RESEARCH REPORT

# Women Are Better than Men at Paying Their Mortgages

*Laurie Goodman* *Jun Zhu* *Bing Bai*

*September 2016*

URBAN INSTITUTE · ELEVATE · THE · DEBATE

## ABOUT THE URBAN INSTITUTE

The nonprofit Urban Institute is dedicated to elevating the debate on social and economic policy. For nearly five decades, Urban scholars have conducted research and offered evidence-based solutions that improve lives and strengthen communities across a rapidly urbanizing world. Their objective research helps expand opportunities for all, reduce hardship among the most vulnerable, and strengthen the effectiveness of the public sector.

# Contents

# Acknowledgments

The Housing Finance Policy Center (HFPC) was launched with generous support at the leadership level from the Citi Foundation and John D. and Catherine T. MacArthur Foundation. Additional support was provided by The Ford Foundation and The Open Society Foundations.

Ongoing support for HFPC is also provided by the Housing Finance Council, a group of firms and individuals supporting high-quality independent research that informs evidence-based policy development. Funds raised through the Council provide flexible resources, allowing HFPC to anticipate and respond to emerging policy issues with timely analysis. This funding supports HFPC's research, outreach and engagement, and general operating activities.

This report was funded by these combined sources. We are grateful to them and to all our funders, who make it possible for Urban to advance its mission.

The views expressed are those of the authors and should not be attributed to the Urban Institute, its trustees, or its funders. Funders do not determine research findings or the insights and recommendations of Urban experts. Further information on the Urban Institute's funding principles is available at www.urban.org/support.

# Women Are Better than Men at Paying Their Mortgages

It's a fact: women on average pay more for mortgages. We are not the first people to have noticed this; a small number of other studies have also pointed it out (e.g., Cheng, Lin, and Liu 2011). One possible explanation is that women, particularly minority women, experience higher rates of subprime lending than their male peers (Fishbein and Woodall 2006; Phillips 2012; Wyly and Ponder 2011). Another explanation is that women tend to have weaker credit profiles (Van Rensselaer et al. 2013). We find that both these explanations are true and largely account for the higher rates.

Looking at loan performance for the first time by gender, however, we find that these weaker credit profiles do not translate neatly into weaker performance. In fact, when credit characteristics are held constant, women actually perform better than men. Nonetheless, since pricing is tied to credit characteristics not performance, women actually pay more relative to their actual risk than do men. Ironically, despite their better performance, women are more likely to be denied a mortgage than men. Given that more than one-third of female only borrowers are minorities and almost half of them live in low-income communities, we need to develop more robust and accurate measures of risk to ensure that we aren't denying mortgages to women who are fully able to make good on their payments.

In this paper, we first describe the data used for our analysis. Next we look at loan characteristics by gender through time. We then focus on loan performance, after which we draw our conclusions.

## Data Sources

The most complete source of loan origination data is public information filed under the Home Mortgage Disclosure Act (HMDA), the federal law that requires all but the smallest lenders to report annually. HMDA data contain information on the race/ethnicity and gender of the borrower and the coborrower, income, year of origination, interest rate, loan amount, loan purpose (purchase, refinance, or home improvement) and census tract of the property. It also contains information on whether the unit is owner occupied and whether it is a government or conventional loan.

However, HMDA data do not include any credit risk–related information like the loan-to-value (LTV) ratio of the property or the borrower's credit score. It also does not include any data on loan performance. By supplementing the HMDA data with proprietary loan-level data from CoreLogic, we can see all these data points and, thereby, obtain a more complete picture of the borrower at origination and observe the actual performance of the loan. CoreLogic covers the overwhelming majority of the mortgages we examined as it contains both loans contributed by a large number of servicers and all mortgage loans contained in private-label securitizations. The CoreLogic data contain extensive information on the loan, property, and borrower characteristics at the time of origination, as well as monthly updates on loan performance after origination. The procedure used to match the two databases is described in Li and colleagues (2014). In short, we match the two datasets by their origination year, loan amount, loan purpose (purchase/refinance), occupancy, lien, loan type (FHA/VA/conventional) and geography.[1]

For the descriptive part of the analysis and for the ordinary least square regressions measuring performance, we use all matches; to demonstrate performance using a hazard model we look only at unique matches.[2]

# Borrower Distribution by Gender over Time

HMDA data report six combinations of borrowers and coborrowers: female only, female-male, female-female, male only, male-female, and male-male. Table 1 shows the distribution of these groups, using data on the more than 60 million mortgages issued between 2004 and 2014. The most common category is male-female (a male borrower and a female coborrower), making up 40.5 percent of the mortgages extended over the 11-year period. Single borrowers are the next most common categories; male-only borrowers make up 28.9 percent of the mortgages and female-only borrowers make up 21.5 percent of the total. A female borrower and male coborrower make up about 7 percent of the total. Borrowers and co-borrowers of the same sex are much less common, making up about 1 percent each for two male and two female borrowers.

Table 1 divides the data into three subperiods to demonstrate how these numbers vary over time. The first period, 2004–07, covers the boom years running up to the financial crises. The second period, 2008–10, covers the financial crises, and the third period, 2011–14, covers the post-crises recovery. The years 2004–07 have a higher percentage of single-borrower mortgages (both female-only and male-only). The female-only share was 23.8 percent during this subperiod versus 21.5 percent for the 11 years. The male-only share was 30.9 percent during this subperiod versus 28.9 percent for the entire period. This discrepancy makes sense because credit was the least constrained during those years, making borrowing easier for those with just one income.

TABLE 1

**Gender Distribution of Mortgage Borrowers**

| Years | Female only | Male only | Male-female | Female-male | Female-female | Male-male |
|---|---|---|---|---|---|---|
| 2004–07 | 23.81 | 30.85 | 36.49 | 6.83 | 0.91 | 1.11 |
| 2008–10 | 19.65 | 26.24 | 44.42 | 7.57 | 0.96 | 1.15 |
| 2011–14 | 19.47 | 27.89 | 43.33 | 7.41 | 0.89 | 1.00 |
| **Total** | **21.46** | **28.89** | **40.48** | **7.18** | **0.91** | **1.08** |

**Source:** Authors' calculations based on matched HMDA and CoreLogic data.

## Borrower and Loan Characteristics by Gender

Table 2 shows the borrower and loan characteristics by six gender categories. Because male-male and female-female borrowers are only about 1 percent of the loans each, we do not discuss these categories in the analysis, though we include the results in the tables.

