Regional Science and Urban Economics 80 (2020) 103390

Contents lists available at ScienceDirect

Regional Science and Urban Economics

journal homepage: www.elsevier.com/locate/regec

Credit risk of low income mortgages

Hamilton Fout [a,*], Grace Li [b], Mark Palim [c], Ying Pan [c]

[a] Fannie Mae and Kansas State University, USA
[b] SunTrust Bank, USA
[c] Fannie Mae, USA

ARTICLE INFO

JEL codes:
G21
R31
R38

Keywords:
Low income lending
Credit risk
Underwriting standards

ABSTRACT

Using Fannie Mae data on purchase mortgage acquisitions, we examine the relative credit performance of low and moderate income homebuyers. We first document the higher observed default rates of low and moderate income borrowers relative to higher income borrowers for three different historical periods. Second, for the loans originated between 2002 and 2007 applying the tighter underwriting standards of the post-crisis period dramatically reduces default risks across income groups, indicating the importance of underwriting standards for sustainable lending and homeownership. Finally, for all but very low income borrowers, credit risk is well accounted for by standard underwriting risk factors.

1. Introduction

Over the last 15 years, the share of purchase loans that have gone to low and moderate income (LMI) households, defined for the purposes of this study as households with income less than or equal to area median income, has declined as average credit scores and down payment percentages have increased given tightened mortgage underwriting standards.[1] Lower income borrowers potentially face a number of barriers to homeownership including a lack of wealth, saving for a down payment, higher debt-to-income ratios and lower credit scores.[2] Given the relatively riskier credit profile of LMI borrowers as well as higher resource constraints, LMI borrowers tend to default more frequently than higher income borrowers. Increasing opportunities for homeownership for LMI borrowers that limits the credit risk exposure associated with these borrowers continues to be an important goal of U.S. housing policy (see for instance the Government Sponsored Enterprise (GSE) housing goals and Community Reinvestment Act (CRA)).

In this paper, we explore the role of underwriting standards in sustainable homeownership for LMI borrowers and the extent to which the credit risk associated with extending mortgage credit to LMI borrowers is predictable at the time of origination in different credit and economic environments. We use a rich data set which contains borrower and loan characteristics as well as loan performance metrics for the purchase loans acquired by Fannie Mae from 2002 to 2013 to examine the performance of LMI loans in three historical time periods: early boom, late boom and after the 2007 housing crisis. We first document, consistent with previous literature (for instance Firestone et al., 2007), that LMI borrowers are more likely to default than higher income borrowers. We also show that underwriting reforms introduced after the crisis were effective in reducing credit risks across all income groups. In particular, we find that the default rates for LMI loans in the two historical periods (early boom and late boom) would have dramatically declined had today's eligibility rules for delivery to Fannie Mae been implemented in these periods, indicating the importance of underwriting standards for sustainable homeownership.

We then turn to understanding the remaining credit risk for the currently eligible population of LMI borrowers in the recent and historical periods. Here, we find that once we control for borrower and credit characteristics, standard underwriting factors along with region and vintage controls are sufficient for measuring the risk associated with most LMI categories. Very low income (VLI, income at or below 50 percent of

* Corresponding author.
E-mail addresses: hamilton_fout@fanniemae.com (H. Fout), ying_pan@fanniemae.com (Y. Pan).

[1] Table 1 shows that the share of LMI purchase mortgages sold to Fannie Mae fell from 45 to 36 percent from the 2002–2004 period to the 2011–2013 time period, as average LTVs fell from 83.5 to 81.0 percent and average credit scores rose from 713 to 757.

[2] In recent years, the median wealth of lower income households has declined, while upper income household wealth has increased (Kochhar and Cilluffo, 2017). Furthermore, over the longer term real wage growth for lower income workers has stagnated relative to higher earners (Shambaugh et al., 2017), while real home price growth has increased, particularly for the lower priced segment of homes where inventory has been more constrained (Source: CoreLogic). These developments further constrain the ability of LMI buyers to find affordable housing and to come up with sufficient down payments.

https://doi.org/10.1016/j.regsciurbeco.2018.07.013
Received 1 November 2017; Received in revised form 25 July 2018; Accepted 26 July 2018
Available online 31 July 2018

area median income) loans, however, present additional risk not fully captured by these factors. In particular, VLI loans default ~25 percent more than expected based strictly on their credit profile. We find this unexplained risk, however, is reasonably stable across the three periods. In robustness tests, we also document the same general conclusion of remaining marginal risks for the VLI segment when our model is estimated only on the first-time homebuyer population. Further robustness tests show that part but not all of the additional risks associated with VLI borrowers is a function of regional factors, potentially including the strength of the labor market or other features of the housing/mortgage markets at the zip code level.

The rest of the paper is organized as follows: Section 2 provides an overview of the relevant economic research in this area, Section 3 explains our data, Section 4 presents empirical results on the importance of eligibility overlays introduced post-crisis. Section 5 presents empirical results on the marginal credit risk of LMI lending with a focus on the ability of standard underwriting factors to explain LMI risks. Section 6 presents robustness tests and Section 7 concludes.

2. Background

There are two strands of economic research that are most relevant to the current paper. The first studies credit risk models and their usage as a tool for gauging the risks of mortgages at the time of underwriting. The second investigates the default and prepayment performance of LMI mortgages.

2.1. Credit risk models

There is a rich literature regarding credit risk management models and how they are used in the underwriting process of a mortgage loan (for instance Quercia and Stegman, 1992 and Avery et al., 1996). Information to assess credit risk is collected, and evaluated in the underwriting process of the loan. During this process, financial institutions assess credit risk using information on a range of risk factors that potentially affect or predict repayment behavior. These factors include the current and past payment behavior of the borrower, loan characteristics including loan type, loan purpose, the loan-to-value (LTV) and debt-to-income (DTI) ratios and the characteristics and value of the property serving as collateral for the loan (Avery et al., 1996; Haughwout et al., 2008; Mayer et al., 2009). Credit risk models are used to quantify the expected future performance of mortgage and other loans based on the information available at origination.

A typical approach in modeling mortgage credit risk involves estimating a logit model to explain some binary outcome of loan performances (e.g. 90 or more days delinquent within two years since origination based on the data observed at underwriting). The general predictive accuracy of the estimated model can be evaluated by the Gini or Kolmogorv-Smirnov (KS) coefficients, which measure the rank-ordering power of the logit model to separate those loans that went delinquent versus those that did not (Mays, 2001; Crook et al., 2007).

There are a number of attributes entering the logit model for prediction of default, such as FICO, CLTV, and DTI. Borrower income, other than its use as an input into the calculation of DTI, is not typically considered as a direct input in risk models. [3] For any nominal level of debt, lower income borrowers will have a higher DTI, which is associated with greater credit risk (Avery et al., 1996; Haughwout et al., 2008). The omission of income, either as a level or a ratio relative to area median income (AMI), is potentially due to concerns surrounding disparate impact on protected classes from including direct income controls in the underwriting process.[4] One of the cases where income is used in underwriting is with loans guaranteed by the US Department of Veterans Affairs (VA), which uses a "residual income" (net after fixed obligations) measure as one of the underwriting factors (Goodman et al., 2015).[5] However, some other underwriting risk factors may be correlated with borrower income. For example, higher income borrowers are typically able to make a larger down payment. Therefore, the CLTV on loans to LMI borrowers tend to be higher, and higher CLTV loans have higher default risk (Kelly, 2008; An et al., 2012; Lam et al., 2013).

Therefore, standard risk management models, although they do not control directly for a borrower's income relative to the area median, still account for part of the additional credit risk of LMI borrowers because of the correlation of relative income with other risk characteristics typically used in underwriting. In this paper, we investigate the extent to which directly controlling for relative income in addition to the usual credit risk factors improves our understanding of past mortgage performance under different underwriting and economic environments.

2.2. Loan performance of low and moderate income mortgages

In addition to the literature on credit risk models, there is a separate strand of the literature that focuses on prepayment and default risk of LMI loans.[6] For example, Archer et al. (1996) find that LMI borrowers are less likely to sell their property and move when facing an income or life event shock. Van Order and Zorn (2003) find that default responses to negative equity are similar across higher income and LMI neighborhoods and the small differences in defaults can be explained by omitted credit history. Deng et al. (1996, 2000), Deng and Gabriel (2006) and Firestone et al. (2007) all find slower voluntary prepayment speed among low income borrowers.

In particular, Deng et al. (1996) investigate a set of loans purchased by Freddie Mac between 1976 and 1983 and create a loan-level set of LMI indicators. They present evidence within a competing hazards framework that default risks decline as household income rises and also that LMI households are more likely to default when faced with negative equity than are higher income households. Deng and Gabriel (2006) also use a proportional hazard model to quantify the prepayment and default risks among Federal Housing Agency (FHA) mortgages originated between 1992 and 1996. They find a significant negative effect of household income on default, after controlling for other borrower and market-level measures. Firestone et al. (2007) analyze loans acquired by Freddie Mac from 1993 to 1997 and find that default probability increases as income relative to area median income declines.[7] Quercia et al. (2002) focus on the performance of a small number of CRA loans (loans made by banks to satisfy the CRA requirement that banks serve the local communities where they obtain deposits) originated in 1998 using a variety of factors including income relative to area median income. They find an insignificant effect of income relative to area median on early delinquencies for the population they investigate.

[3] For instance, Fannie Mae's selling guide lists the following risk factors evaluated as part of the automated underwriting process: credit history, delinquent accounts, installment loans, revolving credit utilization, public records, foreclosures, collection accounts, inquiries, borrower's equity and LTV, liquid reserves, loan purpose, loan term, loan amortization type, occupancy type, DTI, property type, co-borrowers and self-employment (See https://www.fanniemae.com/content/guide/sel053116.pdf, p. 316).

[4] See https://www.federalreserve.gov/boarddocs/caletters/2009/0906/09-06_attachment.pdf (p. iv) for a description of disparate impact.

[5] This residual income measure is similar to DTI, as it compares income relative to debt and other obligations.

[6] Past research on low-income lending has also focused on the equity-building potential of low-income homeownership as well as the relationship between borrowing constraints and homeownership. For example, Painter et al. (2001) examine the determinants of housing tenure choices by racial and ethnic groups. Duca and Rosenthal (1994) and Barakova et al. (2014) analyze the effect of borrowing constraints on homeownership.

[7] In an earlier version of their paper Van Order and Zorn (2002) present further evidence of increased default risks among the LMI borrowers using a similar set of indicators.

Our research extends the literature in a number of ways. First, while the literature described above has focused on the historical period before the recent housing crisis of 2007, we take advantage of a rich dataset of Fannie Mae acquisitions originated between 2002 and 2013 with loan level household income and area income data to investigate relative LMI performance under a variety of underwriting regimes and subsequent housing market environments. Second, we are interested in quantifying the additional default risk (as opposed to prepayment) associated with LMI lending using a set of indicators that allow us to separately measure the relative risk of very low income (VLI, ≤50% of area median income), low income (LI, >50% and ≤80% of area median income) and moderate income (MI, >80% and ≤100% of area median income) borrowers compared to high income (HI, >100% of area median income) borrowers after controlling for a variety of loan-level attributes as well as region and vintage fixed effects.[8] Finally, we focus on the role of underwriting in LMI lending along the following two dimensions: (1) the use of tightened underwriting standards to mitigate credit risks associated with LMI lending and (2) the ability of standard underwriting factors, such as FICO, CLTV and DTI along with region and vintage controls, to sufficiently explain the credit risks presented by LMI loans without the need to explicitly control for relative income.

