The Impact of Interest Rates on House Prices and Housing Demand: 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 2017

**Abstract**

Federal housing policy actively promotes homeownership by subsidizing the cost of home mortgage debt and by guaranteeing many mortgages to households with low down payments or low credit scores. In this paper, we exploit a quasi-experiment – the Federal Housing Administration’s (FHA’s) surprise 50 basis point cut to its mortgage insurance premium in January 2015 – to provide evidence on the impact of government policy on housing demand for a population of households likely to be influenced by policy. The premium cut increased the purchasing power of the median FHA borrower by about 6 percent. Our analysis suggests that households with an FHA mortgage increased the value of the housing they purchased by about 2.5 percentage points relative to a control group of homeowners using financing from Fannie Mae and Freddie Mac, for which policy did not change. After translating the premium cut into the equivalent drop in the mortgage rate, our estimates imply a semi-elasticity of the value of housing purchased to the mortgage rate, conditional on purchasing a home (the intensive margin), of roughly 3.4. 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 show that the premium cut induced about 35,000 households to become first-time homebuyers in the initial year after the cut. This increase, which fell far short of the FHA’s expectation, implies a semi-elasticity of homeownership to interest rates over one year (the extensive margin) of 8.9. Finally, we show that the premium cut benefitted FHA borrowers and home sellers at the expense of home buyers who did not use FHA financing. As a group, the non-FHA buyers paid a hefty tax for each of the 35,000 new FHA buyers.

We thank ATTOM Data Solutions for providing the data used in this paper. We received helpful comments from Chris Herbert, Kevin Park, and other participants at the 2017 Conference on Housing Affordability jointly sponsored by the American Enterprise Institute, the Board of Governors of the Federal Reserve System, the Bank of Israel, Tel Aviv University, and UCLA. The views expressed herein are ours alone and do not represent those of any other individuals or the institutions with which we are affiliated.
1. Introduction

A primary goal of federal housing policy is to encourage homeownership and investment in housing by subsidizing interest costs for homeowners and by making mortgages available to households with low credit scores or low savings. 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 rate of homeownership 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 the impact of exogenous changes in interest rates on housing demand and housing tenure decisions for a key segment of the population that federal housing policy directly targets with its homeownership policy. On January 7, 2015, the Federal Housing Administration (FHA) reduced the 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 purchasing power for those getting an FHA mortgage by about 6 percent.\(^1\) This announcement was a surprise to market participants and was not accompanied by pricing changes at the other major federal agencies that guarantee mortgages, i.e. the government-sponsored enterprises Fannie Mae and Freddie Mac (GSEs) and the Department of Veteran's Affairs (VA). Among the federal guarantee programs, the FHA traditionally has focused the most on lower-income borrowers with relatively weak credit profiles. For this reason, this experiment is directly informative about the impact of mortgage subsidies on the demand for owner-occupied housing for the segment of the population most likely to be on the edge of homeownership.

We use a differences-in-differences (diff-in-diff) approach to analyze the impact of the January 2015 rule change. Even though the premium cut was a surprise that affected a subset of market participants, establishing a compelling diff-in-diff analysis is not straightforward for two reasons. First, the population of FHA homebuyers before the premium cut could have been different than the population of homebuyers after the cut. We account for this by adding controls for a broad set of borrower characteristics in our baseline regressions and our robustness checks. Second, and perhaps more importantly, the FHA policy change indirectly affected other homebuyers via equilibrium effects. We document that the change in policy boosted constant-quality house prices in census tracts with a relatively high share of FHA mortgages; this change in prices reduced the purchasing power of other borrowers in those same areas that did not experience a premium cut. To cleanse our analysis of these equilibrium spillover effects, we carefully define our treatment and control populations. Our treatment group consists of FHA borrowers who live in census tracts where FHA mortgages account for a relatively high proportion of all home purchase mortgages; our control group consists of borrowers who receive a GSE mortgage and who live in census tracts where FHA mortgages are infrequent, accounting for less than 5 percent of all mortgages.

Our dataset contains information on close to the universe of FHA and GSE mortgages originated in the 23 counties included in our data set. The impact of the FHA premium cut is stark enough to see with a simple graph. Figure 1 shows the mean sale price for homes purchased with FHA and GSE financing in 2014 and 2015. Over the year after FHA’s premium cut, the mean sale price of FHA-financed homes

---

\(^1\) 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.
jumped about 6 percent, while the mean sale price for GSE-financed homes, for which policy did not change, edged lower. After a formal statistical analysis that controls for many factors that affect house prices, we show that after the premium cut the prices of FHA-financed homes accelerated about 2.5 percentage points vis-à-vis GSE-financed homes. After translating the premium cut into the equivalent drop in the mortgage rate, this implies a semi-elasticity of housing demand to the mortgage rate of roughly 3.4 among FHA homebuyers, a pool of largely lower- and moderate-income borrowers with relatively weak credit profiles.\(^2\) This figure, which represents the semi-elasticity along the intensive margin of housing demand, provides a calibration target for macroeconomists looking to anchor the interest-rate elasticity of housing demand of existing homebuyers (Chambers et al. (2009), Corbae and Quintin (2015), Davis and van Nieuwerburgh (2015), Favilukis et al. (2017), Greenwald (2017), and others). We show that price pressure from stronger demand fully accounted for the 2.5 percent increase; we find no evidence that the increase arose from FHA buyers opting for homes with more amenities or a better location. To our knowledge, our analysis is the first in the literature to cleanly decompose a policy-induced change in home values into the portions reflecting constant-quality prices and home quality.

![Figure 1. Mean House Price, as a Percent of 2014 Mean: FHA vs GSE](image)

Note: Mean house prices are calculated from the cleaned dataset used for the regression analysis in the paper. See section 2 for details on the dataset.
Source: Authors' calculations using data from ATTOM Data Solutions.

We conclude our paper with an evaluation of the effectiveness of the policy change in boosting homeownership, the extensive margin of housing demand. This expansion was one of the stated purposes of the premium cut. At the time the premium cut was announced, the FHA stated that “these lower premiums will save more than two million FHA homeowners an average of $900 annually and spur

\(^2\) To obtain this estimate of the semi-elasticity, note that 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).
250,000 new homebuyers to purchase their first home over the next three years."\(^3\) Although the number of FHA first-time borrowers increased about 180,000 in the first year after the premium cut, we estimate that the premium cut incentivized only about 35,000 new buyers to enter the market during the first year. The rest of the 180,000 increase 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 existing trend growth in the market, unrelated to the premium cut. The estimate of 35,000 new buyers implies a one-year semi-elasticity of homeownership to interest rates on the extensive margin of 8.9.\(^4\)

These results imply that the premium cut had important distributional consequences. In addition to the direct benefit for FHA borrowers, home sellers 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. We estimate that non-FHA buyers as a group paid a hefty tax—nearly $200,000—for each of the 35,000 new FHA buyers. These distributional results highlight that the premium cut created winners and losers. Focusing only on the potential benefits, as the FHA did when announcing the program, tells only part of the story.

The impact of federal policy on house prices and quantities has been studied extensively; for two recent surveys see Davis and van Nieuwerburgh (2015) and Piazessi and Schneider (2016). Work continues on the relationship of interest rates on Treasuries (Favilukis et al. 2017), monetary policy shocks (Williams 2015), and tax policy (Sommer and Sullivan 2017) to house prices. These studies and most others focus on house prices and quantities at the national or metropolitan level, and many rely on the predictions of calibrated general-equilibrium models.

Besides ours, there are only a few other papers 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 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—and the exact estimate depends on assumptions of the jumbo-conforming spread 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 mortgage interest rates reduces the size of the first-lien mortgage taken out (not house prices or quantities) by 2 to 3 percent, less than our estimate along the intensive margin of roughly 3.4.\(^5\)\(^6\) Note that both Adelino et. al. (2014) and DeFusco and Paciorek (2017) study the behavior of borrowers that can qualify for a


\(^4\) These 35,000 additional new FHA buyers in 2015 represent about 6.5 percent of the 539,000 first-time FHA homebuyers we would have expected to see in 2015 had the FHA not cut its premium. Our estimate of 539,000 expected buyers in 2015 is computed as 660,000 new first-time FHA buyers, less our estimate of 86,000 FHA first-time buyers that would likely have been GSE borrowers had the premium cut not occurred (discussed later), less our estimate of 35,000 new first-time FHA buyers induced by the premium cut. Dividing 6.5 percent by the implied 0.73 percentage point decline in the mortgage rate yields a semi-elasticity of 8.9.

\(^5\) 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.

