Real Denial Rates

A Better Way to Look at Who Is Receiving Mortgage Credit

Laurie Goodman
Urban Institute

Bing Bai
Urban Institute

Wei Li
Federal Deposit Insurance Corporation

July 2018

The authors welcome feedback on this working paper. Please send all inquiries to lgoodman@urban.org.

This working paper has been submitted for publication in *Housing Policy Debate.*

Abstract

The observed mortgage denial rate (ODR), calculated from Home Mortgage Disclosure Act (HMDA) data, is often used to measure credit availability, but it does not account for shifts in applicants' credit profiles. In this paper, we reintroduce the real denial rate (RDR) as a way to account for credit differences and more accurately measure denial rates. We first introduced the RDR in 2014, and this paper updates our previous work with the most recent HMDA data, which we matched with CoreLogic proprietary data to obtain borrower demographic information (e.g., income and race and ethnicity) and mortgage credit characteristics (e.g., loan-to-value ratios, debt-to-income ratios, and credit scores). We account for shifts in applicants' credit profiles by considering only the denial rate of low-credit-profile applicants. This RDR can more accurately portray developments in mortgage credit accessibility. Our RDR results show that conventional mortgages have higher denial rates than government mortgages, racial and ethnic differences are smaller than the ODR indicates but are not eliminated, and small-dollar mortgages have higher RDRs, particularly in the government loan channel.

Real Denial Rates

The traditional mortgage denial rate, calculated from Home Mortgage Disclosure Act (HMDA) data, is often used to measure credit availability across time and across different racial and ethnic groups. But this can be a misleading measure of credit availability, as it depends on both the composition of borrowers who are applying for a mortgage and how tight credit standards are. Thus, higher denial rates can be the result of either a tighter credit environment or an increase in applications by borrowers with weaker credit.

This can best be illustrated by example. When we look at traditional mortgage denial rates, two unintuitive patterns emerge. First, denial rates were higher in 2007 than they were in 2017. If denial rates were a good measure of credit availability, it would suggest that credit was tighter in 2007 than it has been in recent years, which we know is not the case. In 2007, more applicants with weak credit profiles applied for mortgages, so demand was higher. Similarly, government mortgages from the Federal Housing Administration (FHA), the US Department of Veterans Affairs (VA), the U.S. Department of Agriculture's (USDA) Rural Housing Service (RHS), and the U.S. Department of Housing and Urban Development's Office of Public and Indian Affairs appear to have higher denial rates than conventional mortgages, but we know that applicants of government mortgages tend to have weaker credit profiles.

A better measure of the denial rate would hold the credit profile of the application pool constant. But this creates an analytic challenge because researchers can observe information about the credit characteristics only for applicants who receive loans, not those whose applications are denied.[1]

This paper first reviews the methodology for constructing a better measure of the mortgage application denial rate that accounts for shifts in the composition of the applicant

[1] Though a few proprietary mortgage databases, such as CoreLogic's, collect information on originated loans, the Home Mortgage Disclosure Act (HMDA) is the only source of mortgage application data that contains a mortgage applicant's income, loan amount, race or ethnicity, and application outcome. But HMDA data do not have information on common risk factors, such as credit score, loan-to-value ratio, debt-to-income ratio, and loan products. Therefore, an applicant's credit profile is unknown from HMDA data.

pool, which was first presented in Li and Goodman (2014a), using data on mortgage applications through 2013, and updated in Bai, Goodman, and Ganesh (2017). In this paper, we look at five implications of this revised measure, updated with data through 2017.

- The traditional observed denial rate (ODR) understates how difficult it is for borrowers with less-than-perfect credit to get a mortgage relative to our real denial rate (RDR) measure.
- The RDR measure more accurately reflects credit accessibility across time, showing that low-credit-profile borrowers were more likely to get turned down for a mortgage in 2017 than in 2006. The traditional ODR shows the opposite, as borrowers with low credit profiles simply did not apply for mortgages in 2017.
- The RDR also more accurately reflects differences across channels, with the government channel showing a lower RDR than the conventional channel.
- When we look at denial rates by race or ethnicity, denial rates do not disappear but are narrower using the RDR analysis. This suggests that a large component of the racial and ethnic differences in the ODR is because of differences in borrower credit.
- The RDR is higher for small-dollar mortgages (up to $70,000) than for larger loans. The differences are especially large in the government loan market.

Methodology and Data

We limit our universe to single-family (one-to-four-unit), owner-occupied purchase activity, as we are interested in mortgage credit availability to borrowers purchasing a home for personal use.[2] All mortgage loans that are extended go to either high-credit-profile (HCP) borrowers who will never be denied a mortgage and low-credit-profile (LCP) borrowers who might be denied. We define HCP applicants as those whose credit profiles are so strong that

[2] The choice to limit the analysis was also done for consistency over time. The underwriting for non-owner-occupied mortgages is different, as the property's cash flow plays a role. A refinance application is heavily dependent on interest rates. Moreover, various streamlined programs have allowed for loans to refinance that would not meet the criteria for a new loan, on the grounds that the refinance helps the borrowers and reduces the probability of loan default, to the benefit of the holder.

their probability of default is low; for our analysis, we assume it is zero. To calculate the real denial rate, we compare the number of loans denied (who are, by definition, assumed to be all LCP applicants) with LCP applicants who received mortgages. In other words, the RDR controls for applicant credit profiles by excluding HCP borrowers.

To determine whether an originated loan is HCP or LCP, we relied on the credit profiles of the mortgages, just as lenders would when evaluating credit. We first assembled the characteristics for the loans reported in the HMDA database. HMDA contains nearly the entire universe of loans.[3] It includes the applicant's income, loan amount, race or ethnicity, loan purpose, and application outcome. But HMDA does not have information on mortgage credit profile characteristics, such as loan-to-value (LTV) ratio, debt-to-income (DTI) ratio, credit score, documentation type (i.e., full, low, or no documentation), or product type. To gather this information, we matched HMDA to the CoreLogic proprietary database (using both their private-label securities and servicing loan-level databases). This proprietary database contains mortgage credit characteristics on originated loans but lacks demographic information on income and race or ethnicity. Because both databases are anonymized, we matched the data using the databases' common fields, such as geography, loan amount, origination date, loan purpose (e.g., purchase or refinance), loan type, and occupancy. Once we did the matching, we had a rich dataset that contains race or ethnicity, income, LTV ratio, DTI ratio, credit score, documentation type, and whether or not the loan is a risky product. The matching methodology we used and the matching rates are described in the appendix. See also Li, Goodman, Seidman, Parrott, Zhu, and Bai (2014) for more details.

The probability that a consumer is an LCP borrower is based on the historical default rates of mortgages with the same credit characteristics. To determine this, we first analyzed the expected default rates for various combinations of the LTV ratio, DTI ratio, credit score,

[3] HMDA is considered to be the universe of mortgage originations because federal law requires that almost all mortgage applications, except from lenders who make few loans, to be reported in HMDA. See Bhutta, Laufer, and Ringo (2017) for a more complete description. The reporting requirements have changed slightly. In 2016, all depository institutions with more than $44 million in assets that made at least one loan insured or guaranteed by a federal agency were required to report. Nondepository institutions that made more than 100 purchase loans or had assets over $10 million were required to report. In 2017, the reporting requirements were changed so that all institutions that made more than 25 closed-end loans in the preceding two years were required to report their closed-end loans.

documentation type, and whether or not the loan is a risky product.[4] The expected default rates rely on the actual experience of 2001 and 2002 originations (a proxy for a "normal" period in which home prices are rising, which is weighted 90%), and the experience of 2005 and 2006 originations (a proxy for a "stress" period, which is weighed 10%). See Li and Goodman (2014b) for more in-depth discussions on expected default risk for 360 different combinations of LTV ratios, DTI ratios, FICO scores, documentation types, and product types.

Based on expected mortgage default rates, we use the following definitions to construct a look-up table for the probability of a consumer being LCP, for various combinations of LTV ratios, DTI ratios, FICO scores, documentation types, and product types (Appendix Table A1).

- We assign a zero probability of being LCP to consumers who apply for loans without risky features and have a FICO score above 700, an LTV ratio less than 78%, and a DTI ratio less than 30%. This is the lower bound.
- We assign a 100% probability of being LCP to consumers who apply for loans without risky features and who have a FICO score below 580, an LTV ratio greater than 95%, and a DTI ratio greater than 50%. This is the upper bound.
- We do a linear transformation of expected default risk for consumers with credit risk in between the upper and lower bounds and assign a probability of being LCP accordingly.

These definitions are arbitrary, but the conclusions are not sensitive to the definitions, even though the numbers would change under a different weighting scheme.

[4] Loan products without risky features include fixed-rate mortgages and all hybrid adjustable-rate mortgages with an initial fixed-interest-rate period of five years or longer, without any of the following features: prepayment penalty, balloon terms, interest-only terms, and negative amortizations.

Calculating the Real Denial Rate

We illustrate the calculation of the real denial rate in Table 1. Again, the dataset used for this analysis is limited to owner-occupied, single-family properties, and all analyses in this paper refer solely to this universe.

According to HMDA data, there were 6,779,433 mortgage applications in 2006. Lenders denied 1,219,790 and approved 5,559,643.[5] So the traditional ODR is 18%.

Table 1. Calculating the Real Denial Rate

Variable	Variable name	Calculation or data source	2006	2017
Total # of loan applications	A	HMDA	6,779,433	3,809,074
# of loan applications denied by lenders	B	HMDA	1,219,790	394,448
% of loan applications denied by lenders (observed denial rate)	ODR	= B/A	18%	11%
# of loan applications approved by lenders[a]	C	= A – B	5,559,643	3,315,072
% of loans to low credit profiles[b]	D	CoreLogic matched with HMDA	53%	24%
# of approved loan applications by low credit profiles	E	= C×D	2,961,006	811,454
# of approved loan applications by high credit profiles	F	= C–E	2,598,637	2,614,815
# of loan applications by high credit profiles[c]	G	= F	2,598,637	2,614,815
# of denied loan applications by high credit profiles	H	= G–F	0	0
# of loan applications by low credit profiles	I	= A–G	4,180,796	1,194,259
% of loan applications by low credit profiles	J	= I/A	62%	31%
# of denied loan applications by low credit profiles	K	= B	1,219,790	382,805
% of loan applications by low credit profiles denied by lenders (real denial rate)	RDR	=K/I	29%	32%

Sources: HMDA, CoreLogic, and matched HMDA and CoreLogic data.
Notes: HMDA = Home Mortgage Disclosure Act; ODR = observed denial rate; RDR = real denial rate. The analysis is limited to owner-occupied purchase mortgage applications. Loan applications in 2006 and 2017 are used to illustrate how the RDR is calculated. The raw data for other races or ethnicities, channels, and origination years used for calculating the RDR is available upon request.
[a] Includes both originated loans and loan applications approved by the lenders but not accepted by the applicants. The latter accounts for less than 10% of approved applications.
[b] See the Methodology and Data section for the definition of low credit profiles.
[c] Borrowers with high credit profiles have no chance of being denied a loan application.