3. Data

In this paper, we rely on a novel data set based on internal proprietary Fannie Mae data consisting of the entire population of conventional single-family owner-occupied 1–4 unit fixed-rate purchase loans acquired by Fannie Mae during three periods that cover three different underwriting regimes and subsequent economic environments: 2002 to 2004 (early boom), from 2005 to 2007 (late boom), and 2011 to 2013 (post-crisis period).[9] Loan characteristics at origination include CLTV, DTI, number of borrowers on the mortgage, borrower's and co-borrower's FICO scores, loan balance(s), interest rate, income documentation level associated with the loan application, zip code of the property, number of units, loan type (40-year, 30-year or 15-year fixed rate mortgages), whether the loan is negatively amortizing, interest-only or balloon and whether the loan is originated through a third party.[10] The dataset also contains loan performance information, including whether and when a loan was first-time 30, 60 or 90 days past due and whether/when the loan prepaid. Importantly, our data set also includes information on the borrower's income relative to area median, allowing us to identify LMI borrowers.

We use explanatory variables from this set of data to model default, which we define as the case when a loan goes 90 days or more delinquent within 24 months from the first payment date.[11] We use this definition of default for two primary reasons. First, this is a widely-used metric to model credit risk which we use to simply and sufficiently capture the additional risks of LMI lending.[12] Second, the primary interest of our paper is in the role of underwriting in mitigating the additional risk associated with LMI loans and the extent any such risk is stable and predictable across different underwriting regimes and economic environments. Thus, we are interested in the information available at the time of underwriting in predicting early delinquencies as opposed to later delinquency outcomes that are likely to be driven by risk factors that accrue over time and are unavailable at origination (e.g. changes in FICO scores, borrower employment or the future path of home prices).

Given the interest of public policy in sustainable homeownership, we focus on mortgages for primary owner-occupied residences in our empirical results. We exclude refinance loans and government loans from the sample. Thus our data is exclusively conventional, conforming owner-occupied fixed-rate purchase loans. Additionally, we exclude all long-term standby commitments and seasoned loans (first payment date at least one year prior to being acquired by Fannie Mae).[13] In part of the analysis, in order to measure the impact of post-crisis tightening of Fannie Mae's eligibility criteria, we evaluate the performance of a subset of early-boom period loans that would qualify under current eligibility standards. Specifically, this removes loans that have LTV higher than 97 percent, FICO score less than 620, DTI ratio higher than 0.50, loans with 40-year terms, interest-only loans, negative and balloon amortizations and loans with low or no documentation for income.[14]

The loans we study were originated from 2002 to 2013 and thus cover a range of underwriting and economic environments. In particular, housing prices appreciated over the period 2002–2007, followed by a sharp decline of home prices during the second half of 2008, with prices continuing to decline through 2011. Home prices began their recovery in 2012, with the nominal national home price index surpassing its pre-crisis peak level by late 2015 (see for instance the FHFA Purchase-Only Index). Over this period, unemployment declined to 4.5 percent in 2007 followed by a sharp increase to approximately 10 percent in 2010 and a subsequent decline to 5 percent by December 2015. Underwriting standards for conventional single family mortgages also differed significantly in each of these regimes. Two commonly cited measures of mortgage credit availability, the Federal Reserve's Senior Loan Officer Opinion Survey and the median borrower's credit score based on CoreLogic servicer data (Li and Goodman, 2015), both suggest that underwriting standards tightened after 2007. [15]

Based on differences in the macroeconomic and underwriting environment, we define the early-boom sample as mortgages that have a first payment month between July 2002 and July 2004, the late-boom sample as mortgages that have a first payment month between July 2005 and July 2007, and the post-crisis sample as mortgages that have a first payment month between July 2011 and July 2013.[16] The performance window for each loan is the first 24 payment months or the time until

[8] Our definitions of VLI, LI and MI are consistent with FHFA's current definitions of LMI categories used in the measurement of 2015–2017 Enterprise Housing Goals (see http://www.gpo.gov/fdsys/pkg/FR-2015-09-03/pdf/2015-20880.pdf).

[9] Our focus in this paper is on LMI loan performance within the context of conventional conforming lending (i.e. non-government mortgages with balances that conform to the conventional loan limits), and the reliance on Fannie data should not present a significant limitation in generalizing to the broader conventional conforming market. We focus on single family 1–4 unit properties and exclude condos and manufactured housing from the sample. These two property types may potentially be subject to increased unobservable regional risk exposures relative to other property types. We also control for the number of units in our modeling approach.

[10] CLTV represents the combined LTV of the first and any subordinate liens. It is equal to LTV when there are no subordinate liens present.

[11] With this definition of default there is a potential that a portion of the sample may represent mortgage fraud and compromise the ability to model actual mortgage credit performance. In a separate robustness check we drop the loans in our sample that never made a payment (∼2 percent of all defaults) and repeat the modeling analysis. We find in this case that there is no systematic relationship between LMI and potentially fraudulent loans, and our results estimating marginal risks of LMI loans stay the same.

[12] See for instance Haughwout et al. (2008).

[13] See https://www.fanniemae.com/content/guide/selling/e/3/glossary.html for a definition of long term standby commitment.

[14] Ineligible loans also include a set of historical loans that were given contemporaneous designations indicating that the loans contained some unspecified feature that did not meet standard business requirements at the time of acquisition.

[15] It should be noted that the Senior Loan Officer Opinion Survey only applies to mortgages originated by banks and may not fully reflect changes in underwriting standards for the entire mortgage market.

[16] Our interest in this paper is in isolating periods where there are meaningful differences in underwriting environments as well as in the subsequent economic environment. Inevitably, choosing starting and ending points for these periods will involve some degree of judgement. We have investigated the effects on our results of adjusting these time periods, for instance by extending the late-boom period through 2008, and our major results were robust to these changes.

Table 1
Variable definitions and means by sample period (all loans).

Variables	Description	Origination Vintages		
		2002–2004	2005–2007	2011–2013
Default	90 or more days past due within 24 months since first payment date	1.97%	5.92%	0.27%
CLTV	The mortgage's combined total loan to value ratio at origination, %	83.23	85.77	81.10
FICO	The lower of borrower and co-borrower's FICO scores at origination	713.26	714.27	757.22
DTI	The total combined monthly debt to monthly income ratio	0.36	0.40	0.33
Difference of FICOs	The absolute difference of borrower and co-borrower FICO scores	31.15	31.73	24.29
Subordinate Finance	10% or less of subordinate financing	5.30%	6.33%	1.53%
	>10% but ≤15% subordinate financing	5.81%	6.49%	0.50%
	20% or more subordinate financing	3.48%	11.15%	0.45%
	Other type of subordinate financing	1.52%	2.28%	1.83%
Number of Borrower	One Borrower	47.30%	51.88%	48.45%
First-time Homebuyer	First-time homebuyer	26.41%	39.46%	38.44%
Third Party Origination	Broker: The mortgage is initiated through a broker	20.37%	19.19%	8.30%
	Correspondent: The mortgage is initiated through a correspondent	30.95%	38.11%	40.34%
Loan Type	FRM15: 15 and 20 year fixed rate mortgage	9.58%	6.15%	11.33%
	FRM30: 25, 30 and 40 year fixed rate mortgage	88.29%	93.30%	88.67%
Number of Units	One unit	98.07%	98.84%	99.20%
No or Low Documentation	No documentation	1.68%	0.72%	0.00%
	Low or reduced documentation	11.24%	15.16%	0.00%
Borrower's Household Income to Area Median household Income	Income/AMI below or equal to 50%	7.37%	7.23%	6.49%
	Income/AMI greater than 50% and less than or equal to 80%	21.75%	21.01%	17.21%
	Income/AMI greater than 80% and less than or equal to 100%	16.02%	14.83%	12.25%
	Income/AMI greater than 100%	54.85%	56.92%	64.04%
Number of Loans (1,000)		2204	1906	1141

Note: 1. Loan population is conventional single-family owner-occupied 1–4 unit fixed-rate purchase loans acquired by Fannie Mae.
2. Difference in FICOs is summarized for only those loans with FICO scores for both the borrower and the co-borrower.

Table 2
Summary statistics of 2011–2013 sample by relative income status (all loans).

Variables	Description	Income/AMI			
		≤50%	>50% and ≤80%	>80% and ≤100%	>100%
Default	90 or more days past due within 24 months since first payment date	0.71%	0.39%	0.31%	0.18%
CLTV	The mortgage's combined total loan to value ratio at origination, %	77.26	80.82	81.83	81.43
FICO	The lower of borrower and co-borrower's FICO scores at origination	750.22	754.83	756.68	758.67
DTI	The total combined monthly debt to monthly income ratio	0.38	0.35	0.34	0.31
Difference of FICOs	The absolute difference of borrower and co-borrower FICO scores	28.37	26.63	26.24	23.69
Subordinate Finance	10% or less of subordinate financing	0.90%	0.81%	0.87%	1.92%
	>10% but ≤15% subordinate financing	0.40%	0.30%	0.32%	0.60%
	20% or more subordinate financing	2.14%	0.93%	0.36%	0.17%
	Other type of subordinate financing	1.14%	0.76%	0.47%	2.46%
Number of Borrower	One Borrower	84.66%	72.62%	59.49%	36.18%
First-time Homebuyer	First-time homebuyer	64.36%	57.32%	47.96%	28.92%
Third Party Origination	Broker: The mortgage is initiated through a broker	8.06%	8.15%	8.45%	8.34%
	Correspondent: The mortgage is initiated through a correspondent	41.29%	41.79%	40.92%	39.74%
Loan Type	FRM15: 15 and 20 year fixed rate mortgage	7.62%	8.42%	9.31%	12.88%
	FRM30: 25, 30 and 40 year fixed rate mortgage	92.39%	91.59%	90.69%	87.12%
Number of Units	One unit	99.09%	98.87%	98.85%	99.36%
Number of Loans (1,000)		74	196	140	731

Note: 1. Loan population is conventional single-family owner-occupied 1–4 unit fixed-rate purchase loans acquired by Fannie Mae.
2. Difference in FICOs is summarized for only those loans with FICO scores for both the borrower and the co-borrower.

termination, whichever is shorter.