\(^6\) When we re-run our analysis using mortgage debt instead of house values as the dependent variable, our results do not change. This is likely because most FHA borrowers have a loan-to-value ratio at the maximum allowed by the program.
conventional mortgage. The FHA borrowers that we study generally make much smaller down payments and have lower credit scores than borrowers taking out conventional loans. Given the tighter financial constraints facing FHA borrowers, it makes sense that they would be more responsive to federal programs seeking to boost housing demand.

Our analysis complements that of Bhutta and Ringo (2017a) 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, which is known in the literature as the extensive margin of housing demand. Their results on this question appear to be roughly consistent with ours after adjusting for the fact that they measure the effect on all home purchases while we measure the effect on first-time buyers alone. They 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). With very tight inventories of homes for sale, the boost to demand from the premium cut resulted in a substantial rise in home prices and an erosion of purchasing power for buyers taking out conventional loans. Had the same policy change been implemented in a buyer’s market with more plentiful supply, the price increase likely would have been smaller or even non-existent and the number of new homebuyers likely would have been greater than we found.

The rest of the paper is organized as follows. The next section describes the datasets used in the analysis. Section 3 describes how we assess the effect of the premium cut on house prices, while Section 4 presents the results of this assessment and evaluates whether the premium cut achieved FHA’s policy goals. Section 5 concludes.

2. Data

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.

---

7 Additionally, Adelino et. al. (2014) use information about buyers in a tight price range around the conforming loan limit and and DeFusco and Paciorek (2017) limit their analysis to metropolitan areas in California. In our estimates, we use from the entire distribution of FHA mortgages for 23 counties across the United States.

8 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. Conversely, a buyer’s market exists when the inventory of existing homes for sale exceeds six months 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. We obtained data from Zillow for 20 of the 23 counties in our sample, confirming that market conditions were very tight in these counties.

9 Ernest Fisher, the FHA’s first chief economist, long ago recognized the importance of accounting for market conditions. During the buyer’s market that dominated the 1930s, Fisher (1951, pp. 78-9) 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.” In a seller’s market, however, Fisher concluded that liberalization of credit primarily drives up prices. Fisher’s analysis thus foreshadows the issues explored in this paper.
Importantly, the mortgage information includes the loan purpose (which allows us to separate home purchase loans from refinance loans), the loan amount, the exact date of the transaction, and the loan type (whether the loan is FHA, VA, a construction loan, or in a large residual category that includes conventional loans and loans guaranteed by the Rural Housing Service (RHS)). Finally, the data also include an estimated home value for each property, based on ATTOM’s Automated Valuation Model (AVM).

Our data cover 23 counties across 12 states over the period 2013-2015. The counties were selected because of their large FHA loan counts. They also have some of the nation’s largest conventional loan totals by county. The 23 counties contain about 17 percent of all FHA loans and about 13 percent of all conventional loans.

Table 1 presents basic information about the 23 counties in the dataset. 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. Similarly, the percentage of adults with a bachelor’s degree in these counties is centered at about the national average. However, some of the counties, such as Prince George’s County, MD, have low median incomes and low educational attainment relative to the other jurisdictions in their metropolitan area. 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.

To check the coverage of the ATTOM dataset, we 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. Because HMDA does not identify one-unit properties, the comparison covers owner-occupied properties with one to four units.

---

10 Conventional loans consist of those with a GSE guarantee and those held in the private sector with no guarantee. As discussed below, our regression analysis uses only the GSE loans in the broad residual group in the ATTOM data, omitting the private-sector loans and RHS loans. We remove these loans through a matching procedure to other datasets.

11 These data were provided in three distinct files. We merged them based on the property identifier (the SA_PROPERTY_ID field), dropping all unmatched pairs. More than 99 percent of the observations were successfully merged.

12 Counties in Texas and Utah with large FHA loan counts are excluded because the sale price is not disclosed in those states.

13 For more information on the HMDA data, see http://www.consumerfinance.gov/data-research/hmda/learn-more.

14 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.
Table 1. Population and Demographic Characteristics in 2015

<table>
<thead>
<tr>
<th>County, State</th>
<th>Population</th>
<th>Median Household Income</th>
<th>Percent Bachelor’s Degree or Higher</th>
<th>Percent Black or Hispanic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Broward, FL</td>
<td>1,896,425</td>
<td>$53,926</td>
<td>32%</td>
<td>56%</td>
</tr>
<tr>
<td>Clark, NV</td>
<td>2,114,801</td>
<td>$51,552</td>
<td>23%</td>
<td>41%</td>
</tr>
<tr>
<td>Cook, IL</td>
<td>5,238,216</td>
<td>$56,851</td>
<td>37%</td>
<td>49%</td>
</tr>
<tr>
<td>Duval, FL</td>
<td>913,010</td>
<td>$49,554</td>
<td>29%</td>
<td>38%</td>
</tr>
<tr>
<td>El Paso, CO</td>
<td>674,471</td>
<td>$60,109</td>
<td>36%</td>
<td>23%</td>
</tr>
<tr>
<td>Franklin, OH</td>
<td>1,251,722</td>
<td>$53,882</td>
<td>39%</td>
<td>27%</td>
</tr>
<tr>
<td>Gwinnett, GA</td>
<td>895,823</td>
<td>$61,732</td>
<td>35%</td>
<td>46%</td>
</tr>
<tr>
<td>Hillsborough, FL</td>
<td>1,349,050</td>
<td>$51,725</td>
<td>33%</td>
<td>43%</td>
</tr>
<tr>
<td>Kern, CA</td>
<td>882,176</td>
<td>$51,342</td>
<td>16%</td>
<td>58%</td>
</tr>
<tr>
<td>Los Angeles, CA</td>
<td>10,170,292</td>
<td>$59,134</td>
<td>31%</td>
<td>56%</td>
</tr>
<tr>
<td>Macomb, MI</td>
<td>864,840</td>
<td>$54,640</td>
<td>24%</td>
<td>14%</td>
</tr>
<tr>
<td>Maricopa, AZ</td>
<td>4,167,947</td>
<td>$56,004</td>
<td>31%</td>
<td>36%</td>
</tr>
<tr>
<td>Miami-Dade, FL</td>
<td>2,693,117</td>
<td>$43,786</td>
<td>27%</td>
<td>84%</td>
</tr>
<tr>
<td>Orange, FL</td>
<td>1,288,126</td>
<td>$50,720</td>
<td>32%</td>
<td>50%</td>
</tr>
<tr>
<td>Pierce, WA</td>
<td>843,954</td>
<td>$60,167</td>
<td>26%</td>
<td>17%</td>
</tr>
<tr>
<td>Pima, AZ</td>
<td>1,010,025</td>
<td>$47,099</td>
<td>31%</td>
<td>40%</td>
</tr>
<tr>
<td>Prince George's, MD</td>
<td>909,535</td>
<td>$76,741</td>
<td>32%</td>
<td>79%</td>
</tr>
<tr>
<td>Prince William, VA</td>
<td>451,721</td>
<td>$99,766</td>
<td>40%</td>
<td>42%</td>
</tr>
<tr>
<td>Riverside, CA</td>
<td>2,361,026</td>
<td>$58,292</td>
<td>21%</td>
<td>54%</td>
</tr>
<tr>
<td>Sacramento, CA</td>
<td>1,501,335</td>
<td>$58,942</td>
<td>30%</td>
<td>32%</td>
</tr>
<tr>
<td>San Bernardino, CA</td>
<td>2,128,133</td>
<td>$53,803</td>
<td>19%</td>
<td>60%</td>
</tr>
<tr>
<td>San Diego, CA</td>
<td>3,299,521</td>
<td>$67,320</td>
<td>37%</td>
<td>38%</td>
</tr>
<tr>
<td>Wayne, MI</td>
<td>1,759,335</td>
<td>$41,557</td>
<td>23%</td>
<td>45%</td>
</tr>
<tr>
<td><strong>Memo: U.S. total</strong></td>
<td><strong>321,418,821</strong></td>
<td><strong>$55,775</strong></td>
<td><strong>31%</strong></td>
<td><strong>30%</strong></td>
</tr>
</tbody>
</table>

Note: 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. Sources: 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).

