WORKING PAPER

Measuring Mortgage Credit Accessibility

Wei Li *Laurie Goodman* *Ellen Seidman* *Jim Parrott*
Jun Zhu *Bing Bai*

November 2014

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.

ABOUT THE HOUSING FINANCE POLICY CENTER

The Urban Institute's Housing Finance Policy Center (HFPC) was established to provide timely, impartial data and analysis, and to educate policymakers and the public on how the housing finance system affects households, communities and the broader economy. HFPC's goal is to produce analyses and ideas that foster sound public policy, efficient markets, and economic opportunity.

The 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 Foundation.

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.

All funding supports HFPC's research, outreach and engagement, and general operating activities. Funders do not determine research findings or influence scholars' conclusions. Scholars are independent and empowered to share their evidence-based views and recommendations shaped by research. The Urban Institute does not take positions on issues.

Contents

Acknowledgments

This is a draft working paper. Please address correspondence to Wei Li, wli@urban.org.

The Urban Institute thanks The Citi Foundation and The John D. and Catherine T. MacArthur Foundation for providing generous support at the leadership level to launch the Housing Finance Policy Center. Additional support was provided by the Ford Foundation and the Open Society Foundations.

The Urban Institute strives for the highest standards of integrity and quality in its research, analyses, and policy recommendations. Urban scholars believe that independence, rigor, and transparency are essential to upholding those standards. Funders do not determine research findings or influence scholars' conclusions. As an organization, the Urban Institute does not take positions on issues. Urban scholars and experts are independent and empowered to share their evidence-based views and recommendations shaped by research.

The views expressed are those of the authors and should not be attributed to the Urban Institute, its trustees, or its funders.

Abbreviations and Acronyms

ANA	approved but not accepted
AOPR	application-to-origination progression rate
CL	CoreLogic
DAPR	demand-to-application progression rate
DDR	deter and denial rate
DOPR	demand-to-origination progression rate
DTA	debt-to-asset ratio
DTI	debt-to-income ratio
FHA	Federal Housing Administration
FICO	Fair Isaac Credit Score
RD	Rural Development Program (US Department of Agriculture)
FVR	FHA, VA, and RD
GSE	government-sponsored enterprises (Fannie Mae and Freddie Mac)
HCP	high credit profile
HMDA	Home Mortgage Disclosure Act
LCP	low credit profile
LTV	loan-to-value
NHW	non-Hispanic whites
ODR	observed denial rate
PLCPA	proportion of low-credit-profile applicants among all mortgage credit applicants
PLCPB	proportion of low-credit-profile borrowers among all mortgage credit borrowers
PLCPD	proportion of low-credit-profile consumers with credit need among all consumers with credit need
PLS	private-label securities
PP	private-label securities and portfolio loans on the books of financial institutions
RDR	real denial rate
RE	race/ethnicity
SCF	Survey of Consumer Finances
SLO	Senior Loan Officer Opinion Survey
SFLPD	single-family loan performance data from Fannie Mae and Freddie Mac
VA	Department of Veterans Affairs

Executive Summary

Access to mortgage credit is most commonly determined by calculating the denial rate for all mortgage applicants using Home Mortgage Disclosure Act (HMDA) data. Calculating a denial rate based solely on this data, however, has two limitations:

1. HMDA data do not include applicants' credit profiles, so the denial rate calculated from these data merges applicants with strong and weak credit profiles. Because applicants with strong credit essentially have a denial rate of zero, this merger significantly dilutes and masks the only relevant denial rate: that of the weaker-credit-profile applicants.
2. HMDA data fail to properly account for those who want to apply for a loan but do not because they believe their application will be turned down.

We address these limitations by offering a better measure of mortgage credit accessibility: the demand-to-origination progression rate for low-credit-profile consumers. Using this improved measure, we explore several issues critical to credit accessibility including differences among demographic groups, changes over time and credit cycles, and the impact of government support for the single-family owner-occupied mortgage market.

Our analysis results in four findings:

1. government-guaranteed (FHA/VA/RD) lending has reduced racial gaps in loan approvals for low-credit-profile individuals seeking credit, particularly in the years following the collapse of the housing market;
2. up until the collapse, the private market proved particularly accessible for racial minorities with low credit profiles, although this does not take into account the terms or pricing of the loans;
3. contrary to popular opinion, the GSEs extended relatively little credit to weaker credit profile consumers in the lead-up to the crisis; and
4. mortgage credit has been extraordinarily tight since 2009.

Introduction

Access to sustainable mortgage credit can play a critical role in a family's pursuit of financial prosperity. Yet our ability to measure access to such credit has been limited, until now, largely because our measurement tool has been too blunt. This paper addresses the limitations of existing measures of mortgage credit accessibility by proposing better measures. It then uses these improved measures to answer several questions critical to understanding mortgage credit accessibility.

Two Steps and Two Factors Key to Credit Accessibility

Credit accessibility measures the probability that a consumer with a need for credit can secure a loan at a given time, a process that takes two steps and depends on two factors. Consumers who need credit to get a mortgage must advance from demand to application and then from application to origination. The demand-to-application deter rate and the application-to-origination denial rate both depend on two factors—the consumer's credit profile and the market's risk appetite—that represent the demand and supply sides, respectively, of mortgage credit. Credit accessibility intends to measure market tightness, not demand strength. The right measure of credit accessibility would therefore be the two rates while holding the consumer's credit profile constant.

Two Challenges in Assessing Credit Accessibility

Creating such a measure poses two analytic challenges for researchers. First, there are data to calculate the application-to-origination denial rate but none to calculate the demand-to-application deter rate. The issue is analogous to the challenge faced when calculating the unemployment rate: that calculation counts only people looking for jobs, missing those who have given up looking.

Second, while we need to hold the consumer's credit profile constant to measure the mortgage application denial and deter rates, researchers and policymakers have information only about the credit characteristics of those who receive loans, not those whose loan applications are denied, let alone those who want to apply but do not for fear of denial.

By failing to account for both consumers who are denied loans and those who want to apply but don't, most measurements of credit accessibility fall short.

In the first half of this paper, we address these shortcomings and offer two better measures of mortgage credit accessibility: the mortgage application denial and deter rates for consumers with weaker credit profiles. Though our discussion applies mainly to the US single-family, owner-occupied residential mortgage market, our proposed measures can be adapted to other credit fields. Notably, the analyses do not consider the price or terms of the loan.

In the second half of the paper, we use our new measures to answer four significant policy questions about mortgage credit:

1. how big are the accessibility gaps among different demographic groups?
2. how do these gaps change over time, especially over different periods of the credit cycle?
3. how effective have the government's efforts been in reducing these gaps and promoting credit accessibility in general? and
4. how do these measures change over time, especially over different periods of the credit cycles?

Existing Measures of Mortgage Credit Accessibility

The four most commonly cited indicators of mortgage credit accessibility are the borrowers' median credit score at origination[1], the Senior Loan Officer Opinion Survey on Bank Lending Practices (SLO)[2], mortgage application denial rates based on annual Home Mortgage Disclosure Act (HMDA) data,[3] and the relatively new index of credit availability, one by the Mortgage Bankers' Association (MBA). Another upcoming credit availability index developed by the Urban Institute's Housing Finance Policy Center (HFPC) is reviewed here too.

Borrowers' Median Credit Score

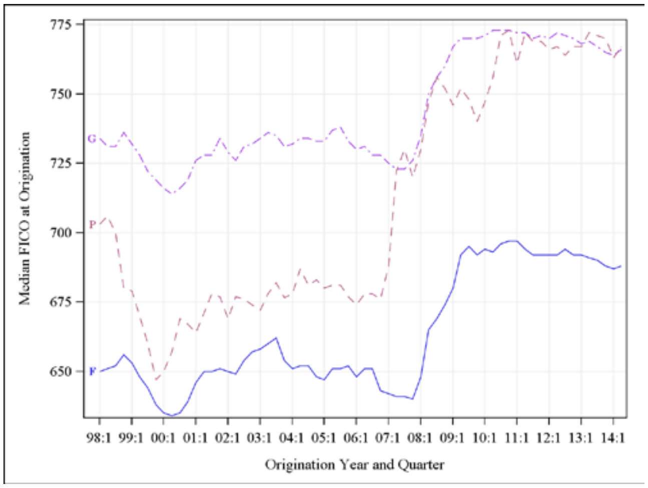
A common metric for credit accessibility is the credit characteristics of loans made to the median borrower. This has some intuitive appeal since it reflects the amount of risk the market is willing and expected to take at a given time. For example, the big jump in borrowers' credit scores after the financial crisis supports the view that the current credit box is too tight.

But this measure has two weaknesses: First, a borrower's credit score alone is insufficient as a measure of credit availability; loan-to-value (LTV), debt-to-income (DTI), and other factors also help a lender determine whether to make a loan to a particular borrower. Second, the results with this measure are counterintuitive: they show a slight increase in the median FICO scores for the private-label and bank portfolio channel between 2000 and 2006, indicating declining credit accessibility over a period widely perceived as driven by increasing accessibility (figure 1.A). The rise was driven by a range of factors having less to do with accessibility of credit in the market than with who was applying for loans and what kinds of loans they were applying for.

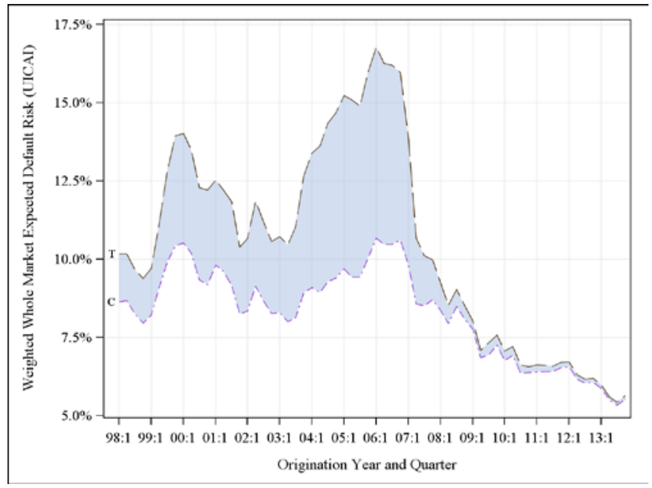
FIGURE 1

Existing Measures of Mortgage Credit Availability

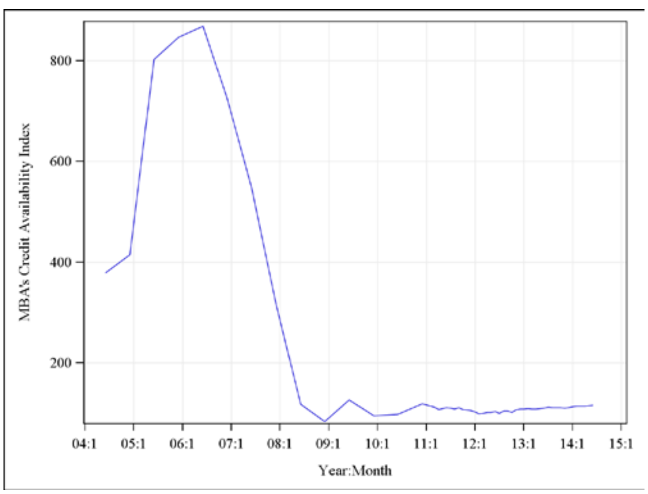
A. Median borrower's credit score (CoreLogic data)



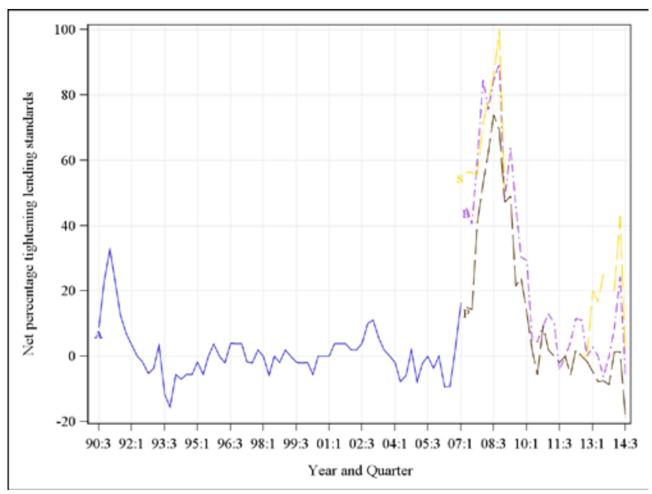
B. HFPC's Credit Availability Index (CoreLogic and HMDA data)



C. MBA's Mortgage Credit Availability Index



D. Federal Reserve's Senior Loan Officer Opinion Survey



HFPC's Credit Availability Index

Conceptually, the market should treat two loans with different combinations of risk factors but equal expected default risk as equally risky. Therefore, instead of examining each risk factor separately, Li and Goodman (upcoming) constructed a credit availability index that distills borrower credit profiles, loan products and terms, and macroeconomic conditions into a measurement of the weighted average probability of default for mortgages originated at a given time (figure 1.B). The index shows more intuitive trends than the median credit scores. However, this index is still imperfect because it is based solely on information about originated loans. This information omits important facts about who and how many applied for credit as well as the bigger pool of consumers who needed credit but were deterred from applying by market tightness.

MBA's Credit Availability Index

Another index is produced monthly by the Mortgage Bankers' Association (MBA).[10] AllRegs[11] scans the credit guidelines for a large number of lenders, which is then aggregated by the MBA into a single number, as shown in figure 1.C. The trend between 2004 and now is quite reasonable, but this index lacks transparency. Though we know, for example, that this index takes many factors into account (such as loan purpose, amortization type, property type, etc.) we do not know how numbers are assigned to these factors. Nor can we assess the formula that converts these many factors into a single index number.[12]

The Senior Loan Officer Opinion Survey

The SLO, usually conducted four times a year by the board of governors of the Federal Reserve System, is designed to measure credit accessibility qualitatively by asking banks to report changes in their lending practices over the previous three months. A senior loan officer at each respondent bank completes this voluntary survey electronically. Currently, up to 60 large domestically chartered commercial banks respond to the SLO.

The SLO can be used to calculate the net share of domestic respondents tightening lending standards for residential mortgage loans: the fraction of banks that reported having tightened standards "considerably" or "somewhat" minus the fraction of banks that reported having eased standards “considerably” or “somewhat,”[4] as shown in figure 1.D. Thus, the SLO provides five categorical measures on lending standard changes perceived by the banks for the past three months. It is unable, however, to measure the change in or the degree of credit accessibility at any given moment.

As also shown in figure 1.D, in the second quarter (Q2) of 2007, the share of banks tightening credit standards for prime mortgages was only 15 percent. This number climbed steadily after that and peaked in Q3 2008 at 74 percent, revealing continuous tightening of mortgage credit after the financial crisis.

The SLO is less successful at capturing the loosening of mortgage credit standards between 2003 and 2007. During these four years, the net share of banks tightening standards was zero or positive for nine quarters and negative for eight quarters,[5] and the magnitude was well below 10 percent for most quarters.

The SLO also fails to recognize the popularity of risky products as a sign of loosening credit standards before the financial crisis. The SLO did not begin asking separately about changes in lending standards for prime, nontraditional,[6] and subprime mortgage loans until Q2 2007. Accordingly, it missed a critical change in credit accessibility: many banks that answered that they did not originate nontraditional or subprime residential mortgages were actually active subprime lenders.[7]

Moreover, the 60 large domestically chartered commercial banks in the SLO’s reporting panel excluded some of the major lenders in the mortgage market before the financial crisis, such as Countrywide, Ameriquest, and New Century.

Mortgage Application Denial Rate Using HMDA Data

In mortgage credit research, especially in the study of racial discrimination in mortgage credit, researchers have traditionally used HMDA data[8] to calculate denial rates as a measure of credit inaccessibility. Most use some version of the following:

Let V be a dummy variable for an incidence of a credit application. It equals 1 when the application is approved and 0 when it is denied. The rate at which credit applications are denied at a given time, t, is

$$D_t = \text{Prob}(V_t = 0) \tag{1}$$

which equals the number of applications denied by the lender, divided by the total number of applications.

This result is usually called the denial rate of credit application or simply the denial rate. To differentiate it from another measure of denial rate developed later in this paper, we will call it the observed denial rate (ODR) or the traditional denial rate measure.

The ODR falls short as a good measure of credit accessibility because it doesn't consider applicants' credit profiles. For example, an increase in the denial rate can reflect either a tightening of the credit environment or an increase in applications by weaker-credit borrowers. To measure pure market tightness rather than demand strength, the denial rate must hold the applicant's credit profile constant. This poses an analytic challenge for researchers: we have information only about the credit characteristics of those who receive loans, not of those whose loan applications are denied. Moreover, ODR does not factor in consumers who would like to apply for a mortgage but do not move forward, assuming that they will be denied.[9] A lender with a low denial rate may either be offering credit broadly or discouraging higher-risk borrowers from even applying. We solve these challenges below.

Developing Better Measures of Credit Accessibility

Steps from Credit Demand to Credit Origination

A prospective borrower goes through two basic steps when securing a mortgage: first from demand to application, then from application to origination (table 1). How quickly a borrower progresses through both these steps depends on the interaction of two factors: the credit supplier's tolerance for risk and the individual's credit profile.

TABLE 1

Steps and Progression Rates from Credit Demand to Credit Origination

Population with Credit Demand (N_t)	
High Credit Profile	*Low Credit Profile*
$N_{H,t}$	$N_{L,t}$
$P_{0,H,t}=N_{H,t}/N_t$	$P_{0,L,t}=N_{L,t}/N_t$

↓

Demand-to-Application Progression Rate (DAPR)	
$P_{1,H,t}$	$P_{1,L,t}$

↓

Subpopulation Applied for Credit	
High Credit Profile	Low Credit Profile
$N_t \times P_{0,H,t} \times P_{1,H,t}$	$N_t \times P_{0,L,t} \times P_{1,L,t}$

↓

Application-to-Origination Progression Rate (AOPR)	
$P_{2,H,t}$	$P_{2,L,t}$

↓

Subpopulation Originated Credit	
High Credit Profile	Low Credit Profile
$N_t \times P_{0,H,t} \times P_{1,H,t} \times P_{2,H,t}$	$N_t \times P_{0,L,t} \times P_{1,L,t} \times P_{2,L,t}$

In a Perfect World, Everyone Has Access to a Loan

In an ideal world of perfect information, with total price flexibility and without statutory or regulatory constraints, lenders would collect the potential borrower's credit profile information, measure the risk of the loan according to the lender's risk standards, and offer a loan priced according to the borrower's risk. In such a world, the interest rate or the price of the loan alone would be enough to clear the market. Thus, no loan application would be turned down because each would have a price that compensated the lender for the risk presented by a particular applicant.[13]

However, lenders do not operate in such a world. An applicant's credit profile is a far-from-perfect measure of the credit risk of the loan, and regulatory and other barriers keep lenders from charging what they think they need to cover the risk. The reason behind this is beyond the scope of this paper, but the result is that lenders sometimes rely on rationing rather than pricing to cover their risk, and borrowers with lower credit profiles are much more likely to be denied a loan than their higher-credit-profile counterparts. In fact, applicants with a strong-enough credit profile are unlikely to be denied a loan at all.

The Denial Rate and Cost of Applying for a Loan Can Dissuade an Applicant

Loan applications are not free. For refinance loans, an appraisal, a preliminary title search, and a credit report must be completed before a lender approves a loan application. Lenders usually charge applicants for the costs of processing these documents. For purchase loans, the costs also include those associated with looking for a home, such as hiring an agent and so on.

So the first decision a potential borrower faces is whether to apply for a loan, given the costs and likelihood of success. Ultimately, credit results in a loan application only when the wealth or the utility a borrower gains from the credit exceeds the cost of the loan application. For some applicants, the probability of being denied a loan is so high that applying does not make economic sense. We call the proportion of people who want to apply for a loan but do not for fear of denial the mortgage application deter rate, which quantifies credit accessibility for demand-to-application step. The deter rate depends on the consumer's expectation of being denied for a loan application, which in turn depends on the consumer's perception of his or her own credit profile. Assuming people understand clearly their own credit profile, individuals who need credit and have high-enough credit profiles are unlikely to be deterred from making a mortgage application. Consequently, the mortgage application deter rate, like the denial rate, is only meaningful for weaker-credit-profile consumers.

In Our Imperfect World, the Government Helps

To promote credit accessibility to applicants with weaker credit profiles, the US government provides various forms of credit support for residential mortgages. This support falls into two broad categories:

1. direct intervention through 100-percent or close-to-100-percent guarantees or insurance by government agencies, such as the Federal Housing Administration (FHA), the Department of Veterans Affairs (VA), and the Rural Development (RD) program at the Department of Agriculture (the FVR channel); and
2. a formerly implicit (now explicit) government guarantee of lending through Fannie Mae and Freddie Mac, the government-sponsored enterprises (GSEs).

A third channel—loans funded through private-label securities (PLS) and loans on the books of financial institutions—make up the remainder of the mortgage market.

The FVR channel traditionally has had the broadest credit box and higher pricing than the other two channels, and it has been used disproportionately by weaker-credit-profile consumers. In the PP loan market, private-sector actors bear all the credit risk; therefore, the availability and pricing of credit through this channel has varied widely over time. The GSE market sits between the FVR and the PP channels in terms of the government's intervention on credit accessibility, but it usually offers better price and loan terms than the private market by setting tighter credit standards.

Guarantors Set Underwriting Standards That Lenders Sometimes Exceed or Fail to Meet

In the two government-backed channels, the mortgage guarantors establish underwriting guidelines for lenders to follow. In theory, if a loan meets these requirements, a lender should make the loan, because the lender will be able to unload the credit risk. Thus, in theory, measuring credit accessibility through these two lending channels should be straightforward: we only need to track the guarantors' underwriting guidelines to determine credit accessibility.

In practice, it's more complicated. The government guarantors agree to take on the credit risk if the lender complies with the guarantor's rules in underwriting the loan. If a lender breaches those rules, the guarantor reserves the right to put the credit risk back on the lender ("put-back"). When put-backs are vague,[14] however, it is difficult for lenders to know if they have fully transferred credit risk to the guarantor. Accordingly, lenders apply their own minimum underwriting standards, which are more conservative than those required by the guarantors ("overlays").

The opposite can also happen when a lender believes that the put-back risk is minimal: a lender may approve a loan that does not actually meet the underwriting guidelines set by the guarantors, assuming that the loan will be guaranteed nonetheless. This clearly happened during the housing boom; indeed, some of the overly aggressive enforcement of put-back rights today can be attributed to an overcorrection from the years when the GSEs took on more credit risk than they had intended or approved.

So tracking guarantors' underwriting guidelines alone is not a sufficient measure of credit accessibility. Credit requirements are determined jointly by guarantors and lenders, producing a functional credit box that is sometimes narrower than allowed by the guarantors, sometimes broader. Our goal is to measure these unobservable credit requirements over time.

If we can determine who needs credit, who applied for a mortgage loan, and whose loan application was approved, we can calculate the progression rates from one group to the next, which measures both the unobservable credit requirements over time and their impact on credit accessibility.

Our information about these three pools, however, is limited. We have complete information about the borrowers who originated loans,[15] limited information about consumers who applied for loans,[16] and no information about consumers who need loans. But a key observation of the mortgage market can help us solve these data limitations: when consumers who need loans have a strong-enough credit profile, they are unlikely to be either deterred from or denied a mortgage application.