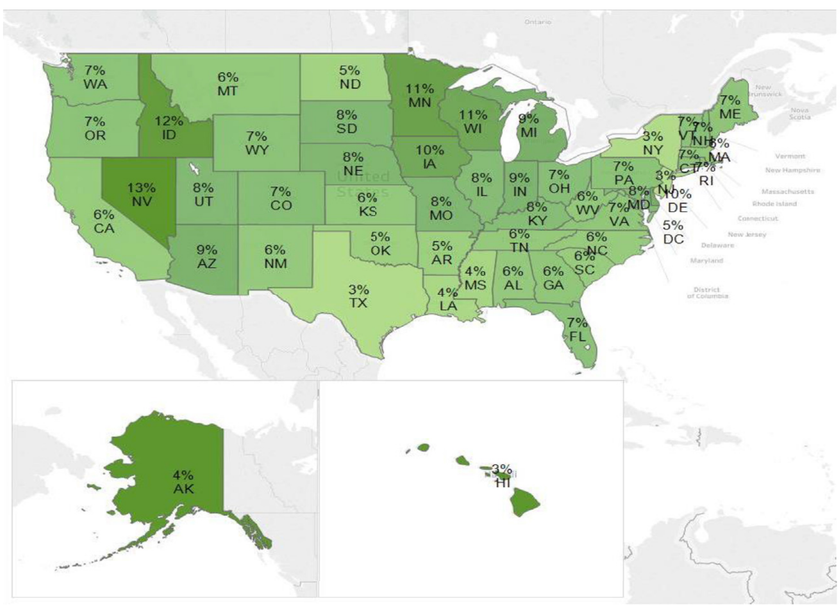
Variable definitions and summary statistics for all variables included are listed in Table 1. Our sample sizes are ~2.2 million loans in the early-boom period, ~1.9 million loans in the late-boom period and ~1.1 million loans in the post-crisis period. In our sample, 1.97 percent of loans originated during 2002–2004, 5.92 percent of loans originated between 2005 and 2007 and 0.27 percent of loans originated between 2011 and 2013 experienced a default. Comparing the means of credit risk factors, we find that the post-crisis period acquisitions have higher FICO scores, lower CLTVs, and lower DTIs than the earlier origination vintages. In the post-crisis period, more loans have a co-borrower present compared with the late-boom period, more loans have no second liens attached, and all loans have full documentation of income. This reflects the dramatic changes in the regulatory and lending environment during and after the financial crisis. Comparing the first two time periods, the late-boom period had higher average CLTVs, DTIs, and higher shares of loans with subordinate financing, third party originators and reduced documentation. However, the worsening of the credit profile evident in the late-boom period in Table 1 does not fully capture the market decline in credit standards for two reasons. First, GSE credit standards were more restrictive than those in the private label securities market prior to the financial crisis. Second, our analysis focuses on the fixed-rate mortgage population and excludes ARM originations, which were used during the housing boom as a means to stretch household budgets in the face of rapid home price appreciation, especially in the PLS market for borrowers who did not meet the prevailing GSE eligibility requirements (Agarwal et al., 2012).

The share of LMI loans in our sample was greatest in the early-boom

Table 3
Correlations between relative income and other risk characteristics in each sample period (all loans).

	2002–2004			2005–2007			2011–2013		
	DTI	CLTV	FICO	DTI	CLTV	FICO	DTI	CLTV	FICO
Panel A: Sample Mean of Variables									
Income/AMI ≤50%	0.45	81.79	706.67	0.47	87.68	688.44	0.38	77.26	750.22
	(0.15)	(19.95)	(66.46)	(0.13)	(19.37)	(72.13)	(0.07)	(16.74)	(46.91)
Income/AMI > 50% and ≤80%	0.41	84.55	710.15	0.44	88.48	701.11	0.35	80.82	754.83
	(0.13)	(16.66)	(62.19)	(0.12)	(17.01)	(68.24)	(0.08)	(14.91)	(43.78)
Income/AMI > 80% and ≤100%	0.38	84.44	712.23	0.41	87.46	709.14	0.34	81.83	756.68
	(0.13)	(15.62)	(59.85)	(0.12)	(16.59)	(65.1)	(0.09)	(14.10)	(42.20)
Income/AMI > 100%	0.33	82.54	715.67	0.37	84.08	723.76	0.31	81.43	758.67
	(0.13)	(15.71)	(57.17)	(0.12)	(16.97)	(58.15)	(0.09)	(13.56)	(39.68)
Panel B: Correlation Coefficient of Variables with Relative Income									
Borrower Income Relative to AMI	−0.32***	−0.09***	0.05***	−0.31***	−0.14***	0.15***	−0.29***	−0.04***	0.03***

Note: 1. Loan population is conventional single-family owner-occupied 1–4 unit fixed-rate purchase loans acquired by Fannie Mae.
2. Panel A: standard deviations are in parentheses.
3. Panel B: *10%,**5%,***1% Significance Levels.



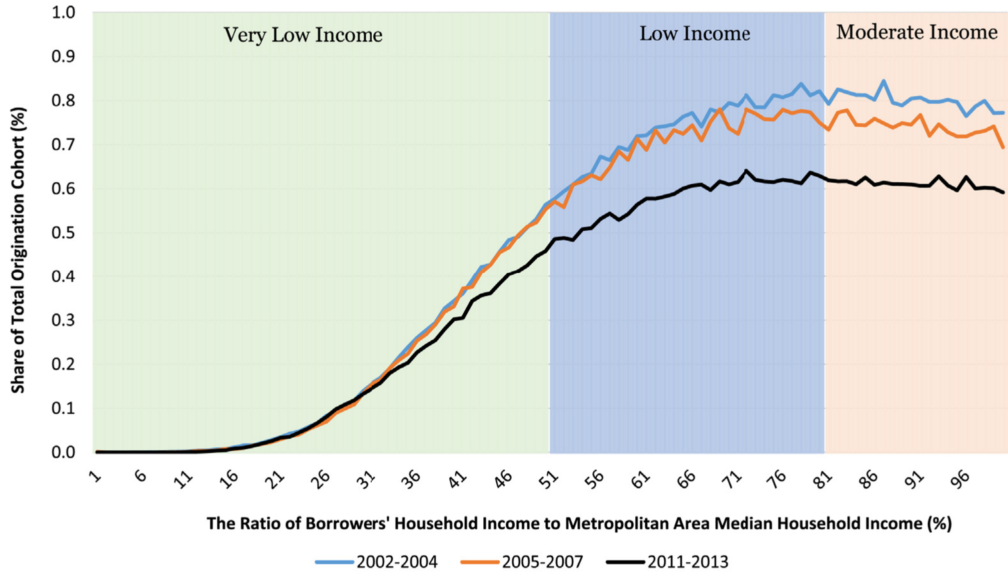
Note: 1. Total loan population is conventional single-family owner-occupied 1-4 unit fixed-rate purchase loans acquired by Fannie Mae.
2. VLI shares vary significantly across states with a range of 3% to 13%. The states with the lowest share of VLI borrowers are HI, NJ, and NY (3%). The states with highest share of VLI borrowers are NV (13%), ID (12%), WI (11%) and MN (11%).

Fig. 1. Share of very low income borrowers by states in the post-crisis period (2011–2013, all loans).

period and lowest in the post-crisis period. In the early-boom period, 7.37 percent of loans were to VLI borrowers, 21.75 percent were to LI borrowers and 16.02 percent were to MI borrowers.[17] In the late-boom period, these values are similar at 7.23 percent, 21.01 percent and 14.83 percent, respectively. In the post-crisis period, however, the share of LMI loans is lower, with a 6.49 percent share of VLI, a 17.21 percent share of LI and a 12.25 percent share of MI borrowers.

In Table 2 we provide a breakdown of the risk characteristics by relative income status of the borrowers for the 2011 to 2013 period. The default rate is 0.71 percent for VLI borrowers, in contrast to 0.39 percent for LI borrowers, 0.31 percent for MI borrowers and 0.18 percent for HI borrowers. In general, credit risk factors improve with the relative income of borrowers. For instance, FICO rises slightly with relative income and DTI and the share of single borrowers fall with relative income. CLTV, however, is the lowest for the lowest relative income group (77.26), rises for the next two income groups (80.82 and 81.83, respectively) and slightly declines for the highest income group (81.43). Table 2 also shows that borrowers with lower relative income are considerably more likely to have one borrower or be a first-time homebuyer in this period. Table 3 provides a further breakdown of the three key credit risk characteristics of DTI, CLTV and FICO by relative income group for each sample period. DTI declines noticeably with relative income in each of the sample periods, but although FICO rises with income in each period, this increase is not very pronounced except for the late-boom period. CLTV does not show a consistent pattern across relative income groups and time periods. The Pearson correlation coefficient estimates in Panel B

[17] One potential issue that needs to be considered is the mismeasurement of income, as this is the primary variable of interest in this paper. One source of inaccurate measurement of income, involves higher income borrowers who potentially only report enough income to qualify, which is not a significant concern for this paper, as we focus on the LMI populations. Another segment with potentially significant mismeasurement of income are those mortgages with low or no documentation of income. As these loans no longer meet eligibility requirements for the GSEs, we remove these loans from much of our later analysis in the paper.



Note: 1. Total loan population is conventional single-family owner-occupied 1-4 unit fixed-rate purchase loans acquired by Fannie Mae. 2. The distribution is across all loans but we show only the LMI cohorts on the chart.

Fig. 2. Distribution of low and moderate income borrowers by relative income (all loans).

show a significant negative correlation between relative income and DTI and a significant positive correlation with FICO. Furthermore, there is also a statistically significant negative relationship between relative income and CLTV in each sample period. Note also that the VLI population generally has a higher standard deviation for these credit risk factors relative to other income groups (with the exception of DTI in the post-crisis period). This indicates greater dispersion of credit risk attributes across the VLI populations.

Fig. 1 presents the share of VLI loans in the post-crisis period over all loans by state. The share of VLI loans varies significantly across states based on income distributions within each state, ranging from 3 percent (Hawaii, New York, New Jersey) to 13 percent (Nevada) in this period, with an average of 6 percent. The LI loan distribution (not shown) has a similar pattern across states, ranging from 11 percent (Texas) to 25 percent (Wisconsin and Vermont), with an average of 17 percent.

Fig. 2 highlights the changing composition of the purchase loans in our sample over time. In particular the share of purchase loans falling into the LMI categories has decreased in the most recent period (2011–2013) compared with the previous time periods (2002–2004 and 2006–2008). This relative shift away from lower income purchase borrowers has occurred as income growth and wealth accumulation for this segment have lagged behind those of the higher income groups (see for instance Kochhar and Cilluffo, 2017 and Shambaugh et al., 2017) and average origination credit scores have increased in the post-crisis period (Table 1), while home price growth has been fastest since 2011 at lower home price tiers where housing inventory has been more constrained.[18] This has potentially undermined the ability of LMI homebuyers to qualify for a mortgage, save for a down payment and find a home at a price point they can qualify for and with a payment they are comfortable with.

This also raises the question of comparability between the LMI borrowers across different time periods, in particular, the extent to which the LMI borrowers in the most recent period are different from those in the two earlier periods, given recent tighter underwriting requirements. To address this concern, we focus on the eligible population under today's standards for much of our later analysis. Later in the paper, we also present evidence that the standard underwriting model sufficiently predicts risk for most borrowers. Furthermore, for the VLI population, where there is some significant unexplained risk outside of the model, the prediction error is stable across alternative time periods. We see this as evidence that, although sensitivities may change over time, the general model structure and drivers of credit risk for LMI borrowers are similar across the time periods we analyze.

4. Underwriting standards and default risks

In this section we present empirical results that focus on the actual default rates and default odds ratios for LMI borrowers relative to those of the HI category.[19] We begin with a comparison of raw default rates across LMI categories and time periods. We then turn to an investigation of the extent to which tighter eligibility standards in the post-crisis period have helped limit the risks across income categories. In particular, following the 2008 financial crisis, Fannie Mae tightened underwriting standards by removing eligibility for purchase loans with the following characteristics: LTV greater than 97 percent; FICO scores less than 620; DTI ratios greater than 0.50; loans with features including interest only/negative amortization, low documentation of income, 40-year terms and balloon payments. We analyze the effectiveness of this tighter underwriting on sustainable homeownership for the LMI borrowers in a subset of the results that follow by applying today's eligibility standards to historical originations from the 2002–2004 and 2005–2007 sample periods and tracking the default behavior of the restricted population across relative income groups.[20]

4.1. Low and moderate income default rates

We first look at raw default rates across LMI categories before controlling for any post-crisis eligibility considerations or other credit risk

[18] Source: CoreLogic (https://www.corelogic.com/insights/corelogic-home-price-insights.aspx).