[5] Our categorization of denials and approvals is as follows: denied = denied; application or preapproval request approved but not accepted = approved; loan originated = approved. We excluded loans purchased by a financial institution. Because only originated HMDA loans can be matched with CoreLogic loans, we assume approved but not originated applications have the same share of LCP applicants as originated loans.

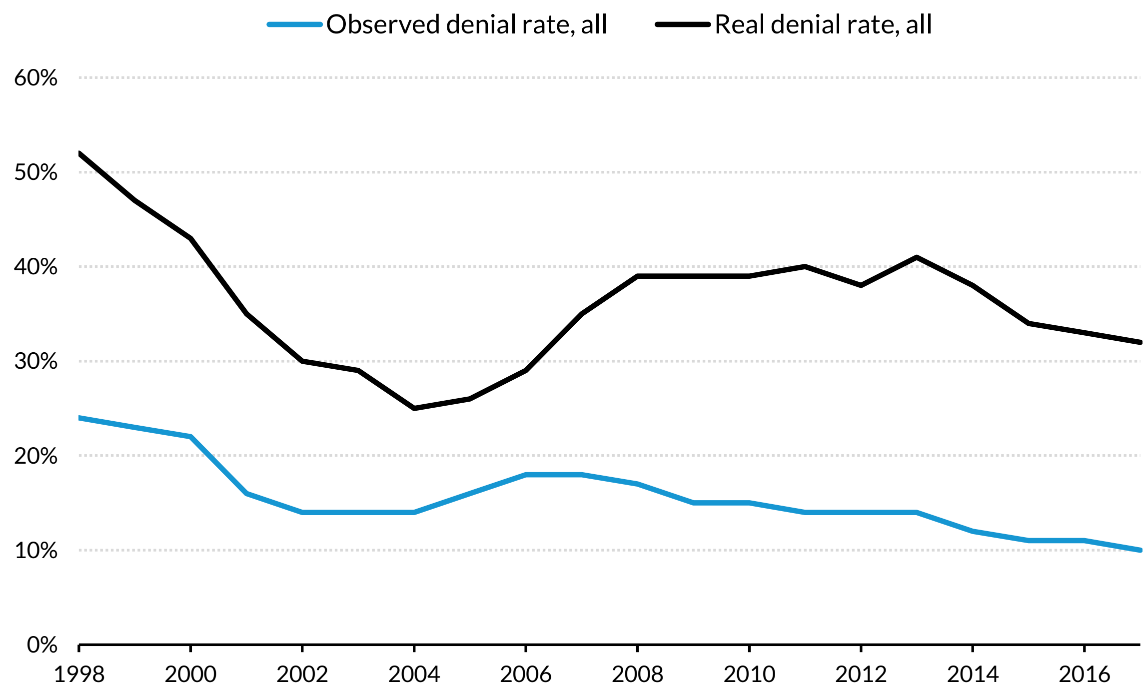
Our matched HMDA and CoreLogic data indicate that of the 5,559,643 approved loans, 2,961,006 (53%) were from LCP consumers and 2,598,637 (47%) were from HCP consumers. Because HCP consumers, by our definition, have a zero probability of default, 4,180,796 (the 6,779,433 applications minus the 2,598,637 from HCP consumers) applications are from LCP borrowers. All denied applications, by definition, come from the LCP pool, so the RDR for LCP applications is 1,219,790 divided by 4,180,796, or 29%.

The difference between the RDR of 29% and the ODR of 18% reflects the fact that, in our calculation of the RDR, we have reduced the denominator to include only the 53% of the applicants who are LCP; that is, we have excluded the 47% of applicants who are HCP. In fact, ODRs understate the difficulty of applicants with marginal credit obtaining a mortgage; the RDR is a more accurate measure.

The Real Denial Rate versus the Observed Denial Rate over Time

Table 1 shows the ODR and RDR calculation for all applicants in 2006 and 2017. The ODR for 2006 (18%) is higher than the ODR in 2017 (11%). This result suggests that credit was tighter in 2006 than in 2017, which runs counter to our expectations. We would expect denial rates to be lower during the housing boom, when lenders approved loans they would not have approved in a tighter lending environment, such as that prevailing a decade after the crisis. Changes in applicants' credit profiles explain the counterintuitive results. In 2006, 62% of loans were to LCP applicants, versus 31% in 2017. In the boom years, more LCP applicants were encouraged to submit applications; thus, there were more rejections. As the credit box tightened after the financial crisis, many LCP borrowers were discouraged from applying, leading to fewer rejections. Figure 1 shows the ODR over time. The rate peaked in 2006 and 2007, the period in which we think of credit as being the loosest, and has come down steadily since then.

Figure 1. Observed versus Real Denial Rates, 1998–2017



Sources: Home Mortgage Disclosure Act, CoreLogic, and the Urban Institute.
Note: Based on owner-occupied purchase mortgage applications.

Because the RDR measures denial rates for only LCP applicants, it reveals a more intuitive pattern. The RDR is 32% in 2017 versus 29% in 2006. More precisely, the RDR rose sharply postcrisis, peaked at 41% in 2013, and has declined over the past few years, reflecting the loosened credit box.

Table 2 shows the share of LCP applicants, which has decreased steadily since the financial crisis. The share of LCP applicants was 49% from 1998 to 2004, 58% from 2005 to 2007, 39% from 2008 to 2010, and 32% from 2011 to 2017.

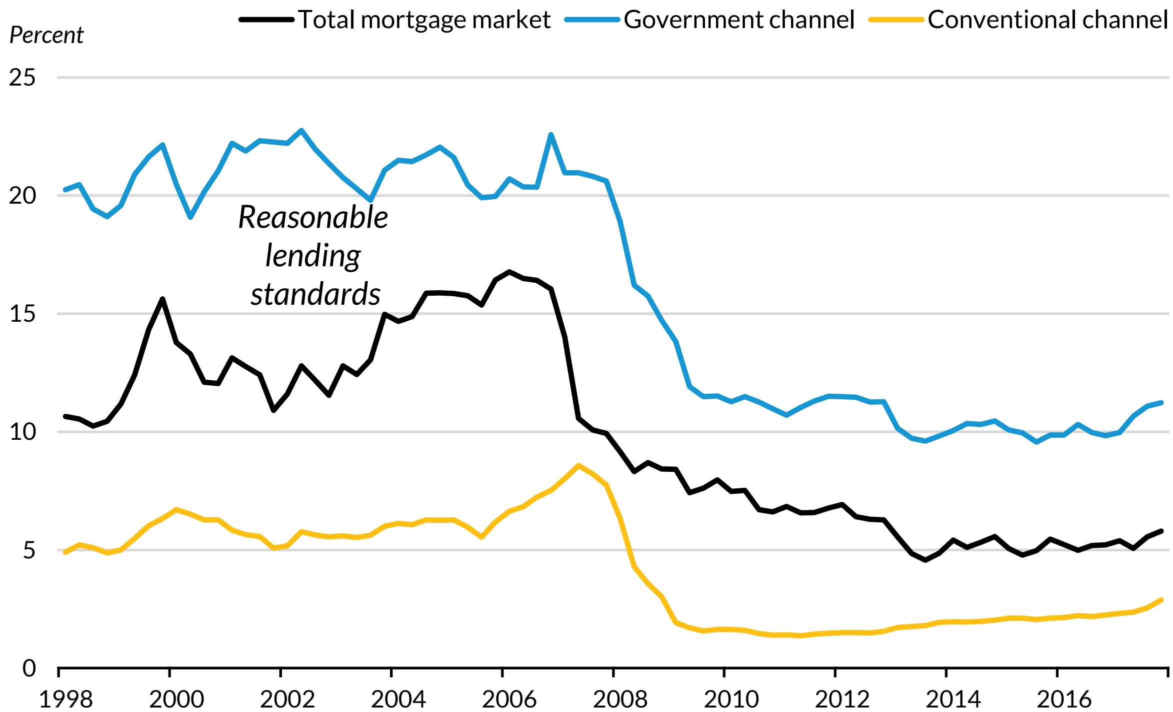
Table 2. Observed Denial Rate versus Real Denial Rate and Share of Low-Credit-Profile Applicants and Borrowers in All Channels

	Denial Rates		LCP Shares	
	ODR all	**RDR all**	**LCP applicants**	**LCP borrowers**
1998	24%	52%	47%	30%
1999	23%	47%	49%	34%
2000	22%	43%	50%	37%
2001	16%	35%	45%	35%
2002	14%	30%	46%	37%
2003	14%	29%	48%	40%
2004	14%	25%	55%	48%
2005	16%	26%	60%	52%
2006	18%	29%	62%	53%
2007	18%	35%	53%	42%
2008	17%	39%	43%	31%
2009	15%	39%	38%	27%
2010	15%	39%	37%	26%
2011	14%	40%	36%	25%
2012	14%	38%	36%	26%
2013	14%	41%	33%	23%
2014	12%	38%	33%	23%
2015	11%	34%	32%	24%
2016	11%	33%	32%	24%
2017	11%	32%	31%	24%
1998–2004	18%	37%	49%	37%
2005–2007	17%	30%	58%	49%
2008–2010	16%	39%	39%	28%
2011–2017	12%	36%	32%	24%

Sources: Home Mortgage Disclosure Act, CoreLogic, and the Urban Institute.
Notes: LCP = low-credit-profile; ODR = observed denial rate; RDR = real denial rate. Based on owner-occupied purchase mortgage applications.

But the RDR indicates the credit box has loosened, a pattern we pick up in the Urban Institute's Housing Credit Availability Index, or HCAI (Urban Institute 2018). This index measures the ex ante probability of default of mortgages underwritten in any given period (Figure 2). The RDR and HCAI show the same pattern: loose credit from 2005 to 2007, a dramatic tightening until 2013, and a marginal loosening since. But the HCAI shows the market is taking less than half the credit risk it was taking before the crisis. The RDR explains why.

Figure 2. Default Risk Taken by the Mortgage Market



Sources: eMBS, CoreLogic, Home Mortgage Disclosure Act, Inside Mortgage Finance, and the Urban Institute.