### BORROWER CHARACTERISTICS

We examined several characteristics of the more than 60 million mortgages originated between 2004 and 2014 for which we were able to match the HMDA and CoreLogic data.

In particular, we examined FICO scores and LTV ratios. A lower LTV at origination indicates that the borrower has put in more money in the form of a down payment and is borrowing a smaller portion of the money needed to buy the house. Thus, a lower LTV presents less risk to lenders because the borrower has more skin in the game. FICO scores give a picture of how well borrowers have paid their bills in the past. A higher credit score indicates a stronger payment history and, thus, a lower-risk borrower. We look at the borrower's debt-to-income (DTI) ratio; this measures the borrower's total payments on all debt including the mortgage, credit cards, auto loans, and student loans relative to income to make sure the share is sustainable. We also look at loan size/income, which measures the amount of debt taken out relative to a borrower's income. Finally, from HMDA data we are able to obtain the median income of the tract in which the property is located relative to the median income of the metropolitan statistical area and the share of the borrowers who live in areas where over 50 percent of the residents are minorities.

TABLE 2

**Summary Borrower Statistics**

| Category | FICO score | LTV | DTI | Loan size ($000s) | Income ($000s) | Loan size/ income | Median income, tract/ MSA (%) | Area minority population >50% (%) | Higher-priced loan (%) | Minority borrower (%) |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | **Full sample** | | | | |
| Female only | 711 | 75.07 | 33.25 | 176.41 | 69.22 | 2.91 | 46.2 | 22.9 | 15.6 | 34.1 |
| Male only | 712 | 77.63 | 33.11 | 202.94 | 94.72 | 2.66 | 44.7 | 20.6 | 15.0 | 32.1 |
| Male-female | 725 | 74.43 | 32.96 | 227.60 | 119.48 | 2.26 | 32.7 | 12.0 | 7.6 | 22.4 |
| Female-male | 718 | 75.86 | 33.29 | 213.24 | 110.23 | 2.25 | 38.0 | 15.3 | 12.6 | 27.5 |
| Female-female | 714 | 76.35 | 33.73 | 210.11 | 105.92 | 2.39 | 44.3 | 22.1 | 12.6 | 32.9 |
| Male-male | 717 | 77.37 | 33.27 | 230.42 | 151.38 | 2.11 | 45.5 | 21.6 | 11.0 | 30.1 |
| **All** | **718** | **75.64** | **33.10** | **208.33** | **101.18** | **2.52** | **39.7** | **17.3** | **11.9** | **28.3** |
| | | | | | | **2004–07** | | | | |
| Female only | 684 | 74.35 | 30.70 | 181.96 | 69.74 | 2.90 | 52.6 | 26.9 | 26.0 | 42.1 |
| Male only | 686 | 76.20 | 30.44 | 203.58 | 93.91 | 2.64 | 52.5 | 24.1 | 25.9 | 38.8 |
| Male-female | 694 | 74.38 | 30.27 | 220.53 | 108.48 | 2.35 | 40.8 | 13.7 | 13.8 | 27.9 |
| Female-male | 686 | 75.48 | 30.74 | 207.82 | 99.72 | 2.35 | 46.0 | 18.3 | 23.5 | 34.7 |
| Female-female | 686 | 75.18 | 30.85 | 205.47 | 98.15 | 2.44 | 51.0 | 24.9 | 21.0 | 39.8 |
| Male-male | 694 | 76.31 | 29.98 | 226.12 | 149.32 | 2.11 | 52.3 | 22.6 | 16.9 | 34.2 |
| *All* | *688* | *75.04* | *30.46* | *205.18* | *94.68* | *2.57* | *47.8* | *20.6* | *21.2* | *35.3* |
| | | | | | | **2008–10** | | | | |
| Female only | 732 | 73.98 | 36.08 | 170.29 | 66.23 | 2.95 | 42.2 | 16.1 | 7.8 | 27.2 |
| Male only | 731 | 76.82 | 36.11 | 197.65 | 91.51 | 2.74 | 40.5 | 14.5 | 7.6 | 25.8 |
| Male-female | 743 | 72.13 | 34.88 | 229.41 | 122.61 | 2.27 | 28.9 | 8.2 | 5.1 | 18.2 |
| Female-male | 735 | 74.33 | 35.55 | 215.78 | 113.17 | 2.26 | 34.5 | 10.6 | 7.5 | 22.7 |
| Female-female | 727 | 76.30 | 36.72 | 212.08 | 107.38 | 2.40 | 42.0 | 17.8 | 8.7 | 28.3 |
| Male-male | 729 | 77.64 | 36.75 | 228.56 | 145.86 | 2.16 | 42.8 | 18.2 | 8.5 | 27.2 |
| *All* | *737* | *74.00* | *35.53* | *208.25* | *102.75* | *2.53* | *35.3* | *11.8* | *6.6* | *22.5* |
| | | | | | | **2011–14** | | | | |
| Female only | 741 | 76.82 | 36.19 | 171.24 | 70.16 | 2.89 | 38.5 | 20.6 | 4.9 | 27.3 |
| Male only | 739 | 80.08 | 35.94 | 204.89 | 97.67 | 2.65 | 35.8 | 19.1 | 4.6 | 27.2 |
| Male-female | 748 | 75.83 | 35.07 | 234.17 | 129.75 | 2.17 | 26.3 | 12.5 | 2.9 | 19.5 |
| Female-male | 744 | 77.19 | 35.34 | 218.17 | 121.34 | 2.13 | 30.6 | 14.7 | 3.8 | 22.9 |
| Female-female | 741 | 77.92 | 36.16 | 215.00 | 115.44 | 2.30 | 36.8 | 21.2 | 4.8 | 27.9 |
| Male-male | 742 | 78.68 | 36.34 | 237.75 | 157.96 | 2.08 | 37.8 | 22.4 | 5.2 | 26.8 |
| *All* | *744* | *77.35* | *35.57* | *212.43* | *108.72* | *2.44* | *31.8* | *16.3* | *3.9* | *23.6* |

**Source:** Authors' calculations based on matched HMDA and CoreLogic data.
**Note:** LTV = loan to value; DTI = debt to income; MSA = metropolitan statistical area.

