# Predatory Lending and Hidden Risks

Sumit Agarwal, Gene Amromin, Itzhak Ben-David,
Douglas D. Evanoff and Souphala Chomsisengphet*

December 12, 2023

**Abstract**

We present novel empirical evidence about the methods and consequences of predatory lending. Rather than focusing on loan or borrower features to identify predatory lending, we analyze a particular business practice in which borrowers are rejected and then approved in rapid succession by the same lender. We show that such borrowers' ex-post characteristics and mortgage contracts are consistent with predatory steering. Specifically, steered borrowers are more likely to come from groups with lower levels of financial sophistication. They are more likely to enter non-amortizing contracts with high margins for originators that are quickly securitized. Steered borrowers default less in boom years when refinancing is easy. However, their performance deteriorates sharply once falling prices trap them in contracts with rising payments, reflecting the long-term costs of predatory lending.

*Keywords:* Mortgages, predatory lending, non-amortizing contracts, financial crisis, household finance.
*JEL Classification:* D12, D18, G21, G18, K2

---

*Agarwal is at the National University of Singapore, Amromin and Evanoff at the Federal Reserve Bank of Chicago, Ben-David at The Ohio State University and NBER, and Chomsisengphet is at the Office of the Comptroller of the Currency. The views expressed in this paper are those of the authors and do not reflect those of the Federal Reserve Bank of Chicago, the Federal Reserve System, the Office of the Comptroller of the Currency, or the US Treasury Department. We thank Caitlin Kearns and Mike Mei for outstanding research assistance and Han Choi for editorial assistance. We also thank Viral Acharya, Gadi Barlevy, Mike Berry, Michael Greff, Jason Keller, Steve Kuehl, Elizabeth Laderman, Geng Li, Leonard Nakamura, Mitchell Petersen, Amit Seru, Greg Udell, Lena Vanterpool, Alicia Williams and Marva Williams for constructive input. The comments of seminar participants at the University of Houston, Rice University, University of Notre Dame, Baruch College, Real Estate Finance & Investment Symposium at Cambridge are well-appreciated. All errors are those of the authors. The views expressed in this paper are those of the authors and do not reflect those of the Federal Reserve Bank of Chicago, the Federal Reserve System, the Office of the Comptroller of the Currency, or the U.S. Treasury Department. Email: bizagarw@nus.edu.sg, gene.amromin@chi.frb.org, ben-david.1@osu.edu, Souphala.Chomsisengphet@occ.treas.gov.

# 1 Introduction

Predatory lending practices were common at the height of the housing market boom of the 2000s (Renuart, 2004; Commission, 2010).[1] Engel and McCoy (2001) define predatory lending as "...onerous lending practices, which are often targeted at vulnerable populations and result in devastating personal losses, including bankruptcy, poverty, and foreclosure." Most of the existing literature identifies predatory lending either by ex-post contract characteristics—such as excessive fees, high interest rates, obscured prepayment penalties, and clauses barring borrowers from seeking judicial redress (Engel and McCoy, 2001; Carr and Kolluri, 2001), or by ex-post mortgage performance (e.g., Engel and McCoy, 2001; White, 2008). These approaches potentially conflate predatory practices, optimal contract choices, and risk exposure.[2] They also offer little systematic evidence about the methods for implementing predatory practices or mechanisms through which such practices inflict costs on borrowers and investors.

In this study, we provide new evidence about predatory mortgages' origination and subsequent performance over the real estate cycle. We identify predatory loans not by the features of the eventual contract but by the process through which a loan was originated. Specifically, we analyze a particular business practice in which borrowers' mortgage applications are initially rejected but approved soon thereafter without meaningful changes in borrower income, loan amount, or property value. A very short time window between the two outcomes (less than two weeks, on average) makes improvements in applicant credit scores unlikely as well. We contrast the outcomes of two sets of observationally equivalent borrowers—those who successfully obtained credit with a new lender and those who got a loan from the same lender to which they originally applied. Closely matching borrower characteristics in the two groups allows us to ascribe the difference in outcomes to lender behavior rather than under-

[1] Also see Luigi Zingales, Financial Stars Behind Bars?, *Project Syndicate*, January 28, 2013 and Rob Garver, Citi Corroborates Two Allegations, *American Banker*, July 30, 2001, page 4.

