller's market to exist when the inventory of existing homes for sale would be exhausted in six months or less at the current sales pace: See http://www.realtor.org/news-releases/2013/04/march-existing-home-sales-slip-due-to-limited-inventory-prices-maintain-uptrend.

[8] During the buyer's market that dominated the 1930s, Ernest Fisher, the FHA's first chief economist, noted that the "longer [loan] terms, lower interest rates, and payment plans [FHA] offered… were a major factor in stimulating the purchase of homes and in reviving the moribund home construction industry" (Fisher 1951, pp, 78-9). In a seller's market, however, Fisher concluded that liberalization of credit primarily drives up prices.

[9] Conventional loans consist of those with a GSE guarantee and those held in the private sector with no guarantee.

county. The 23 counties contain about 17 percent of all FHA loans and about 13 percent of all conventional loans.

The first four columns of Table 1 present basic information about the 23 counties for 2015. These counties have a total population of about 49 million, accounting for 15 percent of the entire U.S. population. Median household income in these counties spans a wide range, with a central tendency that is not noticeably different than the national median of about $56,000 in 2015. Similarly, the percentage of adults with a bachelor's degree in these counties is centered at about the national average. In terms of minority population share, 19 out of the 23 counties have a greater black or Hispanic share than the national average, consistent with the FHA's traditional importance for such borrowers.

The final two columns of Table 1 compare ATTOM's loan counts over the 2013-2015 sample period to the counts in the data collected under the Home Mortgage Disclosure Act (HMDA) over the same period. The loans reported under HMDA represent a near census of the mortgage market.[10] Because HMDA does not identify one-unit properties, the comparison covers primary owner-occupied properties with one to four units.[11] For the 23 counties taken together, first-lien, primary owner-occupied home purchase loans in the ATTOM dataset represent 97 percent of the HMDA count for FHA-guaranteed loans and 99 percent for conventional and RHS loans.[12] Every county has a coverage ratio for FHA loans of at least 90 percent, and the same holds for conventional loans in all but two counties: Cook County, IL (89 percent) and Franklin County, OH (73 percent). The table shows that the coverage ratio relative to HMDA exceeds 100 percent in four counties for FHA loans and eleven counties for conventional and RHS loans. The coverage ratios can exceed 100 percent for at least three reasons. First, HMDA is not a complete loan count because very small lenders are exempt from HMDA reporting. Second, for the properties that sell more than once in our ATTOM dataset, we can only determine whether the buyer is a primary owner-occupant for the most recent sale. We have no information about occupancy status for earlier sales and assume that all the earlier sale(s) were also primary owner-occupied properties, which will overstate the ATTOM loan count for the purpose of matching to HMDA.

---

[10] For more on the HMDA data, see http://www.consumerfinance.gov/data-research/hmda/learn-more.

[11] The HMDA counts we use for the comparison include home improvement loans in addition to home purchase loans for two reasons. First, when we matched ATTOM loans to HMDA based on loan type (FHA or conventional), census tract, loan amount, and sale year, some ATTOM home purchase loans matched to HMDA home improvement loans. And second, many HMDA home improvement loans have loan amounts that are characteristic of HMDA purchase loans.

[12] We include RHS loans in the HMDA count because, as noted above, they can't be separated from conventional loans in the ATTOM data.

Third, we found that the ATTOM data include a small number of land purchase loans taken out by developers, which adds slightly to the overstatement of the ATTOM loan count. In addition to these three explanations, there could be other factors that we haven't identified. Overall, however, the ATTOM dataset conforms well to the loan counts in HMDA.

Table 1. Characteristics of the Counties and Dataset Coverage

	Population and Demographic Characteristics in 2015				***ATTOM Coverage, percent of HMDA***	
County, State	**Population**	**Median Household Income**	**% Bachelor's Degree or Higher**	**Percent Black or Hispanic**	**FHA**	**Conventional and RHS**
Broward, FL	1,896,425	$53,926	32	56	102	111
Clark, NV	2,114,801	$51,552	23	41	94	102
Cook, IL	5,238,216	$56,851	37	49	93	89
Duval, FL	913,010	$49,554	29	38	90	91
El Paso, CO	674,471	$60,109	36	23	102	99
Franklin, OH	1,251,722	$53,882	39	27	94	73
Gwinnett, GA	895,823	$61,732	35	46	91	99
Hillsborough, FL	1,349,050	$51,725	33	43	90	95
Kern, CA	882,176	$51,342	16	58	96	112
Los Angeles, CA	10,170,292	$59,134	31	56	99	98
Macomb, MI	864,840	$54,640	24	14	94	95
Maricopa, AZ	4,167,947	$56,004	31	36	99	107
Miami-Dade, FL	2,693,117	$43,786	27	84	109	117
Orange, FL	1,288,126	$50,720	32	50	94	94
Pierce, WA	843,954	$60,167	26	17	98	103
Pima, AZ	1,010,025	$47,099	31	40	91	102
Prince George's, MD	909,535	$76,741	32	79	95	102
Prince William, VA	451,721	$99,766	40	42	95	90
Riverside, CA	2,361,026	$58,292	21	54	98	105
Sacramento, CA	1,501,335	$58,942	30	32	101	105
San Bernardino, CA	2,128,133	$53,803	19	60	99	97
San Diego, CA	3,299,521	$67,320	37	38	95	93
Wayne, MI	1,759,335	$41,557	23	45	90	104
	---------- Entire United States ----------				**----- All 23 Counties -----**	
	321,418,821	$55,775	31	30	97	99

Population and demographic characteristics. The definition of Black or Hispanic is Hispanic or Latino of any race plus Black or African American alone and not Hispanic or Latino. Median household income is in 2015 dollars. Source: U.S. Census Bureau, 2015 American Community Survey, Tables DP02 (percent with bachelor's degree or higher), DP03 (median household income), and DP05 (population and percent Black or Hispanic).
ATTOM coverage. Comparison is for first-lien, primary owner-occupied purchase loans during 2013-2015 except for Wayne, MI, in 2014 and 2015 and Macomb, MI, in 2015. For Wayne and Macomb, the comparison is for all loans regardless of occupancy status because the occupancy indicator in the ATTOM dataset appears to be unreliable. The HMDA counts for all counties include home improvement loans. Source: Authors' calculations from ATTOM and HMDA data.

Given that our analysis focuses on the effects of the premium cut on housing demand and homeownership, we restrict our sample from the ATTOM dataset to FHA and conventional/RHS loans used to purchase primary owner-occupied one-unit homes between 2013:Q1 and 2015:Q4. We exclude any second liens associated with these home purchases, and we remove distressed sales because the sale prices may not reflect market norms. In creating this target set of loans, we take account of changes in FHA's conforming loan limit during the sample period. On December 31, 2013, the existing FHA loan limit authority under the Economic Stimulus Act of 2008 expired, which reduced the limit in high-cost areas. For some counties in our dataset, the limit fell more than 25 percent. Then, on January 1, 2015, FHA raised the loan limit in five counties in the dataset (Duval County, FL, Franklin County, OH, Gwinnett County, GA, Pierce County, WA, and San Diego County, CA). To avoid any biases from these shifting limits, we exclude loans with an amount that is larger than the county-level 2014 FHA loan limits. This eliminates high-priced FHA sales in 2013 and 2015 that exceed the 2014 limit.[13] We also impose the same county-by-county limits on the sales with conventional financing so that our results reflect the part of the housing market in which FHA operates.

Having defined the target set of loans for our analysis, we screen out loans with missing data as well as outliers. We remove loans without an AVM, sale price or census tract identifier; loans with a combined loan-to-value ratio (CLTV) above 110 percent, where the CLTV incorporates any second liens taken out at the time of origination; loans for which the AVM is more than twice or less than half the sale price; and loans for which the property's AVM/price ratio was in the top and bottom 1 percent within each sale month (this trim is done separately for FHA and conventional/RHS loans). These screens remove a relatively small share of the loan population – about 6 percent of the FHA loans and 7 percent of the conventional/RHS loans. See Appendix A for details.

We then match the resulting dataset to external datasets. We start by matching to loan-level data from HMDA and the Federal Housing Finance Agency (FHFA) for 2013, 2014, and 2015 to add information on borrower gross annual income and other borrower characteristics and to screen out any RHS loans included in the ATTOM data.[14] We initially match from the ATTOM dataset to HMDA and FHFA using the origination year, census tract, FHA versus conventional financing, and loan amount, with

[13] Imposing the 2014 limit results in some bunching of 2014 and 2015 FHA loans at the limit. These borrowers likely would have taken out larger loans had they not been constrained. To eliminate the influence of these borrowers on our results, we drop FHA loans bunched at or within 0.75 percent of the 2014 limit.

[14] Although HMDA captures the vast majority of home mortgage loans, the FHFA dataset includes some additional loans guaranteed by the GSEs. For more on the FHFA data, see http://www.fhfa.gov/DataTools/Downloads/Pages/Single-Family-Census-Tract-File.aspx

supplemental matching on lender information. We use only one-to-one matches. The match is successful for 83 percent of FHA loans and 82 percent of conventional loans in the ATTOM dataset, with similar match rates for each county.

Next, we match the FHA loans in the resulting ATTOM/HMDA/FHFA-matched dataset to FHA's Single-Family Portfolio Snapshot.[15] This step allows us to add an FHA borrower's note rate to the dataset. The match is performed using loan amount, zip code, and origination date, with supplemental matching on lender information. We use only one-to-one matches. The match is successful for over 96 percent of FHA loans.

The final step in the matching process adds the borrower's credit score and, for nearly all loans, debt-to-income ratio. We pick up this information by matching to loan-level data from the American Enterprise Institute's (AEI's) National Mortgage Risk Index (NMRI), discussed in detail below, and to loan-level data released by the GSEs.[16] All of the credit scores and debt-to-income ratios for FHA loans come from matches to the NMRI data. While the NMRI is a near-census of government-guaranteed loans, its geographical detail is limited to the state level. The GSE datasets provide greater geographical detail for the large majority of loans purchased by Fannie Mae or Freddie Mac, which enables additional matches. The matching to the NMRI and GSE datasets uses various combinations of loan type, loan amount, state or zip code, origination date, lender information, LTV, the number of borrowers, and the property type. In the case of the GSE datasets, we also match on the purchasing agency (Fannie or Freddie), and in the case of FHA loans, we also match on the note rate. We use only one-to-one matches. The match is successful for 83 percent of FHA loans and 86 percent of conventional loans. Because the NMRI and GSE datasets only contain loans with a government guarantee, this step eliminates private lenders from our sample. See Appendix B for details on all phases of the matching process.

The final dataset contains 171,000 FHA loans (68 percent of the cleaned ATTOM loans) and 299,000 GSE loans (70 percent of the cleaned ATTOM conventional/RHS loans) originated from 2013:Q1 to 2015:Q4. Although the match process removes roughly 30 percent of the cleaned ATTOM loans, key features of the data are little changed, indicating that the matched loans are representative of the larger

[15] The FHA dataset contains loan-level information on FHA's single-family endorsements, including loan purpose, property type, loan amount, zip code, lender information, interest rate, and other variables. See https://www.hud.gov/program_offices/housing/rmra/oe/rpts/sfsnap/sfsnap.

[16] The GSE datasets are Fannie Mae's Single-Family Loan Performance Data and Freddie Mac's Single-Family Loan-Level Dataset. For more information, see http://www.fanniemae.com/portal/funding-the-market/data/loan-performance-data.html and http://www.freddiemac.com/research/datasets/sf_loanlevel_dataset.html.

loan population. In particular, for FHA loans, the average CLTV in the final dataset is within ¼ percentage point of that in the cleaned ATTOM data, and the average sale price, average loan amount, and average AVM value are all within 1 percent of their counterparts in the cleaned ATTOM data. For conventional loans, the differences are only a shade larger (0.3 percentage point difference in the average CLTV, and differences of 1½ percent or less for the average sale price, loan amount, and AVM value).

2.2. AVMs

The ATTOM AVM in our dataset is an estimate of a property's value at a given time based on a weighted average of four independent submodels. The submodels are a tax-assessed value model, a hedonic model, an appraisal emulation model, and an inflated sale price model; the four submodels are weighted in accord with their estimated accuracy for a specific property.[17] We use ATTOM's AVM for December 2014 as our pre-event "stake in the ground" for home valuation, as that was the last full month before the premium cut was announced and implemented.[18] We assess the accuracy of the December 2014 AVMs by comparing AVM values to reported sale prices for properties that sold in that month. Due to data reporting and collection lags, sales in December 2014 are not known until a subsequent month. Hence, the December 2014 AVM value is calculated independently of the actual December 2014 sale price.[19]

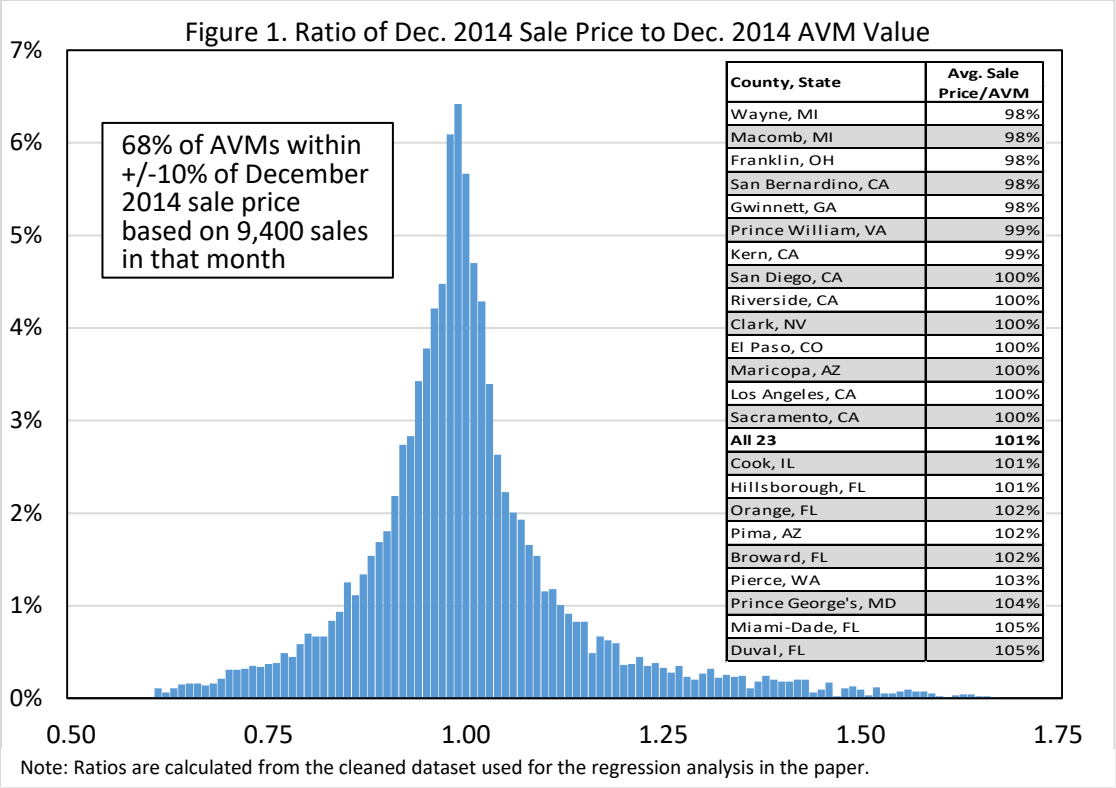
For the roughly 9,400 homes in our final cleaned dataset that sold in December 2014, the histogram in Figure 1 displays the ratio of the home's sales price to its December 2014 AVM value. On average, the sale price was equal to 101 percent of the AVM, and 68 percent of the sale prices fell within +/-10 percent of the AVM. These results also hold with limited variation for the individual counties. The average ratio of sale price to AVM value ranged from 0.98 to 1.05 across the 23 counties, and for 17 counties, at least 60 percent of the sale prices were within +/-10 percent of the AVM. Given these results, we conclude that the ATTOM AVM is a reasonably accurate estimate of property value.

---

[17] For more information, see the DataQuick CMV-Portfolio White Paper posted at http://www.aei.org/wp-content/uploads/2018/01/AVM_White_Paper.pdf. DataQuick is a subsidiary of ATTOM Data Solutions.

[18] Bhutta and Ringo (2017) demonstrate the premium cut was unanticipated, so home sales during or before December 2014 would not have built-in any effect of the cut.

[19] We confirmed directly with ATTOM that the AVM is generated in a real-time fashion and does not include any information revealed in later months.



County, State	Avg. Sale Price/AVM
Wayne, MI	98%
Macomb, MI	98%
Franklin, OH	98%
San Bernardino, CA	98%
Gwinnett, GA	98%
Prince William, VA	99%
Kern, CA	99%
San Diego, CA	100%
Riverside, CA	100%
Clark, NV	100%
El Paso, CO	100%
Maricopa, AZ	100%
Los Angeles, CA	100%
Sacramento, CA	100%
All 23	**101%**
Cook, IL	101%
Hillsborough, FL	101%
Orange, FL	102%
Pima, AZ	102%
Broward, FL	102%
Pierce, WA	103%
Prince George's, MD	104%
Miami-Dade, FL	105%
Duval, FL	105%

Note: Ratios are calculated from the cleaned dataset used for the regression analysis in the paper.

2.3. National Mortgage Risk Index

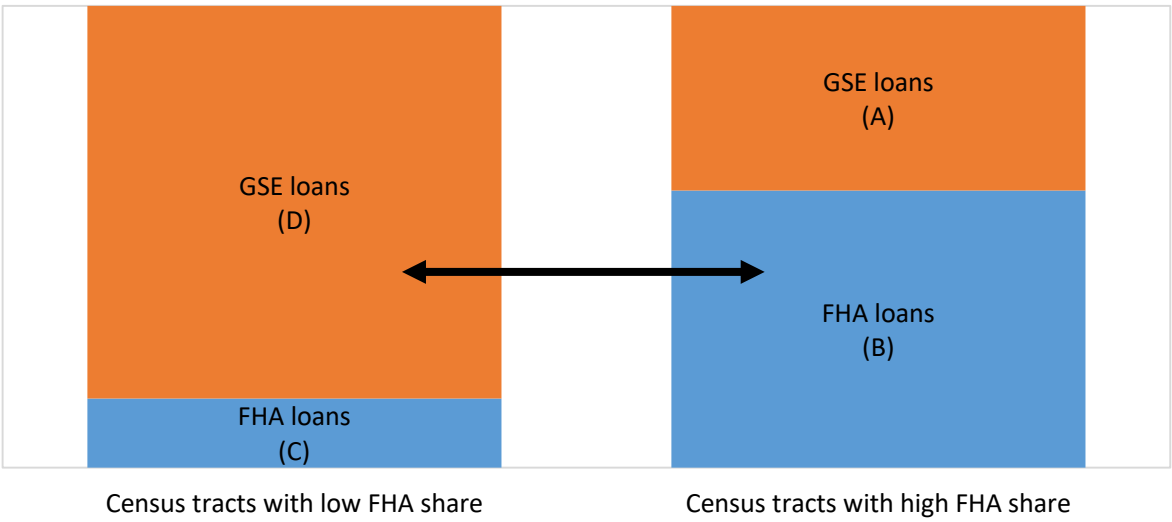
The NMRI dataset maintained at AEI is a near-census of loans that have been acquired and securitized by Fannie Mae or Freddie Mac or guaranteed by the FHA, VA, or RHS. It currently covers more than 19 million home purchase loans dating back to September 2012 with a coverage rate of 99.5 percent. Two features of the NMRI data should be noted. First, the source data for the NMRI include the month of first payment for all loans, while the origination month is not consistently reported. To estimate the origination month, we subtract two months from the first-payment month, which we have found to provide an extremely accurate estimate. Second, the NMRI dataset includes a first-time homebuyer flag provided by the agencies themselves, which we use in our analysis of the number of new homebuyers generated by the FHA premium cut.