Our examination revealed nine critical points about the characteristics of single borrowers, particularly female-only borrowers:

1. ***Single borrowers, particularly women, are more likely to be minorities.*** About 34.1 percent of female-only borrowers are minorities, compared with 32.1 percent of male-only borrowers, 22.4 percent of male-female borrowers, and 27.5 percent of female-male borrowers.

2. ***Single borrowers, particularly women, are more likely to live in areas where the minority population is greater than 50 percent.*** Just under 23 percent of female-only borrowers lived in census tracts that were majority-minority, compared with 20.6 percent of male-only borrowers, 12.0 percent of male-female borrowers, and 15.3 percent of female-male borrowers. This pattern was consistent across subperiods.
3. ***Single borrowers, particularly women, are more likely to live in lower-income areas.*** We looked at the share of borrowers living in census tracts where the median income is lower than the median income of the metropolitan statistical area.[3] Over the entire period, 46.2 percent of female-only borrowers lived in lower-income census tracts, as did 44.7 percent of male-only borrowers, 32.7 percent of male-female borrowers, and 38.0 percent of female-male borrowers. This pattern is consistent across every subperiod.
4. ***Single borrowers, particularly women, are more likely to have higher-priced mortgages.*** About 15.6 percent of female-only borrowers have higher-priced mortgages (using the HMDA definition, which changes over time), versus 15.0 percent of male-only, 7.6 percent of male-female, and 12.6 percent of female-male borrowers.[4] Virtually all subprime mortgages fall into this higher-priced mortgage bucket.
5. ***Single borrowers have lower credit scores.*** Over the entire 2004–14 period, the credit score for female-only borrowers was 711, very similar to the 712 for male-only borrowers. The credit score was a considerably higher 725 for male-female borrowers and 718 for female-male borrowers. In each subperiod, the FICO score for female-only and male-only borrowers was very similar and considerably lower than the credit score for mixed-gender borrowers.
6. ***Single borrowers, particularly women, have lower incomes.*** The average income for female-only borrowers is $69,200, compared with $94,700 for male-only borrowers, $119,500 for male-female borrowers, and $110,200 for female-male borrowers. This ordinal pattern is consistent across every subperiod.
7. ***Single borrowers, particularly women, have smaller mortgages that eat up more of their income.*** The average loan size for female-only borrowers is $176,400; it is $202,900 for male-only borrowers, $222,700 for male-female borrowers, and $213,200 for female-male borrowers. The average loan size–to-income ratio is 2.91 for female-only, 2.66 for male-only, 2.26 for male-female, and 2.25 for female-male borrowers. These relationships hold across every subperiod.
8. ***Single borrowers have more debt as a share of their income.*** Single borrowers have slightly higher debt-to-income ratios than two borrowers. Female-only and male-only DTIs are similar. These patterns hold in each subperiod, but they do not show up as strongly in the total, as there were widespread misstatement of DTIs in 2004–07, when single borrowers were more prevalent.

Accordingly, when looking across groups, we consider DTI a relatively less reliable measure of risk than loan size/income.

9. ***Female-only borrowers own more of their homes than male-only borrowers.*** Female-only borrowers have lower LTVs than male-only and female-male borrowers but higher LTVs than male-female borrowers. The LTV for female-only borrowers is 75.07, versus 77.63 for male-only, 74.43 for male-female, and 75.86 for female-male borrowers. This pattern is consistent across every subperiod.

To summarize: Single borrowers are more likely to be minorities and to live in low-income and minority communities than paired male-female borrowers. They are also more likely to have high-cost mortgages and weaker credit characteristics than male-female borrowers. Specifically, single borrowers have lower credit scores, lower incomes, smaller mortgages that nonetheless take up more of their incomes, and more debt as a share of their incomes; they also own less of their homes. Single women are even more likely than single men to be minorities and live in poor and minority communities and to have weaker credit characteristics. That is, they have much lower incomes and much higher loan size divided by income. The exceptions to the rule that female-only borrowers have weaker credit characteristics than male-only borrowers: female-only borrowers have similar FICO scores and similar DTIs to male-only borrowers and they own more of their own homes than male-only borrowers. This finding is consistent with the results of Van Rensselaer and colleagues (2013) based on borrowers who obtained mortgages from New Century Financial Corporation from 2003 to 2005; their study also finds that female-only borrowers have lower LTVs.

While it is counterintuitive that women have much lower incomes and much higher loan size/income ratios, yet choose to put down a much higher down payment then single men, this finding is consistent with the gender comparison literature, which largely explains the phenomenon through greater risk aversion by women. For example, using data from the 1989 Survey of Consumer Finances, Jianakoplos and Bernasek (1998) find that, after controlling for age, education, children, and homeownership, single women hold a smaller percentage of their wealth in risk assets than single men. Using data from the 1992 and 1995 Survey of Consumer Finances, Sunden and Surrette (1998) examine differences in the allocation of defined contribution plan assets and find that single women are more risk averse than single men. Similarly, Agnew, Balduzzi, and Sunden (2003) show that women invest in fewer risk assets in their retirement plans, and Paley and Do (2010) show that women are less apt to apply for adjustable-rate mortgages than men.

## LOAN CHARACTERISTICS AND PERFORMANCE

Table 3 shows that female-only borrowers and female-male borrowers are more apt than other groups to take out FHA mortgages, and male-only borrowers and male-female borrowers are more apt than other groups to take out VA mortgages.

More important, *single borrowers face higher mortgage interest rates than mixed-gender borrowers*. Female-only borrowers also face higher interest rates than their male counterparts. From 2004 to 2014, the average rate for female-only borrowers was 5.48 percent versus 5.41 percent for male-only borrowers, 5.12 percent for male-female borrowers, and 5.25 percent for female-male borrowers. While we find this true in the aggregate, it is important to look at each subperiod, as we know rates were higher in the earlier period when sole borrowers were more prevalent. We find the same ordinal pattern holds for every subperiod, but the differences are smaller than they are over all 11 years. In the final subperiod, for example, the female-only borrowers rate was 4.01 percent, versus 3.99 percent for male-only, 3.93 percent for male-female, and 3.97 percent for female-male borrowers.

