Low Income Homeownership and the Role of State Subsidies:
A Comparative Analysis of Mortgage Outcomes

Stephanie Moulton[1]
The Ohio State University

Matthew Record[2]
The Ohio State University and San Jose State University

Erik Hembre[3]
University of Illinois, Chicago

April 17, 2018

---

[1] Corresponding Author, Associate Professor, John Glenn College of Public Affairs, The Ohio State University; moulton.23@osu.edu; 614-247-8161

[2] Assistant Professor, Department of Political Science, San Jose State University; matthew.record@sjsu.edu; 408-924-5346

[3] Assistant Professor, Department of Economics, University of Illinois Chicago; ehembre@uic.edu; 312-996-2266

Low Income Homeownership and the Role of State Housing Finance Agencies:
A Comparative Analysis of Loan Performance

Abstract

Between the late 1970s through 2013, state Housing Finance Agencies (HFAs) have financed nearly $300 billion in mortgages to first time homebuyers. Descriptive evidence indicates that these state subsidized programs help financially vulnerable households enter homeownership and retain their homes at higher rates than lower income households purchasing homes in the private market. However, no prior studies have empirically estimated this relationship or the underlying mechanisms. Reduced risk of default is likely due to a combination of more affordable loan terms, value-added service delivery practices, as well as selection into state programs. Using data on first-time homebuyer mortgages securitized by Fannie Mae between 2005 and 2014, we estimate the impact of HFA originations on loan performance. We construct a matched sample of about 500,000 borrowers with similar credit and loan characteristics at the time of origination, one-fifth of whom had loans originated through an HFA. We then estimate the competing hazard of default (or foreclosure) and prepayment. We find that HFA borrowers across our sample period are about 20 to 30 percent less likely to default or foreclose on their mortgages than otherwise similar non-HFA borrowers, with the greatest effects observed during the 2005-2007 mortgage boom. To better understand the mechanisms, we include a vector of time-varying service delivery practices implemented by state HFAs, as well as indicators for structural loan characteristics. These attributes explain three-fourths of the observed HFA effect on mortgage default.

JEL Codes: G21, R31, R38

Key Words: homeownership; mortgage default; affordable mortgages

I. Introduction

Through 2014, state housing finance agencies (HFAs) in the U.S. facilitated home purchases for more than 3.2 million households, the majority of whom were low-to-moderate income (LMI) first-time homebuyers with incomes below 115 percent of the area median (National Council of State Housing Agencies 2015). Unlike higher cost subprime mortgages originated during the housing boom, mortgages originated by state HFAs do not rely on risk based pricing to extend homeownership to LMI households. In fact, prior to 2010, a majority of the mortgages originated by HFAs were financed through tax-exempt mortgage revenue bonds (MRBs), allowing HFAs to charge below market interest rates to borrowers (Moulton and Quercia 2014). Further, HFA originated mortgages typically meet conforming loan standards and are often securitized by Fannie Mae, Freddie Mac, or Ginnie Mae.

If HFA originated mortgages meet conforming loan standards, what, if anything, is the added benefit of HFAs to the mortgage market? A series of studies by the GAO on HFA lending in the 1980s questioned the extent to which HFAs originate mortgages to borrowers who would not have otherwise purchased a home in the conventional market (GAO 1983; 1988). Since that time, additional statutory restrictions have been enacted to increase the targeting of HFA loans to borrowers with incomes below 115 percent of the area median.[4] Aside from targeting LMI first time homebuyers, HFAs tend to originate mortgages at lower costs with additional value added services such as homebuyer education and counseling and preventative servicing practices (Moulton and Quercia 2014). It is possible that these services lead to better loan performance for LMI first-time homebuyers. HFAs were granted an exemption from the Qualified Residential Mortgage (QRM) requirements in part because of the assertion that HFA originated mortgages had inherently lower risk due "government oversight, an important public purpose, strong underwriting, proactive servicing, and a proven track record of safe and sound performance" (NCSHA 2011).[5]

While there have been descriptive reports of HFA loan performance, there have been no rigorous studies comparing the loan performance of HFA originated mortgages to otherwise similar borrowers. This paper seeks to address this gap in the literature through an analysis of home purchase mortgages securitized by Fannie Mae between 2005 and 2014, including loans to LMI borrowers originated through HFA and non-HFA channels. Our paper seeks to inform two primary research questions. How do HFA borrowers compare to conventional market first-time homebuyers? Do HFA mortgages perform better than otherwise similar

---

[4] By federal statute, HFA loans financed with mortgage revenue bonds are limited to borrowers with incomes below 115 percent of area median, with some exceptions for homebuyers located in targeted geographies. HFA loans that are not financed through mortgage revenue bonds do not fall under these federal guidelines.

[5] https://www.ncsha.org/blog/ncsha-supports-bond-exemption-risk-retention-rule

private market mortgages?

We first descriptively compare the characteristics of borrowers purchasing homes through HFAs to other first-time, LMI homebuyers with mortgages securitized by Fannie Mae. We find that while HFA borrowers have similar credit scores and debt to income ratios to non-HFA first-time homebuyers, they have significantly lower incomes and higher loan to value ratios (LTVs). Relative to privately originated LMI mortgages, HFAs serve borrowers at the lower end of the income and wealth distribution. This is important in light of federal policy emphasis to reduce inequities in mortgage access by categories of income and wealth (FHFA 2016). HFAs may provide a mechanism for Fannie Mae to meet affordable housing goals without comprising credit quality.

To analyze loan performance, we construct a matched sample of HFA and non-HFA originations in the Fannie Mae database. As we demonstrate with our descriptive comparison, HFAs tend to serve borrowers who are different than non-HFA borrowers, and failing to account for this selection can bias our results. Using coarsened exact matching, we ensure that the two groups of borrowers are similar at the time of origination on an array of characteristics that prior research has found to be predictive of mortgage default, including credit score, income, LTV, presence of a co-borrower, year of origination and geography (three digit ZIP code). Our resulting matched sample includes about 500,000 borrowers, about one-fifth of whom had loans originated through an HFA.

We find that HFA loans have a lower risk of default and foreclosure than loans originated to otherwise similar LMI borrowers. In our base multinomial logistic model specification, the relative risk of default (vs. prepayment) is about 20 percent lower for HFA borrowers than it is for otherwise similar non-HFA borrowers and the relative risk of foreclosure (vs. prepayment) is about 30 percent lower for HFA borrowers. If part of the HFA effect on loan performance is due to more careful screening of borrowers, we would expect the effect to be greatest during the mortgage boom when conventional underwriting was more lax (2005-2007). Indeed, we find that the HFA effect is strongest prior to the housing bust and the overall tightening of the mortgage market. Over all time periods, we observe slower prepayment speeds for HFA borrowers relative to non-HFA first time homebuyers.

To further explore the HFA effect, we add controls for additional loan attributes that are likely endogenous to the HFA channel including third-party origination, the presence of subordinate financing and whether or not the loan has full documentation. We also merge in data on HFA service delivery practices from 2005 to 2013. These data are collected annually by the National Council of State Housing Agencies. For each year, we construct indicators for (a) whether or not the HFA originates loans internally versus through

private banks and mortgage companies; (b) whether or not the HFA services its own mortgages or contracts with another company that assumes servicing responsibilities, typically a private lender; and (c) whether or not the HFA offers homebuyer education and counseling to its homebuyers. We find that the inclusion of the endogenous loan characteristics and service delivery practices explains three-fourth of the HFA effect for default and about half of the HFA effect for foreclosure.

II. Background and Expectations

LMI Mortgage Lending and Institutional Factors that affect Default

While this study is specific to an analysis of HFA originated loans, it builds from a larger literature on mortgage lending for LMI borrowers, and in particular, the role of the institutional channel to offset risk. A substantial body of literature has investigated the default and prepayment risk of lending to LMI borrowers. A series of studies by Deng and colleagues find that lower income borrowers have higher rates of mortgage default and lower risk of prepayment than higher income borrowers (Deng et al.1996; 2000; Deng and Gabriel 2006). Other studies find that income is not a significant predictor of mortgage default when credit risk characteristics are taken into account (Van Order and Zorn 2003; Fout et al. 2017). In a recent study of Fannie Mae loans, Fout et al. (2017) found that credit risk factors can account for increased rates of default among low income households with incomes between 50 and 80 percent of area median. However, for very low income households (with incomes below 50 percent of area median), they find a direct income effect on mortgage default that cannot be accounted for with standard underwriting variables.

Aside from a direct income effect, LMI borrowers also tend to have lower levels of wealth and thus higher loan to value (LTV) ratios. Previous research has found the LTV ratio to be significant predictor of mortgage default, particularly when house prices fall (Kelly 2008; Guiso et al. 2009; Gerardi et al. 2009; Demyanyk and Hemert 2011; Lam et al. 2013). By definition, higher LTV borrowers are more likely to be in a negative equity situation when house prices fall, as they did during the recent housing bust cycle in the U.S. This circumstance places them at greater risk of foreclosure if they experience a shock that prevents them from making their regular mortgage payments. Some assert that simply having less "skin in the game" makes LMI borrowers with low downpayment loans more likely to default. In a study of FHA loans during the early 2000s, Kelly (2008) finds that borrowers who fund their own downpayment are less likely to default than those who rely on assistance programs or relatives. An analysis of GSE and FHA loans by Lam et al. (2013) finds differential effects of LTV based on the underlying credit risk of the borrowers. An increase in LTV for low credit score borrowers has a much greater impact on the predicted foreclosure rate than the same size increase in LTV for higher credit score borrowers.

While lower incomes and wealth may increase default risk among LMI borrowers, a growing body of research suggests that institutional characteristics, including origination channel and servicing practices, can significantly moderate default risk. For example, borrowers who receive their loans from banks are less likely to default than borrowers who receive their loans from third party mortgage brokers (Alexander et al., 2002; Ding et al., 2011; Jiang et al., 2014; Moulton, 2010; Stegman et al., 2007). In an analysis of CRA-type loans purchased through an affordable lending initiative, Ding, Quercia, Li and Ratcliffe (2011) found that the special program loans had lower rates of default than subprime mortgages, even after matching on borrower characteristics and controlling for structural attributes of the loans.

There are several institutional mechanisms inherent in particular lending channels that have been associated with reduce default risk for LMI borrowers. First, there may be variation in origination practices by lending channel and, in particular, variation in the screening processes used to select borrowers. While certain underwriting criteria can be observed and controlled for in models of default, there may be unobserved differences in the selection of borrowers based on local knowledge or information. Studies have found substantially higher rates of default found for loans with low or no documentation (LaCour-Little 2009; Jiang et al. 2014). In an analysis of HFA originated mortgages in Ohio, Ergungor and Moulton (2014) find that the distance between the borrower and the branch of the originating lender is a significant predictor of default and that this effect is stronger for borrowers with lower credit scores for whom soft information may be more important.

In addition, borrowers may self-select into particular channels based on unobserved factors that are correlated with default. For example, less motivated borrowers (who are more likely to default) may select out of an arduous or complicated screening process associated with particular affordable lending programs. Agarwal et al. (2014) find that a state mandated requirement for education to borrowers receiving risky loans in Illinois indeed reduces the overall default rate, but because borrowers strategically avoid taking risky loans to avoid the requirement rather than a treatment effect of education.

Second, there may be differences in servicing practices by channel that influence default risk post-origination. For example, some lenders engage in preventative servicing, contacting borrowers immediately when a payment is missed prior to severe default (Stegman et al., 2007). There may also be variation in the amount of trust and ongoing communication between lenders and borrowers that influences mortgage performance. Prior research on mortgage modifications finds that one of the primary challenges is getting borrowers at risk of foreclosure to contact their lender (Cutts and Merrill 2008). Specific interventions prior to purchase, such as homebuyer education and counseling, may increase the trust between the lender and borrower. In a study of HFA originated mortgages in Tennessee, Brown (2016) finds that while the

introduction of a homebuyer education and counseling requirement has no impact on the probability of default, it is associated with reduced risk of foreclosure among defaulted borrowers. The explanation for this finding is that borrowers participating in the homebuyer education and counseling are more likely to contact their lender when they experience a hardship and work out a resolution to prevent foreclosure.

Third, structural differences between channels may have a direct impact on the loan performance of LMI borrowers. For example, different structures for providing downpayment assistance may have differential impacts on the cumulative LTV associated with the mortgage and thus a borrower's incentive to default. Further, loan characteristics including the loan terms, interest rates and prepayment penalties that have been associated with default risk may vary systematically by channel. While these factors are often endogenous to the risk characteristics of the borrower, there may be systematic differences in the extent to which different institutional channels implement these mechanisms.

State HFAs and LMI Mortgage Lending

As of 2002, all U.S. states as well as Puerto Rico, the US Virgin Islands, and the District of Columbia have an HFA, and most state HFAs today administer a mortgage program for first-time homebuyers.[6] State HFAs began financing mortgages in the early 1970s through the sale of tax-exempt mortgage revenue bonds (MRBs), which HFAs would then pass along in the form of reduced interest rate mortgages to LMI first-time homebuyers, retaining a portion of the spread to help finance agency operations.[7] By statute, HFA mortgages financed with MRBs are reserved for first-time homebuyers with incomes below 115 percent of area median income, with exceptions for households living in targeted geographic areas. The drop in conventional interest rates following the Great Recession has led HFAs to utilize other non-MRB mortgage backed securities (MBS) to finance mortgages, resulting in interest rates on HFA loans that are comparable to prime conventional mortgage rates (Moulton and Quercia 2014).[8]

---

[6] Not all state HFAs currently administers a first time homebuyer program. This is particularly true of in recent years of HFAs that have been reliant on MRB funding to back originations. In 2013, several states did not originate any MRB backed loans including Alabama, Arizona, Colorado, DC, Idaho, Illinois, New Hampshire, Utah and Wisconsin (NCSHA, 2015). Of those, Arizona and Illinois did not report funding mortgages through other sources either, effectively meaning they did not operate a FTHB single-family origination program in that year.

[7] The spread amount (arbitrage) permitted by statute has changed over time, and is currently limited to 1.25% (Moulton and Quercia 2014). See the Appendix for additional background on HFA interest rates and lending patterns over time.

[8] As of a 2012 survey of state HFAs, nearly 40 percent reported selling non-MRB mortgage backed securities (MBS) directly into the market—a strategy that none of the HFAs reported using in 2006 (Moulton and Quercia 2014). In contrast to MRB financed mortgages that provide revenue to HFAs through the interest rate spread during the life of the loan, non-MRB MBS provide an up-front financial return when the pool of mortgages are sold directly into the MBS market.