[2] Bond, Musto, and Yilmaz (2009) model predatory lending as a reverse asymmetric information problem, where the lender has superior information about the borrower than the borrower has about herself.

lying borrower preferences. Whereas new lenders potentially employ a different underwriting framework, the original lender is likelier to have granted approval by tweaking the contract form or granting a manual exception. This latter case opens a possibility of strategic denial of the first application to steer the borrower into a predatory contract.

Our empirical approach was motivated by background conversations with several mortgage practitioners active during the lending boom preceding the Global Financial Crisis (GFC). These lenders and mortgage officers recounted practices whereby a loan officer could reject an application and then use substantially similar data to get approval by tweaking some terms. The approved contract could be presented at the same time as notifying the borrower of the rejection of the original application. An application rejected through an automated underwriting system would generate a HMDA record, as would an application that eventually got approved. While far from definitive, these anecdotes present a potential path to identifying predatory lending practices.

We used 2003–2006 transaction-level data from HMDA filings to identify borrowers whose rejected applications received subsequent approval without meaningful recorded changes. We found that borrowers who stayed with the original lender—those who were potentially steered—had systematically different demographic characteristics and took out systematically different contracts than those who sought credit with different lenders. In particular, steered borrowers are more likely to be female (primary borrower), have no co-signers, and reside in low-to-moderate income areas. These groups of borrowers have been shown to have lower levels of financial literacy (e.g., Lusardi and Mitchell, 2014) and are thus potentially prone to manipulation by unscrupulous lenders.[3] The effects are well-identified and are economically significant. Each demographic characteristic listed above is associated with a 5 to 10 percentage points higher likelihood of being steered.

We also find that steered borrowers take out mortgage products that deliver high profit margins to lenders. We document that relative to the overall sample mean, steered borrowers

[3]Indeed, Berndt, Hollifield, and Sandås (2016) show that such borrowers paid higher fees for the same loans than their better-educated counterparts.

are 61% more likely to take interest-only mortgages (IOs), 80% more likely to take option ARM mortgages, 71% more likely to take mortgages with prepayment penalties and 33% more likely to take no or low documentation mortgages. Consistent with the idea that originators capitalize on the high margins offered for these products in the secondary market, we report that mortgages of steered borrowers are 47% more likely to be sold to private securitizers.[4]

Furthermore, after controlling for various borrower and loan characteristics, we document that steered loans have an annual percentage rate (APR) of 35–72 basis points higher than that of non-steered loans, which, given the sample average APR of 6.8%, is economically significant.[5] This body of evidence is consistent with findings of legal studies that characterize predatory lending (e.g., Engel and McCoy, 2001, 2006; Eggert, 2001; Azmy, 2005), and thus confirm our identification of steering. It is worth noting that this pattern is observed only for a narrow group of originators, suggesting that a few bad actors practiced this form of predatory lending rather than it being a standard industry practice.

These differences in contract terms may appear inconsequential, with features that defer debt amortization improving near-term affordability. However, we find marked differences in the relative dynamics of mortgages of the steered and non-steered borrowers that reveal the consequences of this predatory lending practice. In a nutshell, steered borrowers refinanced their mortgages much more frequently during the boom years and defaulted much more often in the bust years. This finding can be viewed in charts that depict the refinancing and default rates for steered and non-steered groups over time (Figures 1, panels (a) and (b)). During the boom part of our sample—2004 through early 2007—steered borrowers were more likely to refinance their mortgages, as their conditional quarterly refinancing rates were about 2

[4]Several legal studies argue that the originate-to-distribute model was a prime factor that led lenders to engage in predatory lending (e.g., Engel and McCoy, 2001, 2006; Eggert, 2001; Azmy, 2005). Their basic claim is that the strong demand for mortgages from investors (through securitization) incentivizes mortgage originators to push for mortgages that generate high origination fees. During our sample period, private-label securitizers paid handsome fees for mortgages with high-interest rates and exotic features (e.g., teaser rates and no/flexible amortization).