[19] We calculate the odds ratio of default as the odds of default for a given LMI category over the odds of default for the HI category. For instance, in the post-crisis period, we find a default rate of 0.71 percent for all VLI loans and 0.19 percent for all HI loans. The odds ratio in this case is calculated as $(0.0071/(1–0.0071))/(0.0019/(1–0.0019)) = 3.84$.

[20] All conventional loans eligible for purchase by Fannie Mae during the post-crisis sample conform to this underwriting/eligibility criteria.

Table 4
Default rate and odds ratio by income group: eligible vs. ineligible loans.

	2002–2004		2005–2007		2011–2013	
	Default Rate	Odds Ratio Relative to Hi-inc. Elig.	Default Rate	Odds Ratio Relative to Hi-inc. Elig.	Default Rate	Odds Ratio Relative to Hi-inc. Elig.
Total Loans	1.97%	3.73	5.92%	5.95	0.27%	1.46
Income/AMI ≤50%	4.38%	8.48	12.22%	13.15	0.71%	3.84
Income/AMI >50% and ≤80%	2.89%	5.52	8.50%	8.78	0.39%	2.10
Income/AMI >80% and ≤100%	2.11%	4.00	6.45%	6.52	0.31%	1.70
Income/AMI >100%	1.24%	2.33	4.03%	3.97	0.19%	1
Number of Observations (1,000)	2204		1906		1141	
All Ineligible Loans	3.60%	6.91	9.49%	9.91		
Income/AMI ≤50%	6.32%	12.49	15.47%	17.3		
Income/AMI > 50% and ≤80%	4.59%	8.9	11.76%	12.59		
Income/AMI > 80% and ≤100%	3.67%	7.06	9.73%	10.18		
Income/AMI > 100%	2.47%	4.69	7.10%	7.23		
Number of Observations (1,000)	925		1081			
All Eligible Loans	0.80%	1.49	1.24%	1.19	0.27%	1.46
Income/AMI ≤50%	1.84%	3.47	2.34%	2.27	0.71%	3.84
Income/AMI > 50% and ≤80%	1.23%	2.31	1.68%	1.61	0.39%	2.10
Income/AMI > 80% and ≤100%	0.90%	1.68	1.34%	1.28	0.31%	1.70
Income/AMI > 100%	0.54%	1.00	1.05%	1.00	0.19%	1.00
Number of Observations (1,000)	1279		825		1141	

Note: 1. Total loan population is conventional single-family owner-occupied 1–4 unit fixed-rate purchase loans acquired by Fannie Mae.
2. Non-eligible loans include LTV >97; FICO <620; DTI >0.50, interest only/negative amortization loans, no or low doc loans, 40-year terms, balloons, and loans with designations indicating they contain features that did not meet standard business requirements at the time of acquisition.
3. Default rate is the actual 90 + day delinquency rate within the first 24 month since the first-payment month among each group of loans.

factors. The results for this analysis are shown in the top panel of Table 4. We list the actual default rate for each relative income group for all originations in a given sample period in the first column and the default odds ratio relative to the HI population in the second column. Consistent with past findings in the literature (see Firestone et al., 2007; Avery and Brevoort, 2015), the default rate increases as borrowers' income relative to area median income decreases. Before considering changes in eligibility standards or risks explained by other observable factors, the absolute and relative risk of LMI groups vary greatly over different time periods due to differences in underwriting regime and subsequent housing market experience.

Table 4 shows that the default rate of all loans in the 2002–2004 sample period, when home prices were appreciating, ranges from 4.38 percent to 1.24 percent as relative income increases. For the period with more lax underwriting and a subsequent housing price decline (2005–2007), actual default rates range from 12.22 percent to 4.03 percent. Due to the sharp decline of home prices during this period, even the HI group experienced a default rate more than triple that of their counterparts in the early-boom period. The most recent period is characterized by tighter underwriting relative to the other two periods followed by an improving macroeconomic and housing market environment. The default rate by area median income category ranges from 0.71 percent to 0.19 percent, the lowest among all three time periods.

4.2. Importance of tighter eligibility standards in limiting low and moderate income risk

The middle and bottom panels of Table 4, respectively, present the default rates of the loans that would have been ineligible and those that would remain eligible for delivery to Fannie Mae under current guidelines.[21] As the table indicates, the ineligible loans have a higher default rate than eligible loans for each relative income group in the early-boom and late-boom periods. For example, for the 2005–2007 period, the ineligible VLI loans have a default rate of 15.47 percent, higher than the 2.34 percent for their counterparts in the eligible loan category. The default odds ratio of the ineligible VLI loans relative to the eligible high-income loans in the late-boom period is 17.30, while the odds ratio of the eligible VLI loans relative to the eligible HI loans in the same period is only 2.27. For the 2002–2004 period, the ineligible VLI to the eligible HI default odds ratio is 12.49 versus 3.47 for the eligible VLI to eligible HI default odds ratio. We find a similar pattern for the LI, MI and HI categories, indicating the importance of tighter eligibility criteria in lowering default rates across all income groups in both appreciating and depreciating home price environments. The multifold reduction in the absolute level of defaults across all incomes demonstrates the importance of eligibility criteria in managing default risk and thereby contributing to sustainable homeownership.

Table 5 takes a closer look at the key eligibility changes in reducing default rates for the entire population of purchase loans. Similar to Table 4, the first column under each time period represents the default rate in each period and the second column represents the default odds ratio relative to the eligible HI loans in each period. The top row of numbers for each time period represents the default rate before any currently ineligible loans are removed. Each row removes an additional

[21] Had the tighter eligibility rule in the post-crisis period applied prior to 2002, 42 percent of the loans originated in the early-boom period and 57 percent of the loans originated in the late-boom period would have no longer been eligible.

Table 5
Impact on default odds ratio and remaining loans after removing ineligible loans.

Loans Removed	2002–2004			2005–2007			2011–2013		
	Default Rate	Odds Ratio Relative to Hi-inc. Elig.	Loan Counts (1,000)	Default Rate	Odds Ratio Relative to Hi-inc. Elig.	Loan Counts (1,000)	Default Rate	Odds Ratio Relative to Hi-inc. Elig.	Loan Counts (1,000)
None	1.97%	3.73	2204	5.92%	5.95	1906	**0.27%**	1.46	1141
LTV>97	1.49%	2.81	1997	3.19%	3.11	1464			
FICO<620	1.01%	1.89	1872	2.63%	2.55	1400			
No or Low Documentation	0.90%	1.68	1621	1.72%	1.66	1195			
DTI>0.50	0.84%	1.57	1350	1.55%	1.49	911			
Limited Amortization	0.85%	1.58	1331	1.39%	1.33	870			
All Other Ineligible Loans	0.80%	1.49	1279	1.24%	1.19	825	0.27%	1.46	1141

Note: 1. Total loan population is conventional single-family owner-occupied 1–4 unit fixed-rate purchase loans acquired by Fannie Mae.
2. Non-eligible loans include LTV >97; FICO <620; no or low doc loans, DTI >0.50, interest only, negative amortization, 40-year term, balloons, and loans with designations indicating they contain features that did not meet standard business requirements at the time of acquisition. Eligible loans include all other loans.
3. Exclusions accumulate down the column. For instance in moving from the second to third row, the excluded population now includes LTV>97 and FICO<620.

set of loans based on the indicated criteria, with the third column capturing how many loans remain after the particular eligibility consideration is applied. For instance, removing loans with LTVs greater than 97, removes ~200 thousand loans and drops the early-boom default rate from 1.97 percent to 1.49 percent. Restricting the population to FICO greater than or equal to 620, removes ~100 thousand more loans and further reduces the default rate to 1.01 percent. Other factors that drive the reduction in risk include removing loans with low or no documentation of income and with DTI >0.50. Removing loans with limited amortization features (including interest-only, negatively amortizing, 40-year mortgage and balloon loans), however, does not change default rates very much. Including all post-crisis eligibility requirements drops the default rate to 0.80 percent. A similar pattern holds for the late-boom period, but with greater absolute reductions in risk when excluding LTV >97 (default rates drop from 5.92 percent to 3.19 percent). The FICO eligibility consideration further reduces the default rate to 2.63 percent, while the full documentation eligibility criteria reduces the default rate to 1.72, and the DTI requirement reduces the default rate to 1.55 percent. Removing those loans with limited amortization features in this period now reduces risk meaningfully, with the default rate dropping to 1.39 percent. Finally including the remaining eligibility requirements drops the default rate to 1.24 percent.

One important caveat here is that we assume the eligibility restrictions would have resulted in the ineligible loans never having been made. In fact, these borrowers could have adjusted aspects of their loan profiles, including changing down payments or opting for an eligible product, for instance, to move into the eligible population under today's standards. This would result in including borrowers who historically defaulted at a higher rate and could potentially have presented higher default risks despite the hypothetical improvement in their loan profiles compared with the strictly eligible population. Thus, we should consider the observed effects in Tables 4 and 5 of introducing the current eligibility requirements as upper bounds on the reduction in risk.

5. Marginal credit risk of LMI borrowers

We now turn to the marginal risk associated with LMI lending for the population of loans that are eligible under current underwriting standards. To analyze this marginal risk we start with the following standard logit model of credit risk:

$$\Pr(90 \text{ days delinquent in 24 months}_i) = f(X_i\beta + \textit{fixed effects}_i) \qquad (1)$$

Here, X_i refers to the vector of loan-level characteristics for loan i; β is a vector of parameters; and *fixed effects*$_i$ represents the set of fixed effects we use to control for state and month of acquisition.

We rely on a two-step modeling approach to analyze the extent to which standard underwriting measures adequately account for differences in credit risks across income groups. Specifically, we start by regressing the default outcome on the standard risk characteristics of the loan without controlling for the LMI indicators in the first step according to equation (1). The estimated coefficient vector is denoted as $\widehat{\beta}$, and $X_i\widehat{\beta}$ is the predicted log default odds from the first step.

We then carry $X_i\widehat{\beta}$ as a right-hand side variable in the second step with its coefficient fixed at one and regress the default outcome on $X_i\widehat{\beta}$ along with the relative income indicators VLI_i, LI_i, MI_i and HI_i as follows[22]:

$$\Pr(90 \text{ days delinquent in 24 months}_i) = f\left(X_i\widehat{\beta} + \gamma_1 VLI_i + \gamma_2 LI_i + \gamma_3 MI_i + \gamma_4 HI_i\right) \qquad (2)$$

Here $\gamma_1, \gamma_2, \gamma_3$ and γ_4 are our coefficients of interest, measuring the increase in risk for the respective relative income. We estimate two versions of equation (2). The first excludes the HI category in the estimation, so that coefficients on the remaining LMI indicators can be converted to odds ratios for each of the LMI groups relative to the HI group, and the significance of the coefficient provides the significance for the odds ratio. If the contribution of a given LMI categorical indicator results in an odds ratio that is significant and greater than one, then we can conclude that lending to this LMI group is on average more risky than lending to borrowers with an income above 100 percent of area median income after fully controlling for the standard risk characteristics and state and time fixed effects. The second version of equation (2) we estimate includes all of the relative income indicators, including the HI group. This version allows us to test directly whether the standard underwriting model fully explains default risks or whether any particular relative income group has significant remaining risks after controlling for the standard underwriting variables.[23]

We evaluate the general rank-ordering power of the estimated models using the Gini coefficient. We also include AIC as a standard measure of the goodness-of-fit for each of the models.