3. Empirical design

3.1. Treatment and control groups

Figure 2 illustrates how we define the treatment and control groups for our diff-in-diff analysis. The treatment group consists of FHA borrowers in census tracts with a relatively high FHA share of loans, the blue portion of the right bar marked “B” in the figure. Our baseline analysis uses a minimum FHA share of 20 percent to define the treatment group, but we conduct robustness tests with a variety

Figure 2. Schematic for Treatment and Control Groups



of other threshold values. The control group consists of borrowers with GSE loans in census tracts with a low FHA share of mortgages, the orange portion of the left bar marked "D". These borrowers are least likely to have been affected by the FHA premium cut, since they used GSE financing to purchase a house in a tract with few FHA mortgages. We set the maximum FHA share at 5 percent in the baseline analysis, but we also present results with a maximum share of 15 percent. The diff-in-diff compares the change in behavior of the "B" and "D" groups of borrowers before and after the premium cut. Except as noted below, the control group excludes GSE borrowers in census tracts with a relatively high share of FHA loans, the orange bar marked "A". Because these borrowers are likely to have been indirectly affected by the FHA premium cut, their inclusion in the control group would muddy the experimental design. Similarly, the treatment group omits the very small number of FHA borrowers in census tracts with an FHA share of 5 percent or less, the blue bar marked "C".[20] Finally, as a reminder, all loans used in the analysis have loan amounts below the 2014 FHA limit for their county. This restriction ensures both that the results are not biased by changes in the FHA limits over time and that the GSE loans in the control group come from the same strata of the market as the FHA loans.

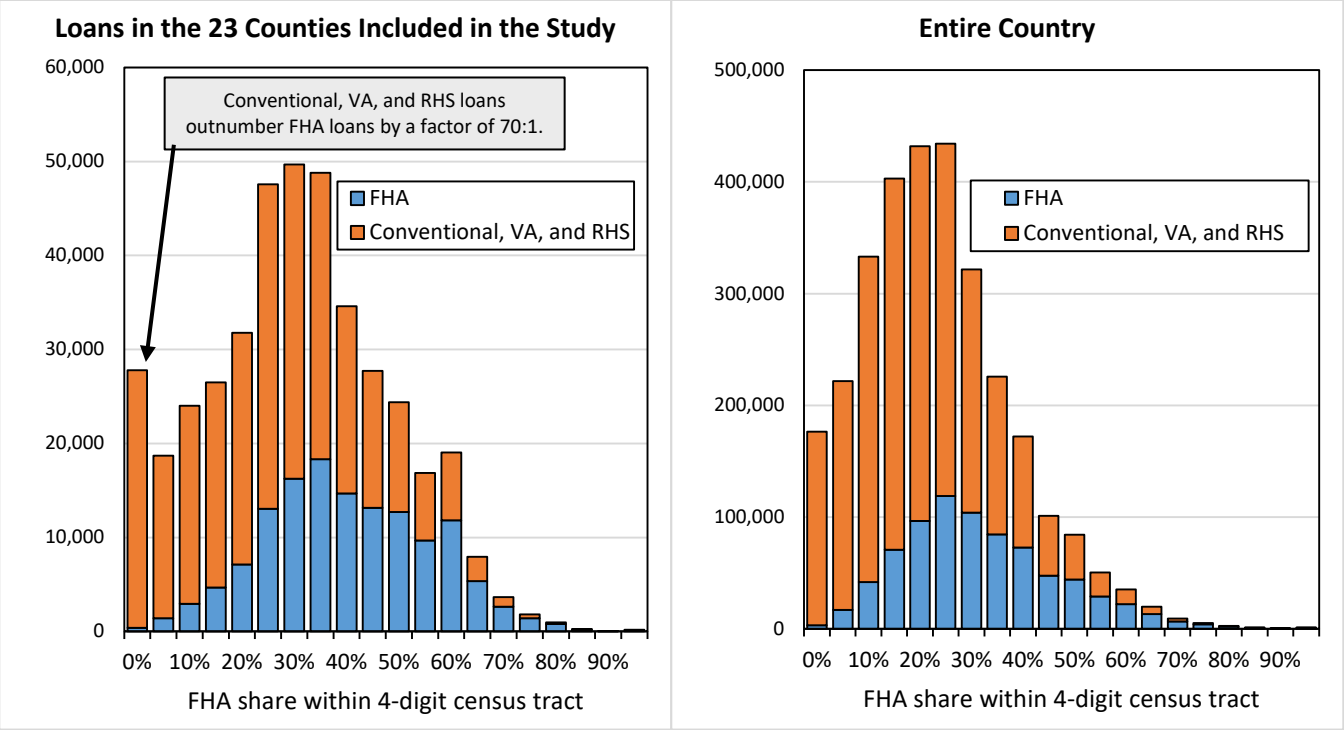
This approach to identification assumes that the premium cut had no effect on the housing market in the control tracts, where nearly all borrowers use GSE mortgages. We believe this is a reasonable assumption. Any effect in these GSE-heavy tracts is likely to have been extremely small, if

[20] The FHA loans in these tracts represent less than 1 percent of the total number of FHA loans in our final dataset. As a practical matter, including these scattered FHA loans in the treatment group would have no material effect on our results. But, conceptually, it is cleaner to exclude them in order to focus on tracts that have enough FHA loans to potentially have market effects.

not zero, because the treatment and control tracts are not close substitutes from the point of view of prospective homebuyers. In particular, it is unlikely that the premium cut caused homebuying demand in the control tracts to weaken because of a shift to treatment tracts. Had this occurred, the diff-in-diffs would not measure the absolute impact of the premium cut on house prices and quality, but rather the impact relative to a potentially affected control group. We provide evidence against this concern in Appendix G. Notably, this appendix shows that the characteristics of homebuyers in the treatment and control tracts are very different (mean income, for example, is twice as high in the control tracts) and that the two groups of tracts are geographically separated.

Figure 3 shows the distribution for 2015 of FHA purchase loans and the combination of conventional, VA, and RHS purchase loans by tract-level FHA share for the 23 counties in our data set (left panel) and the entire country (right panel).[21] The left-most orange bar in the left panel contains the census tracts in the control group in our baseline regressions. In those tracts, conventional, VA, and RHS loans outnumber FHA loans by a factor of 70 to 1, indicating that the premium cut should have little

Figure 3. Number of 2015 Purchase Loans by FHA Share



Note: Loans with loan amounts above the 2014 FHA loan limit are excluded from the bars. The FHA share for each census tract equals the number of 2015 FHA purchase loans in the tract divided by the number of 2015 total of conventional, FHA, VA, and RHS loans in the tract, using only loans with amounts below the 2014 FHA loan limit.
Source: HMDA (2015)

[21] This figure includes all 2015 HMDA purchase loans with amounts below the 2014 FHA loan limits without imposing the restrictions we put in place to generate our estimation sample.

effect on these borrowers. Figure 3 also shows the distribution for the 23 counties has more weight in tracts with relatively high FHA shares than does the distribution for the country as a whole. This reflects the fact that the 23 counties in our analysis have a relatively large volume of FHA lending. In effect, we are oversampling FHA loans to increase the size of the treatment group and improve the precision of the results.

3.2. Overall effect

We estimate the overall effect of the premium cut on house prices with the following regression:

$$ln(price) = b_0 + b_1 FHA + b_2 Q + b_3(FHA * Q) + b_4 C + b_5(FHA * C) + b_6' X + b_7'(FHA * X) + e \quad (1)$$

where *price* represents the sale price for homes financed with FHA and GSE loans, *FHA* is a dummy variable for FHA-financed purchases, *Q* is a set of quarterly dummy variables for the period 2013:Q1 to 2015:Q4 (with 2014:Q4 as the omitted quarter), *C* is a set of dummy variables for the 23 counties in the analysis (with Wayne County, MI omitted), and *X* is a vector of borrower characteristics that includes the credit score used to underwrite the loan, the natural log of the borrower's gross annual income rounded to the nearest $1,000, the borrower and co-borrower's race, ethnicity, and gender, and whether pre-approval was sought for the loan. We estimate equation 1 as an OLS regression and calculate clustered standard errors at the county level. As mentioned above, the loans used to estimate equation 1 are FHA loans in the treatment group (blue bar "B" in Figure 3) and GSE loans in the control group (orange bar "D"). We weight the loan data so that each combination of origination year, 4-digit census tract, and loan type (FHA vs. GSE) is representative of the HMDA distribution for the 23 counties, subject to a maximum upweighting by a factor of three to avoid giving heavy influence to thinly populated cells (see Appendix C for additional details).

The county dummy variables account for differences in price levels across the counties, and *X* controls for the composition of FHA and GSE borrowers.[22] The quarterly dummies represent the average price change compared to the omitted dummy for 2014:Q4. We use the coefficients on the quarterly dummies to measure the effect of the premium cut through a diff-in-diff calculation over three

[22] In equation 1, the coefficients on the borrower controls are allowed to differ for GSE and FHA borrowers (b_6 applies to GSE borrowers and b_6+b_7 to FHA borrowers). We adopt this specification because the data reject the restriction of a common set of coefficients on borrower controls across both groups. This specification controls for changes in the mix of GSE borrowers over time and separately for changes in the mix of FHA borrowers, but does not directly control for differences between the two groups of borrowers. However, as a robustness test, we pool the GSE and FHA borrowers (by constraining b_7 to be zero) and find that the results are little changed. See Table 7 below for details.

progressively longer treatment periods. The varying treatment periods allow us to measure how quickly any treatment effects become apparent.

We consider three alternative endpoints for the treatment period: 2015:Q2, 2015:Q3, and 2015:Q4. Table 2 summarizes the difference-in-difference calculations for the three alternative endpoints. As shown, we use the same set of quarters to compute both parts of the diff-in-diff calculation to prevent the results from being affected by seasonality. Because FHA's premium cut occurred during January 2015, an event window ending in 2015:Q1 does not lend itself to a clean difference-in-difference calculation.

Table 2. Difference-in-Difference Calculations Over Various Periods

End of Treatment Period	**Diff-in-Diff Calculation**
2015:Q2	(2013:Q4-2014:Q2) vs. (2014:Q4-2015:Q2)
2015:Q3	(2013:Q4-2014:Q3) vs. (2014:Q4-2015:Q3)
2015:Q4	(2013:Q4-2014:Q4) vs. (2014:Q4-2015:Q4)

3.3. Constant-quality price effect and quality effect

The overall effect of the premium cut on sale prices reflects both the change in price holding house characteristics constant (the "constant-quality price effect") and the change in the characteristics of the homes that are purchased (the "quality effect").

To capture the constant-quality price effect of FHA's premium cut, we re-estimate equation 1 with the addition of the December 2014 AVM value for FHA borrowers and separately for GSE borrowers as right-hand-side variables. The AVM serves as a summary measure of house quality, and adding the AVM turns equation 1 into a parsimonious hedonic regression. The price changes calculated from the quarterly dummies can be interpreted as hedonic price indices that measure the degree to which the additional demand induced by the premium cut boosted home prices, holding quality constant.[23]

We estimate the effect of the premium cut on the quality of homes purchased with the following regression:

$$ln(AVM) = b_0 + b_1 FHA + b_2 Q + b_3(FHA * Q) + b_4 C + b_5(FHA * C) + b_6' X + b_7'(FHA * X) + e \quad (2)$$

[23] To be clear, the estimating equation is the same as (1) but with the AVM included as an additional regressor. The dependent variable continues to be the sale price of the house.

where AVM represents the home's December 2014 AVM value which, as mentioned, will be free of any price effects due to the premium cut. Equation 2 is identical to equation 1 except that the December 2014 AVM estimate of the home's value is the dependent variable. The AVM can be thought of as a quality index that aggregates the value of various house characteristics and location. Equation 2 estimates how this index of quality evolves over time for home sales in the treatment group versus sales in the control group. Sale prices themselves play no role in this equation. We estimate equation 2 as an OLS regression with clustered standard errors at the county level and measure the treatment effect of the premium cut using the diff-in-diff calculation described above.

3.4. Comments on selection

Generally speaking, there are two groups of borrowers in the mortgage market. The first group consists of borrowers with high credit scores and sufficient funds for a sizable down payment. These borrowers typically opt for GSE or other conventional financing. The second group consists of borrowers with weaker credit profiles and less money for a down payment, for whom an FHA loan is usually the most attractive option.

One concern about selection focuses on borrowers near the boundary between these two groups, specifically the borrowers who would have chosen a GSE loan before the premium cut but crossed over to the FHA market as a consequence of the premium cut.[24] For selection to affect our results, these "switchers" to FHA loans would have had to behave differently than the non-switching population of FHA borrowers *after controlling for income, credit score, and demographic characteristics*. The fact that we control for these borrower characteristics reduces the risk that this element of selection biases our results. Additionally, as a robustness check, we remove the borrowers who are most likely to be switchers and find the results are essentially unchanged.

A second concern about selection arises from the possibility that the typical FHA borrower selects into an FHA loan based on *unobserved* characteristics. To examine this issue, we test whether borrowers sort into FHA or GSE loans based on observable pricing differences between the two types of loans. As shown in Section 4 below, the pricing of GSE loans relative to FHA loans depends systematically on the borrower's credit score and LTV. In particular, the top panel of Table 11 in that section displays the pricing difference in 2014 for 40 combinations of credit score and LTV buckets. For the test at hand, we aggregated the 40 combinations into three groups: those for which (1) the GSE loan had monthly payments at least $50 below the FHA loan, (2) the FHA loan had monthly payments at least

[24] We show later that the FHA premium cut likely induced some borrowers to choose an FHA loan instead of GSE financing with a loan-to-value ratio of 95 percent or a bit less.

$50 below the GSE loan, and (3) the difference was less than $50. Using the NMRI data for primary owner-occupied purchase loans in 2014, we then calculated the GSE share of the combined GSE and FHA loan count in each group. For the group with the GSE pricing advantage, 96 percent of the loans were GSE loans; conversely, for the group with the FHA pricing advantage, only 4 percent of the loans were GSE loans. And for the group with small pricing differences, the loans were split about 50/50 between GSE and FHA loans. Thus, we find that observable pricing differences explain the sorting extremely well, leaving little role for unobservable characteristics.

A final concern is that even though borrowers select into FHA or GSE mortgages based on observable variables, the borrowers themselves might be fundamentally different in unobserved ways. In Appendix D we discuss identification with unobservable characteristics. The existence of such characteristics *per se* does not pose a problem. Rather, the issue concerns changes in unobservables during the time window used to estimate the diff-in-diff regression. We show that unbiased estimation of the effect of the premium cut requires that any changes in unobservable variables from the pre-event period to the post-event period be the same for the FHA and GSE borrowers. For example, if FHA borrowers boosted their expectation of future income growth relative to that of GSE borrowers at exactly the time of the premium cut, the coefficient estimates we report would represent the combined impact of the premium cut and the relative change in expected income growth. Since we study behavior using a relatively narrow window of time around the premium cut, we assume any change in unobservable variables for either group of borrowers is small, which implies that the difference in such changes between the two groups is unlikely to influence our estimates. Appendix D also discusses pre-event trends for FHA and GSE borrowers, showing that identical pre-trends for the two groups are both unlikely and not required for identification.

4. Results

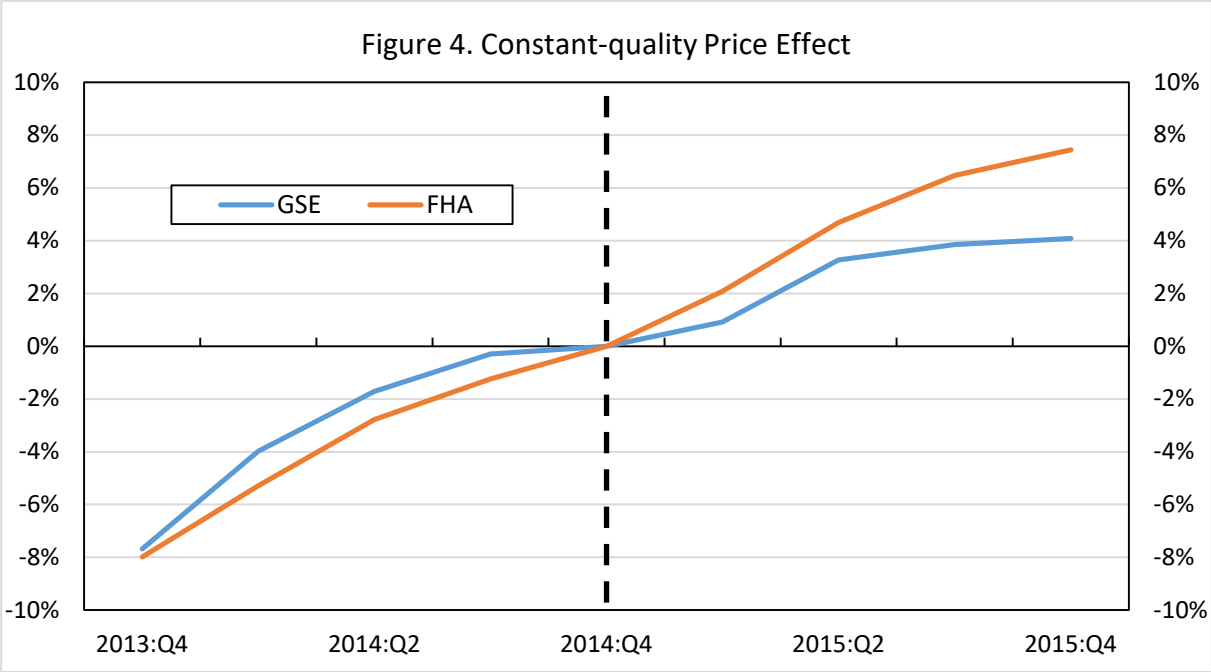
4.1. Constant-quality price effect

We start by describing the results of the regressions for the constant-quality price effect, i.e. equation 1 with the December 2014 AVM value as an additional right-hand-side variable. The estimated coefficients on the quarterly dummies and the *FHA**Q* interaction dummies trace out the implied path of constant-quality prices for the control and treatment groups.[25] In our baseline regressions, the control

[25] Among the results for variables other than the quarterly dummies, income and credit score have positive coefficients, as expected, in equation 1 and all other regressions estimated in the paper. The full set of estimated coefficients from all regression results we report is posted at http://www.aei.org/fha-premium-cut-paper-regression-tables/.

group consists of GSE loans in census tracts with an FHA share of less than 5 percent and the treatment group consists of FHA loans in census tracts where the FHA share is at least 20 percent. There are 189,199 loans in this baseline sample.