[5]For example, Agarwal, Rosen, and Yao (2016) find that a significant fraction of consumers refinances their mortgage at an interest rate differential of 40 basis points.

percentage points higher, on average (relative to the mean refinancing rate of 3.2% for non-steered borrowers). However, the difference in the likelihood of refinancing collapses in later years once the real estate market stalls and the bust begins.

Default rates exhibit the reverse pattern (Figure 1, panel (b)). In the first half of the sample (2004-early 2007), the default likelihood of steered borrowers is somewhat lower than that of non-steered borrowers. In later years, however, the quarterly default rates of the steered zoom were 2 percentage points above that of the non-steered. This is similar in magnitude to the effect of predatory lending on default documented by Agarwal, Amromin, Ben-David, Chomsisengphet, and Evanoff (2014). This reversal in fortune coincides almost perfectly with steered borrowers' disappearance of higher refinancing rates.

Our results about the refinancing and default patterns over the cycle offer a new perspective on the costs of predatory lending. From the borrowers' viewpoint, exotic mortgage features appeared appealing at first because they allowed lower mortgage payments in the short run, and thus borrowers performed better on average. These very features encourage borrowers to refinance their mortgages (and thus let lenders target them a second time to collect origination fees) when house values increase. However, when market prices stall, borrowers cannot refinance anymore and are trapped in contracts with rising payments and lower equity values, leading to a default. Our results thus mirror Gennaioli, Shleifer, and Vishny (2012) who argue that financial innovation introduces hidden risks into financial products exposed during a crisis. The main message of our study is that predatory lending practices may appear innocuous during good times but reveal their true costs when the market trends downwards.

Our study directly adds to the literature about predatory lending, mortgage fraud, and financial literacy. Most of the literature discusses predatory lending either at a theoretical level or at an empirical level, however, in aggregate. Eggert (2001), Engel and McCoy (2001). Engel and McCoy (2006), and Reiss (2005) define predatory lending, discuss the extent of the phenomena, and hypothesize about its potential sources. Several studies examine the effects of anti-predatory laws on the supply of credit and the mortgage performance (e.g., Elliehausen and Staten, 2004; Harvey and Nigro, 2004; Ho and Pennington-Cross, 2006; Bostic, Engel, McCoy, Pennington-Cross, and Wachter, 2008; Mayer, Pence, and Sherlund, 2009; Bostic, Chomsisengphet, Engel, McCoy, Pennington-Cross, and Wachter, 2012; Agarwal et al., 2014; Di Maggio, Kermani, and Korgaonkar, 2019). A closely related set of work assesses which groups are more susceptible to mortgages with predatory loan terms (see Immergluck and Smith, 2003; Bocian, Ernst, and Li, 2008; Gurun, Matvos, and Seru, 2016; Agarwal, Ambrose, and Yao, 2020a).

Predatory lending is part of a broader set of mortgage fraud practices that flourished during the boom years. Ben-David (2011) documents that borrowers inflated home prices to borrow greater amounts from lenders. Garmaise (2015), Jiang, Nelson, and Vytlacil (2014), Piskorski, Seru, and Witkin (2015), and Griffin and Maturana (2016) report systematic misreporting in mortgage applications. In addition, there are studies about the importance of financial literacy and financial advice in borrower decision-making. Agarwal, Ben-David, and Yao (2017) show that borrowers' decisions regarding mortgage features (mortgage points) are almost arbitrary, likely because of poor financial literacy. Agarwal, Amromin, Ben-David, Chomsisengphet, and Evanoff (2020b) find that providing counseling sessions attempting to warn borrowers from risky mortgages or predatory lending does not change their choice. Agarwal, Amromin, Ben-David, Chomsisengphet, and Evanoff (2010) find that long-term financial education programs help households make better financial decisions and perform better on loans.