After controlling for the variability in the applicant mix through the boom and bust, the RDR analysis shows that the real denial rates were similar to what they were in the pre-bubble period (that is, 36% for 2011–17 is similar to 37% for 1998–2004). Table 2 shows that the share of LCP applicants is lower, as fewer marginal applicants are applying for loans. From 2011 to 2017, 32% of applicants were LCP, but from 1998 to 2004, 49% of applicants were LCP.

The RDR More Accurately Reflects Credit Differentials by Channel

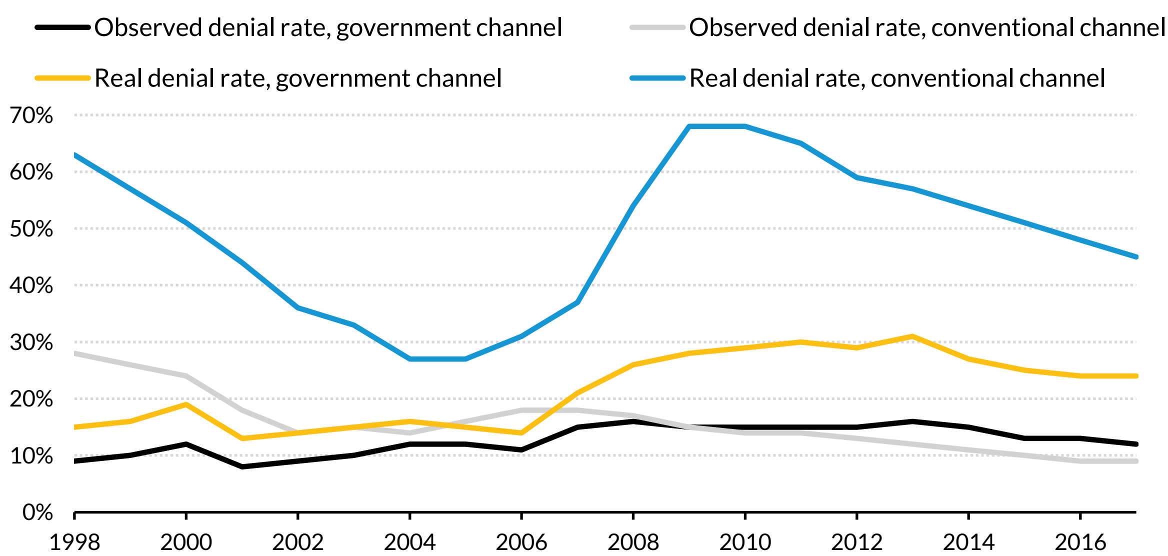
At the point of loan application, a borrower chooses either a government mortgage or a conventional mortgage. The government channel includes loans insured by the Federal

Housing Administration, the U.S. Department of Veterans Affairs, the U.S. Department of Agriculture, or the Office of Public and Indian Affairs within the U.S. Department of Housing and Urban Development. The conventional channel includes executions by the government-sponsored enterprises (GSEs), bank portfolio, and private-label securities. In the post-bubble years, as the private-label securities market has all but disappeared, the GSEs (Fannie Mae and Freddie Mac) are the main issuers in the conventional market.

Because of its low–down payment requirements, the government channel has traditionally been used to a disproportionate extent by low- and moderate-income borrowers and minority consumers, and we would assume it would be easier to qualify for a government loan than for a conventional loan. Therefore, we would assume denial rates in the government channel would be lower than in the conventional channel.

The ODR measure in Figure 3 confirms this was the case before the financial crisis. After the crisis, an ODR analysis suggests that the conventional channel had lower denial rates than the government channel.

Figure 3. Observed versus Real Denial Rates in the Government and Conventional Channels



Sources: Home Mortgage Disclosure Act, CoreLogic, and the Urban Institute.
Note: Based on owner-occupied purchase mortgage applications.

Credit profile changes in the loan applicant pool explain these counterintuitive results. Table 3 shows that the average share of LCP applicants in the conventional channel are 45% in the pre-bubble years of 1998 to 2004, 56% in the bubble years of 2005 to 2007, 25% in the crisis years of 2008 to 2010, and 20% in the postcrisis years of 2011 to 2017. Low-credit-profile shares in the government channel were 65, 77, 55, and 52% in those periods, respectively. Following the crisis, the conventional channel changed its pricing to be more risk based, while the government channel does not use risk-based pricing. The GSEs imposed loan-level pricing adjustments, a system of up-front risk-based charges. The private mortgage insurers recalibrated their risk models to reflect greater differentiation by risk bucket. (The GSEs, by charter, cannot be in a first-loss position on any loan with an LTV ratio over 80%; further credit enhancement is required. Private mortgage insurance comprises the overwhelming majority of this additional credit enhancement). Moreover, the GSEs and their regulator, the Federal Housing Finance Agency (FHFA), imposed risk-based capital charges on the mortgage insurers with their adoption of Private Mortgage Insurer Eligibility Requirements. These rules went into effect in 2015. These requirements, which must be adhered to for a mortgage insurer to do business with the GSEs, further increased the risk-based adjustments. It is now more economical for LCP borrowers to apply for mortgages through the government channel rather than through the conventional channel, leading to few LCP applicants in the conventional channel.

Table 3. Share of Low-Credit-Profile Applicants and Borrowers by Channel

	Low-Credit-Profile Applicants			Low-Credit-Profile Borrowers		
	Government	**Conventional**	**All**	**Government**	**Conventional**	**All**
1998	57%	44%	47%	53%	23%	30%
1999	60%	46%	49%	56%	27%	34%
2000	62%	47%	50%	57%	30%	37%
2001	63%	40%	45%	59%	27%	35%
2002	67%	41%	46%	64%	31%	37%
2003	70%	44%	48%	66%	35%	40%
2004	76%	53%	55%	73%	45%	48%
2005	78%	58%	60%	74%	50%	52%
2006	79%	60%	62%	76%	51%	53%
2007	73%	50%	53%	68%	39%	42%
2008	62%	32%	43%	54%	17%	31%
2009	52%	22%	38%	44%	8%	27%
2010	51%	21%	37%	43%	8%	26%
2011	51%	21%	36%	42%	8%	25%
2012	54%	21%	36%	45%	10%	26%
2013	53%	21%	33%	43%	10%	23%
2014	54%	20%	33%	46%	10%	23%
2015	52%	20%	32%	45%	11%	24%
2016	51%	20%	32%	44%	11%	24%
2017	51%	20%	31%	45%	12%	24%
1998–2004	65%	45%	49%	61%	31%	37%
2005–2007	77%	56%	58%	73%	47%	49%
2008–2010	55%	25%	39%	47%	11%	28%
2011–2017	52%	20%	32%	45%	11%	24%

Sources: Home Mortgage Disclosure Act, CoreLogic, and the Urban Institute.
Note: Based on owner-occupied purchase mortgage applications.

The RDR consistently conforms to our intuition. Referring back to Figure 3, the conventional channel consistently has a higher RDR than the government channel, but the two curves have the least differential during the bubble years. This makes sense because during the bubble years, conventional underwriting standards declined as nontraditional products (e.g., interest-only mortgages, 40-year mortgages, mortgages with negative amortization, and mortgages with an initial “teaser” payment and a reset period shorter than five years) composed a significant portion of originations. In contrast, nontraditional products remained a small part of government origination.

Loosening by Channel: Timing Differences

Figure 3 shows that although the conventional channel continues to have higher RDRs than the government channel, both have experienced declines since 2013. Moreover, the RDR for conventional loans started to drop in 2012, while the RDR for government loans started to drop in 2014. The same loosening patterns can be seen in our HCAI (Figure 2), which shows that the credit availability for GSE loans began rising in 2012, while government loans did not experience any increase until 2014.

Why did GSE loans show a loosening in credit earlier? Lenders were putting overlays on both the GSE and FHA credit box. That is, lenders were imposing more stringent standards on GSE and FHA loans than what the agencies required, as the lenders were afraid they would be forced to repurchase the GSE loans—a frequent occurrence in 2009–12—or would be sued by the government for triple damages for defective FHA loans. The GSEs took steps earlier to assure lenders that the lender would be responsible for defects in the manufacturing of the loan but would not be responsible for subsequent borrower performance. Starting in 2012, the GSEs and their conservator and regulator, the FHFA, clarified the standards for mortgage repurchases. This included the introduction of sunsets for the representation and warranty provisions in 2012, the relaxation of the sunset eligibility requirement and the clarification of life-of-loan exclusions in 2014, the introduction of the taxonomy that grades defects and spells out the level of severity necessary for repurchase in 2015, and the independent resolution process in 2016. In 2016, Fannie Mae released its Day 1 Certainty program, waiving certain representations and warranties at origination. Freddie Mac followed with its Loan Advisor Suite in 2017, which performs a similar function. (For a complete description of the steps the GSEs and the FHFA took, see Goodman, 2017).