[22] Since the coefficient of $X_i\widehat{\beta}$ is fixed at one, the second step essentially regresses the residual from the first step on the LMI indicators. For more technical details and an example using the two-step approach see: https://support.sas.com/documentation/cdl/en/statug/63033/HTML/default/viewer.htm#statug_logistic_sect064.htm.

[23] The results in Tables 6 and 7 are based on the first approach to estimating equation (2) that excludes the HI group, while Fig. 3 and Tables A1–A4 are based on the second approach which includes the HI group.

Table 6
Default odds ratio controlling for additional underwriting variables (eligible loans only).

	LMI Only	State Fixed Effect	CLTV	Num. Borr.	FICO	Full Set
2002–2004 Originations						
Income/AMI ≤50%	3.47***	3.25***	3.27***	1.78***	1.60***	1.46***
Income/AMI >50% and ≤80%	2.31***	2.25***	2.08***	1.29***	1.28***	1.18***
Income/AMI >80% and ≤100%	1.68***	1.68***	1.55***	1.15***	1.16***	1.09***
Income/AMI >100%	1.00	1.00	1.00	1.00	1.00	1.00
Gini Coefficient without income		0.238	0.469	0.561	0.725	0.738
Gini Coefficient with income	0.232	0.335	0.511	0.569	0.729	0.740
-2LogL	116,776	114,936	109,950	107,538	97,878	96,830
AIC	116,785	114,945	109,959	107,546	97,887	96,838
Number of Observations (1,000)	1279	1279	1279	1279	1279	1278
2005–2007 Originations						
Income/AMI ≤50%	2.27***	2.54***	2.95***	1.84***	1.34***	1.30***
Income/AMI > 50% and ≤80%	1.61***	1.76***	1.77***	1.21***	1.07***	1.03
Income/AMI > 80% and ≤100%	1.28***	1.36***	1.32***	1.03	0.98	0.95
Income/AMI > 100%	1.00	1.00	1.00	1.00	1.00	1.00
Gini Coefficient without income		0.248	0.458	0.530	0.693	0.717
Gini Coefficient with income	0.053	0.316	0.494	0.539	0.694	0.718
-2LogL	109,396	106,768	101,762	100,086	91,840	89,450
AIC	109,405	106,777	101,771	100,094	91,849	89,458
Number of Observations (1,000)	825	825	825	825	825	825
2011–2013 Originations						
Income/AMI ≤50%	3.84***	3.88***	4.00***	2.41***	1.84***	1.57***
Income/AMI > 50% and ≤80%	2.10***	2.10***	2.05***	1.36***	1.22***	1.10**
Income/AMI > 80% and ≤100%	1.70***	1.70***	1.65***	1.24***	1.20***	1.12**
Income/AMI > 100%	1.00	1.00	1.00	1.00	1.00	1.00
Gini Coefficient without income		0.156	0.258	0.416	0.679	0.694
Gini Coefficient with income	0.228	0.298	0.378	0.449	0.689	0.700
-2LogL	41,836	41,590	41,176	40,570	36,934	36,640
AIC	41,845	41,599	41,185	40,579	36,943	36,649
Number of Observations (1,000)	1141	1141	1141	1141	1141	1141

Note: 1. Population is eligible loans only and excludes LTV >97; FICO <620; DTI >0.50, interest only/negative amortization loans, no or low doc loans, 40-year terms, balloons, and loans with designations indicating they contain features that did not meet standard business requirements at the time of acquisition.
2. Controls accumulate across the row. For instance in moving from the third to fourth column, the control variables now include LMI, state fixed effects and CLTV.
3. Full set of explanatory variables also includes DTI, first-time homebuyer indicator, subordinate financing indicators, product type, number of units, third party originator indicators and acquisition date (measured by month) fixed effects. Model uses two-step logit to predict 90-day delinquency in first 24 months.
4. *10%,**5%,***1% Significance Levels.

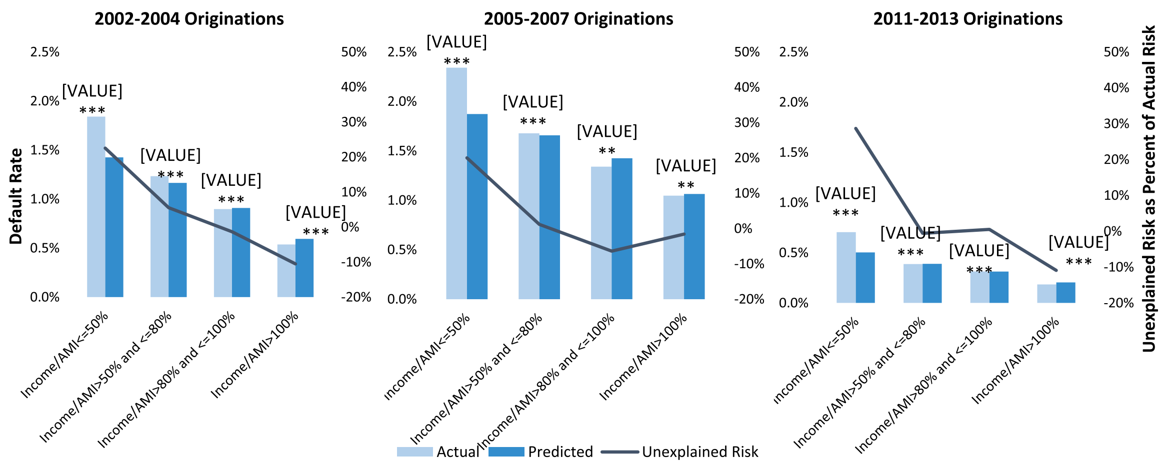
5.1. Importance of standard underwriting factors in explaining low and moderate income risk

Table 6 shows the relative odds ratios for each LMI category after introducing a cumulative set of controls by estimation time period for the eligible population.[24] The first column of numbers for the first four rows of each estimation period shows the odds ratios for the case with only the LMI indicators.[25] As controls are introduced, to the extent that riskier loan characteristics are more likely in the LMI groups, this should lower the odds ratios associated with the LMI categories. The final column of numbers in the first four rows of each estimation period represents the residual risks associated with each LMI category after controlling for all of the other explanatory variables (including DTI, FICO, CLTV, subordinate financing indicators, number of borrowers, third party originator indicators, first-time homebuyer indicator, term and number of units associated with each loan and state and time fixed effects).[26] One important takeaway is that after controlling for these standard risk characteristics, the relative risk of LMI lending versus HI lending declines significantly.

In particular, LMI loans have between 1.09 and 1.46 times the unexplained default risk after accounting for underwriting, geographic and time controls relative to HI loans in the early-boom period. The results are similar for the post-crisis period, with 1.10–1.57 times the unexplained default risk for LMI loans relative to HI loans. For the late-boom period, only the VLI group has significant additional risks at 1.30 times the relative default risk of the HI loans. The control variables that account for the biggest reduction in marginal risk for the VLI borrowers in the early-boom period are the state-level effect, the number of borrowers and FICO. In other words, the lower credit scores and higher frequency of having only one borrower for this sample period, once accounted for, reduce the marginal risk of these borrowers. For the other LMI borrowers, their higher CLTVs seem to have a greater effect on their higher defaults relative to the HI borrowers, with number of borrowers still an important factor, but FICO less so. In the late-boom and post-crisis periods, the number of borrowers remains important in explaining the higher default rates of the LMI population, while FICO also seems to be an important driver of relative LMI defaults.

We also provide Gini coefficients for the models in Table 6 to gauge the additional value of including LMI indicators in sorting out the credit risks at the time of underwriting. The models with a full set of controls including LMI indicators, standard underwriting factors and region and vintage controls have much higher Gini coefficients (0.700–0.740) than models with just LMI indicators (0.053–0.232), indicating the relative

[24] The results presented here and in the remainder of the paper only include loans eligible for delivery to Fannie Mae under current underwriting standards.
[25] These results are equivalent to the odds ratios for the eligible loans in Table 3.
[26] The complete set of respective coefficients for the early-boom, late-boom and post-crisis time periods are available in Tables A1–A3 for the eligible population in the appendix. We also present credit risk coefficients in Table A4 for the entire population (including both eligible and ineligible loans) for purposes of reference. We also provide a discussion of the parameter estimates in the appendix.



Note:
1. Population is eligible loans only and excludes LTV > 97; FICO < 620; DTI > 0.50, interest only/negative amortization loans, no or low doc loans, 40-year terms, balloons, and loans with designations indicating they contain features that did not meet standard business requirements at the time of acquisition.
2. Light blue bars show actual default rates and dark blue bars show predicted default rates. The curves show the unexplained risk as the percent of actual default rate (values on right axis).
3. The predicted default is produced using the first-step logit model in the last column of Table 6, which uses the full list of control variables (CLTV, FICO, DTI, number of borrowers, the first-time homebuyer indicator, subordinate financing indicators, product type, number of units, third party originator indicators, state fixed effects and acquisition date fixed effects) but does not include the relative income variables.
4. *10%,**5%,***1% Significance Levels. Based on two-step logistic regression coefficients in Tables A1 – A3.

Fig. 3. Comparison of actual and predicted default by relative income group (eligible loans only).

importance of factors outside of LMI indicators in explaining credit risk. Furthermore, the Gini coefficients do not substantially change when adding relative income controls to the full set of standard underwriting variables (increases range from 0.001 to 0.006), indicating the limited additional gain in including LMI indicators in ranking overall credit risks in a standard underwriting model.

Fig. 3 provides additional evidence on the performance of the standard underwriting credit risk model in predicting default risk. In particular, the left bar for each time period and relative income group shows the average actual default rate for the corresponding group of loans; the right bar shows the average predicted default rate using a traditional credit risk model that includes only the standard underwriting factors and vintage and region controls but no relative income controls. The difference of the two bars is the default risk that is unexplained by the standard underwriting model. The curves displays the unexplained risk as the percent of actual default risk (calculated as one minus the ratio of predicted over actual default rates). Comparing the performance of the LMI loans to their predicted performance based on a fully specified underwriting model allows us to gauge the reasonableness of such a model in explaining the observed behavior of these loans. Note that for all sample periods the VLI group has a larger portion of credit risk unexplained by the model.[27] For the 2002–2004 period, the actual default rate of VLI loans is ~23 percent greater than predicted, for the 2005–2007 period the corresponding value is ~20 percent and for the 2011–2013 period, the corresponding value is ~29 percent. Thus, despite substantial variation in default rates, as Fig. 3 illustrates the ratio is reasonably stable across the three periods, with the lowest percentage of unexplained risk coming in the late-boom period and the highest percentage of unexplained risk coming in the recent low default post-crisis period. For the remaining LMI categories, this ratio is small as a percentage of additional unexplained risk relative to total risk. For instance, the additional unexplained risk for the LI borrowers is ~6 percent in the early-boom period, ~1 percent in the late-boom period and ~-1 percent (meaning over-predicted by the standard underwriting model) in the post-crisis period.