Calculating the Real Denial Rate and Deter Rate

When consumers with credit needs have strong-enough credit profiles, they are unlikely to be either deterred from or denied a mortgage application (see our mathematical explanation in box 1). This fact has the following implications:

1. Since high-credit-profile, or HCP, consumers with credit need have deter and denial rates of zero, deter and denial rates are only meaningful for low-credit-profile, or LCP, consumers who need credit.

2. Including both HCP and LCP consumers in the calculation underestimates both rates.

3. We can observe the number and credit profiles of HCP borrowers, allowing us to calculate the number and credit profiles of all HCP applicants and HCP consumers who need credit.

With these implications, we can calculate denial and deter rates just for LCP consumers, as shown in box 2.

BOX 1

At any given time t, there are distinct boundaries of credit requirements set in each of the three channels—FVR, PP, and GSE—determined partly by the guarantor and partly by the lenders. Call these three levels of credit requirements $A_{1,t}$, $A_{2,t}$, and $A_{3,t}$, for the FVR, PP, and GSE channels, respectively. The credit requirements tend to be highest for GSE loans and lowest for FVR loans, or $A_{1,t} \leq A_{2,t} \leq A_{3,t}$.[a] The cost of credit tracks this order, with higher credit requirements generally translating into a lower cost of credit.

We assume that an individual accurately assesses his or her credit profile, which we will call X_t. If we let the individual's perception of the three levels of credit requirement be $\tilde{A}_{1,t}$, $\tilde{A}_{2,t}$ and $\tilde{A}_{3,t}$, and $\tilde{A}_{1,t} \leq \tilde{A}_{2,t} \leq \tilde{A}_{3,t}$; then an economically rational individual will apply for a loan only when $X_t \geq \tilde{A}_{1,t}$. To capture the variation of $\tilde{A}_{i,t}$ among individuals, assume $\tilde{A}_{i,t}$ follows a bounded random distribution: $\tilde{A}_{1,t} \in (-\infty, A_{2,t})$, $\tilde{A}_{2,t} \in (A_{1,t}, A_{3,t})$, and $\tilde{A}_{3,t} \in (A_{2,t}, \infty)$.[b] Therefore, if an individual's credit profile is equal to or greater than $A_{2,t}$, then it is always economically rational for him or her to apply for a loan. If applicants have a credit profile equal to or greater than $A_{3,t}$, they are unlikely being either denied or deterred by a lender in any of the three channels.

Let

$$A_3 = \max_t\{A_{3,t}\} \tag{2}$$

If we define individuals with a credit profile equal to or greater than A_3 as the HCP group, and everyone else as the LCP group:[c]

$$C_t = \begin{cases} 1, if\ X_t \geq A_3 \\ 0, if\ X_t < A_3 \end{cases} \tag{3}$$

Then HCP consumers will always have zero chance being either denied or deterred—that is, $P_{1,H,t} = 1$ and $P_{2,H,t} = 1$, as defined in table 1.

a. This order may change over time. It is generally right for our study period (1998–2012).
b. In other words, we assume that the mortgage market including its three channels is efficient with sending consumers signals about their lending standards. Therefore, consumer's perception of these lending standards may vary at individual level but the variations fall within the boundaries defined here. The mortgage origination process is dominated by brokers, which help to achieve the efficiency of signaling of lending standards. However, in reality, when broker's incentive is not to help consumers make the best decisions, consumers may be misled with wrong choice of channels. In the past, there is documented evidence of predatory lending practices by brokers, especially steering consumers into subprime loans, see Ernst, Bocian, and Li (2008).
c. See appendix A2.2 for empirically defining LCP applicants and borrowers.

BOX 2

From table 1, we can derive that the ODR is equal to

$$D_t = \frac{P_{0,L,t} \times P_{1,L,t} \times (1 - P_{2,L,t})}{1 - P_{0,L,t} + (P_{0,L,t} \times P_{1,L,t})} \quad (4)$$

D_t is known from HMDA data. For the definitions of $P_{i,j,t}$, see table 1.

Let B_t be the proportion of LCP borrowers out of the pool of all borrowers (PLCPB). Similarly from table 1, we have

$$B_t = Prob(C_t = 0 | V_t = 1) = \frac{P_{0,L,t} \times P_{1,L,t} \times P_{2,L,t}}{1 - P_{0,L,t} + (P_{0,L,t} \times P_{1,L,t} \times P_{2,L,t})} \quad (5)$$

B_t is known from CoreLogic (CL) data.[a]

Let Q_t be the proportion of LCP applicants out of all applicants (PLCPA). Combining equations (4) and (5), we have

$$Q_t = Prob(C_t = 0 | Y_t = 1) = B_t + D_t - B_t D_t \quad (6)$$

Here, Y_t is a dummy variable for an incidence of a credit decision by an individual who wants credit. It equals 1 if the individual applies for credit and 0 otherwise. Q_t is solvable with the matched HMDA and CL data.

Then the denial rate just for LCP applicants is equal to

$$D_{L,t} = Prob(V_t = 0 | C_t = 0, Y_t = 1) = \frac{D_t}{Q_t} \quad (7)$$

We call $D_{L,t}$ the real denial rate, or RDR, which is also solvable with the matched HMDA and CL data. 1-$D_{L,t}$ is the application-to-origination progression rate (AOPR) for LCP applicants (see table 1):

$$P_{2,L,t} = Prob(V_t = 1 | C_t = 0, Y_t = 1) = 1 - \frac{D_t}{Q_t} \quad (8)$$

Meanwhile, the demand-to-application progression rate (DAPR) for LCP individuals with credit need is given by

$$P_{1,L,t} = Prob(Y_t = 1 | C_t = 0) = \frac{1 - P_{0,L,t}}{P_{0,L,t}} \times \frac{Q_t}{1 - Q_t} \quad (9)$$

$$P_{0,L,t} = Prob(C_t = 0) \quad (10)$$

$P_{0,L,t}$ is the proportion of LCP individuals who want credit out of all individuals who want credit (PLCPD).

Then the overall demand-to-origination progression rate (DOPR) for LCP individuals is given by

$$P_{L,t} = Prob(V_t = 1 | C_t = 0) = P_{1,L,t} \times P_{2,L,t} = \frac{Q_t \times (1 - P_{0,L,t})}{(1 - Q_t) \times P_{0,L,t}} \times (1 - \frac{D_t}{Q_t}) \qquad (11)$$

a. CoreLogic, Inc., is a US corporation providing financial, property and consumer information, analytics and business intelligence. According to CoreLogic's data dictionary, its loan database covers approximately 75 percent to 90 percent of all residential mortgages including both outstanding and terminated loans, although the percentage varies by market. As of March 2013, the database covers approximately 85 percent of all outstanding residential mortgages.

Results

Denial Rate for All Applicants

The most widely cited measure of credit accessibility is the observed denial rate, which is simply the number of denied loan applications divided by the total number of applications. The ODR, which is calculated with HMDA data and published annually by the Federal Reserve Board when new data become available, measures credit accessibility for all applicants including both low- and high-credit-profile groups (see Bhutta and Ringo 2014 for the latest article). The results presented below allow us to compare this often-quoted ODR with the RDR, which measures denial rate only for low-credit-profile applicants.

Loan Application Outcomes

While there are four possible outcomes[17] when a loan application is made, our analysis excludes one outcome and regroups the remaining three into either Denied or Approved.

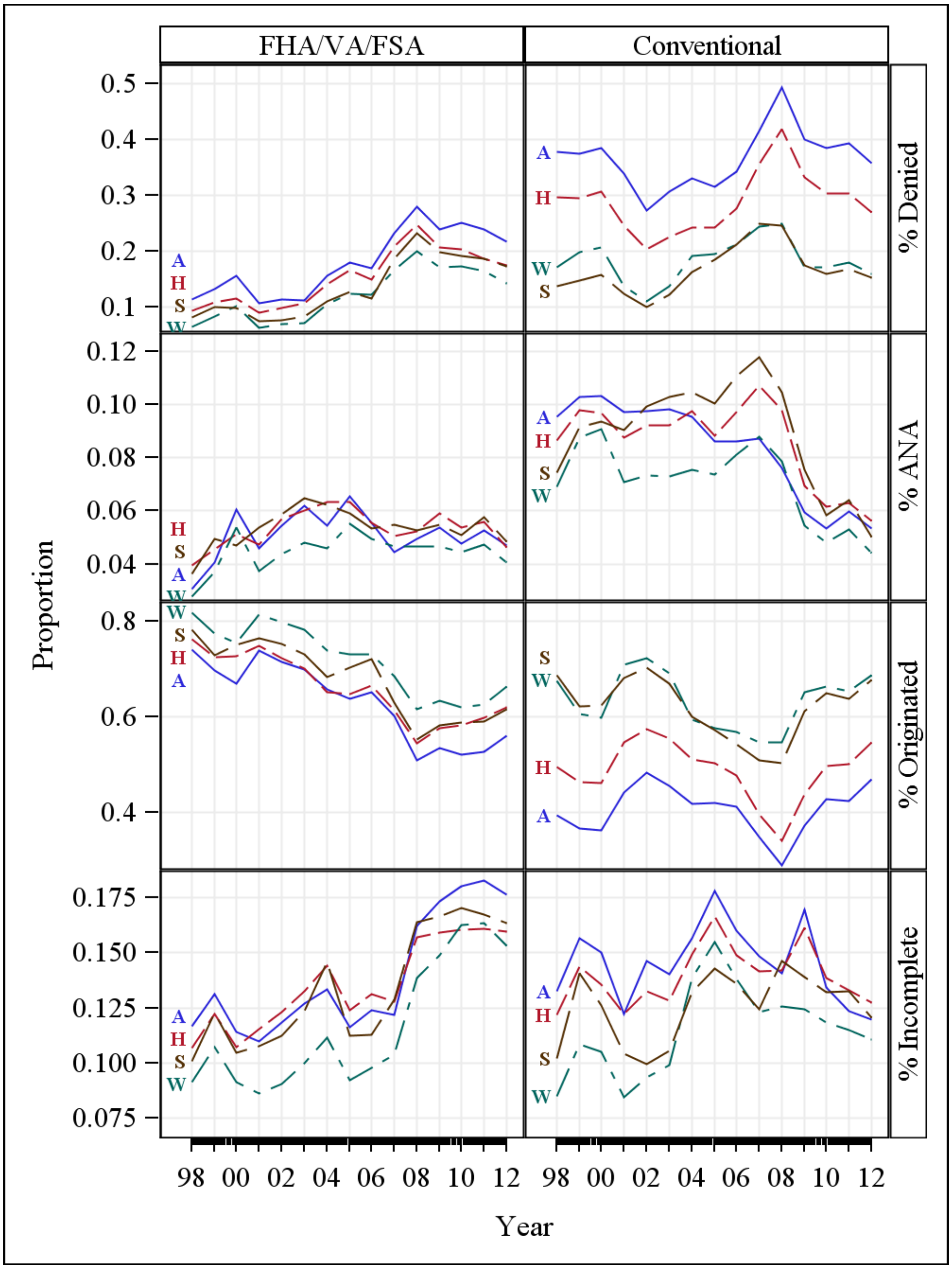
The four potential outcomes are dealt with as follows:

1. Application denied = Denied;
2. Application approved but not accepted = Approved;
3. Loan originated = Approved; and
4. Application withdrawn by applicant or file closed for incompleteness[18] = Excluded from the analysis.

Figure 2 and table 2 show the distribution of these four outcomes over time by race and ethnicity[19] and lending channel.[20] Incomplete application rates range from 8 percent to 18 percent of total applications, and vary over time by race, ethnicity, and channel.[21] Non-Hispanic whites tend to have a lower incomplete application rate than other minority groups. In addition, incomplete application rates for FVR loans increase over time. The rates of application approved but not accepted range from 4 percent to 10 percent for conventional loans and from 3 percent to 7 percent for FVR loans.

FIGURE 2

Outcomes of Mortgage Credit Applications



Source: HMDA data.
A = black; H = Hispanic; S = Asian; W = non-Hispanic white; ANA = approved but not accepted

TABLE 2

Outcomes of Mortgage Credit Applications

Channel	Race & Ethnicity	Origination Year														
		98	99	00	01	02	03	04	05	06	07	08	09	10	11	12
		% Application Denied														
FHA & VA & RD	Black	11	13	16	11	11	11	16	18	17	23	28	24	25	24	22
	Hispanic	9	11	11	9	10	11	14	17	15	21	25	21	20	19	17
	Non-Hispanic white	6	8	10	6	7	7	10	12	12	17	20	17	17	16	14
	Asian	8	10	10	7	8	8	11	13	11	19	23	20	19	19	17
Conventional	Black	38	38	38	34	27	31	33	32	34	42	49	40	38	39	36
	Hispanic	30	29	31	24	20	23	24	24	28	36	42	33	30	30	27
	Non-Hispanic white	17	20	21	14	11	14	19	20	21	24	25	17	17	18	16
	Asian	14	15	16	12	10	12	16	18	21	25	25	17	16	17	15
		% Application Approved but not Accepted														
FHA & VA & RD	Black	3	4	6	5	5	6	5	7	6	4	5	5	5	5	5
	Hispanic	4	5	5	5	6	6	6	6	6	5	5	6	5	6	5
	Non-Hispanic white	3	4	5	4	4	5	5	6	5	5	5	5	4	5	4
	Asian	4	5	5	5	6	6	6	6	5	5	5	5	5	6	5
Conventional	Black	10	10	10	10	10	10	10	9	9	9	8	6	5	6	5
	Hispanic	9	10	10	9	9	9	10	9	10	11	10	7	6	6	6
	Non-Hispanic white	7	9	9	7	7	7	8	7	8	9	8	5	5	5	4
	Asian	7	9	9	9	10	10	10	10	11	12	10	8	6	6	5
		% Application Approved and Originated														
FHA & VA & RD	Black	74	70	67	74	71	70	66	64	65	60	51	53	52	53	56
	Hispanic	76	72	73	75	72	70	65	65	66	61	54	58	58	60	62
	Non-Hispanic white	82	77	75	81	80	78	74	73	73	68	61	63	62	62	66
	Asian	78	73	75	76	75	73	68	70	72	63	55	58	59	59	62
Conventional	Black	39	37	36	44	48	46	42	42	41	35	29	37	43	42	47
	Hispanic	50	46	46	55	57	55	51	50	48	40	34	44	50	50	55
	Non-Hispanic white	68	61	60	71	72	69	59	58	57	55	55	65	66	65	69
	Asian	69	62	62	68	70	67	60	57	54	51	50	61	65	64	68
		% Application Incomplete or Withdrawn														
FHA & VA & RD	Black	12	13	11	11	12	13	13	12	12	12	16	17	18	18	18
	Hispanic	11	12	11	12	12	13	14	12	13	13	16	16	16	16	16
	Non-Hispanic white	9	11	9	9	9	10	11	9	10	10	14	15	16	16	15
	Asian	10	12	10	11	11	12	15	11	11	13	16	17	17	17	16
Conventional	Black	13	16	15	12	15	14	16	18	16	15	14	17	13	12	12
	Hispanic	12	14	14	12	13	13	15	17	15	14	14	16	14	13	13
	Non-Hispanic white	8	11	11	8	9	10	14	15	14	12	13	12	12	11	11
	Asian	10	14	13	10	10	11	13	14	14	12	15	14	13	13	12

Source: HMDA.

The Traditional Measure Shows Higher Denial Rates for Blacks and Hispanics

Table 3 and the upper left panel of figure 3 show the observed/traditional denial rate measure for all applicants in four race and ethnic groups from 1998 to 2012. The ODR is consistently highest for black applicants, followed by Hispanics. Non-Hispanic white and Asian applicants tend to have about the same ODR over the 15-year period. On average, the ODR for black applicants is about two times higher than for non-Hispanic whites and Asians. The pattern over time is about the same for all four groups.

TABLE 3

Four Denial or Deter Rates Measuring Credit Accessibility

Channel[a]	Race and ethnicity	Origination Year														
		98	99	00	01	02	03	04	05	06	07	08	09	10	11	12
		Observed Application-to-Origination Denial Rate (ODR) for All Applicants (%)														
FHA & VA & RD	Black	13	15	18	12	13	13	18	20	19	26	33	29	31	29	26
	Hispanic	10	12	13	10	11	12	16	19	17	24	29	25	24	22	21
	NHW	7	9	11	7	8	8	12	14	14	19	23	20	21	20	17
	Asian	9	11	11	8	9	9	13	14	13	21	28	24	23	22	21
	All[b]	8	11	12	8	9	9	14	15	15	21	26	22	22	21	19
Conventional	Black	44	44	45	39	32	36	39	38	41	49	57	48	44	45	41
	Hispanic	34	34	36	28	23	26	28	29	32	41	49	40	35	35	31
	NHW	19	22	23	15	12	15	22	23	25	28	29	19	19	20	18
	Asian	15	17	18	14	11	14	19	21	24	28	29	20	18	19	17
	All	21	25	26	18	14	17	25	26	28	32	33	22	21	22	20
All[c]	Black	42	43	44	37	31	35	39	38	40	48	55	44	42	42	38
	Hispanic	32	33	34	27	23	25	28	29	32	41	47	37	33	33	30
	NHW	18	22	23	15	12	15	22	23	25	28	28	19	20	20	18
	Asian	15	17	18	14	11	14	19	21	24	28	29	20	19	19	17
	All	21	24	26	17	14	17	24	26	28	32	33	22	21	22	20
		Application-to-Origination Denial Rate (RDR) for LCP Applicants (%)														
FHA & VA & RD	Black	18	20	22	16	17	17	24	26	25	33	45	60	65	65	66
	Hispanic	17	18	17	14	16	18	23	27	23	31	46	63	67	66	65
	NHW	14	16	17	11	13	13	19	21	20	27	40	60	65	65	62
	Asian	18	20	17	14	14	17	21	24	21	32	50	70	74	72	72
	All	15	17	18	13	14	15	20	23	21	29	42	60	66	65	64
GSE	Black	91	85	73	85	89	94	86	85	79	83	96	99	99	98	97
	Hispanic	95	91	75	91	93	96	90	89	84	88	98	100	99	99	98
	NHW	92	87	66	84	89	95	89	88	81	82	96	99	99	99	97
	Asian	94	89	69	91	94	98	93	93	85	90	98	100	99	99	99
	All	92	85	65	82	87	95	86	85	80	83	97	99	99	99	97
Portfolio & PLS	Black	64	60	56	53	40	56	52	54	60	74	93	98	98	98	96
	Hispanic	69	61	54	51	45	62	56	62	68	82	94	99	99	99	97
	NHW	57	48	43	21	19	53	44	47	55	73	92	97	99	98	95
	Asian	69	63	48	46	50	69	58	65	70	85	95	99	99	99	97
	All	66	59	52	48	47	64	56	57	62	77	93	98	99	98	96
Conventional[d]	Black	74	69	64	67	66	74	63	63	67	80	95	98	97	97	97
	Hispanic	84	76	67	74	73	80	68	71	74	86	96	99	99	99	98
	NHW	78	68	55	61	64	79	65	65	67	80	94	98	98	98	97
	Asian	85	75	60	74	76	85	73	75	77	89	97	99	99	99	99
	All	78	69	58	64	66	79	66	66	69	81	94	98	98	98	97

Channel[a]	Race and ethnicity	98	99	00	01	02	03	04	05	06	07	08	09	10	11	12
		Origination Year														
All[c]	Black	66	62	59	58	57	67	60	62	65	76	79	83	84	85	88
	Hispanic	67	62	56	58	59	70	64	69	72	82	82	86	87	88	88
	NHW	69	61	50	53	56	71	61	62	64	74	77	85	88	88	89
	Asian	75	66	54	66	70	81	70	74	76	87	88	93	94	94	95
	All	69	61	52	55	57	70	62	64	66	76	78	86	88	88	89
		Demand-to-Application Deter Rate for LCP Consumers (%)														
All	Black	52	34	12	49	69	71	53	41	40	36	50	75	78	78	83
	Hispanic	73	68	57	76	82	84	78	71	68	60	71	84	86	87	89
	NHW	69	41	13	59	72	72	42	43	39	42	55	77	78	77	81
	Asian	83	72	60	79	84	83	69	53	46	45	67	81	83	82	85
	All	69	45	20	62	74	74	48	46	41	41	57	80	81	80	83
		Overall Demand-to-Origination Deter/Denial Rate for LCP Consumers (%)														
FHA & VA & RD	Black	60	47	31	57	74	76	64	56	54	57	73	90	92	93	94
	Hispanic	78	74	64	80	85	87	83	79	75	73	84	94	96	96	96
	NHW	73	50	27	64	75	76	53	54	51	58	73	91	92	92	93
	Asian	86	78	67	82	86	85	75	64	57	63	83	94	96	95	96
	All	72	53	33	65	77	78	60	58	55	59	75	91	93	93	94
GSE	Black	96	90	76	92	96	98	93	91	87	89	98	100	100	100	99
	Hispanic	99	97	89	98	99	99	98	97	95	95	99	100	100	100	100
	NHW	98	92	70	94	97	99	94	93	88	89	98	100	100	100	99
	Asian	99	97	88	98	99	100	98	97	92	94	99	100	100	100	100
	All	97	91	72	93	96	99	93	92	88	90	98	100	100	100	100
Portfolio & PLS	Black	82	73	61	76	81	87	77	73	76	83	96	100	100	100	99
	Hispanic	92	88	80	88	90	94	90	89	90	93	98	100	100	100	100
	NHW	87	70	50	68	77	87	68	70	72	85	96	99	100	100	99
	Asian	95	90	79	89	92	95	87	84	84	92	98	100	100	100	100
	All	89	77	61	79	86	91	77	77	78	87	97	100	100	100	99
Conventional	Black	88	79	68	83	89	93	83	78	80	87	97	99	99	99	99
	Hispanic	96	93	86	94	95	97	93	92	91	94	99	100	100	100	100
	NHW	93	81	61	84	90	94	80	80	80	88	97	100	100	100	99
	Asian	97	93	84	94	96	97	91	88	88	94	99	100	100	100	100
	All	93	83	66	85	91	94	83	82	82	89	98	100	100	100	99
All	Black	84	75	63	79	87	90	81	77	79	84	90	96	96	97	98
	Hispanic	91	88	81	90	93	95	92	91	91	93	95	98	98	98	99
	NHW	90	77	57	81	87	92	78	78	78	85	89	97	97	97	98
	Asian	96	90	82	93	95	97	91	88	87	93	96	99	99	99	99
	All	90	79	62	83	89	92	80	80	80	86	91	97	98	98	98

Sources: Matched loan-level HMDA-CoreLogic data and Survey of Consumer Finance.

Note: NHW = non-Hispanic white.

a. For the ODR analysis, we rely on HMDA alone, which only allow us to separate the conventional channel from the FVR channel; for the other three rates, we use CL data to calculate PLCPB, which allows us to separate GSE loans from PP loans. The FVR, the conventional and the "All" categories are common in both analysis.