Figure 4 plots the constant-quality price series for FHA-financed and GSE-financed homes. It is important to note that the diff-in-diff calculation operates in growth rates and not in levels. From 2013:Q4 to 2014:Q4, constant-quality prices for FHA- and GSE-financed homes increased at essentially the same rate. After the premium cut, a significant gap emerged in the growth rates, with the constant-quality price of homes purchased by FHA borrowers rising at a faster rate. Table 3 shows the diff-in-diff results for the three treatment periods. Averaging over these treatment periods, we estimate the premium cut induced a statistically significant 2.8 percentage point increase in constant-quality prices for FHA-financed homes.[26]



[26] Two features of Figure 4 deserve comment. First, there is a noticeable seasonal pattern – especially for GSE loans – with the largest price increases in the first half of the year. This reflects strong demand during the spring buying season and is a well-known pattern in house prices. Second, the gap that opens up after the premium cut owes to a deceleration in GSE prices rather than an acceleration in FHA prices. Of course, in a diff-in-diff analysis, the existence of a gap is all that matters. Nonetheless, as a check on our GSE price series, we constructed a parallel series from FHFA's tract-level house price indices for GSE-financed transactions and found that this parallel series shows the same deceleration as in Figure 4.

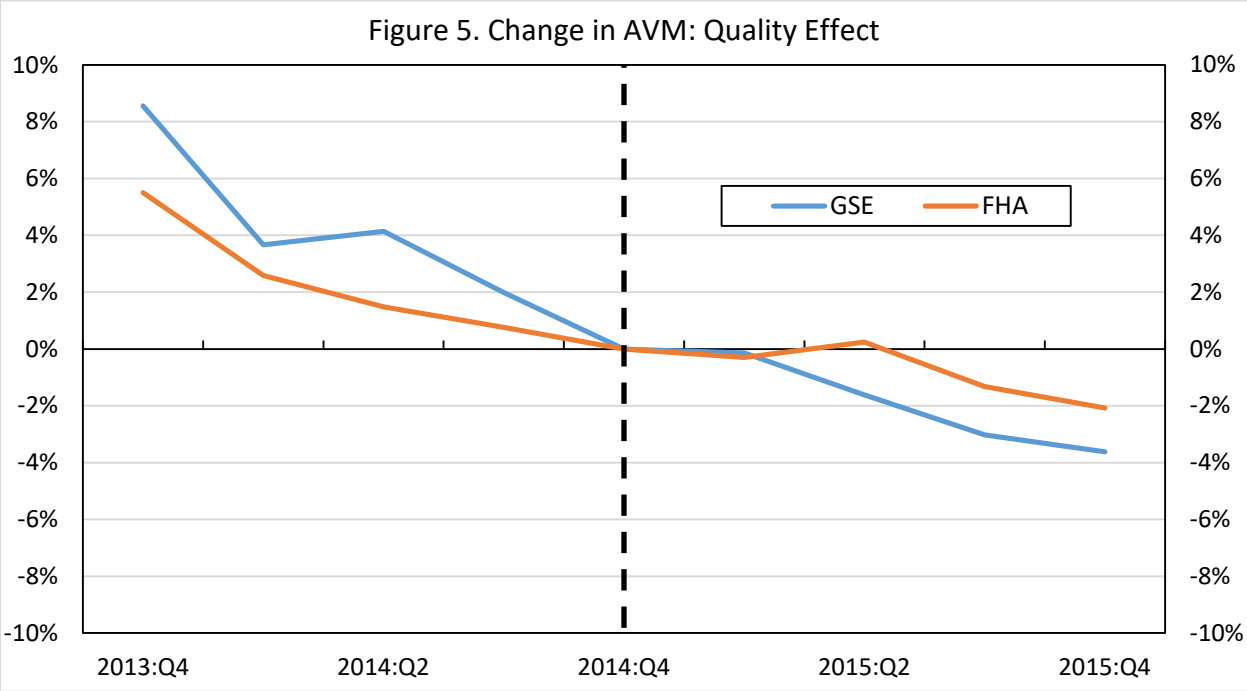
Table 3. Constant-quality Price Effect

End of Treatment Period	Diff FHA	Diff GSE	Diff-in-Diff
2015:Q2	-0.5 ppt	-2.7 ppts**	2.2 ppts**
2015:Q3	-0.3 ppt	-3.5 ppts**	3.3 ppts**
2015:Q4	-0.6 ppt	-3.6 ppts**	3.0 ppts**
Average over treatment periods	-0.4 ppt	-3.3 ppts**	2.8 ppts**

* and ** denotes significance at the 10 percent and 5 percent levels, respectively. Diff-in-diff results may not equal the difference between the two columns to the left due to rounding.

4.2. Quality effect

Equation 2 measures the extent to which FHA borrowers opted to buy higher-quality homes relative to GSE borrowers after the premium cut. If the quality of homes purchased over time had remained constant, the regression would show a flat time path for the quarterly dummies and the *FHA*-by-quarter dummies. Figure 5 shows, however, that the time path of these dummy variables declines for both GSE and FHA borrowers over the sample period. This decline in quality likely stems from the sharp rise in constant-quality home prices shown in Figure 4. For homes financed with FHA loans, constant-quality prices jumped more than 15 percent between 2013:Q4 and 2015:Q4; the increase for GSE-financed homes, though not quite as large, was still about 12 percent. With incomes rising at a much slower pace, many homebuyers would have had to shift to less expensive homes over this period.



Given that quality was declining throughout the sample period, what can we say about the effect of the premium cut on the quality of homes purchased by FHA borrowers? Our diff-in-diff approach tests whether the rate of quality decline of FHA-financed homes changed after the premium cut differently than that for GSE-financed homes, after controlling for important borrower characteristics. For example, the 2015:Q3 row of Table 4 shows that the quality change for FHA-financed homes accelerated by 3.4 percentage points from 2013:Q4-2014:Q3 to 2014:Q4-2015:Q3, as quality declined more slowly over the second period. The equivalent figure for GSE-financed homes is 3.6 percentage points, leading us to estimate a 0.1 percentage point decline in quality (after rounding) of FHA-financed

Table 4. Change in AVM: Quality Effect

End of Treatment Period	Diff FHA	Diff GSE	Diff-in-Diff
2015:Q2	4.3 ppts**	2.8 ppts**	1.5 ppts
2015:Q3	3.4 ppts**	3.6 ppts**	-0.1 ppt
2015:Q4	3.4 ppts**	4.9 ppts**	-1.5 ppts
Average over treatment periods	3.7 ppts**	3.8 ppts**	-0.1 ppt

* and ** denotes significance at the 10 percent and 5 percent levels, respectively. Diff-in-diff results may not equal the difference between the two columns to the left due to rounding.

homes relative to GSE-financed homes in response to the premium cut. Although the results vary across the three treatment windows, we find no statistically significant evidence of a different pattern between FHA-financed and GSE-financed homes in any of the three treatment windows or in the average of the three. Therefore, we cannot reject the null hypothesis that the FHA premium induced no change in the quality of homes purchased.

4.3. Overall effect

Figure 6 and Table 5 show the same information for the total price changes, which includes changes in both constant-quality prices and home quality. These results are freely estimated and are not constrained to be the sum of the price and quality effects estimated earlier. As Figure 6 shows, the purchase price of GSE-financed homes was essentially flat both before and after the premium cut, whereas the price of FHA-financed homes was rising before the cut and then accelerated after the cut. Averaging over the three treatment periods, Table 5 indicates that the FHA premium cut induced a 2.5 percentage point increase in the market price of homes purchased by FHA borrowers relative to GSE borrowers. This increase is statistically significant at the 5 percent level.

Our estimate of an overall price effect of 2.5 percentage points implies a semi-elasticity of housing demand along the intensive margin to the mortgage rate of about 3.4 (calculated as 2.5 divided

by 0.73 percent, the decline in the mortgage rate with the same effect on monthly payments as the premium cut). This figure is within the very wide range of 1.2 to 9.1 estimated by Adelino et al. (2014). We can also compare our results to those obtained by DeFusco and Paciorek (2017). They estimate the semi-elasticity of mortgage debt to the mortgage interest rate, and find that a one percentage point increase in the mortgage rate reduces the size of the first-lien mortgage taken out by 2 to 3 percent. To obtain directly comparable results, we replace the dependent variable in equation 1, the natural log of house price, with the natural log of the first lien amount. We find a statistically significant average effect over the three treatment periods of 2.5 percentage points, the same as the overall price effect, which leaves the semi-elasticity at 3.4. Although our point estimate is above the range from DeFusco and Paciorek (2017), the difference is not statistically significant.

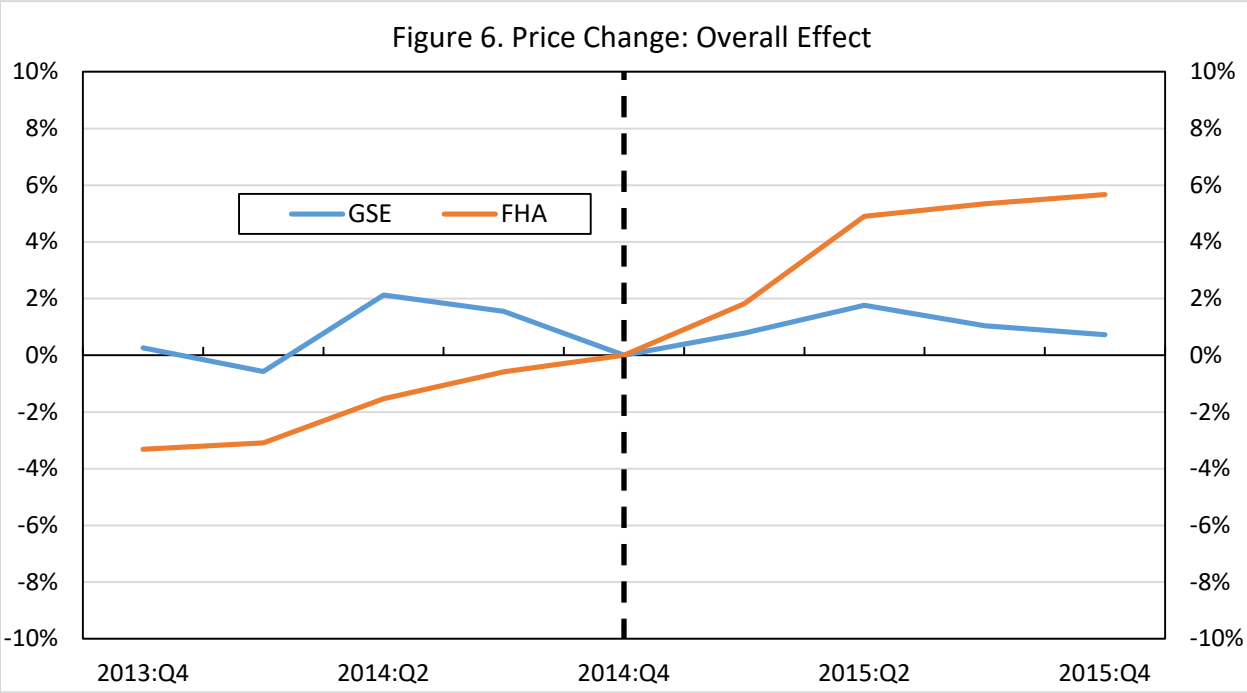


Table 5. Price Change: Overall Effect

End of Treatment Period	Diff FHA	Diff GSE	Diff-in-Diff
2015:Q2	3.1 ppts**	-0.1 ppt	3.2 ppts**
2015:Q3	2.6 ppts**	-0.2 ppt	2.9 ppts**
2015:Q4	2.4 ppts**	1.0 ppt	1.4 ppts
Average over treatment periods	2.7 ppts**	0.2 ppt	2.5 ppts**

* and ** denotes significance at the 10 percent and 5 percent levels, respectively. Diff-in-diff results may not equal the difference between the two columns to the left due to rounding.

Table 6 summarizes our results and converts the effects to dollar terms based on the FHA's average 2014 home price of $207,000 in the 23 counties in our dataset, calculated using HMDA data. Averaging over the three treatment periods, we find that the FHA premium cut induced a statistically significant 2.8 percentage point increase in the constant-quality price of homes purchased by FHA borrowers relative to GSE borrowers, no significant change in the quality of homes purchased, and a statistically significant overall price effect of 2.5 percentage points. In levels, the premium cut raised the average sale price of FHA-financed home between $5,100 (overall effect) and $5,800 (sum of price and quality effects) vis-à-vis homes with GSE financing.

Table 6. FHA Price Acceleration in 2015 vis-à-vis GSE Loans

End of Treatment Period	Constant-quality Price Effect		Quality Effect		Overall Effect	
	in ppts	in dollars	in ppts	in dollars	in ppts	in dollars
2015:Q2	2.2 ppts**	$4,500**	1.5 ppts	$3,000	3.2 ppts**	$6,600**
2015:Q3	3.3 ppts**	$6,700**	-0.1 ppt	-$300	2.9 ppts**	$5,900**
2015:Q4	3.0 ppts**	$6,300**	-1.5 ppts	-$3,100	1.4 ppts	$2,800
Average over treatment periods	2.8 ppts**	$5,900**	-0.1 ppt	-$100	2.5 ppts**	$5,100**

Note: Constant-quality price effect, quality effect, and overall effect are estimated separately and are not constrained to add up. Effects in dollars are based on 2014 average FHA purchase price of $207,000 in the 23 sample counties. * and ** denotes significance at the 10 percent and 5 percent levels, respectively.

4.4. Pre-event trends

In diff-in-diff analyses, it is common to test whether the treatment and control groups have the same trends before the event of interest as a way to provide evidence on the comparability of the two groups. To conform with this common practice, we tested the null hypothesis of parallel pre-trends by checking whether the quarterly dummies for GSE and FHA borrowers in Figures 4, 5, and 6 are equal for every quarter in the pre-event period. The results show that we cannot reject the hypothesis of parallel pre-trends for the constant-quality price regression but can reject it for the quality regression and the overall price regression.[27] Thus, our diff-in-diff findings for the quality effects and overall price effects do not satisfy this standard diagnostic.

We would note, however, that tests of pre-trends are only an indicator, and not a definitive test, of the key condition that must hold for unbiased estimation. As we show in Appendix D, this condition is that unobservable variables that influence housing choices – for example, expected income growth -- do not change *differentially* between the treatment and control groups during the estimation window

[27] Full results are available from the authors on request.

around the premium cut. If there were such a differential change in unobservables, the estimated treatment effect could reflect that change rather than the cut itself. Persistent differences in unobservables across the treatment and control groups can generate non-parallel pre-trends, but do not result in biased estimates.

To limit the threat to identification in our diff-in-diffs, we include many control variables that account for important observable differences across the treatment and control groups. This limits the range of possible unobserved factors. In addition, we study a narrow time window around the surprise premium cut, so that any changes to unobservables are likely to be small. We believe these safeguards minimize the risk of biased results.

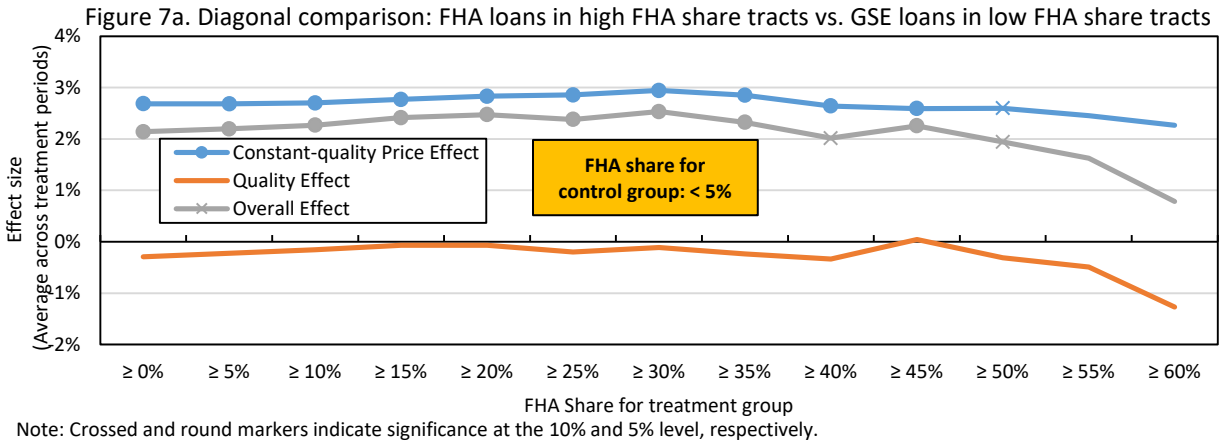
4.5. Robustness checks

Figures 7a and 7b show our results for constant-quality price effects (blue), quality effects (orange) and overall effects (grey) when we vary the definition of the treatment group but keep the baseline definition of the control group. Each point on these graphs represents a separate regression estimate, with a solid circle indicating that the estimate is statistically significant at the 5 percent level and a cross indicating significance at the 10 percent level. In Figure 7a, the treatment group consists of all borrowers in census tracts with an FHA share indicated on the x-axis. Moving to the right along the x-axis defines the treatment group to have a progressively higher minimum FHA share. The takeaway from Figure 7a is that our main result – the overall effect of the FHA premium cut is due exclusively to the constant-quality price effect – is not sensitive to the threshold value of the FHA share for inclusion in the treatment group.