More broadly, our paper contributes to the growing literature that finds evidence linking the real estate bubble in the early 2000s to misaligned incentives of intermediaries. Keys, Mukherjee, Seru, and Vig (2010) show that securitization led to lax screening by lenders, and Agarwal, Ben-David, and Yao (2015) find that appraisers cave to pressures from borrowers and lenders and inflate home prices. Agarwal and Ben-David (2018) find that volume-based incentives to intermediaries worsen the information asymmetry in lending.

# 2 Background, Hypothesis Development, and Empirical Design

## 2.1 Background: What is Mortgage Steering?

Steering is a well-known term in the industry, which means an intermediary or a seller pushes a particular product that may or may not be optimal for the customer. Traditionally, in the real estate world, regulators focused on market steering by realtors, who might restrict neighborhoods shown to certain potential home buyers. Such behavior can result in taste-based or statistical discrimination and distort the spatial patterns of housing demand by white and minority homebuyers in such a way as to perpetuate neighborhood segregation (Ondrich, Ross, and Yinger, 2003). Such practices are illegal based on the Fair Housing Amendments Act of 1988 and numerous state laws.[6]

During the housing boom of the 2000s, a different form of steering in housing markets—namely, credit steering—emerged. Here, the borrowers would be encouraged to obtain credit from a particular lender or in a particular type of mortgage contract. Engel and McCoy (2001) discuss how lenders may steer prime borrowers into high-cost mortgages. Also, Mac (1996) finds that in the early 1990s, 10% to 35% of subprime borrowers had credentials that should have qualified them for prime loans, and Barr (2005) argues that some subprime borrowers "may have been steered to higher cost lenders." Such behavior could be helpful for borrowers if it made it possible to obtain credit they may not have otherwise received and if that credit was accurately priced based on their credentials. However, credit steering could also be a predatory lending strategy. The concern is that the lender may not have the borrower's best interest in mind and may "gouge" them—whether through higher interest rates, excess fees, or contract features that increase the value of the loan to the originator

[6]Steering practices also exist in other domains, e.g., brokers pushing consumers into high-fee mutual funds (Egan, 2019).

but that may be unnecessary or non-transparent to the borrower.[7] In our empirical part, we impose no ex-ante value judgment on the value of customer steering and let the data speak.

## 2.2 Hypotheses Development

Mortgage steering is likely to occur at the first application a potential borrower makes. However, it is difficult for an outside observer to identify mortgage choices resulting from steering because of two major hurdles. First, one needs to separate cases in which lenders steered borrowers into a product from cases in which borrowers expressed demand for the product. Second, assessing the optimality of a selected product is problematic in itself, as the econometrician does not observe the complete set of borrower characteristics and constraints. An ideal empirical setting to detect steering activity would be to observe borrowers demanding one product and measure whether lenders concur or try to market a different product with unambiguously inferior features for the borrower.

Given the absence of transaction-level negotiation data, we developed a novel methodology to identify loans that were likely steered. Specifically, we argue that borrowers whose mortgage applications were rejected by a lender but were approved shortly thereafter by the same lender or its close affiliate are more likely to have been steered into a suboptimal mortgage product. We test our conjecture by comparing this group of borrowers to a group of observationally similar borrowers whose application was initially rejected but whose subsequent application was approved by lenders unaffiliated with the original lender.[8]

As an illustrative example, suppose a borrower enters a lending institution seeking a mortgage, and their loan application is evaluated. Their application is rejected outright if judged to be a poor credit risk. However, if their credit risk is acceptable, they might still

[7] Renuart (2004) argues that steering may have played a larger role in mortgage rate determination than did borrower risk.