Van Rensselaer and colleagues (2013) show that as result of their weaker characteristics, female New Century Financial borrowers paid higher mortgage rates than men: 13 basis points (bps) higher in 2003, 8 bps higher in 2004, and 12 bps higher in 2005, numbers very consistent with ours. Haughwout, Mayer, and Tracy (2009) find that after controlling for characteristics, female borrowers in subprime deals do not pay more than male borrowers. However, Fishbein and Woodall (2006) find that women are more apt to obtain a subprime mortgage, a finding consistent with the work of Cheng, Lin, and Liu (2011) that women pay higher mortgage rates because of lender selection.

Table 3 also sheds light on mortgage performance by looking at the proportion of loans that go 90 days or more delinquent, including loans that are in REO or foreclosure before the 90-day delinquency point. We refer to this proportion as the default rate. *The data show that female-only borrowers actually default less than male-only borrowers, and that two mixed-gender borrowers default less than sole borrowers.* Because the default rate was so much higher in the first subperiod, 2004–07 (22.7 percent) than in the subsequent periods (7.8 percent in 2008–10 and 1.7 percent in 2011–14), a small shift in weighting will distort the picture of performance over time. It is thus important to look at performance by period of origination, rather than over the entire period. For mortgages originated from 2004 to 2007, the default rate for female-only borrowers was 24.6 percent, compared with 25.4 percent for male-only borrowers, 19.3 percent for male-female borrowers, and 22.1 percent for female-male borrowers. For mortgages originated in 2008–10, the female-only default rate was again lower than the male-only rate (9.6 percent versus 9.7 percent), while the mixed-gender default rates were considerably lower than the female-only rate. The same pattern holds for mortgages originated in 2011–14.

The fact that female-only borrowers have, on average, slightly weaker credit characteristics than male-only borrowers but default less indicates that women perform better given their credit characteristics. Later in the paper we employ a more rigorous regression analysis comparing female-only to male-only borrowers to provide more conclusive evidence.

TABLE 3

**Loan Characteristics and Performance by Gender of Borrower(s)**

| Category | FHA loans (%) | VA loans (%) | Interest rate (%) | Share of loans delinquent 90+ days (%) |
|---|---|---|---|---|
| | | **Full sample** | | |
| Female only | 13.6 | 1.1 | 5.48 | 14.71 |
| Male only | 12.4 | 4.6 | 5.41 | 14.67 |
| Male-female | 8.6 | 3.7 | 5.13 | 9.56 |
| Female-male | 12.9 | 1.9 | 5.26 | 11.56 |
| Female-female | 19.1 | 0.9 | 5.33 | 12.85 |
| Male-male | 18.9 | 1.1 | 5.34 | 12.48 |
| **All** | **11.3** | **3.2** | **5.30** | **12.35** |
| | | **Loans originated 2004–07** | | |
| Female only | 4.6 | 0.4 | 6.49 | 24.60 |
| Male only | 4.0 | 1.5 | 6.46 | 25.35 |
| Male-female | 4.1 | 1.6 | 6.24 | 19.27 |
| Female-male | 5.3 | 0.8 | 6.38 | 22.05 |
| Female-female | 9.2 | 0.3 | 6.37 | 22.60 |
| Male-male | 9.2 | 0.4 | 6.30 | 21.76 |
| *All* | *4.4* | *1.2* | *6.38* | *22.66* |
| | | **Loans originated 2008–10** | | |
| Female only | 27.2 | 1.1 | 5.32 | 9.61 |
| Male only | 25.0 | 4.7 | 5.31 | 9.67 |
| Male-female | 15.3 | 3.1 | 5.15 | 5.90 |
| Female-male | 24.0 | 1.7 | 5.23 | 7.68 |
| Female-female | 33.6 | 0.7 | 5.32 | 10.01 |
| Male-male | 33.9 | 1.0 | 5.35 | 8.97 |
| *All* | *21.2* | *3.0* | *5.23* | *7.83* |
| | | **Loans originated 2011–14** | | |
| Female only | 19.8 | 2.2 | 4.01 | 2.14 |
| Male only | 17.6 | 8.8 | 3.99 | 2.20 |
| Male-female | 9.6 | 6.4 | 3.94 | 1.21 |
| Female-male | 15.3 | 3.3 | 3.97 | 1.43 |
| Female-female | 22.9 | 1.7 | 3.99 | 1.77 |
| Male-male | 22.7 | 2.1 | 4.01 | 1.58 |
| *All* | *14.5* | *5.9* | *3.97* | *1.69* |

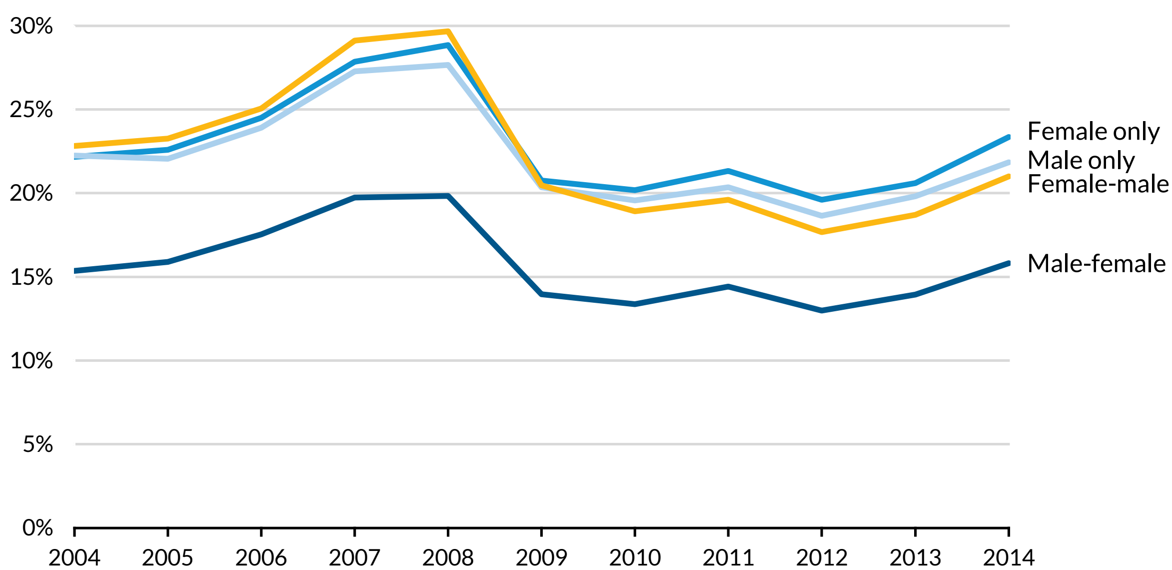
**Source:** Authors' calculations based on matched HMDA and CoreLogic data.
**Note:** FHA = Federal Housing Administration; VA = Department of Veterans Affairs.