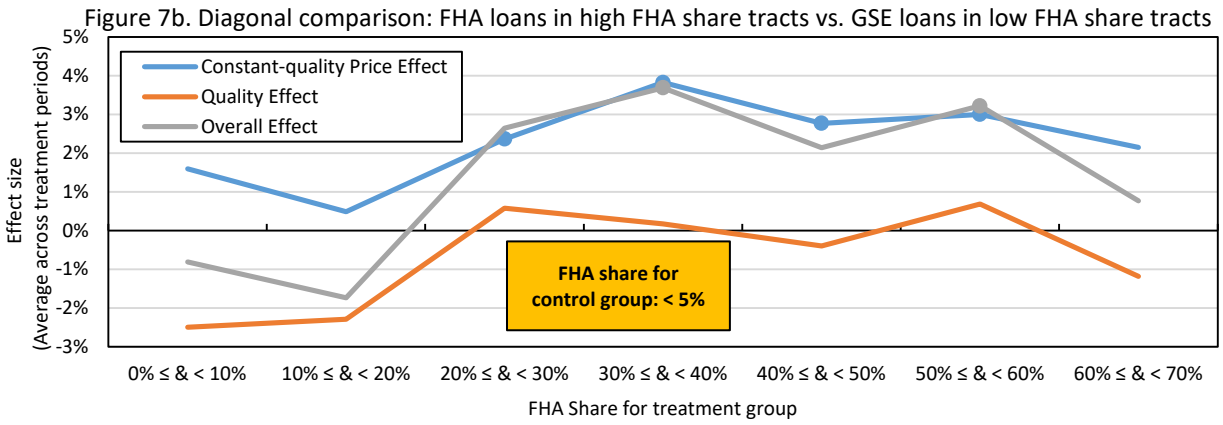
Figure 7b shows our results when we consider non-overlapping sets of census tracts for the treatment group. For example, the results for the x-axis at "20% ≤ & < 30%" show our regression results when the treatment group is limited to households with FHA mortgages living in census tracts with an FHA share between 20 percent and 30 percent.[28] This figure highlights that the point estimate of the constant-quality price effect is 2 percentage points or greater as long as the treatment group consists of census tracts with an FHA share of at least 20 percent. Moreover, these effects are statistically significant for any grouping of tracts with FHA shares between 20 percent and 60 percent. For tracts with FHA shares below 20 percent, the constant-quality price effects become small and insignificant,

[28] The highest FHA-share bucket in Figure 7b (and Figure 8b below) is 60 to 70 percent because, as shown in Figure 3, very few census tracts have FHA shares above 70 percent.

presumably because the FHA borrowers benefitting from the premium cut are not numerous enough to affect the local market.

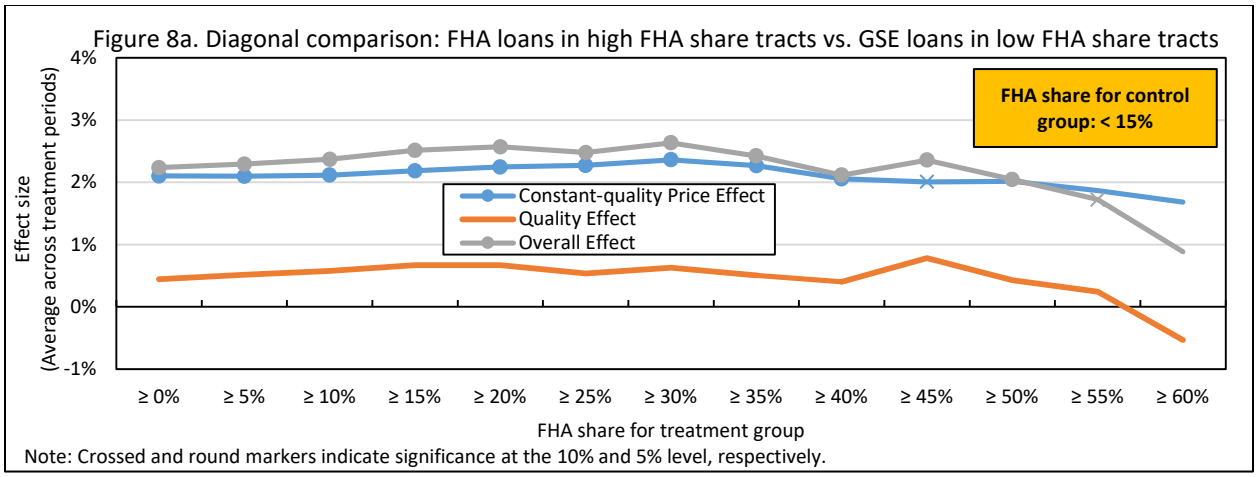


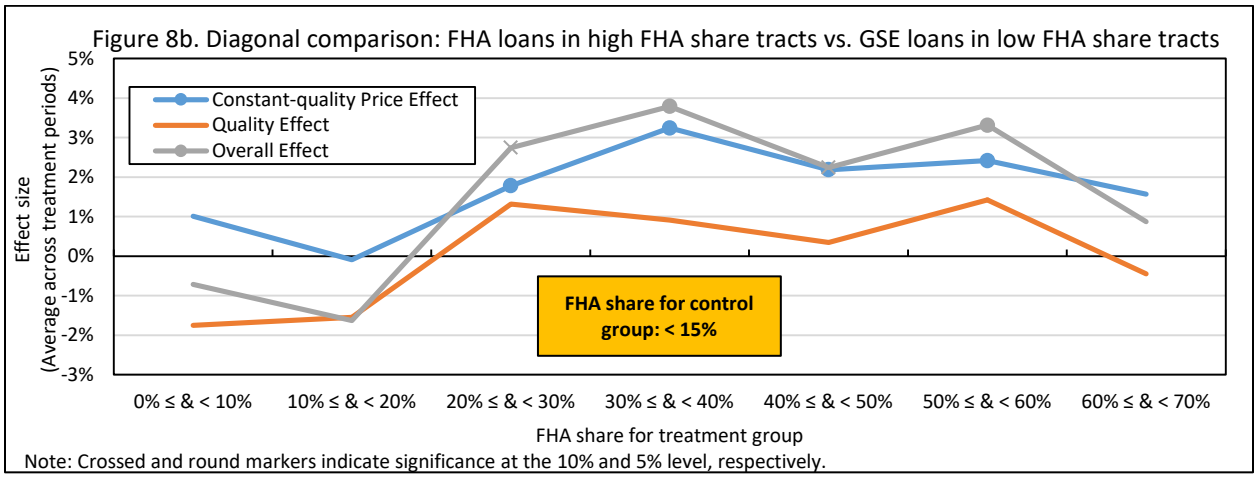
Note: Crossed and round markers indicate significance at the 10% and 5% level, respectively.



Note: Crossed and round markers indicate significance at the 10% and 5% level, respectively.

Figures 8a and 8b display the same information as 7a and 7b except that we change the control group to borrowers with GSE mortgages located in census tracts with an FHA share of less than 15 percent. Qualitatively, the results are the same as in Figures 7a and 7b. The key difference is that the constant-quality price effect falls to about 2 percentage points. We believe the effect declines because the control group now includes census tracts that are not as insulated from the behavior of FHA borrowers as in the baseline.





Our next robustness tests address concerns about the influence of unobservable variables, possible changes in the FHA and GSE groups after the premium cut, and the validity of using GSE borrowers as the control group. Beginning with the possible influence of unobservables, we run two alternatives to the baseline regression that vary the set of control variables. The first alternative removes all the borrower controls, leaving only the quarterly diff-in-diff time dummies and the county dummies as controls. The second alternative goes in the opposite direction by augmenting the baseline regression with controls for additional characteristics – the borrower's debt-to-income ratio (DTI) and the combined loan-to-value ratio including any second liens (CLTV).[29] If the results are stable across these very different alternative sets of control variables, that would provide some confidence that the results would remain stable if it were possible to include unobservable controls.

[29] This analysis controls for the DTI and CLTV through risk bucketing. The DTI buckets are defined as follows, in percent: 1-33, 34-38, 39-43, 44-50, and ≥ 51. The CLTV buckets, also in percent, are 1-60, 61-70, 71-75, 76-80, 81-85, 86-90, 91-95, and ≥ 96.

To test for the effects of changes in the FHA and GSE groups, we exclude borrowers most likely to have switched from GSE financing to an FHA loan after the premium cut, whose behavior conceivably could differ from standard FHA borrowers. We characterize this part of the market by using the loan-level risk ratings available as part of the NMRI data. The NMRI estimates the probability of default under stressed conditions akin to the 2007-08 financial crisis.[30] We exclude FHA loans with stressed default rates between 8 and 18 percent, as both the FHA and the GSEs have sizable market shares for these loans. The FHA dominates the market for loans with stressed default rates above 18 percent, while the GSEs dominate the market for loans with stressed default rates below 8 percent.

The final two tests in the table focus on the use of GSE borrowers as a control group. In the first test, we exclude the GSE borrowers with a credit score or income above the 95th percentile of the FHA distribution for their county and the quarter during which their purchase occurred. This exclusion removes GSE borrowers with a financial profile that is seldom observed among the FHA borrowers. The second test retains the full GSE control group but replaces the two parallel sets of controls for GSE and FHA borrower characteristics with a common set of controls (that is, we constrain the coefficient vector b_7 in equations 1 and 2 to be zero). Although the data reject this constraint, as noted above, pooling the GSE and FHA borrowers in this way allows us to see whether directly controlling for observable differences between the two groups would affect our results.

The results from these tests are shown in Table 7. The first row of the table repeats our baseline results for reference. Row 2 removes all the borrower controls from the baseline regression, row 3 adds the controls for DTIs and CLTVs to the baseline regression, row 4 excludes loans with stressed default rates between 8 and 18 percent (what we call "poachable" loans), row 5 excludes the GSE borrowers with high credit scores or income, and row 6 imposes a common set of borrower controls. As shown, these alternatives do not change any of the core takeaways from our baseline results. The constant-quality price effect and the overall price effect always remain statistically significant and lie in a fairly narrow range around their baseline values. And in every case, the quality effect continues to be insignificant. These results provide some assurance that our baseline findings are not driven by unobservables, by changes in the mix of FHA and GSE borrowers or loan characteristics due to the premium cut, or by the use of GSE borrowers as the control group.

[30] For more on the NMRI's methodology see http://www.housingrisk.org/wp-content/uploads/2016/10/Housing-Risk-NMRI-methodology-Oct-2016.pdf.

Table 7. FHA Price and Quality Acceleration in 2015 vis-à-vis GSE Loans , Alternative Control Variables and Sets of Loans: FHA loans in high FHA share tracts vs. GSE loans in low FHA share tracts

	Average across treatment periods			Number of loans
	Constant-Quality Price Effect	Quality Effect	Overall Effect	
1. Baseline	2.8 ppts**	-0.1 ppt	2.5 ppts**	189,199
2. Baseline excluding all borrower controls	2.6 ppts**	0.3 ppt	2.8 ppts**	189,199
3. Baseline plus controls for CLTV and DTI buckets	2.8 ppts**	0.5 ppt	3.0 ppts**	183,180
4. Baseline excl. poachable loans[1]	2.8 ppts**	-0.3 ppt	2.3 ppts**	144,739
5. Baseline excl. high income/high credit score GSE borrowers[2]	3.4 ppts**	-0.8 ppt	2.3 ppts**	170,174
6. Baseline with common GSE/FHA borrower controls[3]	2.5 ppts**	0.4 ppt	2.9 ppts**	189,199

* and ** denote significance at the 10 percent and 5 percent levels, respectively.
[1] Poachable loans are loans with an NMRI value between 8% and 18%. This range represents the range of greatest competition between FHA and the GSEs. For more on the NMRI methodology, see http://www.housingrisk.org/wp-content/uploads/2017/01/Housing-Risk-NMRI-methodology-January-2017-FINAL.pdf.
[2] Excludes GSE borrowers with a credit score or income above the 95th percentile of the distribution for FHA borrowers in their county and quarter of purchase.
[3] Pools FHA and GSE borrowers to estimate coefficients on borrower characteristics (the vector $\boldsymbol{X}$ in eqns. 1-3) by constraining $\boldsymbol{b_7}$ to be zero.

4.6. Spillovers to GSE borrowers

Our results so far have compared FHA borrowers in FHA-rich census tracts to GSE borrowers in tracts with few FHA loans. This raises the question: how did the premium cut affect GSE borrowers in census tracts with a substantial FHA presence? To address this question, we redefine the control group to consist of GSE borrowers in the same census tracts as the treatment group of FHA borrowers. In terms of Figure 2, the control group is now represented by the orange box A, while the treatment group continues to be the blue box B. This "within-market" regression design removes the geographic separation between the treatment and control groups in our baseline set-up. Both groups come from census tracts with an FHA share of purchase loans of at least 20 percent.

Table 8 shows that these within-market results are strikingly different than the baseline results. The constant-quality price effect is now very small and statistically insignificant. That is, the premium cut is estimated to have had essentially the same effect on constant-quality prices for both FHA borrowers and GSE borrowers in FHA-rich areas. In these areas, FHA borrowers – armed with the extra purchasing power from the premium cut – pushed up the prices of *all* homes, including those bought by households using GSE mortgages. This spillover makes sense, as FHA and GSE borrowers would have competed with each other for the same homes in a small geographic area.

The price spillover to GSE borrowers helps explain the results for the within-market quality effect. In contrast to the baseline results, we find that the premium cut had a positive impact on the quality of homes purchased by FHA borrowers relative to GSE borrowers in the same census tracts. Because GSE borrowers faced higher market prices but had no additional purchasing power, they reduced the quality of the homes they bought relative to FHA borrowers. This result documents that the premium cut had adverse effects on non-FHA borrowers in areas with a substantial FHA presence.[31,32]

Table 8. FHA Price and Quality Comparison in 2015 vis-à-vis GSE Loans: Baseline Results and Within-Market Results

	Average across treatment periods			**Number of loans**
	Constant-Quality Price Effect	**Quality Effect**	**Overall Effect**	
1. Baseline	2.8 ppts**	-0.1 ppt	2.5 ppts**	189,199
2. Within-Market	0.4 ppt	2.5 ppts**	2.4 ppts**	351,927
** denotes significance at the 5 percent level.				

4.7. Placebo test

We consider a placebo test for both our baseline and within-market results. If the significant effects we have found are real, and not simply random noise, these effects should go away if we instead use "fake" policy dates when the premium was unchanged. Finding such dates is difficult because the FHA premium also changed in April 2013, not long before the January 2015 cut that we analyze. Using quarterly data, that leaves only 2013:Q3 to 2014:Q4 as a possible window for a placebo test. With this constraint, we can run one – and only one – placebo test. We assume a fake policy change at the end of 2014:Q3 and then use 2014:Q3 to 2014:Q4 to measure the policy effect, with 2013:Q3 to 2013:Q4 as the pre-event comparison period. We use the same set of quarters in both periods, as we do for all the other results in the paper, because of the differential seasonal effects for GSE and FHA borrowers.

[31] These within-market results are robust to the same tests shown in Table 7 for the baseline results. In particular, the constant-quality price effect is small and insignificant at the 5 percent level in every version of the regression, while the quality effect is large and significant at the 5 percent level in every version except the regression that excludes poachable loans, where the effect drops to 1.8 percentage points and is significant at the 10 percent level.

[32] The adverse quality effects for GSE borrowers were largest in the tracts with very high FHA shares. In these tracts, the average income of mortgage borrowers – including GSE borrowers – is much lower than in tracts with little FHA presence. These lower-income GSE borrowers would have had fewer resources to avoid moving down-market in the face of higher constant-quality prices.

In this placebo test, the key significant results in the paper – the constant-quality price effect in the baseline set-up and the quality effect in the within-market analysis – become insignificant. This contrast provides additional assurance that the significant effects we find are real.[33]

4.8. Impact on homeownership

We now estimate the number of new homebuyers attributable to the premium cut. When announcing the cut, FHA projected it would add 250,000 new first-time homebuyers over the following three years, roughly 83,000 per year over 2015-2017. NMRI data show that FHA's volume of home purchase loans rose by 217,000 from 2014 to 2015 and that 176,000 of the added 2015 loans went to first-time buyers. But only part of this increase represents new demand drawn into the market by the lower premium. Some FHA buyers entered the market in 2015 because an improving economy was raising incomes and others would have bought homes with financing from other government agencies had FHA not cut its premium.

We disentangle the sources of new FHA business after the premium cut as follows. First, we estimate a multinomial logit regression model using the NMRI data to predict the extent to which the FHA "poached" homebuyers in 2015 from other government agencies. We then estimate the amount of new business FHA would have picked up in 2015 based on prevailing market trends. The number of new homebuyers that FHA attracted through the premium cut equals the total increase in FHA purchase loans less the sum of poached loans and new business due to market trends. This estimate includes both first-time buyers and repeat buyers. In the final step, we separate out the first-time buyers.

We estimate the multinomial logit model using loans originated in 2014, the year prior to FHA's premium cut, to establish the baseline relationship between the agency guaranteeing the loan and each loan's observable characteristics.[34] The estimation uses the NMRI data for primary owner-occupied home purchase loans, as FHA does not guarantee investor or second-home loans. Based on the model's fit for 2014 and observable characteristics for loans made in 2015, we then predict loan volume for each agency in 2015, the year after the premium cut. The key assumption is that the characteristics of loans made by each guarantee agency would have been stable from 2014 to 2015 had the premium cut not

[33] The placebo test also yields an anomalous result: the insignificant quality effect in the baseline set-up using the actual premium cut becomes significant at the fake policy date. If it were possible to run placebo tests for a number of fake dates, we could judge whether this finding is an outlier. The fact that we cannot do this highlights the limitation of being able to conduct only a single placebo test.

[34] The explanatory variables in the multinomial logit model are borrower credit score, state in which the property is located, price of the property, number of borrowers, a first-time buyer dummy, and whether the loan was originated by a bank, credit union, nonbank, or state housing finance agency. For more details and an assessment of model fit, see Appendix E. As discussed there, analysis using holdout samples shows that the model accurately predicts loan counts by agency.

occurred. The difference between each agency's actual 2015 loan count and the predicted count is our estimate of FHA poaching.

Table 9 summarizes the results of this exercise. As shown in the second-to-last line in the table, we estimate that FHA poached about 117,000 primary owner-occupied home purchase loans in 2015 from the other agencies, with the bulk coming from the GSEs. When scaled by their vastly different sizes, the GSEs lost about 5 percent of their volume of primary owner-occupied purchase loans in 2015, RHS lost about 23 percent, and VA lost about 2 percent.