[8] Withdrawn applications may offer another path to steering. In the current version of the paper, we do not analyze this possibility since the "applications withdrawn" reporting code in HMDA is notoriously difficult to implement consistently. For instance, the borrower has to request a withdrawal before a credit decision is made expressly. Simply stopping communications does not qualify as withdrawal, yet some lenders report it as such.

be told that they do not qualify for the specific loan they applied for but might qualify for another mortgage product. If steering is to occur, it would be initiated once the lender has determined that the loan applicant is an acceptable credit risk.

In such cases, a loan officer's job is to determine if a borrower can be convinced to take an alternative loan product—one that enhances the loan officer's compensation and/or the risk-adjusted profitability of the organization. In making this decision, the officer has to consider the risk of the borrower rejecting an alternative offer and seeking credit elsewhere. Consequently, this decision would be influenced by the perceived financial sophistication of the loan applicant.[9] The likelihood of rejecting an alternative offer might also be affected by how quickly that offer is presented to the applicant following the initial rejection.

This example relates to anecdotal evidence we gathered in written and verbal correspondence with mortgage loan officers active during the pre-GFC housing boom. At least some mortgage originators actively managed the type of contracts for which their clients would be approved, whether or not they applied for those contracts in the first place. These originators engaged in extensive manual exceptions and strategically timed presentation of alternative offers.

The above description can be used to develop our hypotheses. First, steered borrowers are likely to be less financially sophisticated, e.g., come from certain socioeconomic groups with weaker financial literacy backgrounds. Second, steered borrowers take loan products that are considered to have high profit margins for mortgage lenders (e.g., prepayment penalty, option ARM) and carry higher interest rates than non-steered loans. Third, steered loans are sold to private-label securitizers, who pay a high fee for exotic loan products with said features to monetize these margins. Finally, if steered loans maximize originator profits instead of fitting borrower credit needs, we anticipate their performance to differ.

[9]For most mortgage loans, not just steered loans, there would be asymmetric information advantages for the lender. The lender operates daily in the mortgage markets and is closely aware of the matching of customer credit qualifications and alternative mortgage products. Many borrowers do not follow the mortgage markets nearly as closely nor understand the credit-qualification-to-product matches. However, the lending officer who intends to steer the applicant inappropriately would be looking for applicants with a below-average level of financial sophistication.

## 2.3 Research Design

To implement the identification strategy, we focus on a subset of lenders who are organized under bank holding companies (BHCs), and thus are likely to be more closely affiliated with each other. For these lenders, we can observe the original borrower demand in the form of a mortgage application. Since we cannot identify the steered borrowers directly in the data, we develop an algorithm to detect steering. To do this, we consider mortgage loan applications that are denied by one lender only to be approved within a relatively short period without material changes in the key observable loan application variables.[10] Instances in which the approving and rejecting lenders are the same bank (or its close affiliate) are tagged as 'steered.' These borrowers form our steered group. The borrowers originally rejected but later approved by an unaffiliated lender fall into the group of potential controls. To make these two groups comparable, we use several approaches to construct matched samples that achieve tight covariate balance in terms of a wide array of observable borrower characteristics.[11] The comparability of these two groups is further enhanced by the requirement that both comprise rejected applications approved within a short period thereafter.

Next, we evaluate observed demographic characteristics (not used in sample construction) in these two groups to gauge whether specific borrowers were more likely to have been successfully kept by the original rejecting lender. Finally, we analyze whether there are meaningful differences in outcomes between the two groups; these include APR on the mortgage, the type of mortgage and various mortgage characteristics granted, and mortgage performance captured by the refinancing propensity and delinquency rate.

It is important to emphasize that we are not attempting to identify all instances of credit steering. The focus here is on one specific practice and one specific group of mortgage appli-

[10] As described in Section 3.3, we attempt to ensure that these pairs of applications are by the same applicants and are backed by the same property by requiring very tight matches on a set of applicant, loan, and property characteristics.

[11] We elaborate on the mechanics of constructing matched control samples in Section 3.2.

cants. All of them get turned down for credit but are approved shortly thereafter, strongly suggesting that their credit profile at the time of original application was not disqualifying. Next, we describe our data and methodology in more detail.