In summary, the measures of relative income we introduce to the standard underwriting framework do not help with the rank-ordering power of the model (as measured by the Gini coefficient), indicating that relative income for most loans is not a key missing variable in standard underwriting models. Furthermore, standard underwriting factors are reasonably sufficient in identifying the key risks associated with credit performance for most LMI segments as shown by the multi-fold decline in the odds ratio within relative income segments and across the segments (Table 6) and the predicted versus actual loan performance (Fig. 3). For the VLI segment, these models can help rank-order and account for a substantial portion of the risks of this group, but significant unexplained risks remain. In general, VLI loans experience default ~25 percent more than would be expected based on their credit profile alone. This error, however, is reasonably stable across the three periods, suggesting a relatively predictable degree of unexplained risk.

6. Robustness tests

Our baseline results focus on differences in loan performance controlling for a set of predictive underwriting factors as well as relative income, adopting a logit-based credit scoring approach.[28] We rely on a wide set of variables observable to the underwriter at time of loan origination and model default within 24 months controlling for the location of the loan only with state-level controls. Thus, we ignore potentially important regional variation within a state that can impact marginal LMI risks, including future home price changes, local labor market dynamics and spillover effects on loan performance from riskier lending or foreclosure contagion. Another limitation for our baseline results is that the standard logit model used for our baseline results ignores the competing risk of prepayment. Finally, our baseline specification only controls for the first-time homebuyer status of the borrower(s)

[27] This additional risk can potentially arise from the correlation of relative income with other risk factors not typically controlled for in the underwriting process, for instance industry of employment. This factor could potentially result in LMI borrowers being more vulnerable to shocks that result in a decline in income and an increased difficulty in making mortgage payments.

[28] Credit scoring models are often developed on 24 months of loan performance data (see, for example, Mays (2004) for additional details.

Table 7
Default odds ratio with additional zip code fixed effects (eligible loans only).

	2002–2004		2005–2007		2011–2013	
	State Fixed Effect	Zip Code Fixed Effect	State Fixed Effect	Zip Code Fixed Effect	State Fixed Effect	Zip Code Fixed Effect
Income/AMI ≤50%	1.46***	1.26***	1.30***	1.25***	1.57***	1.40***
Income/AMI > 50% and ≤80%	1.18***	1.08***	1.03	1.04	1.10**	1.01
Income/AMI > 80% and ≤100%	1.09***	1.03***	0.95	0.98	1.12**	1.02
Income/AMI > 100%	1.00	1.00	1.00	1.00	1.00	1.00
-2LogL	96,830	65,066	89,450	53,612	36,640	22,222
AIC	96,838	65,075	89,458	53,621	36,649	22,231
Number of Zip Codes	26,024	7144	24,396	4973	23,737	6116
Number of Observations (1,000)	1279	1071	825	612	1141	948

Note:: 1. Population is eligible loans only and excludes LTV >97; FICO <620; DTI >0.50, interest only/negative amortization loans, no or low doc loans, 40-year terms, balloons, and loans with designations indicating they contain features that did not meet standard business requirements at the time of acquisition.
2. The control variables in the first-step logit regression include CLTV, FICO, DTI, number of borrowers, the first-time homebuyer indicator, subordinate financing indicators, product type, number of units, third party originator indicators and acquisition date (measured by month) fixed effects.
3. The zip code fixed effect model removes zip codes that have 5° or fewer loan observations. The odds ratios estimated using the model with state fixed effects are robust to removing those zip codes.
4. *10%,**5%,***1% Significance Levels.

Table 8
Default and prepayment odds ratio – 24-month panel multinomial logit (eligible loans only).

	2002–2004		2005–2007		2010–2013	
	Default	Prepayment	Default	Prepayment	Default	Prepayment
24-Month Panel Multinomial Logit Controlled for CLTV						
Income/AMI ≤50%	1.49***	0.79***	1.73***	0.78***	1.88***	0.62***
Income/AMI > 50% and ≤80%	1.25***	0.93***	1.27***	0.86***	1.22***	0.78***
Income/AMI > 80% and ≤100%	1.15***	1.00***	1.04***	0.92***	1.20***	0.89***
Income/AMI > 100%	1.00	1.00	1.00	1.00	1.00	1.00
-2LogL	3,779,346		1,641,787		1,846,711	
AIC	3,779,782		1,642,225		1,847,151	
Number of Observations (1,000)	873		319		414	
24-Month Panel Multinomial Logit Controlled for MTMCLTV						
Income/AMI ≤50%	1.46***	0.77***	1.68***	0.77***	1.87***	0.63***
Income/AMI > 50% and ≤80%	1.24***	0.92***	1.23***	0.86***	1.22***	0.79***
Income/AMI > 80% and ≤100%	1.15***	0.99***	1.02***	0.93***	1.20***	0.90***
Income/AMI > 100%	1.00	1.00	1.00	1.00	1.00	1.00
-2LogL	3,776,666		1,637,390		1,845,602	
AIC	3,777,096		1,637,820		1,846,034	
Number of Observations (1,000)	873		319		414	

Note: 1. Population is eligible loans only and excludes LTV >97; FICO <620; DTI >0.50, interest only/negative amortization loans, no or low doc loans, 40-year terms, balloons, and loans with designations indicating they contain features that did not meet standard business requirements at the time of acquisition.
2. The control variables in the 24-month panel multinomial logit regressions include FICO, DTI, loan age, actual UPB, rate incentive, number of borrowers, the first-time homebuyer indicator, subordinate financing indicators, product type, number of units, third party originator indicators, state fixed effects and acquisition date (measured by month) fixed effects. In the top panel we include CLTV, and in the bottom panel we include the ratio of LTV to CLTV as well as MTMLTV in order to control for MTMCLTV.
3. *10%,**5%,***1% Significance Levels.

by including an indicator variable. In this section we present a series of robustness tests intended to address the above limitations of our baseline credit scoring approach.

In the first robustness test, we introduce zip-level fixed effects into our baseline credit scoring model in an attempt to disentangle regional drivers of loan performance from the marginal risk of LMI borrowers. Table 7 presents the results of this exercise and shows that including zip-level controls takes away some, but not all, of the marginal risks associated with LMI borrowing. For instance, the odds ratio for VLI borrowers declines from 1.46 to 1.26 in the early-boom period, from 1.30 to 1.25 in the late-boom period and 1.57 to 1.40 in the post-crisis period. Declines for other relative income groups show a similar pattern. The results here indicate that the local housing market and employment conditions explain part of the LMI marginal risk but the neighborhood effect does not fully explain all of the marginal risk of lower income borrowers.

In the second robustness check, we build a panel data set of monthly loan performance from the loans used in our baseline credit scoring approach and estimate a competing risk model of the probability of default versus prepayment over the same 24-month horizon. Since we have as many as 15 years of loan performance history for the oldest (2002) cohort, resulting panel data sets are extremely large. For computational ease, therefore, we use 100 percent of the event loan months (months in which the loan experiences either default or prepayment) and an equally sized random sample of non-event loan months.[29,30]

We augment the origination CLTV with a time-varying market-to-

[29] Default is defined as the first instance of 90 days delinquency, consistent with the baseline credit scoring approach.

[30] Coslett (1981) shows that such a procedure applied within a logit framework does not bias coefficient estimates except for that of the intercept term, for which there is a known correction formula dependent solely on the random sampling percentage.

Table 9
First-time homebuyer default and prepayment odds ratio – 24-month panel multinomial logit (eligible loans only).

	2002–2004		2005–2007		2011–2013	
	Default	Prepayment	Default	Prepayment	Default	Prepayment
Income/AMI ≤50%	1.63***	0.73***	1.53***	0.58***	1.56***	0.56***
Income/AMI > 50% and ≤80%	1.23***	0.89***	1.19***	0.74***	1.11***	0.78***
Income/AMI > 80% and ≤100%	1.21***	0.97***	1.01***	0.88***	1.24**	0.87***
Income/AMI > 100%	1.00	1.00	1.00	1.00	1.00	1.00
-2LogL	812,180		447,754		593,304	
AIC	812,604		448,178		593,732	
Number of Observations (1,000)	193		94		149	

Note: 1. Population is eligible loans only and excludes LTV >97; FICO <620; DTI >0.50, interest only/negative amortization loans, no or low doc loans, 40-year terms, balloons, and loans with designations indicating they contain features that did not meet standard business requirements at the time of acquisition.
2. The control variables in the 24-month panel multinomial logit regressions include FICO, DTI, MTMLTV, the ratio of LTV and CLTV, loan age, actual UPB, rate incentive, number of borrowers, the first-time homebuyer indicator, subordinate financing indicators, product type, number of units, third party originator indicators, state fixed effects and acquisition date (measured by month) fixed effects.
3. *10%,**5%,***1% Significance Levels.

market LTV (MTMLTV), utilizing Fannie Mae's zip code level house price index. This incorporates cross-sectional variation in house price paths as a control for relative prepay and default behavior. In addition, for purposes of the prepayment function, we add a measure of the refinancing incentive, defined as the original note minus the time-varying market rate on 30-year or 15-year fixed rate loans, observed monthly. We also add time-varying variables such as unpaid principal balance and time since loan origination, measured in months.

Table 8 presents the results for the panel regression. The top panel shows the results when controlling only for origination CLTV and not MTMLTV. Here, the default odds ratio relative to the HI group is 1.49 in the early-boom period, 1.73 in the late-boom period and 1.88 in the post-crisis period. The pattern is similar for other LMI categories. We also confirm the finding in the existing literature that lower income households have lower prepayment risk (see for example Firestone et al., 2007). This finding supports our restriction of the performance window to 24-months to maintain comparability of predicted default rates across relative income groups. In particular, since higher income loans prepay faster than LMI loans, the higher income loans are more frequently removed from the potential default population. To the extent that some of these higher income prepayments represent refinances that could ultimately result in default, the default odds ratio of the LMI loans can be potentially overstated, with the risk of overstatement rising as the performance window increases.

In the second panel of Table 8, we add in a control for MTMLTV to account for changes in the home price over the 24-month period. In comparing the results with the top panel, the default risk to the LMI groups change only slightly. For instance, the VLI odds ratio drops to 1.46 in the early-boom period, to 1.68 in the late-boom period and to 1.87 in the post-crisis period. When combining the results from Tables 7 and 8, we see that adding zip-level regional effects explains part of the observed marginal riskiness of LMI borrowers but not all. Furthermore, controlling for relative differences in home prices as embodied in the MTMLTV does not change the estimated marginal risk of LMI borrowers.[31] This suggests that some other feature of the regional market, for instance labor market dynamics or delinquency/foreclosure contagion, is behind some of the difference in observed marginal risk for LMI borrowers.

We also estimate the multinomial logit model only on first-time homebuyers. The results of this exercise are shown in Table 9. The default odds ratios in this table indicate that the marginal risk of being an LMI borrower is lower in the late-boom and post-crisis period for the first-time homebuyers as opposed to the entire eligible population (including both first-time and repeat homebuyers). In particular, the relative odds of default for the VLI first time homebuyer is 1.53 in the late-boom period (versus 1.68 for all eligible loans, lower panel of Table 8) and 1.56 in the post-crisis period (versus 1.87 for all eligible loans). For the early-boom period, however, the relative marginal default risk of the VLI borrowers is higher for first-time homebuyers (1.63) compared with the entire population (1.46). The relative risk of prepayment for the LMI segment versus HI borrowers, on the other hand, is lower for first-time homebuyers versus repeat buyers in all periods. For instance, the VLI prepayment odds ratio is 0.73 in the early-boom period, compared with 0.77 for the entire population in the same period. In summary, our earlier documented conclusion based on our baseline specification of remaining marginal default risks for LMI borrowers, particularly VLI borrowers, also holds for first-time homebuyers across our three time periods.