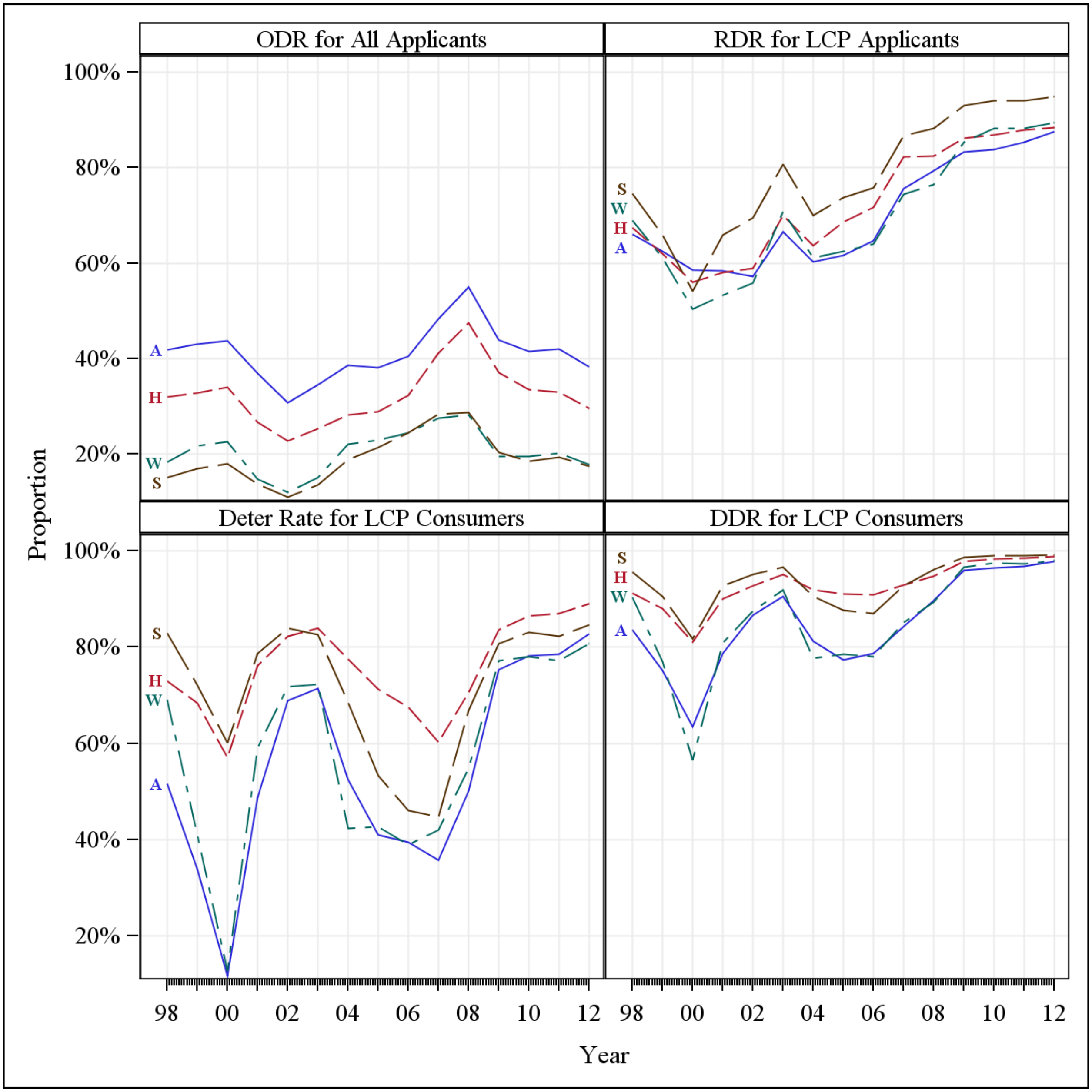
b. Results combined four RE groups.

c. Results combined the FVR and conventional channels.

d. Under HMDA, once a loan is originated, there is a field called type of purchaser; the possible values are GSE, PP, others, and not sold. Results shown here combined all four types.

FIGURE 3

Four Denial/Deter Rates by Race and Ethnicity



Sources: HMDA and CoreLogic.

Notes: A = black; H = Hispanic; S = Asian; W = non-Hispanic white. LCP = low credit profile. ODR = observed denial rate for all applicants; RDR = real denial rate for LCP applicants; Deter Rate = demand-to-application deter rate for LCP applicants; DDR = overall demand-to-origination deter/denial rate.

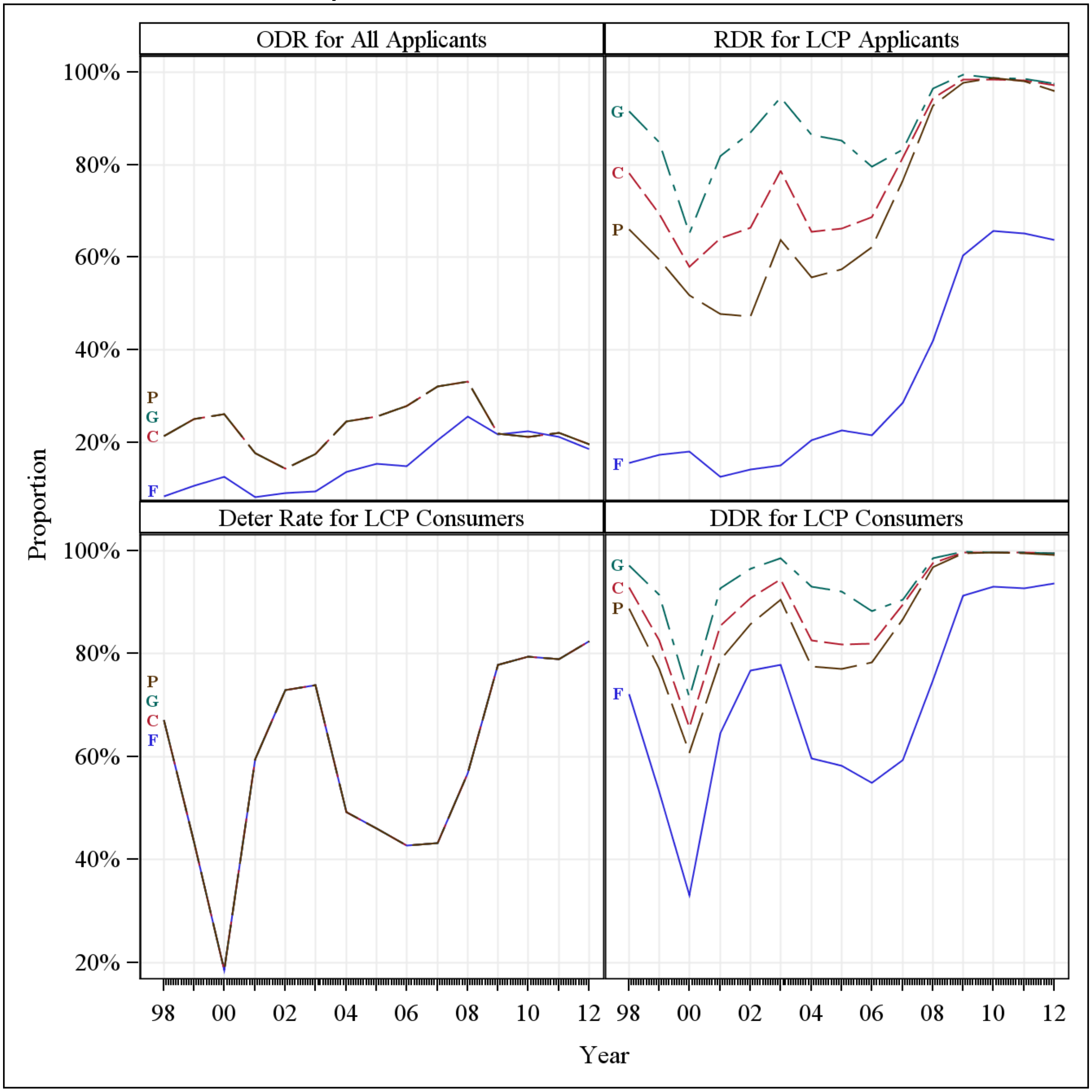
Does this finding indicate that black and Hispanic applicants have less access to mortgage credit than their white and Asian counterparts? No. The ODR alone does not allow us to draw this conclusion, since it doesn't consider the credit profile of the applicants. For example, if the share of low-credit-profile applicants is two times higher for black applicants than for white applicants, then the observed higher denial rate for the former could be explained entirely by the higher proportion of black LCP applicants.

The Traditional Measure Masks Distinctions between Channels

Table 3 and the upper left panel of figure 4 show observed denial rates by channel. It is consistently higher for applications submitted to conventional[22] channels than for applications submitted to FVR channels. Between 1998 and 2001, the former is about 2.3 to 2.6 times higher than the latter. Between 2002 and 2008, the former is about 1.3 to 1.9 times higher than the latter. After the financial crisis, the two show about the same level of observed denial rate. It is unsurprising that conventional channels have a higher denial rate than FVR channels. However, the observed disparity in denial rates between the two channels could be even higher if applicants' credit profile is considered because LCP applicants are more likely to apply for loans through the FVR channel than through the conventional channel. Once again, this finding highlights the limitation of using the ODR to measure and compare credit accessibility among different demographic groups and channels.

FIGURE 4

Four Denial/Deter Rates by Channels

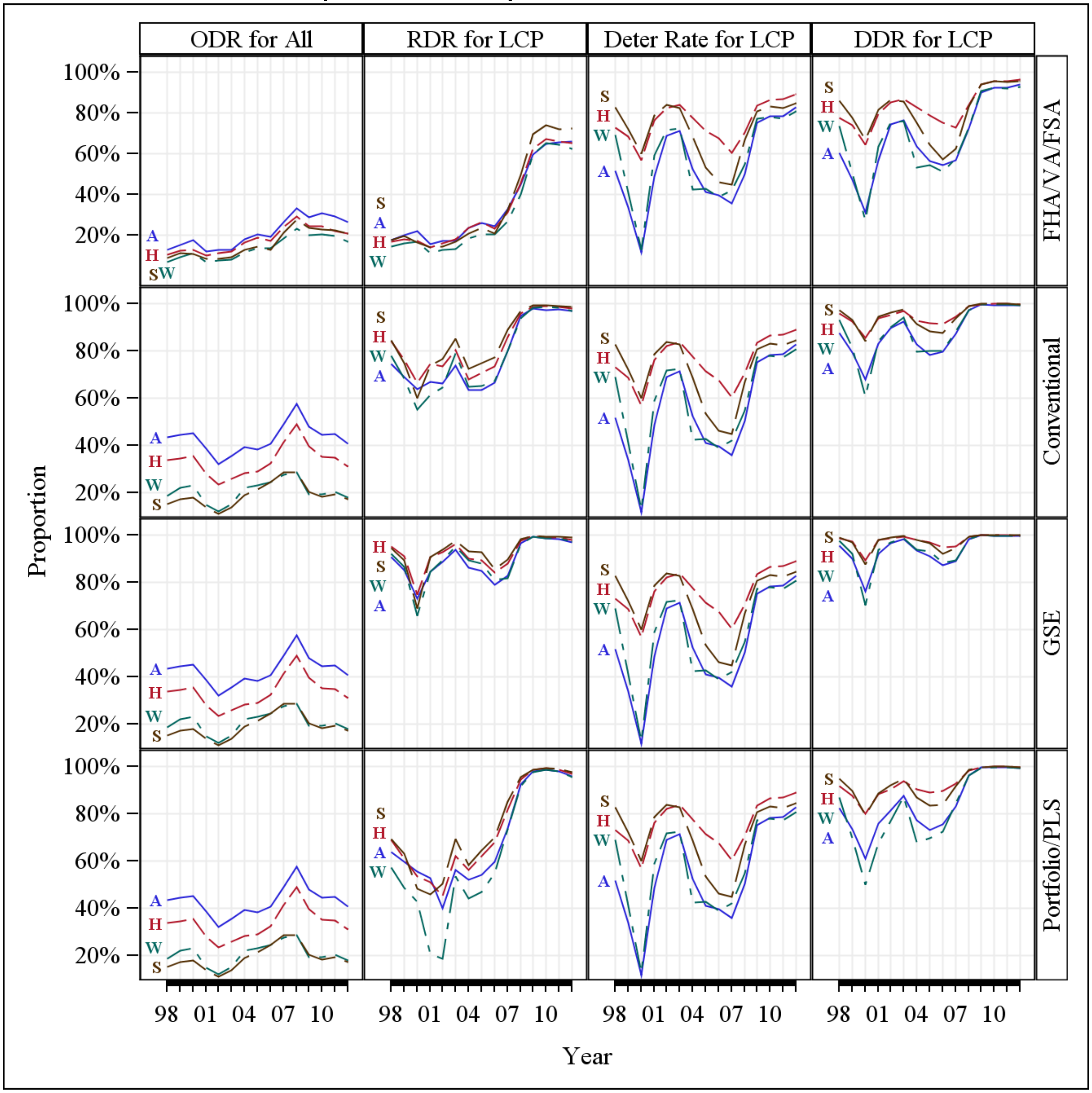


Sources: HMDA and CoreLogic.
Notes: C = conventional (GSE, portfolio, and PLS); F = FHA/VA/RD; G = GSEs; P: portfolio/PLS. LCP = low credit profile. ODR = observed denial rate for all applicants; RDR = real denial rate for LCP applicants; Deter Rate = demand-to-application deter rate for LCP applicants; DDR = overall demand-to-origination deter/denial rate.

The left-most panel of figure 5 and table 3 shows observed denial rates by race and ethnicity and channel. The racial difference on the ODR is smaller for applications to the FVR channel than for applications to the conventional channel. Within the FVR channel, on average over the 15-year period, the ODR is approximately 21, 18, 14, and 16 percent, respectively, for black, Hispanic, non-Hispanic white and Asian applicants. These numbers are 43, 33, 21, and 19 percent, respectively, within the conventional channel.

FIGURE 5

Four Denial/Deter Rates by Race/Ethnicity and Channels



Sources: HMDA and CoreLogic.
Notes: A = black; H = Hispanic; S = Asian; W = non-Hispanic white. LCP = low credit profile. ODR = observed application-to-origination progression rate for all applicants; RDR = real denial rate for LCP applicants; Deter Rate = demand-to-application deter rate for LCP applicants; DDR = overall demand-to-origination deter/denial rate.

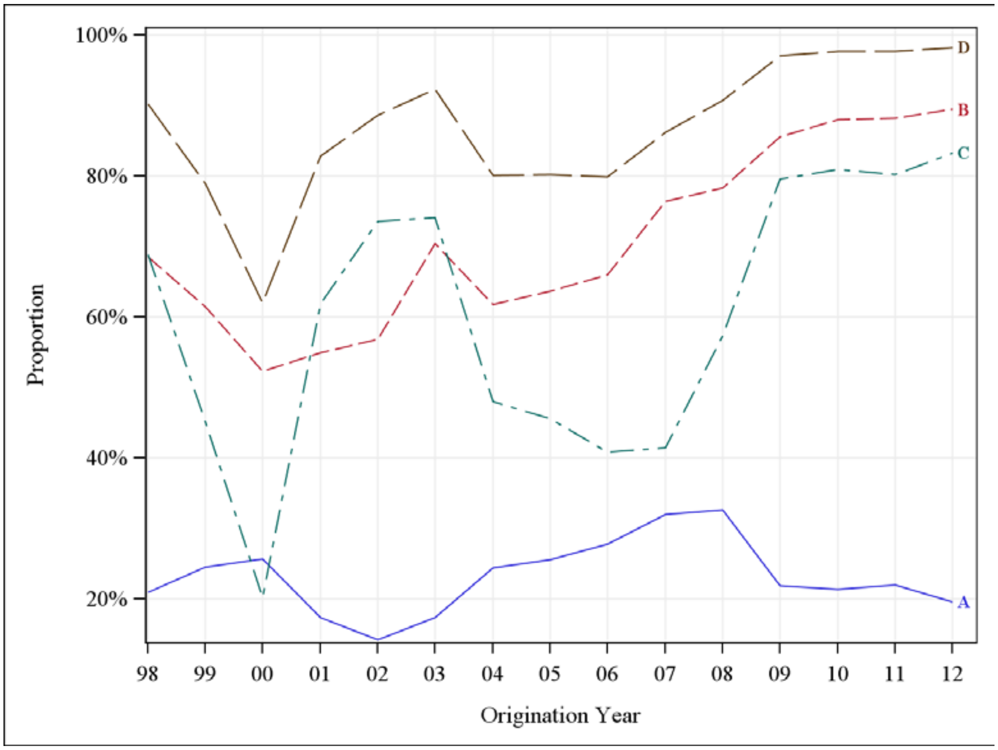
These patterns suggest that within the FVR channel, applicants' credit profile may be more homogenous across different racial and ethnic groups than in the conventional channels. These numbers also show that the channel difference in ODR varies by racial and ethnic groups: the ODR for blacks who applied for FVR loans is about two times higher than for those who applied for conventional loans, whereas the same ratio is much lower for non-Hispanic white and Asian applicants. To fully understand the gaps on credit accessibility among different demographic groups and channels, we have to control for applicant's credit profile.

The Traditional Measure Shows Counterintuitive Results

The observed denial rate's pattern over time is also counterintuitive because it correlates positively to the strength of the housing market. Curve A in figure 6 shows ODRs for all race and ethnic groups and all channels combined together. Rates are high from 1998 to 2000, drop as the market stalls from 2001 and 2003, and increase again as the market takes off in 2004. Rates then increase each year through the boom, reaching a peak in 2007 and 2008, falling again as the bottom falls out of the market in 2009.

FIGURE 6

Four Denial/Deter Rates for All Race/Ethnicity and Channels Combined



Sources: HMDA and CoreLogic.
Notes: A = observed denial rate for all applicants; B = real denial rate for LCP applicants; C: demand-to-application deter rate for LCP applicants; D: overall demand-to-origination deter/denial rate.

The US mortgage market reached the lowest ODR of our study period (14 percent) in 2002 and the highest (33 percent) in 2008. This counterintuitive pattern over time is observed for each race and ethnicity and for each channel. At its peak between 2007 and 2008, the ODR is about 55, 47, 28, and 29 percent, respectively, for black, Hispanic, non-Hispanic white, and Asian applicants (table 3). We would expect denial rates to be lower during the housing boom, with lenders approving loans they don't normally approve, and higher after the financial crisis as the credit box tightened.

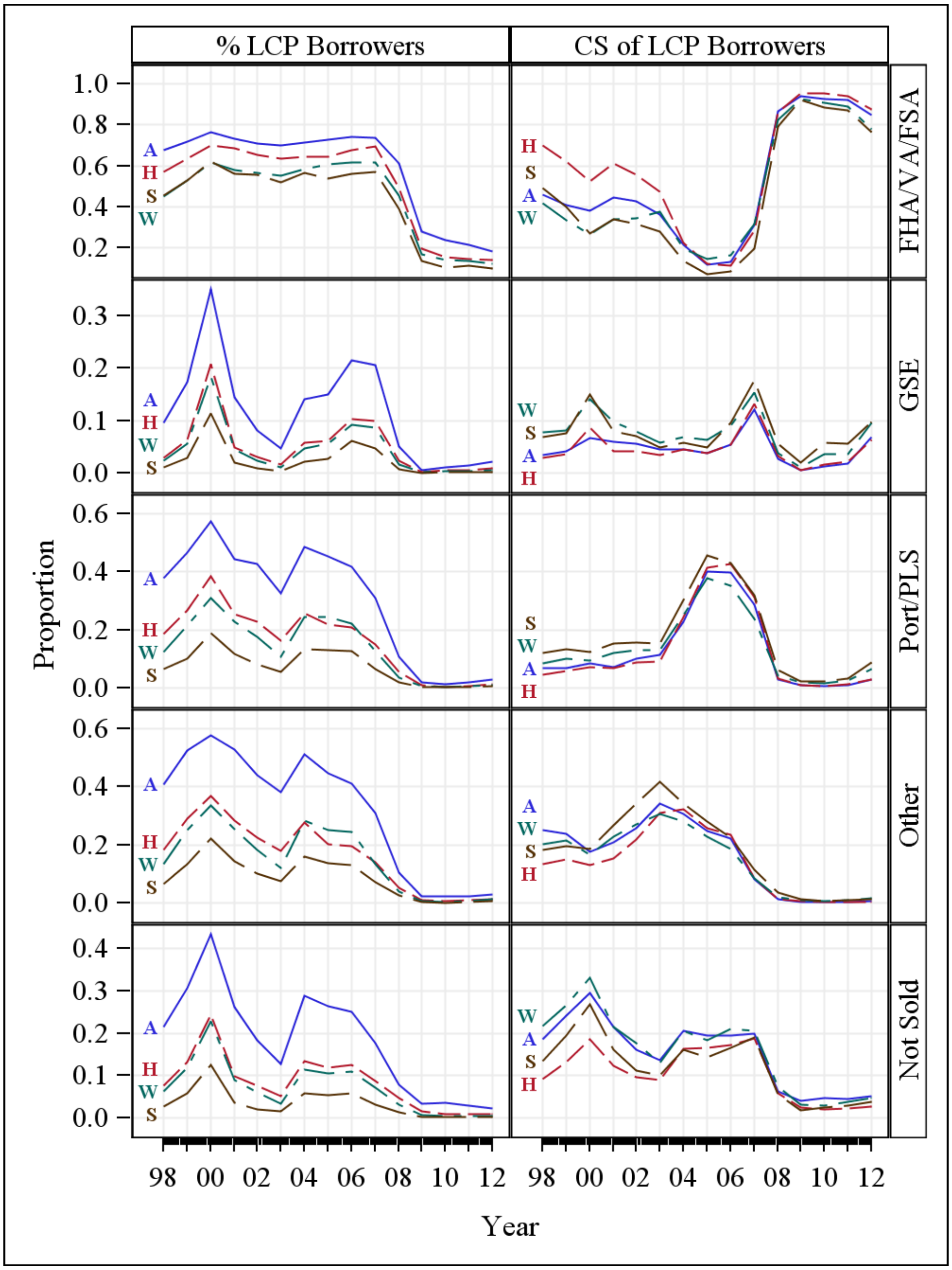
The counterintuitive results may be explained by changes in loan applicants' credit profiles. In the boom years, more LCP consumers were encouraged to submit applications; thus, more of them were rejected. As the credit box tightened after the financial crisis, many LCP consumers were discouraged from applying at all, leading to fewer rejections. These findings highlight that the ODR is a far-from-perfect measure of credit accessibility, especially for examining changes over time. If we control for the applicants' credit profiles, taking potential but discouraged applications into consideration, we may find totally different patterns.

Proportion of LCP Applicants

While CoreLogic data provide credit profiles for borrowers, neither HMDA nor CoreLogic data provide the credit profiles of applicants. Equation 6 in box 2 shows how we can calculate the proportion of LCP applicants in the broader pool of applicants (PLCPA), using the ODR and the proportion of LCP borrowers in the pool of borrowers (PLCPB). The left panels of figure 7 and table 4 show the PLCPB for all four racial and ethnic groups by channels over the past 15 years. PLCPA is shown in the left panel of figure 8 and table 4.