# Denial Rates

Women default less despite slightly weaker credit characteristics. This suggests that we are not adequately capturing all the dimensions of credit that predict performance. Unfortunately, lenders are clearly relying on these imperfect predictions in their denial rate decisions: a higher percentage of female-only borrowers are denied mortgages than their male-only counterparts. Figure 1 shows the mortgage denial rates through time for female-only, male-only, male-female, and female-male borrowers. After 2009, the female-only category has a consistently higher denial rate than the other categories; before 2009, the two highest categories were female only and female-male.

FIGURE 1

**Denial Rates over Time**




**Source:** Authors' calculations based on matched HMDA and CoreLogic data.

# Race and Ethnicity of Sole Borrowers

Before we run the regressions we need to look at one last item: the race and ethnicity of female-only versus male-only borrowers. Differences in race and ethnicity between female- and male-only categories could influence our results. To some extent, race picks up wealth, a variable not captured in the standard measures of mortgage creditworthiness. According to the 2013 Federal Reserve Survey of Consumer Finances the median white family has a net worth of $142,000, while the median nonwhite or

Hispanic family has a net worth of $18,100 (Bricker et al. 2014). Since wealth can serve as a cushion to temper income volatility, and thus reduce the risk of default, the racial characteristics of borrowers could be important in explaining differences in default rates.

Table 4 shows the distribution by of borrowers by race and ethnicity. Generally, female-only borrowers are marginally less likely to be white than male-only borrowers (65.4 percent versus 67.5 percent). If nonwhite, female-only borrowers are more likely to be African American (11.8 percent) than their male counterparts (7.1 percent). Female-only borrowers are less like to be Hispanic (9.3 percent versus 11.4 percent) or Asian (5.4 percent versus 6.1 percent) than male-only borrowers.

Looking across the subperiods, we see find that in the 2004-2007 period, there were proportionately fewer white borrowers than in the other two periods. This is true for both female -only borrowers (57.9 percent versus 72.8 percent for the latter two periods) and male-only borrowers (61.2 percent versus 73.8 percent in the latter two periods). The 2004–07 period had a considerably higher proportion of female-only and male-only African American, Hispanic, and other borrowers than the latter periods. The only nonwhite group that increased its mortgage borrowing in the latter two periods was Asians.

TABLE 4

**Distribution of Mortgage Borrowers by Race/Ethnicity (percent)**

| Years | Asian | Black | Hispanic | Others | White |
|---|---|---|---|---|---|
| | | | **Female only** | | |
| 2004–07 | 4.86 | 14.74 | 11.66 | 10.86 | 57.88 |
| 2008–10 | 5.23 | 9.71 | 6.27 | 5.94 | 72.85 |
| 2011–14 | 6.20 | 8.52 | 7.48 | 5.04 | 72.77 |
| **Total** | **5.35** | **11.84** | **9.34** | **8.11** | **65.36** |
| | | | **Male only** | | |
| 2004–07 | 4.80 | 11.54 | 13.26 | 10.62 | 59.78 |
| 2008–10 | 5.92 | 7.31 | 7.19 | 5.99 | 73.60 |
| 2011–14 | 7.21 | 6.53 | 8.30 | 5.17 | 72.79 |
| **Total** | **5.80** | **9.12** | **10.51** | **7.98** | **66.59** |

**Source:** Authors' calculations based on matched HMDA and CoreLogic data.
**Note:** Numbers may not total 100 because of rounding.

# Regression Results: Ordinary Least Squares Analysis

Do female borrowers actually perform better than their male counterparts after controlling for other variables? To take a look at this question, we confine our analysis to sole borrowers. Couples with more than one borrower often have more than one income, providing a cushion against default that is absent with solo borrowers. Our results indicate the answer is a resounding yes: female performance is much stronger.

The regressions take the form:

$$D_i = f(\beta_1{}' X_i + \gamma_1 Gender_i + \rho_1 Race_i + \varepsilon_{1i}) \quad \text{(C1)}$$

$$D_i = f(\beta_2{}' X_i + \gamma_2 Gender_i + \rho_2 Race_i + \theta_2 Gender_i * Race_i + \varepsilon_{2i}) \quad \text{(C2)}$$

where *Gender* is a binary variable that is set to 1 if a borrower for a loan *i* is a woman and 0 otherwise. *X* is a vector of explanatory variables for loan *i*. *Race* is a category variable using white as the base. *D* indicates the default probability calculated from the HMDA and CoreLogic database matching process.

We first run ordinary least squares regressions, using our complete set of matched HMDA-CoreLogic data. That is, this regression is run using our 30-million-plus matches on female-only (13 million) and male-only (17 million) borrowers. These ordinary least squares variables measure the likelihood that a borrower ever goes delinquent by 90 days or more. To expand the size of the matched database beyond unique matches, we assigned weights to each HMDA-CoreLogic loan pair, to reflect how close the match is, and supplemented information in either database with information from the other using this weight. Where there is a unique match, the dependent variable will be 0 or 1. If there is not a unique match, the variable will be between 0 and 1, computed using the aforementioned weights; more details about the matching method can be found in Li and colleagues (2014). We used only unique matches for the hazard analysis discussed on pages 13–15.

We have two different specifications for the model. The first, equation C1, includes gender, race/ethnicity, and the mortgage variables that should matter for performance. The second, equation C2, adds the interactions between gender and race/ethnicity, to capture the gender differences by race and ethnicity. Since we are trying to measure the performance of female-only versus male-only borrowers, it is important that we control for all other factors that influence credit performance. Hence we include not only borrower- and mortgage-specific variables, but also characteristics of that census tract. These mortgage variables, which are listed in table 5, include term (longer-term mortgages are more likely to default), FICO (high-FICO mortgages are less likely to default), LTV (higher-LTV mortgages are more likely to default), interest rate (higher–interest rate mortgages are more likely to default), number of units (more units are more likely to default), documentation (reduced documentation is more likely to default), loan size/income (higher loan size/income is more likely to default), purchase indicator (purchase loans are less likely to default than refinanced mortgages), owner-occupied indicator (owner occupants are less likely to default than investors), census population in tract (more-populated tracts are less likely to default), tract to MSA income <100 percent indicator (if yes, more likely to default), and minority >50 percent indicator (if yes, more likely to default). We also include agency fixed effects (the supervisory agency of the institution reporting the loan),[5] loan type fixed effects (conventional, government), occupancy-type fixed effects (owner occupied, second, investor), purchaser fixed effects,[6] issue-year fixed effects (2004–14), and state fixed effects.