In results not shown in the interest of brevity, we also find that our measure of marginal LMI risks are robust to removing loans with subordinate financing, interacting FICO and CLTV in the logit specification and when removing the time fixed effects. We have also compared the results of our sequential two-step regression framework, where we model the LMI component in a second step after controlling for the standard underwriting factors, to an approach where we simultaneously estimate loan performance on LMI and other explanatory factors. The estimates of marginal risks of LMI in this case were consistent with what we find in our baseline two-step estimation approach.

7. Conclusion

This research uses Fannie Mae's conventional conforming purchase mortgage loan-level acquisitions data to examine the credit risk associated with LMI borrowers and the extent to which tighter eligibility standards reduced the risk of extending mortgage credit to LMI homebuyers relative to the early-boom standards.

We first document the higher observed default rates for LMI borrowers relative to HI borrowers across three different time periods. For instance, the VLI loans originated during the late-boom period had a default rate of 12.22 percent, compared with a default rate of 4.03 percent for the corresponding HI loans (default odd ratio of 3.32). Second, we show that eligibility standards are extremely important for sustainable lending for borrowers across all relative income groups. In particular, after removing the historical loans from the late-boom period originations that do not meet today's eligibility standards, the default rate for the late-boom period drops to 2.34 percent for the VLI population and to 1.05 percent for the HI population (default odd ratio of 2.27).

We further show that even after controlling for current eligibility

[31] We find in our sample of eligible loans, that the LMI borrowers have a slightly higher average home price appreciation in the early-boom and post-crisis periods compared with higher income borrowers, and slightly less home price depreciation in the late-boom period. This is consistent with the recently observed faster increase in home prices and tighter inventories in the lower price tiers over the last few years (Source: CoreLogic).

restrictions, LMI loans typically have a riskier credit profile, for instance lower FICOs and higher DTIs along with a greater share of loans to one borrower. Much of the higher observed default rate for the LMI group can be explained by these and other standard underwriting factors. For instance, the default odds ratio for VLI loans relative to HI loans drops to 1.30 in the late-boom period once we control for both current eligibility restrictions and credit risk factors. The corresponding default odds ratio for the LI and MI populations are now insignificantly different from one in the late-boom period.

Thus, we find that standard credit variables are good predictors of default across most relative income groups. For the VLI segment, however, while these models account for a substantial portion of the default risk for this group, significant unexplained risks remain. In particular, VLI loans default ∼25 percent more frequently than would be expected given their observed credit profile. This error, however, is reasonably stable, ranging from ∼20 to ∼29 percent in the three historical time periods analyzed here.

Coming out of the financial crisis, LMI borrowers have encountered many obstacles when it comes to homeownership. These include slower wealth and income growth compared with higher income groups, faster home price appreciation, more restricted supply in lower price segments of the market and tighter underwriting standards in the post-crisis period. For market participants interested in extending credit to LMI borrowers, we show here that standard underwriting models do a good job of rank-ordering risks for most segments of the market and where the model does not perform as well (VLI borrowers), the error has been stable across different time periods. In order to fully address the issue of housing affordability facing today's LMI borrowers, however, also requires a careful consideration of the constraints facing affordable housing supply.

Acknowledgement

The authors thank Anthony Sanders for his in-depth review and invaluable feedback on an earlier version of this draft, as well as Neil Bhutta, Steve Laufer, Jonathan Lawless, Michael Lacour-Little, Raven Malloy, Carlos Perez, Spencer Perry, Daniel Ringo, Eric Rosenblatt, Ozge Oundee Savascin, and Yi Song for their helpful comments and suggestions. Any opinions expressed are those of the authors and not those of Fannie Mae.

Appendix

Tables A1–A3 present the complete set of estimated parameters for the logistic default models used to generate the primary results in the paper by time period for the sample of loans eligible for delivery to Fannie Mae under today's underwriting guidelines. We also present coefficient estimates for the entire population of loans (both eligible and ineligible) for the two earlier time periods in Table A4 along with a set of indicators capturing the ineligible loan features. The coefficients in these tables are estimated using a two-step procedure, as described in Section 5. The first step regresses default outcomes on the variables listed in the top panel of the table (either a constant term or a full set of standard underwriting risk factors). The second step effectively regresses the residuals from the first step on the relative income indicators, with the constant term suppressed at zero. Intuitively, the model in the first column of Tables A1–A3 first computes the mean default rate of the sampled loan population and then derives the deviation from the mean by income group. The model in the second column of Tables A1–A3 and both columns of Table A4 first predict default rate based on underwriting risk factors and then estimates the difference between the actual and predicted default risk by income group.

The primary variables of interest in this paper are the LMI indicators. As discussed in Section 5, accounting for a full set of standard underwriting factors dramatically reduces unexplained risk from the first step. For example, the coefficient for the VLI indicator falls from to 0.85 to 0.27 in the early-boom period, from 0.65 to 0.24 in the late-boom period and from 0.97 to 0.35 in the post-crisis periods. Note that the estimated coefficients on the four relative income indicators decline incrementally from a positive value for VLI to a negative value for HI. The positive value indicates that the first-step model underestimates the default risk of the corresponding relative income group and the negative value indicates an overestimation from the first step.

In general, the estimates for the standard underwriting characteristics shown in Tables A1–A3 are consistent with other works in the credit risk literature. Default risk increases with CLTV and DTI and decreases with FICO. The FICO splines used in the estimation model are defined based on the minimum of the borrower's and co-borrower's FICO scores. In the case where the two FICO scores are different, the difference in FICOs captures the creditworthiness of the higher scoring borrower. The coefficient of this variable is consistently estimated at −0.01 across all three periods, which means every additional point of FICO score for the higher scoring borrower reduces the default odds ratio by one percent.

The estimates for the categorical variables in the model are mostly robust over time and consistent with intuition. For example, one-borrower loans are riskier than multiple-borrower loans, as multiple borrower loans likely come with some diversification in the source of income. Thirty year fixed rate mortgages are riskier than their 15-year counterparts, indicating a potential clientele effect for the 15-year product. After controlling for the borrower equity position through CLTV, mortgages that have subordinate financing are less risky than those that do not, potentially indicating a more resourceful borrower able to avoid the mortgage insurance premiums typically associated with LTVs in excess of 80 percent. Loans originated through brokers or correspondents are riskier than loans originated by lenders.

The coefficient on the first-time homebuyer indicator is positive and small in both the early-boom and post-crisis periods, indicating increased risk, but is quite negative in the late-boom period. The change in sign across time periods arises partly because we have restricted our estimation set to only those loans eligible for delivery under today's guidelines. This removes the riskiest first-time homebuyers from the late-boom period.[32]

Given the focus of our analysis on the eligible population and sustainable LMI lending under today's underwriting standards, we do not present the marginal credit risk of ineligible LMI loans in the text of the paper. Table A4 presents the full set of loans (including the ineligible set) for the early-boom and late-boom periods. The non-LMI coefficients are generally robust to the inclusion of ineligible loans in the estimation set when compared with the corresponding values in Tables A1 and A2. The estimated marginal risk of the LMI loans declines slightly when adding in the ineligible loans and controlling for the ineligible loan features. For instance, VLI drops from 0.27 to 0.20 in the first period and from 0.24 to 0.17 in the second period. Table A4 also shows that, not surprisingly, most ineligible loan features increase the likelihood of default for a given loan, with the largest marginal effects coming from less than full documentation and FICO <620. On the other hand, limited amortization features, such as interest-only, negative amortization, 40-year term and balloon payments, marginally improve a loan's credit performance in the early-boom period. This improvement in marginal performance is likely driven by the lower monthly payments for these products in the first two years of a loan's life when compared with a

[32] As the results in Table A4 show, when we include all loans (both eligible and ineligible) in the late-boom period, the first-time homebuyer coefficient is still negative but much smaller in magnitude.

traditional 30-year amortizing product. In the looser underwriting environment and subsequent home price decline of the late-boom period, however, these limited amortization features increase the likelihood of default, potentially driven at least partially by clientele effects not controlled by standard underwriting factors or relative income.

Table A1
Default regression results: Eligible loans only, sample period 2002–2004.

Dependent Variable:		90 or more days delinquency in the first 24 months			
Sample Period:		**2002–2004 Originations**			
		(1)		(2)	
		Estimate	StdErr.	Estimate	StdErr.
		First Step			
Intercept		−4.82***	0.01	−3.22***	0.36
Baseline Controls (Linear Splines or Continuous)					
CLTV Splines	≤60			−0.02***	0.00
	>60 and ≤70			0.04***	0.01
	>70 and ≤80			0.03***	0.01
	>80 and ≤90			0.05***	0.00
	>90 and ≤150			0.08***	0.00
FICO Splines	>620 and ≤660			−0.02***	0.00
	>660 and ≤720			−0.02***	0.00
	>720 and ≤760			−0.03***	0.00
	>760 and ≤850			−0.00***	0.00
FICO Diff	Continuous			−0.01***	0.00
DTI Splines	≤0.36			1.04***	0.20
	>0.36 and ≤0.45			1.95***	0.34
Baseline Controls (Categorical)					
One Borrower				1.04***	0.03
First-time Homebuyer				0.03***	0.02
Subordinate Finance	≤10pct Subfin			−0.71***	0.07
(Excluded = No Subfin)	15pct Subfin			−0.74***	0.05
	≥20pct Subfin			−0.47***	0.07
	Other Subfin			−0.34**	0.14
Third Party Origination	Broker			0.40***	0.03
(Excluded = No TPO)	Correspondent			0.17***	0.02
FRM15				−0.42***	0.05
One Unit				−0.69***	0.05
Relative Income Controls (Categorical)		***Second Step***			
Income/AMI	**≤50%**	**0.85*****	**0.03**	**0.27*****	**0.03**
	>50% and ≤80%	**0.44*****	**0.02**	**0.06*****	**0.02**
	>80% and ≤100%	**0.12*****	**0.02**	**−0.01*****	**0.02**
	>100%	**−0.40*****	**0.02**	**−0.10*****	**0.02**
-2LogL		116,776		96,830	
AIC		116,785		96,838	
Number of Observations (1,000)		1279		1278	

*10%, **5%, ***1% Significance Levels.

Table A2
Default regression results: Eligible loans only, sample period 2005–2007.