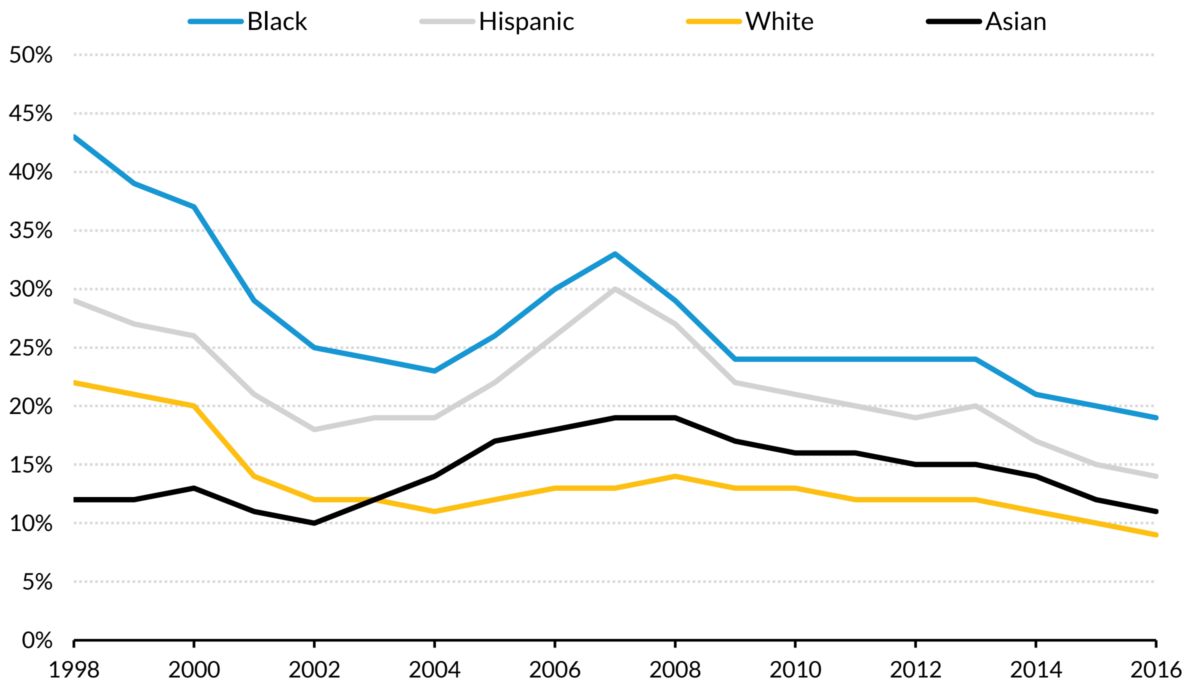
The FHA, the largest participant in the government channel, has lagged behind the GSEs and the FHFA in its attempts to reduce lender uncertainty. The FHA published the Single-Family Housing Policy Handbook in March 2015, the FHA defect taxonomy in June 2015, and the Supplemental Performance Metric in August 2015. The handbook puts together more than 900 FHA-issued mortgagee letters, eliminating inconsistencies. The Supplemental Performance Metric assures lenders they will not be shut down if they have a riskier book of business than their peers. But these measures are not enough to counteract the effect of the False Claims Act, a powerful tool the U.S. Department of Justice has used to pursue expensive claims against mortgage originators (Goodman, 2015). The FHA's most promising tool to reassure lenders is the defect taxonomy, which outlines various errors at origination and grades their severity (FHA Office of Single Family Housing, 2015). But the taxonomy does not provide remedies for each error type and contains no mechanism to allow the FHA to rely on the taxonomy or tie it to actions against mortgage lenders under the False Claims Act.

Much of the opening of the government credit box has been because of the increasing role of nonbank originators. Since 2013, the nonbank share of government originations has increased from 35% to 81%, and the nonbank share of GSE originations has increased from 35% to 56% (Goodman et al., 2018). These originators are less concerned about the reach of the False Claims Act because they have less at stake—less of an established reputation, usually in only one business so that there is no reputational impact on other activities, and less capital. The nonbanks have opened up the credit box. The median bank FICO score for government loans in March 2018 was 696, and the median nonbank score was 674. The median bank DTI ratio for government loans in March 2018 was 40.8%, and the median nonbank DTI ratio was 42.6% (Ginnie Mae, 2018a; Goodman et al. 2018).

The RDR Shows Smaller Gaps by Race and Ethnicity in Denial Rates

The ODR indicates that denial rates are consistently highest for blacks and Hispanics and are lower for non-Hispanic whites (hereafter "whites") and Asians (Figure 4A). In 2017, the ODR indicates that black applicants had twice the denial rates as white applicants, Hispanic applicants had 1.4 times the denial rate, and Asian applicants had 1.2 times the denial rate. Some news articles have used these ODRs at the local level to allege redlining in mortgage lending (Glantz & Martinez 2018). Our RDR analysis shows that this claim is a gross oversimplification.

Figure 4A. Observed Denial Rates by Race and Ethnicity



Sources: Home Mortgage Disclosure Act, CoreLogic, and the Urban Institute.
Note: Based on owner-occupied purchase mortgage applications.

The differences in denial rates across groups primarily reflect differences in credit characteristics. In 2017, the average LCP share was 47% among black applicants, 40% among Hispanic applicants, 29% among white applicants, and 24% among Asian applicants (Appendix Tables A2 and A3).

More intuitively, Table 4 shows the median characteristics by racial and ethnic group for approved loans in 2017, based on the HMDA-CoreLogic matched data. The median black and Hispanic borrowers have lower FICO scores, higher LTV and DTI ratios, and lower incomes than their white counterparts. The median FICO score is 697 for black borrowers, 708 for Hispanic borrowers, and 738 for white borrowers. The median LTV ratio for black owner occupants is 97%, the median for Hispanics is 95%, and the median for whites is 89%. These figures apply only to borrowers who got mortgages approved and originated because we cannot observe credit profiles for denied applicants. Median applicants are likely to have weaker credit profiles, as applicants tend to have more LCP consumers than borrowers for all races and ethnicities (Appendix Table A3).

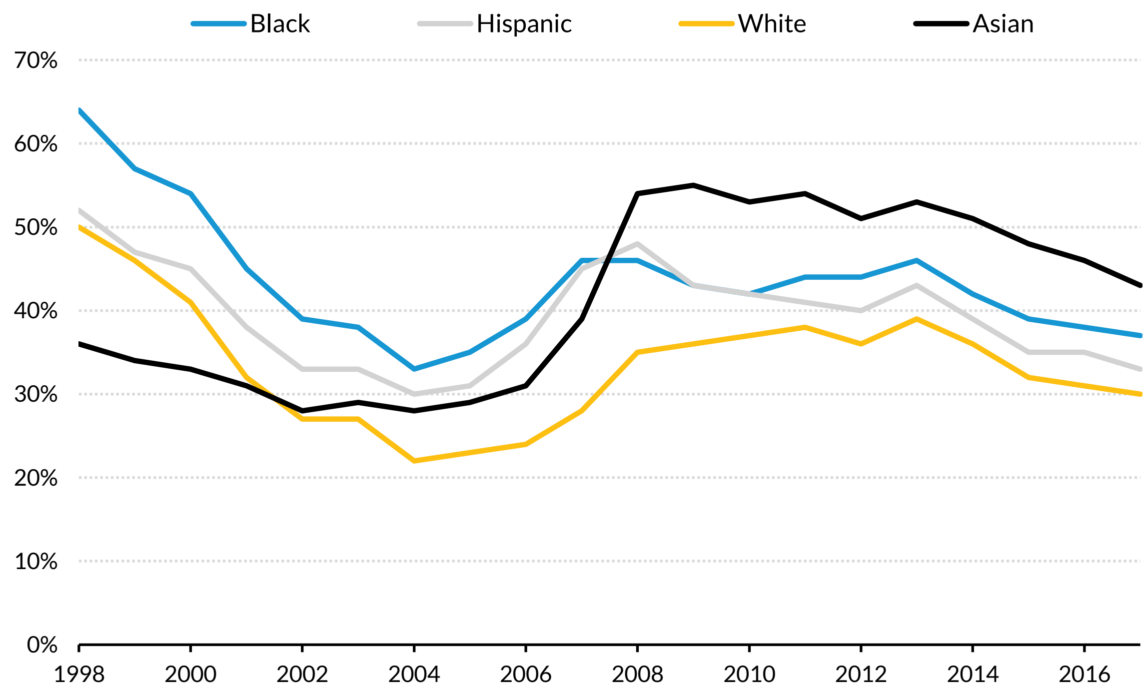
Table 4. Borrower Characteristics by Race and Ethnicity for 2017 Purchase Mortgage Originations

	Black	Hispanic	White	Asian	All
Median FICO score	696.7	708.2	738.0	751.3	733.9
Median LTV ratio	96.7	95.0	88.7	81.8	89.8
Median DTI ratio	39.0	39.0	36.4	37.0	37.0
Income	$67,000	$65,000	$80,000	$102,000	$78,000
Income/area median income	0.9	1.0	1.2	1.3	1.1

Sources: Home Mortgage Disclosure Act, CoreLogic, and the Urban Institute.
Notes: DTI = debt-to-income; LTV = loan-to-value. Based on owner-occupied purchase mortgage originations.

Given the differences in credit characteristics, the RDR shows smaller racial and ethnic gaps (Figure 4B). In 2017, black applicants had 1.2 times the denial rate of white applicants, Hispanic applicants had 1.1 times the denial rate, and Asian applicants had 1.4 times the denial rate. Our RDR analysis shows that once we control for differences in credit characteristics, racial and ethnic differences in denial rates get smaller. We controlled for loan-to-value ratio, credit score, debt-to-income ratio, and product and documentation type, but we did not control for income and did not have data on assets or reserves, which are factors in underwriting. Thus, we would not have completely controlled for credit differentials.

Figure 4B. Real Denial Rates by Race and Ethnicity

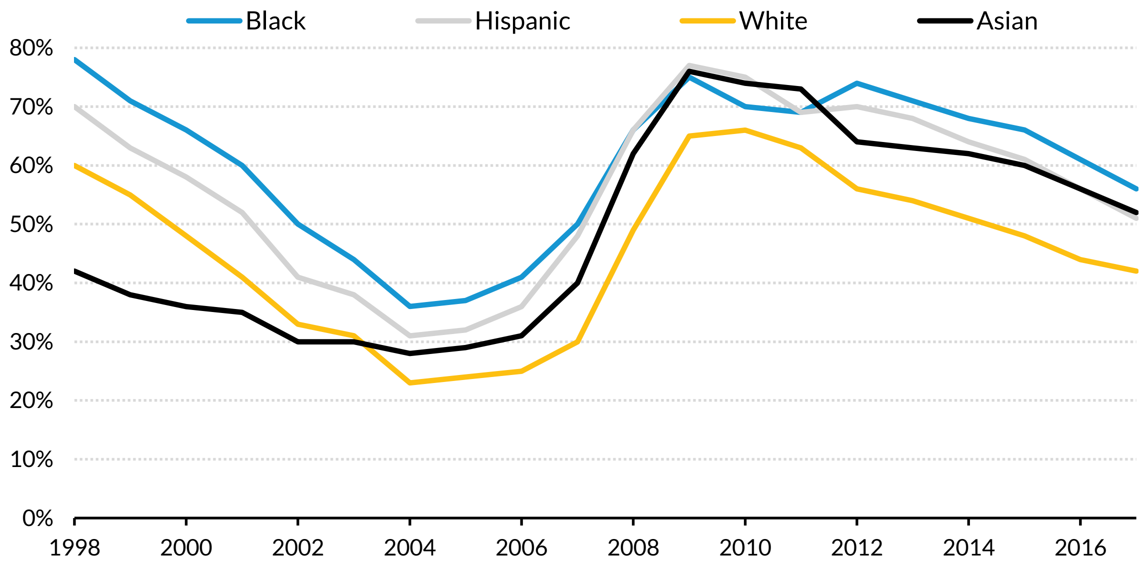


Sources: Home Mortgage Disclosure Act, CoreLogic, and the Urban Institute.
Note: Based on owner-occupied purchase mortgage applications.