Table 2 shows the coverage for the ATTOM dataset in the aggregate and for each county compared to HMDA. 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.\(^{15}\) 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 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

\(^{15}\) We include RHS loans in the HMDA count because, as noted above, they can’t be separated from conventional loans in the ATTOM data.
Table 2. ATTOM Coverage Relative to HMDA, 2013-2015

<table>
<thead>
<tr>
<th>County, State</th>
<th>FHA</th>
<th>Conventional &amp; RHS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Broward, FL</td>
<td>102%</td>
<td>111%</td>
</tr>
<tr>
<td>Clark, NV</td>
<td>94%</td>
<td>102%</td>
</tr>
<tr>
<td>Cook, IL</td>
<td>93%</td>
<td>89%</td>
</tr>
<tr>
<td>Duval, FL</td>
<td>90%</td>
<td>91%</td>
</tr>
<tr>
<td>El Paso, CO</td>
<td>102%</td>
<td>99%</td>
</tr>
<tr>
<td>Franklin, OH</td>
<td>94%</td>
<td>73%</td>
</tr>
<tr>
<td>Gwinnett, GA</td>
<td>91%</td>
<td>99%</td>
</tr>
<tr>
<td>Hillsborough, FL</td>
<td>90%</td>
<td>95%</td>
</tr>
<tr>
<td>Kern, CA</td>
<td>96%</td>
<td>112%</td>
</tr>
<tr>
<td>Los Angeles, CA</td>
<td>99%</td>
<td>98%</td>
</tr>
<tr>
<td>Macomb, MI</td>
<td>94%</td>
<td>95%</td>
</tr>
<tr>
<td>Maricopa, AZ</td>
<td>99%</td>
<td>107%</td>
</tr>
<tr>
<td>Miami-Dade, FL</td>
<td>109%</td>
<td>117%</td>
</tr>
<tr>
<td>Orange, FL</td>
<td>94%</td>
<td>94%</td>
</tr>
<tr>
<td>Pierce, WA</td>
<td>98%</td>
<td>103%</td>
</tr>
<tr>
<td>Pima, AZ</td>
<td>91%</td>
<td>102%</td>
</tr>
<tr>
<td>Prince George's, MD</td>
<td>95%</td>
<td>102%</td>
</tr>
<tr>
<td>Prince William, VA</td>
<td>95%</td>
<td>90%</td>
</tr>
<tr>
<td>Riverside, CA</td>
<td>98%</td>
<td>105%</td>
</tr>
<tr>
<td>Sacramento, CA</td>
<td>101%</td>
<td>105%</td>
</tr>
<tr>
<td>San Bernardino, CA</td>
<td>99%</td>
<td>97%</td>
</tr>
<tr>
<td>San Diego, CA</td>
<td>95%</td>
<td>93%</td>
</tr>
<tr>
<td>Wayne, MI</td>
<td>90%</td>
<td>104%</td>
</tr>
<tr>
<td><strong>All 23 counties</strong></td>
<td>97%</td>
<td>99%</td>
</tr>
</tbody>
</table>

Note: Comparison is for first-lien, primary owner-occupied purchase loans 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.

owner-occupant for the most recent sale. We have no information about occupancy status for earlier sales, and assumed that the earlier sale(s) were also primary owner-occupied properties, which will overstate the ATTOM loan count for the purpose of matching to HMDA. 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 that contribute to the coverage ratios above 100 percent. Overall, however, the ATTOM dataset conforms well to the loan counts in HMDA.

We restrict our sample from the ATTOM dataset to first-lien, primary owner-occupied purchase loans for one-unit homes sold between 2013:Q1 and 2015:Q4 with FHA or conventional/RHS financing. To create this sample, 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. Thus, the top end of the price distribution for FHA-financed sales in 2014 is truncated relative to the 2013 distribution, which imparts a downward bias to the price change between the two years. 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), which biases up the price change from 2014 to 2015 in those counties. To avoid these biases, our sample excludes any loan with an amount that is larger than the county-level 2014 FHA loan limits. This ceiling eliminates the high-priced FHA sales in 2013 and 2015 that exceed the 2014 limit. We also impose the same county-by-county limits on the sales with conventional financing to focus on the part of the market in which FHA operates.

We then screen out loans with missing data as well as outliers. Among these screens, we remove loans without an AVM or a sale price, properties with a loan-to-value ratio (LTV) above 110 percent, and loans for which the property’s AVM was in the top and bottom 1 percent for their county and sale month (this AVM trim is done separately for FHA and conventional loans). We also remove distressed sales because their sale price may not reflect market norms. These screens remove roughly 24 percent of all loans, mostly due to the elimination of distressed sales. See Appendix A for details on the data cleaning and trimming.

As a final step, we match the resulting dataset to external datasets to add information about borrower characteristics. The first round of matches is to loan-level data from HMDA and the Federal Housing Finance Agency (FHFA) for 2013, 2014, and 2015 to add information on borrower income and other borrower characteristics and to screen out any RHS loans included in the ATTOM data.

Although HMDA captures the vast majority of home mortgage loans, the FHFA dataset includes some additional loans guaranteed by the GSEs. The HMDA and FHFA data both report the borrower’s gross annual income. 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 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. A higher match rate could be achieved if the origination month, not just the year, were included in the publicly available HMDA and FHFA data files, but that information is not disclosed.

The second round matches the FHA loans in the resulting ATTOM/HMDA/FHFA-matched dataset to FHA’s Single-Family Portfolio Snapshot. This step allows us to add an FHA borrower’s note rate to the dataset, which we use in the final round of matching. 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.

---

16 Imposing the 2014 limit results in some bunching of 2014 and 2015 FHA loans at the limit. These borrowers would have likely taken out greater loan amounts were they not constraint by the loan limit. To eliminate the influence of these constrained borrowers, we drop FHA loans bunched at the 2014 limit or within 0.75 percent of the limit.
17 For more on the FHFA dataset, see http://www.fhfa.gov/DataTools/Downloads/Pages/Single-Family-Census-Tract-File.aspx
18 The FHA dataset contains loan-level information on FHA’s single-family endorsements, including loan purpose, property type, loan amount, zip code, lender and sponsor information, interest rate, and other variables. For more information, see https://www.hud.gov/program_offices/housing/rmra/oe/rpts/sfsnap/sfsnap.
The third and final round of matches is to loan-level data from the American Enterprise Institute’s (AEI’s) National Mortgage Risk Index (NMRI), discussed in detail later, and to loan-level data released by the GSEs. While the NMRI is a near-census of government-guaranteed loans, its geographical detail is limited to the state level. The GSE datasets contain the large majority of loans purchased by Fannie Mae or Freddie Mac at greater geographical detail (3-digit zip level), which enables additional matches. Because the NMRI and GSE datasets only contain loans with a government guarantee, this step eliminates private lenders from our sample.

The third-round match picks up the borrower’s credit score and, for nearly all loans, debt-to-income ratio. The match is performed using various combinations of loan type, loan amount, state or either 3-digit or 5-digit zip code, origination date, lender information, LTV, the number of borrowers, and the property type. In the case of conventional loans, 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. 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. The premium cut enabled FHA to attract borrowers who likely would have otherwise taken out a conventional mortgage, as we show in section 4. Having data on borrower characteristics such as credit score, income, and race and ethnicity allows us to effectively control for changes in the distributions of FHA and GSE borrowers.

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. ATTOM’s AVMs are proprietary data, and ATTOM does not disclose the data or assumptions underlying each model. Parts of our analysis hinge on the accuracy of the AVM. Although we cannot assess ATTOM’s AVM methodology, we can check the accuracy of ATTOM’s AVMs against subsequent sale data, as described below.

We use ATTOM’s AVM for December 2014 because it represents the most recent estimate of property value that would be unaffected by the FHA premium cut. December 2014 was the last full month before the premium cut was announced and implemented. Moreover, as Bhutta and Ringo (2017) demonstrate, the cut was unanticipated, so home sales during or before December 2014 would not have built in any effect of the cut. In subsequent months, sale prices could have been boosted by the premium cut, and these higher prices would have been reflected in higher AVMs. Therefore, we use the December 2014 AVM as our pre-event “stake in ground” for home valuation. Importantly, this stake-in-ground approach can be applied to any date in the range of our data.

---

19 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.

20 In an intermediate step, we match the FHA data onto the NMRI, which adds the zip code for the matched loans to the NMRI dataset.

21 Previous research shows that FHA competes to some degree with conventional lending activity. See, for example, Bhutta and Ringo (2016), Nothaft (2014), and Ambrose, Pennington-Cross, and Yezer (2002).

22 For more information, see the DataQuick CMV-Portfolio White Paper (http://textlab.io/doc/169451/dataquick-cmv-portfolio%E2%84%A2-white-paper). DataQuick is a subsidiary of ATTOM Data Solutions.
the-ground is a real-time AVM that does not include any information revealed in later months, a fact that we confirmed directly with ATTOM.

We assess the accuracy of the December 2014 AVMs by comparing the AVM value for properties that sold in that month to the reported sale price. 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.

For the roughly 9,400 homes in our final cleaned dataset that sold in December 2014, the histogram in Figure 3 displays the ratio of the home’s sales price to its December 2014 AVM value. As shown, most of the sale prices lie within a fairly narrow band around the AVM. On average, the sale price equaled 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 sufficiently accurate reflection of property value to be useful in the analysis that follows.