TABLE 5

**Cross-Section Results**

| | C1 | | C2 | |
|---|---|---|---|---|
| **Parameter** | **Estimate** | ***T*-statistic** | **Estimate** | ***T*-statistic** |
| Intercept | 6.65E-01 | 504.5 | 6.66E-01 | 504.6 |
| Female | -1.89E-03 | -25.9 | -6.74E-04 | -7.5 |
| Asian | 1.44E-03 | 9.0 | 1.04E-04 | 0.5 |
| Black | 3.38E-02 | 239.0 | 3.80E-02 | 193.3 |
| Hispanic | 2.94E-02 | 223.9 | 3.16E-02 | 196.2 |
| Other | 7.29E-03 | 52.8 | 7.90E-03 | 43.6 |
| Asian × female | | | 3.51E-03 | 11.2 |
| Black × female | | | -8.05E-03 | -31.4 |
| Hispanic × female | | | -5.58E-03 | -22.8 |
| Other × female | | | -1.47E-03 | -5.4 |
| Term | 1.23E-04 | 142.7 | 1.24E-04 | 142.9 |
| FICO | -1.08E-03 | -1,101.9 | -1.08E-03 | -1,102.1 |
| LTV | 1.21E-03 | 423.9 | 1.21E-03 | 424.2 |
| Interest rate | 1.18E-02 | 245.2 | 1.17E-02 | 244.9 |
| Units | 1.33E-02 | 75.8 | 1.33E-02 | 75.6 |
| Documentation | -7.52E-03 | -59.5 | -7.48E-03 | -59.1 |
| Size/income | 2.39E-04 | 19.0 | 2.37E-04 | 18.9 |
| Purchase indicator | -1.03E-03 | -11.9 | -1.05E-03 | -12.2 |
| Census population in tract | -7.38E-07 | -39.7 | -7.38E-07 | -39.7 |
| Tract to MSA income <100% indicator | 1.09E-02 | 136.7 | 1.09E-02 | 136.4 |
| Minority >50% indicator | 1.62E-02 | 148.8 | 1.62E-02 | 149.3 |
| Loan type fixed effects | Yes | | Yes | |
| Occupancy fixed effects | Yes | | Yes | |
| Agency fixed effects | Yes | | Yes | |
| Purchase type fixed effects | Yes | | Yes | |
| Year fixed effects | Yes | | Yes | |
| State fixed effects | Yes | | Yes | |
| Observations | 30,571,572 | | 30,571,572 | |
| $R^2$ | | 0.39026 | | 0.39031 |

The regression results are shown in table 5. Regression C1 shows that holding all else constant, a female-only borrower has a .002 (0.2 percent) lower probability of default than a male-only borrower, and the *t*-statistic is a very significant -25.9. That is, if a male-only borrower has a 6 percent probability of default, a female-only borrower with the same characteristics would be expected to default at a 5.8 percent rate. All minority borrowers have a higher probability of default than white borrowers. All the mortgage variables have the expected sign and are significant.

Regression C2 includes terms that allow for interaction between race/ethnicity and gender. In this specification, a female-only borrower has a .0007 (0.07 percent) lower probability of default, all else constant, with a *t*-statistic of -7.5. The cross terms indicate that African American, Hispanic, and other-race females have a lower default rate than their male counterparts. Asians are the only minority group where women default at a higher rate than men. Again, all the mortgage variables have the expected sign and are significant.

# Regression Results: Hazard Model

Cross-sectional regression analysis can be used to establish the relationship between default risk and credit factors, but it has a fixed time horizon (time from loan origination to the last observation, a period which is shorter for recent loans, longer for loans originated at earlier dates) and hence ignores the life-course of a mortgage. Therefore, we also run a hazard model, which considers the default probability in any given month as a function of the age of the mortgage as well as borrower and loan credit factors. We focus on the hazard rate of ever going 90 days or more delinquent (D90+) for female-only and male-only borrowers. Equation C3 is our baseline empirical specification—a hazard model showing whether a loan goes D90+ for the first time in month t during the period spanning the loan origination and the end of our performance period. It includes age, the credit variables, race and gender. C4 considers age, the credit variables, race and gender, and adds the interaction effect between race and gender.

$$h(t) = f(\beta_3{}'X_{it} + \gamma_3 Gender_{it} + \rho_3 Race_{it} + \varepsilon_{3it}), t = 1, \dots .T \quad \text{(C3)}$$

$$h(t) = f(\beta_4{}'X_{it} + \gamma_4 Gender_{it} + \rho_4 Race_{it} + \theta_4 Gender_{it} * Race_{it} + \varepsilon_{4it}), t = 1, \dots .T \quad \text{(C4)}$$

This hazard model is run using only the unique CoreLogic/HMDA matches. These are loans where we are highly confident we have paired each HMDA loan with its CoreLogic match; there is no uncertainty about gender or any other characteristics. For the hazard model, which measures the probability of default in only one period, we must use only unique matches, as each observation must have a 0 (no default) or a 1 (goes 90+ DQ) value in each period. Once the loan goes 90+ DQ, subsequent performance is not considered in the regression. For this regression, we include the origination variables we included in the ordinary least squares regression plus three time-varying covariates, which vary from month to month: mark-to-market LTV, loan age, and loan age squared (to allow the loan age variable to enter nonlinearly).

Table 6 shows the results of the hazard analysis, including the list of independent variables. This analysis substantiates the results of the first ordinary least squares analysis. Female-only borrowers are .0281 (or 2.81 percent) less likely to default in any given period, and the results are highly significant. These results are very consistent with the results in the ordinary least squares regression. The race/ethnicity variables indicate that nonwhite borrowers (African American, Asian, Hispanic, and other) are more likely to default than white borrowers. Again, the mortgage variables all have the correct sign.