The RDR is highest for Asians because Asians frequent the conventional channel more than other groups, and the conventional channel has higher denial rates.

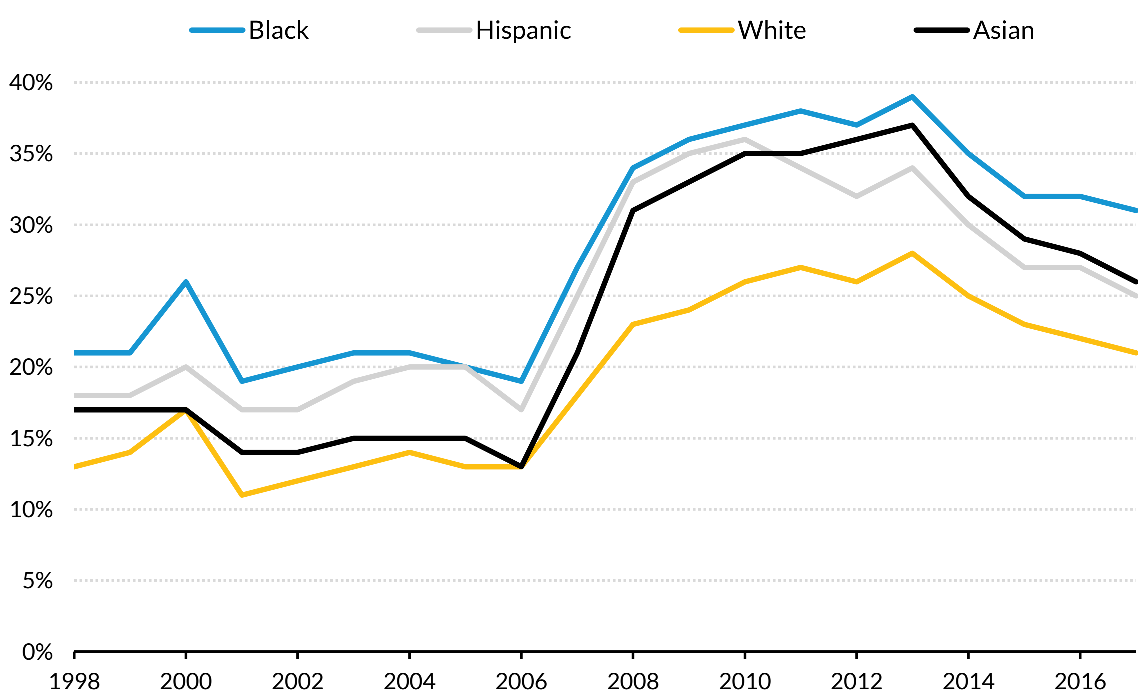
Figure 5 shows the RDR by channel and race and ethnicity. In both channels, Asian applicants have RDRs just above but close to those for Hispanic applicants in 2017 (52% for Asians and 51% for Hispanics in the conventional channel; 26% for Asians and 25% for Hispanics in the government channel). The discrepancy in figure 4B arises because 84% of the Asian applicants use the conventional channel compared with 48% of Hispanic applicants. We believe this is because Asian borrowers tend to live in high-cost coastal areas that rely more heavily on conventional loans (which have higher denial rates than government loans).

Figure 5A. Real Denial Rates by Race and Ethnicity in the Conventional Channel



Sources: Home Mortgage Disclosure Act, CoreLogic, and the Urban Institute.
Note: Based on owner-occupied purchase mortgage applications.

Figure 5B. Real Denial Rates by Race and Ethnicity in the Government Channel



Sources: Home Mortgage Disclosure Act, CoreLogic, and the Urban Institute.
Note: Based on owner-occupied purchase mortgage applications.

The appendix contains detailed tables for ODRs, RDRs, and shares of LCP applicants and borrowers, sorted by channel and race or ethnicity. (Appendix Tables A4 and A5 contain the results for the conventional channel, and Appendix Tables A6 and A7 contain the results for the government channel.) It is clear from these tables that the real denial rates show smaller racial and ethnic gaps than do the observed denial rates, controlling for channel; the convergence is more pronounced in the conventional channel. For example, comparing Hispanics with whites in the conventional channel, the RDR ratio is 1.2, much closer than the 1.6 ODR ratio.

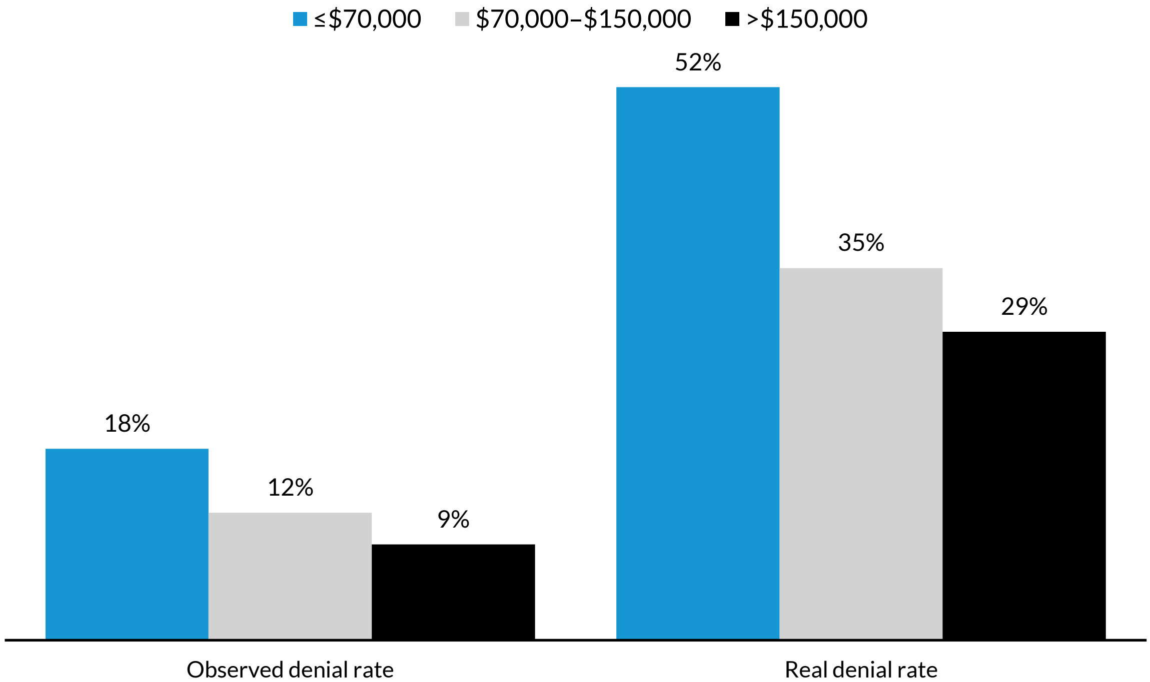
The RDR Shows Higher Denial Rates for Small-Dollar Mortgages

Recent research has documented the dearth of small-dollar mortgages for low-cost single-family residential home purchases, limiting affordable homeownership opportunities for creditworthy families living in low-cost, underserved housing markets. McCargo, Bai, George, and Strochak (2018) reveal that only a quarter of homes sold for $70,000 or less were financed through a mortgage, while almost 80% of homes worth between $70,000 and $150,000 were bought with a mortgage in 2015.

Figure 6 shows that the ODR for small-dollar mortgages (up to $70,000) is 18%, double that for larger loans (more than $150,000) in 2017.[6] There is little variation in applicants' credit profile compositions by loan size: the share of LCP applicants was 34% for loans up to $70,000, 35% for loans between $70,000 and $150,000, and 30% for loans more than $150,000. After controlling for applicant credit profiles, the RDR gap remains large across the three loan size buckets. In 2017, the RDR for small loans (up to $70,000) is 52%, compared with 29% for large loans (more than $150,000).

[6] Earlier years show denial rate patterns similar to 2017. Data are available upon request.

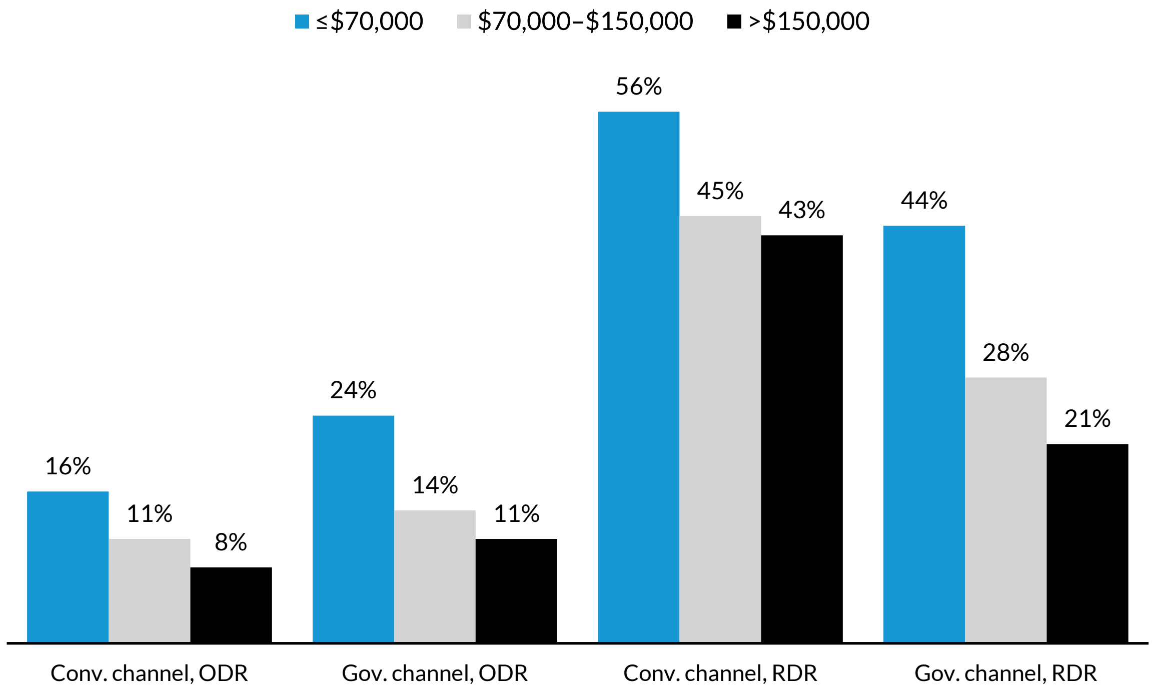
Figure 6. Denial Rates by Loan Size among 2017 Purchase Mortgage Applications



Sources: Home Mortgage Disclosure Act, CoreLogic, and the Urban Institute.
Note: Based on owner-occupied purchase mortgage applications.