# 3 Data, Coverage across Data Sets, and Descriptive Statistics

## 3.1 Data Sources

We identify steered and non-steered loan samples based on the Home Mortgage Disclosure Act (HMDA) data. This source provides the loan application date, the date that a decision is made on the application, and the type of decision made (e.g., accept or reject the loan application). The HMDA dataset provides limited information on affiliation structure, the qualifications of the borrower, or (if a loan originated) the characteristics of the loan. Therefore, we obtain additional information from mortgage servicing sources, the Bank Holding Company Structure files, and Bank Call Reports.

McDash Analytics (McDash) provides loan-level information collected from residential mortgage servicers on loans packaged into government agency and non-agency mortgage-backed securities and loans held on lenders' portfolios. The McDash data provides extensive information about the loan, property, and borrower characteristics at the time of mortgage origination. Property-related variables include appraisal amount, geographic location, and property type (single-family residence, condo, or other type of property). Loan characteristics include origination amount, term to maturity, lien position, loan type (i.e., whether or not the loan is conventional), loan purpose (purchase or refinance), and the coupon rate on the mortgage. Credit-risk-related variables include the borrower's debt-to-income ratio, FICO credit score, loan-to-value (LTV) ratio at origination, and the level of documentation provided. The McDash data coverage has grown over time, including 9 of the top 10 mortgage servicers by 2003. Since servicers only provide information on active loans when they start reporting data to McDash, the McDash database includes relatively few loans that originated in the late 1990s and the early 2000s.

Two caveats are in place. First, due to data limitations (McDash) we only consider banks and ignore potential steering to and from credit unions, savings and loans, and mortgage companies.[12] Second, McDash Analytics data contains a smaller share of sub-prime loans; hence, the effects we document exist in the prime market and among highly regulated lenders. These results are likely to represent conservative estimates of the steering behavior consequences.

Beyond the McDash information available at origination, the dataset also contains dynamically updated loan information, enabling one to monitor refinancing activity and loan performance. Variables of interest include interest rates (which change for adjustable-rate mortgages (ARMs) and have the potential to change with loan modifications), delinquency status (current, 31–60 days delinquent, 61–90 days delinquent, over 91 days delinquent, foreclosure, real estate owned by the lender (REO), or paid off), investor type (held in portfolio, private securitization, or "public" securitization via the housing GSEs),[13] and the actual unpaid principal balance as well as the scheduled principal balance if the borrower pays according to the original terms of the loan.

## 3.2 Sample Construction

To identify the set of loans to study, we start with HMDA loan application data for 1998–2006. The HMDA data encompass nearly all mortgage lending activity each year, with some exceptions for small and rural institutions that do not fall under the mandatory filing

[12] Demyanyk and Loutskina (2016) emphasize the importance of mortgage companies in originating riskier mortgages during this period under HMDA. Mortgage reporting may be done by the mortgage company or the money-center bank that acquires the loan, often under a standing contract. We might capture the latter type of transaction as "steered" but will likely miss the former.

[13] The public securitizations can be through Government National Mortgage Association (Ginnie Mae), Federal National Mortgage Association (Fannie Mae), Federal Home Loan Mortgage Corporation (Freddie Mac), Ginnie Mae via buyout loans, Local Housing Authority, or Federal Home Loan Banks).

requirements. Since the HMDA dataset includes the exact action taken and the date of that action for each application, we can determine whether a withdrawal or denial precedes the origination of a nearly identical loan by either the same or a different lender in the same U.S. Census tract. To develop our sample, we impose rather strict criteria on pairs of applications. These applications are allowed a difference in action date of no more than 60 days. They are required to match on applicant race, applicant sex, loan type (conventional or backed by the Federal Housing Administration (FHA) or administered by the U.S. Department of Veterans Affairs (VA)), loan purpose, Census tract, and occupancy type.[14] We also match iteratively on the loan amount and applicant income—by identifying and removing the sample pairs with no difference in amount or income and then increasing the window by $1,000 and matching again. We continue this process up to a maximum differential of $5,000.[15] This matching process produces approximately 3.4 million unique pairs of loan applications. Each is denied at first, but is subsequently approved within a short time window and without substantial changes in application data.