TABLE 6

**Hazard Results**

| | C1 | | | C2 | | |
|---|---|---|---|---|---|---|
| **Parameter** | **Estimate** | **Hazard** | ***T*-st** | **Estimate** | **Hazard** | ***T*-st** |
| Intercept | -4.80E+00 | -9.92E-01 | -55.4 | -4.80E+00 | -9.92E-01 | -55.4 |
| Female | -2.85E-02 | -2.81E-02 | -4.8 | -2.50E-02 | -2.47E-02 | -3.4 |
| Asian | 3.82E-03 | 3.83E-03 | 0.3 | -2.67E-02 | -2.63E-02 | -1.3 |
| Black | 2.37E-01 | 2.67E-01 | 23.9 | 2.56E-01 | 2.91E-01 | 20.2 |
| Hispanic | 1.88E-01 | 2.07E-01 | 18.1 | 1.96E-01 | 2.16E-01 | 15.8 |
| Other | 3.67E-02 | 3.74E-02 | 3.2 | 3.07E-02 | 3.11E-02 | 2.2 |
| Asian × female | | | | 7.26E-02 | 7.53E-02 | 2.4 |
| Black × female | | | | -4.15E-02 | -4.06E-02 | -2.3 |
| Hispanic × female | | | | -2.30E-02 | -2.27E-02 | -1.1 |
| Other × female | | | | 1.60E-02 | 1.61E-02 | 0.7 |
| Term | 4.34E-03 | 4.35E-03 | 56.7 | 4.35E-03 | 4.35E-03 | 56.7 |
| FICO | -1.05E-02 | -1.05E-02 | -217.8 | -1.05E-02 | -1.05E-02 | -217.7 |
| LTV | 1.82E-02 | 1.83E-02 | 90.6 | 1.82E-02 | 1.83E-02 | 90.6 |
| Interest rate | 6.50E-02 | 6.72E-02 | 25.4 | 6.50E-02 | 6.72E-02 | 25.4 |
| Units | 5.19E-02 | 5.32E-02 | 7.6 | 5.16E-02 | 5.30E-02 | 7.6 |
| Documentation | -3.01E-01 | -2.60E-01 | -43.8 | -3.00E-01 | -2.59E-01 | -43.8 |
| Size/income | 2.48E-03 | 2.48E-03 | 13.1 | 2.48E-03 | 2.48E-03 | 13.1 |
| Purchase indicator | -2.13E-01 | -1.92E-01 | -33.5 | -2.13E-01 | -1.92E-01 | -33.6 |
| Census population in tract | -4.53E-06 | -4.53E-06 | -2.6 | -4.57E-06 | -4.57E-06 | -2.6 |
| Tract to MSA income <100% indicator | 4.90E-02 | 5.02E-02 | 8.0 | 4.90E-02 | 5.02E-02 | 8.0 |
| Minority >50% indicator | 6.10E-02 | 6.29E-02 | 7.0 | 6.13E-02 | 6.33E-02 | 7.1 |
| mtm_ltv | 2.00E-03 | 2.01E-03 | 72.2 | 2.00E-03 | 2.01E-03 | 72.2 |
| age_loan | 4.35E-02 | 4.44E-02 | 101.0 | 4.35E-02 | 4.44E-02 | 101.0 |
| $age\_loan^2$ | -5.10E-04 | -5.10E-04 | -99.8 | -5.10E-04 | -5.10E-04 | -99.8 |
| Loan type fixed effects | | Yes | | | Yes | |
| Occupancy fixed effects | | Yes | | | Yes | |
| Agency fixed effects | | Yes | | | Yes | |
| Purchase type fixed effects | | Yes | | | Yes | |
| Year fixed effects | | Yes | | | Yes | |
| State fixed effects | | Yes | | | Yes | |
| Observations | | 70,053,560 | | | 70,053,560 | |
| -2 log likelihood | | 1,609,708 | | | 1,609,694 | |

The second regression allows for interaction between race and gender. Again the results are very consistent with the ordinary least squares results. Female-only borrowers are .025 (or 2.5 percent) less likely to default in any one period, and the results are highly significant. The race/ethnicity variables indicate that nonwhite borrowers are more likely to default than white borrowers. The interaction terms show that female-only African American and Hispanic borrowers default less than their male counterparts; this pattern is reversed for Asian borrowers. Again, the mortgage variables all have the correct sign.

# Conclusion

Female-only borrowers pay more for their mortgages, both because they have generally weaker credit characteristics and because a higher percentage of these mortgages have been subprime. Our analysis shows, however, that these weaker credit characteristics do not accurately predict how well women pay their mortgages. Holding all credit characteristics constant, female-only borrowers default less than their male counterparts. This finding is true for white, African American, and Hispanic women. The bottom line: single women with mortgages are doing a better job of paying their mortgages than their credit characteristics predict. Because the higher price they pay for their mortgages is based on their credit characteristics when they take out the loan, this means single women borrowers are paying too much for their mortgages.

This inequality does not translate into a significant amount that single women overpay for their mortgages: less than $150 per female-only borrower per loan. The important issue, however, is that the dimensions we rely on to assess credit risk today do not adequately capture all the differences. This omission has real consequences. Women generally are denied for mortgages more often despite their superior payment performance. Given that more than one-third of single-women borrowers are minorities and almost half of them live in low-income communities, we need to develop more robust and accurate measures of risk to ensure that we aren't denying mortgages to women who are fully able to make good on their payments.