FIGURE 7

Percentage of LCP Borrowers out of All Borrowers and Channel Share (CS) of LCP Borrowers among All Borrowers

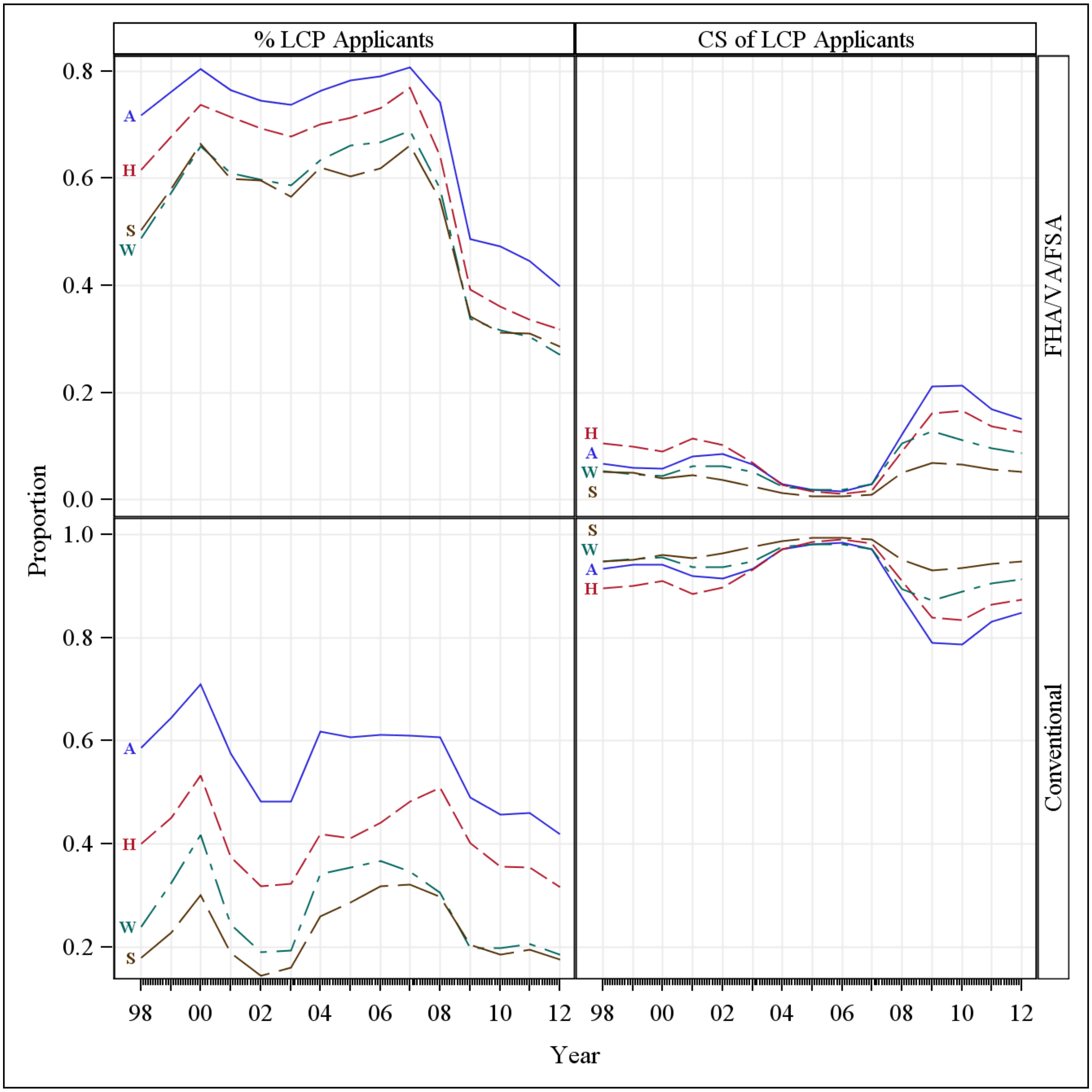


Sources: HMDA and CoreLogic.

Notes: A = black; H = Hispanic; S = Asian; W = non-Hispanic white. LCP = low credit profile.

FIGURE 8

% LCP Applicants out of All Applicants and Channel Share (CS) among LCP Applicants



Sources: HMDA and CoreLogic.
Notes: A = black; H = Hispanic; S = Asian; W = non-Hispanic white. LCP = low credit profile.

As we discussed under "Steps from Credit Demand to Credit Origination," potential borrowers with high credit profiles are unlikely to be deterred from or denied a loan application. So, the change in the share of LCP consumers from the applicant pool to the borrower pool solely reflects denials of LCP applicants. If LCP applicants have no chance of being denied, then PLCPA would equal PLCPB. Holding the ODR constant, a higher PLCPA will always lead to a higher PLCPB. This mathematical relationship between PLCPA, ODR, and PLCPB allow us to calculate the percentage of LCP applicants. We illustrate this calculation with some examples below.

TABLE 4

Percentage of LCP Borrowers and Applicants out of All Borrowers and Applicants, and Channel Shares among LCP Borrowers and Applicants

		Origination Year															
Channel	Race and ethnicity	98	99	00	01	02	03	04	05	06	07	08	09	10	11	12	Mean
		% LCP Applicants of All Applicants															
	Black	72	76	80	76	75	74	76	78	79	81	74	49	47	45	40	66
	Hispanic	62	68	74	72	69	68	70	71	73	77	64	39	36	34	32	57
FHA & VA &RD	Non-Hispanic white	49	57	66	61	60	59	63	66	67	69	58	34	32	30	27	48
	Asian	50	58	66	60	60	57	62	60	62	66	56	34	31	31	29	44
	All	54	62	70	65	64	62	67	69	69	72	61	36	34	32	29	51
	Black	59	65	71	58	48	48	62	61	61	61	61	49	46	46	42	58
	Hispanic	40	45	53	38	32	32	42	41	44	48	51	40	36	35	32	41
Conventional	Non-Hispanic white	24	32	42	24	19	19	34	35	37	35	30	20	20	21	19	28
	Asian	18	23	30	19	14	16	26	29	32	32	30	20	19	19	18	23
	All	27	36	45	28	22	22	38	39	41	39	35	22	22	22	20	31
	Black	59	65	71	59	50	49	62	61	61	61	62	49	46	46	42	59
	Hispanic	42	47	55	40	34	33	42	41	44	49	52	40	36	35	32	42
All	Non-Hispanic white	25	33	42	25	20	20	35	36	37	35	32	21	21	21	19	28
	Asian	19	24	31	19	15	16	26	29	32	32	30	21	19	20	18	23
	All	28	37	46	29	23	23	38	39	41	40	37	23	23	23	21	32
		% Channel among All LCP Applicants															
	Black	7	6	6	8	8	7	3	2	2	3	12	21	21	17	15	6
FHA &	Hispanic	10	10	9	11	10	7	3	1	1	2	9	16	17	14	13	6
VA &	Non-Hispanic white	5	5	4	6	6	5	2	2	2	3	11	13	11	10	9	5
RD	Asian	5	5	4	5	4	2	1	1	1	1	5	7	6	6	5	3
	All	6	5	5	7	7	6	2	2	2	3	10	14	12	10	9	5
	Black	93	94	94	92	92	93	97	98	98	97	88	79	79	83	85	94
	Hispanic	90	90	91	89	90	93	97	99	99	98	91	84	83	86	87	94
Conventional	Non-Hispanic white	95	95	96	94	94	95	98	98	98	97	89	87	89	90	91	95
	Asian	95	95	96	95	96	98	99	99	99	99	95	93	94	94	95	97
	All	94	95	95	93	93	94	98	98	98	97	90	86	88	90	91	95
		% LCP Borrowers of All Borrowers															
	Black	68	72	76	73	71	70	71	73	74	74	61	28	24	22	18	57
FHA &	Hispanic	57	63	70	68	65	63	64	65	68	70	49	19	16	15	14	47
VA &	Non-Hispanic white	45	53	62	58	57	55	59	61	62	62	45	17	14	13	12	39
RD	Asian	45	53	62	56	56	52	56	54	56	57	39	14	11	11	10	31
	All	50	57	65	62	60	58	61	63	64	64	48	18	15	14	13	42
	Black	10	17	35	14	8	5	14	15	22	21	5	1	1	1	2	11
	Hispanic	3	6	21	5	3	2	6	6	10	10	2	0	1	1	1	5
GSE	Non-Hispanic white	2	6	18	4	2	1	5	6	9	9	2	0	0	0	1	3
	Asian	1	3	11	2	1	0	2	3	6	5	1	0	0	0	0	2
	All	3	6	19	5	2	1	5	6	10	10	2	0	0	0	1	4

Channel	Race and ethnicity	Origination Year															
		98	99	00	01	02	03	04	05	06	07	08	09	10	11	12	Mean
Portfolio & PLS	Black	38	47	57	44	43	33	49	45	42	31	11	2	1	2	3	39
	Hispanic	19	27	38	26	23	16	26	22	21	15	6	1	0	1	1	19
	Non-Hispanic white	13	22	31	23	18	11	25	25	22	13	4	1	0	1	1	15
	Asian	7	10	19	12	8	6	13	13	13	7	2	0	0	0	1	8
	All	14	23	33	23	19	12	26	26	24	14	4	1	0	1	1	17
All	Black	37	45	55	42	33	26	42	38	37	30	32	16	14	13	9	34
	Hispanic	23	30	40	26	21	15	22	19	19	15	19	10	8	7	5	19
	Non-Hispanic white	10	18	29	15	11	7	18	18	18	13	12	4	3	3	3	12
	Asian	6	11	19	8	5	4	10	10	10	6	5	2	1	2	1	6
	All	12	20	32	17	13	9	20	19	20	14	14	5	4	4	3	14
		% Channel among All LCP Borrowers															
FHA & VA & RD	Black	82	79	71	77	73	70	44	21	23	43	94	99	98	97	90	59
	Hispanic	90	87	77	85	81	79	44	22	19	39	93	99	98	97	91	63
	Non-Hispanic white	72	65	53	61	62	66	38	25	27	45	91	97	94	93	83	55
	Asian	72	66	50	59	58	58	27	12	14	28	87	95	92	91	81	46
	All	76	70	59	66	66	68	40	23	25	43	92	97	95	94	85	56
GSE	Black	6	8	13	11	10	9	9	7	9	17	3	1	1	2	7	9
	Hispanic	4	5	13	6	6	6	9	7	9	18	4	1	2	2	7	8
	Non-Hispanic white	14	16	28	17	14	10	13	11	15	22	4	1	4	4	10	14
	Asian	10	12	27	14	13	10	12	8	16	26	6	2	6	6	10	13
	All	11	13	24	15	12	9	12	10	13	20	4	1	3	3	9	12
Portfolio & PLS	Black	12	13	16	12	17	22	47	72	68	40	3	1	1	1	3	32
	Hispanic	6	8	10	10	13	16	47	72	72	43	4	1	1	1	3	30
	Non-Hispanic white	15	19	19	22	24	23	49	64	58	34	5	2	2	3	7	31
	Asian	18	22	23	27	29	32	61	79	71	46	7	2	2	3	9	41
	All	13	17	18	19	21	22	48	67	62	36	4	2	2	2	6	31

Source: Matched loan-level HMDA-CoreLogic data.

In 2007, the PLCPB for FVR loans is 74 and 62 percent, respectively, for black and non-Hispanic white borrowers (table 4). Under the ODR, their denial rates are 26 and 19 percent, respectively (table 3), putting the denial rate of FVR loans for blacks at about 1.4 higher than the denial rate for non-Hispanic whites.

Plugging these numbers into equation 6 in box 2, however, allows us to calculate the PLCPA: 81 and 69 percent, respectively, for black and non-Hispanic white borrowers. So, for FVR loans, blacks have a higher proportion of LCP applicants than whites—specifically, about 1.2 times higher—similar to the difference in the denial rate as calculated by the traditional measure.

Black and Hispanic Groups Have More LCP Applicants

In fact, over our 15-year study period, the share of low-credit-profile applicants has been consistently higher in the black and Hispanic groups than in the Asian and non-Hispanic white groups. The average

PLCPAs for black, Hispanic, non-Hispanic white, and Asian FVR loan applicants over the 15 years are about 66, 57, 48, and 44 percent , respectively (table 4 and figure 8, left panel); the PCLPAs for black, Hispanic, non-Hispanic white, and Asian conventional loan applicants are 58, 41, 28, and 23 percent. So, the PLCPAs for black FVR and conventional loan applicants are about 1.4 and 2.1 times higher, respectively, than for non-Hispanic white applicants; they are about 1.5 and 2.5 times higher, respectively, than the PLCPAs for Asian FVR and conventional loan applicants. The above ratios are about 1.2, 1.3, 1.5, and 1.8 times higher, respectively, for Hispanic FVR and conventional loans applicants than for white and Asian applicants. This PLCPA pattern is reasonable; as shown in table A.6, the proportion of LCP consumers who want credit (PLCPD)—whether they apply for a loan or not—is higher for blacks and Hispanics than for non-Hispanic whites and Asians.

Within each race and ethnic group, the PLCPA is always higher for FVR channels than for conventional channels. On average over the 15-year period, the PLCPA is 1.1 for black applicants, 1.4 for Hispanics, 1.7 for non-Hispanic whites, and 1.9 times higher for Asians. The bottom half of table 4 also shows the share among LCP applicants and borrowers. Among black, Hispanic, non-Hispanic white, and Asian LCP applicants, only 6 percent, 6 percent, 5 percent, and 3 percent actually applied for FVR loans. Yet FVR loans account for 57, 47, 39, and 31 percent, respectively, of the loans received by black, Hispanic, non-Hispanic white, and Asian LCP borrowers.

The trend over time for PLCPA is consistent with our expectations of change over different periods of the credit cycles for all four racial and ethnic groups (left panels of figure 8 and table 4). We see a peak between 1998 and 2001 (PK1), followed by a trough between 2002 and 2003 (TR1), followed by another peak between 2004 and 2008 (PK2), and another trough from 2009 until now (TR2). The average PLCPA in PK1 is 72, 57, 39, and 29 percent, respectively, for black, Hispanic, non-Hispanic white and Asian applicants; in the same order of race and ethnicity, these numbers in PK2 are 67, 49, 39, and 31 percent; in TR1, they are 57, 41, 24, and 18 percent; and in TR2, they are 48, 37, 23, and 21 percent. On average, the peak-to-trough change is about 10 to 20 percentage points for all four racial and ethnic groups.

The period between 1998 and 2001 is noteworthy. This was the dot.com boom, during which mortgage interest rates[23] were considerably higher than in later periods, leading to lower mortgage originations.[24] The share of both LCP borrowers and applicants is higher in 2000 than periods before or after, suggesting that lenders loosened their underwriting standards to generate more business, encouraging more LCP consumers to apply for mortgage credit (see table 4).

Denial Rate for LCP Applicants

Our real denial rate, or RDR, improves on the traditional measure of credit accessibility by considering variations in applicants' credit profiles. Using the new measure, we see higher denial rates, more intuitive patterns of denial, and much less variation between racial and ethnic groups.

As expected, because the RDR measures only LCP applicants,[25] it reveals a much higher denial rate—on average about three times higher (71 percent vs. 23 percent)—than the ODR (table 3 and figure 6).

Different patterns also emerge. While the ODR tends to be higher during the housing boom years, the RDR is much lower, a more intuitive result. The RDR increases consistently between 1998 and 2012 except for two dips between 2000 and 2002 and between 2004 and 2006. After the financial crisis, the RDR reaches its highest level: an average 88 percent of LCP applicants were denied loans between 2009 and 2012. This rate is even higher for loans sold to the GSEs. Fully 99 percent of LCP applicants were denied GSE loans after 2009. Accessibility is higher in the FVR channels, with a 64 percent RDR for applications between 2009 and 2012—significantly better than the RDR for the conventional channels.

However, the denial rate nearly tripled in the FVR channel, from 19 percent between 1998 and 2007 to 64 percent between 2007 and 2009. (See upper right panel of figure 4 and table 3.) On average over the 15 years, the RDR is 32 percent for FVR channels, 89 percent for GSE channels, and 71 percent for PP channels. The RDR for the conventional channels is about two to three times higher than for the FVR channel (see table 3 and figure 4).

Denial Rates Show Minimal Variation by Race and Ethnicity

The RDR varies much less among the four race and ethnic groups than the ODR (see figure 3 and table 3). On average over the 15 years and in all channels combined, the RDR is 70 percent for black LCP applicants, 73 percent for Hispanic LCP applicants, 70 percent for non-Hispanic white LCP applicants, and 79 percent for Asian LCP applicants. For applications to the FVR channels, the averages are 35 percent, 34 percent, 31 percent and 36 percent, in the same order of race and ethnicity. For applications to the GSE channel, the averages are 89 percent, 92 percent, 90 percent, and 93 percent.

Under the traditional measure of accessibility, blacks appear to have a higher denial rate than non-Hispanic whites. But by looking exclusively at LCP applicants, we see that blacks and non-Hispanic whites tend to have the same denial rate in both the FVR and conventional channels. The real denial rate reveals that a high proportion of all low-credit-profile applicants are unable to access the mortgage market, no matter their race or ethnicity.

The RDR is also an imperfect measure because it groups all LCP applicants together and makes no distinctions among them. In addition, the RDR, like the ODR, fails to take into account another important way that people fail to access mortgage credit: they are deterred from applying in the first place, even though they need a mortgage. We propose a way to assess this deter rate below.

Deter Rate for LCP Consumers

Calculating the deter rate for low-credit-profile consumers uses the same mathematical logic as the denial rate for LCP applicants. We start by assuming that HCP consumers are unlikely to be deterred from applying for necessary loans. Accordingly, any consumers that are deterred from applying for loans must have a low credit profile. Mathematically, this means that the difference in the percentage of LCP consumers in the demand pool and in the application pool solely reflects a deterrence effect on LCP consumers. If LCP consumers have no chance of being deterred, then PLCPD would equal PLCPA; otherwise PLCPD is always higher than PLCPA. This mathematical relationship between PLCPD and PLCPA allows us to calculate the deter rate for LCP consumers with credit need.[26] We illustrate this calculation with some examples below.

According to the Survey of Consumer Finances, in 2004, 79 percent of black applicants who wanted mortgages had low credit profiles, compared with 50 percent of non-Hispanic whites who wanted mortgages. Plugging these numbers into equation 9 in box 2 gives us their deter rates: 53 percent for blacks and 42 percent for non-Hispanic whites.

Intuitive Trends over Time

Curve C in figure 6 and table 3 shows the deter rate for all four racial and ethnic groups between 1998 and 2012. One minus the deter rate is the demand-to-application progression rate (DAPR) for LCP individuals. Appendix table A.3 shows the DAPR for the same groups and time periods.[27] The general trend for the deter rate over different periods of the credit cycle is intuitive: it is lowest during boom years and highest when the market slows. The deter rate is 69 percent in 1998, 45 percent in 1999, and

20 percent (its lowest level) in 2000. The rate reaches 74 percent in 2003, descends to 41 percent in 2007, and starts ascending in 2008, reaching 83 percent by 2012.

Significant Variation by Race and Ethnicity

Low-credit-profile consumers from the four racial and ethnic groups show different deterrence responses to the credit cycle (table 3 and lower-left panel of figure 3). In 2000, black and non-Hispanic white LCP consumers had the lowest deter rates of 12 and 13 percent, respectively. In the same year, their Hispanic and Asian counterparts had much higher deter rates: 57 and 60 percent, respectively. While deter rates for each group changed over the credit cycle, the pattern of black and non-Hispanic white rates being lower than Hispanic and Asian rates remains consistent. On average over the 15 years, Hispanic and Asian LCP consumers have deter rates of 76 and 71 percent, respectively, compared with 55 and 57 percent, respectively, for black and non-Hispanic white LCP consumers. Hispanics have the least variation in their deter rate over the period. The racial gap on deter rate deserves further study and is discussed further below. After the financial crisis, the deterrence rate for all four racial and ethnic groups peaked at around 80 percent; in other words, more than 80 percent of LCP consumers who had some credit need were deterred from seeking a mortgage, at what should have been a very attractive time to buy a home (table 3).

Demand-to-Origination Progression Rate for LCP Consumers

So far we have examined deterrence and denial rates for low-credit-profile consumers. One minus the deter rate is the DAPR, and one minus the denial rate is the application-to-origination progression rate (AOPR). These two steps in the process give us the overarching demand-to-origination progression rate (DOPR) for LCP consumers. The DOPR is a complete measure of credit accessibility—although not of the quality of credit available—for LCP consumers.

The general trend of credit accessibility as measured by DOPR over time appears to follow the credit cycle (table A.3 and curve D in figure A.3). The DOPR is at 10 percent in 1998, peaks at 38 percent in 2000, falls to 8 percent by 2003, increases to 20 percent from 2004 through 2006, then drops to 14 percent in 2007 and to single digits thereafter, staying steady at 2 percent since 2010. These results reveal extreme variability in credit accessibility over different periods of the credit cycle: the DOPR peak is about 19 times higher than the lowest DOPR over the 15 years. The results show that credit has been tightest since 2009. With only about a 2 percent overall progression rate since the financial crisis, the mortgage credit market has virtually closed its door to LCP consumers. In contrast, during the 2004–07 housing boom years, about one in five LCP consumers with credit need were able to obtain mortgage credit, although the quality of those loans varied greatly. The same trend holds across all racial and ethnic groups and loan channels.

Over the 15-year period, the DOPR averages 15 percent for black LCP consumers, 7 percent for Hispanic LCP consumers, 15 percent for non-Hispanic white LCP consumers, and 6 percent for Asian LCP consumers (table A.3 and figure A.4). Blacks and non-Hispanics white have about the same DOPR, which is more than two times higher than their Hispanic and Asian counterparts. This pattern is the same when we examine by channels. In the same order of race and ethnicity, the DOPR averages 32 percent, 17 percent, 32 percent, and 20 percent for FVR channels; 6 percent, 2 percent, 6 percent, and 3 percent for GSE channels; and 16 percent, 7 percent, 18 percent, and 8 percent for PP channels. Since, as discussed under "Denial Rate for LCP Applicants," the racial gap in application deterrence is much more significant than the racial gap in application denial, application deterrence clearly drives the racial gap in the DOPR.

When all four racial and ethnic groups are combined, the average DOPR over the 15-year period is 30, 6, and 13 percent, respectively, for FVR, GSE, and PP channels. In other words, LCP consumers with credit need are five times more likely to obtain mortgage credit through the FVR channels than through GSE channels (table A.3 and figure A.5). The DOPR is highest in 2000 and 2006 for FVR (67 percent and 45 percent) and GSE loans (28 percent and 12 percent), and in 2000 and 2005 for PP loans (39 percent and 23 percent). Over the 15-year period, all three channels reached their lowest DOPR between 2009 and 2012: about 7 percent for FVR channels, 0.3 percent for GSE channels, and 0.4 percent for the PP channels.

In general, as a measurement of credit accessibility, the DOPR is both more granular and more comprehensive than traditional methods. The DOPR reveals extreme variability in access among groups and economic conditions, which deserve further study.

Contributions to Credit Accessibility for LCP Consumers by Channels

Another way to measure credit accessibility within the three channels is to calculate the share of LCP loans originated by each channel (see table 4 and the right panels of figure 7). Our analysis shows that, through time, on average, LCP loans by the GSEs accounted for only 12 percent of all LCP loans. The GSEs had the highest share of LCP loans (24 and 20 percent) in 2000 and 2007. In 2000, they accounted for 13 percent of LCP loans to blacks and Hispanics, 28 percent of LCP loans to non-Hispanic whites, and 27 percent of LCP loans to Asians. In 2007, as the housing market was imploding, these numbers were 17 percent, 18 percent, 22 percent and 26 percent. Even in their best years, GSE loans accounted for only 13 percent to 28 percent of total LCP loans. On average over the period, GSE loans accounted for only 9 percent, 8 percent, 14 percent, and 13 percent, respectively, of the LCP loans to black, Hispanic, non-Hispanic white, and Asian borrowers (see table 4 and the right panels of figure 7).

In 2009 the GSEs guaranteed the smallest share of loans to LCP consumers, accounting for only 1 percent of all LCP loans to blacks, Hispanics, and non-Hispanic whites and 2 percent of all LCP loans to Asians. Since the total number of loans to LCP borrowers during 2009 was small, this means that the GSEs almost completely stopped making loans to LCP consumers. The GSE's share of loans to LCP borrowers inched up in 2009 and 2012 yet remained below 7 percent for blacks and Hispanics and below 10 percent for non-Hispanic whites and Asians (see table 4).