The results of the multinomial logit model also allow us to estimate the market trend – that is, the amount by which FHA's purchase loan count would have risen from 2014 to 2015 had the premium cut not occurred. The starting point for this calculation is in the final row of Table 7. As shown, we estimate that the purchase loan count for the non-FHA agencies as a group would have increased 13.5 percent from 2014 to 2015 had they not lost loans to FHA through poaching induced by the premium cut. This figure provides an anchor to estimate FHA's trend growth. In particular, if FHA's share of the agency purchase loan market would have been unchanged in 2015 in the absence of the premium cut, its trend growth in that year would have matched the 13.5 percent trend growth of the non-FHA agencies. Alternatively, if FHA's market share would have risen in 2015, its trend growth would have exceeded 13.5 percent, and vice versa. Using the regression analysis detailed in Appendix F, we estimate that FHA's share of agency purchase loans would have been unchanged from 2014 to 2015 had the premium cut not occurred. With trend growth of 13.5 percent from a 2014 base of 594,000 FHA home purchase loans, we estimate that FHA's purchase loan count would have risen by about 80,000 in 2015 without the premium cut.

Table 9. Summary of Poaching Calculation (all loan counts in thousands)

First-lien, primary owner-occupied purchase loans		GSEs	VA	RHS	GSE+VA+RHS
2014	Actual (from NMRI)	1,288	281	136	1,706
2015	Actual (from NMRI)	1,384	319	116	1,819
	Predicted (from model)	1,461	325	151	1,937
2015 FHA Poaching: 2015 Predicted - 2015 Actual		76	6	35	117
2015 Trend: (2015 Predicted / 2014 Actual) – 1		13.4%	15.5%	10.6%	13.5%

Note: GSE+VA+RHS totals may not equal the sum of individual agency figures due to rounding. All figures in the table pertain to primary owner-occupied purchase loans.

Table 10 assembles the results so far. Of the 217,000 home purchase loans that FHA added in 2015, 54 percent were poached from other agencies; 37 percent represented market trend growth

unrelated to FHA's premium cut; and only about 9 percent, or 20,000, were new purchase loans resulting from the premium cut. This figure includes both first-time buyers and repeat buyers.

Table 10. Disaggregation of 2015 Increase in FHA Purchase Loan Volume

	Count in thousands	**% of 2015 FHA increase**
Total	217	100%
Poaching	117	54%
Trend	80	37%
New Homebuyers	20	9%

Note: All figures in the table pertain to FHA purchase loans taken out by first-time buyers and repeat buyers.

To focus on first-time buyers, we estimate a separate regression – also detailed in Appendix F – for the share of FHA's home purchase loans taken out by these buyers. The results of this regression indicate that the premium cut had no significant effect on FHA's first-time buyer share in 2015. First-time buyers accounted for 81.6 percent of FHA's home purchase loans in both 2014 and 2015. We use this constant first-time buyer share to apportion the 20,000 new homebuyers in 2015 into first-time buyers and repeat buyers, which yields about 17,000 new first-time buyers. These 17,000 new buyers represent about 3.1 percent of the 549,000 first-time buyers we would have expected FHA to underwrite in 2015 had the premium cut not occurred, implying a one-year semi-elasticity of homeownership to interest rates of about 4.2 for first-time buyers using FHA financing.[35]

There is a wide confidence band around our estimate of first-time buyers generated by the premium cut, owing to uncertainty about how FHA's share of the purchase loan market would have changed from 2014 to 2015 had the FHA premium remained constant. The number of new buyers induced by the premium cut in 2015 cannot be estimated precisely because normal fluctuations in FHA's share of the agency purchase loan market are large relative to the incremental effect of the premium cut. As discussed in Appendix F, we estimate that the 95 percent confidence band for this change in the share, which is centered at zero, runs from -0.8 percentage point to +0.8 percentage point. Given this result, the implied 95 percent confidence band for new FHA homebuyers in 2015 runs from a bit below zero to about 48,000. For first-time buyers alone, the upper end of the confidence band would be somewhat lower – about 39,000 if we scale down the figure for all buyers by FHA's first-time buyer share. This imprecision carries over to our estimate of the semi-elasticity of homeownership to interest rates, for which the 95 percent confidence band ranges from roughly zero to ten. Even so, we can

[35] The estimate of 549,000 FHA first-time buyers in 2015 is computed by grossing up the actual number of FHA first-time buyers in 2014 (484,000) by 13.5 percent, the estimated trend growth rate for FHA's purchase loans from 2014 to 2015 in the absence of the premium cut. The semi-elasticity of 4.2 equals the 3.1 percent increase in first-time buyers divided by the 0.73 percentage point decline in the mortgage rate that equates to the premium cut.

conclude that the premium cut was much less effective at bringing in new first-time buyers than FHA expected.

4.9. Wealth transfers

The premium cut redistributed wealth among participants in the housing market. Existing homeowners benefitted from the rise in constant-quality home prices, which boosted their housing wealth. First-time homebuyers using FHA financing came out ahead as well, due to the premium cut.[36] In contrast, non-FHA first-time buyers were made worse off, as they faced higher house prices but did not benefit from the cut in monthly mortgage insurance premiums.

While it is impossible to do a complete analysis of transfers without a fully-specified macro model of the housing market, we can use our results to discuss the cost borne by first-time non-FHA homebuyers. Recall that Figure 7b showed a statistically significant increase in constant-quality house prices for FHA borrowers in census tracts with an FHA share between 20 percent and 60 percent; the average constant-quality price increase across the census-tract buckets in this range was 3 percent. This constant-quality price increase affected not only FHA borrowers but other borrowers in these tracts too.

In 2015, there were 1.22 million non-FHA home purchase borrowers in the country with loan amounts below FHA's 2014 loan limit in tracts with an FHA share of at least 20 percent and below 60 percent.[37] We estimate that about 500,000 of these borrowers were first-time homebuyers.[38] With an average constant-quality price increase of 3 percent in these tracts as a result of the premium cut, each of these 500,000 non-FHA first-time homebuyers paid approximately $6,200 extra per house, a total extra payment of about $3.1 billion.[39]

With regard to possible effects of the premium cut on rents, we lack the data needed to examine this issue. That said, we show that the premium cut failed to create many new first-time homebuyers,

---

[36] As noted in the introduction, the premium cut increased the purchasing power of the typical FHA borrower by about 6 percent, which was only partly offset by the higher constant-quality house prices induced by the cut.

[37] The 1.22 million borrowers are represented by the orange bars in the right panel of Figure 3 for FHA shares from 20 percent to 60 percent. The number would be greater if we included cash sales, which we cannot determine with our data sources.

[38] To estimate the number of first-time homebuyers, we first compute the first-time buyer share in 2015 for each government agency using the NMRI data. For private-sector conventional loans, we assume the same first-time buyer share as for GSE loans. We then multiply these shares by the respective counts of conventional, FHA, VA, and RHS home purchase loans in HMDA and sum the results to arrive at roughly 500,000 first-time buyer loans. We focus exclusively on first-time buyers because the effect on repeat buyers, who benefit from rising prices when selling their home, is ambiguous.

[39] We derive the $6,200 extra per house as follows: Using HMDA data for 2015, the average loan amount for all non-FHA home purchase loans in tracts with an FHA share between 20 and 60 percent was $175,000. Assuming an average down payment of 15 percent, the $175,000 loan amount implies an average home price of $206,000, and 3 percent of that amount is about $6,200.

which implies the pool of households seeking rental units did not change much. With little change in rental demand, we would expect rents to have been little affected. Further research on this topic, however, would be useful.

4.10. Why didn't the FHA poach more buyers?

Our analysis suggests that about half of the increase in FHA homebuyers after the premium cut consisted of buyers who otherwise would have used financing from other government agencies. This begs the question: Why didn't the FHA poach even more buyers? The answer is that FHA mortgages are more expensive than GSE mortgages for most buyers, even after the FHA premium cut.

Table 11 summarizes the relative cost of GSE and FHA mortgages before the premium cut (top panel) and afterwards (middle panel). Each cell shows the savings in monthly mortgage payment from using a GSE mortgage instead of an FHA mortgage when the borrower is financing the purchase of a $250,000 home. All FHA mortgages are assumed to have an interest rate of 4 percent and a 3.5 percent down payment. Because nearly all FHA borrowers roll the upfront 1.75 percent mortgage insurance premium into the loan amount, the effective LTV is 98.2 percent. The monthly payment includes the applicable annual insurance premium. The GSE loans we consider have a range of possible LTVs (shown by the rows in the table) and have an interest rate equal to 4 percent plus a loan-level pricing adjustment (LLPA).[40] For GSE loans with LTVs above 80 percent, the monthly payment includes the required private mortgage insurance premium. Given the difference in LTVs across the GSE and FHA loans, the GSE loans are smaller than the FHA loans throughout the table. The larger FHA loan amounts reflect a realistic feature of the choice between GSE and FHA loans. Borrowers opting for FHA loans generally choose the maximum LTV in order to minimize their down payment.

[40] The assumed 4 percent interest rate for FHA loans equals the median note rate for all 30-year fixed-rate FHA purchase loans in 2015, calculated from the NMRI data. For GSE loans, the note rate shown in the NMRI data embeds the effect of the LLPA, which varies by credit score and loan-to-value ratio. We reverse engineer the unadjusted GSE note rate for each primary owner-occupied 30-year fixed-rate purchase loan by subtracting out the effect of the LLPA, which is converted from a loan price adjustment to a rate effect with a factor of 5:1 (meaning a 100 basis point LLPA adjustment equals 20 basis points in additional note rate). The average of the reverse-engineered pre-LLPA GSE note rates is about 4 percent. Also note that loan pricing for GSE mortgages is computed assuming an 80 percent LTV for the cells in the LTV-bucket rows marked "1-80%" and a 97 percent LTV for the rows marked "≥ 96%".

Table 11. GSE versus FHA Pricing and Changes in GSE Volume

Monthly GSE pricing advantage over FHA *before* FHA premium cut (2014)

		Credit score bucket							
		< 640	640-659	660-679	680-699	700-719	720-739	740-759	≥ 760
LTV bucket	1-80%	$342	$342	$366	$390	$396	$407	$413	$413
	81-85%	$209	$232	$245	$270	$276	$299	$312	$312
	86-90%	$80	$120	$133	$160	$167	$204	$220	$220
	91-95%	($79)	($13)	$1	$29	$36	$103	$126	$126
	≥ 96%	($183)	($127)	($127)	($91)	($91)	($34)	($17)	($17)

Monthly GSE pricing advantage over FHA *after* FHA premium cut (2015)

		Credit score bucket							
		< 640	640-659	660-679	680-699	700-719	720-739	740-759	≥ 760
LTV bucket	1-80%	$242	$242	$266	$289	$295	$307	$313	$313
	81-85%	$108	$132	$144	$169	$176	$199	$212	$212
	86-90%	($20)	$20	$33	$59	$66	$103	$119	$119
	91-95%	($179)	($113)	($99)	($71)	($64)	$3	$26	$26
	≥ 96%	($284)	($227)	($227)	($192)	($192)	($135)	($118)	($118)

Change in distribution of GSE loans by LTV and credit score bucket: 2013 to 2014 vs 2014 to 2015

		Credit score bucket							
		< 640	640-659	660-679	680-699	700-719	720-739	740-759	≥ 760
LTV bucket	1-80%	-0.2%	-0.2%	-0.3%	-0.1%	-0.1%	0.2%	0.5%	2.9%
	81-85%	0.0%	0.0%	-0.1%	-0.1%	0.0%	0.0%	0.0%	0.1%
	86-90%	-0.1%	-0.2%	-0.2%	-0.3%	-0.3%	-0.1%	0.0%	0.4%
	91-95%	-0.1%	-0.4%	-0.8%	-1.2%	-1.1%	-1.2%	-0.9%	-1.2%
	≥ 96%	0.0%	0.1%	0.2%	0.5%	0.6%	0.9%	1.0%	1.9%

Note: Pricing calculation assumes a house price of $250,000, an interest rate of 4.0% for FHA and 4.0% for the GSEs plus a loan-level pricing adjustment. FHA loan has an LTV of 96.5% plus 1.75% upfront mortgage insurance premium; the monthly payment includes the applicable annual insurance premium (1.35% before the premium cut and 0.85% afterward). GSE loan has an LTV in the bucket shown; for loans with LTVs above 80%, monthly payment includes required private mortgage insurance premium. Cells with red font indicate a pricing advantage for FHA over GSE loans for a given LTV and credit score combination. Cells shaded in yellow indicate LTV and credit score combinations for which the pricing advantage changed from the GSEs to FHA after FHA's January 2015 premium cut. Cells shaded in red indicated LTV and credit score combinations where the GSEs lost significant market share from 2014 to 2015 as compared to 2013 to 2014 based on loan counts for primary owner-occupied purchase loans.

Cells with red font in the top and middle panels indicate that the FHA loan has a lower monthly payment than the GSE loans in that cell. The cells highlighted in yellow in the middle panel mark the buyers who would have found it advantageous to have taken out a GSE mortgage before the premium cut and an FHA mortgage after the premium cut. These borrowers either had (i) LTVs between 86 and 90 percent and credit scores below 640 or (ii) LTVs between 91 and 95 percent and credit scores between 660 and 719. However, FHA financing may have also become attractive to other borrowers,

especially those with a LTV between 91 and 95 percent and credit scores above 719. For these borrowers, the GSE pricing advantage narrowed after the premium cut to just $3 or $26 per month. Accordingly, FHA may have become the preferred financing option for income-constrained borrowers in these cells given its wider credit box, which allows for DTIs up to 57 percent, whereas the GSEs limit DTIs to a maximum of 50 percent. Borrowers using less leverage and having good credit continued to have lower monthly mortgage payments with a GSE mortgage than with an FHA mortgage after the premium cut.

The bottom panel of the table shows the relative change in the composition of GSE mortgage borrowers. Each cell is computed using NMRI data as

$$\Delta GSE\%_{2015} - \Delta GSE\%_{2014}$$

where $\Delta GSE\%_{2015}$ is the change between 2014 and 2015 in the percentage of total GSE loans accounted for by loans in that cell and $\Delta GSE\%_{2014}$ is the change between 2013 and 2014.[41] The cells in the bottom panel highlighted in red show the buckets of borrowers whose representation in the total pool of GSE loans declined the most after the premium cut. These cells are all for LTVs between 91 and 95 percent and for credit scores above 660 – many of the same cells where the FHA gained a pricing advantage relative to GSE mortgages after the premium cut.[42]

4.11. Supply response

Our analysis has examined the effect of the premium cut over a period of four quarters or less. We chose a relatively short post-event period to exclude subsequent developments in the mortgage market that could contaminate our results. Nonetheless, the increase in constant-quality home prices that we find in FHA-rich localities could have prompted additional housing construction over a longer period. If so, it is possible that the longer-term effect of the premium cut on house prices could have been smaller than what we find through the end of 2015.

To explore this possibility, Table 12 shows the estimated number of newly-constructed housing units relative to the stock of existing units in the census tracts that comprise our treatment and control groups.[43] The estimates are based on public records data obtained from First American, supplemented by data from Zillow. As shown, the rate of new home construction in the control-group tracts was

[41] We study *the change in the change* in the distribution because credit standards were declining throughout this period, causing the shares in the higher-risk cells of the table to trend up.
[42] In 2015, the GSEs initiated an effort to guarantee more loans with 3 percent down payments, explaining the increase in the share of these loans (bottom row, bottom panel).
[43] The series for new construction and the existing stock pertain to one-unit structures.

essentially flat over 2013-2017 at about 0.2 percent per year. In contrast, construction in the treatment-group tracts picked up in 2015 and 2016 before dropping back in 2017 to a pace that was still above the 2013 and 2014 levels. The final column quantifies the extent of the increase in the treatment-group tracts relative to the control-group tracts.

Table 12. New Home Construction as a Share of Housing Stock, One-Unit Structures

Year	Treatment-group tracts	Control-group tracts	Difference
2013	.46%	.20%	.26 ppt
2014	.42%	.17%	.25 ppt
2015	.48%	.18%	.30 ppt
2016	.55%	.17%	.38 ppt
2017	.50%	.19%	.31 ppt

Note: Census tracts in Kern County, CA, and Pima County, AZ, are omitted as these counties are not included in the source data. New home construction and the housing stock both pertain to one-unit structures. The housing stock is measured as of December 2018. The new construction shares for 2017 have been increased by 20 percent from those shown in the source data as a rough adjustment for the incomplete reporting of recent home sales in the public records data.
Source: Authors' calculations using public records data from First American on new home construction and the housing stock, supplemented by data from Zillow on new home construction.

The results in Table 12 are consistent with a housing supply response to the premium cut. However, these results do not imply that the rise in constant-quality prices we that measured for 2015 necessarily overstates the full effect of the premium cut. First, some of the supply response occurred in 2015, and thus would have been built into our estimate of the price increase. Second, the supply response in the treatment-group tracts in 2016 and 2017 could have been driven, at least in part, by demand-boosting factors other than the premium cut.

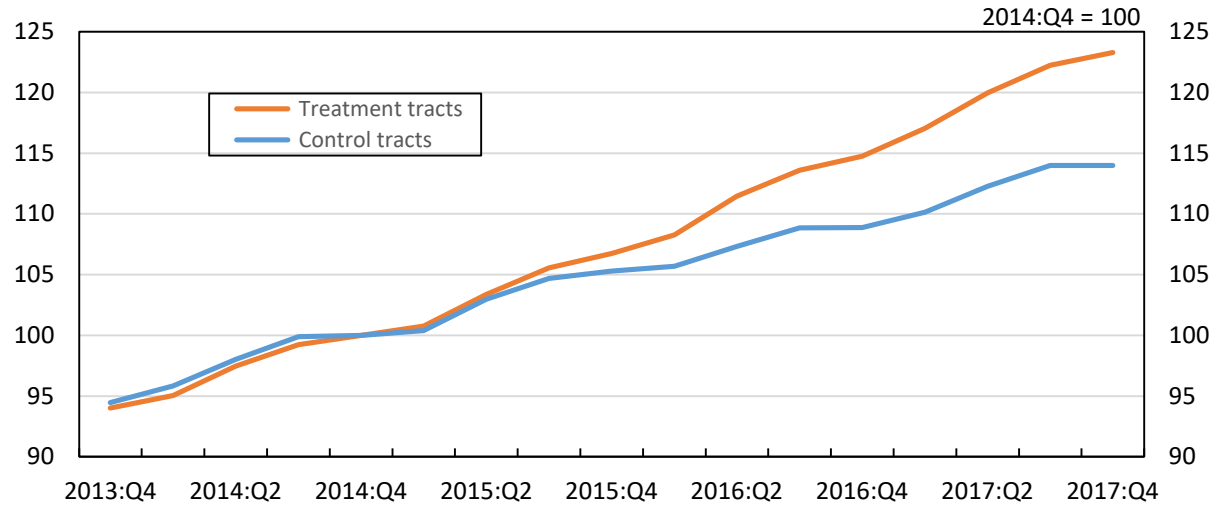
If the premium cut were the only factor that increased demand in the treatment-group tracts, we would expect the price rise in these tracts relative to the control-group tracts to have narrowed in 2016 and 2017 as the new supply came on line. To assess whether this was the case, Figure 9 displays constant-quality house price indices through 2017 that we constructed from public records data obtained from First American. Like the price indices calculated for our diff-in-diff analysis, these indices are estimated from hedonic regressions that use property-level AVMs as the control for quality, but there are some differences in methodology.[44]

Figure 9 shows that the two price indices were essentially identical during the full pre-event period for our diff-in-diff analysis (2013:Q4-2014:Q4) and then begin to diverge in 2015. Although the

[44] In particular, the indices in Figure 9 do not include controls for borrower and loan characteristics due to the limited information in the public records data. They also use all home sales in the treatment-group and control-group census tracts, rather than only FHA-financed sales in the treatment-group tracts and GSE-financed sales in the control-group tracts. In this sense, the indices in Figure 9 are more like generic house price indices than the customized indices we use for the diff-in-diff analysis.

gap between the two series in 2015:Q4 is not as wide as in our diff-in-diff analysis, the pattern is qualitatively the same. The gap widens much further in 2016 and 2017, indicating that the supply

Figure 9. Constant-quality House Price Index



Note: The indices are based on hedonic regressions of sale price for homes sold in the treatment and control census tracts on tract dummies, quarterly dummies, and the property-level AVM in 2014:Q4.
Source: Authors' calculations using public records data from First American.