What's behind the high RDRs for small-dollar mortgages? Figure 7 shows that small-dollar mortgages have higher RDRs than larger mortgages within each channel. The RDR for conventional mortgages is 56% for mortgages up to $70,000, 1.3 times higher than the 43% for mortgages more than $150,000. The gap between the two loan size groups is more pronounced in the government channel: 44% for the small-loan group versus 21% for the large-loan group. Small loans in the government channel have an RDR 2.09 times those for larger loans. The RDR for small-dollar government mortgages is slightly higher than the RDR for conventional mortgages over $150,000, despite that fact that overall, the conventional channel has considerably higher RDRs than the government channel.

Figure 7. Denial Rates by Loan Size and Channel among 2017 Purchase Mortgage Applications



Sources: Home Mortgage Disclosure Act, CoreLogic, and the Urban Institute.
Notes: ODR = observed denial rate; RDR = real denial rate. Based on owner-occupied purchase mortgage applications.

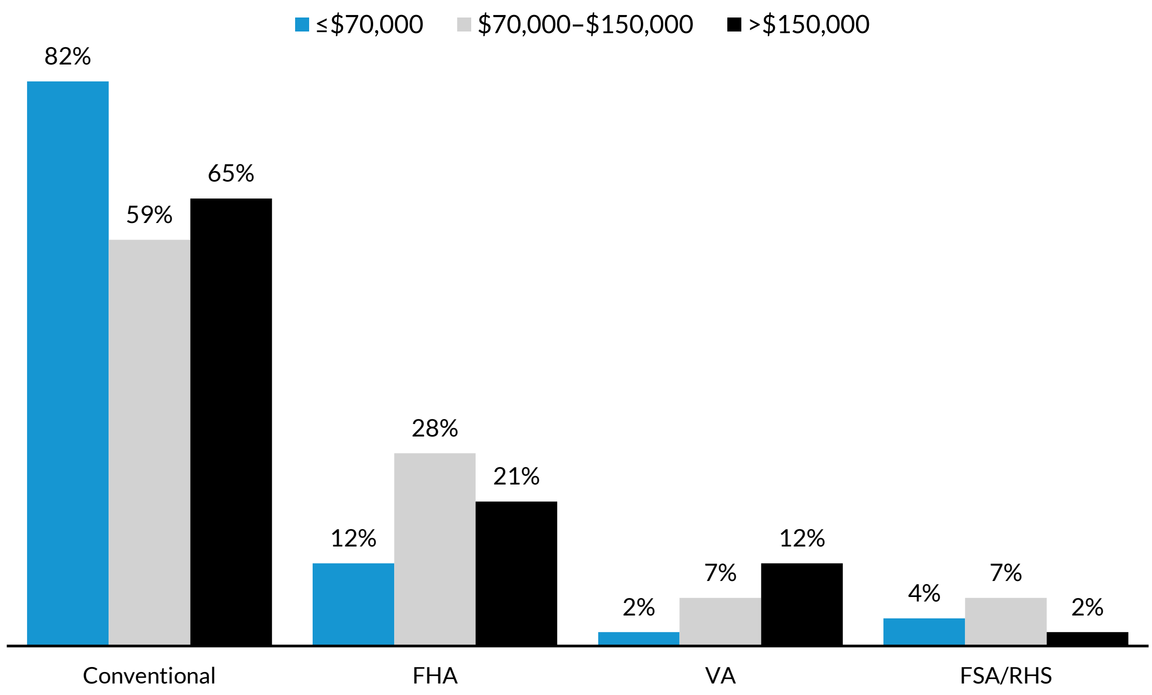
Moreover, small loans are overrepresented in the conventional channel, which has higher RDRs. Figure 8 shows that conventional loans compose 82% of small-dollar applications but only 65% of mortgage applications more than $150,000. The FHA serves 12% of the small-dollar mortgage market, 28% of the $70,000–$150,000 market, and 21% of the over-$150,000 market. Part of the reason for this is that the fixed costs of originating a loan and servicing a loan makes small loans less attractive to originate (McCargo et al., 2018). But banks, which have moved away from the FHA market[7] and are more likely to originate conventional loans, have certain requirements under the Community Reinvestment Act (CRA). They must meet the credit needs of the communities in which they operate, including the needs of low- and moderate-income neighborhoods, as long as it can be done in

[7] The nonbank share of FHA purchase mortgage originations has risen from 34% in 2013 to 83% in the first quarter of 2018. See Ginnie Mae (2018b, p. 31).

a manner consistent with safe and sound operations. Many of these small loans would "count" for CRA purposes, giving banks an incentive that compensates them for the higher fixed costs of origination and servicing. But nonbanks have no CRA obligations. In addition, many small-balance conventional loans are held in the portfolios of lending institutions, often small banks and credit unions that serve rural communities.

The largest gaps in market share of small loans relative to large loans are in Veterans Administration lending. The VA financed only 2% of small-dollar purchase mortgages and 12% of purchase loans more than $150,000. Lending for low-cost properties might be particularly affected by the VA's residual income test, which could eliminate many low-income borrowers (more apt to need a smaller loan to buy a less expensive home) who are either discouraged from applying or fail to qualify under this test.

Figure 8. Market Share by Channel and Loan Size among 2017 Purchase Mortgage Applications



Sources: Home Mortgage Disclosure Act, CoreLogic, and the Urban Institute.
Notes: FHA = Federal Housing Administration; FSA/RHS = Farm Service Agency or Rural Housing Service; VA = U.S. Department of Veterans Affairs. Based on owner-occupied purchase mortgage applications.

The higher RDRs for small loans come from the fact that (1) most of these loans are conventional, which have higher RDRs than the government channel, and within the conventional channel, small loans have a higher denial rate than large loans, and (2) within the government channel, the denial rate for loans up to $70,000 is more than twice that for loans more than $150,000.

Conclusion

This analysis, using recent and historical HMDA data matched with CoreLogic data, has shown that the real denial rate is higher than the traditional observed denial rate. By considering only borrowers with low credit profiles, the real denial rate reduces the distortion varying credit profiles have on the observed denial rate. It shows a more intuitive pattern over time. Our analysis indicates that the real denial rate peaked in 2013 and has been dropping ever since. But the share of low-credit-profile borrowers has been steadily declining. Thus, the increase in credit availability reflects a decrease in the denial rate of those applying, not a broadening of the applicant pool. That is, although the denial rate among low-credit applicants has declined to precrisis levels, low-credit applicants account for a smaller share of applicants in 2017 (31%) than they did from 1998 to 2004 (49%).
Our real denial rate shows that government loans have lower denial rates than conventional loans. The denial rates for conventional loans started to decrease before denial rates for FHA loans, as the GSEs were more aggressive in assuring lenders that they are responsible only for loan manufacturing defects, not the borrower's subsequent performance. These findings are consistent with our Housing Credit Availability Index.

Our updated results also show that the racial and ethnic disparities in traditional observed denial rates are in large part because of differences in credit profiles. These differences include differences in credit scores and in debt-to-income and loan-to-value ratios. Using observed denial rates to judge whether redlining has occurred, as many recent news articles have done, is misleading. When one constructs real denial rates, as we have, the differences by race and ethnicity are narrower. And we have not accounted for all the factors that go into loan underwriting, which could explain some of the remaining differentials.

Finally, we demonstrate that real denial rates are higher for small-dollar mortgages (up to \$70,000) than for larger loans. This is especially true in the government channel, which is underrepresented in the small-dollar mortgage market.

Appendix

Table A1. Probability of Being a Low-Credit-Profile Borrower, Calculated from the Expected Default Risk of the Loans

		Nonrisky Loan Products							Risky Loan Products						
		Avg.	>740	(700,740]	(660,700]	(620,660]	(580,620]	≤580	Avg.	>740	(700,740]	(660,700]	(620,660]	(580,620]	≤580
DTI	CLTV	18%	6%	16%	24%	38%	57%	80%	81%	49%	76%	87%	93%	95%	96%
Average	Average														
	(0,68]	3%	0%	2%	5%	13%	22%	39%	21%	4%	18%	28%	38%	42%	58%
	(68,78]	9%	2%	9%	16%	25%	33%	58%	51%	21%	45%	60%	72%	70%	84%
	(78,82)	12%	4%	12%	19%	28%	34%	54%	74%	40%	65%	79%	94%	91%	99%
	[82,90]	18%	7%	17%	24%	32%	43%	69%	89%	62%	89%	95%	88%	95%	100%
	(90,95]	19%	7%	15%	21%	32%	44%	69%	82%	58%	76%	80%	87%	100%	100%
	>95	56%	29%	42%	52%	70%	93%	100%	97%	86%	94%	99%	100%	100%	100%
Low or no doc.	Average	30%	13%	30%	41%	55%	73%	93%	83%	57%	82%	89%	94%	90%	93%
	(0,68]	4%	0%	5%	11%	21%	34%	59%	22%	5%	20%	29%	38%	40%	58%
	(68,78]	16%	5%	19%	29%	41%	49%	82%	54%	26%	50%	64%	75%	68%	83%
	(78,82)	21%	10%	22%	33%	49%	53%	80%	78%	48%	71%	84%	98%	84%	99%
	[82,90]	35%	19%	36%	46%	53%	62%	86%	92%	71%	96%	100%	91%	99%	100%
	(90,95]	36%	21%	36%	42%	50%	61%	99%	84%	67%	83%	84%	90%	100%	100%
	>95	69%	51%	65%	65%	71%	88%	100%	100%	100%	100%	100%	100%	100%	100%
Full doc. & ≥50	Average	17%	5%	13%	23%	40%	52%	62%	88%	43%	71%	87%	95%	99%	97%
	(0,68]	3%	0%	0%	5%	14%	23%	27%	35%	9%	24%	37%	50%	63%	61%
	(68,78]	9%	1%	8%	14%	26%	34%	41%	60%	15%	45%	63%	84%	84%	84%
	(78,82)	11%	3%	10%	18%	29%	34%	37%	72%	31%	52%	63%	79%	100%	100%
	[82,90]	19%	7%	13%	24%	38%	47%	56%	91%	50%	78%	83%	85%	97%	100%
	(90,95]	22%	10%	16%	26%	39%	52%	61%	91%	60%	71%	81%	99%	100%	100%
	>95	53%	23%	33%	51%	78%	100%	100%	98%	74%	90%	100%	100%	100%	100%
Full doc. & [40,50)	Average	17%	4%	11%	21%	35%	54%	71%	86%	40%	65%	85%	93%	98%	98%
	(0,68]	2%	0%	0%	4%	11%	18%	27%	27%	3%	15%	28%	38%	50%	56%
	(68,78]	7%	0%	5%	12%	21%	28%	40%	51%	13%	31%	47%	64%	77%	87%
	(78,82)	10%	2%	8%	15%	23%	34%	45%	73%	27%	47%	65%	85%	100%	100%
	[82,90]	15%	4%	9%	18%	28%	40%	64%	85%	37%	58%	74%	78%	90%	100%
	(90,95]	17%	6%	11%	19%	31%	43%	60%	83%	43%	59%	70%	81%	100%	100%
	>95	51%	20%	30%	47%	73%	100%	100%	96%	65%	83%	100%	100%	100%	100%
Full doc. & [30,40)	Average	12%	1%	6%	15%	29%	50%	69%	74%	26%	51%	72%	89%	97%	97%
	(0,68]	1%	0%	0%	1%	9%	17%	27%	19%	1%	11%	21%	40%	43%	50%
	(68,78]	4%	0%	2%	7%	16%	27%	42%	38%	9%	23%	35%	55%	76%	85%
	(78,82)	5%	0%	4%	10%	17%	27%	40%	57%	18%	37%	52%	76%	100%	100%
	[82,90]	8%	1%	5%	12%	21%	33%	53%	76%	27%	50%	68%	76%	84%	100%
	(90,95]	12%	2%	6%	14%	25%	37%	55%	74%	33%	53%	65%	77%	100%	100%
	>95	41%	12%	21%	38%	66%	98%	100%	89%	47%	69%	91%	100%	100%	100%
Full doc. & (0,30)	Average	6%	0%	3%	9%	22%	41%	63%	58%	13%	35%	60%	83%	94%	96%
	(0,68]	1%	0%	0%	0%	6%	15%	30%	11%	0%	6%	11%	31%	45%	58%
	(68,78]	2%	0%	0%	4%	11%	22%	39%	26%	3%	13%	30%	45%	64%	86%
	(78,82)	2%	0%	1%	5%	12%	21%	34%	42%	9%	26%	45%	70%	98%	98%
	[82,90]	5%	0%	3%	8%	18%	29%	51%	71%	22%	46%	59%	71%	86%	100%
	(90,95]	9%	0%	4%	11%	23%	35%	55%	68%	23%	43%	63%	76%	100%	100%
	>95	36%	8%	18%	32%	59%	92%	100%	84%	33%	59%	86%	100%	100%	100%