We use the December 2014 AVM values for two separate parts of our research. In the first part, they serve as a measure of house quality that is unaffected by the premium cut. By including this AVM-based control for quality in a standard house-price regression, the estimated coefficients on the monthly
dummy variables trace out the effect of the premium cut on constant-quality home prices. We compare this price effect for FHA-financed transactions before and after the premium cut to that for the control group of GSE loans, which had no change in policy. In the second part of our analysis, we use the AVM values to capture changes in the quality of the cohort of homes sold each month. To illustrate, if the homes sold in April 2015 had higher December 2014 AVM values than did the homes sold in March 2015, this would indicate that sales had shifted toward higher quality homes between March and April. We use the December 2014 AVMs in this way to track quality change for FHA-financed transactions relative to the control group of GSE loans before and after the FHA premium cut.

National Mortgage Risk Index

For the analysis in section 4, we use the NMRI data mentioned above. 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 over 14 million home purchase loans dating back to September 2012 with a coverage rate of 99.5 percent.23

Two features of the NMRI data should be noted. First, the source data for the NMRI report 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; we have found this rule to be an extremely accurate estimate of the origination month. Second, the NMRI dataset includes a first-time homebuyer flag provided by the agencies themselves, which we use to estimate the extent of FHA’s success in drawing in first-time buyers after its premium cut.

3. Empirical Design

This section describes our empirical framework for measuring the impact of FHA’s premium cut on house prices and house quality.

Treatment and Control Groups

Figure 4 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” of 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 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, the control group excludes GSE borrowers in census tracts with a relatively high share of FHA loans, the orange bar marked “A,” as these borrowers are likely to have been indirectly affected by the FHA premium cut. Similarly, the treatment group omits the very small number of FHA borrowers in census tracts with a low FHA share, the blue bar marked “C”. 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.

**Figure 4. Schematic for Treatment and Control Groups**

Figure 5 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). The left-most orange bar in the 23-county panel represents the control group in our baseline regressions, i.e., conventional loans in census tracts with an FHA share of less than 5 percent. In those tracts, conventional, VA, and RHS loans outnumber FHA loans by a factor of 70 to 1, indicating that any effect of the premium cut should be nil, as required for this to be a suitable control group. The other notable feature of Figure 5 is that 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 outcome reflects our experimental design. We chose the 23 counties because they had a large volume of FHA lending, which enhances the statistical power of the analysis. To generalize the results to the entire county, we simply map the results obtained with different treatment-group thresholds to the fraction of loans nationally in the relevant FHA-share buckets.

---

Note that the bars in this figure include all 2015 HMDA purchase loans with loan amounts below the 2014 FHA loan limits without imposing the restrictions we put in place to generate our estimation sample.
Figure 5. 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 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)

Overall Effect

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

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

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 Broward County, FL 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 4) and GSE loans in the control group (orange bar “D”). We weight the loan data so that each combination of origination year, 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.

The county dummy variables account for differences in price levels across the counties, and \(X\) controls for the composition of FHA and GSE borrowers. 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 progressively longer treatment periods. The varying treatment periods allow us, first, to measure how quickly any treatment effects become apparent, and second, to account for possible spillovers from FHA sales to those with GSE financing that might damp the effects as time passes.\textsuperscript{25}

We consider three alternative endpoints for the treatment period: 2015:Q2, 2015:Q3, and 2015:Q4. Table 3 summarizes the difference-in-difference calculations for the three alternative endpoints. As shown, we use the same set of quarters in both the pre-event and post-event periods 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 and is not used in our analysis.

<table>
<thead>
<tr>
<th>End of Treatment Period</th>
<th>Diff-in-Diff Calculation: Pre-event period vs. post-event period</th>
</tr>
</thead>
</table>

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”).

**Constant-quality Price Effect**

To capture the pure 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, capturing the effect of many individual characteristics. Adding the AVM turns equation 1 into a parsimonious hedonic regression. In this augmented version of equation 1, the price changes calculated from the quarterly dummies can be interpreted as hedonic price indexes that measure the average change in sale prices over time after controlling for house quality. We use these indices to estimate the degree to which the additional demand induced by the premium cut boosted the prices of homes that people bought, holding quality constant.

**Quality Effect**

Finally, we estimate the effect of the premium cut on the quality of homes sold with the following regression:

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

where \(AVM\) represents the home’s December 2014 AVM value. Equation 2 is identical to equation 1 other than substituting the AVM for the sale price as the dependent variable. As described above, the AVM value serves as a proxy for house quality; we use the December 2014 value because it is a pre-event valuation that will not impound any price effects from the premium cut. Parallel to equation 1,

\textsuperscript{25} For example, the price effects potentially could shrink over time as FHA transactions get used as “comps” for appraisals throughout the market or by sellers when deciding the listing price of their home.
we estimate equation 2 as an OLS regression and calculate clustered standard errors. We measure the
treatment effect of the premium cut through the same diff-in-diff calculation as above.

**Is Selection an Issue?**

The potential exists for selection to affect our results, as the pools of borrowers in the control and
treatment groups are not fixed over time. To explore this issue, note that – roughly 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 generally opt for GSE or other
conventional financing. The second group consists of borrowers with weaker credit profiles and less
money for a down payment; these borrowers typically take out an FHA loan.

Concerns about selection focus 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 because of the reduced premium. Later in the paper, we show that the FHA
premium cut induced some borrowers with loan-to-value ratios between 90 and 95 percent to choose
an FHA loan instead of GSE financing. For selection to affect our results, these “switchers” 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 important
borrower characteristics reduces the risk that selection would bias our results. In addition, as a
robustness check, we remove the borrowers who are most likely to be “switchers” and find only minor
changes from our baseline results.

**4. Results**

**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 addition of the December 2014 AVM value as a 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. In our baseline regressions, the control
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. Later on, we show results with different definitions of the
treatment and control groups.

Figure 5 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, Figure 5 shows that constant-quality prices for FHA-financed homes increased at
essentially the same rate as GSE-financed homes. After the premium cut, a significant gap emerged in
the growth rates, with the constant-quality price of homes purchased by FHA borrowers rising relative
to the price of homes purchased by borrowers using GSE loans. Table 4 shows the diff-in-diff results for
various treatment periods. Focusing on the 2015:Q4 row, the constant-quality price of homes backed by
FHA mortgages increased 0.6 percentage point less rapidly from 2014:Q4 to 2015:Q4 than from 2013:Q4
to 2014:Q4; for GSE mortgages, the growth rate of prices from 2014:Q4 to 2015:Q4 slowed much more –
by 3.6 percentage points – relative to the prior four quarters. Thus, for the treatment period ending in

---

26 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. These coefficients are
highly significant in all cases except for the credit score on FHA loans in the constant-quality price regression.
Borrower race tends to be significant, while the other demographic characteristics are less often significant. The
full set of estimated coefficients from equation 1, along with all other regression results, is available on request.
2015:Q4, we estimate the premium cut boosted constant-quality prices of FHA-financed homes by 3.0 percentage points more than GSE financed homes. Averaging over the three treatment periods we consider, we estimate the premium cut induced a statistically significant 2.8 percentage point increase in constant-quality prices for FHA-financed homes.

Table 4. Constant-quality Price Effect

<table>
<thead>
<tr>
<th>End of Treatment Period</th>
<th>Diff FHA</th>
<th>Diff GSE</th>
<th>Diff-in-Diff</th>
</tr>
</thead>
<tbody>
<tr>
<td>2015:Q2</td>
<td>-0.5 ppt</td>
<td>-2.7 ppts**</td>
<td>2.2 ppts**</td>
</tr>
<tr>
<td>2015:Q3</td>
<td>-0.3 ppt</td>
<td>-3.5 ppts**</td>
<td>3.3 ppts**</td>
</tr>
<tr>
<td>2015:Q4</td>
<td>-0.6 ppt</td>
<td>-3.6 ppts**</td>
<td>3.0 ppts**</td>
</tr>
<tr>
<td>Average over treatment periods</td>
<td>-0.4 ppt</td>
<td>-3.3 ppts**</td>
<td>2.8 ppts**</td>
</tr>
</tbody>
</table>

* and ** denotes significance at the 10 percent and 5 percent levels, respectively.

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. As noted above, the dependent variable in the regression — the December 2014 AVM value — serves as a summary measure of quality that is unaffected by the price changes induced by the premium cut. Barring changes in the mix of homes sold each month, the quality of homes purchased over time would remain constant, which would show up in the regression as a flat time path for the quarterly dummies and the FHA* Q interaction dummies.

Figure 6, however, shows a decline in the quality of the homes purchased for both groups of borrowers over the sample period. This decline in quality likely stems from the sharp rise in constant-quality home prices.
prices displayed in Figure 5. 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.