If the GSEs are not supporting lending to LCP borrowers, who is? The answer varies by time, race and ethnic group.

During the boom years, the portfolio and private-label security channel was the major mortgage credit provider to LCP consumers (see table 4). (Again, this measure does not take the quality of loans into consideration.) In 2005, the year with the highest share of PP loans, the PP channel accounted for 72 percent of all LCP loans to blacks and Hispanics, 64 percent to non-Hispanic whites, and 79 percent to Asians. The PP channel was the dominant source of mortgage credit to LCP consumers in the boom, accounting for 48, 67, 62, and 36 percent, respectively, of all loans to LCP borrowers in 2004, 2005, 2006 and 2007.

In all other years, the PP channel played a smaller role. For Asian and non-Hispanic white LCP borrowers, the PP channel accounted for about 15–32 percent of all LCP loans between 1998 and 2003. Over the same period, PP loans accounted for about 12–22 percent of LCP loans to blacks. This number is about 6–16 percent for Hispanic LCP borrowers, a smaller share than the FVR channel. Since the financial crisis, PP lending to LCP borrowers, like GSE lending, has just about dried up (table 4 and the right panels of figure 7).

Except during the housing boom, the FVR channel has been the major channel for lending to LCP consumers; it has played an especially important role since the financial crisis. From 2008 onward, FVR loans have made up of about 93 percent of all LCP loans. This number is even higher for black and Hispanic LCP borrowers. Nevertheless, it is important to remain aware that, as shown in table 6, only very low levels of demand for lending by LCP consumers are being met at all (table 4 and the right panels of figure 7).

Discussion

Better Measures of Credit Accessibility

This paper measures credit accessibility by three progression rates: the rate from demand to application (DAPR), from application to origination (AOPR), and from demand to origination (DOPR). While these rates do not take loan quality into account, they do offer a more robust measure of accessibility that takes into account the credit profiles of consumers and the impact when consumers are deterred from applying for loans. These measures thus provide policymakers with powerful new tools to assess credit accessibility over time and by demographic group.

First, these measures show a more intuitive trend over time than that shown by traditional measure of credit, the observed denial rate. The ODR is higher in the housing boom years than in the non-boom years, which is inconsistent with the observation that lenders loosened their underwriting standards during the boom. The DAPR, AOPR, and DOPR show a trend consistent with a loosening of credit standards during the boom, with a peak between 1998 and 2001, a trough between 2002 and 2003, another peak between 2004 and 2008, and another trough from 2009 to the present.

Second, these measures reveal unmet demand through various lending cycles. For example, between 1998 and 2012, the highest DOPR is about 15 to 19 times higher than its lowest value for any of the four racial and ethnic groups, showing how much more demand was met in some periods than others. Whether this demand should be more effectively met to promote homeownership and grow the economy, or constrained to manage risk in the system, is a critical decision for policymakers. To decide, however, policymakers need a better measure of where demand is not being met.

Racial Gaps in Credit Accessibility

For many years, many researchers have made the case that credit accessibility is far from uniform among different demographic groups (Munnell et al. 1996), in particular that minorities have not had the same level of access to the credit market as have their white counterparts. Our results based on the ODR supports this view. Between 1998 and 2012, for example, the ODR for black applicants is, on average, about two times higher than for non-Hispanic whites and Asians. We also find, however, that a higher percentage of minority applicants have weaker credit profiles than non-Hispanic white applicants, suggesting that a closer look at the numbers is warranted.

If we look only at applicants with weaker credit profiles, we find different and surprising results. First, when only applicants with weaker credit profile are considered, the differences among the denial rates of the four race and ethnicity groups almost disappear. Second, of the four groups considered, Hispanics and Asians with weaker credit profiles are deterred from applying for a mortgage more frequently than blacks and non-Hispanic whites.

Blacks and non-Hispanic whites with low credit profiles who want credit progress at about the same rate from demand to origination (DOPR)—a rate about two times higher than that of their Hispanic and Asian counterparts. This racial difference in overall progression rates is driven by the racial gap in deter rates.

The findings of lower progression rates for Hispanics and Asians deserve further study. Are the lower rates the result of the multiple incomes many Hispanic families rely upon creating questions for lenders? Do concerns about immigration status deter mortgage applications? Are cultural differences at play, such as a tendency to borrow less and save more than the majority population?

It is also important to note that a higher progression rate does not necessarily lead to a better financial outcome. On the contrary, LCP consumers who borrowed at the peak of the housing boom encountered a declining housing market and, hence, a higher risk of foreclosure. This is confirmed by our results on the rising role of private-label securities (PLS) market on credit accessibility to LCP consumers during the boom. By not taking the price and quality of loans into account, our measures of higher credit accessibility are implicitly giving the market credit for oversupply of bad products. All these issues deserve further study.

Credit Accessibility by Channels

Our research also reveals interesting, mixed results about the government's success in promoting credit accessibility to disadvantaged groups. The FVR channel serves low-credit-profile borrowers reasonably well, especially blacks and Hispanics. But the GSE channel does not serve LCP borrowers of any race or ethnicity with particular vigor.

Of the three channels considered, denial rates for LCP applications through the GSE and PP channels are three and two times higher than for LCP applications through the FVR channels. LCP consumers that need a loan are five times more likely to obtain a mortgage through FVR channels than through GSE channels. Over time, on average, the GSEs satisfied only 12 percent of the demand for loans from LCP consumers. Even in their best years, GSE loans met only about 20 percent of the demand for loans by LCP consumers.

During the boom years of 2004–07, the PP channel was the dominant source of mortgage credit to LCP consumers,[28] accounting for more than 60 percent of all loans to LCP borrowers. In all other years, the PP channel played a smaller role, accounting for less than 20 percent of all loans to LCP borrowers. Except for the housing boom years, the major channel for lending to LCP consumers has been through the FVR.

The FVR role has been especially important since the financial crisis. Since 2008, FVR loans have made up about 93 percent of all loans to LCP consumers and an even higher percentage for black and Hispanic LCP consumers. Despite the presence of FVR loans, very little of the demand for lending by LCP consumers is being met.

Effectiveness of Post-Crisis Efforts to Open the Credit Box

All three progression rates show that the mortgage credit market has been tighter since 2009 than in any prior period since 1998. For individuals with weaker credit profiles, the average DOPR between 1998 and 2001 was about ten times higher than the current 2 percent rate. The DOPR was 4.5 times higher than today in 2002 and 2003 and seven times higher from 2004 to 2008.

Since the financial crisis, the mortgage credit market has virtually closed its doors to weaker credit profile consumers, reducing the number of buyers and, thereby, lengthening the housing recovery. When affordable but rising home prices make homeownership an attractive proposition, this closed door means an excellent wealth-building tool is out of reach for many consumers with lower credit profiles.

A major factor behind the current tight credit box is the diminished PLS market, which before the financial crisis accounted for more than 80 percent of all mortgages to borrowers with weaker credit profiles. While the quality of many of those mortgages was questionable; the extreme reduction in this segment has vastly diminished mortgage availability. With the PLS channel all but closed, the FVR channel remains the primary channel through which these borrowers receive loans. However, even the FVR channel is producing very few loans to LCP borrowers, and the GSE channel is producing fewer still.

Conclusions

Access to credit is central not just to the nation's housing policy, but also to its broader economic agenda. If access is too broad, the system takes on too much risk; If access is too narrow, demand is choked off and economic health falters. The ability to accurately measure access to credit allows policymakers to strike this delicate balance.

This paper provides a method of measuring credit accessibility that addresses several shortcomings of traditional methods. Credit accessibility is measured by the rates at which prospective borrowers progress through the process of getting a loan: from demand to application (the DAPR); from application to origination (the AOPR); and, combining both, from demand to origination (the DOPR). This result is an analysis that is both more granular and more comprehensive, revealing temporal and demographic trends that raise as many questions as they answer.

Appendix: Data and Definitions

TABLE A.1

Demand-to-Application Progression Rates (%)

Origination year	Black	Hispanic	Non-Hispanic white	Asian
98	48(35, 62)	27(18, 37)	31(29, 33)	17(9, 26)
99	66(53, 80)	32(21, 42)	59(55, 63)	28(15, 41)
00	88(72, **)	43(28, 58)	87(81, 94)	40(21, 59)
01	51(41, 62)	24(16, 32)	41(38, 44)	21(11, 31)
02	31(25, 37)	18(13, 23)	28(26, 31)	16(10, 22)
03	29(23, 34)	16(12, 21)	28(25, 30)	17(11, 24)
04	47(38, 57)	22(16, 29)	58(53, 62)	31(20, 43)
05	59(46, 72)	29(22, 35)	57(53, 62)	47(33, 60)
06	60(47, 74)	32(25, 40)	61(56, 66)	54(39, 69)
07	64(50, 79)	40(31, 48)	58(53, 63)	55(40, 71)
08	50(40, 59)	29(23, 36)	45(42, 48)	33(24, 42)
09	25(20, 29)	16(13, 20)	23(21, 24)	19(14, 25)
10	22(18, 26)	14(10, 17)	22(21, 23)	17(12, 22)
11	22(18, 26)	13(10, 16)	23(21, 24)	18(13, 23)
12	17(14, 20)	11(8, 13)	19(18, 20)	15(11, 20)

Sources: Matched loan-level HMDA-CoreLogic data and Survey of Consumer Finance.

Note: 95% confidence intervals in parentheses.

TABLE A.2

Overall Demand-to-Origination Progression Rates (%)

Orig.	FHA & VA & RD				GSE				Portfolio & PLS			
year	Black	Hispanic	NHW	Asian	Black	Hispanic	NHW	Asian	Black	Hispanic	NHW	Asian
98	40(29, 51)	22(15, 30)	27(25, 28)	14(7, 21)	4(3, 6)	1(.9, 2)	2(2, 3)	.9(.5, 1)	18(13, 22)	8(5, 11)	13(12, 14)	5(3, 8)
99	53(42, 64)	26(17, 35)	50(46, 53)	22(12, 33)	10(8, 12)	3(2, 4)	8(7, 9)	3(2, 4)	27(21, 32)	12(8, 17)	30(28, 33)	10(5, 15)
00	69(56, 78)	36(23, 48)	73(67, 78)	33(18, 49)	24(19, 27)	11(7, 15)	30(27, 32)	12(7, 18)	39(32, 44)	20(13, 27)	50(46, 54)	21(11, 30)
01	43(34, 52)	20(13, 27)	36(34, 39)	18(10, 27)	8(6, 10)	2(1, 3)	6(6, 7)	2(1, 3)	24(19, 29)	12(8, 16)	32(30, 35)	11(6, 17)
02	26(20, 31)	15(11, 19)	25(23, 27)	14(9, 19)	4(3, 4)	1(.9, 2)	3(3, 3)	1(.6, 1)	19(15, 23)	10(7, 12)	23(21, 25)	8(5, 11)
03	24(19, 28)	13(10, 17)	24(22, 26)	15(9, 20)	2(1, 2)	.6(.4, .8)	1(1, 1)	.4(.3, .6)	13(10, 15)	6(4, 8)	13(12, 14)	5(3, 7)
04	36(29, 44)	17(12, 22)	47(43, 51)	25(16, 34)	7(5, 8)	2(2, 3)	6(6, 7)	2(1, 3)	23(18, 27)	10(7, 13)	32(29, 35)	13(8, 18)
05	44(34, 53)	21(16, 26)	46(42, 49)	36(25, 46)	9(7, 11)	3(2, 4)	7(6, 8)	3(2, 4)	27(21, 33)	11(9, 13)	30(28, 33)	16(12, 21)
06	46(35, 56)	25(19, 30)	49(45, 52)	43(31, 55)	13(10, 15)	5(4, 6)	12(11, 13)	8(6, 10)	24(19, 30)	10(8, 13)	28(25, 30)	16(12, 21)
07	43(34, 53)	27(21, 33)	42(39, 46)	37(27, 48)	11(9, 14)	5(4, 6)	11(10, 11)	6(4, 7)	17(13, 21)	7(6, 9)	15(14, 17)	8(6, 11)
08	27(22, 33)	16(12, 20)	27(25, 29)	17(12, 21)	2(1, 2)	.7(.6, .9)	2(2, 2)	.6(.4, .7)	4(3, 4)	2(1, 2)	4(3, 4)	2(1, 2)
09	10(8, 12)	6(5, 8)	9(9, 10)	6(4, 7)	.1(.1, .2)	.1(.0, .1)	.1(.1, .1)	.1(.0, .1)	.5(.4, .6)	.2(.2, .3)	.6(.5, .6)	.2(.2, .3)
10	8(6, 9)	4(3, 5)	8(7, 8)	4(3, 6)	.3(.2, .3)	.1(.1, .2)	.3(.3, .3)	.1(.1, .2)	.3(.3, .4)	.1(.1, .1)	.3(.3, .3)	.1(.1, .2)
11	7(6, 9)	4(3, 5)	8(8, 9)	5(4, 6)	.4(.3, .4)	.1(.1, .2)	.3(.3, .4)	.1(.1, .2)	.5(.4, .6)	.2(.1, .2)	.5(.4, .5)	.2(.1, .2)
12	6(5, 7)	4(3, 5)	7(7, 8)	4(3, 5)	.5(.4, .6)	.2(.2, .3)	.5(.5, .5)	.2(.1, .2)	.7(.6, .9)	.3(.3, .4)	.9(.8, .9)	.4(.3, .5)

Sources: Matched loan-level HMDA-CoreLogic data and Survey of Consumer Finance.

Note: 95% confidence intervals in parentheses. NHW = non-Hispanic white.

TABLE A.3

Four Progression Rates Measuring Credit Accessibility

Channel	Race and ethnicity	Origination Year 98	99	00	01	02	03	04	05	06	07	08	09	10	11	12
		Observed Application-to-Origination Progression Rate for All Applicants (%)														
FHA & VA & RD	Black	87	85	82	88	87	87	82	80	81	74	67	71	69	71	74
	Hispanic	90	88	87	90	89	88	84	81	83	76	71	75	76	78	79
	NHW	93	91	89	93	92	92	88	86	86	81	77	80	79	80	83
	Asian	91	89	89	92	91	91	87	86	87	79	72	76	77	78	79
	All	92	89	88	92	91	91	86	85	85	79	74	78	78	79	81
Conventional	Black	56	56	55	61	68	64	61	62	59	51	43	52	56	55	59
	Hispanic	66	66	64	72	77	74	72	71	68	59	51	60	65	65	69
	NHW	81	78	77	85	88	85	78	77	75	72	71	81	81	80	82
	Asian	85	83	82	86	89	86	81	79	76	72	71	80	82	81	83
	All	79	75	74	82	86	83	75	74	72	68	67	78	79	78	80
All	Black	58	57	56	63	69	65	61	62	60	52	45	56	58	58	62
	Hispanic	68	67	66	73	77	75	72	71	68	59	53	63	67	67	70
	NHW	82	78	77	85	88	85	78	77	75	72	72	81	80	80	82
	Asian	85	83	82	86	89	86	81	79	76	72	71	80	81	81	83
	All	79	76	74	83	86	83	76	74	72	68	67	78	79	78	80
		Application-to-Origination Progression Rate for LCP Applicants (%)														
FHA & VA & RD	Black	82	80	78	84	83	83	76	74	75	67	55	40	35	35	34
	Hispanic	83	82	83	86	84	82	77	73	77	69	54	37	33	34	35
	NHW	86	84	83	89	87	87	81	79	80	73	60	40	35	35	38
	Asian	82	80	83	86	86	83	79	76	79	68	50	30	26	28	28
	All	85	83	82	87	86	85	80	77	79	71	58	40	34	35	36
GSE	Black	9	15	27	15	11	6	14	15	21	17	4	.5	1	2	3
	Hispanic	5	9	25	9	7	4	10	11	16	12	2	.3	1	1	2
	NHW	8	13	34	16	11	5	11	12	19	18	4	.5	1	1	3
	Asian	6	11	31	9	6	2	7	7	15	10	2	.3	.8	.8	1
	All	8	15	35	18	13	5	14	15	20	17	3	.5	1	1	3
Portfolio & PLS	Black	36	40	44	47	60	44	48	46	40	26	7	2	2	2	4
	Hispanic	31	39	46	49	55	38	44	38	32	18	6	1	.8	1	3
	NHW	43	52	57	79	81	47	56	53	45	27	8	3	1	2	5
	Asian	31	37	52	54	50	31	42	35	30	15	5	1	.9	.9	3
	All	34	41	48	52	53	36	44	43	38	23	7	2	1	2	4
Conventional	Black	26	31	36	33	34	26	37	37	33	20	5	2	3	3	3
	Hispanic	16	24	33	26	27	20	32	29	26	14	4	1	1	1	2
	NHW	22	32	45	39	36	21	35	35	33	20	6	2	2	2	3
	Asian	15	25	40	26	24	15	27	25	23	11	3	.7	.8	.9	1
	All	22	31	42	36	34	21	34	34	31	19	6	2	2	2	3
All	Black	34	38	41	42	43	33	40	38	35	24	21	17	16	15	12
	Hispanic	33	38	44	42	41	30	36	31	28	18	18	14	13	12	12
	NHW	31	39	50	47	44	29	39	38	36	26	23	15	12	12	11
	Asian	25	34	46	34	30	19	30	26	24	13	12	7	6	6	5
	All	31	39	48	45	43	30	38	36	34	24	22	14	12	12	11
		Demand-to-Application Progression Rate for LCP Consumers (%)														
All	Black	48	66	88	51	31	29	47	59	60	64	50	25	22	22	17
	Hispanic	27	32	43	24	18	16	22	29	32	40	29	16	14	13	11
	NHW	31	59	87	41	28	28	58	57	61	58	45	23	22	23	19
	Asian	17	28	40	21	16	17	31	47	54	55	33	19	17	18	15
	All	31	55	80	38	26	26	52	54	59	59	43	20	19	20	17
		Overall Demand-to-Origination Progression Rate for LCP Consumers (%)														
FHA & VA & RD	Black	40	53	69	43	26	24	36	44	46	43	27	10	8	7	6
	Hispanic	22	26	36	20	15	13	17	21	25	27	16	6	4	4	4
	NHW	27	50	73	36	25	24	47	46	49	42	27	9	8	8	7
	Asian	14	22	33	18	14	15	25	36	43	37	17	6	4	5	4

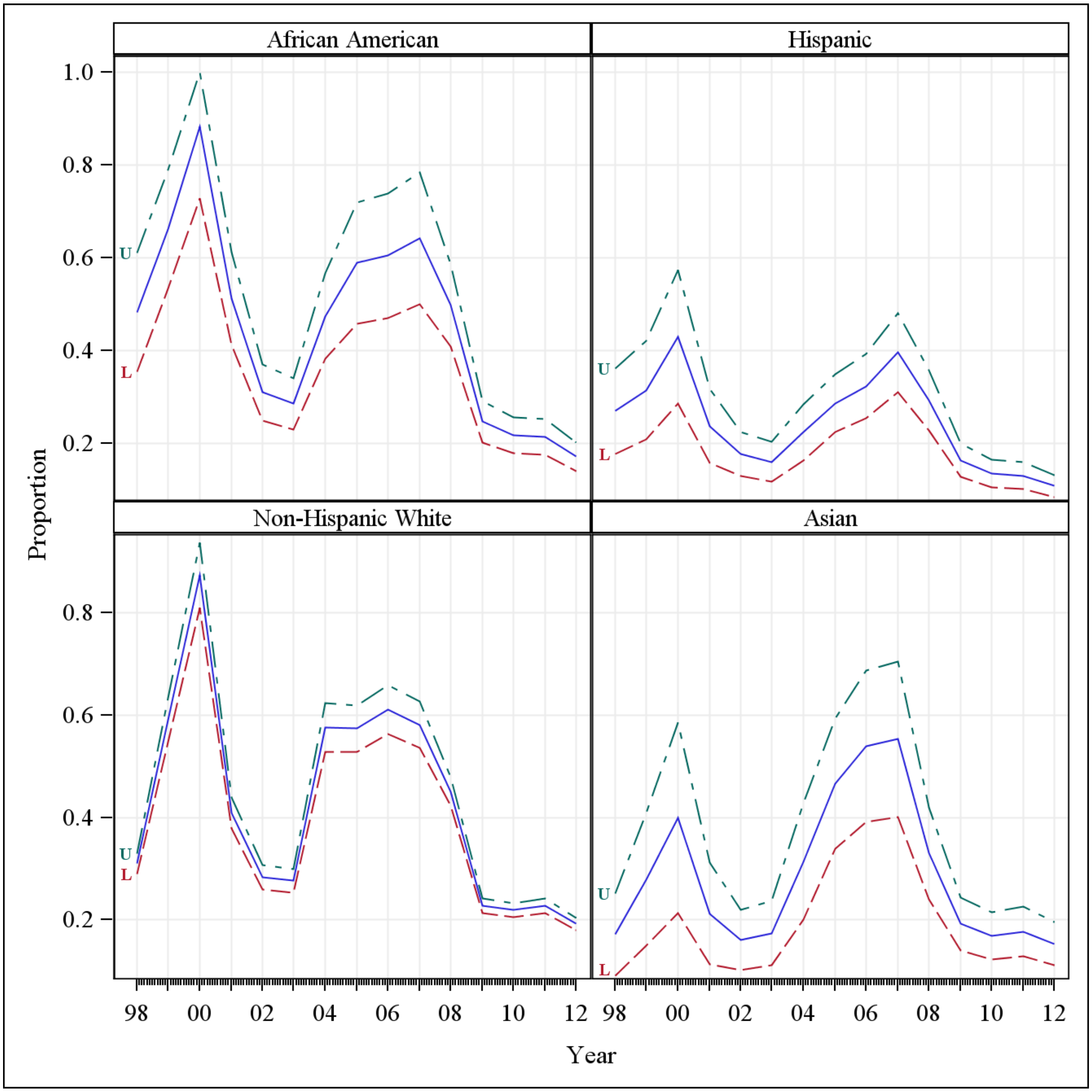
Channel	Race and ethnicity	Origination Year														
		98	99	00	01	02	03	04	05	06	07	08	09	10	11	12
	All	28	47	67	35	23	22	40	42	45	41	25	9	7	7	6
GSE	Black	4	10	24	8	4	2	7	9	13	11	2	.1	.3	.4	.5
	Hispanic	1	3	11	2	1	.6	2	3	5	5	.7	.1	.1	.1	.2
	NHW	2	8	30	6	3	1	6	7	12	11	2	.1	.3	.3	.5
	Asian	.9	3	12	2	1	.4	2	3	8	6	.6	.1	.1	.1	.2
	All	3	9	28	7	4	1	7	8	12	10	2	.1	.3	.3	.4
Portfolio & PLS	Black	18	27	39	24	19	13	23	27	24	17	4	.5	.3	.5	.7
	Hispanic	8	12	20	12	10	6	10	11	10	7	2	.2	.1	.2	.3
	NHW	13	30	50	32	23	13	32	30	28	15	4	.6	.3	.5	.9
	Asian	5	10	21	11	8	5	13	16	16	8	2	.2	.1	.2	.4
	All	11	23	39	21	14	9	23	23	22	13	3	.5	.3	.4	.7
Conventional	Black	12	21	32	17	11	7	17	22	20	13	3	.5	.6	.6	.6
	Hispanic	4	7	14	6	5	3	7	8	9	6	1	.2	.2	.2	.2
	NHW	7	19	39	16	10	6	20	20	20	12	3	.4	.3	.4	.6
	Asian	3	7	16	6	4	3	9	12	12	6	1	.1	.1	.2	.2
	All	7	17	34	15	9	6	17	18	18	11	2	.3	.3	.4	.5
All	Black	16	25	37	21	13	10	19	23	21	16	10	4	4	3	2
	Hispanic	9	12	19	10	7	5	8	9	9	7	5	2	2	2	1
	NHW	10	23	43	19	13	8	22	22	22	15	11	3	3	3	2
	Asian	4	10	18	7	5	3	9	12	13	7	4	1	1	1	.8
	All	10	21	38	17	11	8	20	20	20	14	9	3	2	2	2

Sources: Matched loan-level HMDA-CoreLogic data and Survey of Consumer Finance.