To determine whether a relationship exists between the two original (rejecting) and ultimate (approving) lenders, we match the HMDA lender identifier for each application to its highest holder (i.e., the highest bank holding company) in the BHC Structure data and Call Reports. Following this merge, the sample size declines to 1.35 million records of which 244,621 are loans originated by the original rejecting institution or lenders affiliated with it (i.e., 'steered').[16]

Since HMDA data do not include information on key risk characteristics of the borrower (such as the FICO score), loan terms, or loan performance, we match the originated loan in each pair of applications to mortgage-level data from McDash, which collect loan characteristics at origination from mortgage servicers and track the performance of these loans over

[14] Results were robust when a slightly shorter or longer timeframe was used.

[15] The thought is that the borrower may receive a slightly different loan amount or report a marginally different income based on the interaction with the initial lender.

[16] Due to proprietary data restrictions, the process of merging HMDA and mortgage servicer data requires replacing lender identifiers with randomly generated numbers. Thus, while the resulting analysis can incorporate lender-fixed effects, including lender-specific characteristics is not feasible.

time. The approved HMDA loan applications in our sample are matched to the mortgage-level data on the origination date, zip code, loan amount, loan type, loan purpose, occupancy type, and lien. This step substantially reduces the sample size, as McDash data do not have universal coverage, and mortgage servicer data (particularly information on loan origination dates) may not coincide with the regulator-collected data. Moreover, as the servicer data are concentrated in the latter part of our HMDA sample, the merged dataset becomes heavily weighted towards the 2003–2006 period (over 98% of observations are in this period). We end up with 303,368 unique loan originations, of which 90,349 fit the definition of a 'steered' transaction.

Next, we create two control samples (non-steered sample). Both control groups consist of borrowers whose applications were also initially denied (potentially in an attempt to steer), but then approved within a short time by another lender not affiliated with the holding company that originally denied the loan. The samples differ in the technique used to match them to the steered sample.

The first control sample is a propensity score matched (PSM) sample. Specifically, we perform a nearest neighbor propensity score match, with each loan in the steered sample cutoff matched with replacement to a similar non-steered loan. The match criterion is the conditional treatment probability from a logit model, where the independent variables include the log income, the log home value, FICO score at origination, and loan-to-value (LTV) at origination. We require the potential non-steered loans to be in the same state, originated within 90 days, be issued for the same purpose (purchase or refi), have the same occupancy status (owner or investor), and be of the same type (conventional or FHA) as a given steered loan. From the resulting sample of potential controls, we choose a loan with the smallest difference in the propensity score, subject to an absolute threshold of 0.05. The resulting propensity-score matched sample contains 71,682 steered loans and an equivalent number of non-steered loans.[17]

[17]The more lenient PSM approach generates a larger sample but also increases the possibility of pairwise mismatches in steered and non-steered loans.

The second control group is a strict matching (SM) sample based on each characteristic. That is, for each steered loan, we find a non-steered counterpart that is very close in each of the following: applicant income, loan amount, FICO score, LTV ratio, and origination date, while matching exactly on loan purpose, loan type, occupancy type, and state. We require that applicant's income and loan amount be within 25%, FICO score within 25 points, LTV ratio within 5 percentage points, and the origination date within 90 days. Not surprisingly, this approach results in a smaller final sample of 13,252 steered loans and 13,252 non-steered loans.

In Table 1, we also present t-tests for differences in means between groups. Ideally, we would have non-significant t-statistics for differences of each of the matched variables (the first seven variables). Unfortunately, some differences are significant. In the propensity-score matched sample, FICO, LTV, and loan amounts are higher for the steered group. In the strict-matched sample, steered borrowers have higher incomes and loan amounts. While these differences are statistically significant and point to less-than-perfect matching, some mitigating aspects exist. First, some significant differences (e.g., FICO score) are economically negligible. Second, the steered and non-steered groups are different along different dimensions across the two matching techniques, yet the empirical results shown later are very similar. Third, the direction of the biases varies; for some variables, the steered borrowers are more leveraged or borrow larger amounts, while other variables indicate that they may be more financially stable, e.g., have higher income.