Table 5. Change in AVM: Quality Effect

<table>
<thead>
<tr>
<th>End of Treatment Period</th>
<th>Diff FHA</th>
<th>Diff GSE</th>
<th>Diff-in-Diff</th>
</tr>
</thead>
<tbody>
<tr>
<td>2015:Q2</td>
<td>4.3 ppts**</td>
<td>2.8 ppts**</td>
<td>1.5 ppts</td>
</tr>
<tr>
<td>2015:Q3</td>
<td>3.4 ppts**</td>
<td>3.6 ppts**</td>
<td>-0.1 ppt</td>
</tr>
<tr>
<td>2015:Q4</td>
<td>3.4 ppts**</td>
<td>4.9 ppts**</td>
<td>-1.5 ppts</td>
</tr>
<tr>
<td>Average over treatment periods</td>
<td>3.7 ppts**</td>
<td>3.8 ppts**</td>
<td>-0.1 ppt</td>
</tr>
</tbody>
</table>

* and ** denotes significance at the 10 percent and 5 percent levels, respectively.

We estimate that the FHA premium cut did not change the quality of homes purchased by FHA borrowers. This may seem counterintuitive at first, as the blue line (GSE-financed homes) declines at a more rapid rate than the orange line (FHA-financed homes). However, the diff-in-diff keys off changes in the rate of quality before and after the premium cut. Focusing on the 2015:Q3 row of Table 5, we can see that the growth of quality for FHA-financed homes was 3.4 percentage points greater from 2014:Q4 to 2015:Q3 than it was from 2013:Q4 through 2014:Q3. 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 homes relative to GSE-financed homes in response to the premium cut. For the other treatment periods, the diff-in-diff quality effects are further from zero but are not statistically significant. Averaging over the results over all three treatment periods, we estimate that quality of FHA-financed homes declined 0.1 percentage point relative to GSE mortgages after the premium cut. This
result is not statistically significant, so we cannot reject that the FHA premium induced no change in the quality of homes purchased.

**Overall Effect**

Figure 7 and Table 6 show the same information as the previous two figures and tables for the total price changes, inclusive of 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.\(^\text{27}\) Averaging over the three treatment periods, Table 6 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 those purchased by GSE borrowers. This increase is statistically significant at the 5 percent level. The results vary over the three treatment periods, with significant increases in the prices of FHA-financed homes over the two shorter periods and a smaller insignificant rise over the longest period.

![Figure 7. Price Change: Overall Effect](image)

<table>
<thead>
<tr>
<th>End of Treatment Period</th>
<th>Diff FHA</th>
<th>Diff GSE</th>
<th>Diff-in-Diff</th>
</tr>
</thead>
<tbody>
<tr>
<td>2015:Q2</td>
<td>3.1 ppts **</td>
<td>-0.1 ppt</td>
<td>3.2 ppts **</td>
</tr>
<tr>
<td>2015:Q3</td>
<td>2.6 ppts **</td>
<td>-0.2 ppt</td>
<td>2.9 ppts **</td>
</tr>
<tr>
<td>2015:Q4</td>
<td>2.4 ppts **</td>
<td>1.0 ppt</td>
<td>1.4 ppts</td>
</tr>
<tr>
<td>Average over treatment periods</td>
<td>2.7 ppts **</td>
<td>0.2 ppt</td>
<td>2.5 ppts **</td>
</tr>
</tbody>
</table>

* and ** denotes significance at the 10 percent and 5 percent levels, respectively.

\(^{27}\) The figure shows a seasonal pattern of price changes, with prices tending to rise during the spring buying season and contracting in the fall. This pattern is consistent with other datasets that report seasonally unadjusted data.
Summary of Effects

Table 7 pulls together all the results presented so far. 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 and no significant change in the quality of homes purchased. The overall effect on home price, which captures the impact on constant-quality prices and home quality combined, shows that the premium cut boosted the price of homes purchased by FHA borrowers relative to GSE borrowers by 2.5 percentage points, on average, over the three treatment periods; this overall price effect is significant at the 5 percent level.

Table 7. FHA Price Acceleration in 2015 vis-à-vis GSE Loans

<table>
<thead>
<tr>
<th>End of Treatment Period</th>
<th>Constant-quality Price Effect</th>
<th>Quality Effect</th>
<th>Overall Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>in ppts</td>
<td>in dollars</td>
<td>in ppts</td>
</tr>
<tr>
<td>2015:Q2</td>
<td>2.2 ppts**</td>
<td>$4,500**</td>
<td>1.5 ppts</td>
</tr>
<tr>
<td>2015:Q3</td>
<td>3.3 ppts**</td>
<td>$6,700**</td>
<td>-0.1 ppt</td>
</tr>
<tr>
<td>2015:Q4</td>
<td>3.0 ppts**</td>
<td>$6,300**</td>
<td>-1.5 ppts</td>
</tr>
<tr>
<td>Average over treatment periods</td>
<td>2.8 ppts**</td>
<td>$5,900**</td>
<td>-0.1 ppt</td>
</tr>
</tbody>
</table>

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 mean FHA purchase price of $207,000 in the 23 sample counties. * and ** denotes significance at the 10 percent and 5 percent levels, respectively.

Table 7 converts our estimated effects to dollar terms based on the FHA’s mean 2014 home price of $207,000 in the 23 counties in our dataset. FHA’s premium cut raised the mean sale price of FHA-financed home between $5,100 and $5,800 vis-à-vis homes with GSE financing. The lower amount reflects the average of the overall effect over the three treatment periods (the right-most column), while the upper amount reflects the average of the sum of the pure price effect and the quality effect over the three periods, which is not constrained to sum to the overall effect. This increase reflects the rise in constant-quality prices, as the quality effect is insignificant. We conclude the FHA premium cut raised market prices for FHA buyers without changing the quality of the stock being purchased.

As we noted in the introduction, the overall price effect of 2.5 percentage points (when averaged over the three treatment periods) implies a semi-elasticity of housing demand along the intensive margin to the mortgage rate of about 3.4. This figure is within the very wide range 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 (rather than housing demand per se) 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 replaced 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. This estimate implies a semi-elasticity of about 3.4, slightly greater than DeFusco and Paciorek (2017) of 2 to 3. Our estimate likely is larger because the FHA borrowers that we study face tighter financial constraints than the GSE borrowers in their analysis.

---

28 This figure is calculated from 2014 HMDA data.
Robustness Checks

Figures 8a and 8b 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 8a, the treatment group consists of all borrowers in census tracts with an FHA share greater than what is indicated on the x-axis. For example, the points shown on the graph at the x-axis value of 25 percent are for a treatment group residing in a census tract with an FHA share of at least 25 percent; similarly, the points shown at the x-axis value of 20 percent are for a treatment group residing in a census tract with at least a 20 percent FHA share. Thus, all the data used to generate the estimates at the x-value of 25 percent are included for the estimates at the x-value of 20 percent, plus the data in tracts with FHA shares between 20 and 25 percent. The main takeaway from Figure 8a 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.

Note: Crossed and round markers indicate significance at the 10% and 5% level, respectively.
Figure 8b 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. This figure highlights that the point estimate of the constant-quality price effect is 2 percent or higher as long as the treatment group comes from 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, presumably because the FHA borrowers benefitting from the premium cut are not numerous enough to affect the local market.

Figures 9a and 9b display the same results as 8a and 8b 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 (as compared to 5 percent in the baseline). Qualitatively, the results are the same as in Figures 8a and 8b, with the overall effect of the FHA premium cut due exclusively to constant-quality price effects. The

---

29 The highest FHA-share bucket in Figure 8b (and Figures 9b and 10b below) is 60 to 70 percent because, as shown in Figure 5, very few census tracts have FHA shares above 70 percent.
key difference is that the constant-quality price effect from the FHA premium cut becomes somewhat smaller. It averages about 2 percentage points in Figure 9a with this alternate control group, compared to almost 3 percentage points in Figure 8a. Intuitively, the effect has become smaller because the control group now includes census tracts that are not as insulated from FHA borrowers, which allows greater spillovers to the control group. Thus, we conclude that the control group needs to be essentially free of FHA borrowers – as it is in our baseline – to accurately capture the effects of the premium cut.

Figures 10a and 10b show an additional insight that comes from our experimental design. In these figures, we specify that households in both the control and treatment groups reside in census tracts with a relatively high FHA share. In other words, if the treatment group is given by the blue box B in Figure 4, the control group is given by the orange box A. This combination removes the separation between the treatment and control groups in our baseline set-up, and the results are strikingly different than before. With this mixing of the treatment and control groups, Figures 10a and 10b show that the FHA premium cut had very little effect on constant-quality prices, but rather induced a large increase in home quality and total price paid.