As public and quasi-public entities, HFAs are exposed to more political pressures than private sector lenders and thus may have more incentives than private lenders to reduce default risk of the loans they originate.[9] There are several institutional mechanisms associated with HFA lending that may offset the default risk of these mortgages. Some of these mechanisms are direct practices that can be observed and measured; others are embedded in the structural environment and incentives for HFA lending participants (both borrowers and lenders).

First, HFA origination practices may incentivize more careful screening of borrowers. While some HFAs originate mortgages directly, others originate mortgages through networks of approved private lending institutions in their state including banks and mortgage brokers. Regardless of the initial source of the origination, all HFAs monitor and ultimately approve loans to be financed with HFA funds, adding an incentive for careful screening (Ergungor and Moulton 2014). HFAs may add overlays to conventional underwriting standards to directly offset risk. For example, 75 percent of HFAs reported having minimum credit score requirements in 2012, the majority of which were higher than those set by FHA or the GSEs (Moulton and Quercia 2014).

Second, HFAs often require the completion of additional services or paperwork that may directly and indirectly be associated with lower default rates for borrowers. For example, the majority of HFAs offer or require homebuyer education and counseling for at least a portion of their borrowers prior to purchase. A 2010 survey of HFAs found that while about one-third of HFAs required homebuyer education and counseling for all borrowers, more than 80 percent required at least a portion of their borrowers to participate in these services (Dylla and Caldwell-Taugtes 2012). While such requirements may directly reduce default risk (e.g., borrowers may be more likely to proactively reach out for help with their mortgage if they experience a financial shock), these requirements may also act as additional sorting mechanisms to identify more motivated borrowers who are less likely to default.

Third, HFA servicing practices post-purchase may affect the probability of borrower default, or in the case of default, the likelihood that a borrower will cure (vs. experience foreclosure). The majority of HFAs have centralized servicing where the HFA either services loans in-house, or they contract with a single Master-Servicer who purchases the mortgage servicing rights (Moulton and Quercia 2014). Centralized servicing may allow for increased agency monitoring of loan performance and servicing practices. Further, nearly half of HFAs in 2012 identified additional preventative servicing practices that were provided to HFA

---

[9] The origination of mortgages through MRBs or non-MRB MBS generates an independent revenue stream to the HFA that can help fund agency operations. Negative publicity about the program (e.g., through high rates of default) may be viewed as a threat to the independence of the revenue stream and accumulated capital reserves.

borrowers above and beyond what would typically be provided to borrowers in the private market, including early default counseling and assistance with loan modification or short-sale in the case of pending foreclosure (Moulton and Quercia 2014).

Finally, there may be structural differences in HFA originated mortgages that affect default risk. The lower interest rates on HFA loans reduces the monthly mortgage payment, increasing the likelihood that borrowers can afford their monthly payment and lowering default risk (Ding et al. 2011). Aside from lower interest rates, HFA programs increase mortgage affordability by providing downpayment assistance (DPA) to borrowers through grants, forgivable loans or secondary financing. Lack of money for downpayment has been identified as a primary barrier to home purchase (Fuster and Zafar 2016). The proportion of HFA mortgages with DPA has fluctuated over time; in 2006, approximately 50 percent of HFA loans were reported to have some form of DPA compared to 70 percent in 2012 (Moulton and Quercia).

The provision of DPA may be associated with increased risk of default if borrowers with less of their own funds invested in the purchase are more likely to default (Kelly 2008). However, it is difficult to disentangle this sort of strategic behavior from unobserved selection of borrowers with lower levels of wealth into DPA programs; borrowers with lower wealth may have less of a financial cushion to buffer financial shocks thus placing them at greater risk of default. Holding constant observable risk factors, borrowers who are approved for downpayment assistance programs may actually be at lower risk of default due to additional screening mechanisms that are required to obtain the assistance (Fout et al. 2017; Leventis 2014). In an analysis of FHFA regulated loans, Leventis (2014) finds that the relationship between second liens and default is sensitive to the inclusion of covariates and varies over time. After accounting for higher default that is due to higher loan to value ratios, the receipt of secondary financing is associated with lower risk of default.

III. Data and Methods

Data

The data for this analysis come from two primary sources. The first source is the population of 30 year fixed-rate, single family, one unit dwelling or condominium, conventional, owner occupied purchase loans originated between 2005 and 2014 that were acquired by Fannie Mae.[10] To more accurately reflect HFA

---

[10] The majority of HFA loans in the Fannie Mae database are 30 year fixed rate mortgages (FRMs), with a small proportion (about 5%) that are 15 or 20 year FRMs. We exclude those borrowers with mortgages that are not 30 year FRMs from our population for this analysis. We further limit the population to borrowers purchasing one-unit

borrower characteristics, we limit this population to loans originated to first-time homebuyers[11] with household incomes less than $200,000 per year.[12] These restrictions result in a population of 1,059,250 loans, of which 126,193 are originated through HFAs.

The Fannie Mae dataset includes details on each loan and borrower at the time of origination and data on loan performance. Loan data at the time of origination include the original loan balance and loan terms, the loan to value ratio, indicators for secondary financing, and originator type. Crucially, these data include an indicator for whether or not the loan was financed through an HFA. Borrower data at the time of origination include borrower and co-borrower's FICO scores, income and geographic indicators at the ZIP code level. Regarding loan performance, these data include the date of the first 60 or 90 day delinquency and the date of foreclosure or prepayment (if applicable).

Our second source of data is the National Council of State Housing Agencies (NCHSA) annual State HFA Factbook. The NCHSA Factbook compiles self-reported survey data collected from each of the 56 member agencies yearly regarding agency and program operations.[13] Included in the Factbook are details about the single family lending programs administered by state HFAs, such as volumes of originated loans, average characteristics of borrowers served, and specific lending practices. We code the Factbook data to identify HFA service delivery practices from 2005 through 2013.[14] Specifically, we code three practices that we expect may influence loan performance: whether or not the state HFA provides any homeownership education and counseling in a given year, whether or not the state HFA operated a direct lending program in a given year (versus originating all loans through participating lenders), and whether or not the state HFA provides the majority of its loan servicing in house in a given year (versus contracting out servicing to a Master Servicer or other private lenders).

We supplement these datasets with annual data on house prices from the Federal Housing Finance Agency

---

properties, representing 99% of the HFA observations in the Fannie Mae dataset. Finally, we exclude borrowers from Puerto Rico for data comparability.

[11] While the majority of HFA originated mortgages are to first-time homebuyers, we exclude the small subset of mortgages that was not originated to first time homebuyers (less than 10 percent of the total HFA originations), thus increasing comparability to the non-HFA first time homebuyer mortgages.

[12] The household income threshold of $200,000 is typically higher than the standard 115 percent of area median income threshold used for most HFA mortgages; however, we wanted to ensure the ability to identify an appropriate comparison sample for the small subset of HFA borrowers with incomes exceeding the 115 percent threshold (e.g., HFA loans to borrowers in targeted geographies).

[13] 50 U.S. States, District of Columbia, Puerto Rico, The U.S. Virgin Islands, the New York State Department of Housing and Community Renewal, the New York City Housing Development Corporation, and the New York State Housing Finance Agency which was once, but no longer is, treated as distinct from the State of New York Mortgage Agency (SONYMA).

[14] While we primarily rely on the Factbook data to identify servicing practices, we validate the information reported in the Factbook with the raw data from the Moulton and Quercia (2014) HFA survey.

at the ZIP code level (see Bogin et al. 2016), and county level unemployment rates (quarterly) from the Bureau of Labor Statistics. We include the monthly 30 year fixed rate mortgage (FRM) interest rate from Freddie Mac's Primary Mortgage Market Survey data and the annually adjusted consumer price index (CPI) from the Organization for Economic Co-operation and Development.

The Fannie dataset includes the universe of HFA originated mortgages that are securitized through Fannie Mae; however, this represents only a portion of HFA originated mortgages. The total volume of HFA originations and the proportion that are securitized by Fannie Mae varies substantially over time in response to changes in the macro-economy and mortgage environment. Figure 1 plots proportion of HFA originations securitized by Fannie Mae from 2004-2014, relative to the total number of HFA single family originations during the same period per data compiled from the State HFA Factbooks.[15] For reference, the graph also plots the proportion of FHA loans originated by HFAs during the same years.

The vertical bars indicate the total single family loan volume across HFAs in a given year. Overall, HFA single family loan volume spiked in 2007 to a high of nearly 130,000 originations in that year, and dropped to a low of about 40,000 originations in 2009. In 2014, total single family production by HFAs was estimated to be around 73,000 originations (including MRB and other single family loans, excluding refinancing). [16] The proportion of HFA mortgages that were securitized by Fannie Mae follows a similar trend, making up nearly one-third of HFA originations in 2007, and dropping to around 5 percent of HFA originations in 2011, and increasing to nearly 25 percent of originations by 2014. As might be expected, the FHA origination trend moves in the opposite direction to the proportion of mortgages securitized by Fannie Mae, and follows the growth in FHA lending to LMI borrowers post 2008 in the general mortgage market.[17]

In addition to variation in HFA volume and securitization patterns over time, there is significant variation in patterns of HFA loans by state. Particular state HFAs are more likely to securitize loans through Fannie Mae than others, this affecting the distribution of loans (by state) in our dataset.[18] Our weighted sample

---

[15] For the years 2011 through 2014, HFA "Total Loans" include other non-MRB single family loans in addition to MRB loans. Prior to 2010, the total includes only MRB loan volume, as data on other single family loans is not reported in the Factbook until 2011 (when non-MRB financing strategies began to grow). Not graphed in Figure 1 are the proportion of HFA mortgages securitized by Freddie Mac and those held in an HFA's portfolio.

[16] These volume numbers are derived from the HFA Annual Factbook, and likely undercount non-MRB HFA originations by HFAs. The Factbook data does not report non-MRB loans until 2011, and even then, it is not clear that all non-MRB HFA loans are being reported (e.g., in some recent years, the number of HFA loans in the Fannie database for a particular state exceeds the number reported in the Factbook).

[17] See Appendix Figure A2 for a comparison of FHA and conventional mortgages over time in the full population of LMI mortgage lending.

[18] See Appendix Table A1 for the distribution of HFA loans in the Fannie Mae database by state and year, relative to the total distribution of HFA loans by state and year.

regressions (with exact matching on geography, as described below) account for the dominance of particular states with HFA loans in the Fannie Mae database. This helps to ensure that the results are generalizable within the Fannie Mae population of HFA loans. However, the results of the analysis may be less generalizable to the population of HFA loans in the U.S. as a whole.

Sample Construction

It is likely that first-time homebuyers who select into HFA loans are different than non-HFA first time homebuyers. Failing to address these differences may bias the estimates of the HFA effect. To help account for this, we construct a matched sample of HFA and non-HFA borrowers in the Fannie dataset using coarsened exact matching (CEM). This procedure is conceptually similar to propensity score matching but uses ex ante definition of match parameters and therefore relies more heavily on theory for generating matches. This technique has been found to improve the balance, error, and efficiency compared to traditional propensity score matching methods (Iacus et al.2012). While CEM only matches on observable characteristics, any matching approach assumes that some of the unobservable characteristics that may affect selection into an HFA loan as well as loan outcomes are correlated with the observed characteristics used for matching.

Following the CEM procedure, variables to be used for matching are first "coarsened" into categories following user defined parameters. Observations are then assigned into strata that correspond to the exact combination of categories for the matching variables. Rather than selecting a single match for each treatment observation, we allow for multiple matches, as multiple treatment and comparison observations could be grouped in the same stratum. Weights are assigned by stratum to account for this, where each treatment observation (HFA borrower) is given a weight equal to "1' and comparison observations are weighted within the stratum to equal the total number of treatment observations within the stratum. Observations in stratum with *only* treatment or *only* comparison group observations are given a weight of "0". These CEM generated weights are then included in any descriptive or empirical analyses of the sample.

Six loan origination variables are used for matching. We required an exact match on the year of origination, whether or not there was a co-borrower on the loan application and the 3-digit ZIP code of the property being purchased.[19] Matching on three digit ZIP code helps account for spatial differences in the distribution of HFA loans that could otherwise influence observed outcomes. The spatial variation between HFA and

---

[19] In an alternative specification, we conducted the match using the 5-digit zip code. This process yielded substantially fewer matched observations. The match on 3-digit zip code was thus selected as the preferred sample for this analysis.

non-HFA loans in the Fannie database appears to be relatively minimal within a given MSA where HFA loans are originated (see Appendix for further discussion).

In addition to the exact matching criteria, observations are matched on coarsened groupings of FICO credit score, combined loan-to-value ratio and household monthly income.[20] The combination of exact and coarsened variables results in 28,600 unique strata used for matching.

Applying these CEM parameters results in a total matched sample of 484,699 first-time homebuyer loans, of which 95,461 are originated through HFAs. This represents about 76% of the 126,209 HFA first-time homebuyer loans. We drop observations with missing values for credit score, debt to income ratio, mortgage payment or geographic location (county) at the time of origination, resulting in a final sample size of 477,181 observations including 93,741 HFA loans.

Borrower and Loan Characteristics for the Unmatched Population

Table 1 provides a comparison of means for key variables used in our analysis before constructing the matched sample, first for the overall sample and then by origination cohort. The comparison of the unmatched sample demonstrates that HFA borrowers are significantly different than other first-time homebuyers on a number of characteristics. All means are statistically different at the $p<0.05$ unless noted in the table (with a † symbol). These differences inform our first research question about the types of borrowers served by HFAs, relative to first time homebuyers in the full Fannie Mae population of borrowers.

First, with regard to loan performance, HFA loans are less likely to be ever 60 or 90 days late within the first 24 months after origination and are significantly less likely to foreclose or prepay than non-HFA loans to first-time homebuyers. This difference is most pronounced during the mortgage boom period from 2005-2007, where nearly 15 percent of loans originated to non-HFA first time homebuyers experience a 90 day default, compared to only 8 percent of HFA loans to first-time homebuyers. However, by the final cohort observed in the sample (2012-2014), the HFA difference in loan performance reverses, where HFA loans appear slightly more likely to experience a 60 or 90 day delinquency, although the proportion of loans experiencing such outcomes in the first 24 months is very low (only about 1 percent of the sample). The observed differences between HFA and non-HFA loan performance in this table do not account for other

---

[20]For creating coarsened buckets, the FICO cut-points were set at scores of 580, 660, 720, 800 and 950, meaning that those observations with FICOs below 580 and above 950 were not eligible for the match. We set the CLTV cut-points at 50%, 80%, 97% and 106%. Finally monthly income cut-points were set at $0, $2,000, $3,500, $5,000 and $9,000 per month.

differences between borrowers that may be associated with better (or worse) loan outcomes.