Note: NHW = non-Hispanic white.

FIGURE A.1

Demand-to-Application Progression Rates with Confidence Intervals

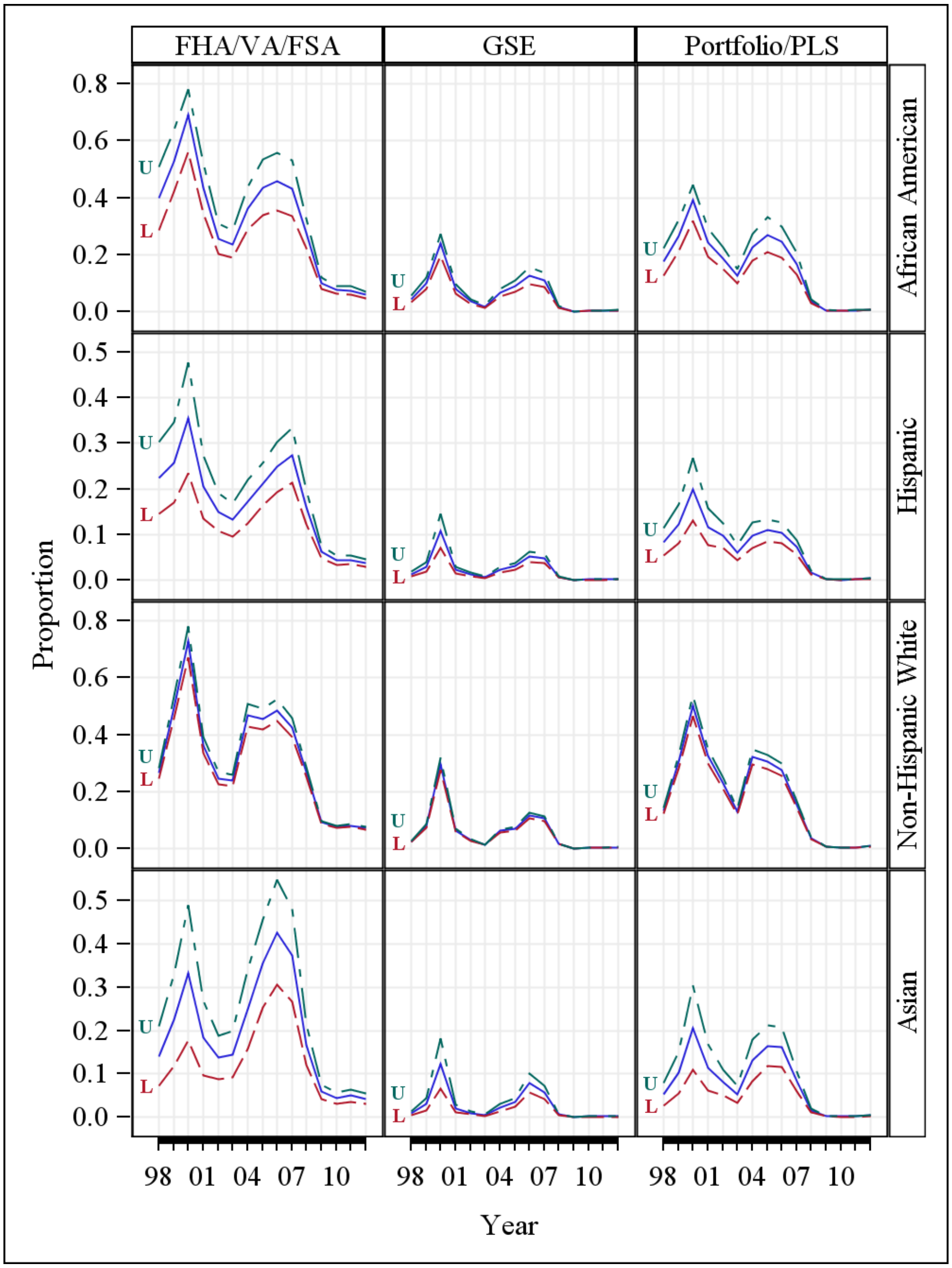


Sources: HMDA and CoreLogic.

Note: L = lower 95% confidence limit; U = upper 95% confidence limit.

FIGURE A.2

Over Demand to Origination Progression Rate with Confidence Intervals

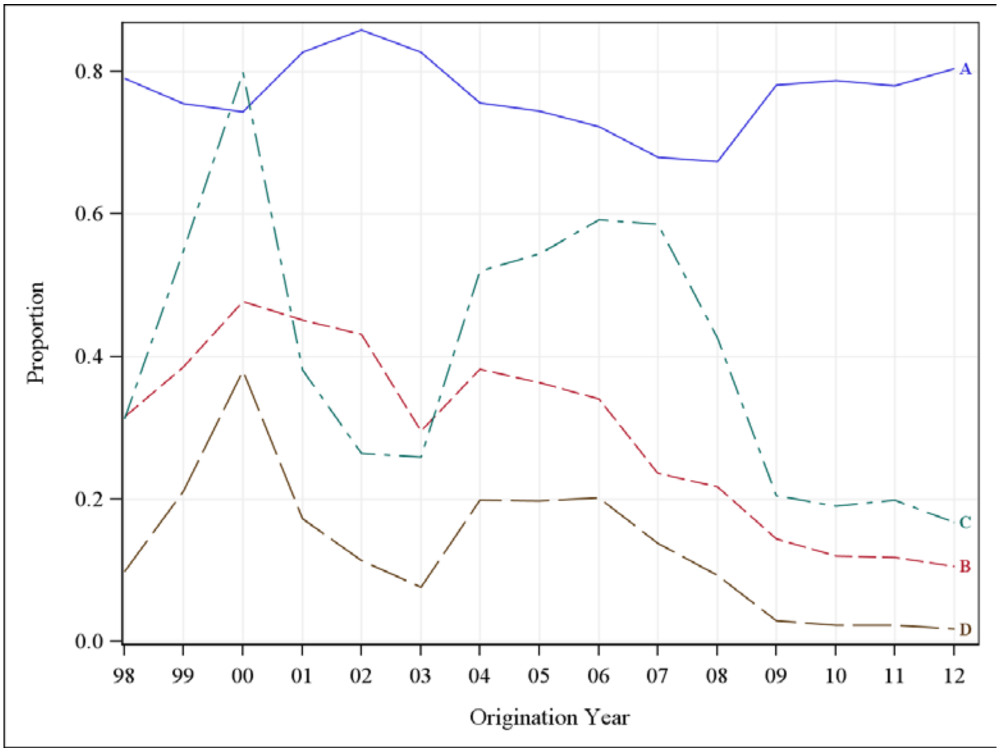


Sources: HMDA and CoreLogic.

Note: L = lower 95% confidence limit; U = upper 95% confidence limit.

FIGURE A.3

Measuring US Mortgage Credit Accessibility by Four Progression Rates

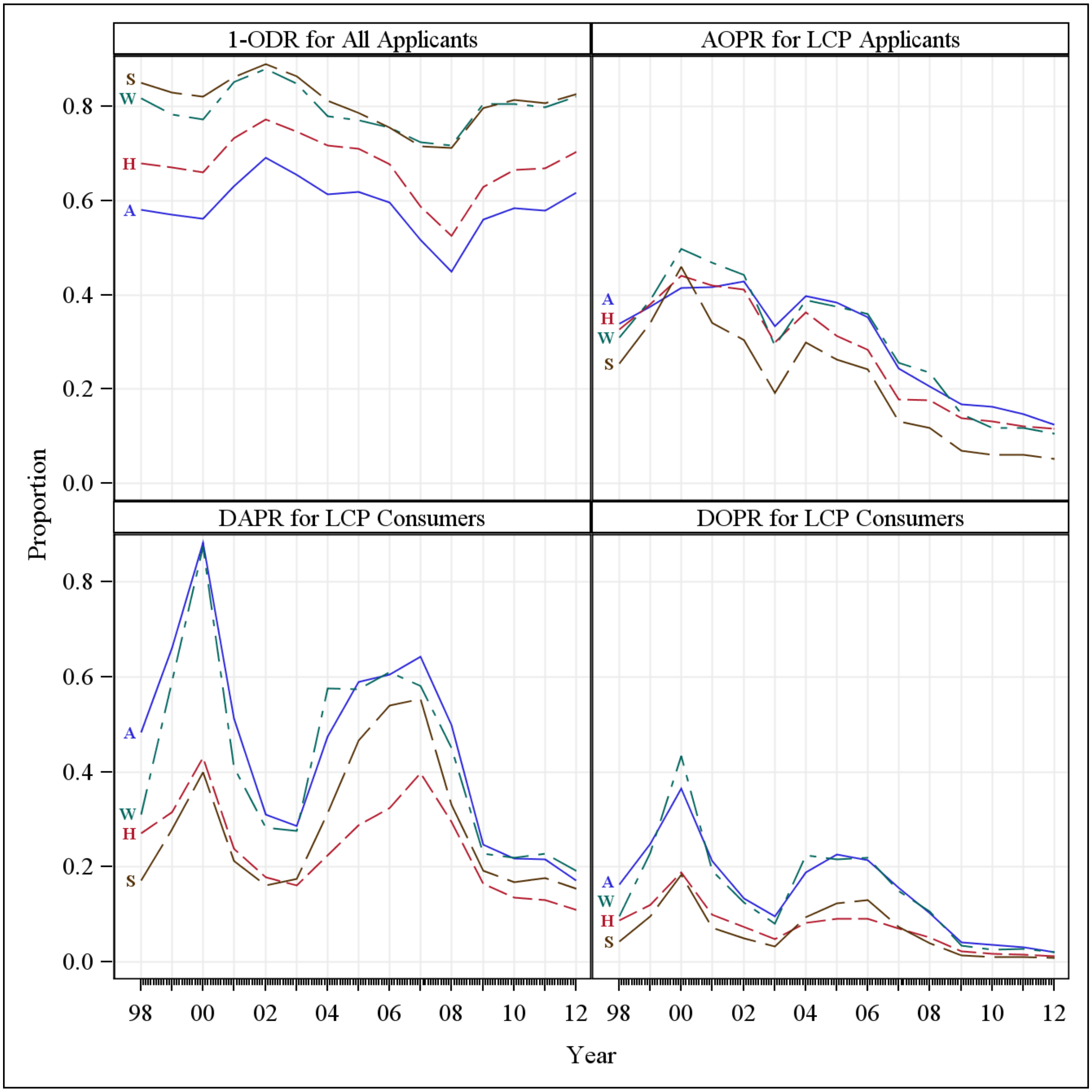


Sources: HMDA and CoreLogic.

Notes: A = observed application-to-origination progression rate for all applicants; B = application-to-origination progression rate for LCP applicants; C = demand-to-application progression rate for LCP applicants; D = overall demand-to-origination progression rate.

FIGURE A.4

Four Progression Rates by Race and Ethnicity

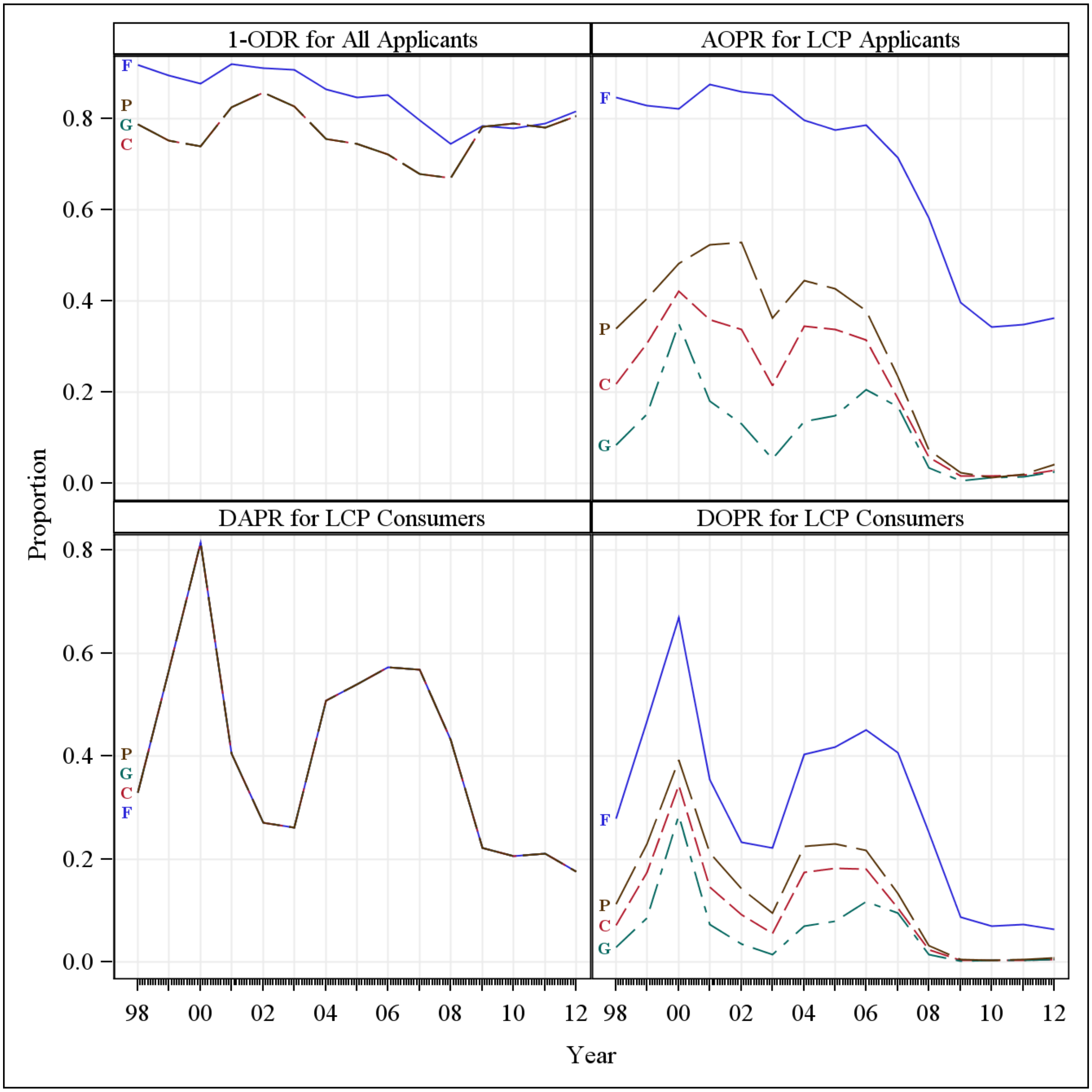


Sources: HMDA and CoreLogic.

Notes: A = black; H = Hispanic; S = Asian; W = non-Hispanic white. LCP = low credit profile. ODR = observed application-to-origination progression rate for all applicants; AOPR = application-to-origination progression rate for LCP applicants; DAPR = demand-to-application progression rate for LCP applicants; DOPR = overall demand to origination progression rate.

FIGURE A.5

Four Progression Rates by Channels

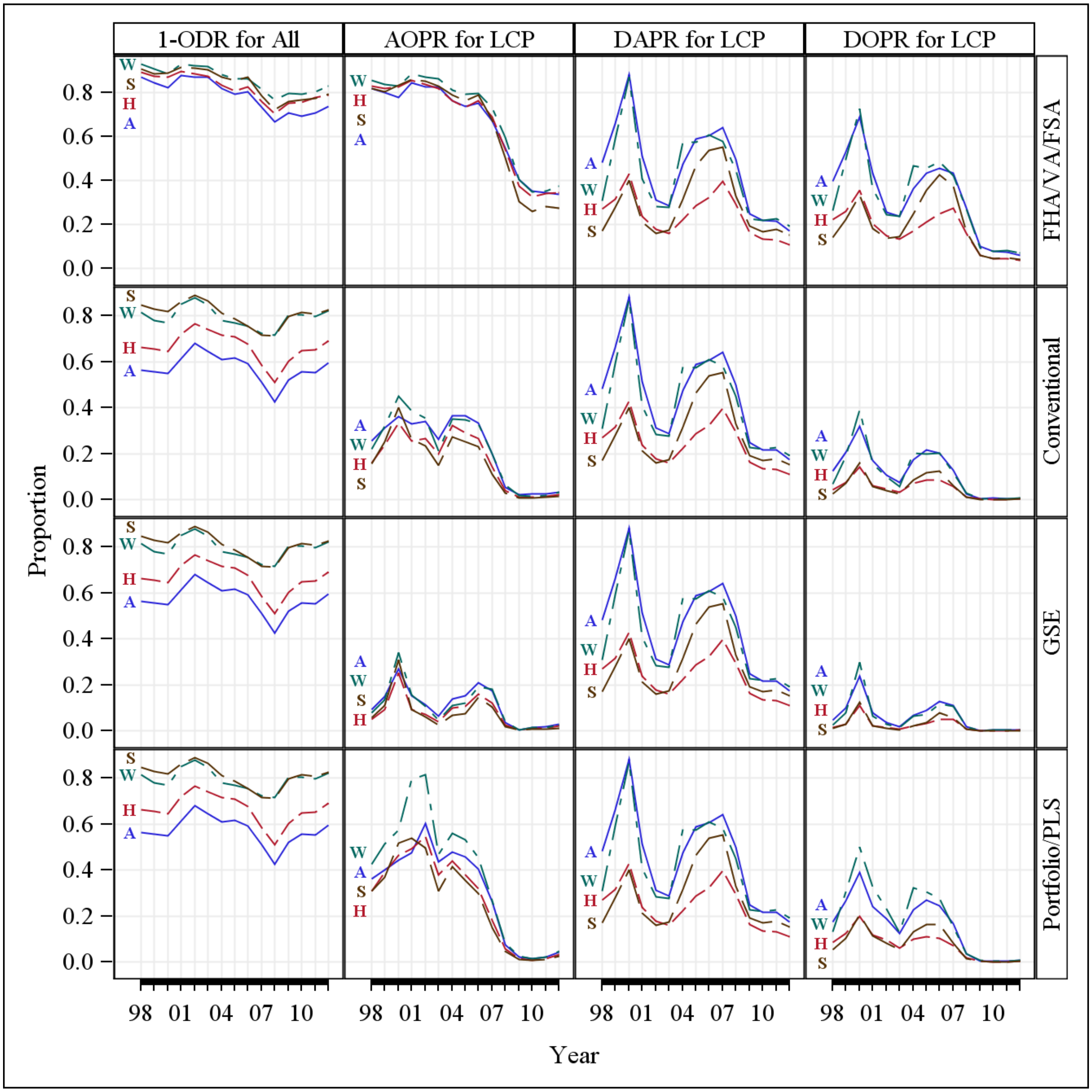


Sources: HMDA and CoreLogic.

Notes: C = conventional (GSE, portfolio, and PLS); F = FHA/VA/RD; G = GSE; P = portfolio/PLS. LCP = low credit profile. ODR = observed application-to-origination progression rate for all applicants; AOPR = application-to-origination progression rate for LCP applicants; DAPR = demand-to-application progression rate for LCP applicants; DOPR = overall demand to origination progression rate.

FIGURE A.6

Four Progression Rates by Race/Ethnicity and Channel

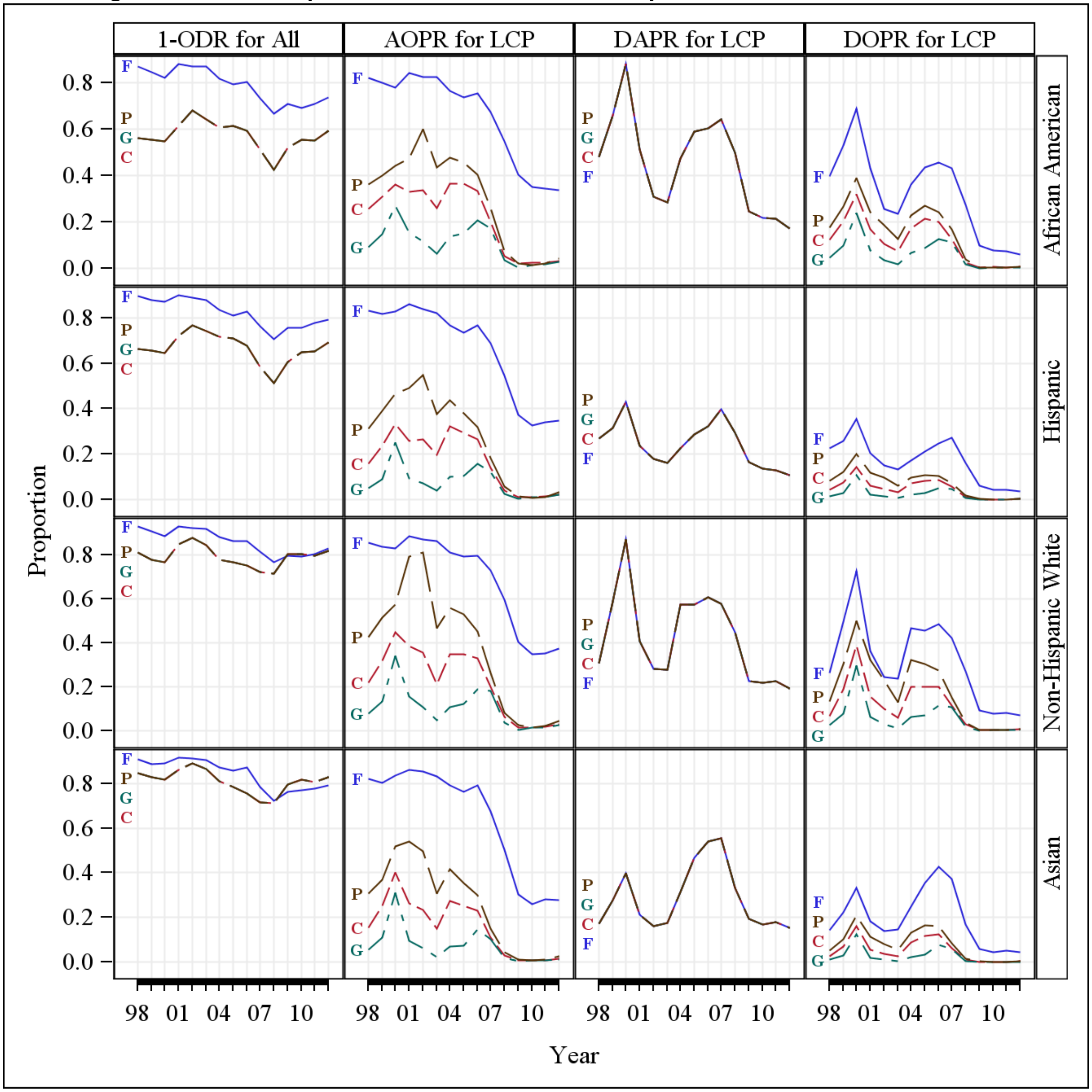


Sources: HMDA and CoreLogic.