Figures 10a and 10b show the “within-market” effects of the premium cut – that is, the equilibrium outcome from interactions between FHA and GSE borrowers. After the premium cut, purchasing power increased for FHA borrowers. In areas with a relatively high share of FHA loans, FHA borrowers pushed up the prices of all homes, including those bought by households using GSE mortgages. This explains
the nearly complete absence of statistically significant effects for constant-quality prices in a diff-in-diff framework. All prices rose, leaving no contrast. In addition, because the GSE borrowers faced higher market prices but had no additional purchasing power, they reduced the quality of the homes they bought. This explains why Figures 10a and 10b show the premium cut increased the quality of FHA borrowers relative to GSE borrowers. Interestingly, the quality effects are 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 likely would have had fewer resources to avoid moving down-market in the face of higher constant-quality prices. These equilibrium spillovers from FHA borrowers to GSE borrowers in FHA-rich areas are an important part of the mosaic created by the premium cut.

As the final set of robustness tests, we have redone our baseline analysis in two ways to address possible concerns that our results could have been affected by changes in loan characteristics and the FHA borrower pool after the premium cut. The first variant on the baseline adds controls for the borrower’s debt-to-income ratio (DTI) and loan-to-value ratio including any second liens (the combined loan-to-value ratio, or CLTV).\(^\text{30}\) This specification controls for potential changes in loan risk factors that might not be captured by the borrower characteristics included in the baseline regression. The second variant retains the baseline specification but excludes what we call “poachable” loans. These are borrowers in the contested part of the market that are the most likely to have switched from GSE financing to an FHA loan after the premium cut and whose behavior conceivably could differ from standard FHA borrowers. We characterize this part of the market by using the loan-level risk ratings from AEI’s National Mortgage Risk Index.\(^\text{31}\) The NMRI estimates the probability of default under stressed conditions akin to the 2007-08 financial crisis. FHA and the GSEs both have sizable market shares for loans with estimated stressed default rates between 8 and 18 percent; these are the poachable loans that we exclude. 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.

Table 8. 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

<table>
<thead>
<tr>
<th></th>
<th>Average across treatment periods</th>
<th>Number of loans</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Constant-Quality Price Effect</td>
<td>Quality Effect</td>
</tr>
<tr>
<td>1. Baseline</td>
<td>2.8 ppts**</td>
<td>-0.1 ppt</td>
</tr>
<tr>
<td>2. Baseline plus controls for CLTV and DTI buckets</td>
<td>2.8 ppts**</td>
<td>0.5 ppt</td>
</tr>
<tr>
<td>3. Baseline excl. poachable loans(^1)</td>
<td>2.8 ppts**</td>
<td>-0.3 ppt</td>
</tr>
</tbody>
</table>

* 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.

\(^\text{30}\) 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.

The results are shown in Table 8. The first row of the table repeats our baseline results for reference. Row 2 adds the controls for DTIs and CLTVs, while row 3 excludes the poachable loans. As shown, the results for the constant-quality price effect is completely unaffected by either of these changes. In addition, while the size of the quality effect varies somewhat across the three rows, it remains insignificant in all cases. This variation in the quality effect carries over to the overall price effect, which is somewhat larger relative to baseline in row 2 and slightly smaller in row 3. On balance, these results confirm our baseline findings.

Table 9 presents parallel results for the within-market comparison of the effects of the premium cut, i.e. the comparisons of boxes A and B in Figure 4. The first row repeats our baseline results, while row 2 adds CLTV and DTI as controls and line 3 excludes poachable loans. In all three rows, the sample consists of FHA and GSE loans in census tracts with an FHA share of at least 20 percent.

In the two alternative specifications, the constant-quality price effect becomes significant at the 10 percent level but remains quite small, showing that almost all of the house price increase induced by the premium cut – if not the full amount – spilled over to GSE borrowers. The changes in the quality effect from the baseline are more notable. When controlling for the DTI and CLTV (row 2), the quality effect increases from 2.5 percentage points to 2.9 percentage points. As noted above, this quality effect likely reflects the move down-market by GSE borrowers who were forced to compete with FHA borrowers benefitting from the premium cut. These GSE borrowers may have attempted to mitigate the reduction in house quality by taking on more debt. When we control for the two indicators of greater debt, the DTI and CLTV, we take away this avenue for mitigating the move down-market and find a larger quality effect. The slightly larger overall effect in row 2 relative to the baseline largely mirrors the change in the quality effect.

Furthermore, in results not shown, the size of the quality effect in row 2 increases when we raise the minimum FHA share from 20 percent to higher minimum shares. In census tracts with greater concentrations of FHA loans, the GSE borrowers tend to be less well-off relative to GSE borrowers in tracts with lower FHA shares, and thus may have less wherewithal with which to avoid a decline in house quality in response to higher constant-quality prices.

Table 9. FHA Price and Quality Acceleration in 2015 vis-à-vis GSE Loans, Alternative Control Variables and Sets of Loans: FHA and GSE loans in high FHA share tracts

<table>
<thead>
<tr>
<th></th>
<th>Average across treatment periods</th>
<th>Number of loans</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Constant-quality Price Effect</td>
<td>Quality Effect</td>
</tr>
<tr>
<td>1. Baseline</td>
<td>0.4 ppt</td>
<td>2.5 ppts**</td>
</tr>
<tr>
<td>2. Baseline plus controls for CLTV and DTI buckets</td>
<td>0.5 ppt*</td>
<td>2.9 ppts**</td>
</tr>
<tr>
<td>3. Baseline excl. poachable loans1</td>
<td>0.6 ppt*</td>
<td>1.8 ppts*</td>
</tr>
</tbody>
</table>

* 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.

This mechanism may also explain the lower quality and overall price effects in line 3. The poachable GSE loans are disproportionally in tracts with relatively high FHA loan shares. When we omit these loans, the remaining GSE borrowers tend to live in tracts with a limited FHA presence. These low-risk borrowers may have greater financial wherewithal to compensate for higher prices by using cash reserves, which
would work to reduce the size of the quality effect. Therefore, the most adverse impact of the premium cut on house quality appears to have hit less well-off GSE borrowers.

Impact on Homeownership

The next part of our analysis examines the degree to which FHA’s premium cut introduced new homebuyers to the market. As we noted in the introduction, FHA estimated that the premium cut would add 250,000 new homebuyers over the following three years, roughly 83,000 per year. NMRI data show that FHA’s volume of home purchase loans rose 217,000 from 2014 to 2015 and that 179,000 of the added loans went to first-time buyers. But not all of these new homebuyers were drawn into the market by the lower premium. Some of FHA’s new borrowers potentially would have used financing from other government agencies, and instead used an FHA mortgage as a result of the premium cut. As mentioned earlier, we call these “poached” buyers. In addition, some of the added borrowers would have entered the market regardless of the lower premium because an improving economy was raising incomes and lowering unemployment across the nation.

To disentangle the three causes of new FHA business after the premium cut — poaching, market trends, and new homebuyers — we estimate a multinomial logit regression model using the NMRI data to predict the extent of FHA’s poaching, which we then use to estimate the amount of new FHA business stemming from market trends. Having estimated both of these effects, we set the number of new homebuyers that FHA attracted equal to the residual.

The assumption for the multinomial logit model is that the characteristics of loans made by each of the government guarantee agencies should be fairly stable from one year to the next, holding all else equal. This allows us to estimate an aggregate loan count for each agency had FHA not cut its premium. The difference between the actual and the predicted loan counts provides an estimate of the extent of FHA’s poaching from the other agencies.

The multinomial logit model allows us to establish a relationship between observable loan characteristics, such as borrower credit score, and the agency guaranteeing the loan. For this analysis, we use NMRI data for primary owner-occupied homes bought by first-time buyers — the group most relevant to policy makers aiming to broaden homeownership. We estimate the 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. Based on the model’s fit for 2014 and observable characteristics for loans made in 2015, we then predict loan volume for each guarantee agency in 2015, the year after the premium cut.

The predicted results for 2015 represent our estimate of an agency’s loan count had FHA’s premium cut never occurred. The difference between the actual 2015 loan count and the predicted count represents the amount of FHA’s poaching.

The results of the multinomial logit model also allow us to estimate the market trend, which represents the growth in loan volume that would have occurred regardless of FHA’s premium cut. We estimate the market trend by comparing the 2015 logit-predicted loan volume for those agencies whose policies did not change with their 2014 actual loan total. Using the predicted loan volume controls for any FHA

---

32 The full set of explanatory variables for the model are borrower credit score, state in which the property is located, price of the property, number of borrowers, and whether the loan was originated by a bank or nonbank. For more on the model, see Appendix D.