Sources: Home Mortgage Disclosure Act, CoreLogic, and the Urban Institute.
Notes: CLTV = combined loan-to-value ratio; DTI = debt-to-income ratio. Based on owner-occupied purchase mortgage applications.

Table A2. Observed Denial Rate and Real Denial Rate in All Channels

	Observed Denial Rate				Real Denial Rate			
	Black	Hispanic	White	Asian	Black	Hispanic	White	Asian
1998	43%	29%	22%	12%	64%	52%	50%	36%
1999	39%	27%	21%	12%	57%	47%	46%	34%
2000	37%	26%	20%	13%	54%	45%	41%	33%
2001	29%	21%	14%	11%	45%	38%	32%	31%
2002	25%	18%	12%	10%	39%	33%	27%	28%
2003	24%	19%	12%	12%	38%	33%	27%	29%
2004	23%	19%	11%	14%	33%	30%	22%	28%
2005	26%	22%	12%	17%	35%	31%	23%	29%
2006	30%	26%	13%	18%	39%	36%	24%	31%
2007	33%	30%	13%	19%	46%	45%	28%	39%
2008	29%	27%	14%	19%	46%	48%	35%	54%
2009	24%	22%	13%	17%	43%	43%	36%	55%
2010	24%	21%	13%	16%	42%	42%	37%	53%
2011	24%	20%	12%	16%	44%	41%	38%	54%
2012	24%	19%	12%	15%	44%	40%	36%	51%
2013	24%	20%	12%	15%	46%	43%	39%	53%
2014	21%	17%	11%	14%	42%	39%	36%	51%
2015	20%	15%	10%	12%	39%	35%	32%	48%
2016	19%	14%	9%	11%	38%	35%	31%	46%
2017	18%	13%	9%	11%	37%	33%	30%	43%

Sources: Home Mortgage Disclosure Act, CoreLogic and the Urban Institute.
Note: Based on owner-occupied purchase mortgage applications.

Table A3. Share of Low-Credit-Profile Applicants and Borrowers in All Channels

	Applicants				Borrowers			
	Black	**Hispanic**	**White**	**Asian**	**Black**	**Hispanic**	**White**	**Asian**
1998	68%	57%	44%	33%	43%	39%	28%	24%
1999	68%	58%	47%	36%	48%	42%	32%	27%
2000	68%	58%	48%	39%	50%	44%	35%	30%
2001	64%	55%	42%	36%	49%	43%	33%	28%
2002	63%	56%	43%	37%	51%	46%	35%	30%
2003	65%	59%	45%	42%	53%	49%	38%	34%
2004	70%	64%	51%	51%	61%	56%	45%	43%
2005	72%	69%	55%	57%	63%	60%	49%	49%
2006	76%	73%	56%	58%	65%	63%	49%	49%
2007	71%	66%	47%	48%	56%	52%	39%	36%
2008	63%	56%	39%	35%	48%	40%	29%	20%
2009	57%	51%	35%	30%	43%	38%	26%	17%
2010	56%	50%	34%	29%	43%	37%	24%	16%
2011	55%	48%	33%	29%	41%	36%	23%	16%
2012	56%	49%	33%	30%	41%	36%	24%	17%
2013	52%	46%	31%	28%	37%	32%	21%	15%
2014	51%	44%	30%	27%	37%	33%	21%	15%
2015	50%	44%	30%	26%	38%	33%	22%	15%
2016	49%	41%	30%	25%	37%	32%	22%	15%
2017	47%	40%	29%	24%	36%	31%	22%	15%

Sources: Home Mortgage Disclosure Act, CoreLogic, and the Urban Institute.
Note: Based on owner-occupied purchase mortgage applications.

Table A4. Observed Denial Rate and Real Denial Rate in the Conventional Channel

	Observed Denial Rate				Real Denial Rate			
	Black	Hispanic	White	Asian	Black	Hispanic	White	Asian
1998	54%	38%	25%	12%	78%	70%	60%	42%
1999	49%	35%	24%	13%	71%	63%	55%	38%
2000	46%	33%	22%	13%	66%	58%	48%	36%
2001	37%	25%	15%	11%	60%	52%	41%	35%
2002	30%	21%	13%	10%	50%	41%	33%	30%
2003	27%	21%	13%	12%	44%	38%	31%	30%
2004	24%	19%	11%	14%	36%	31%	23%	28%
2005	26%	22%	13%	17%	37%	32%	24%	29%
2006	31%	26%	14%	18%	41%	36%	25%	31%
2007	35%	31%	13%	19%	50%	48%	30%	40%
2008	35%	31%	14%	19%	66%	66%	49%	62%
2009	33%	27%	13%	16%	75%	77%	65%	76%
2010	30%	24%	12%	15%	70%	75%	66%	74%
2011	29%	23%	12%	15%	69%	69%	63%	73%
2012	28%	21%	11%	14%	74%	70%	56%	64%
2013	26%	20%	11%	14%	71%	68%	54%	63%
2014	22%	17%	10%	13%	68%	64%	51%	62%
2015	21%	16%	9%	12%	66%	61%	48%	60%
2016	20%	14%	8%	11%	61%	56%	44%	56%
2017	18%	13%	8%	10%	56%	51%	42%	52%

Sources: Home Mortgage Disclosure Act, CoreLogic, and the Urban Institute.
Note: Based on owner-occupied purchase mortgage applications.

Table A5. Share of Low-Credit-Profile Applicants and Borrowers in the Conventional Channel

	Applicants				Borrowers			
	Black	**Hispanic**	**White**	**Asian**	**Black**	**Hispanic**	**White**	**Asian**
1998	69%	55%	41%	30%	32%	26%	22%	20%
1999	68%	55%	44%	33%	39%	31%	26%	23%
2000	69%	56%	45%	37%	42%	35%	29%	27%
2001	61%	49%	38%	33%	39%	32%	26%	24%
2002	59%	50%	38%	34%	42%	37%	29%	27%
2003	61%	55%	41%	40%	47%	43%	33%	32%
2004	68%	62%	49%	51%	57%	54%	42%	42%
2005	71%	68%	54%	57%	61%	59%	47%	49%
2006	75%	73%	54%	58%	64%	63%	47%	49%
2007	69%	65%	45%	47%	52%	49%	36%	35%
2008	53%	47%	28%	30%	28%	23%	17%	14%
2009	44%	35%	20%	22%	16%	11%	8%	6%
2010	43%	32%	19%	20%	19%	11%	7%	6%
2011	43%	33%	19%	21%	19%	13%	8%	7%
2012	38%	30%	20%	22%	14%	11%	10%	9%
2013	36%	29%	20%	21%	14%	12%	10%	9%
2014	33%	27%	19%	21%	13%	12%	10%	9%
2015	31%	26%	18%	19%	14%	12%	11%	9%
2016	32%	26%	19%	19%	16%	13%	11%	10%
2017	32%	26%	19%	19%	17%	15%	12%	10%

Sources: Home Mortgage Disclosure Act, CoreLogic, and the Urban Institute.
Note: Based on owner-occupied purchase mortgage applications.

Table A6. Observed Denial Rate and Real Denial Rate in the Government Channel

	Observed Denial Rate				Real Denial Rate			
	Black	**Hispanic**	**White**	**Asian**	**Black**	**Hispanic**	**White**	**Asian**
1998	13%	11%	7%	9%	21%	18%	13%	17%
1999	14%	11%	8%	10%	21%	18%	14%	17%
2000	17%	12%	11%	10%	26%	20%	17%	17%
2001	12%	11%	7%	9%	19%	17%	11%	14%
2002	14%	12%	8%	9%	20%	17%	12%	14%
2003	15%	14%	9%	10%	21%	19%	13%	15%
2004	17%	16%	10%	12%	21%	20%	14%	15%
2005	17%	16%	10%	11%	20%	20%	13%	15%
2006	16%	14%	10%	10%	19%	17%	13%	13%
2007	22%	19%	13%	15%	27%	25%	18%	21%
2008	24%	22%	13%	19%	34%	33%	23%	31%
2009	22%	20%	12%	17%	36%	35%	24%	33%
2010	22%	20%	13%	18%	37%	36%	26%	35%
2011	22%	19%	13%	18%	38%	34%	27%	35%
2012	23%	19%	13%	19%	37%	32%	26%	36%
2013	23%	19%	14%	20%	39%	34%	28%	37%
2014	21%	17%	13%	17%	35%	30%	25%	32%
2015	19%	15%	11%	15%	32%	27%	23%	29%
2016	18%	14%	11%	14%	32%	27%	22%	28%
2017	17%	13%	11%	13%	31%	25%	21%	26%

Sources: Home Mortgage Disclosure Act, CoreLogic, and the Urban Institute.
Note: Based on owner-occupied purchase mortgage applications.