Notes: A = black; H = Hispanic; S = Asian; W = non-Hispanic white. LCP = low credit profile. ODR = observed application-to-origination progression rate for all applicants; AOPR = application-to-origination progression rate for LCP applicants; DAPR = demand-to-application progression rate for LCP applicants; DOPR = overall demand to origination progression rate.

FIGURE A.7

Four Progression Rates by Channel and Race/Ethnicity



Sources: HMDA and CoreLogic.

Notes: C = conventional (GSE, portfolio, and PLS); F = FHA/VA/RD; G = GSE; P = portfolio/PLS. LCP = low credit profile. ODR = observed application-to-origination progression rate for all applicants; AOPR = application-to-origination progression rate for LCP applicants; DAPR = demand-to-application progression rate for LCP applicants; DOPR = overall demand to origination progression rate.

Data

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 except some small lenders that are exempt. CoreLogic covers the vast majority 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 supplemented information in either database with information from the other using this weight.

We match every HMDA loan to every CoreLogic loan to create a Cartesian product of the two databases. We then filter out those matches where the common fields between the two databases are inconsistent with each other. First, if a matched pair of loans originated from different years, the pair is dropped out of the matched loan database. Second, if a matched pair of loans has a loan amount difference equal or greater than \$2,000, the pair is dropped out. Third, matched loan pairs pass through a "geographic filter"—that is, a matched pair of loans with properties from different geographic locations is dropped out of the matched loan database. Because HMDA reports data by census tract and CoreLogic by zip code, the geographic filter is not straightforward. To solve this issue, we used HUD'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 ith HMDA loan from census tract X_i matched to the jth CoreLogic loan from zip code Y_j, $i = 1,\ldots I$, $j = 1,\ldots J$. 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

$$P_i = {Z_{ij}}/{X_i} \qquad \text{(A.1)}$$

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

$$P_j = {Z_{ij}}/{Y_j} \qquad \text{(A.2)}$$

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

$$P_{ij} = {Z_{ij}^2}/{X_i Y_j} \quad \text{(A.3)}$$

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

For the other common variables between HMDA and CL, we adopted a fuzzy matching algorithm to filter out inconsistent matches. The other common variables are loan type (FHA, VA, or conventional), loan purpose (purchase, refinance etc.), occupancy, lien,[29] and type of purchaser (FNMA, GNMA, PLS,[30] portfolio, etc.). However, for the same common variable, HMDA and CoreLogic may be coded differently. Moreover, both data sources have missing values. Missing values act as a wild card 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 non-match, and a fuzzy match. A perfect match is assigned a weight of 1; a perfect non-match is assigned a weight of 0; and a fuzzy match, with equal likelihood of a perfect match and non-match, is assigned a weight of 0.5.

Suppose 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,…K, then the probability that HMDA loan *i* and the CoreLogic loan *j* are a true match is given by

$$Q_{ij} = P_{ij} \times \prod_{k=1}^{K} W_{ijk} = \left({Z_{ij}^2}/{X_i Y_j} \right) \times \left(\prod_{k=1}^{K} W_{ijk} \right) \quad \text{(A.4)}$$

For HMDA loan *i*, any supplemental information obtained from the CoreLogic loan *j* is weighted by Q_{ij}.

Descriptive Statistics of the Matched Dataset

Table A.4 shows the matching rate between HMDA and CoreLogic loans. In the final dataset of the matched loan pairs after passing various filters, there are a total of 115 million HMDA loans originated between 1998 and 2012, each matched with at least one CoreLogic loan and with matching weights greater than zero, or about 65 percent of all HMDA loans find at least a match with CoreLogic loans. Similarly, there are total of 122 million CoreLogic loans, each matched with at least one HMDA loan and with matching weights greater than zero, or about 82 percent of all CoreLogic loans find at least one match with HMDA loans.

TABLE A.4

Match Rates between HMDA and CoreLogic (CL) Loans

Origination year	# of HMDA loans	# of matched HMDA loans	HMDA match rate (%)	# of CL loans	# of matched CL loans	CL match rate (%)
1998	12,274,581	8,013,060	65	10,700,894	8,699,511	81
1999	10,227,513	6,601,824	65	8,698,013	6,984,571	80
2000	8,138,120	5,155,089	63	6,190,200	4,948,062	80
2001	13,671,112	8,794,802	64	11,250,945	9,186,323	82
2002	16,112,008	10,561,827	66	14,020,880	11,623,991	83
2003	21,420,330	20,338,350	95	20,275,898	19,195,788	95
2004	15,028,550	9,411,268	63	14,728,992	12,242,631	83
2005	15,621,943	9,585,980	61	16,174,302	13,112,808	81
2006	13,970,183	8,487,539	61	13,548,855	10,732,886	79
2007	10,441,545	5,982,624	57	8,212,581	6,353,149	77
2008	7,177,262	3,366,157	47	4,664,236	3,277,240	70
2009	8,950,936	4,576,022	51	5,887,877	4,424,057	75
2010	7,863,337	3,935,358	50	5,067,038	3,738,019	74
2011	7,095,262	3,372,866	48	4,232,761	3,018,644	71
2012	9,783,966	6,843,212	70	4,856,565	4,210,597	87
Total	177,776,648	115,025,978	65	148,510,037	121,748,277	82

To assess how well the two datasets matched each other, Table A.5 shows the distribution of some common variables reported in the original HMDA data, and the matched HMDA and CoreLogic data. The results are very close, showing that the original HMDA loans are well represented by the matched dataset on each single common variable. To make the matched loans representative to the original HMDA loans on any combination of important variables, each matched loan is weighted to reflect the same distribution as the original HMDA loans on the combination of the following variables: year of origination, Bureau of Economic Analysis regions, loan purpose (home purchase, home improvement and refinance), occupancy (owner-occupied or not), race, ethnicity, and borrower income, and channel (GSE, FVR, and PP). Therefore, with this weighting, the matched loans should perfectly represent the whole population of the original HMDA loans, in terms of the distribution on the above variables.

TABLE A.5

Comparing the Distribution of Common Variables between the Original HMDA and the HMDA Matched with CoreLogic

		Original HMDA															HMDA Matched with CoreLogic														
Variables	**Categories**	**98**	**99**	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**98**	**99**	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
Loan amount	≤50K	25	28	31	19	14	11	15	17	19	16	13	7	7	8	7	26	29	31	18	14	10	12	14	17	13	6	2	2	3	2
	(50,100]K	32	31	28	26	25	25	22	19	19	18	17	15	16	18	16	34	33	30	29	27	25	23	20	20	19	17	15	15	18	15
	(100,200]K	32	30	30	37	39	40	35	32	30	32	35	39	38	37	38	31	29	28	36	38	41	36	33	31	33	39	41	41	40	40
	(200,300]K	7	7	8	12	14	15	15	16	15	16	19	21	20	19	21	6	7	7	11	14	15	16	16	15	17	20	22	22	20	22
	(300,400]K	2	2	2	3	4	5	7	8	8	8	9	10	10	9	11	2	2	2	2	4	5	7	8	8	8	9	11	10	10	11
	>400K	2	2	2	3	4	4	6	8	9	9	8	8	9	9	9	1	2	2	3	4	4	6	8	9	10	8	9	10	10	9
BEA region	Far west	19	18	17	20	21	22	23	22	20	18	15	17	19	19	20	17	16	15	17	19	22	25	24	22	21	19	21	22	23	22
	Great Lakes	19	18	18	19	18	17	15	14	14	14	16	16	16	16	16	23	22	22	23	22	17	17	16	16	16	17	18	17	17	16
	Mideast	13	14	13	12	13	14	14	14	14	14	14	14	13	14	12	15	16	15	14	15	14	16	16	17	17	16	16	16	16	12
	New England	5	5	5	5	6	6	5	5	5	5	5	6	6	6	6	6	6	5	7	7	6	6	6	6	6	6	7	7	7	5
	Plains	7	7	7	7	7	7	6	6	6	6	8	8	8	8	8	6	6	7	7	6	6	6	6	6	6	7	8	8	8	7
	Rocky Mountain	5	5	4	5	4	4	4	4	4	5	5	5	5	5	5	4	4	4	4	3	4	4	4	4	4	5	5	5	5	5
	Southeast	23	24	25	23	22	22	24	25	26	27	26	23	22	23	23	21	23	24	21	20	22	18	20	21	21	21	18	17	17	22
	Southwest	9	9	10	9	9	9	10	10	11	11	11	10	10	10	10	7	8	8	7	7	9	8	8	9	9	9	7	7	8	10
Borrower income	Low	7	9	9	7	7	7	7	6	6	5	7	8	7	8	8	8	10	10	8	7	7	7	6	6	5	7	7	7	8	7
	Medium low	15	17	17	15	15	15	16	15	14	14	15	16	15	15	14	16	18	19	16	16	15	17	16	14	14	16	16	15	15	13
	Medium	28	28	28	28	27	27	28	27	26	25	26	26	25	24	24	28	29	29	28	28	27	29	28	27	26	28	27	26	25	24
	Medium high	27	25	25	27	27	27	26	27	27	26	26	26	26	25	26	27	25	24	26	26	27	26	27	27	27	26	26	26	26	27
	High	23	21	21	23	24	24	22	26	29	29	26	25	27	28	28	21	19	19	21	22	23	21	24	27	27	23	24	26	27	28
Channel	FVR	19	22	22	18	13	11	9	6	6	10	30	33	30	30	27	17	19	21	17	12	10	8	5	5	8	29	30	28	23	18
	GSE	45	38	35	47	51	51	34	25	23	35	40	42	41	43	51	47	40	36	49	52	51	34	25	23	36	45	47	41	50	61
	PP	36	40	43	35	36	39	57	70	71	55	29	25	29	28	23	36	41	43	34	36	38	58	70	72	56	26	23	31	27	21
Gender	Male	79	76	74	76	76	74	70	69	68	69	71	73	73	73	74	78	75	73	75	75	74	70	68	67	68	70	72	73	73	74
	Female	21	24	26	24	24	26	30	31	32	31	29	27	27	27	26	22	25	27	25	25	26	30	32	33	32	30	28	27	27	26
Occupancy	Owner	93	92	92	92	92	92	89	88	88	88	88	92	91	89	89	94	93	93	93	92	93	91	90	89	90	92	96	95	93	92
	Non-owner	6	7	7	7	7	7	10	12	12	11	11	7	8	10	11	5	6	6	6	7	7	9	10	10	9	8	4	4	7	8
	Other	1.0	1.0	1.0	1.0	1.0	0.9	0.7	0.6	0.5	0.5	0.6	0.4	0.4	0.4	0.5	0.8	1.0	0.9	0.9	1.0	0.6	0.4	0.3	0.2	0.2	0.3	0.2	0.1	0.1	0.1
Purpose	Purchase	37	48	59	36	32	26	43	47	48	45	44	31	32	34	28	36	46	57	34	30	26	41	46	47	43	43	27	28	28	20
	Home improvement	8	9	11	6	4	3	6	7	8	9	8	4	4	5	4	7	9	10	5	4	2	5	6	7	7	4	2	2	2	1
	Refinance	55	43	30	58	64	71	51	46	44	46	48	64	63	61	68	57	46	32	60	66	72	54	49	46	49	53	71	71	71	79
Race/ Ethnicity	African American	6	7	8	6	6	6	8	9	10	9	7	5	4	4	4	7	8	9	6	6	6	9	10	11	9	7	5	4	4	4
	Asian	3	3	4	4	5	5	6	5	5	5	5	5	6	6	7	3	3	3	4	5	5	6	5	5	5	5	6	7	8	8
	Hispanic	6	8	9	8	8	8	13	14	15	12	9	7	7	7	7	5	6	7	7	7	8	13	14	15	12	8	6	6	6	6
	Non-Hispanic white	84	82	80	82	82	81	73	72	70	75	80	83	83	82	82	85	82	80	83	82	81	73	71	69	74	79	83	83	82	82

Calculate Variances and Confidence Intervals of the Estimates Using the Matched Data

Variance comes when we use samples to estimate population measures. The matched HMDA and CoreLogic data are a sample of the population of all residential mortgages. However, the size of the sample is very large, covering about 65 percent of all residential mortgages originated between 1998 and 2012. Moreover, each matched loan is weighted to reflect the same distribution as the original HMDA loans on the combination of the major important variables reported under HMDA. Therefore, the whole population of the original HMDA loans is well represented by the matched data.

However, in addition to sample error, there is error associated with uncertainty of whether the match is a true match or not, which we call matching error. Since the core of our matching methodology is to assign a probability of a true match to each matched loan pairs, we are able to calculate this matching error.

Let X_{ij} be the value of a scalar quantity of a loan matched between HMDA loan *i* and CoreLogic loan *j* on a specific loan or borrower characteristics of interest (e.g., FICO score, LTV). Then the following equation estimates the mean of *X* for all US residential mortgages:

$$\overline{\mathrm{X}} = \frac{1}{I}\sum_{i=1}^{I}\left\{\frac{1}{\sum_{j=1}^{J}(Q_{ij})}\sum_{j=1}^{J}\left[Q_{ij} * X_{ij}\right]\right\} \tag{A.5}$$

The variance of individual values of X among all the matched loans is given by

$$\hat{\sigma}^2 = \frac{1}{I}\sum_{i=1}^{I}\left\{\frac{1}{\sum_{j=1}^{J}(Q_{ij})}\sum_{j=1}^{J}\left[Q_{ij} * \left(X_{ij} - \overline{\mathrm{X}}\right)^2\right]\right\} = \hat{\sigma}_{sample}^2 + \hat{\sigma}_{matching}^2 \tag{A.6}$$

where

$$\hat{\sigma}_{sample}^2 = \frac{1}{I}\sum_{i=1}^{I}(X_{i.} - \bar{X})^2 \tag{A.7}$$

$$\hat{\sigma}_{matching}^2 = \frac{1}{I}\sum_{i=1}^{I}\left\{\frac{1}{\sum_{j=1}^{J}(Q_{ij})}\sum_{j=1}^{J}\left[Q_{ij} * \left(X_{ij} - X_{i.}\right)^2\right]\right\} \tag{A.8}$$

Where,

$$X_{i.} = \frac{\sum_{j=1}^{J}[Q_{ij} * X_{ij}]}{\sum_{j=1}^{J}(Q_{ij})} \tag{A.9}$$

Equations A.7 and A.8 decompose equation A.6 into two sources of variations: variation among the independent sample of HMDA loans, and variation among CoreLogic loans matched to each HMDA loan. The former captures the sample variance, and the latter captures the matching variance.

The variance of $\overline{\mathrm{X}}$ is given by

$$\widehat{Var}(\overline{\mathrm{X}}) = \frac{\hat{\sigma}^2}{I^2} \sum_{i=1}^{I} \left\{ \frac{\sum_{j=1}^{J} Q_{ij}^2}{\left[\sum_{j=1}^{J}(Q_{ij})\right]^2} \right\} \tag{A.10}$$

So the standard error of $\overline{\mathrm{X}}$ is given by

$$se(\overline{\mathrm{X}}) = \frac{\hat{\sigma}}{I} \sqrt{\sum_{i=1}^{I} \left\{ \frac{\sum_{j=1}^{J} Q_{ij}^2}{\left[\sum_{j=1}^{J}(Q_{ij})\right]^2} \right\}} \tag{A.11}$$

Under perfect match conditions, which result in all unique matches, then,

$$\hat{\sigma}_{matching}^2 = 0, \text{ and}$$

$$\sum_{i=1}^{I} \left\{ \frac{\sum_{j=1}^{J} Q_{ij}^2}{\left[\sum_{j=1}^{J}(Q_{ij})\right]^2} \right\} = \sqrt{I}$$

Equation A.8 becomes

$$se_{perfect}(\overline{\mathrm{X}}) = \frac{\hat{\sigma}_{sample}}{\sqrt{I}} \tag{A.12}$$

which is the standard formula for standard error calculations.

Under imperfect match conditions, which result in multiple matches to each HMDA loan, then,

$$\hat{\sigma}_{matching}^2 > 0, \text{ and}$$

$$\sum_{i=1}^{I} \left\{ \frac{\sum_{j=1}^{J} Q_{ij}^2}{\left[\sum_{j=1}^{J}(Q_{ij})\right]^2} \right\} < \sqrt{I}$$

So,

$$se(\overline{\mathrm{X}}) < \frac{\sqrt{\left(\hat{\sigma}_{sample}^2 + \hat{\sigma}_{matching}^2\right)}}{\sqrt{I}} \tag{A.13}$$

It is true that

$$\lim_{I \to \infty} se(\overline{\mathrm{X}}) = 0 \tag{A.14}$$

Equation A.14 shows that using the matched HMDA and CoreLogic data to estimate measures of the residential mortgage market at such large geographic areas as the national or state level, since the sample size is large, standard errors of the estimate will be small enough to be neglected. However, for estimates at smaller geographic levels such as counties, zip codes and census tracts, since the sample size is small, it is recommended to use the above equations to calculate standard errors of the estimates and assign confidence intervals.

Defining LCP Consumers, Applicants, and Borrowers

Defining LCP Consumers at the Demand Step

The Survey of Consumer Finances (SCF) helps us find the proportion of individuals who want credit that are LCP (PLCPD). The SCF is a triennial statistical survey of the balance sheet, income, and other demographic characteristics of families in the United States. The SCF also gathers information on the use of financial institutions. About 4,500 to 6,500 families are interviewed in the main study. For details of the survey, see Kennickell, McManus, and Woodburn (1996); Kennickell (1998); Lindamood, Hanna, and Bi (2007); and Bricker and colleagues (2012).

The SCF has asked the following two questions triennially, beginning in 1995:

1. Have you (and your {husband/wife/partner}) applied for any type of credit or loan in the last five years?
2. Was there any time in the past five years that you or your (husband/wife/partner) thought of applying for credit at a particular place, but changed your mind because you thought you might be turned down?

Individuals who answered yes to either question are considered consumers who want credit. Individuals who answered no to both questions are considered consumers who do not want credit.

Next we discuss an algorithm to define LCP consumers in SCF. The SCF allows us to calculate the following measures of an individual's financial status: total gross household income from all sources, total assets, total debts, net worth, total liquid assets, total monthly payments due to revolving debts, and total monthly rents for renters. For a detailed definition of these measures, see Bricker and colleagues (2012). For renters, we arbitrarily assign an expected house price of four times their household income; their total liquid assets divided by the expected house price give us the percentage of down payment a renter can afford; further, a renter's total monthly debt payments plus total monthly rent, when divided by monthly household income, gives us an estimate of debt-to-income (DTI). For owners, DTI equals total monthly debt payments (including mortgage debt) divided by total monthly household income; the debt-to-asset (DTA) ratio equals total debt divided by total assets. If the consumer's net worth is $500,000 or more, or the consumer's household income is greater than $150,000, the consumer is treated as high-credit-profile or HCP. If the consumer's income is $30,000 or less, the consumer is treated as low-credit-profile or LCP. Otherwise, LCP or HCP is defined in table A.6. Any combination of the consumer's financial status in table A.6 with a value of "0" is defined as an LCP consumer. With this definition, table A.7 shows LCP consumers who want credit as a percentage of all consumers who want credit by race and ethnicity, which is used as a proxy for $P_{0,L,t}$:

TABLE A.6

Defining Low Credit Profile Consumers Using Survey of Consumer Finance Data

	Renters			Owners		
Annual household income ($)	**% down payment**	**DTI**	**LCP or HCP**	**DTA**	**DTI**	**LCP or HCP**
30,000–60,000	<20%	Any	0	>40	Any	0
	≥20%	>30%	0	≤40	>40%	0
		≤30%	1		≤40%	1
60,000–90,000	<15%	Any	0	>50	Any	0
	≥15%	>40%	0	≤50	>50%	0
		≤40%	1		≤50%	1
90,000–120,000	<10%	Any	0	>60	Any	0
	≥10%	>50%	0	≤60	>60%	0
		≤50%	1		≤60%	1
120,000–150,000	<10%	Any	0	>70	Any	0
	≥10%	>60%	0	≤70	>70%	0
		≤60%	1		≤70%	1

Source: Survey of Consumer Finance.
Notes: For renters, we assign an expected house price of four times household income; their total liquid assets divided by the expected house price give us the percentage of down payments the renters can afford; renter's total monthly debt payments plus total monthly rents, divided by monthly household income, gives us an estimate of DTI. For owners, DTI equals total monthly debt payments (including mortgage debt) divided by total monthly household income; the debt-to-asset (DTA) ratio equals total debt divided by total assets. If the consumer's net worth is $500,000 or more, or the consumer's household income is greater than $150,000, the consumer is high credit profile; if the consumer's income is $30,000 or less, the consumer is low credit profile; otherwise, LCP or HCP is defined in this table. Any combination of the consumer's financial status with a value of "0" is low credit profile.

TABLE A.7

Share of Low Credit Profile Consumers out of All Consumers with Credit Demand by Race and Ethnicity

Origination year	Black	Hispanic	Non-Hispanic white	Asian and others
1995	80% (76%, 84%)	79% (73%, 84%)	57% (55%, 59%)	70% (62%, 78%)
1998	78% (74%, 83%)	77% (71%, 83%)	54% (52%, 55%)	60% (49%, 70%)
2001	77% (74%, 80%)	78% (73%, 84%)	48% (47%, 50%)	55% (45%, 65%)
2004	79% (76%, 82%)	78% (73%, 83%)	50% (47%, 52%)	54% (45%, 63%)
2007	73% (69%, 78%)	72% (67%, 76%)	50% (48%, 52%)	47% (40%, 54%)
2010	82% (79%, 85%)	82% (79%, 85%)	56% (55%, 58%)	59% (53%, 66%)

Source: Survey of Consumer Finance.
Notes: LCP consumers are defined in table 3. Numbers in the parenthesis are 95% confidence intervals of the estimate. Details on the calculation of the confidence intervals are described in the paper.

Table A.7 shows that within each racial and ethnic group the proportion of low-income consumers who want credit is stable over different periods from 1990 to 2010. For example, between 1990 and 2010, for the measure of the percentage of LCP consumers with credit demand, the difference between a survey year with a minimum measure and a survey year with a maximum measure is about 9, 10, 9, and 23 percentage points, respectively, for black consumers, Hispanic consumers, non-Hispanic white consumers, and consumers from other races including Asians.[31,32]

Defining LCP Applicants and Borrowers at the Application and Origination Steps

A consumer's credit profile is usually measured along a number of different dimensions such as loan-to-value ratio (LTV), credit score (FICO), DTI, but the final goal is to combine them to predict the consumer's credit risk. Different combinations of these risk factors with equal levels of default risk would be treated as the same level of credit risk by the lenders. Empirically, we can look at historical mortgage default rates and assign a cut-off point on default rate. We treat as low credit profile any combination of the common risk factors with a historical default rate greater than the cut-off point. Specifically, we used Fannie Mae and Freddie Mac Single Family Loan Performance Data (SFLPD) to measure the compensation effects on default among LTV, DTI, and FICO, and to empirically define LCP consumers. SFLPD includes a subset of Fannie and Freddie 30-year, fully amortizing, full-documentation, single-family, conventional fixed-rate mortgages, which makes the dataset reflective of current GSE underwriting guidelines.[33]

Table A.8 shows the compensation effects on default among the three major risk factors: combined LTV, FICO score, and backend DTI. In the table, "1" means that for loans with a combined LTV less than the specified value, a FICO score greater than the specified value, and a backend DTI less than the specified value, the cumulative default rate[34] of the loan should be less than 7 percent. Loans/borrowers that meet these criteria are called HCP loans/borrowers. In the table, "0" means the opposite of "1,"—that is, for loans with a combined LTV greater than the specified value, a FICO score less than the specified value, and a backend DTI greater than the specified value, the cumulative default rate of the loan should be greater than 7 percent. Loans/borrowers that meet these criteria are called LCP loans/borrowers. Total LCP loans compose about 10 percent of all SFLPD loans. In other words, borrowers with credit risk equal to or greater than the values set by table A.8 constitute about 10 percent of all GSE loans originated between 2000 and 2003 in the SFLPD database.