In general, HFA borrowers have lower incomes, lower original loan balances, borrow more money relative to the value of their homes and have credit scores that are very slightly lower than the population of Fannie Mae first time borrowers. These trends continue across all periods studied, though the gap in average credit scores between HFA and non-HFA borrowers increases after 2008. The ratio of a borrower's income to the area median income (of the ZIP code) is also lower for HFA borrowers relative to non-borrowers (70 percent compared to 84 percent). HFA borrowers tend to be slightly younger than other Fannie Mae first-time homebuyers (32 years of age compared with 35), with about 30 percent having co-borrowers. The majority of borrowers (more than 80 percent) purchase single-unit homes, versus a single family property within a condominium structure.

We also compare means for loan characteristics that are likely endogenous to the HFA lending channel; that is, HFA loans are likely to systematically differ on these characteristics because of the structure of the HFA loan program. Given that HFAs often structure DPA in the form of a second lien, it is not surprising that a substantially higher proportion of HFA loans have sub-financing (community seconds) than do first time homebuyer loans in Fannie Mae's general population. This is particularly acute in the periods post the 2008 contraction of the mortgage market. Nearly 50 percent of HFA borrowers had community seconds during the 2012 to 2014 period, compared with only 3 percent of non-HFA first time homebuyers during the same period. HFA loans are also substantially less likely to be originated through broker channels across all time periods than non-HFA loans, a difference that is particular substantial during the boom years of 2005 to 2007. Finally, all HFA loans are full-documentation loans across all time periods, contrasted to only 86% of non-HFA first time homebuyer loans from 2005 to 2007. By the last period in the data (2012 to 2014), all loans in the sample population had full documentation.

Another endogenous loan characteristic is the mortgage interest rate. Although the credit and income of borrowers with HFA loans tend to be lower, the interest rates are also significantly lower, averaging 5.5 percent for HFA borrowers and 5.9 percent for non-HFA borrowers. This reflects a fundamental difference that HFA mortgage interest rates are not determined by risk based pricing but instead are the same for borrowers in a particular program at a particular point in time. This means that interest rates for HFA borrowers do not vary systematically with the risk of the borrower, as they do for non-HFA loans. Further, practices regarding mortgage insurance (MI) differ between HFA and non-HFA loans which impacts the endogeneity of interest rates. Mortgages originated with LTVs above 80 percent are typically required to include MI coverage to offset the risk of loss to investors should the loan default; the premium for this coverage is typically paid by borrowers as part of their monthly mortgage payments. For non-HFA

mortgages, mortgages originated with lower MI coverage carry higher interest rates. However, state HFAs may participate in risk sharing programs or other strategies to reduce the required MI coverage amounts for LMI mortgages, with no direct link to the interest rate charged on the loan.[21] The interest rate differential is thus a structural feature of HFA loans. We explore the effects of the interest rate differential in alternative model specifications.

Figure 2 graphs the interest rates for HFA and non-HFA loans to borrowers with incomes below 115 percent of area median income within a narrow band of credit scores (620 to 680) over our sample period. As demonstrated in Figure 2, the spread between the HFA and non-HFA rate is greatest during the pre-crisis years, when HFAs relied primarily on MRBs to finance loans and prime interest rates were higher than in the crisis period. For example, in 2006, the average interest rate on an HFA loans to borrowers with credit scores between 620 and 680 was 5.95 percent, compared to an interest rate of 6.93 for non-HFA borrowers in the same credit window. While the gap in interest rates has narrowed, it will likely widen again when interest rates rise again in the private market and MRBs regain traction as a financially viable strategy for HFAs to use to finance mortgages (Moulton and Quercia 2014).

Borrower and Loan Characteristics for the Matched Sample

The differences observed between HFA and non-HFA borrowers highlight the importance of our matched sample. In order to assess the accuracy of the matching process, we compare the differences in the means of each matching variable between the HFA and non-HFA originations at baseline. Table 2 reports the comparison of means for the matched sample. Ideally, the resulting sample will be completely balanced with no significant differences in baseline characteristics between groups. However, with large sample sizes, it is expected that small differences may be statistically significant, and thus it is also important to consider the economic significance of the difference. We also compare differences in means for other model variables that will be included as control variables in the regression analysis.

As reported in Table 2, the differences in means for two of the matching variables (income and combined loan to value ratio, or CLTV) are statistically significant; however, the magnitude of these differences is small and if anything would make it more difficult to detect an HFA effect (e.g., lower incomes and higher CLTVs are likely associated with a higher risk of default). Specifically, even after matching, HFA borrowers have slightly lower incomes by about $150, and have slightly higher CLTVs (95.8 percent

---

[21] In the Fannie Mae databased, loans originated by non-HFAs without MI coverage have systematically higher interest rates than those originated with MI coverage; no difference in interest rates is observed for HFA originations based on MI coverage amount.

compared to 95.0).

Overall, all borrowers in the matched sample have low incomes, high LTVs, and relatively high debt to income ratios. With regard to income, the average borrower in our sample has an income of about $4,200 per month. The majority of the loans in this sample meet the Federal Housing Finance Agency's (FHFA) standards for single-family affordable housing goals based on area median income (AMI). For example, in 2014, the FHFA set a goal for 7 percent of Fannie Mae home purchase mortgages to go to borrowers with incomes below 50 percent of AMI, and 23 percent of mortgages to go to borrowers with incomes below 80 percent of AMI (FHFA 2014). In the HFA matched sample used for this analysis, 21 percent of borrowers have incomes below 50 percent AMI, 50 percent have incomes between 50-80 percent AMI, 18 percent have incomes between 81 and 100 percent, and only 12 percent have incomes greater than 100 percent AMI. Thus, based on income alone, nearly three-fourths of the mortgages in the HFA matched sample help meet Fannie Mae's affordable housing goals. It is likely that a proportion of the remaining loans help meet affordable goals based on borrowers living in low income or minority geographic areas.

Methods

We estimate the loan performance of LMI first-time homebuyers in the Fannie Mae database, comparing the performance of HFA originated mortgages to otherwise similar non-HFA mortgages. As is standard in the mortgage literature, we estimate the competing risks of prepayment and default (or foreclosure) using a multinomial logit specification (Pennington-Cross and Chomsisengphet 2007; Ding et al. 2011). The data are constructed as a panel with each borrower-month constituting an observation. The multinomial logit specification restricts the sum of the probabilities of default, prepayment and remaining active on the loan in each period to one, therefore directly controlling for the competing risks. The likelihood function is as follows:

$$\ln L = \sum_{t} \sum_{i} \sum_{j} d_{ijt} \ln \pi \, (y_{it} = j)$$

$$\pi(y_{it} = j) = e^{\alpha z_{it}} \, / \, (1 + \sum_{k=1}^{2} e^{\alpha z_{it}}) \text{ for } j = 1,2$$

$$\pi(y_{it} = j) = 1 \, / \, (1 + \sum_{k=1}^{2} e^{\alpha z_{it}}) \text{ for } j = 0$$

where d_{ijt} is an indicator variable equal to 1 if outcome j occurs for loan i at time t and 0 otherwise, α are the coefficients to be estimated and z represents the explanatory variables. Our primary explanatory variable is an indicator for whether or not the loan was originated through an HFA. Other explanatory variables include those collected at the time of origination such as FICO credit score, the presence of a co-borrower, the difference in FICO between the borrower and co-borrower, monthly mortgage payment, debt to income ratio (monthly debt payments divided by income), borrower age, income, housing unit type and the three year average change in the house price index at the time of loan origination.[22] Time varying explanatory variables include the monthly mark-to mark combined loan to value (CLTV) ratio[23], the monthly 30 year conventional interest rate[24], annual inflation, the quarterly unemployment rate in the county and the time since origination (in months). All model specifications include observation year and state fixed effects. Standard errors are clustered by borrower. Summary statistics for all model variables as used in the multinomial logistic regression specification are included in Appendix Table A2. Observations are weighted with the Coarsened Exact Matching (CEM) weights, discussed below.

As an alternative specification, we add controls for structural loan characteristics, including the subordinate financing (separating between community second liens and other second liens), whether or not the loan was originated through a broker or correspondent (vs. bank) and whether or not the loan was underwritten with full documentation. We also include a vector of service delivery processes undertaken by state housing finance agencies in the given year of loan origination: homeownership education and counseling, direct originations, and direct servicing. Importantly, these practices vary over time within a state. The service delivery indicators allow us to better unpack the HFA effect, and are only coded "1" for HFA originations in states that adopt the particular practices in a given year. We include the HFA indicator as well as state dummy indicators in the service delivery regressions, similar to a difference in difference estimation of the change in service delivery practices within a state on loan performance for HFA borrowers, relative to the loan performance for non-HFA borrowers in the same state.

---

[22] We test for splines in several of the continuous variables to allow for nonlinearities, including: credit score (<620; 621-660; 661-720; 721-760; and >760); debt to income ratio (<36; 36.1-45%; <45%); and combined loan to value ratio (<60%; 60.1-70%; 70.1-80%; 80.1-90%; and >90%). Only DTI shows nonlinear trends and thus our primary specification only preserves the splines for DTI.

[23] Fannie Mae data includes the mark-to-market LTV for the first mortgage each month, and the combined LTV as of the time of origination. In our primary specifications, we estimate the mark-to-market combined LTV in each month by adding the amount of secondary financing at the time of origination (derived from the combined LTV) to the outstanding mortgage balance each month. In an alternative specification, we calculate the combined LTV excluding the balance on community second liens in the 6th+ year after origination, as many community second liens are forgiven after 5 years. The HFA effects are econometrically and statistically identical to this specification. Results are available from the authors upon request.

[24] Board of Governors of the Federal Reserve System (US), 30-Year Conventional Mortgage Rate©, retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/MORTG, October 3, 2017.

To explore the relationship between HFA interest rates and loan performance, we estimate an alternative specification that replaces the single HFA indicator with three HFA indicators corresponding to the interest rate of the HFA loan. We code each HFA loan as being 0.5 percent or more below market interest rate, 0.05 percent or more above market interest rate, or at within 0.05 percent of the market interest rate. As indicated in Table 2, there is substantial variation in the proportion of HFA loans originated at below market interest rates over time, with more than one-third below market rate for the 2005 to 2007 origination cohort, dropping to about 10 percent for the 2008 to 2011 cohort and less than 5 percent for the 2012 to 2014 cohort. There is also an increase in the proportion originated above the market interest rate over time.

While our primary specifications estimate the model for the entire sample period, we also estimate sub-sample regressions for three cohorts of originations corresponding to the pre-crisis (2004-2007), crisis (2008-2011), and post-crisis (2012-2014) periods.

IV. Results

Matched Sample Outcomes: Comparison of Means

The matching process helps to balance the sample of HFA and non-HFA borrowers on observable characteristics. We thus begin by comparing the unconditional (weighted) means for loan performance outcomes for the HFA and non-HFA matched sample. Table 3 reports differences in loan performance and survival time.

Overall, HFA borrowers have statistically significant lower rates of 60 and 90 day default or foreclosure than non-HFA borrowers through 2011, this being more economically significant in the first 24 months of origination than over the lifetime of the loan. Interestingly, across all periods, HFA borrowers are less likely to prepay their mortgages than non-HFA first time homebuyers; by the last period in the data (2012-2014), this rate of prepayment is half that of other LMI first time homebuyers (10 percent compared to more than 20 percent prepayment rates for non-HFA borrowers). This may indicate that HFA borrowers are less likely to refinance their loans when it may be "in the money" to do so, either because of lack of information, transaction costs associated with refinancing, or barriers presented by higher rates of subordinate financing. This is similar to prior research that finds that FHA borrowers have slower prepayment speeds than conventional borrowers (Deng and Gabriel 2002); however, in this case, both the HFA and non-HFA borrowers have conventional loans securitized by Fannie Mae. The average survival time of HFA borrowers differs from non-HFA borrowers: across all periods, HFA borrowers retain their

mortgage for three-quarters of a year longer than otherwise similar non-HFA borrowers.

Figure 3 graphs the unconditional cumulative 90-day default hazard for HFA and non-HFA borrowers, showing that the gap in the default hazard between HFA and non-HFA borrowers is persistent (and increases slightly) over time. This is important, as it suggests that the HFA effect is not simply the result of a temporary delay that deteriorates over time.

Multinomial Logit Results of Competing Risks

Table 4 presents the results of the multinomial logit regression that models that risk of 90 day default (model 1) or foreclosure (model 2) relative to prepayment.[25] Coefficients are exponentiated to allow for interpretation, as the estimated change in the risk of default relative to prepayment or remaining current (the relative risk ratio) for a one-unit change in the explanatory variable. Loan outcomes are tracked until the last period the sample was observed (October 1, 2016).

The HFA indicator demonstrates a negative association with the risk of default or foreclosure. Specifically, the risk of default is about 20 percent lower (relative risk ratio of 0.773) and the risk of foreclosure is about 30 percent lower (0.686) for first-time homebuyers with loans originated through HFAs. The relative risk of prepayment is about also 30 percent lower.

Other model covariates have the expected signs, where borrowers with higher incomes, higher credit scores, and co-borrowers are less likely to default or foreclose, and those with higher mortgage payments, debt to income ratios, and mark-to-market CLTVs are more likely to default or foreclose. Macroeconomic variables generally have the expected relationships, where an increase in the county unemployment rate is associated with an increase in default, and the three year average change in house prices at the ZIP code level is associated with reduced risk of default.

Table 5 reports the estimates for the models that include structural loan characteristics and service delivery attributes. Results are shown for all origination years. All other variables (not shown) included in the base specification (see Table 4) are also included in the regressions, in addition to state and year fixed effects. Note that service delivery variables are only available through 2013; thus, originations for 2014 are excluded from these regressions, resulting in a smaller sample size (N=419,161).

Models (1) and (3) add the structural loan characteristics, including the presence of sub-financing

---

[25] We also estimated a similar model predicting 60 day rather than 90 day default. The results are similar to the 90 day results and are available from the authors upon request.

(community second or other second lien), whether or not the loan was originated through a broker or correspondent channel and whether or not the loan had full documentation at the time of origination. As expected, mortgages originated through third party originators (broker or correspondent channels) have a higher risk of default or foreclosure, while those with full documentation have a lower probability of default or foreclosure and higher risk of prepayment. Interestingly, the sub-financing variables (community seconds and other second liens) are associated with a lower risk of default or foreclosure in this sample population. This finding contrasts with other studies that have found higher rates of default among loans with secondary financing (e.g. Demyanyk and Van Hemert 2011). However, this finding is consistent with recent studies using Fannie Mae or FHFA data that also control for CLTV, debt to income and cohort effects (Fout et al. 2017; Leventis 2014). Community second liens are categorized by Fannie Mae as those provided by public, nonprofit or employer sources (Fannie Mae 2016); a lower probability of default with this form of secondary financing may be due to low or reduced interest rates and/or forgiveness of loan principal after a period of time.