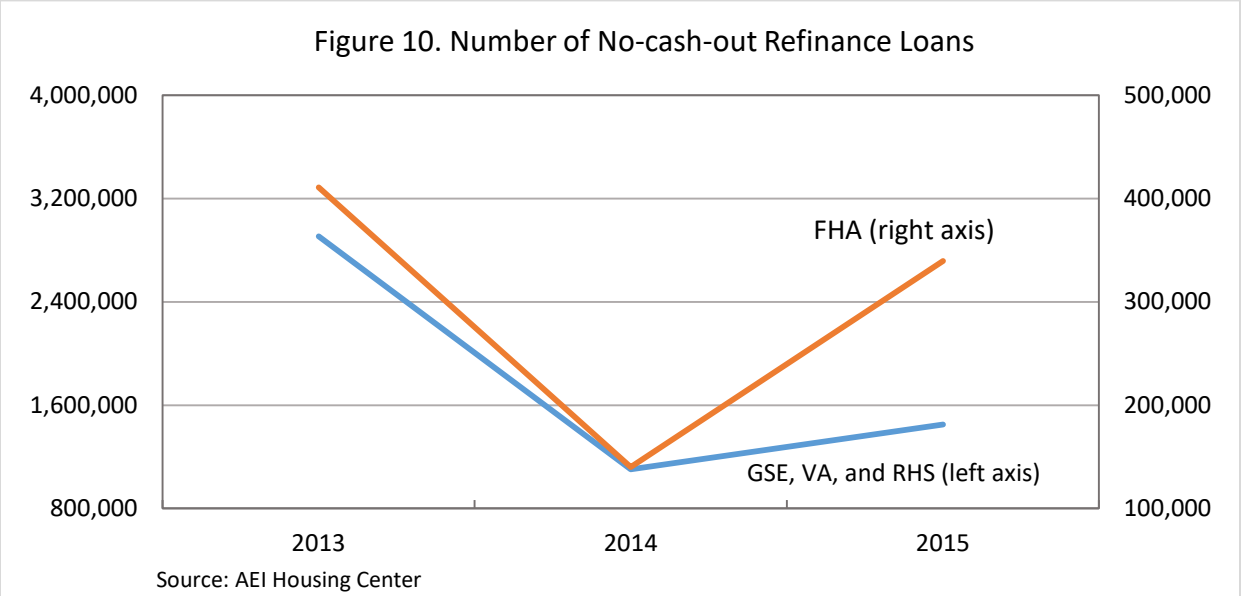
response in the treatment-group tracts was dwarfed by factors boosting demand. A full investigation of these demand influences is beyond the scope of this paper, but the GSEs' efforts to regain market share from FHA likely was a key factor. Since 2015, the GSEs have greatly expanded their acquisition of home purchase loans with CLTVs above 95 percent. In 2015, their acquisitions of these loans amounted to only 8.5 percent of FHA's volume, but this share rose to more than 14 percent in 2016 and nearly 26 percent in 2017, according to the NMRI data. The increased availability of high-CLTV loans would have expanded the supply of mortgage credit in treatment-group tracts relative to control-group tracts.

To summarize, the premium cut does appear to have spurred some housing supply response in the treatment-group tracts. However, it is impossible to gauge the size of the supply response due to the premium cut *per se*, or its effect on house prices beyond 2015, because other factors caused prices to rise in 2016 and 2017 in these tracts. The bottom line is that one cannot conduct a clean analysis of the premium cut that far removed in time from the event itself.

4.12. Other impacts

The FHA premium cut induced many existing FHA borrowers to refinance into a new FHA loan in 2015, as demonstrated in Figure 10. The reduction in monthly payments due to the lower premium may

have reduced default risk for these borrowers. To the extent that it did, the FHA premium may have indirectly supported house prices and homeownership among existing FHA borrowers.



5. Conclusion

We exploit a surprise cut in the FHA's mortgage insurance premium to estimate the impact of a change in mortgage interest rates on borrower behavior for the segment of the mortgage market whose homeownership decisions are most likely to be influenced by interest rates and government policy.

Relative to a control group of GSE borrowers in census tracts with few FHA mortgages, we find that FHA borrowers increased the value of their purchased housing by 2.5 percentage points, which represents almost half of the increase in purchasing power afforded by the premium cut itself. This estimate implies an interest-rate semi-elasticity of housing demand of about 3.4 on the intensive margin. We show that the rise in spending reflected an increase in constant-quality home prices, with no significant change in the quality of housing purchased by FHA buyers.

We also estimate the impact of the premium cut on homeownership. After accounting for aggregate trends and the likelihood that many new FHA borrowers would have used other financing had the premium cut not occurred, we estimate that the change in FHA policy created about 17,000 new first-time homebuyers in the first year after the cut. However, the 95 percent confidence band around this estimate is wide, with a top end of about 39,000. The baseline estimate of 17,000 implies a one-year semi-elasticity of homeownership to interest rates of 4.2 for FHA first-time buyers, with a 95 percent confidence band that runs from roughly zero to ten. Even the top end of the wide confidence band is well below the response that the FHA expected from the premium cut.

We show that the premium cut had significant distributional effects. The rise in constant-quality home prices implies that existing homeowners received windfall gains, while non-FHA first-time buyers bore a significant cost. These unintended distributional effects likely owe, at least in part, to the tight housing market at the time of the premium cut. With limited supply, the boost to demand from the cut ended up raising home prices. In a softer market, the same policy change could have been more effective at spurring homeownership, highlighting the need to take account of market conditions when designing policy.

References

ADELINO, M, A. SCHOAR, AND F. SEVERINO (2014). "Credit Supply and House Prices: Evidence from Mortgage Market Segmentation," working paper. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1787252

BHUTTA, N., S. LAUFER, AND D. RINGO (2017). "Residential Mortgage Lending in 2016: Evidence from the Home Mortgage Disclosure Act Data, *Federal Reserve Bulletin* 103(6). https://www.federalreserve.gov/publications/files/20yyyy16_HMDA.pdf

BHUTTA, N., AND D. RINGO (2017). "The Effect of Interest Rates on Home Buying: Evidence from a Discontinuity in Mortgage Insurance Premiums, " Finance and Economics Discussion Series 2017-086, Board of Governors of the Federal Reserve System. https://www.federalreserve.gov/econres/feds/files/2017086pap.pdf

_____ (2016). "Changing FHA Mortgage Insurance Premiums and the Effects on Lending," *FEDS Notes*, September 29. https://www.federalreserve.gov/econresdata/notes/feds-notes/2016/changing-fha-mortgage-insurance-premiums-and-the-effects-on-lending-20160929.html.

CHAMBERS, M., C. GARRIGA, AND D. SCHLAGENHAUF (2009). "The Loan Structure and Housing Tenure Decisions in an Equilibrium Model of Mortgage Choice," *Review of Economic Dynamics* 12(3), 444-468.

CORBAE, D., AND E. QUINTIN (2015). "Leverage and the Foreclosure Crisis," *Journal of Political Economy* 123(1), 1-65.

DAVIS, M.A., AND S. VAN NIEUWERBURGH (2015). "Housing, Finance, and the Macroeconomy," in G. Duranton, J.V. Henderson, and W. Strange, eds., *Handbook of Regional and Urban Economics, Volume 5A*, 754-811. Amsterdam: Elsevier.

DEFUSCO, A. A., AND A. PACIOREK (2017). "The Interest Rate Elasticity of Mortgage Demand: Evidence from Bunching at the Conforming Loan Limit," *American Economic Journal: Economic Policy* 9(1), 210-240

FAVILUKIS, J., S. C. LUDVIGSON, AND S. VAN NIEUWERBURGH (2017). "The Macroeconomic Effects of Housing Wealth, Housing Finance, and Limited Risk Sharing in General Equilibrium," *Journal of Political Economy* 125(1), 140-223.

FISHER, E. M. (1951). *Urban Real Estate Markets: Characteristics and Financing*. National Bureau of Economic Research. http://papers.nber.org/books/fish51-1.

FUSTER, A., AND B. ZAFAR (2015). "The Sensitivity of Housing Demand to Financing Conditions: Evidence from a Survey," Federal Reserve Bank of New York Staff Reports No. 702, August. https://www.newyorkfed.org/medialibrary/media/research/staff_reports/sr702.pdf

GETE, P., AND F. ZECCHETTO (2017). "Distributional Implications of Government Guarantees in Mortgage Markets." http://faculty.georgetown.edu/pg252/Gete-Zecchetto-guarantees.pdf. Forthcoming in the *Review of Financial Studies*.

GREENWALD, D. L. (2018). "The Mortgage Credit Channel of Macroeconomic Transmission," MIT Sloan Research Paper No. 5184-16, revised January 18, 2018. https://ssrn.com/abstract=2735491.

JESKE, K., D. KRUEGER, AND K. MITMAN (2013). "Housing, Mortgage Bailout Guarantees and the Macro Economy," *Journal of Monetary Economics* 60(8), 917-935.

PIAZZESI, M., AND M. SCHNEIDER (2016). "Housing and Macroeconomics," NBER Working Paper No. 22354. https://ideas.repec.org/p/nbr/nberwo/22354.html

SOMMER, K., AND P. SULLIVAN (2018). "Implications of U.S. Tax Policy for House Prices, Rents, and Homeownership," *American Economic Review* 108(2), 241-274.

WILLIAMS, J. C. (2015). "Measuring Monetary Policy's Effect on House Prices," FRBSF Economic Letter 2015-28, August 31, 2015. http://www.frbsf.org/economic-research/publications/economic-letter/2015/august/measuring-monetary-policy-effect-on-house-prices-speech/

Appendix A: Data Cleaning and Trimming

We perform two rounds of data cleaning and trimming. The first round limits the ATTOM dataset to first-lien, FHA-guaranteed and conventional home purchase loans for owner-occupied properties. Our comparison of loan counts to HMDA is done with this set of loans. The second round then eliminates transactions with loan amounts above the 2014 county-level FHA conforming loan limits, transactions with missing data or with data values that could distort the results, and transactions for properties with more than one unit. This dataset is then matched to other datasets to add a variety of borrower characteristics.

A.1. First-round cleaning

1) Purchase: A loan is assumed to be a purchase loan if the field SR_TRAN_TYPE is coded "R" (Resale) or "S" (Subdivision). Loans coded "L" (Refinance or Equity) or "C" (Construction) are omitted.
2) Loan type: Loan type is based on the value in the field SR_LOAN_TYPE_1. An FHA loan is coded "F", a VA loan is coded "V", a Construction loan is coded "C". Loans with a blank value are assumed to be conventional loans or Rural Housing Services loans. We omit the VA and Construction loans, as well as loans with indeterminate loan type.
3) Owner-Occupied: All FHA loans are assumed to be for owner-occupied properties. For conventional loans, we keep loans for which the field SA_SITE_MAIL_SAME is coded "Y" (Yes), which indicates owner occupancy. Repeat sales, which represent roughly 4 percent of the sample, are missing the SA_SITE_MAIL_SAME variable. All repeat sales are assumed to be owner-occupied.[45]
4) Residential properties: we keep loans for which the field USE_CODE_STD has a use code of "condominium, PUD", "cooperative", "duplex", "multi-family dwelling (2-4 units)", "quadraplex", "single family residence", "timeshare", "triplex", "mobile home", "miscellaneous residential", and "multi-family residence (5+ units)". We omit loans with all other use codes. We keep all 1-4 unit properties to ensure an apples-to-apples comparison to HMDA. In the second round cleaning we remove the 2-4 unit properties to create the regression dataset that is limited to one-unit properties.
5) Cash sales: A sale is assumed to be a cash sale if the value of the mortgage in the field SR_LOAN_VAL_1 is zero. Cash sales are excluded.

[45] In the case of Wayne County, MI, in 2014 and 2015, as well as Macomb County, MI, in 2015, the SA_SITE_MAIL_SAME variable is defective, and we do not eliminate loans based on occupancy status. Since HMDA data show that over 90 percent of the home purchase loans in these counties for these years are for primary owner-occupied properties, our decision not to screen on occupancy status introduces little error in the data.

6) Duplicates: We remove all records that have an identical sale price, filing date, loan amount, and property location.

A.2. Second-round cleaning

1) 2nd or 3rd liens: we exclude transactions where one or both of the subordinate liens are larger than the 1st lien.
2) Upper limit on loan amount: we exclude sales with a loan amount above the respective county's 2014 FHA conforming loan limit (plus the upfront FHA mortgage insurance premium, which is generally rolled into the loan). This limit is applied to FHA and conventional loans alike across the entire sample period. We also remove the FHA loans that bunch at the applicable 2014 limit or are within 0.75 percent of the limit. This filter ensures that are results are not driven by bunching of FHA loans. (We do not apply this filter to conventional loans because there is no bunching of these loans at the county-level FHA limits).
3) Upper limit on combined loan-to-value ratio (CLTV): we exclude loans with a CLTV above 110 percent. These loans likely contain data errors or represent developer purchases of multiple properties.
4) Missing AVM and/or sale price: we exclude loans for properties that do not have both an AVM value and a reported sale price, as both are necessary for our regression analysis.
5) Outliers: we eliminate loans for which the AVM is more than double or less than half of the actual sale price. In addition, we eliminate loans for which the AVM/price ratio is in the top or bottom 1 percent of values within each sale month. We do this trim separately for FHA-financed and conventionally-financed properties.
6) Distressed sales: we eliminate all distressed sales because their sale price may not be a true reflection of the property's value.
7) Missing census tract: we eliminate loans with a missing census tract identifier.
8) One-unit properties: we keep loans for which the field USE_CODE_STD has a use code of "condominium, PUD", "cooperative", "single family residence", or "mobile home". We omit loans with all other use codes.

The cleaned ATTOM data consists of 249,000 FHA and 424,000 conventional loans for the years 2013-2015.

Appendix B: Matching ATTOM Loans to External Data Sources that Report Borrower Characteristics

Borrower characteristics are not reported in the ATTOM data. We add them to our dataset by matching ATTOM loans to external data sources that report borrower income, credit score, demographic characteristics, and whether the borrower sought pre-approval for the loan. We match the ATTOM data to five different datasets with multiple individual matching rounds described in detail below. We use the cleaned and matched loans for the analysis in this paper.

B.1. First-round match

This round matches ATTOM data to HMDA and FHFA data to add all the borrower characteristics listed above other than credit score.

First step: We start by matching the cleaned ATTOM loans to HMDA using sale year, census tract, loan amount, and loan type (FHA or conventional). Since the first-lien loan amount is rounded to the nearest $1,000 in HMDA, we also round the first-lien loan amount in the ATTOM data. This first match step yields about 421,000 one-to-one matches out of the 674,000 cleaned loans in our dataset.[46]

Second step: Many of the loans were not deemed to match in the first round because the matches were not one-to-one – that is, more than one ATTOM loan matched to one or more HMDA loans, or vice versa. For the unmatched loans, we add lender name to the fields used in the first round to look for additional one-to-one matches. We did not use lender name in the first round because the lender names are not always reported in a consistent manner in ATTOM and HMDA. Consequently, requiring matches on lender name in the first round could have incorrectly eliminated true matches. To conduct the second match step, we use the loans that matched in the first step to establish the most common associations between ATTOM's lender code and the respondent identification code in HMDA.[47] Using this additional field for the unmatched loans resulted in about 113,000 additional matches.

Third step: For the conventional loans that remain unmatched, we do a supplemental match to loan-level FHFA data that reports borrower income. This FHFA dataset (available at http://www.fhfa.gov/DataTools/Downloads/Pages/Single-Family-Census-Tract-File.aspx) has comprehensive coverage of mortgages acquired by the GSEs. The dataset includes loans originated by lenders that are not subject to HMDA reporting and thus can augment the HMDA matches. We match from the ATTOM data to the FHFA dataset using sale year, census tract, and loan amount (rounded to

[46] The 674,000 count for cleaned loans is 1,000 greater than the sum of 249,000 FHA loans and 424,000 conventional loans mentioned above due to rounding.

[47] Respondents report loans in HMDA that they either originated or purchased. We use the ID code for the respondent that reported originating the loan.

the nearest $1,000) for the conventional loans that remain unmatched after the first- and second-round matching. This ATTOM to FHFA match resulted in about 20,000 additional exact, one-to-one matches.

With this three-step matching procedure, we obtain 206,000 FHA and 348,000 conventional matches for the years 2013-2015, which represents a match rate of 83 percent for FHA loans and 82 percent for conventional loans.

B.2. Second-round match

This round matches the ATTOM/HMDA/FHFA-matched loans to FHA's Single-Family Portfolio Snapshot (available at https://www.hud.gov/program_offices/housing/rmra/oe/rpts/sfsnap/sfsnap) to add the note rate for FHA borrowers. Although this round does not add any borrower characteristics, adding another loan variable increases the number of uniquely identifiable FHA loans for the third-round FHA match below. The FHA Snapshot dataset is made available each month by the Department of Housing and Urban Development and it contains loan-level information on all of FHA's single-family endorsements including the exact loan amount, the property's zip code, the property type, the loan type, and the originating lender and the lender sponsoring the mortgage.[48] Similar to the HMDA matching, the matching of FHA loans is performed over two steps using only one-to-one matches between the datasets.