TABLE A.8

Define Low Credit Profile Borrowers by Default Rate

	Backend DTI<40				Backend DTI ≥ 40			
Combined	FICO				FICO			
LTV	≤ 620	620–640	640–660	> 660	≤ 620	620–640	640–660	>660
0–60	1	1	1	1	1	1	1	1
60–70	1	1	1	1	0	1	1	1
70–79	0	1	1	1	0	1	1	1
79–8)	0	1	1	1	0	1	1	1
81–85	0	0	1	1	0	0	0	1
85–90	0	0	1	1	0	0	0	1
90–95	0	0	0	1	0	0	0	1
>95	0	0	0	1	0	0	0	1

Source: Fannie Mae and Freddie Mac single-family loan performance data.
Note: Any combination of the combined LTV, backend DTI, and FICO scores with a value of "0" are defined as low-credit-profile borrowers.

We only used loans originated between 2000 and 2003 for the calibration of high and low credit profiles. Performance of these vintages is more reflective of normal market conditions than loans originated between 2004 and 2008, which suffered very substantial home price depreciation early in their life. For loans originated after 2009, we are concerned that it is too soon to observe lifetime cumulative default rates.

In addition to the 7 percent default rate, we repeated our analysis using 6 and 8 percent default rates as cutting points to define LCPs. Table A.9 shows the results.

TABLE A.9

RDR by Different LCP Definitions

		DR=6%				DR=7%				DR=8%			
Channel	Year	A	H	S	W	A	H	S	W	A	H	S	W
F	1998	16	14	14	11	18	17	18	14	19	18	19	15
F	1999	18	15	16	13	20	18	20	16	21	19	21	17
F	2000	20	15	14	15	22	17	16	17	23	18	17	18
F	2001	14	12	12	9	15	14	14	11	16	15	15	12
F	2002	16	14	12	11	17	16	14	13	18	17	16	14
F	2003	16	15	14	11	17	18	16	13	19	20	19	15
F	2004	22	21	17	16	23	23	21	18	26	27	24	21
F	2005	24	24	20	18	26	26	23	20	28	29	26	22
F	2006	22	21	18	18	24	23	21	20	26	25	23	22
F	2007	30	28	28	24	33	31	32	27	34	33	34	28
F	2008	40	39	42	34	45	45	49	40	47	48	51	42
F	2009	47	46	53	43	59	62	69	59	62	65	72	62
F	2010	51	49	57	50	65	67	74	65	66	69	76	67
F	2011	53	48	56	50	65	66	72	65	69	68	74	67
F	2012	52	47	55	46	66	65	72	62	68	67	75	65
G	1998	87	92	91	89	90	95	94	91	94	97	97	95
G	1999	81	86	83	82	84	91	89	86	90	94	93	91
G	2000	68	68	60	60	72	74	68	65	79	82	78	75
G	2001	80	86	85	79	84	90	90	83	89	94	94	90
G	2002	84	89	90	85	88	93	94	89	91	95	96	92
G	2003	91	93	96	93	94	96	98	95	95	97	98	97
G	2004	81	85	89	85	86	90	93	89	89	93	95	92
G	2005	79	83	87	83	84	89	92	87	88	93	95	91
G	2006	73	77	80	75	79	84	85	80	84	88	89	85
G	2007	77	82	84	76	82	87	89	81	86	91	92	86
G	2008	94	96	97	94	96	97	98	96	98	98	99	98
G	2009	99	99	99	99	99	100	100	99	100	100	100	100
G	2010	98	98	99	98	99	99	99	98	99	99	99	99
G	2011	97	98	99	98	98	99	99	98	99	99	99	99
G	2012	95	96	98	96	97	98	99	97	97	98	99	98
P	1998	59	62	59	47	62	67	68	54	75	79	81	73
P	1999	56	53	52	40	59	59	61	46	70	71	74	64
P	2000	52	48	42	37	55	53	47	41	65	64	60	54
P	2001	47	40	33	10	51	49	43	18	63	64	63	41
P	2002	34	32	35	5	38	43	47	15	54	62	67	43
P	2003	48	47	54	38	55	61	68	51	67	74	79	68
P	2004	47	45	45	34	51	55	57	43	62	68	69	56
P	2005	49	51	53	37	53	61	64	46	62	72	73	57
P	2006	54	59	61	46	59	67	69	54	66	77	78	63
P	2007	68	75	78	65	73	81	84	73	79	86	89	79
P	2008	90	92	93	88	92	94	95	91	95	96	97	94
P	2009	97	98	98	96	98	98	99	97	98	99	99	98
P	2010	97	99	98	98	98	99	99	99	99	99	99	99
P	2011	96	97	98	96	98	99	99	98	98	99	99	98
P	2012	94	94	95	93	96	97	97	95	96	97	98	96

Notes: DR = default rate used to define the combinations of FICO, LTV and DTI for LCP. A = black; H = Hispanic; S = Asian; W = non-Hispanic white. F = FVR; G = GSE; P = PP.

Confidence Intervals of DAPR, DOPR, and PLCPD

In this section, we discuss ways to assign confidence intervals to the estimates of the three progression rates and other population measures of all US residential mortgages. In general, variances are inevitable when we use samples to estimate population measures such as mean, proportion or total of the population on a specific characteristic. HMDA is almost the population of all residential mortgages, so estimates calculated with the HMDA alone such as the ODRs should be treated as the true population value without variance.

For the estimates calculated with the matched HMDA and CoreLogic data such as real denial rates for LCP applicants, AOPRs for LCP applicants, and the percent of LCP applicants out of all applicants (PLCPA), see the earlier "Calculate Variances and Confidence Intervals" section for the details of assigning confidence intervals to these estimates. In general, for reasons discussed under that section, variances on these estimates are small enough to be neglected, if the estimation is at a large geographic level such as national or state level, which will rely on a large sample size.

For the other two progression rates, the DAPR, and the DOPR, in addition to the matched HMDA and CoreLogic data, we have to use SCF to get the credit demand information for LCP consumers. According to SCF, estimates of the US population measures using the survey have two types of errors:[35] sampling error and imputation error. Sampling error comes from the fact that the survey is just a sample of the US population with sample size of about 4,500 to 6,500 families. Imputation error results from assigning imputed values to missing values.[36] To allow researchers to calculate sample errors, SCF provides 999 bootstrap replicates. In table 4, numbers in parentheses show the 95 percent confidence interval for each estimate calculated using the bootstrap replicates. To test how sensitive the calculated $P_{0,L,t}$ affects our estimates on DAPR and the overall DOPR, using the bootstrap replicates, for each imputed LCP ID, we calculated 999 PLCPDs. Then plugging them into equations (9) and (11), and using the formulas provided by Kennickell, McManus, and Woodburn (1996) we are able to calculate the standard errors of the measures of the two progression rates. With the calculated standard errors, we are able to calculate the 95 percent confidence intervals of these measures.

The formulas for standard errors are:

$$SE = \left\{(6/5)SX^2_{imputation} + SX^2_{sample}\right\}^{1/2} \tag{A.15}$$

Where the imputation variance SX^2_{imp} is given by

$$SX^2_{imputation} = \frac{1}{4}\sum_{i=1}^{5}(\hat{P}_{0,L,t,i,.} - \hat{P}_{0,L,t,.,.})^2 \tag{A.16}$$

And the sampling variance SX^2_{samp} is given by

$$SX^2_{sample} = \frac{1}{(5\times 998)} \sum_{i=1}^{5} \left[\sum_{j=1}^{999} (P_{0,L,t,i,j} - \hat{P}_{0,L,t,i,.})^2 \right] \quad \text{(A.17)}$$

Where

$\hat{P}_{0,L,t,i,.} = \frac{1}{999} \sum_{j=1}^{999} \left(P_{0,L,t,i,j} \right)$ and,

$$\hat{P}_{0,L,t,.,.} = \frac{1}{5} \sum_{i=1}^{5} \left(\hat{P}_{0,L,t,i,.} \right)$$

Notes

1. See, for example, page 14 http://www.urban.org/publications/413234.html.
2. See detailed information at http://www.federalreserve.gov/boarddocs/snloansurvey/about.htm.
3. For each mortgage loan application, four possible outcomes are reported in HMDA data: application denied, application approved but not accepted, loan originated, and application withdrawn by applicant or file closed for incompleteness.
4. The survey asks the following question: "Over the past three months, how have your bank's credit standards for approving applications from individuals for mortgage loans to purchase homes changed?" Banks may respond that their lending standards "remained basically unchanged," "tightened considerably," "tightened somewhat," "eased considerably," or "eased somewhat."
5. A positive percentage means tightening and zero means unchanged lending standards in the past three months. A negative percentage means loosening lending standards in the past three months.
6. The nontraditional category of residential mortgages includes adjustable-rate mortgages with multiple payment options, interest-only mortgages, and Alt-A products such as mortgages with limited income verification and mortgages secured by non-owner-occupied properties.
7. Notice the disconnection before and after Q2 2007 in the upper-left panel of figure 1.
8. HMDA is considered the "universe" of mortgage loans because federal law requires that almost all mortgage originations, except some small lenders, be reported in HMDA. See Avery, Brevoort, and Canner (2007) and McCoy (2007) for detailed discussion on HMDA's coverage of residential mortgages.
9. Munnell and colleagues (1996) are among the few researchers who have attempted to overcome this problem.
10. See http://www.mbaa.org/ResearchandForecasts/MCAI.htm for a more complete description.
11. AllRegs is a publisher of underwriting and loan product guidelines for Fannie Mae, Freddie Mac, Federal Home Loan Bank of Chicago, Wells Fargo Home Mortgage, JPMorgan Chase, and other mortgage lenders.
12. It is unclear what the pool of surveyed lenders looks like and how each lender's answers are weighted relative to other lenders' answers.
13. This perfect world still wouldn't create 100 percent origination rates because applicants still have to choose whether to accept or refuse the credit at the given price.
14. According to Parrott and Zandi (2013) and Goodman and Zhu (2013), the way the put-back rules are written and enforced creates this problematic vagueness.
15. Mortgage loan databases sold by CoreLogic and other commercial data providers contain detailed borrower and loan information including FICO, LTV, and DTI.
16. HMDA contains mortgage applicant's information including their income, loan amount, race and ethnicity, and the outcome of the application, but no information on common risk factors such as credit score, LTV, and DTI.
17. Starting in 2004, HMDA data report additional outcomes for loan applications associated with certain types of requests for preapproval of home purchase loans; see Avery, Canner, and Cook (2005) for details. In this paper, outcomes of a preapproval request and loan application are combined; denied includes both denied preapproval requests and applications denied by financial institution; approved but not accepted includes both approved preapproval requests and applications approved but not accepted. Loans purchased by a financial institution at a HMDA reporting year are excluded from analysis.
18. In this paper, we call the fourth outcome "application incomplete or withdrawn."
19. See Avery, Brevoort, and Canner (2007) for race and ethnicity definition issues. From 1990 to 2003, race and ethnicity were reported jointly in one of six possible categories: white, black, Hispanic, Asian or Pacific Islander, American Indian and Alaska Native, and "other." Since 2004, race and ethnicity have been reported separately; moreover, applicants are allowed to choose more than one racial category. In this paper, we adopt a

hierarchical approach to define race and ethnicity jointly. For HMDA data reported between 1998 and 2003, the jointly reported field on applicant's race and ethnicity is used directly for the definition. For HMDA data reported between 2004 and 2012, we adopted the same approach as used by Avery, Canner, and Cook (2005) and Avery, Breevort, and Canner (2006): black trumps Hispanic, Hispanic trumps Asian, Asian trumps other minorities, and other minorities trumps white, in any one of the five race fields and one ethnicity field. Co-applicant's race and ethnicity are ignored when defining applicant's race and ethnicity.

20. At the time of application, applicants choose either the conventional or agency channel. The agency loan channel is FHA, VA, and RD loans. Any other loan belongs to the conventional channel.

21. The incomplete application rate equals to the number of applications withdrawn by applicant or files closed for incompleteness, divided by total number of applications.

22. Within the conventional channel, after the lender originates the loan, it can be securitized with GSE or PLS, or be kept in the financial institution's portfolio. HMDA data do not specify which of these channels the loan will go to at the application stage, so the ODR of the conventional channel applies to both GSE and PP.

23. The average 30-year fixed-rate mortgage rate during 2000 was about 7–8 percent, according to Freddie Mac's Mortgage Rate Survey.

24. The origination volume in 2000 is less than half the annual volume thereafter, according to Inside Mortgage Finance.

25. In table 3, we have a more complete channel for the RDR analysis than for the ODR analysis. For the ODR analysis, we rely on HMDA alone; for the RDR analysis we are able to layer in our CoreLogic Data, which allows us to separate GSE loans from PP loans. The FVR, conventional, and "all" categories are common in both analyses.

26. We are unable to calculate a separate deter rate for each channel because we are unable to identify which channel deterred potential consumers with credit need. Instead, an annual deter rate is calculated for each race and ethnicity for all channels. Consequently, the annual deter rate for FVR, GSE, and PP are identical within each racial and ethnic group.

27. As discussed in the appendix, measures of PLCPD contain both sampling and imputation errors. We therefore conducted a sensitivity analysis and assigned 95 percent confidence intervals to our calculated DAPR and DOPR measures. In appendix figures A1 and A2 and appendix tables A2 and A3, L and U stand for the lower and the upper 95 percent confidence limits, respectively. The results indicate that the width of the interval depends largely on the sample size of each race and ethnic group in the SCF. For example, non-Hispanic whites have the narrowest confidence interval, whereas Asians have the widest confidence interval.

28. Again, this measure does not take the quality of loans into consideration. Unfortunately, the PP channel produced many "affordability products" that ultimately had high foreclosure rates (Ding et al. 2011).

29. Lien status is available only for HMDA loans originated after 2004.

30. PLS as a separate type of purchaser is available only for HMDA loans originated after 2004.

31. If it is true that within a demographic group, the proportion of low-income consumers who want credit to all consumers who want credit is constant over a relatively short period, it will make our attempt to calculate DAPR much easier. First, the ratio between two DAPRs from two different time points for LCP individuals who want credit is given by

$$\frac{P_{1,L,t_1}}{P_{1,L,t_2}} = \frac{Q_{t_1}}{(1-Q_{t_1})} \times \frac{(1-Q_{t_2})}{Q_{t_2}} \times \frac{(1-P_{0,L,t_1})}{P_{0,L,t_1}} \times \frac{P_{0,L,t_2}}{(1-P_{0,L,t_2})} \tag{A.15}$$

With the above assumption, equation 12 becomes

$$\frac{P_{1,L,t_1}}{P_{1,L,t_2}} = \frac{Q_{t_1}}{(1-Q_{t_1})} \times \frac{(1-Q_{t_2})}{Q_{t_2}} \tag{A.16}$$

which can be solved empirically using the matched HMDA and CoreLogic data alone, without using numbers presented in table 4. Similarly, with this assumption, the ratio between two DOPRs from two different time points is simplified as

$$\frac{P_{L,t_1}}{P_{L,t_2}} = \frac{1-Q_{t_2}}{1-Q_{t_1}} \times \frac{Q_{t_1}-D_{t_1}}{Q_{t_2}-D_{t_2}} \tag{A.17}$$

which can also be solved empirically using the matched HMDA and CoreLogic data alone.

32. There was a much larger, 23 percentage point difference between a survey year with a minimum measure and a survey year with a maximum measure for Asians and other races. We believe this is because this category is a mixture of all races other than blacks, Hispanics, and whites, which inherently increase the variation of the measure. For sample size purposes, SCF doesn't allow researchers to break this category further into smaller groups.
33. For more details about what loans are included in the dataset, see the documentation at https://loanperformancedata.fanniemae.com/lppub-docs/lppub_glossary.pdf and http://www.freddiemac.com/news/finance/pdf/user_guide.pdf.
34. Default is defined as 180 days or more delinquent, including various stages of foreclosure and termination due to foreclosure, by December 2013.
35. In addition to the two types of errors, SCF presents two limitations when we use it to calculate PLCDP. First, the credit demand is not limited to housing credit demand; second, the survey asks for consumers' experience over the past five years instead of annually, and the survey is taken every three years.
36. The SCF provides five imputations for each missing value.

References

Avery, R. B., G. B. Canner, and R. E. Cook. 2005. "New Information Reported under HMDA and Its Application in Fair Lending Enforcement." *Federal Reserve Bulletin* 91: 344.

Avery, R. B., K. P. Brevoort, and G. B. Canner. 2006. "Higher-Priced Home Lending and the 2005 HMDA Data." *Federal Reserve Bulletin* 92: A123–66.

———. 2007. "Opportunities and Issues in Using HMDA Data." *Journal of Real Estate Research* 29 (4): 351–80.

Bricker, J., A. B. Kennickell, B. Moore, and J. Sabelhaus. 2012. "Changes in U.S. Family Finances from 2007 to 2010: Evidence from SCF." *Federal Reserve Bulletin* 98 (2).

Bhutta, N., and D. R. Ringo. 2014. "The 2013 Home Mortgage Disclosure Act Data." *Federal Reserve Bulletin* 100 (4).

Ding, L., R. G. Quercia, W. Li, and J. Ratcliffe. 2011. "Risky Borrowers or Risky Mortgages: Disaggregating Effects Using Propensity Score Models." *Journal of Real Estate Research* 33 (2): 245–77.

Ernst, K., D. Bocian, and W. Li. 2008. *Steered Wrong: Brokers, Borrowers, and Subprime Loans*. Durham, NC: Center for Responsible Lending.

Goodman, L. S., and J. Zhu. 2013. "Reps and Warrants." Washington, DC: Urban Institute.

Kennickell, A. B. 1998. "Multiple Imputation in SCF." In *Proceedings of the Section on Business and Economic Statistics*.

Kennickell, A. B., D. A. McManus, and R.L. Woodburn. 1996. *Weighting Design for the 1992 SCFs*. Washington, DC: Board of Governors of the Federal Reserve System.

Li, W. and Goodman, L. S. 2014. "Measuring Credit Availability Using Ex-Ante Probability of Default" Washington, DC: Urban Institute.

Lindamood, S., S. D. Hanna, and L. Bi. 2007. "Using SCF: Some Methodological Considerations and Issues." *Journal of Consumer Affairs* 41 (2): 195–222.

McCoy, P. A. 2007. "The Home Mortgage Disclosure Act: A Synopsis and Recent Legislative History." *Journal of Real Estate Research* 29 (4): 381–97.

Munnell, A. H., G. M. Tootell, L. E. Browne, and J. McEneaney. 1996. "Mortgage Lending in Boston: Interpreting HMDA Data." *American Economic Review*, 25–53.

Parrott, J., and M. Zandi. 2013. *Opening the Credit Box.* New York and Washington, DC: Moody's Analytics and the Urban Institute.

About the Authors

Wei Li is a senior research associate in the Housing Finance Policy Center at the Urban Institute, where his research focuses on the social and political aspects of the housing finance market, and their implications for urban policy. He is also a quantitative research methodologist with a deep understanding of cost-benefit analysis, program evaluation, and causal inference in social and political science. Before joining Urban, Li was a principal researcher with the Center for Responsible Lending, where he wrote numerous publications on the housing finance market and created and managed the nonprofit organization's comprehensive residential mortgage database. Li's work has been published widely in various academic journals and covered in the *Wall Street Journal*, *Washington Post*, and *New York Times* as well as other print and broadcast media. Li received his PhD in environmental science, policy, and management and an MA in statistics from the University of California at Berkeley.

Laurie Goodman is the center director of the Housing Finance Policy Center at the Urban Institute. This Center is dedicated to providing policy makers 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 2008-2013, Laurie was a Senior Managing Director at Amherst Securities Group, L.P., a boutique broker dealer specializing in securitized products. Her strategy effort became known for its analysis of housing policy issues. From 1993-2008, Laurie was head of Global Fixed Income Research and Manager of U.S. Securitized Products Research at UBS and predecessor firms. The UBS Securitized Products Research group was ranked #1 by Institutional Investor for 11 straight years. Prior to that, Laurie held positions as a senior fixed income analyst, a mortgage portfolio manager, and a Senior Economist at the Federal Reserve Bank of New York.

Laurie has a BA in Mathematics from the University of Pennsylvania, and a MA and PhD in Economics from Stanford University. She has published more than 200 articles in professional and academic journals and co-authored and co-edited five books. Laurie was inducted into the Fixed Income Analysts Hall of Fame in 2009. She serves on the Board of Directors of MFA Financial, and is a member of the Bipartisan Policy Center's Housing Commission, the Federal Reserve Bank of New York's Financial Advisory Roundtable, as well as the New York State Mortgage Relief Incentive Fund Advisory Committee.

Ellen Seidman is a Senior Fellow at the Urban Institute, focusing on housing finance and community development. During 2012 and 2013, she was a visiting scholar with the Community Development Department of the Federal Reserve Bank of San Francisco, and has been a Research Fellow of the Filene Research Institute since 2011. In 2012, she was appointed to the Consumer Advisory Board of the Consumer Financial Protection Bureau. She is the 2013-2014 NYU Stern-Citi Leadership & Ethics Distinguished Fellow. From 1997 through 2001, she was the Director of the Office of Thrift Supervision, the federal regulator of the savings and loan industry, and concurrently served on the Board of Directors of the Federal Deposit Insurance Corporation and as Chair of the Board for the Neighborhood Reinvestment Corporation.

From 2002 through 2010, Ms. Seidman held various positions at ShoreBank Corporation and its affiliates. From 2007 to 2010, she was a scholar in the Asset Building Program at the New America Foundation, a Washington, D.C. think tank. Ms. Seidman has also been Senior Counsel to the Democratic Staff, House Financial Services Committee (2001-02), Special Assistant to the President for Economic Policy (1993-1997), and has held senior positions at Fannie Mae, the Department of the Treasury and the Department of Transportation. She was a founder of and serves on the board of the Center for Financial Services Innovation and is a member of several other boards, including Coastal Enterprises Inc. (CEI), City First Bank of DC and the CDFI Assessment and Ratings System (CARS). Ms. Seidman received her M.B.A. in Finance and Investments from George Washington University, her J.D. from Georgetown University Law Center, and her A.B. from Radcliffe College.

Jim Parrott is a Senior Fellow at the Urban Institute and owner of Falling Creek Advisors, which provides financial institutions with strategic advice on housing finance issues. Jim spent several years in the White House as a Senior Advisor at the National Economic Council (NEC), where he led the team of advisors charged with counseling the cabinet and President on housing issues. He was on point for developing the Administration's major housing policy positions, articulating and defending those positions with Congress, the press and public, and counseling White House leadership on related communications and legislative strategy. Prior to his time with the NEC, Jim was a senior advisor to Secretary Donovan at the Department of Housing and Urban Development. He has a J.D. from Columbia University School of Law, an M.A. from the University of Washington and a B.A. from the University of North Carolina.

Jun Zhu is a senior financial methodologist in the Housing Finance Policy Center. 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. While at Freddie Mac, she also served as 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, 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.

URBAN
INSTITUTE · ELEVATE · THE · DEBATE

2100 M Street NW
Washington, DC 20037

www.urban.org