# Notes

1. The HMDA data contain information at the census tract level, while CoreLogic data contain information at the zip code level. We use a cross-walk to ensure geographic correspondence. When we find more than one unique match, which is often the case, we apply fuzzy matching and weight the matches.
2. For these loans, we are highly confident we have paired each HMDA loan with its CoreLogic match; there is no uncertainty about gender or any other characteristics.
3. Census tracts generally contains 1,200 to 8,000 people, with an optimal size of about 4,000 people. Thus, they are relatively homogenous areas.
4. We use the HMDA definition of higher-priced loans, a definition that is not consistent over time. From 2004 to 2009, loans were considered higher priced if the difference between the annual percentage rate (APR) on the loan and the yield on Treasury securities with a similar maturity was at least 3 percentage points for a first-lien loan and 5 percentage points for a subordinate lien (see "Frequently Asked Questions about the New HMDA Data," Office of the Comptroller of the Currency, April 3, 2006, https://www.occ.gov/news-issuances/news-releases/2006/nr-ia-2006-44a.pdf). Effective for loans closed after January 1, 2010, a loan is considered higher priced if the difference between the APR and that of a survey-based estimate of APRs offered on prime mortgage loans of a comparable mortgage type is at least 1.5 percentage points for a first-lien loan and 3.5 percentage points for a subordinate lien (see "What Is a 'Higher-Priced Mortgage Loan'?", Consumer Financial Protection Bureau, accessed August 23, 2016, http://www.consumerfinance.gov/askcfpb/1797/what-higher-priced-mortgage-loan.html).
5. Office of the Comptroller of the Currency, Federal Reserve System, Federal Deposit Insurance Corporation, the National Credit Union Administration, or, for nondepositories, the Department of Housing and Urban Development or the Consumer Financial Protection Bureau.
6. Fannie Mae, Freddie Mac, Ginnie Mae, private securitization, commercial bank, savings bank or savings association portfolio, life insurance company, credit union, mortgager bank or finance company portfolio, affiliate institution, and other.

# References

Agnew, Julie, Pierluigi Balduzzi, and Annika Sunden. 2003. "Portfolio Choice and Trading in a Large 401(k) Plan." *American Economic Review* 93 (1): 193–215.

Bricker, Jesse, Lisa J. Dettling, Alice Henriques, Joanne W. Hsu, Kevin B. Moore, John Sabelhaus, Jeffrey Thompson, and Richard A. Windle, 2014. "Changes in U.S. Family Finances from 2010 to 2013: Evidence from the Survey of Consumer Finances. *Federal Reserve Bulletin* Vol. 100, No. 4. Washington DC: Board of Governors of the Federal Reserve System. http://www.federalreserve.gov/pubs/bulletin/2014/pdf/scf14.pdf.

Cheng, Ping, Shenguo (Len) Lin, and Yingchun Liu. 2011. "Do Women Pay More for Mortgages?" *Journal of Real Estate Finance and Economics* 43 (4): 423–40.

Fishbein, Allen J., and Patrick Woodall. 2006. "Women Are Prime Targets for Subprime Lending: Women Are Disproportionately Represented in High Cost Mortgage Market." Washington, DC: Consumer Federation of America.

Haughwout, Andrew, Christopher J. Mayer, and Joseph S. Tracy. 2009. "Subprime Mortgage Pricing: The Impact of Race, Ethnicity and Gender on the Cost of Borrowing." Staff Report 368. New York: Federal Reserve Bank of New York.

Jianakoplos, Nancy Ammon, and Alexandra Bernasek. 1998. "Are Women More Risk Averse?" *Economic Inquiry* 36 (4): 620–30.

Li, Wei, Laurie Goodman, Ellen Seidman, Jim Parrott, Jun Zhu, and Bing Bai. 2014. "Measuring Mortgage Credit Accessibility." Washington, DC: Urban Institute. http://urbn.is/2bt4MES.

Paley, Irina, and Chau Do. 2010. "Gender and Mortgage Choice: More Evidence on Gender Based Risk Aversion." Washington, DC: US Office of the Comptroller of the Currency.

Phillips, Sandra. 2012. "The Subprime Mortgage Calamity and the African American Woman." *Review of Black Political Economy* 39:227–37.

Sunden, Annika E., and Brian J. Surette. 1998. "Gender Differences in the Allocation of Assets in Retirement Savings Plans." *American Economic Review* 88 (2): 207–11.

Van Rensselaer, Kristy N., Tanja F. Blackstone, Jerry Crabb, and Bruce L. Gordon. 2013. "Mortgage Pricing and Gender: A Study of New Century Financial Corporation." *Academy of Accounting and Financial Studies Journal* 18 (4): 95–100.

Wyly, Elvin, and C. S. Ponder. 2011. "Gender, Age, and Race in Subprime America." *Housing Policy Debate* 21:529–64.

# About the Authors

**Laurie Goodman** is the director of the Housing Finance Policy Center at the Urban Institute. The center is dedicated to providing policymakers with data-driven analysis of housing finance policy issues that they can depend on for relevance, accuracy, and independence. Before joining Urban in 2013, Goodman spent 30 years as an analyst and research department manager at a number of Wall Street firms. From 2008 to 2013, she was a senior managing director at Amherst Securities Group, LP, where her strategy effort became known for its analysis of housing policy issues. From 1993 to 2008, Goodman was head of global fixed income research and manager of US securitized products research at UBS and predecessor firms, which were ranked number one by *Institutional Investor* for 11 straight years. Before that, she was a senior fixed income analyst, a mortgage portfolio manager, and a senior economist at the Federal Reserve Bank of New York. Goodman was inducted into the Fixed Income Analysts Hall of Fame in 2009.

Goodman is on the board of directors of MFA Financial, is an advisor to Amherst Capital Management, and is a member of the Bipartisan Policy Center's Housing Commission, the Federal Reserve Bank of New York's Financial Advisory Roundtable, and Fannie Mae's Affordable Housing Advisory Council. She has published more than 200 journal articles and has coauthored and coedited five books.

Goodman has a BA in mathematics from the University of Pennsylvania and an MA and PhD in economics from Stanford University.

**Jun Zhu** is a senior research associate in the Housing Finance Policy Center at the Urban Institute. She designs and conducts quantitative studies of housing finance trends, challenges, and policy issues.

Before joining Urban, Zhu worked as a senior economist in the Office of the Chief Economist at Freddie Mac, where she conducted research on the mortgage and housing markets, including default and prepayment modeling. She was also a consultant to the Treasury Department on housing and mortgage modification issues.

Zhu received her PhD in real estate from the University of Wisconsin–Madison in 2011.

**Bing Bai** is a research associate with the Housing Finance Policy Center at the Urban Institute, where he helps build, manage, and explore data to analyze housing finance trends and related policy issues. Formerly an economic modeling senior at Freddie Mac, Bai conducted research on housing and mortgage markets and developed models to evaluate foreclosure alternatives for nonperforming mortgage loans. He holds a PhD in economics from Clemson University.

## Statement of Independence

URBAN INSTITUTE · ELEVATE · THE · DEBATE

2100 M Street NW
Washington, DC 20037

www.urban.org