33 FHA does not guarantee investor or second home loans. Therefore, it cannot poach these loans.

34 See Appendix D for an assessment of the logit model fit. As discussed there, analysis using holdout samples shows that the model accurately predicts loan counts by agency.
poaching that occurred. For FHA, we assume the same market-trend growth rate between 2014 and 2015 minus an adjustment factor of 3.7 percentage points. The adjustment factor is based on a comparison of market growth between FHA and the other agencies in the lead-up to FHA’s premium cut. By adjusting FHA’s growth rate down, we lower the number of borrowers attributed to trend growth and increase the number of new homebuyers, which represents the best-case scenario for FHA. For more on the adjustment factor, see Appendix E.

Table 10 summarizes the results of this exercise. The 2014 and 2015 actual loan totals for the GSEs, VA, and RHS come from the NMRI dataset, while the 2015 predicted loan totals are obtained from the multinomial logit model. By subtracting the 2015 actual from the predicted, we estimate the extent of FHA’s poaching in 2015. As shown in the second-to-last line in the table, we estimate that FHA poached about 86,000 first-time buyer 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 10 percent of their first-time buyer business in 2015, RHS lost about 21 percent, and VA lost about 6 percent. With these results, the final line of the table shows that we estimate these agencies taken together would have increased their first-time buyer volume by about 18 percent in 2015 had FHA not poached from them.

Table 10. Summary of Poaching and Trend Calculation

<table>
<thead>
<tr>
<th>First-lien, primary owner first-time buyer loans</th>
<th>GSEs</th>
<th>VA</th>
<th>RHS</th>
<th>GSE+VA+RHS</th>
</tr>
</thead>
<tbody>
<tr>
<td>2014 Actual</td>
<td>541,394</td>
<td>140,206</td>
<td>104,028</td>
<td>785,618</td>
</tr>
<tr>
<td>2015 Actual</td>
<td>590,319</td>
<td>161,516</td>
<td>90,729</td>
<td>842,564</td>
</tr>
<tr>
<td>Predicted (from model)</td>
<td>645,702</td>
<td>170,339</td>
<td>112,911</td>
<td>928,952</td>
</tr>
<tr>
<td>2015 Trend: (2015 Predicted / 2014 Actual) – 1</td>
<td>19.3%</td>
<td>21.5%</td>
<td>8.5%</td>
<td>18.2%</td>
</tr>
</tbody>
</table>

Lastly, we calculate how many new homebuyers FHA drew into the market. The number of new buyers is the residual obtained by subtracting the loans FHA poached and the loans that FHA would have made due to the existing momentum in the housing market from FHA’s 2015 actual loan total. Table 11 summarizes the results, which show that of the nearly 180,000 first-time buyers that FHA added in the first year after its premium cut, almost half were poached from other agencies. Close to 60,000, or one-third, represented market trend growth unrelated to FHA’s premium cut, and only around 35,000, or one-fifth, were new homebuyers who could opted for homeownership because of the cut. This total falls far short of the 83,000 marginal borrowers per year that FHA predicted before its policy change.

Table 11. Disaggregation of 2015 Increase in FHA First-Time Buyer Volume

<table>
<thead>
<tr>
<th>Total</th>
<th>% of 2015 FHA increase</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>179,049</td>
</tr>
<tr>
<td>Poaching</td>
<td>86,389</td>
</tr>
<tr>
<td>Trend</td>
<td>58,254</td>
</tr>
<tr>
<td>New Homebuyers</td>
<td>34,406</td>
</tr>
</tbody>
</table>

As noted in the introduction, the estimated 35,000 new buyers in 2015 represent about 6.5 percent of the 539,000 first-time borrowers we would have expected the FHA to underwrite had the premium cut not occurred. The estimate of 539,000 equals the total number of first-time FHA borrowers, 660,000, less the 86,000 buyers poached from the GSEs, less the 35,000 new buyers induced by the premium cut. This 6.5 percent increase implies a semi-elasticity of homeownership to interest rates over one year of about 8.9 (6.5 percent divided by the implied 0.73 percentage point decline in the mortgage rate).
Wealth Transfers

The premium cut redistributed wealth among participants in the housing market. FHA borrowers gained, as the increase in their purchasing power of about 6 percent exceeded the 2.8 percent rise in constant-quality home prices that they faced relative to the control group of GSE borrowers. Sellers of homes in markets with a significant FHA share gained as well, as the prices of their homes increased. In contrast, GSE borrowers residing in markets with a significant FHA share of mortgages were hurt by the premium cut, as the rise in constant-quality home prices eroded their purchasing power.

To put some numbers on these wealth transfers, recall that Figure 8b showed a statistically significant increase in house prices for FHA borrowers in census tracts with an FHA share between 20 percent and 60 percent. The price increases affected not only GSE and FHA borrowers, but also borrowers with VA, RHS, or private-sector financing in these markets. 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. Assuming all these borrowers faced a 2.8 percentage increase in house prices resulting from the equilibrium effects of the premium cut (about $6,000 extra per house for these borrowers), then in the aggregate the cost of the houses purchased by these borrowers increased nearly $7 billion. Given our estimate that the FHA premium cut generated 35,000 new homeowners, non-FHA borrowers paid almost $200,000 per new FHA homeowner – quite a hefty tax.

Why Didn’t the FHA Poach More Buyers?

Our analysis suggests that nearly half of the increase in first-time homebuyers after the premium cut was poached 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 12 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 CLTV is 98.2 percent. The GSE loans we consider have a range of possible CLTVs (shown by the rows in the table) and have an interest rate equal to 4 percent plus a loan-level pricing adjustment (LLPA). Given the difference in CLTVs across the GSE and FHA loans, the GSE

---

35 The 1.22 million borrowers are represented by orange bars in the right panel of Figure 5 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.

36 We obtain the $6,000 estimate as follows. Using HMDA data for 2015, the average loan amount for all non-FHA home purchase loans 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 2.8 percent of that amount is about $6,000.

37 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 bp LLPA adjustment equals 20 bp 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
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 opt for the maximum CLTV in order to minimize their down payment.

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 relatively small pool of buyers that 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) CLTVs between 86 and 90 percent and credit scores below 640 or (ii) CLTVs 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 CLTV 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 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 mortgages. Each cell is computed using NMRI data as:

\[ \Delta \text{GSE}\%_{2015} - \Delta \text{GSE}\%_{2014} \]

where \( \Delta \text{GSE}\%_{2015} \) is the change between 2014 and 2015 in the percentage of total GSE loans accounted for by loans in that cell and \( \Delta \text{GSE}\%_{2014} \) is the change between 2013 and 2014.\(^\text{38}\) The cells of the table 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 CLTVs 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.\(^\text{39}\)

\[ \text{an 80 percent CLTV for the cells in the CLTV-bucket rows marked “1-80%” and a 97 percent CLTV for the rows marked “≥96%”}. \]

\( \text{38} \) We study the change in the change in the distribution because credit standards were declining throughout this period, which caused the shares in the higher-risk cells of the table to trend up.

\( \text{39} \) In 2015, the GSEs pushed to increase lending with 3 percent down payments for high-credit-score borrowers, explaining the increase in the share of these loans in the bottom row of the bottom panel.
Table 12:  
GSE pricing advantage over FHA before FHA MIP cut

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

GSE pricing advantage over FHA after FHA MIP cut

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

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

<table>
<thead>
<tr>
<th>CLTV bucket</th>
<th>Credit score bucket</th>
<th>&lt; 640</th>
<th>640-659</th>
<th>660-679</th>
<th>680-699</th>
<th>700-719</th>
<th>720-739</th>
<th>740-759</th>
<th>≥ 760</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-80%</td>
<td></td>
<td>-0.2%</td>
<td>-0.2%</td>
<td>-0.3%</td>
<td>-0.1%</td>
<td>-0.1%</td>
<td>0.2%</td>
<td>0.5%</td>
<td>2.9%</td>
</tr>
<tr>
<td>81-85%</td>
<td></td>
<td>0.0%</td>
<td>0.0%</td>
<td>-0.1%</td>
<td>-0.1%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.1%</td>
</tr>
<tr>
<td>86-90%</td>
<td></td>
<td>-0.1%</td>
<td>-0.2%</td>
<td>-0.2%</td>
<td>-0.3%</td>
<td>-0.3%</td>
<td>-0.1%</td>
<td>0.0%</td>
<td>0.4%</td>
</tr>
<tr>
<td>91-95%</td>
<td>-0.1%</td>
<td>-0.4%</td>
<td>-0.8%</td>
<td>-1.2%</td>
<td>-1.1%</td>
<td>-1.2%</td>
<td>-0.9%</td>
<td>-1.2%</td>
<td></td>
</tr>
<tr>
<td>≥96%</td>
<td></td>
<td>0.0%</td>
<td>0.1%</td>
<td>0.2%</td>
<td>0.5%</td>
<td>0.6%</td>
<td>0.9%</td>
<td>1.0%</td>
<td>1.9%</td>
</tr>
</tbody>
</table>