Models (2) and (4) add service delivery practices. Recall that these service delivery practices are only coded "1" for HFA originations in a particular state, as non-HFA originations in a given state would not be affected by the HFA's service delivery practices. Direct servicing of loans by the HFA in a given year is associated with reduction in the relative default risk of about 20 percent and reduced prepayment risk of 8 percent. The operation of a homeownership education and counseling program in a given year is associated with a 10 percent reduction in default or foreclosure risk, and increase in the risk of prepayment. This is consistent with prior studies that have found that homebuyer education and counseling increases the borrowers' exercise of the prepayment option when it is in the money to do so (Quercia and Spader 2008). Together, the endogenous covariates and HFA servicing practices explain about 75 percent of the lower risk of default for HFA originated mortgages and about 40 percent of the lower risk of foreclosure.

Table 6 presents the HFA estimates by origination cohort. Models (1) and (2) are estimates from the base specification, while models (3) and (4) are estimates from the specification with loan structure and servicing delivery practices as controls. The results indicate some deterioration in the HFA effect on loan default and foreclosure during our sample period. With either specification, the HFA effects are not statistically significant for the third origination cohort (2012-2014). However, it is important to keep in mind the low overall rate of default and foreclosure for this cohort which reduces the power to detect significant effects. With regard to prepayment, the HFA effects are robust over time, indicating a persistent reduced prepayment risk of about 30 percent across all origination cohorts.

Table 7 presents the estimates for specification that replaces the single HFA indicator with three indicators

corresponding to the interest rate on the HFA loan. These models also control for structural loan attributes and servicing characteristics. HFA loans originated with below market interest rates (defined as 0.5 percent or more below the prime interest rate) have significantly lower risk of default and foreclosure in all periods except the final origination cohort. HFA loans originated at market rate have lower risk of foreclosure, but do not have significantly lower risk of default. HFA loans originated above market rate are not significantly less likely to default or foreclose than non-HFA loans. However, they are still less likely to prepay than non-HFA loans, although the prepayment effect is greatest for HFA loans with below market interest rates as would be expected—borrowers with low interest rates at origination are less likely to benefit from refinancing for lower rates.

Robustness Tests

We estimate several alternative specifications to probe the robustness of our results. First, rather than treating the outcome as the competing risk of default or prepayment using MNL panel estimation, we model only the hazard of default using a Cox proportional hazard model without time varying covariates (See Appendix Table A3). The HFA effect estimated through the hazard model is nearly identical to the estimates from the MNL specification; HFA originations are associated with a 25 percent reduction in the risk of default in the base specification, but only a 7 percent reduction in default risk after including origination and service delivery variables. The size and significance of the origination and service delivery variables are also similar to the estimates from the MNL specification.

As a second alternative specification, we limit our sample to those loans that are not originated through brokers. Recall that in our pre-match comparison (Table 1), non-HFA borrowers were significantly more likely to have broker originated mortgages than HFA borrowers—a difference that remained after our CEM process (Table 3). This difference is in a sense endogenous, as HFA loans are more likely to be originated through banks and other conventional lenders; however, it is important to understand the proportion of the HFA effect that may be explained by the broker difference. Further, simply controlling for broker loans in the model may miss unobserved differences that are associated with the types of borrowers served by brokers that could thereby bias our other model coefficients. To account for this, we drop broker originations from our full sample and re-run the CEM procedure excluding broker originated mortgages. This exclusion results in a matched sample of 218,456 loans, of which 92,247 are HFA originations. We re-estimate our MNL specification using the new CEM weights. Our results are nearly identical to our original specification. In the base model not controlling for structural loan characteristics or servicing practices, HFA originations are associated with a 20 percent reduction in the risk of default (0.809 relative risk ratio) and a 30 percent reduction in the risk of prepayment (0.695 relative risk ratio).

As a third robustness test, we limit our sample to loans that have information on whether or not the borrower received downpayment assistant (DPA) from a third party. Unfortunately, this variable is only available in the Fannie Mae database for about one-fourth of the observations, and thus we do not include it in our primary specifications. We rerun the CEM procedure excluding observations missing the information about whether or not they have DPA, and then re-estimate our MNL models controlling for DPA. Of the 188,850 observations with downpayment information, about 36 percent of HFA borrowers receive DPA, compared to 22 percent of non-HFA borrowers. The results indicate that the receipt of DPA is associated with slight 5 percent reduction in the risk of default or foreclosure, this reducing the HFA effect slightly by about 3 percent.

V. Conclusions

Our analysis indicates that HFAs serve primarily low-income first-time homebuyers with higher loan to value ratios than the typical first-time homebuyer in the Fannie Mae portfolio. Our empirical results suggest HFAs do so with less risk of mortgage default, particularly during the boom and bust periods from 2005 to 2011. We find that three-fourths of this effect on reduced default and nearly half of the effect for reduced foreclosure are related to observed HFA origination and service delivery practices. Lower risks of prepayment among HFA originated mortgages are persistent across all origination cohorts from 2005 through 2014 and are not well explained by variation in observed origination or service delivery practices.

The findings from this study contribute to an understanding of different ways to reduce the risk of lending to LMI first-time homebuyers, beyond underwriting. As demonstrated by the recent mortgage crisis, careful underwriting clearly has an important role to play in reducing default risk. Since 2009, default rates for both HFA and non-HFA mortgage borrowers have dropped substantially. In our sample of Fannie Mae first-time homebuyers, fewer than 1 percent of loans originated between 2012 and 2014 ever experiencing a 90 day default, relative to 20 percent of loans originated between 2005 and 2007. The average credit scores for first time homebuyers during this same period increased by nearly 50 points, with an average score of about 700 for the 2005 to 2007 origination cohort, to an average score of nearly 750 for the 2012 to 2014 origination cohort. Similarly, the average CLTV declined from 93 percent to 87 percent during our study period. However, tightening credit score and LTV requirements too much may make it difficult for otherwise creditworthy LMI first-time homebuyers to enter the market.

Unpacking the HFA effect on loan performance can offer insights for alternative strategies to reduce the default risk of lending to this population of borrowers. Part of the observed HFA effect could be related to more careful screening of borrowers. Recall that both HFA and non-HFA borrowers in our sample met

Fannie Mae conforming loan guidelines, and thus differences due to screening would be due to the use of additional information in the underwriting process or more careful monitoring of loan originators (to ensure accuracy of provided information). The fact that all HFA loans in our study required full documentation and that full documentation is associated with reduced default supports the careful screening mechanism. If the HFA effect were driven entirely by screening, we might expect to find that HFAs that directly originate their own mortgages (versus relying on private mortgage originators) would have borrowers with significantly lower rates of default. However, the indicator for direct lending by HFAs is associated with a slight increase in the risk of default, perhaps due to willingness to serve relatively riskier borrowers when lending directly.

Aside from direct originations, it could be that having a third-party HFA monitoring the origination practices of private lenders leads to more careful screening by the lenders. Given claims of lax screening in the private market during the boom period, it would make sense that HFAs (with more careful screening of loan originations) would have lower rates of default during this period. Loans originated by broker and correspondent channels are consistently associated with higher risk of default in our study, adding further support to the importance of monitoring.

Another aspect of HFA lending that may differ from the private market is servicing practices. Research suggests that preventative servicing practices, such as ongoing monitoring of borrower payments and early contact for delinquencies, can contribute to reduced rates of mortgage default (Stegman et al. 2007). On a recent survey of state HFAs, many HFAs indicated in engaging in at least some type of preventative servicing for their pool of borrowers—even more so if they directly serviced their own loans (Moulton and Quercia 2014). In this study, we find that first-time homebuyers with loans originated by state HFAs that service the majority of their own loans have a reduced probability of default. To the extent that these HFAs engage in preventative servicing practices, these practices may directly contribute to the observed HFA loan performance effect. If non-HFA mortgage servicers begin to adopt preventative servicing practices as part of their standard operating procedures, the gap in performance between HFA and non-HFA originations may be reduced.

A third component of HFA lending examined in this study is the provision of homebuyer education and counseling. Unfortunately, we cannot directly measure whether or not an individual borrower completed homebuyer education and counseling—we can only observe if the borrower received a loan from a state HFA that provided such services when the loan was originated. In our study, we find a correlation between the presence of a homebuyer education and counseling program and reduced default during the boom period, in line with prior studies of HFA homeownership programs (Brown 2016). However, this effect

reverses for originations during the bust period (2008-2011) perhaps driven by the endogenous choice of HFAs to adopt homebuyer education and counseling programs in response to rising default rates. Our data is not well suited to estimate a treatment effect of homebuyer education and counseling.

Another explanation for the HFA effect is borrower selection. It is likely that there are additional transaction costs associated with HFA loans; even without the transaction cost of completing HFA loan requirements, adding another entity into the loan review process can extend the timeframe for loan closing (and increase paperwork). First-time homebuyers who are willing to persist through a more complex process may also be more motivated to make payments. Our matching strategy helps reduce observed differences that lead to selection, but we cannot fully account for unobserved differences such as borrower motivation, that may contribute to the observed HFA effect. It is likely that a portion of the HFA effect that we cannot explain with our model covariates is due to unobserved borrower selection.

Structural differences between HFA and non-HFA mortgages also contribute to lower rates of default. By design, we held constant the type of mortgage included in this study to be 30 year conventional fixed rate mortgages. Had we included other mortgage types, it is likely that the HFA differential with loan performance would be even larger, as the majority of HFA originations are 30 year conforming loans. Further, all HFA loans have full documentation and a smaller proportion are originated through broker channels; attributes which contribute to (but do not fully explain) the HFA effect. HFA originated mortgages also have systematically lower interest rates than non-HFA originated mortgages. On average, the interest rate differential is about 0.4 percentage points; in a regression adjusted estimate controlling for credit score, income, debt, CLTV, year and state, we estimate this differential to be 0.44 percentage points. HFA originated mortgages are also much more likely to have subordinate financing and third-party funded DPA—both of which have been attributed with higher rates of mortgage default in other studies. In this study, we observe a negative relationship between subordinate financing in the form of community seconds or other second liens and mortgage default. After accounting for higher CLTV ratios, subordinate financing could indicate something unobserved about borrowers that makes them less risky (Leventis 2014).

Taken together, our results suggest that the reduced mortgage default of HFA originations is likely due to a bundle of observed and unobserved service delivery practices, loan characteristics, and borrower selection. While these results are informative, there are limitations. Most notably, our results lack generalizability for HFA loans that are not securitized by Fannie Mae. Ideally, we would have data on the universe of HFA loans, this allowing us to model selection into the HFA channel. Our CEM strategy helps to increase the comparability between HFA and non-HFA originations within the Fannie Mae database on observed characteristics; however, there may still be unobserved characteristics that are associated with

selection into the HFA channel and mortgage default. Future research with more complete data on the universe of HFA originations is needed to help unpack the selection effects.

Table 1: Borrower and Loan Characteristics for the Unmatched Population of First Time Homebuyers

	All Years		2005-2007		2008-2011		2012-2014	
	Non-HFA	HFA	Non-HFA	HFA	Non-HFA	HFA	Non-HFA	HFA
Ever 60 days late in 24 months	11.4%	7.4%	17.9%	11.3%	7.8%	6.7%	0.9%	1.3%
Ever 90 days late in 24 months	9.2%	5.4%	14.5%	8.3%	6.4%	5.0%	0.5%	0.8%
Loan Foreclosed by 10/1/2016	13.5%	9.5%	22.8%	16.4%	5.7%†%	5.7%†	0.07%†	0.07%†
Loan Prepaid by 10/1/2016	49.9%	39.3%	56.6%	52.5%	67.9%	50.9%	23.7%	9.2%
Borrower Characteristics								
Income	5,079	4,120	5,027	3,886	5,130	4,325	5,155	4,447
Income as % AMI	83.8%	69.0%	84.2%	67.1%	82.7%	69.4%	83.9%	72.6%
Original Loan Balance	179,781	145,101	175,329	138,875	190,075	154,157	182,204	154,029
Mortgage Payment	1,006	817	1,122	827	1,110	885	897	782
Frontend Ratio	21.5%	20.5%	23.1%	22.0%	22.3%†	21.1%†	17.8%	18.0%
DTI	39.6%	38.8%	42.4%	41.2%	39.2%†	38.8%†	34.3%	35.2%
CLTV	89.8	95.6	93.0	95.9	84.7	94.4	86.7	96.5
LTV	87.6	93.5	89.5	94.6	83.9	90.9	86.4	93.2
Age	34.9	32.2	34.8	31.9	35.8	32.6	34.4	32.4
Borrower FICO	720	717	698	702	740	725	749	736
Borrower - Co-borrower FICO	7.5	9.9	7.0	9.4	6.0	10.4	9.6	10.8
Co-Borrower	27.1%	29.2%	26.1%	27.8%	25.1%	31.0%	30.4%†	30.6%†
Single Unit	82.0%	86.4%	83.5%	88.0%	78.7%	81.6%	81.0%	85.6%
Δ3-Year Unemployment (MSA)	0.4%	0.2%	0.8%	0.7%	1.1%	0.8%	-0.8%	-0.8%
Δ 3-Year HPI (Zip)	0.2	0.8	-0.2†	-0.2†	-5.7	-4.2	4.9	4.1
Broker Originator	21.2%	4.1%	28.0%	5.4%	19.0%	3.6%	8.9%	2.3%
Correspondent Originator	36.6%	53.6%	38.1%	26.5%	35.8%	64.9%	34.2%	94.3%
Community Second	4.0%	24.8%	6.1%	12.6%	1.8%	24.2%	1.3%	47.4%
Other Second Lien	8.7%	1.0%	14.0%	1.5%	3.5%	0.8%	1.6%	0.1%
Full Documentation	91.9%	100.0%	86.4%	100.0%	96.0%	100.0%	100.0%	100.0%
Interest Rate	5.9%	5.5%	6.7%	6.0%	5.8%	5.6%	4.3%	4.5%
N	933,057	126,193	502,503	65,223	172,700	18,358	249,934	38,773

Dollar values constant 2014 dollars. Means differences between HFA and non-HFA observations are statistically significant at the 5% level unless otherwise noted.