First step: We start by matching the FHA loans in the ATTOM/HMDA/FHFA-matched dataset to the FHA Snapshot dataset using 5-digit zip code, the exact loan amount, and a loan's origination date. Since the FHA dataset reports endorsement date, which has to occur after a loan is originated, we allow this date to fall as much as 5 months after the origination date.[49] We begin by checking for matches with the same endorsement and origination date and subsequently check for endorsement dates in the month after origination, then 2 months after origination, up to a maximum of 5 months after origination. We introduce little error by extending the time period for the matches because of the highly unique nature of the other matching parameters. Once a loan has been matched, we remove the loan from the ATTOM/HMDA/FHFA and FHA datasets and continue with the next stage of matching. This first matching step yields about 186,000 one-to-one matches out of the 206,000 FHA loans in the ATTOM/HMDA/FHFA-matched dataset.

Second step: Similar to the earlier matching from ATTOM to HMDA, some of the loans were not deemed to match in the first step because the matches were not one-to-one. For these unmatched

[48] The FHA Snapshot codebook defines a sponsoring lender as a lender that underwrites the loan for the originator while also deciding "whether the borrower represents an acceptable credit risk for HUD."

[49] We chose 5 months as the maximum difference between the origination date and the endorsement based on tests that showed very few matches at longer time lags.

loans, we add the lender name and, when available, the sponsor name to the fields used in the first step and repeat the matching process. Similar to the HMDA matching, we did not use lender name in the first round because of inconsistencies in lender-name reporting between the datasets. Consequently, requiring matches on lender name in the first round could have incorrectly eliminated true matches. Using this additional field for the unmatched loans resulted in about 12,000 additional FHA matches.

With this two-step matching procedure, we add the note rate to a total of 198,000 FHA loans in the ATTOM/HMDA/FHFA-matched dataset for the years 2013-2015, which represents a match rate of 96 percent.

Our next and final round of matching adds a borrower's credit score to the dataset. The matching is performed separately for conventional and FHA loans. The resulting matched loans include all the borrower characteristics for the analysis in this paper. This matching also adds the debt-to-income ratio for the vast majority of loans. These loans are used for the robustness analysis.

B.3. Third-round match for loans with conventional financing

First step: We start by matching the ATTOM/HMDA/FHFA-matched loans with conventional financing to AEI's National Mortgage Risk Index (NMRI) dataset.

The matching is done in multiple stages. The first match stage is based on geography (state), origination month, first-lien loan amount, LTV, originating lender (or the lender buying the loan as reported in HMDA through the Action Type variable)[50], number of borrowers (which is assumed to be 2 if the gender of a of the second borrower is reported in HMDA or FHFA data), property type (single family, condo, coop, mobile home), and purchaser type (Fannie Mae or Freddie Mac). We include only one-to-one matches between loans in the ATTOM/HMDA/FHFA dataset and the NMRI dataset. Once a loan has been matched, we remove it from both datasets before proceeding to the next match step.

For the remaining (unmatched) loans, we gradually loosen the matching requirements by removing lender, number of borrowers, property type, or purchaser type to allow for potential data inconsistencies between the datasets, while keeping a core set of matching fields (state, origination month, loan amount, and LTV). Since the origination date is reported differently in the ATTOM data (recordation date of the sale) and the NMRI (first payment date of the loan, which we assume to be 2 months after loan origination), we furthermore allow for one-to-one matches that fall within plus/minus 2 months of the origination date. Likewise, due to rounding of the LTV, we allow for one-to-one matches when the LTV differs by 1 percentage point across the two datasets.

[50] The NMRI data generally record the lender selling the loan to the GSEs, which can be either the originating lender or a lender that bought the loan from the originator.

These steps yield about 241,000 one-to-one matches out of the 348,000 conventional loans in the ATTOM/HMDA/FHFA-matched dataset.

Second step: For those loans not matched, we attempt a match to Fannie Mae's Single Family Loan Performance Data and Freddie Mac's Single Family Loan-Level Dataset (GSE data). These datasets are released quarterly by the GSEs to track the performance of loans they acquire. Both datasets identify the loan's 3-digit zip code, which is more granular than the state-level location indicator in the NMRI data.[51] We match based on zip code, first-lien loan amount, origination month, LTV, property type, purchaser type, lender, and number of borrowers. Similar to step one, we include only one-to-one matches between loans in ATTOM/HMDA/FHFA and the GSE data, and we remove all matched loans from both datasets before continuing with the next matching stage. We gradually loosen the matching requirements by removing lender, number of borrowers, property type, and purchaser type in subsequent matching rounds to allow for inconsistencies in data reporting between the datasets.

Since the origination date for the Freddie data has to be imputed from the monthly reporting period and the loan age, we allow one-to-one matches that fall within plus/minus 1 month of the recordation date in the ATTOM data. Using the GSE datasets resulted in 50,000 additional one-to-one matches of conventional loans.

Third step: In the first two steps of matching to the NMRI or GSE data, we required one-to-one matches. In the third step, we allow for many-to-many matches if the duplicative loans fall within the same credit score or DTI bucket. We match on the same variables and stages described in the first-step match. This yields another 8,000 matches based on credit score buckets and 7,000 matches based on DTI buckets.

Because the NMRI and GSE datasets both exclude loans without a government guarantee, the matching process just described removes all private-sector loans from the set of conventional loans in the ATTOM/HMDA/FHFA dataset. Thus, the conventional loans used in our analysis are all GSE loans.

B.4. Third-round match for loans with FHA financing

First step: We start by matching the FHA loans in the ATTOM/HMDA/FHFA-matched dataset to a subset of the NMRI dataset. We generate this subset by matching the NMRI data to the FHA Single-Family Snapshot Portfolio. The NMRI dataset does not include the 5-digit zip, the exact loan amount, or

[51] One minor limitation is that the GSE data exclude adjustable-rate mortgages (ARMs). ARMs represented only about 3 percent of GSE home purchases loans in 2013-2015. For more information on the datasets, refer to: http://www.fanniemae.com/portal/funding-the-market/data/loan-performance-data.html and http://www.freddiemac.com/research/datasets/sf_loanlevel_dataset.html.

the property type for FHA loans. These variables are added from the FHA Single-Family Snapshot dataset for half of all FHA purchase loans for the years 2013-2015. Matching is done in multiple stages. The first match stage is based on geography (5-digit zip), origination month, the exact first-lien loan amount, and note rate. As with the third-round matching for conventional loans, this matching allows us to pick up the borrower's credit score and, for many loans, the debt-to-income ratio as well.

Since the origination date is reported differently in the ATTOM data (recordation date of the sale) and the NMRI (first payment date of the loan, which we assume to be 2 months after loan origination[52]), we allow for one-to-one matches with origination dates that differ by up to 6 months in either direction. We begin by checking for matches with the same origination month and subsequently adjust the imputed NMRI origination month by one month at a time up to the plus/minus 6-month limit. Because we require loans to also match on the exact loan amount, note rate, and 5-digit zip code, this tolerance range for the origination month should introduce little error. We include only one-to-one matches between loans in the ATTOM/HMDA/FHFA-matched dataset and the NMRI. Once a loan has been matched, we remove it from both datasets and continue with the next stage of matching.

In subsequent stages, we include all NMRI loans by matching on the loan amount truncated to the nearest $1,000, state instead of zip code, LTV (to 1 decimal point, where available, otherwise the rounded LTV), the lender name, number of borrowers (which is assumed to be 2 if the gender of a second borrower is reported in HMDA or FHFA data), and the property type (single family, condo, coop, mobile home).[53] Because of the coarser geographical detail and the truncated loan amount, these stages yield fewer matches than the matches based on the subset of the NMRI matched to the FHA Snapshot data. In later stages, we gradually remove variables as matching criteria to account for data inconsistencies between the datasets. For these matching steps, we narrow the tolerance band around the origination date to plus/minus 3 months of the ATTOM-reported origination date to reduce the possibility of false positive matches. For all these stages, we include only one-to-one matches between loans in ATTOM and the NMRI. Once a loan has been matched, we remove it from the ATTOM/HMDA/FHFA-matched dataset and the NMRI dataset and continue with the next stage of matching.

---

[52] For FHA loans originated starting in 2015, the origination date is mostly reported in the NMRI data. In those cases, we use the origination date and not the first payment date with the two-month adjustment.

[53] The NMRI data generally record the lender selling the loan to the GSEs, while the ATTOM data include the lender originating the loan. If the selling lender and loan originator differ (i.e. because of a broker or correspondent agreement), we are unable, as a rule, to match these loans. However, if a given originator in the ATTOM data sells over 80 percent of its loans to a single selling lender in the NMRI data, we assume for matching purposes that the loan originator is the selling lender and declare a match on the lender name.

This first-step matching yields about 161,000 one-to-one matches out of the 206,000 FHA loans in the ATTOM/HMDA/FHFA-matched dataset.

Second step: In this step we allow for many-to-many matches if the duplicative loans fall within the same credit score or DTI bucket. We match on geography (state), the truncated loan amount, the note rate, and lender, allowing the imputed NMRI origination date to differ from the ATTOM origination date by up to plus/minus 6 months. In a second stage, we remove lender and allow for matches to have NMRI and ATTOM origination dates that differ by up to plus/minus 3 months. This yields another 10,000 matches based on credit score buckets and 11,000 matches based on DTI buckets.

For our final dataset, we only include loans with a reported credit score. This dataset consists of 171,000 FHA and 299,000 GSE loans for the years 2013-2015. Because some loans in the NMRI are missing a reported DTI and because of the many-to-many matches, which may result in matches based on credit score but not on DTI, not every loan with a reported credit score has a reported DTI. Thus, the dataset for the robustness check is slightly smaller than the dataset used for the main analysis. It consists of 165,000 FHA and 294,000 GSE loans for the years 2013-2015.

Appendix C: Weighting and FHA Shares by Census Tract

Due to our removal of some loans through the matching process, our dataset may not be representative of the full set of loans for the 23 counties. We correct for this by weighting the data by loan type (FHA or conventional), 4-digit census tract, and sale year using HMDA purchase-loan data. The weighting ensures that our regression dataset is representative of the HMDA data for each combination of loan type, census tract, and sale year, subject to one adjustment. We cap the weight for any combination at 3 to avoid giving heavy influence to thinly populated cells. This cap applies to about 8 percent of the loans in the regression dataset.

We also use HMDA data to compute the 2015 FHA loan share of all loans within any given 4-digit census tract. Because the changes in home prices and quality that we seek to measure depend on the concentration of FHA loans benefitting from the premium cut – including loans poached from other agencies – the post-event year is the relevant year in which to measure the FHA share. The 2015 FHA shares by census tract are then merged onto our final dataset.

Appendix D: Notes on Identification and Pre-trends

D.1. Identification

Suppose borrowers labeled with an *i* subscript take out FHA loans. For convenience in exposition, assume these borrowers choose growth *g* of their housing stock each period. This growth is a function

of the borrowing cost *r* (which includes the mortgage insurance premium), a set of observable characteristics *X* (such as income) and a set of unobservable characteristics *Z* (such as expected future income growth):

$$g_{it} = \alpha + \beta r_{it-1} + \Gamma X_{it-1} + Z_{it-1}. \quad \text{(D.1)}$$

Note that g_{it} denotes growth between *t-1* and *t*, explaining the *t-1* subscripts on the right-hand-side variables. Assume that the borrowing cost at *t-1* for FHA borrowers is a function of observables

$$r_{it} = \Pi X_{it-1} \quad \text{(D.2)}$$

and substitute equation D.2 into equation D.1 to get

$$g_{it} = \alpha + (\beta\Pi + \Gamma)X_{it-1} + Z_{it-1}. \quad \text{(D.3)}$$

We cannot identify β, the parameter of interest, as there is no variation in borrowing cost independent of *X*.[54]

Now suppose the premium cut occurred before households chose g_{it+1}. The surprise cut studied in this paper was the same for all FHA homebuyers and hence was independent of *X*. Let *p* denote the change in borrowing cost such that

$$r_{it} = \Pi X_{it-1} - p,$$

which implies

$$g_{it+1} = (\alpha - \beta p) + (\beta\Pi + \Gamma)X_{it} + Z_{it}. \quad \text{(D.4)}$$

The presence of the constant term α prevents us from identifying β using data only on FHA borrowers since *p* does not vary across borrowers.

Denote the Δ symbol as a first difference, i.e. the period *t+1* value less the period *t* value. Then (after suppressing the redundant time subscripts) equation D.4 less equation D.3 can be written as

$$\Delta g_i = -\beta p + (\beta\Pi + \Gamma)\Delta X_i + \Delta Z_i. \quad \text{(D.5)}$$

Consider now a version of equation D.5 that holds for GSE borrowers, labeled with a *j* subscript. Note that since GSE borrowers did not experience a premium cut, they have no βp term:

$$\Delta g_j = (\beta\Pi + \Gamma)\Delta X_j + \Delta Z_j.^{55} \quad \text{(D.6)}$$

[54] We abstract from the small variation in borrowing cost caused by search frictions as presumably we would not want to identify β based on small differences in *r* across people.

[55] For convenience we hold fixed the coefficients linking observables to interest rates and housing demand across equations D.5 and D.6. Changing this assumption does not affect the intuition.

Then, the "diff-in-diff" of the FHA borrowers less the GSE borrowers can be written as

$$\Delta g_i - \Delta g_j = -\beta p + (\beta\Pi + \Gamma)(\Delta X_i - \Delta X_j) + (\Delta Z_i - \Delta Z_j), \quad \text{(D.7)}$$

where $\Delta Z_i - \Delta Z_j$, as an unobserved variable, is part of the regression error term. Unbiased estimation of β requires that the premium cut be uncorrelated with $\Delta Z_i - \Delta Z_j$. That is, the estimation of β will be confounded if unobserved characteristics of FHA and GSE borrowers change differentially during the time window used to estimate the diff-in-diff regression. To minimize this risk, our estimation strategy uses a relatively short window so that any changes to unobservables are likely to be small. If ΔZ_i and ΔZ_j are both small, the difference between them will be as well.

D.2. Pre-trends

One question that arises is whether we should expect growth of the housing variables to be identical between the FHA and GSE borrowers prior to the premium cut. The short answer is "no." Given our framework and letting α* denote the intercept of an equation like D.3 for GSE borrowers, $g_{it} - g_{jt}$ is the expression for pre-trends, which equals

$$g_{it} - g_{jt} = (\alpha - \alpha^*) + (\beta\Pi + \Gamma)\left(X_{it-1} - X_{jt-1}\right) + \left(Z_{it-1} - Z_{jt-1}\right). \quad \text{(D.8)}$$

Even if the FHA and GSE borrowers have the same values of the unobservable variables, as long as their observables differ they will have different pre-trends. Identical pre-trends should be expected in a diff-in-diff environment only when individuals are randomly assigned to the two groups. This assumption does not hold for FHA and GSE borrowers as we know they sort into mortgages based on observables. For our estimates of β to be unbiased we only require, as noted above, that any change in unobservables for the FHA and GSE borrowers be identical during the time window around the premium cut. Common pre-trends are not needed for unbiased estimation of β.

The pre-trends we present in the paper are calculated from quarterly dummy variables after controlling for a large set of observables. Because we have controlled for the influence of the X variables in equation D.8, the pre-trends we calculate will differ solely because of differences in unobservables. What must hold for unbiased estimation is that the unobserved variables do not change differentially for FHA and GSE borrowers within the estimation window.

Appendix E: Multinomial Logit Model

Using the NMRI data for home purchase loans, we estimate a multinomial logit regression to model the federal agency that guaranteed the loan (the GSEs, FHA, VA, or RHS). The explanatory variables for the logit are the borrower's credit score, a first-time buyer dummy, the home's sale price

(represented by a set of price buckets), a dummy variable for the number of borrowers, the state in which the property is located, and lender type (bank, credit union, nonbank, or state housing finance agency). The probabilities of the four agencies are additive for each loan and equal 100 percent. We estimate the model on primary owner-occupied purchase loans originated in 2014, the year prior to FHA's premium cut. We then use the estimated model coefficients, combined with the values of the explanatory variables for loans originated in 2015, to predict the loan totals for each agency in that year. The predicted loan totals are the sum of the estimated logit probabilities for each agency.

To assess the validity of the model, we conduct the following holdout analysis for the 2014 loans: We randomly sample 1 million of the nearly 2.3 million 2014 loans and estimate the model described above with these loans. We then use the model results to predict the agency probabilities for the loans held out of the regression. We repeat this process 500 times, which generates a set of 500 predicted aggregate loan counts for the holdout sample for each agency.

The results indicate that the model provides a very good fit for the agency distribution in the holdout sample. The 95 percent confidence band for the prediction error for FHA's loan count in 2014 ranges from -600 to +1,000 loans. As discussed in section 4, we estimate that FHA poached about 117,000 loans from other agencies, where the poached volume is calculated as FHA's actual volume in 2015 minus its predicted volume from the logit regression. Accordingly, the confidence band for FHA's prediction error implies a 95 percent confidence band for the number of poached loans that runs from 116,000 to 118,000. If the structure of the model had changed from 2014 to 2015 (apart from the effect of the premium cut), this confidence band would understate the true confidence band for the amount of FHA poaching. Nonetheless, the very tight band generated by the 2014 holdout exercise implies that the change in model structure would have to have been quite substantial for the confidence band to be wide.

Appendix F: FHA Market Share and First-time Buyer Share

As described in section 4, our estimate of new first-time buyers brought into the market by FHA's premium cut depends on two essential building blocks. The first is an estimate of what FHA's share of the agency purchase loan market would have been in 2015 had the premium cut not occurred. This is needed for our calculation of market trend growth in FHA's purchase loan count. The second building block is an estimate of how the premium cut affected the share of first-time buyers among FHA's purchase loans, which we need to convert results for the combination of first-time and repeat buyers into an estimate of new first-time buyers alone. This appendix covers both issues after describing the

historical data on FHA's share of the purchase loan market that we use in the analysis of FHA's market trend growth.

F.1. Historical data on FHA's market share

Monthly HMDA data provided by Bhutta, Laufer, and Ringo (2017) are the starting point for calculating FHA's historical market share. The HMDA data show first-lien, primary owner-occupied home purchase loan counts for 1-4 unit properties for FHA, VA, RHS, and the conventional market. As a shorthand, we will refer to first-lien, primary owner-occupied home purchase loans simply as "purchase" loans.