Dependent Variable:		90 or more days delinquency in the first 24 months			
Sample Period:		**2005–2007 Originations**			
		(1)		(2)	
		Estimate	StdErr.	Estimate	StdErr.
		First Step			
Intercept		−4.38***	0.01	−4.87***	0.42
Baseline Controls (Linear Splines or Continuous)					
CLTV Splines	≤60			0.01**	0.01
	>60 and ≤70			0.06***	0.01
	>70 and ≤80			0.03***	0.01
	>80 and ≤90			0.06***	0.00
	>90 and ≤150			0.05***	0.01
FICO Splines	>620 and ≤660			−0.03***	0.00
	>660 and ≤720			−0.02***	0.00
	>720 and ≤760			−0.02***	0.00
	>760 and ≤850			−0.02***	0.00
FICO Diff	Continuous			−0.01***	0.00
DTI Splines	≤0.36			0.95***	0.25
	>0.36 and ≤0.45			3.68***	0.33
Baseline Controls (Categorical)					
One Borrower				0.78***	0.03
First-time Homebuyer				−0.17***	0.02
Subordinate Finance	≤10pct Subfin			−0.48***	0.05

(*continued on next column*)

Table A2 (*continued*)

Dependent Variable:		90 or more days delinquency in the first 24 months			
Sample Period:		**2005–2007 Originations**			
		(1)		(2)	
		Estimate	StdErr.	Estimate	StdErr.
(Excluded = No Subfin)	15pct Subfin			−0.43***	0.04
	≥20pct Subfin			−0.15***	0.05
	Other Subfin			−0.54***	0.09
Third Party Origination	Broker			0.37***	0.03
(Excluded = No TPO)	Correspondent			0.09***	0.02
FRM15				−0.61***	0.07
One Unit				−0.67***	0.06
Relative Income Controls (Categorical)		***Second Step***			
Income/AMI	**≤50%**	**0.65*** **	**0.04**	**0.24*** **	**0.04**
	>50% and ≤80%	**0.31*** **	**0.02**	**0.01*** **	**0.02**
	>80% and ≤100%	**0.08*** **	**0.03**	**−0.07** **	**0.03**
	>100%	**−0.17*** **	**0.01**	**−0.02** **	**0.01**
-2LogL		109,396		89,450	
AIC		109,405		89,458	
Number of Observations (1,000)		825		825	

*10%, **5%, ***1% Significance Levels.

Table A3
Default regression results: Eligible loans only, sample period 2011–2013.

Dependent Variable:		90 or more days delinquency in the first 24 months			
Sample Period:		**2011–2013 Originations**			
		(1)		(2)	
		Estimate	StdErr.	Estimate	StdErr.
		First Step			
Intercept		−5.91***	0.02	−6.11***	0.60
Baseline Controls (Linear Splines or Continuous)					
CLTV Splines	≤60			0.00***	0.01
	>60 and ≤70			0.02***	0.01
	>70 and ≤80			0.00***	0.01
	>80 and ≤90			0.03***	0.01
	>90 and ≤150			0.09***	0.01
FICO Splines	>620 and ≤660			−0.02***	0.00
	>660 and ≤720			−0.03***	0.00
	>720 and ≤760			−0.02***	0.00
	>760 and ≤850			−0.01***	0.00
FICO Diff	Continuous			−0.01***	0.00
DTI Splines	≤0.36			3.47***	0.44
	>0.36 and ≤0.45			3.82***	0.64
Baseline Controls (Categorical)					
One Borrower				0.92***	0.06
First-time Homebuyer				0.02***	0.04
Subordinate Finance	≤10pct Subfin			−0.31	0.20
(Excluded = No Subfin)	15pct Subfin			−0.34	0.25
	≥20pct Subfin			−0.34**	0.16
	Other Subfin			−0.31*	0.18
Third Party Origination	Broker			0.17***	0.07
(Excluded = No TPO)	Correspondent			0.16***	0.04
FRM15				−0.33***	0.08
One Unit				0.27	0.21
Relative Income Controls (Categorical)		***Second Step***			
Income/AMI	**≤50%**	**0.97*** **	**0.04**	**0.35*** **	**0.04**
	>50% and ≤80%	**0.36*** **	**0.04**	**−0.01*** **	**0.04**
	>80% and ≤100%	**0.15*** **	**0.05**	**0.01*** **	**0.05**
	>100%	**−0.38*** **	**0.03**	**−0.11*** **	**0.03**
-2LogL		41,836		36,640	
AIC		41,845		36,649	
Number of Observations (1,000)		1141		1141	

*10%, **5%, ***1% Significance Levels.

Table A4
Default regression results: All loans, sample periods 2002–2004 and 2005–2007.

Dependent Variable:		90 or more days delinquency in the first 24 months			
Sample Period:		**2002–2004**		**2005–2007**	
		Estimate	StdErr.	Estimate	StdErr.
		First Step			
Intercept		−4.17***	0.23	−4.92***	0.19
Baseline Controls (Linear Splines or Continuous)					
CLTV Splines	≤60	−0.01***	0.00	0.01***	0.00
	>60 and ≤70	0.04***	0.01	0.05***	0.01
	>70 and ≤80	0.03***	0.00	0.05***	0.00
	>80 and ≤90	0.05***	0.00	0.06***	0.00
	>90 and ≤150	0.06***	0.00	0.06***	0.00
FICO Splines	>620 and ≤660	−0.02***	0.00	−0.02***	0.00
	>660 and ≤720	−0.02***	0.00	−0.01***	0.00
	>720 and ≤760	−0.02***	0.00	−0.02***	0.00
	>760 and ≤850	−0.00**	0.00	−0.01***	0.00
FICO Diff	Continuous	−0.01***	0.00	−0.01***	0.00
DTI Splines	≤0.36	0.84***	0.12	1.01***	0.12
	>0.36 and ≤0.45	1.61***	0.20	2.69***	0.14
Baseline Controls (Categorical)					
One Borrower		0.81***	0.01	0.59***	0.01
First-time Homebuyer		0.03**	0.01	−0.02***	0.01
Subordinate Finance	≤10pct Subfin	−0.65***	0.04	−0.48***	0.03
(Excluded = No Subfin)	15pct Subfin	−0.67***	0.04	−0.32***	0.02
	≥20pct Subfin	−0.34***	0.04	−0.16***	0.02
	Other Subfin	−0.28***	0.08	−0.51***	0.04
Third Party Origination	Broker	0.27***	0.01	0.36***	0.01
(Excluded = No TPO)	Correspondent	0.11***	0.01	0.26***	0.01
FRM15		−0.54***	0.03	−0.70***	0.04
One Unit		−0.42***	0.03	−0.55***	0.03
LTV>97 Indicator		0.06***	0.02	0.21***	0.02
FICO<620 Indicator		0.56***	0.02	0.58***	0.01
No or Low Documentation		0.94***	0.04	1.28***	0.04
DTI>0.50 Indicator		0.11***	0.02	0.13***	0.01
Limited Amortization		−0.14**	0.06	0.37***	0.01
All Other Ineligible Loans		0.36***	0.02	0.31***	0.01
Relative Income Controls (Categorical)		***Second Step***			
Income/AMI	**≤50%**	**0.20*** **	**0.01**	**0.17*** **	**0.01**
	>50% and ≤80%	**0.02** **	**0.01**	**−0.00*** **	**0.01**
	>80% and ≤100%	**0.01** **	**0.01**	**−0.03*** **	**0.01**
	>100%	**−0.10*** **	**0.01**	**−0.04*** **	**0.01**
-2LogL		328,628		617,124	
AIC		328,636		617,133	
Number of Observations (1,000)		2177		1904	

*10%, **5%, ***1% Significance Levels.

References

Agarwal, S., Ambrose, B.W., Chomsisengphet, S., Sanders, A.B., 2012. Thy Neighbor's mortgage: does living in a subprime neighborhood affect One's probability of default? R. Estate Econ. 40 (1), 1–22.

An, X., Deng, Y., Rosenblatt, E., Yao, V.W., 2012. Model stability and the subprime mortgage crisis. J. R. Estate Finance Econ. 45 (3), 545–568.

Archer, W.R., Ling, D.C., McGill, G.A., 1996. The effect of income and collateral constraints on residential mortgage terminations. Reg. Sci. Urban Econ. 26 (3), 235–261.

Avery, R.B., Bostic, R.W., Calem, P.S., Canner, G.B., 1996. Credit risk, credit scoring, and the performance of home mortgages. Fed. Reserv. Bull. 82, 621.

Avery, R.B., Brevoort, K.P., 2015. The subprime crisis: is government housing policy to blame? Rev. Econ. Stat. 352–363.

Barakova, I., Calem, P.S., Wachter, S.M., 2014. Borrowing constraints during the housing bubble. J. Hous. Econ. 4–20.

Coslett, S.R., 1981. Efficient estimation of discrete-choice models. In: Manski, C.F., McFadden, D. (Eds.), Structural Analysis of Discrete Data and Econometric Applications. MIT-Press, Cambridge, Mass.

Crook, J.N., Edelman, D.B., Thomas, L.C., 2007. Recent developments in consumer credit risk assessment. Eur. J. Oper. Res. 183 (3), 1447–1465.

Deng, Y., Gabriel, S.A., 2006. Risk-based pricing and the enhancement of mortgage credit availability among underserved and higher credit-risk populations. J. Money Credit Bank. 38 (6), 1431–1460.

Deng, Y.H., Quigley, J.M., Van Order, R., 1996. Mortgage default and low down payment loans: the costs of public subsidy. Reg. Sci. Urban Econ. 263–285.

Deng, Y., Quigley, J.M., Van Order, R., 2000. Mortgage terminations, heterogeneity and the exercise of mortgage options. Econometrica 275–307.

Duca, J.V., Rosenthal, S.S., 1994. Borrowing constraints and access to owner-occupied housing. Reg. Sci. Urban Econ. 24 (3), 301–322.

Firestone, S., Van Order, R., Zorn, P., 2007. The performance of low-income and minority mortgages. R. Estate Econ. 35 (4), 479–504.

Goodman, L.S., Seidman, E., Zhu, J., 2015. VA loans outperform FHA loans—why? and what can we learn? J. Fixed Income 39–51.

Haughwout, A., Peach, R., Tracy, J., 2008. Juvenile delinquent mortgages: bad credit or bad economy? J. Urban Econ. 64 (2), 246–257.

Kelly, A., 2008. 'Skin in the game': zero down payment mortgage default. J. Hous. Res. 17 (2), 75–99.

Kochhar, R., Cilluffo, A., 2017. How Wealth Inequality Has Changed in the U.S. Since the Great Recession, by Race, Thenicity and Income. Pew Research Center.

Lam, K., Dunsky, R.M., Kelly, A., 2013. Impacts of down payment underwriting standards on loan performance — evidence from the GSEs and FHA portfolios. In: Office of Policy Analysis and Research, Federal Housing Finance Agency Working Paper.

Li, W., Goodman, L., 2015. Measuring Mortgage Credit Availability Using Ex-ante Probability of Default. Urban Institute, Washington, DC.

Mayer, C., Pence, K., Sherlund, S.M., 2009. The rise in mortgage defaults. J. Econ. Perspect. 27–50.

Mays, E., 2004. Credit Scoring for Risk Managers: the Handbook for Lenders. Thomson/South-Western.

Mays, E., 2001. Handbook of Credit Scoring. Global Professional Publishing, p. 382.

Painter, G., Gabriel, S., Myers, D., 2001. Race, immigrant status, and housing tenure choice. J. Urban Econ. 49 (1), 150–167.

Quercia, R.G., Stegman, M.A., Davis, W.R., Stein, E., 2002. Performance of community reinvestment loans: implications for secondary market purchases. In: Low-Income Homeownership: Examining the Unexamined Goal, pp. 348–374.

Quercia, R.G., Stegman, M.A., 1992. Residential mortgage default: a review of the literature. J. Hous. Res. 341.

Shambaugh, J., Nunn, R., Liu, P., Nantz, G., 2017. Thirteen Facts about Wage Growth. The Hamilton Project.

Van Order, R., Zorn, P., 2002. Performance of low-income and minority mortgages. In: Van Order, R., Zorn, P. (Eds.), Low-income Homeownership: Examining the Unexamined Goal, pp. 322–347.

Van Order, R., Zorn, P., 2003. Income, location and default: some implications for community lending. R. Estate Econ. 385–404.