†HFA and non-HFA observations are not statistically different from one another at the 5% level

Table 2: Comparison of Means, Matched Sample

	All Years			2005-2007			2008-2011			2012-2014		
	Non-HFA	HFA		Non-HFA	HFA		Non-HFA	HFA		Non-HFA	HFA	
Variables used for Matching												
Monthly Income	4,273	4,120	**	4,020	3,881	**	4,577	4,406	**	4,707	4,537	**
CLTV	95.0	95.8	**	96.3	96.7	**	92.5	93.7	**	93.3	94.9	**
FICO	714	714		699	701	**	729	727	**	741	739	**
Co-borrower	26.1%	26.1%		25.7%	25.7%		26.0%	26.0%		27.2%	27.2%	
Other Independent Variables												
Income as % AMI	70.9%	68.2%	**	69.1%	66.5%	**	72.4%	69.6%	**	74.3%	71.6%	
Original Loan Balance	152,508	149,582	**	142,659	141,572	**	168,347	161,379	**	167,251	162,836	**
Mortgage Payment	903	849	**	918	844	**	985	930	**	828	822	
Frontend Ratio	21.9%	21.1%	**	23.6%	22.2%	**	22.1%	21.5%	**	17.9%	18.4%	**
Age	33.3	31.8	**	33.3	31.6		33.5	32.3		33.1	32.1	
Borrower-Co-borrower FICO	7.7	9.4	**	7.5	9.3	**	6.7	9.0	**	8.7	9.9	**
DTI	41.2%	39.7%	**	44.0%	41.8%	**	40.3%	39.3%	**	35.3%	35.2%	
Single Unit	85.0%	85.0%		87.9%	87.6%		80.1%	78.1%		80.9%	82.8%	**
Δ Three Year Unemployment (MSA)	36.0%	36.1%		73.6%	73.4%		90.5%	89.9%		-74.8%	-75.1%	
Δ Three Year HPI (Zip)	26.2%	26.4%		-58.0%	-58.5%		-421.9%	-421.0%		405.2%	407.0%	
Broker Originator	21.9%	3.4%	**	28.6%	4.0%	**	19.0%	3.9%	**	7.8%	1.7%	**
Correspondent Originator	36.2%	49.4%	**	37.5%	26.1%	**	33.7%	64.7%	**	34.3%	96.4%	**
Community Second	9.1%	16.6%	**	8.1%	13.2%	**	5.7%	19.0%	**	13.5%	23.7%	**
Other Second Lien	7.8%	1.0%	**	10.7%	1.5%	**	5.0%	0.8%	**	2.2%	0.1%	**
Full Documentation	95.7%	100.0%	**	93.4%	100.0%	**	98.3%	100.0%	**	100.0%	100.0%	
Interest Rate	6.0%	5.5%	**	6.7%	6.0%	**	5.9%	5.7%	**			
Direct Lending		30.5%			34.4%			42.6%			16.4%	
Direct Servicing		25.0%			15.0%			44.1%			39.6%	
Homeownership Counseling		59.0%			71.4%			67.3%			28.9%	
.5% + Below Market Rate		21.6%			31.5%			9.8%			4.2%	
At Market Rate		62.8%			66.9%			70.2%			49.7%	
.5% + Above Market Rate		15.6%			1.6%			20.0%			46.1%	
N	383,440	93,741		232,897	56,307		44,113	12,143		105,221	24,206	

Dollar values in 2014 constant dollars. Weighted means regression test for significant differences; ** p< 0.05

Table 3: Comparison of Outcomes, Matched Sample

	All Years			2005-2007			2008-2011			2012-2014		
	Non-HFA	HFA		Non-HFA	HFA		Non-HFA	HFA		Non-HFA	HFA	
Ever 60 days late within 24 months	11.4%	8.3%	***	16.5%	11.8%	***	9.0%	6.8%	***	1.05%	1.10%	
Ever 90 days late within 24 months	9.1%	6.1%	***	13.2%	8.7%	***	7.6%	5.2%	***	0.60%	0.64%	
Ever 60 days late	24.6%	22.3%	***	36.3%	32.7%	***	16.9%	16.4%		1.81%	1.41%	**
Ever 90 days late	22.6%	20.1%	***	33.7%	29.7%	***	15.3%	14.1%		1.01%	0.86%	
Loan Foreclosed by 10/1/2016	14.1%	11.2%	***	21.6%	17.1%	***	7.9%	6.0%	***	0.08%	0.05%	
Loan Prepaid by 10/1/2016	49.1%	41.9%	***	56.9%	52.2%	***	64.5%	55.2%	***	22.3%	10.3%	***
Survival time (months)	48.7	57.4	***	55.2	67.5	***	49.5	57.0	***	32.3	32.7	***
N	383,440	93,741		232,846	56,307		44,112	12,143		105,219	24,205	

Weighted means regression test for significant differences
* p< 0.05

Table 4: Competing Risk of 90 Day Default , Foreclosure or Prepayment

| | Default vs. Prepay | | Foreclose vs. Prepay | |
| | (1) | | (2) | |
	Default	Prepay	Foreclose	Prepay
HFA Flag	0.773***	0.695***	0.686***	0.731***
	(0.009)	(0.006)	(0.010)	(0.006)
Monthly Income (thousands)	0.838***	1.080***	0.903***	1.109***
	(0.007)	(0.006)	(0.010)	(0.006)
Mortgage Payment (hundreds)	1.099***	1.063***	1.055***	1.043***
	(0.003)	(0.002)	(0.003)	(0.002)
DTI<36 (spline)	1.005***	1.002	1.009***	1.003**
	(0.002)	(0.001)	(0.002)	(0.001)
DTI 36.1-45 (spline)	1.006***	1.003***	1.009***	1.004***
	(0.002)	(0.001)	(0.002)	(0.001)
DTI>45 (spline)	1.006***	1.003***	1.008***	1.003***
	(0.001)	(0.001)	(0.001)	(0.001)
Borrower FICO Score at Origination	0.989***	1.003***	0.993***	1.005***
	(0.000)	(0.000)	(0.000)	(0.000)
Co-borrower	0.935***	1.079***	0.933**	1.081***
	(0.022)	(0.018)	(0.031)	(0.017)
Borrower FICO – Co-Borrower FICO	0.999**	1.000	1.000	1.000
	(0.000)	(0.000)	(0.001)	(0.000)
Borrower Age	1.006***	0.985***	0.998**	0.984***
	(0.001)	(0.001)	(0.001)	(0.001)
Borrower Age (missing)	0.859***	0.983	0.855***	0.994
	(0.027)	(0.021)	(0.034)	(0.021)
Single Unit	1.029	1.130***	0.871***	1.078***
	(0.021)	(0.016)	(0.021)	(0.015)
Δ Three Year HPI (Zip)	0.991***	1.017***	1.002	1.016***
	(0.002)	(0.001)	(0.002)	(0.001)
Mark-to-Market CLTV	1.008***	0.989***	1.008***	0.986***
	(0.000)	(0.000)	(0.000)	(0.000)
30-Year FRM Interest Rate	1.246***	0.558***	1.028	0.538***
	(0.027)	(0.008)	(0.026)	(0.007)
CPI (Inflation)	0.918***	0.924***	1.015*	0.931***
	(0.004)	(0.004)	(0.009)	(0.004)
Unemployment Rate	1.058***	0.921***	1.041***	0.925***
	(0.004)	(0.003)	(0.006)	(0.003)
Months Since Origination	1.038***	1.027***	1.066***	1.022***
	(0.002)	(0.001)	(0.003)	(0.001)
Months Since Origination (squared)	1.000***	1.000***	1.000***	1.000***
	(0.000)	(0.000)	(0.000)	(0.000)
Constant	0.041***	0.006***	0.000***	0.004***
	(0.012)	(0.001)	(0.000)	(0.000)
Observations	19,345,412	19,345,412	22,885,452	22,885,452
Unique Borrowers	477,181	477,181	477,181	477,181

Estimates from multinomial logit panel regression. Clustered standard errors by borrower in parentheses. Coefficients are exponentiated and represent the relative risk ratio. All models include state and year fixed effects

*** p<0.01, ** p<0.05, * p<0.1

Table 5: Competing Risks, Service Delivery Variables

	(1)		(2)		(3)		(4)	
	Default	Prepay	Default	Prepay	Foreclose	Prepay	Foreclose	Prepay
HFA Flag	0.865***	0.693***	0.944***	0.671***	0.764***	0.718***	0.810***	0.689***
	(0.012)	(0.007)	(0.021)	(0.010)	(0.014)	(0.007)	(0.023)	(0.011)
Broker Originator	1.289***	0.987	1.295***	0.984	1.277***	0.938***	1.288***	0.934***
	(0.024)	(0.016)	(0.024)	(0.016)	(0.031)	(0.014)	(0.031)	(0.014)
Correspondent Originator	1.159***	1.015	1.166***	1.010	1.130***	0.985	1.140***	0.978*
	(0.021)	(0.012)	(0.021)	(0.012)	(0.026)	(0.012)	(0.027)	(0.012)
Community Second	0.854***	1.028	0.857***	1.038*	0.786***	1.074***	0.782***	1.083***
	(0.028)	(0.020)	(0.029)	(0.020)	(0.029)	(0.020)	(0.030)	(0.021)
Other Second Lien	0.947**	1.101***	0.949**	1.096***	0.851***	1.122***	0.853***	1.119***
	(0.024)	(0.020)	(0.024)	(0.020)	(0.026)	(0.020)	(0.026)	(0.020)
Full Documentation	0.648***	1.109***	0.648***	1.110***	0.783***	1.187***	0.781***	1.189***
	(0.020)	(0.025)	(0.020)	(0.025)	(0.029)	(0.027)	(0.029)	(0.027)
Direct Lending			1.027	1.072***			1.045	1.080***
			(0.024)	(0.017)			(0.030)	(0.017)
Direct Servicing			0.810***	0.923***			1.014	0.945***
			(0.027)	(0.017)			(0.045)	(0.017)
Homeownership Counseling			0.895***	1.034**			0.896***	1.036**
			(0.019)	(0.016)			(0.024)	(0.015)
Borrower-Year Observations	19,345,412	19,345,412	18,769,516	18,769,516	22,885,452	22,885,452	22,297,374	22,297,374
Unique Borrowers	477,181	477,181	419,161	419,161	477,181	477,181	419,161	419,161

Estimates from multinomial logit panel regression. Clustered standard errors by borrower in parentheses. Coefficients are exponentiated and represent the relative risk ratio. All models include covariates in Table 4, including state and year fixed effects. Models 2 and 4 exclude loans originated in the year 2014 due to missing service delivery variables in 2014 (and thus have fewer unique borrowers and observations).
*** p<0.01, ** p<0.05, * p<0.1

Table 6: Competing Risks, Compare HFA by Origination Cohort

	Base Model Specification				With Loan Structure and Servicing Controls			
Panel A: 2005-2007	(1)		(2)		(3)		(4)	
	Default	Prepay	Foreclose	Prepay	Default	Prepay	Foreclose	Prepay
HFA Dummy	0.768***	0.689***	0.688***	0.730***	0.949**	0.616***	0.809***	0.631***
	(0.009)	(0.007)	(0.011)	(0.007)	(0.022)	(0.011)	(0.023)	(0.011)
Observations	14,774,518	14,774,518	18,086,801	18,086,801	14,774,518	14,774,518	18,086,801	18,086,801
Panel B: 2008-2011								
HFA Dummy	0.802***	0.717***	0.702***	0.737***	0.916	0.712***	0.745*	0.740***
	(0.039)	(0.016)	(0.051)	(0.016)	(0.108)	(0.044)	(0.126)	(0.046)
Observations	2,389,546	2,389,546	2,603,670	2,603,670	2,389,546	2,389,546	2,603,670	2,603,670
Panel C: 2012-2014								
HFA Dummy	1.342	0.695***	0.697	0.701***	0.945	0.743**		
	(0.249)	(0.033)	(0.727)	(0.033)	(0.419)	(0.104)		
Observations	2,012,810	2,012,810	2,014,352	2,014,352	1,605,452	1,605,452		

Estimates from multinomial logit panel regression. Clustered standard errors by borrower in parentheses. Coefficients are exponentiated and represent the relative risk ratio. All models include covariates in Table 4, including state and year fixed effects. Model 4 is excluded for the final cohort due to very low probability of foreclosure and high correlation between the HFA indicator and loan structure covariates in the final period.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 7: Competing Risks, Compare HFA Effects by Interest Rate and Cohort

	All Years		2005-2007		2008-2011		2012-2013	
Panel A: Default vs. Prepayment								
	Default	Prepay	Default	Prepay	Default	Prepay	Default	Prepay
HFA (below market rate)	0.882***	0.546***	0.891***	0.518***	0.703**	0.551***	0.676	0.449***
	(0.022)	(0.010)	(0.023)	(0.010)	(0.106)	(0.040)	(0.471)	(0.119)
HFA (at market rate)	0.976	0.758***	0.985	0.725***	0.881	0.702***	0.904	0.796
	(0.023)	(0.013)	(0.024)	(0.014)	(0.107)	(0.045)	(0.435)	(0.123)
HFA (above market rate)	1.204***	0.900***	1.176***	0.828***	1.068	0.855**	1.423	1.149
	(0.054)	(0.026)	(0.068)	(0.045)	(0.134)	(0.059)	(0.741)	(0.204)
Observations	18,769,516	18,769,516	14,774,518	14,774,518	2,389,546	2,389,546	1,605,452	1,605,452
Panel B: Foreclose vs. Prepayment								
	Foreclose	Prepay	Foreclose	Prepay	Foreclose	Prepay		
HFA (below market rate)	0.749***	0.568***	0.749***	0.536***	0.566**	0.578***		
	(0.024)	(0.010)	(0.024)	(0.010)	(0.126)	(0.042)		
HFA (at market rate)	0.848***	0.770***	0.851***	0.731***	0.708*	0.727***		
	(0.025)	(0.013)	(0.026)	(0.014)	(0.125)	(0.046)		
HFA (above market rate)	0.973	0.902***	0.933	0.795***	0.876	0.879*		
	(0.057)	(0.026)	(0.067)	(0.043)	(0.151)	(0.059)		
Observations	22,297,374	22,297,374	18,086,801	18,086,801	2,603,670	2,603,670		

Estimates from multinomial logit panel regression. Clustered standard errors by borrower in parentheses. Coefficients are exponentiated and represent the relative risk ratio. All models include service delivery variables from Table 5 as well as covariates in Table 4, including state and year fixed effects. Foreclosure models are excluded for the final cohort due to very low probability of foreclosure in the final period.