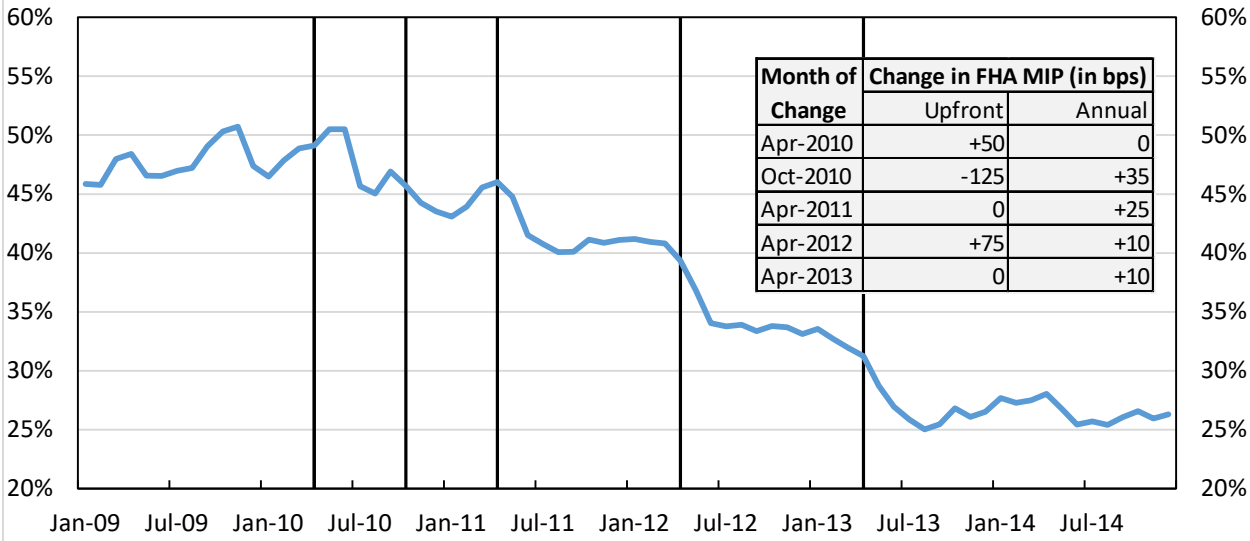
Because the HMDA data do not separate GSE loans from the rest of the conventional market, we use other data sources to estimate the GSE purchase loan count. The loan-level Fannie and Freddie datasets described in the main text provide a census of fully-documented, fully-amortizing, fixed-rate loans outside of specified affordable housing programs. These datasets exclude (1) loans that have less than full documentation or that do not fully amortize each month (such as interest-only loans), (2) high-LTV loans guaranteed under specified affordable housing programs, and (3) adjustable-rate mortgages (ARMs). To deal with the first of these exclusions, we begin our historical dataset in 2009. Starting in that year, the purchase loans guaranteed by the GSEs all have had full documentation of income and have been fully amortizing, thus rendering the exclusion irrelevant.

To account for the omitted high-LTV loans, we compare the monthly count of GSE fixed-rate purchase loans with an LTV greater than 95 percent from the Fannie/Freddie datasets to the corresponding near-universe count from the NMRI data for the overlap period that begins in September 2012. We use the monthly ratio of the NMRI count to the Fannie/Freddie count to gross-up the count of high-LTV loans. Prior to September 2012, we gross-up the Fannie/Freddie count of loans with an LTV greater than 95 percent by 20 percent, which is roughly the average ratio for the overlap months.

To account for the missing ARMs, we turn to CoreLogic's Loan Level Market Analytics (LLMA) dataset, which contains data from loan servicers for a subset of the market. For each month, we compute the ratio of GSE ARMs to GSE fixed-rate mortgages and gross-up the GSE counts based on this ratio.

Combining the resulting GSE purchase loan count with the HMDA purchase loan counts for FHA, VA, and RHS, we compute the monthly FHA share of agency purchase loans from January 2009 to December 2014. Figure F.1 plots the monthly FHA share.

Figure F.1: FHA Share of Agency Purchase Loan Count



Month of Change	Change in FHA MIP (in bps)	
	Upfront	Annual
Apr-2010	+50	0
Oct-2010	-125	+35
Apr-2011	0	+25
Apr-2012	+75	+10
Apr-2013	0	+10

Note: Data refer to first-lien, primary owner-occupied, 1-4 unit agency purchase loans. Vertical lines drawn at dates of changes in the FHA mortgage insurance premium.
Source: HMDA, NMRI, Fannie Mae's Single Family Loan Performance Data and Freddie Mac's Single Family Loan-Level Dataset, and CoreLogic.

F.2. Estimating FHA's 2015 market share had the premium cut not occurred

We use the 2009-2014 data to estimate the factors influencing the FHA share of agency purchase loans. Then, given the regression results, we predict the FHA share of the agency purchase loan market in 2015 had there been no premium cut. The regression we estimate is as follows:

$$MKT_{FHA} = b_0 + b_1 Month + b_2 Payroll + \sum_{i=1}^{5} b_{3,i}\, PhaseIn_i + \sum_{i=1}^{5} b_{4,i}\, Regime_i + e \qquad \text{(F.1)}$$

where MKT_{FHA} represents the FHA share of agency purchase loans, *Month* represents a set of monthly dummies to capture seasonal variation (these dummies are constrained to sum to zero), and *Payroll* represents the percent change in nonfarm payroll employment from twelve months earlier, which we include as a proxy for the effect of the state of the economy on the mortgage market. We omit time subscripts to keep the notation simple. The other variables in the regression account for the five changes in FHA's mortgage insurance premia from 2009 to 2014 (see Bhutta and Ringo 2016). Figure F.1 shows that the FHA share adjusts relatively quickly to a premium change and then settles at a new permanent level. To account for this pattern, $PhaseIn_i$ is a dummy for the month of the ith premium change and the following two months, the typical adjustment period, while $Regime_i$ is a dummy for all subsequent months of the ith premium regime. The dates of the five premium changes are listed in the

inset box to Figure F.1.[56] We estimate equation F.1 as a constrained OLS regression and calculate robust standard errors.

The regression confirms what can be seen from Figure F.1 – that the main determinants of the FHA purchase-loan share are changes in the FHA premia, which affect the pricing of FHA loans relative to other agency loans. The regression also finds a slight seasonal pattern. The FHA share declines in the summer months as repeat buyers with school-age children, who rely more on the GSEs than FHA, complete transactions before the beginning of the new school year. Finally, the 12-month change in nonfarm payroll employment is not a significant determinant of the FHA share.

Based on these findings, we use equation F.1 to calculate the fitted monthly FHA shares for 2014 and the predicted monthly shares for 2015. Because the final FHA premium change in our historical sample period occurred in April 2013, all of 2014 is covered by the last regime dummy. We apply this dummy to 2015 as well, as our objective is to estimate the 2015 FHA share had the January 2015 premium cut not occurred. Note also that we ignore the monthly dummies because they wash out, by definition, over each calendar year. Thus, the fitted 2014 FHA shares and the predicted 2015 shares are calculated as

$$MKT_{FHA} = b_0 + b_2 Payroll + b_{4,5} Regime_5 \qquad \text{(F.2)}$$

We average the 12 monthly shares for each year to arrive at an average FHA share for 2014 and a predicted share for 2015. We find that the predicted 2015 share is unchanged from the fitted 2014 share – an unsurprising result since the only time variation in equation F.2 comes from the change in payroll employment, which has a coefficient very close to zero.

We also estimate the 95 percent confidence interval around the estimate of no change in the FHA share. Using equation F.1, the change in the FHA share from the fitted 2014 value to the unknown 2015 value, denoted ΔMKT_{FHA}, is

$$\Delta MKT_{FHA} = b_2(Payroll_{2015} - Payroll_{2014}) + e \qquad \text{(F.3)}$$

where $Payroll_{2015}$ is the average 12-month change in payroll employment for the months of 2015, $Payroll_{2014}$ is defined analogously for 2014, all other explanatory variables drop out of the

[56] In addition to these five changes, in June 2013 FHA mandated that borrowers would have to pay annual insurance premiums over the entire life of the loan. Previously, the annual premium had been cancelled when the loan balance dropped to 78 percent of the home purchase price. Because this change followed so closely after the April 2013 increase in the annual premium, we roll it into premium regime that began in April 2013.

equation because they are the same in 2014 and 2015, and the error term e is nonzero in 2015 but is set to zero in 2014 because we are computing the change from the fitted 2014 value.

As shown by equation F.3, the two sources of uncertainty for the change in the FHA share are the confidence band around b_2 and the regression error term in 2015. To construct the 95 percent confidence interval for ΔMKT_{FHA} from equation F.3, we use b_2 +/- 2 times its standard error and add on the 95 percent confidence interval for the average value of the error term in 2015. For each month in 2015, we randomly sample from a normal distribution with mean zero and variance equal to the squared RMSE of the estimated regression in F.1. We average the 12 monthly errors to arrive at an estimated annual error and repeat this process 500 times. The bounds for the 95 percent confidence interval of the error distribution equal the 2.5th and the 97.5th percentile of the simulated distribution. Plugging this error band into equation F.3 along with the +/- 2 standard error range for b_2, the resulting 95 percent confidence band for the change in the 2015 FHA share ranges from -0.8 percentage point to +0.8 percentage point.

F.3. Estimating FHA's first-time buyer share after the January 2015 premium cut

FHA's monthly production reports show the share of FHA purchase loans that went to first-time buyers starting in June 2012.[57] We use these data to estimate a regression that relates the first-time buyer share to the same factors as in equation F.1 (with necessary modifications due to the change in the sample period):

$$FTB_{FHA} = b_0 + b_1 Month + b_2 Payroll + \sum_{i=5}^{6} b_{3,i}\, PhaseIn_i + \sum_{i=5}^{6} b_{4,i}\, Regime_i + e \qquad \text{(F.4)}$$

where FTB_{FHA} is FHA's first-time buyer share as reported in the monthly production reports, and *Month, Payroll*, $PhaseIn_i$, and $Regime_i$ are as defined in equation F.1. The two premium regimes within the June 2012 to December 2015 estimation period are the one that begin in April 2013, which was regime 5 in equation F.1, and the January 2015 premium cut, which we label as regime 6. We estimate equation F.4 as a constrained OLS regression and calculate robust standard errors.

We find the coefficient on $Regime_6$, the fully phased-in first-time buyer share after the January 2015 cut, is not significantly different from the coefficient on $Regime_5$, the fully phased-in share after the April 2013 change, which applied throughout 2014. We therefore conclude that the January 2015 premium cut had no effect on FHA's first-time buyer share in 2015.

[57] These reports are posted at https://www.hud.gov/program_offices/housing/hsgrroom/fhaprodrpt. The monthly first-time buyer share is reported directly starting in April 2013, and we back out the monthly shares for June 2012 to March 2013 from the reported monthly year-to-date totals for the previous fiscal year.

Appendix G: Characteristics and Location of Treatment-Group Versus Control-Group Tracts

In this appendix, we document that homebuyers in the FHA-rich treatment-group Census tracts have very different characteristics than buyers in the control-group tracts. In addition, the two sets of tracts tend to be geographically separated from each other.

As shown in Table G.1, homebuyers in the treatment tracts have mean income that is only half that of buyers in the control tracts. They also have substantially lower credit scores and purchase homes that cost less than half as much as in the control tracts.

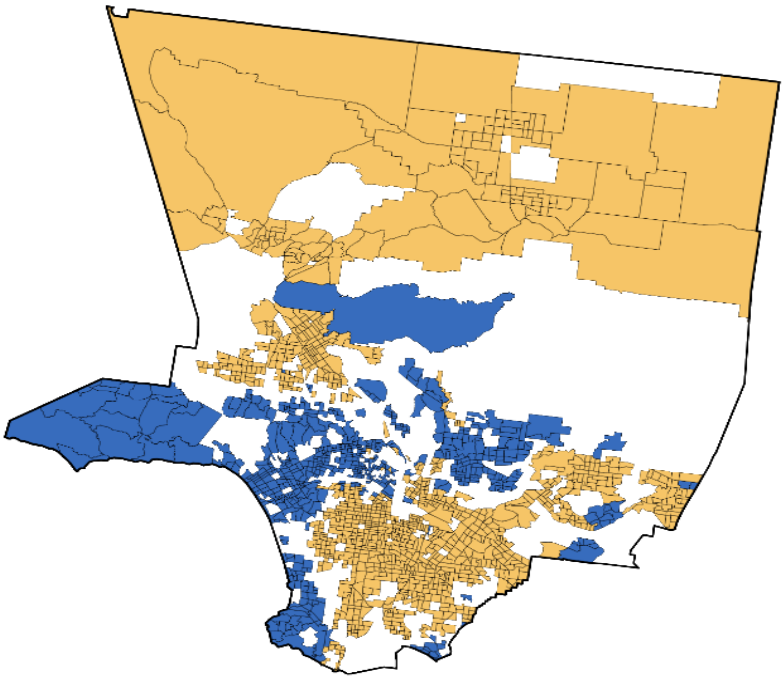
Table G.1. Mean Income, Credit Score, and House Price, Treatment- and Control-Group Tracts
(Based on dataset of home sales used in the paper)

	Treatment-group tracts	Control-group tracts
Mean income	$66,200	$133,700
Mean credit score	683	757
Mean house price	$213,100	$453,000

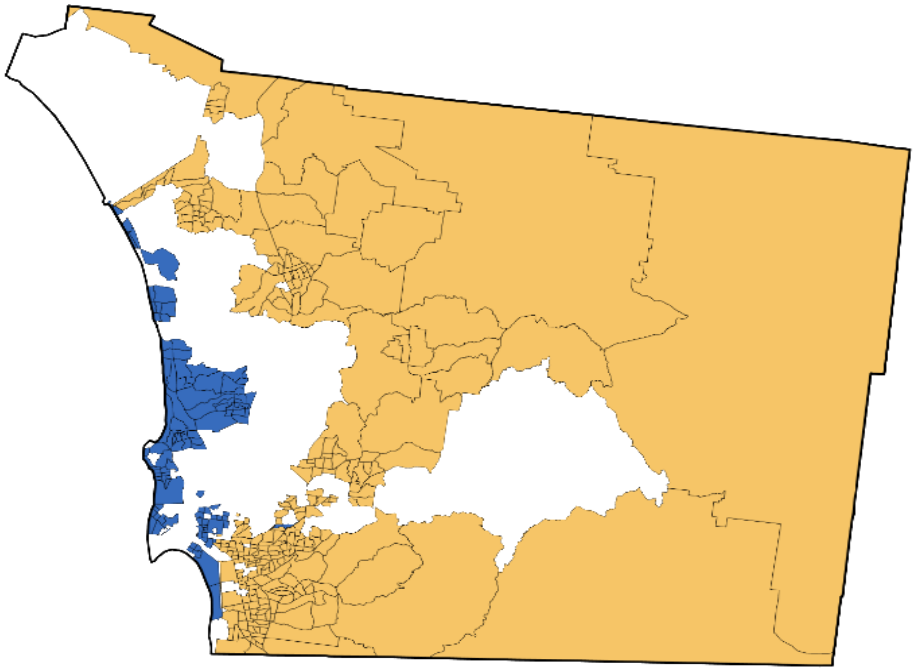
The maps below shows the location of control-group and treatment-group tracts in four counties in our dataset that have a large number of both types of tracts. These maps show that the tracts are geographically separated. The maps for the other counties in the dataset are similar.

Figure G.1. Maps of Treatment-Group and Control-Group Census Tracts
(Color key: orange – treatment-group tracts; blue – control-group tracts; white – all other tracts)

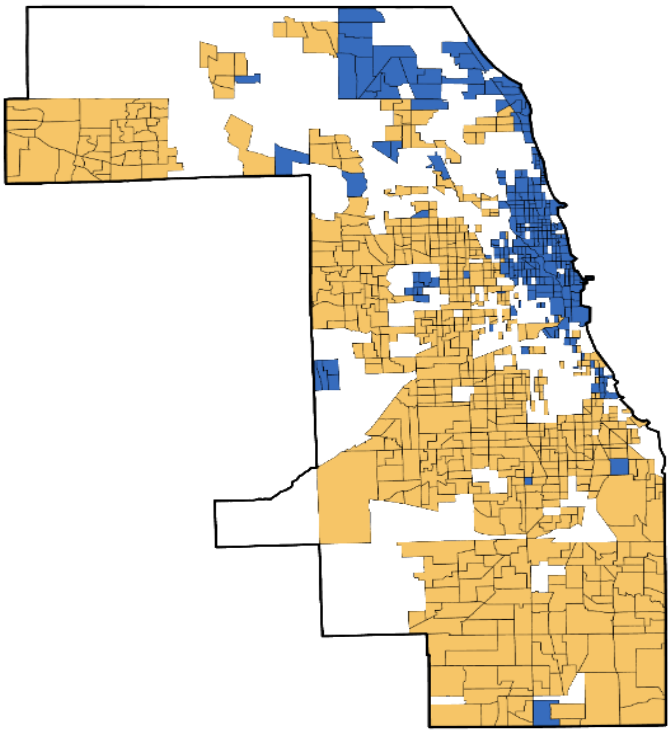
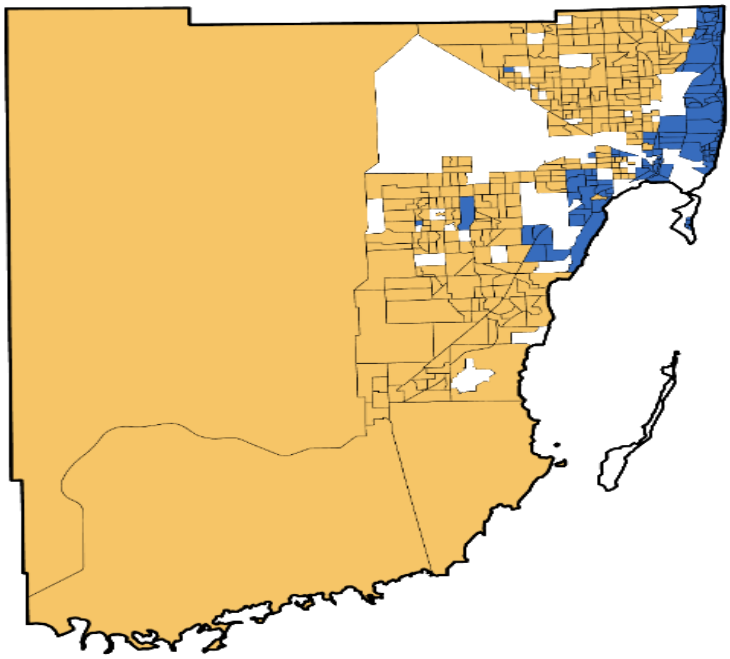
Los Angeles County, CA



San Diego County, CA



Cook County, IL



Miami-Dade County, FL