Table A7. Share of Low-Credit-Profile Applicants and Borrowers in the Government Channel

	Applicants				Borrowers			
	Black	**Hispanic**	**White**	**Asian**	**Black**	**Hispanic**	**White**	**Asian**
1998	62%	58%	56%	57%	56%	53%	53%	52%
1999	65%	61%	59%	60%	59%	56%	56%	55%
2000	66%	62%	61%	60%	59%	57%	57%	56%
2001	66%	64%	62%	62%	62%	59%	59%	59%
2002	70%	69%	66%	67%	66%	65%	63%	64%
2003	73%	72%	68%	70%	68%	67%	65%	66%
2004	80%	78%	75%	76%	76%	74%	72%	73%
2005	82%	80%	76%	76%	78%	77%	73%	73%
2006	83%	81%	78%	77%	80%	78%	75%	75%
2007	79%	77%	71%	71%	74%	71%	67%	66%
2008	70%	66%	59%	61%	61%	57%	53%	52%
2009	60%	58%	50%	52%	49%	47%	43%	42%
2010	60%	57%	49%	51%	48%	46%	41%	41%
2011	59%	55%	48%	51%	47%	44%	41%	41%
2012	62%	58%	51%	54%	50%	48%	44%	43%
2013	60%	57%	51%	54%	47%	46%	42%	43%
2014	60%	57%	52%	55%	50%	49%	44%	45%
2015	59%	56%	50%	53%	49%	48%	44%	44%
2016	57%	53%	50%	50%	48%	46%	43%	42%
2017	56%	53%	49%	51%	47%	46%	43%	43%

Sources: Home Mortgage Disclosure Act, CoreLogic, and the Urban Institute.
Note: Based on owner-occupied purchase mortgage applications.

Matching HMDA and CoreLogic Loans

To obtain borrower credit profile information, we matched HMDA origination data to CoreLogic's proprietary loan-level databases (using both their private-label securities and servicing databases), which provide complementary information. HMDA is considered the "universe" of mortgage loans, as federal law requires that almost all mortgage originations be reported in HMDA (only a few small lenders are exempt). CoreLogic covers most of the residential mortgage market over the study period. To expand the size of the matched database beyond unique matches, we assigned weights to each matched HMDA-CoreLogic loan pair to reflect how close the match is, and we supplemented information in either database with information from the other using this weight.

We matched every HMDA loan to every CoreLogic loan to create a Cartesian product of the two databases. We first looked at each HMDA loan, filtering out CoreLogic loans where the common fields between the two databases were inconsistent with each other. First, if the loans were originated in different years, we did not include the pair in the matched loan database. Second, if a HMDA loan and a CoreLogic loan had a loan amount difference of at least $2,000, we dropped the pair. Third, loan pairs passed through a "geographic filter"–that is, a pair of loans with properties from different geographic locations was dropped out of the matched loan database. We required the census tract information from HMDA to be consistent with the zip code information from CoreLogic. We were left with loan pairs from the same issue year for the same amount with a geographic match. Because HMDA reports data by census tract and CoreLogic by zip code, the geographic filter is not straightforward. To solve this issue, we used the U.S. Department of Housing and Urban Development's zip code and census tract "cross-walk" file to match CoreLogic loans in a zip code to HMDA loans in a census tract and to assign geographic weights to the matched loans.

Suppose the *i*th HMDA loan from census tract X_i matched to the *j*th CoreLogic loan from zip code Y_j, $i = 1,\cdots I$, $j = 1,\cdots J$. X_i and Y_j overlap at Z_{ij}. Let X_i, Y_j, and Z_{ij} also denote the number of residential properties in each of the areas. The probability that the HMDA loan *i* is in Z_{ij} is given by

$$P_i = {}^{Z_{ij}}\!/_{X_i} \tag{A.1}$$

assuming that *i* has an equal chance of being located anywhere in X_i. Similarly, the probability that CoreLogic loan *j* is in Z_{ij} is given by

$$P_j = {}^{Z_{ij}}\!/_{Y_j} \tag{A.2}$$

The joint probability that both the HMDA loan *i* and the CoreLogic loan *j* are in Z_{ij} is given by

$$P_{ij} = {}^{Z_{ij}^2}\!/_{X_i Y_j} \tag{A.3}$$

which is the geographic weight for the matched loan pair of HMDA loan *i* and CoreLogic loan *j*.

For the other common variables between HMDA and CoreLogic, we adopted a fuzzy matching algorithm to filter out inconsistent pairs. The other common variables are loan type (e.g., FHA, VA, or conventional), loan purpose (e.g., purchase or refinance), occupancy, lien, and type of purchaser (e.g., Fannie Mae, Ginnie Mae, private-label securities, or portfolio). But for the same common variable, HMDA and CoreLogic might be coded differently. Moreover, both data sources have missing values. Missing values are wild cards and could expand the range of matches. So we adopted a fuzzy matching algorithm for this step. Any match on a common variable between a HMDA loan and a CoreLogic loan is in one of three matching categories: a perfect match, a perfect nonmatch, and a fuzzy match. A perfect match is assigned a weight of 1, a perfect nonmatch is assigned a weight of 0, and a fuzzy match, with equal likelihood of a perfect match and nonmatch, is assigned a weight of 0.5.

The fuzzy matching approach can generate multiple CoreLogic matches for a given HMDA loan. We need to determine which is the most likely and assign weights accordingly. If the weight assigned to the match between the *i*th HMDA loan and *j*th CoreLogic loan on the *k*th common variable is W_{ijk}, $k = 1,\cdots K$, the probability that HMDA loan i and CoreLogic loan j are a true match is given by

$$Q_{ij} = P_{ij} \times \prod_{k=1}^{K} W_{ijk} \tag{A.4}$$

For HMDA loan i, any supplemental information obtained from the CoreLogic loan j is weighted by Q_{ij}. For example, if a HMDA loan has two matched CoreLogic loans, with weights q1 and q2, and credit scores cs1 and cs2, respectively, the inferred credit score for the HMDA loan would be calculated by (q1*cs1+q2*cs2)/(q1+q2).

Matching Rate

Table A.8 shows the matching rate between HMDA and CoreLogic loans for 2012 through 2017. In 2017, there are 3,426,269 owner-occupied purchase mortgage originations in the original HMDA data. After passing through the matching steps described above, there are 2,345,951 HMDA loans, each matched with at least one CoreLogic loan. Sixty-eight percent of HMDA loans find at least one match with CoreLogic loans in 2017, and 29% of these are unique matches.

Table A.8. Match Rates between HMDA and CoreLogic Loans, 2012–17

Origination year	HMDA loans	Matched HMDA loans	Match rate
2012	2,210,872	1,762,760	80%
2013	2,548,282	1,965,170	77%
2014	2,646,439	1,959,384	74%
2015	3,005,118	2,188,192	73%
2016	3,315,072	2,405,776	73%
2017	3,426,269	2,345,951	68%

Source: HMDA, CoreLogic, and the Urban Institute.
Notes: HMDA = Home Mortgage Disclosure Act. The data are limited to owner-occupied purchase mortgage originations with nonmissing race and ethnicity information.

We used a two-step weighting approach to make the matched loans representative of the mortgage market on credit score composition and representative of the original HMDA loans on any combination of important variables (e.g., origination year, race or ethnicity, income, loan amount, and channel). Each matched loan is first weighted to reflect the same credit score distribution of the mortgage market for that year (weighted separately for the conventional and government channels) and then weighted to reflect the same joint distribution as the original HMDA loan characteristics above.

References

Bai, B., Goodman, L., & Ganesh, B. (2017). *Lower-credit mortgage applicants are dropping out of the market: Evidence from the latest real denial rate update*. Washington, DC: Urban Institute.

Bhutta, N., Laufer, S., & Ringo D. R. (2017). *Residential lending in 2016: Evidence from the Home Mortgage Disclosure Act data*. Washington, DC: Board of Governors of the Federal Reserve System.

FHA Office of Single Family Housing. (2015). *FHA's single family housing loan quality assessment methodology (defect taxonomy)*. Washington, DC: Author.

Ginnie Mae. (2018a). The Ginnie Mae Global Market Analysis Report. Retrieved from https://www.ginniemae.gov/data_and_reports/reporting/Pages/global_market_analysis.aspx.

Ginnie Mae. (2018b). Global markets analysis report: A monthly publication of Ginnie Mae's Office of Capital Markets. Washington, DC: Author.

Glantz, A., & Martinez E. (2018, February 15). For people of color, banks are shutting the door to homeownership. Reveal. Retrieved from https://www.revealnews.org.

Goodman, L. S. (2015). *Wielding a heavy enforcement hammer has unintended consequences for the FHA mortgage market*. Washington, DC: Urban Institute.

Goodman, L. S. (2017). Quantifying the tightness of mortgage credit and assessing policy actions. *Boston College Journal of Law and Social Justice*, 37(2), 235–265.

Goodman, L., McCargo, A., Golding, E., Parrott, J., Pardo, S., Hill, T., …Reyes, A. (2018). *Housing Finance at a Glance: A Monthly Chartbook, April 2018*. Washington, DC: Urban Institute.

Li, W., & Goodman L. S. (2014a). *A better measure of mortgage application denial rates*. Washington, DC: Urban Institute.

Li, W., & Goodman L. S. (2014b). *Measuring credit availability using ex-ante probability of default*. Washington, DC: Urban Institute.

Li, W., Goodman, L., Seidman, E., Parrott, J., Zhu, J., & Bai, B. (2014). *Measuring mortgage credit accessibility*. Washington, DC: Urban Institute.

McCargo, A., Bai, B., George, T., and Strochak, S. (2018). *Small-dollar mortgages for single-family residential properties*. Washington, DC: Urban Institute.

Urban Institute. (2018, July 12). *Housing Credit Availability Index*. Retrieved from https://www.urban.org/policy-centers/housing-finance-policy-center/projects/housing-credit-availability-index.