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Figure 1: HFA Loan Volume
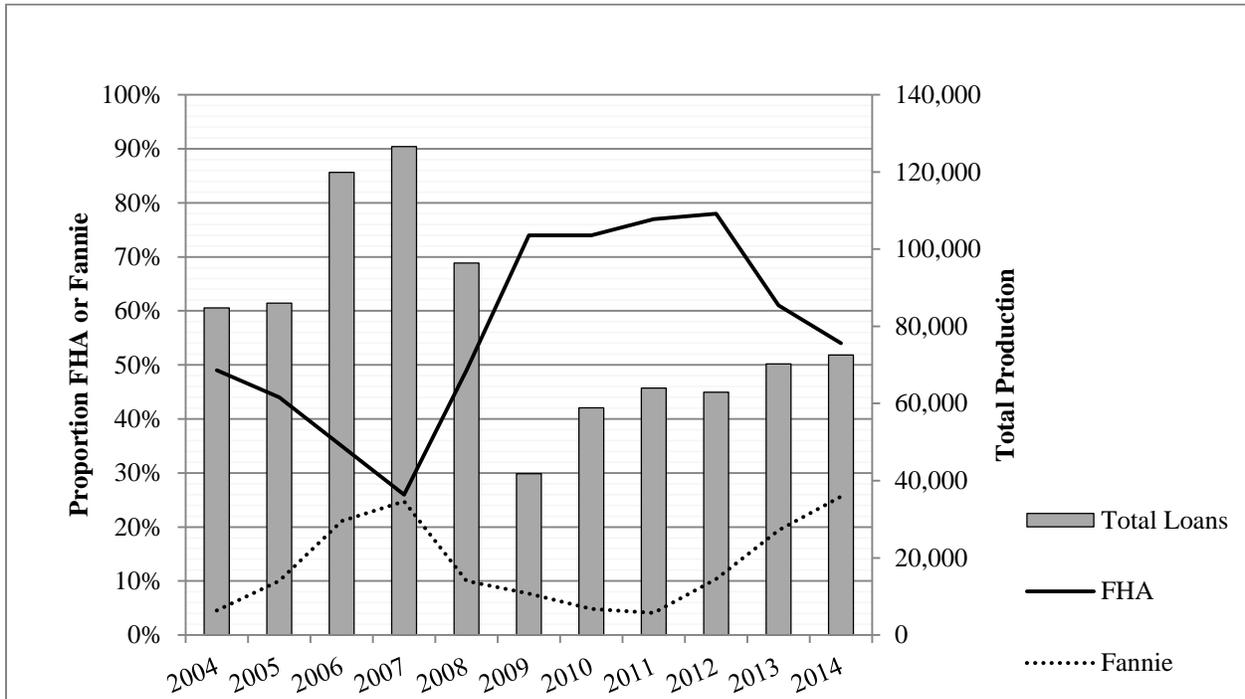
Source: Author's calculations from National Council of State Housing Agency Factbook data and Fannie Mae Single Family Lending data. Total loans include HFA single family, home purchase loans (excludes refinancing). For the years 2011 through 2014, total loans include other non-MRB single family loans in addition to MRB loans. Prior to 2010, the total includes only MRB loan volume.

Figure 2: HFA Interest Rate Differential
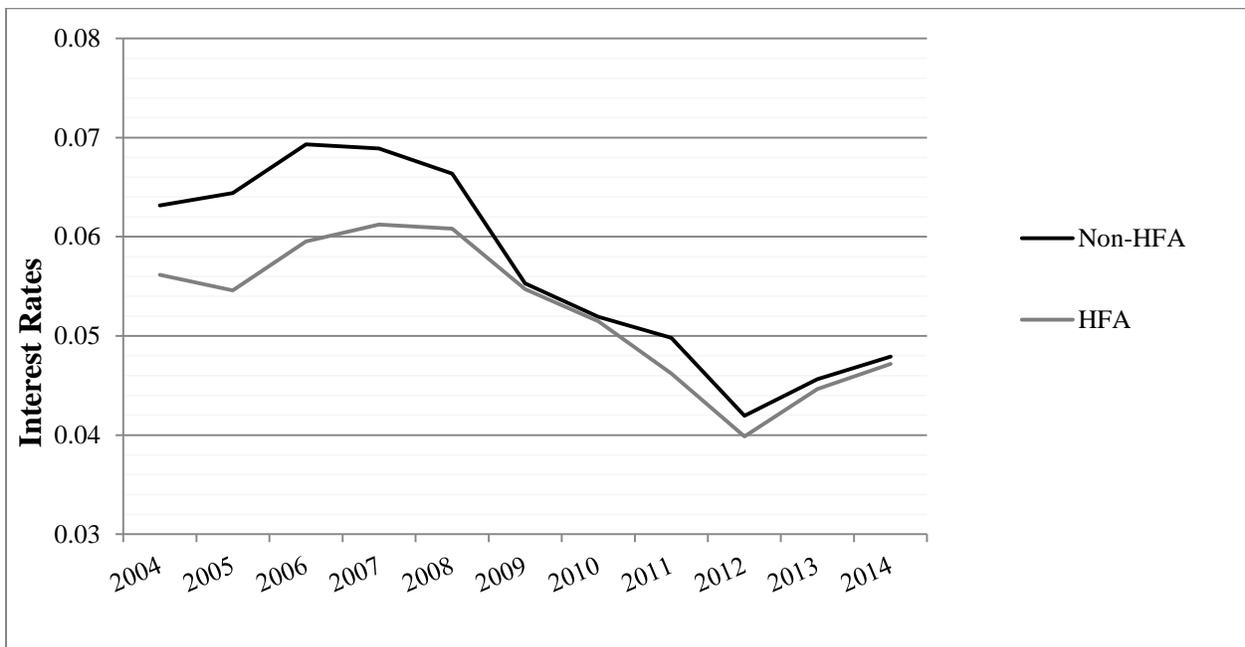
Figure 3: Cumulative Hazard Rate of Default, Matched Sample (weighted)
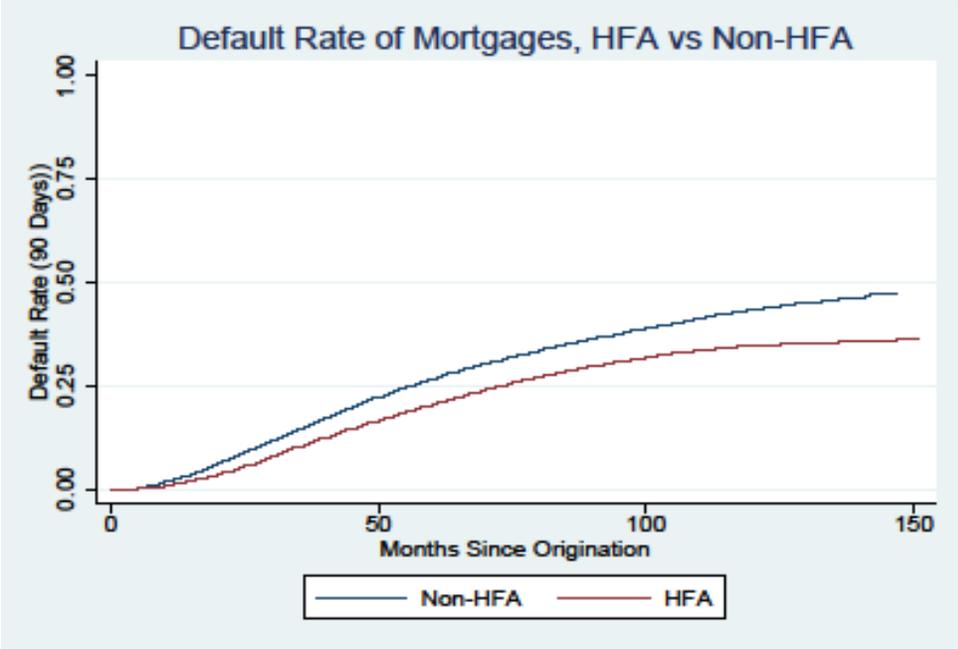
Source: Author's calculations from Fannie Mae Single Family Loan data

References

Agarwal, S., Amromin, G., Ben-David, I., Chomsisengphet, S., & Evanoff, D. D. (2014). "Predatory lending and the subprime crisis." *Journal of Financial Economics* 113(1), 29-52.

Alexander, WP, SD Grimshaw, GR McQueen, and BA Slade. (2002). "Some loans are more equal than others: Third-party originations and defaults in the subprime mortgage industry." *Real Estate Economics* 30:667–697.

Bogin, A., Doerner, W.M., and Larson, W. (2016). "Local house price dynamics: New indices and stylized facts." *Federal Housing Finance Agency*. Working Paper 16-01

Brown, S. R. (2016). "The influence of homebuyer education on default and foreclosure risk: A natural experiment." *Journal of Policy Analysis and Management* 35(1), 145-172.

Cutts, A. C. , & Merrill, W. (2008). "Interventions in mortgage defaults: Problems and practices to prevent home loss and lower costs." In N. P. Retsinas & E. S. Belsky (Eds.), *Borrowing to live: Consumer and mortgage credit revisited* (pp. 203–254). Washington, DC: Brookings Institution Press.

Demyanyk, Y., & Van Hemert, O. (2011). "Understanding the subprime mortgage crisis." *Review of Financial Studies* 24(6), 1848-1880.

Deng, Y.H., & Gabriel, S. A. (2006). "Risk-based pricing and the enhancement of mortgage credit availability among underserved and higher credit-risk populations." *Journal of Money, Credit, and Banking* 38(6), 1431-1460.

Deng, Y.H., Quigley, J.M., & Van Order, R. (1996). "Mortgage default and low downpayment loans: the costs of public subsidy." *Regional Science and Urban Economics* 263- 285.

Deng, Y.H., Quigley, J. M., & Van Order, R. (2000). "Mortgage terminations, heterogeneity and the exercise of mortgage options." *Econometrica* 275-307.

Deng, Y., & Gabriel, S. (2002). "Modeling the performance of FHA-insured loans: Borrower heterogeneity and the exercise of mortgage default and prepayment options." *Report to Office of Policy Development and Research,* United States Department of Housing and Development, May.

Ding, L, Quercia, RG, Li, W and Ratcliffe, J. (2011). "Risky borrowers or risky mortgages disaggregating effects using propensity score models." *Journal of Real Estate Research* 33:245–277.

Dylla, D. and Caldwell-Tautges, D. (2012). "Winning strategies: An analysis of state housing finance agency support for homeownership education and counseling services." *Doug Dylla Consulting Services*, February.

Ergungor, E. and Moulton, S. (2014). "Beyond the transaction: Banks and mortgage default of low-income homebuyers." *Journal of Money, Credit and Banking* 46: 1721-1751.

Fannie Mae. (2016). "Community Seconds." Fact Sheet for Mortgage Lenders. Available Online at: https://www.fanniemae.com/content/fact_sheet/community-seconds-fact-sheet.pdf

Federal Housing Finance Agency (FHFA). (2014). Annual Housing Report. Available Online at: https://www.fhfa.gov/AboutUs/Reports/ReportDocuments/Annual_Housing_Report_2014.pdf

Federal Housing Finance Agency (FHFA). (2016). Annual Housing Report. Available Online at: https://www.fhfa.gov/AboutUs/Reports/ReportDocuments/Annual-Housing-Report-2016.pdf

Fout, H., Li, G. and Palim, M. (2017). "Credit risk of low income mortgages." Economic and Strategic Research, Fannie Mae.

Fuster, A,, and Zafar, B. (2016). "To buy or not to buy: Consumer constraints in the housing market." *American Economic Review* 106(5): 636-40.

Gerardi, K., Shapiro, A., and Willen, P. (2009). "Decomposing the foreclosure crisis: House price depreciation versus bad underwriting." Federal Reserve Bank of Atlanta Working Paper.

Guiso, L., Sapienza, P and Zingales, L. (2009). "Moral and social constraints to strategic default on mortgages." NBER Working Paper Series w15145. National Bureau of Economic Research.

Iacus, S. M., King, G., & Porro, G. (2012). "Causal inference without balance checking: Coarsened exact matching." *Political Analysis 20*(1), 1–24.

Jiang, W., Nelson, A. A., & Vytlacil, E. (2014). "Liar's loan? Effects of origination channel and information falsification on mortgage delinquency." *Review of Economics and Statistics* 96(1), 1-18.

Kelly, A. (2008). ""Skin in the game": Zero downpayment mortgage default." *Journal of Housing Research* 17:75–98.

LaCour-Little, M. 2009. "The pricing of mortgages by brokers: an agency problem?" *Journal of Real Estate Research* 31:235–263.

Lam, K., Dunsky, R. M., & Kelly, A. (2013). "Impacts of down payment underwriting standards on loan performance—evidence from the GSEs and FHA portfolios." FHFA Working Paper, 13-3.

Leventis, Andrew. (2014). "The Relationship between Second Liens, First Mortgage Outcomes, and Borrower Credit: 1996-2010." FHFA Working Paper 14-3. Available at SSRN: https://ssrn.com/abstract=2498532

Moulton, S. (2010). "Originating lender localness and mortgage sustainability: An evaluation of delinquency and foreclosure in Indiana's mortgage revenue bond program." *Housing Policy Debate* 20:581–617.

Moulton, S. and Quercia, R. (2014). "Access and sustainability for first time homebuyers: The evolving role of state housing finance agencies." with Roberto Quercia. In *Homeownership Built to Last: Balancing Access, Affordability, and Risk after the Housing Crisis.* Brookings Institution Press with the Joint Center for Housing Studies of Harvard University.

National Council of State Housing Agencies (NCSHA). (2011). "NCSHA supports bond exemption in risk retention rule." Blog Post. https://www.ncsha.org/blog/ncsha-supports-bond-exemption-risk-retention-rule August 2, 2011.

National Council of State Housing Agencies (NCSHA). (2015). State HFA Factbook: 2013 NCSHA Annual Survey Results. Washington, D.C.

Pennington-Cross, A., & Chomsisengphet, S. (2007). Subprime refinancing: Equity extraction and mortgage termination. *Real Estate Economics*, 35(2), 233-263.

Quercia, R., & Spader, J. (2008). "Does homeownership counseling affect the prepayment and default behavior of affordable mortgage borrowers?" *Journal of Policy Analysis and Management*, 27(2), 304-325.

Stegman, MA, RG Quercia, J Ratcliffe, L Ding, and WR Davis. (2007). "Preventive servicing is good for business and affordable homeownership policy." *Housing Policy Debate* 18:243–278.

United States General Accounting Office (GAO). (1983). "The costs and benefits of single family mortgage revenue bonds: Preliminary report." GAO/RCED-83-145. Washington, D.C.: U.S. General Accounting Office, April 1983.

United States General Accounting Office (GAO). (1988). "Homeownership: Mortgage bonds are costly and provide little assistance to those in need." GAO/RCED-88-111. Washington, DC: U.S. General Accounting Office, March 1988.

Van Order, R., & Zorn, P. 2003. "Income, location and default: Some implications for community lending." *Real Estate Economics* 385 - 404.