Data are for primary owner-occupied purchase loans only.
Pricing calculation assumes a house price of $250,000, an interest rate of 4.0% for FHA and 4.0% for GSE plus loan level pricing adjustment. FHA loan assumes a CLTV of 96.5% plus 1.75% upfront MIP.
Cells with red font indicate a pricing advantage for FHA over GSE loans for a given CLTV and credit score combination.
Cells shaded in yellow indicate CLTV and credit score combination for which the pricing advantage changed from the GSEs to FHA after FHA's MIP cut.
Cells shaded in red indicated CLTV and credit score combination where the GSEs lost significant market share from 2014 to 2015 as compared to 2013 to 2014.
5. Conclusion

We exploit a surprise cut in the FHA mortgage 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. We find that, relative to a control group of GSE borrowers in census tracts with a very low share of FHA mortgages, FHA borrowers increased the value of their purchased housing by 2.8 percentage points, accounting for about 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 conventional financing had the premium cut not occurred, we estimate that the change in FHA policy created about 35,000 new homebuyers in the first year after the cut, implying a semi-elasticity of homeownership to interest rates over one year of 8.9. To conclude the paper, we explore the distributional effects of the premium cut, showing that it benefitted FHA borrowers and home sellers at the expense of home buyers who did not use FHA financing. As a group, the non-FHA buyers incurred a cost of nearly $200,000 in the form of higher house prices for each of the 35,000 new FHA buyers.

References


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 used to match to HMDA and Federal Housing Finance Agency (FHFA) data to add borrower characteristics and income to the ATTOM dataset.

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.40
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.
6) Duplicates: We remove all records that have an identical sale price, filing date, loan amount, and property location.

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 below 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).

40 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.
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 is in the top or bottom 1 percent of values within each county and 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.

**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.41

**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.42 Using this additional field for the unmatched loans resulted in about 113,000 additional matches.

---

41 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 solely to rounding.

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

**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](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 mortgagee.\(^43\) 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.\(^44\) 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 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.

---

\(^{43}\) 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.”

\(^{44}\) 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.
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 add 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.

**Third-round match for loans with conventional financing:**

**First-step match:** 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)\(^{45}\), 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 the loan from both datasets.

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.

These steps yield about 241,000 one-to-one matches out of the 348,000 conventional loans in the ATTOM/HMDA/FHFA-matched dataset. A higher match rate could be achieved with an alternative to the NMRI data that included a finer level of geography than state.

**Second-step match:** 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.\(^{46}\) 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

---

\(^{45}\) The NMRI data generally records 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.

\(^{46}\) One minor limitation is that the Freddie data (but not the Fannie data) exclude adjustable-rate mortgages (ARMs). ARMs represented only about 4 percent of Freddie’s 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](http://www.fanniemae.com/portal/funding-the-market/data/loan-performance-data.html) and [http://www.freddiemac.com/research/datasets/sf_loanlevel_dataset.html](http://www.freddiemac.com/research/datasets/sf_loanlevel_dataset.html).
one-to-one matches between loans in ATTOM/HMDA/FHFA and the GSE data, and we remove all matched loans from either dataset 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 match:** 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.

**Third-round match for loans with FHA financing:**

**First-step match:** 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 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), 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

---

47 For FHA loans originated starting in 2015, 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.
second borrower is reported in HMDA or FHFA data), and the property type (Single Family, Condo, Coop, Mobile Home). Because of the coarser geographical detail and the truncated loan amount, these stages yield fewer matches than the matches based on the subset of NMRI matched to the FHA Snapshot data. In later stages, we gradually remove the variables as matching criterion 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.

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 match: 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. In a second stage, we remove lender and allow for matches to fall within plus/minus 3 months of the origination date. 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.

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 the loan originator is the selling lender and declare a match on the lender name.
Appendix D: 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, the home’s sale price (represented by a set of price buckets), the state in which the property is located, and lender type (bank or nonbank). The probabilities of the four agencies are additive for each loan and equal 100 percent. We estimate the model on 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 500,000 of the nearly 1.2 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 is +/-2,500 loans. As discussed in section 4, we estimate that FHA poached about 86,000 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 83,500 to 88,500. 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 E: FHA market trend adjustment factor

As described in section 4, we estimate the market trend growth in FHA’s first-time buyer volume from 2014 to 2015 as the growth over this period for the other government agencies combined (the GSEs, VA, and RHS) minus an adjustment factor of 3.7 percentage points. We use an adjustment factor because the growth in FHA’s volume had been lagging behind that of the other agencies combined in the period leading up to the premium cut. Here we describe the rationale for this value of the adjustment factor and discuss the effect of selecting different values.

Figure E1, which uses data from the NMRI, plots the percentage point difference in year-over-year first-time buyer volume growth for FHA and the combination of the other guarantee agencies. The solid black line at 0 percent provides the benchmark of no difference in year-over-year growth. The figure shows that first-time buyer volume had been growing more slowly at FHA than for the rest of the agency market before the premium cut. However, the gap had been narrowing as shown by the upward sloping blue line, which crossed the solid black line in December 2014, indicating faster FHA first-time buyer volume growth than the rest of the market.
For the six months prior to FHA’s premium cut, FHA’s gap appears to have stabilized. Averaging the year-over-year growth rate differential for those six months yields the market trend adjustment factor of 3.7 percentage points (the dashed black line).

We perform sensitivity analysis around this adjustment factor to ascertain its potential impact on the estimated number of new homebuyers that FHA added in 2015. Table E1 summarizes the results of this analysis. Scenario 1 represents our baseline case, in which FHA’s first-time buyer growth in 2015, absent the premium cut, would have been 3.7 percentage points slower than for the rest of the agency market. This scenario implies that the premium cut resulted in about 35,000 new homebuyers. Scenario 2 assumes an adjustment factor of zero, or equal growth rates for FHA and the rest of the agency market. In this scenario, FHA only adds about 20,000 new homebuyers, as more of FHA’s actual growth in 2015 is attributed to market trend growth and less to the stimulative effect of the premium cut. Scenario 3 assumes an adjustment factor of half of 3.7 percentage points — the middle case between Scenarios 1 and 2. In this scenario, FHA adds 27,000 new homebuyers. Lastly, for illustrative purposes, Scenario 4 assumes no trend growth at all in FHA’s first-time buyer volume in 2015. This is an extremely unlikely scenario because first-time buyer volume for the rest of the agency market grew about 18 percent (after adding back the roughly 86,000 loans that FHA is estimated to have poached). The implied adjustment factor of 18 percentage points is about five times the size of our baseline adjustment factor. In this highly unlikely scenario, FHA’s premium cut would have added almost 93,000 new homebuyers, which is barely more than the 83,000 FHA would need to add for three years in a row to reach its stated goal of 250,000 new homebuyers.
Table E1. Disaggregation of 2015 Increase in FHA First-Time Buyer Volume, by Scenario

<table>
<thead>
<tr>
<th>Scenario 1: FHA Market Trend (FHA = Rest of agency market - market trend adj. factor)</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Poaching</td>
<td>86,389</td>
<td>48%</td>
</tr>
<tr>
<td>Trend</td>
<td>58,254</td>
<td>33%</td>
</tr>
<tr>
<td>New homebuyers</td>
<td>34,406</td>
<td>19%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Scenario 2: No adjustment (FHA = Rest of agency market)</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Poaching</td>
<td>86,389</td>
<td>48%</td>
</tr>
<tr>
<td>Trend</td>
<td>73,073</td>
<td>41%</td>
</tr>
<tr>
<td>New homebuyers</td>
<td>19,587</td>
<td>11%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Scenario 3: Middle case (FHA = Rest of agency market - 1/2 market trend adj. factor)</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Poaching</td>
<td>86,389</td>
<td>48%</td>
</tr>
<tr>
<td>Trend</td>
<td>65,596</td>
<td>37%</td>
</tr>
<tr>
<td>New homebuyers</td>
<td>27,064</td>
<td>15%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Scenario 4: No FHA growth (purely illustrative)</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Poaching</td>
<td>86,389</td>
<td>48%</td>
</tr>
<tr>
<td>Trend</td>
<td>-</td>
<td>0%</td>
</tr>
<tr>
<td>New homebuyers</td>
<td>92,660</td>
<td>52%</td>
</tr>
</tbody>
</table>