Appendix

HFA Borrower Demographics

Figure A1 shows average demographic characteristics of the population of HFA borrowers between 2004 and 2013, based on state level data reported in the annual HFA Factbook. On average over time, state HFAs report that approximately 60 percent of their borrowers have low income, defined as households with incomes below 80 percent of area median. The share of HFA loans to very low income borrowers (with incomes below 60 percent of area median) increased to roughly one-third of all borrowers following the contraction of the private mortgage market in 2009 and 2010, dropping to just over 20 percent in 2013. This is perhaps not surprising given the increased reliance on non-MRB MBS beginning in 2012, with fewer income restrictions than HFA mortgages backed by MRBs. About one-fourth of borrowers served by HFAs are minority households and about one-third of borrowers are female headed households, these proportions also peaking during the 2009 and 2010 contraction of the private mortgage market.

HFA Borrowers in the Fannie Mae Dataset

The Fannie Mae dataset includes only a portion of total HFA loans (as summarized in Figure 1). The proportion of HFA loans securitized by Fannie Mae varies over time, following trends in the mortgage market for LMI borrowers generally. Figure A2 employs HMDA data from 2004 through 2014 to graph the total number of loans to LMI borrowers (with incomes below 115% of AMI) in a given year, overlaid with the proportion of loans that were FHA or conventional. Trends in the securitization of LMI loans in the population over time are similar to those observed for HFA lending (Figure 1). However, the proportion of FHA mortgages is generally much higher in HFA population than the overall LMI population.

There is variation between state HFAs in the extent to which they securitize loans through Fannie Mae. Table A1 shows the total number of HFA loans originated by state and year and the number of HFA loans securitized by Fannie Mae for the state and year. Several states emerge as dominant HFA lenders in the sample in different time periods. Prior to 2009, states with a large share of HFA mortgages securitized by Fannie Mae included Florida, Ohio, Iowa and Illinois—although many states were represented in the Fannie Mae HFA database. After 2009, fewer states securitized a large portion of their loans through Fannie Mae, with Massachusetts emerging as the most dominant originator.

HFA Service Delivery Practices

Figure A3 summarizes the distribution of HFA service delivery practices over time. Most state HFAs report providing homebuyer education and counseling, with approximately 80 percent reporting the provision of such services since 2008. By contrast, the proportion of state HFAs originating mortgages directly has declined over time, from nearly 50 percent in 2004 to 26 percent in 2013. The proportion of state HFAs directly servicing the majority of their loans in house is around 30 percent, with some growth in recent years.

The dominance of particular states in the Fannie Mae dataset necessarily implies that the practices of the HFAs in those particular states will have more influence over the analysis. Figure A4 graphs the distribution of HFA loans in the Fannie Mae database by states with particular service delivery practices. The trends in this graph differ substantively from the trends in Figure A3 of service delivery practices by state for the entire population of HFAs. The proportion of HFA loans in the Fannie Mae database originated by states with homeownership counseling peaks at 90 percent in 2008, but then drops to fewer than half of the originations in 2009 through 2012. By contrast, the proportion of HFA loans in the Fannie database with direct servicing is much higher post 2009 than in the general HFA population, with nearly 80 percent of loans originated through state HFAs with direct servicing. Much of this reflects the dominance of a few states in the Fannie database (particularly, Massachusetts) in the post 2009 period. Our weighted sample regressions (with exact matching on geography, as described below) account for the dominance of particular states with HFA loans in the Fannie Mae database. This helps to ensure that the results are generalizable within the Fannie Mae population of HFA loans. However, the results of the analysis may be less generalizable to the population of HFA loans in the U.S. as a whole.

Our empirical approach to match on three digit ZIP code accounts for geographic differences to the extent that HFA and non-HFA loans are prevalent within the same three-digit ZIP codes and are distributed in similar neighborhoods. We find evidence that this is the case. For example, Figure A5 plots the locations of HFA and non-HFA originated mortgages within the Boston MSA in 2012, overlaid on ZIP code demographics for the percent black. The concentration of loans to first time homebuyers in particular ZIP codes appear to be relatively evenly distributed for the HFA and non-HFA samples.

Supplemental Tables

Descriptive statistics for MNL model variables are included in Table A2. The results of the Cox proportional hazard model are included in Table A3.

Table A1: HFA Loan Volume from NCSHA Factbook, compared to Fannie Mae HFA Volume

	2005 Total	FNMA	2006 Total	FNMA	2007 Total	FNMA	2008 Total	FNMA	2009 Total	FNMA	2010 Total	FNMA	2011 Total	FNMA	2012 Total	FNMA	2013 Total	FNMA
AK	942	1	1,328	0	1,403	0	1,383	0	302	0	2,078	2	1,305	0	1,082	0	1,850	0
AL	0	2	695	142	2,167	1,026	657	157	0	0	0	0	998	0	892	0	782	28
AR	0	26	1,085	141	1,331	301	1,272	90	269	0	513	0	593	0	599	0	540	0
AZ	116	114	280	198	366	51	156	2	6	0	108	0	368	0	132	0	207	0
CA	5,327	27	6,140	59	5,217	1,236	4,070	248	168	13	7	7	1,394	57	0	44	0	43
CO	2,390	6	2,422	72	2,934	140	2,816	92	203	0	0	0	2,337	0	2,135	4	2,415	545
CT	3,558	0	4,010	0	0	0	2,794	0	1,976	0	1,963	0	1,612	0	1,445	0	1,733	91
DC	0	1	150	123	273	27	31	6	0	0	36	0	44	0	0	0	54	41
DE	394	247	1,089	795	2,117	1,305	1,281	318	525	0	626	27	876	50	407	43	737	54
FL	1,163	413	2,723	1,839	4,704	3,371	3,390	905	2,578	113	4,369	51	3,712	19	2,784	13	2,036	82
GA	447	54	1,125	201	1,381	475	987	99	449	2	1,118	1	1,417	1	1,005	0	1,465	0
HI	13	41	90	38	29	29	0	0	0	0	0	0	0	0	22	30	107	78
IA	2,487	1,317	3,006	2,124	3,681	2,351	2,047	614	672	127	1,100	127	1,013	171	2,102	282	1,402	546
ID	1,230	3	1,734	10	4,098	626	2,199	1,168	289	403	207	286	2,750	180	5,075	970	3,842	1,403
IL	1,662	115	1,647	346	0	1,210	1,089	15	17	0	311	0	1,002	182	1,697	260	0	963
IN	1,682	194	2,430	721	2,691	176	2,997	510	507	4	1,082	4	1,269	11	1,549	7	1,196	16
KS	N/A	82	N/A	468	N/A	0	N/A	0	N/A	0	N/A	0	N/A	0	N/A	0	N/A	0
KY	3,803	797	5,221	2,004	4,095	1,441	2,582	82	1,751	14	2,069		1,030		442	0	2,449	427
LA	428	95	1,556	631	2,080	63	1,609	4	437	35	444	8	505	2	95	3	208	17
MA	896	109	1,341	295	1,263	825	1,183	729	744	1,896	836	1,180	1,137	884	3,041	2,647	4,110	2,958
MD	887	6	2,591	14	4,064	1	2,340	0	702	0	707	0	1,595	0	1,731	17	1,799	283
ME	1,019	0	1,094	0	961	0	1,163	0	1,002	0	916	0	795	0	318	0	610	0
MI	1,150	0	2,396	0	2,014	0	4,426	0	899	0	947	0	1,099	0	0	0	0	0
MN	2,845	284	2,784	501	3,683	150	2,419	0	935	42	2,287	210	2,320	127	2,403	293	3,329	889
MO	2,409	351	3,407	1,129	3,078	969	1,482	175	1,644	59	2,818	10	3,308	42	4,470	25	4,801	27
MS	1,749	5	2,700	269	2,982	1,164	859	134	198	0	587	0	621	1	495	1	6	0
MT	1,822	5	1,758	17	1,751	15	778	30	255	8	0	0	270	0	0	0	568	0
NC	1,445	0	2,173	0	3,185	0	1,266	0	280	0	317	0	478	0	730	0	1,232	0
ND	1,249	0	1,324	1	1,662	0	1,656	0	1,620	0	1,388	0	1,110	0	885	0	1,072	0
NE	1,542	435	3,688	2,599	4,766	3,745	2,182	884	424	0	1,585	0	1,504	88	1,150	80	2,219	671

State																		
NH	1,212	0	1,228	0	1,380	0	759	0	421	0	527	0	613	3	495	25	689	114
NJ	483	0	1,159	0	2,163	0	1,780	0	728	0	579	0	777	0	693	0	1,132	0
NM	1,058	134	1,901	695	2,311	1,175	1,760	475	1,417	100	1,159	59	1,012	112	895	145	1,147	236
NV	0	3	222	65	0	169	495	64	218	0	553	2	0	0	0	0	56	4
NY	3,206	0	3,399	0	2,672	0	4,001	0	1,136	0	2,240	0	1,683	0	797	0	1,717	4
OH	5,027	1,285	9,918	6,338	7,965	5,594	7,038	1,811	2,296	23	3,234	51	3,448	129	3,342	97	3,871	388
OK	1,281	308	2,155	655	1,845	767	751	24	787	0	1,122	0	1,213	0	1,441	0	586	0
OR	1,195	0	1,171	0	1,381	0	1,598	0	381	0	37	0	620	0	434	0	419	0
PA	5,678	1,040	6,660	654	6,960	118	4,839	251	2,905	2	6,527	0	4,361	19	3,076	461	1,656	1,174
PR	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
RI	838	0	1,088	0	1,449	0	928	0	718	0	409	0	354	0	425	101	484	197
SC	802	0	1,852	0	1,489	0	1,286	0	535	0	1,043	0	706	0	365	0	298	0
SD	2,245	27	2,495	6	2,792	0	2,432	0	1,730	0	1,700	0	1,447	0	808	84	1,486	244
TN	2,431	0	3,270	0	4,647	0	2,885	0	2,446	0	2,650	0	2,222	0	2,240	0	2,071	0
TX	1,924	349	2,531	1,065	2,636	1,234	1,409	73	273	2	1,248	14	1,920	15	2,100	0	1,827	0
UT	1,991	192	2,030	233	2,193	250	1,915	294	959	111	1,130	0	1,852	0	2,686	0	2,685	203
VA	5,114	0	6,166	1	5,348	0	4,954	0	3,964	1	3,409	0	2,445	0	3,422	80	3,771	325
VT	707	0	935	0	993	0	495	0	95	33	135	18	221	66	407	115	360	104
WA	1,004	528	1,755	658	1,870	1,292	794	381	684	141	1,370	96	1,124	76	489	57	2,753	351
WI	4,131	0	4,559	149	4,705	0	2,748	0	0	15	700	671	374	369	651	586	1,126	969
WV	1,473	16	1,488	11	1,564	62	897	41	252	73	0	24	450	14	851	47	581	34
WY	1,563	0	1,951	0	2,272	0	1,547	0	1,016	0	690	0	662	0	633	0	754	118
Total	86,018	8,612	119,920	25,267	126,611	31,358	96,425	9,676	41,813	3,217	58,881	2,848	63,966	2,618	62,936	6,517	70,238	13,627

Table A2: MNL Model Variables, Descriptive Statistics

	mean	sd	min	max
HFA Flag	0.22	0.41	0	1
Monthly Income (thousands)	3.57	1.4	0.31	9
Mortgage Payment (hundreds)	7.57	3.5	0.61	38.52
DTI<36 (spline)	9.47	13.98	0	36
DTI 36.1-45 (spline)	13.09	19.02	0	45
DTI>45 (spline)	18.65	26	0	100
Borrower FICO Score at Origination	712.26	55.59	434	950
Co-borrower	0.25	0.43	0	1
Borrower FICO - Co-Borrower FICO	7.63	21.45	0	327
Borrower Age	33.22	10.79	18	99
Borrower Age (missing)	0.07	0.26	0	1
Single Unit	0.85	0.36	0	1
Δ Three Year HPI (Zip)	-0.19	5.4	-38.8	31.88
Mark-to-Market CLTV	98.56	21.61	0.09	862.17
30-Year FRM Interest Rate	5	0.96	3.36	6.6
CPI (Inflation)	2.09	1.64	-1.96	5.5
Unemployment Rate	6.65	2.41	1.7	28.9
Months Since Origination	33.32	25.24	0	133
Months Since Origination (squared)	1747.61	2310.61	0	17689
Broker Originator	0.19	0.4	0	1
Correspondent Originator	0.36	0.48	0	1
Community Second	0.1	0.3	0	1
Other Second	0.08	0.27	0	1
Full Documentation	0.96	0.2	0	1
Direct Lending	0.07	0.26	0	1
Direct Servicing	0.05	0.21	0	1
Homeownership Counseling	0.15	0.35	0	1
Year of Origination	2007.2	2.15	2004	2014
Year of Observation	2009.96	2.45	2004	2016

N= 19,345,412; summary statistics are weighted using the CEM weights

Table A3: Cox Proportional Hazard Model, 90 Day Default

	(1)	(2)	(3)
HFA Flag	0.748***	0.841***	0.929***
	(0.009)	(0.011)	(0.019)
Monthly Income (thousands)	0.851***	0.850***	0.849***
	(0.006)	(0.006)	(0.006)
Mortgage Payment	1.083***	1.078***	1.079***
	(0.003)	(0.003)	(0.003)
DTI<36	1.404*	1.372*	1.305
	(0.252)	(0.244)	(0.230)
DTI 36.1-45	1.529***	1.489***	1.436***
	(0.204)	(0.197)	(0.190)
DTI>45	1.503***	1.497***	1.452***
	(0.152)	(0.149)	(0.145)
Borrower Age	1.007***	1.007***	1.007***
	(0.001)	(0.001)	(0.001)
Borrower Age (missing)	0.864***	0.883***	0.892***
	(0.026)	(0.027)	(0.027)
Borrower FICO Score at Origination	0.989***	0.989***	0.989***
	(0.000)	(0.000)	(0.000)
Borrower FICO - Co-Borrower FICO	0.999*	0.999	0.999
	(0.000)	(0.000)	(0.000)
Co-borrower	0.927***	0.951**	0.950**
	(0.022)	(0.022)	(0.022)
CLTV	1.033***	1.038***	1.038***
	(0.001)	(0.001)	(0.001)
Single Unit	0.932***	0.928***	0.939***
	(0.018)	(0.018)	(0.017)
Δ Three Year Unemployment (Zip)	1.164***	1.168***	1.173***
	(0.023)	(0.023)	(0.023)
Δ Three Year HPI (Zip)	0.979***	0.979***	0.979***
	(0.002)	(0.002)	(0.002)
Broker Originator		1.258***	1.264***
		(0.024)	(0.024)
Correspondent Originator		1.150***	1.157***
		(0.020)	(0.020)
Community Second		0.782***	0.778***
		(0.024)	(0.023)
Other Second Lien		0.919***	0.921***
		(0.023)	(0.023)
Full Documentation		0.552***	0.553***
		(0.017)	(0.017)
Direct Lending			1.053**
			(0.023)
Direct Servicing			0.801***

			(0.027)
Homeownership Counseling			0.872***
			(0.018)
Observations	477,181	477,181	419,161

Robust standard errors in parentheses. Coefficients are exponentiated and represent the hazard ratio. All models include non-time varying covariates in Table 4, including state and year of origination fixed effects

*** p<0.01, ** p<0.05, * p<0.1

Figure A1: State-Level HFA Indicators of Borrower Need, All HFA Lending
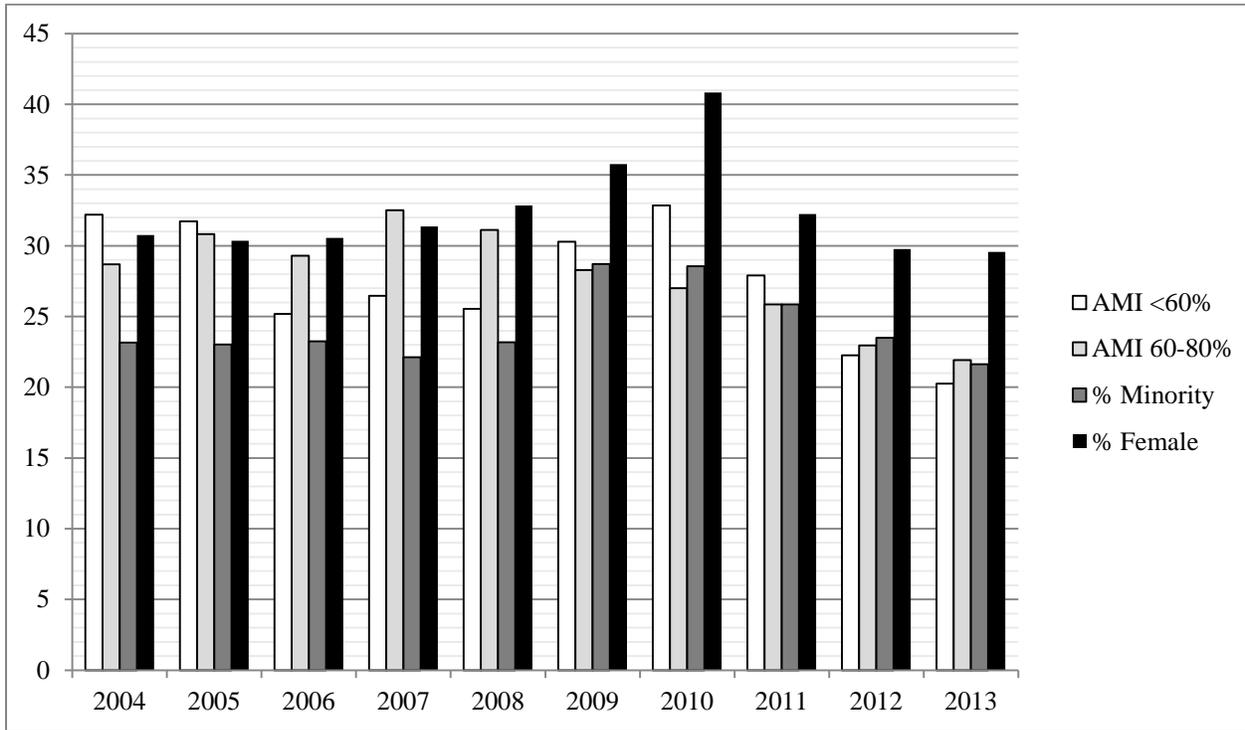
Source: Author's calculations from National Council of State Housing Agency Factbook data

**Figure A2: HMDA Loans to LMI Borrowers (Incomes below 115% AMI)**
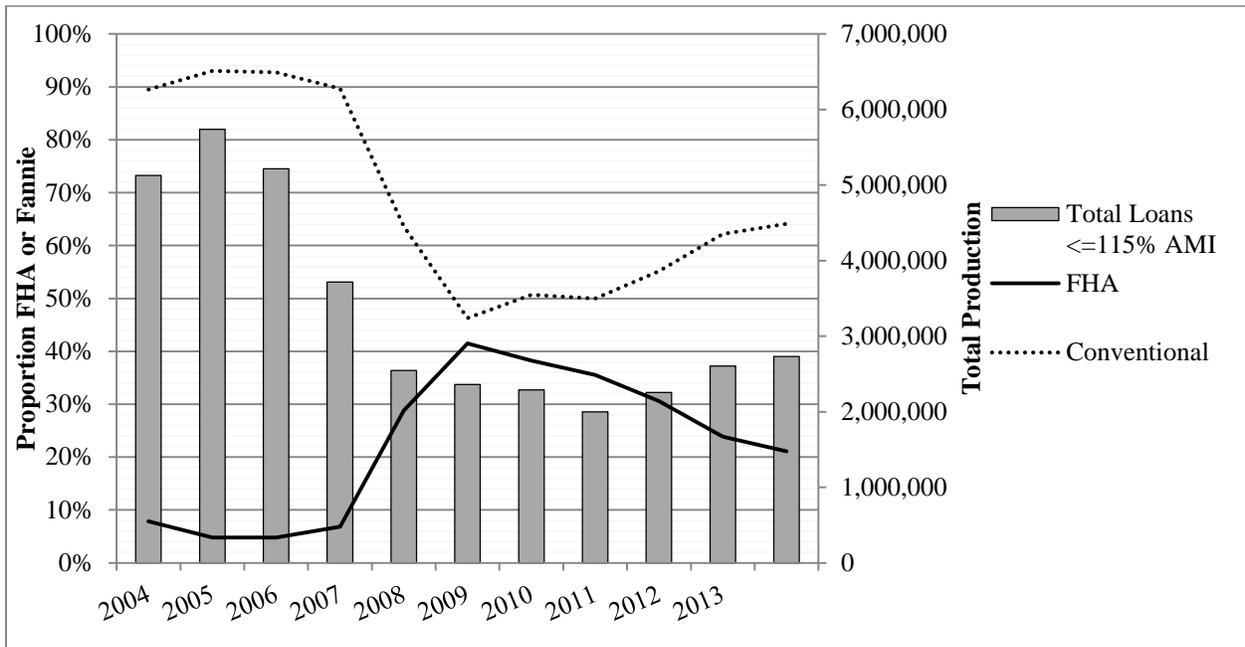
Source: Authors' calculations from HMDA data

Figure A3: HFA Service Delivery Practices at State Level, All HFA Lending
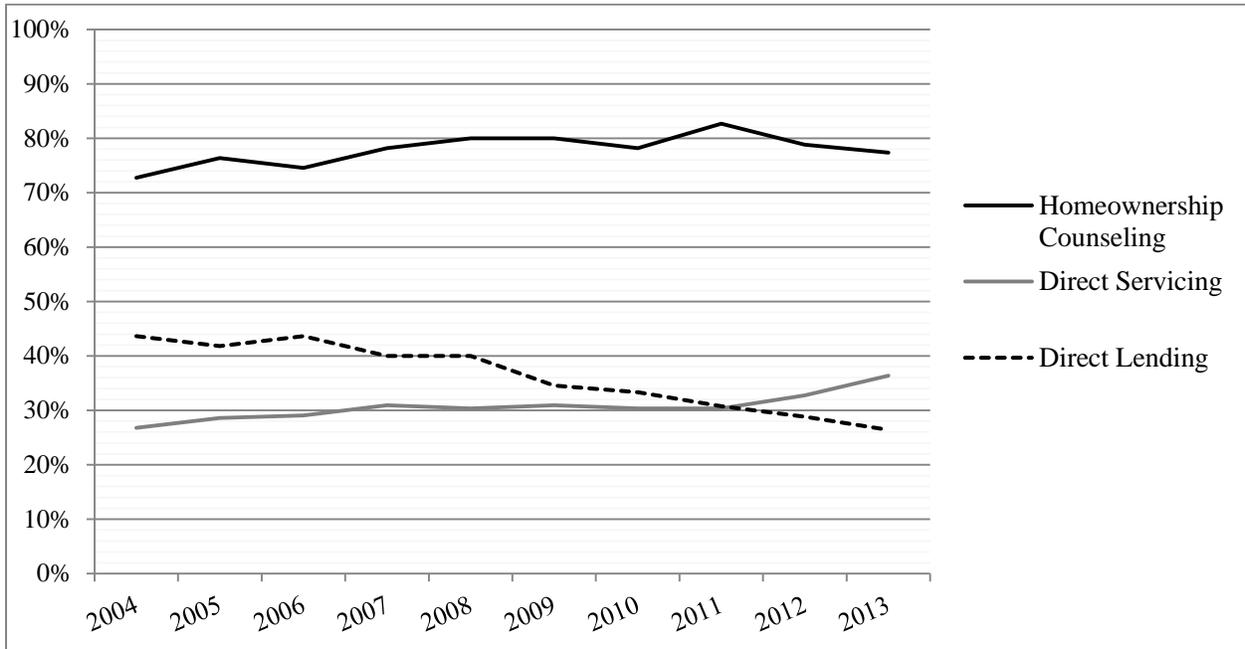
Source: Author's calculations from National Council of State Housing Agency Factbook data

**Figure A4: HFA Service Delivery Practices at Loan Level, Fannie Mae HFA Loans**
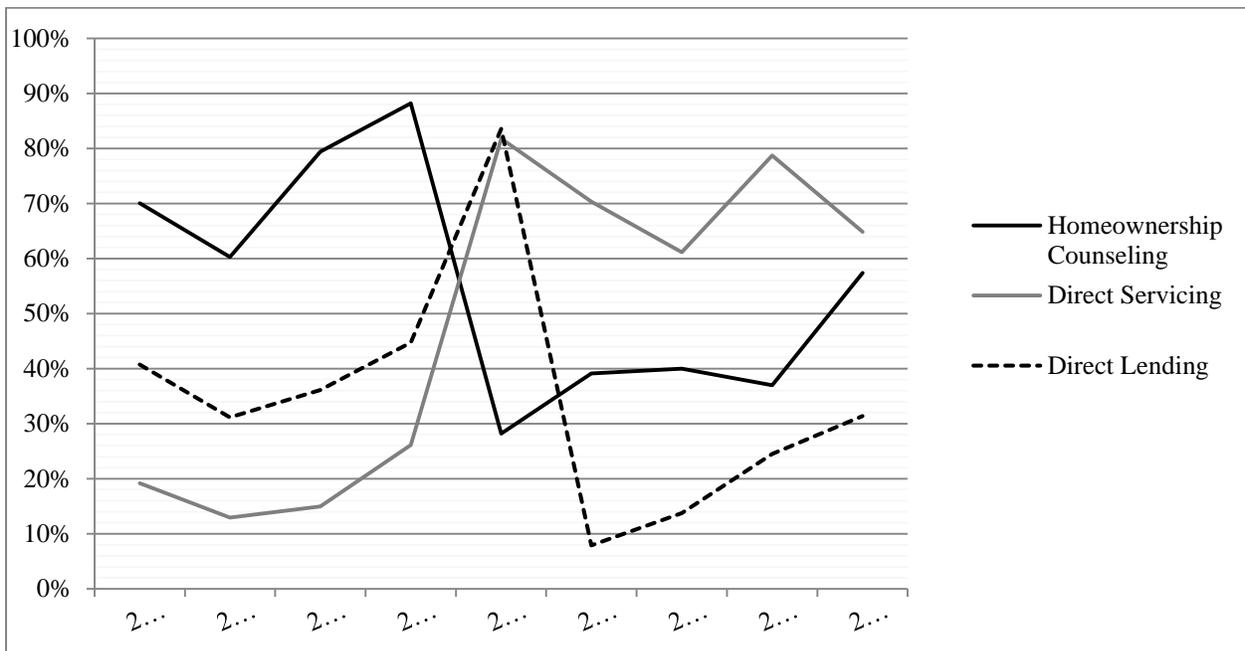
Source: Author's calculations from Fannie Mae Single Family Loan data combined with NCSHA Factbook data on service delivery practices by year.

Figure A5: Spatial Distribution of Fannie Mae First-Time Homebuyer Loans, Boston MSA 2012

Non-HFA Originations
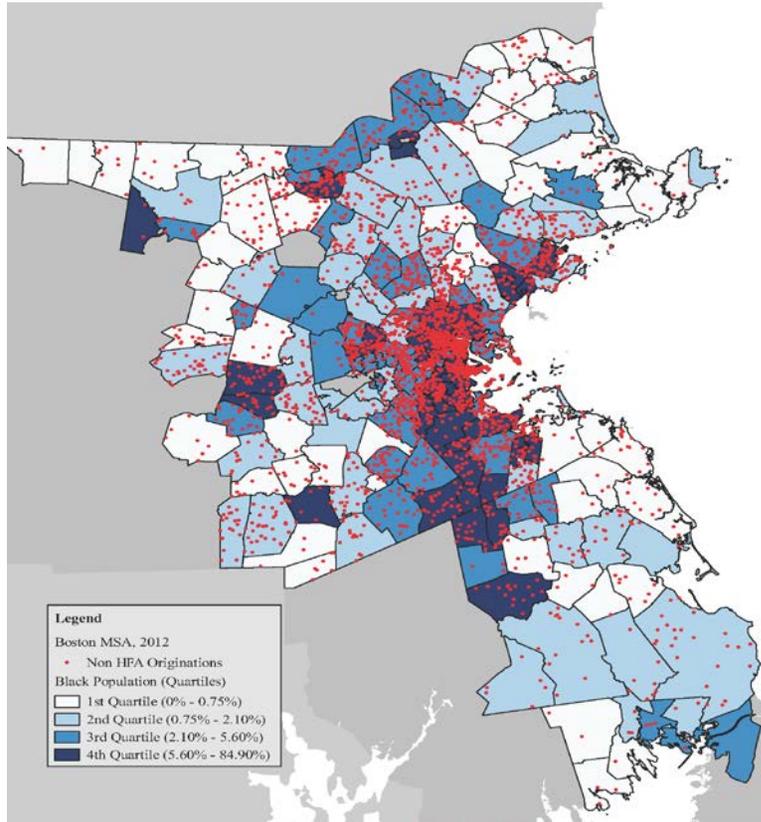
HFA Originations

Source: Author's calculations from Fannie Mae Single Family Loan data