In addition to the data sources discussed above, we use the CoreLogic Home Price Index (HPI) to compute local changes in home prices. HPI data are available at the zip code level for 57.3% of the U.S. population. For observations for which zip-code-level data are unavailable, we use data at the Core Based Statistical Area (CBSA) level, which are available for 83.9% of the U.S. population. Finally, we use the 2000 Census to identify census tracts in the low-to-moderate income (LMI) category.[18]

[18]LMI areas are defined as those census tracts in which the median family income is less than 80 percent of the area median income.

## 3.3 Descriptive Statistics

Table 1 presents summary statistics for the resulting pair of steered and non-steered samples. The left-hand panel presents characteristics of the propensity score matching approach, and the right-hand panel is based on the strict matching approach.

### Table 1. Summary Statistics

The table provides summary statistics for the analysis used in the study. The first sample is based on a propensity-matching algorithm of loans rejected by one lender and eventually approved by an affiliate. The second sample is based on exact matching criteria. See Section 3 for details on data sources and sample construction. Data sources: Home Mortgage Disclosure Act (HMDA), 1998–2006; McDash Analytics; CoreLogic. Variable definitions are provided in Appendix A.

| | Propensity Score Matching | | | | | Strict Matching | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Steered | | Non-steered | | | Steered | | Non-Steered | | |
| Variables | $N = 71,682$ | | $N = 71,682$ | | | $N = 13,252$ | | $N = 13,252$ | | |
| | Mean | StDev | Mean | StDev | $t$-Diff | Mean | StDev | Mean | StDev | $t$-Diff |
| Matching variables | | | | | | | | | | |
| FICO score | 711.2 | 49.0 | 708.7 | 59.6 | 8.7 | 709.2 | 51.8 | 709.0 | 52.5 | 0.3 |
| LTV ratio (%) | 68.8 | 21.6 | 65.8 | 22.2 | 26.5 | 70.7 | 20.4 | 70.8 | 20.3 | −0.5 |
| Borrower income ($1,000s) | 124.5 | 97.2 | 124.8 | 100.7 | −0.7 | 83.5 | 74.3 | 74.7 | 51.5 | 11.2 |
| Loan amount ($1,000s) | 277.2 | 205.1 | 262.7 | 199.9 | 13.6 | 185.1 | 139.8 | 177.5 | 132.2 | 4.5 |
| I(Refinancing) | 0.41 | 0.49 | 0.41 | 0.49 | 0.0 | 0.58 | 0.49 | 0.58 | 0.49 | 0.0 |
| I(Owner-occupied) | 0.81 | 0.39 | 0.81 | 0.39 | 0.0 | 0.95 | 0.22 | 0.95 | 0.22 | 0.0 |
| I(Conventional (non-FHA)) | 1.00 | 0.07 | 1.00 | 0.07 | 0.0 | 0.99 | 0.08 | 0.99 | 0.08 | 0.0 |
| Outcome variables of interest | | | | | | | | | | |
| Initial interest rate (%) | 6.96 | 1.32 | 6.59 | 1.98 | 42.3 | 6.73 | 1.35 | 6.44 | 1.58 | 15.9 |
| 90-day delinquency w/ 2 years | 0.063 | 0.243 | 0.077 | 0.266 | −10.4 | 0.043 | 0.202 | 0.048 | 0.213 | −1.9 |
| I(Interest Only) | 0.32 | 0.47 | 0.16 | 0.37 | 69.2 | 0.27 | 0.45 | 0.09 | 0.29 | 39.0 |
| I(Option ARM) | 0.38 | 0.49 | 0.16 | 0.37 | 96.3 | 0.27 | 0.44 | 0.08 | 0.27 | 40.7 |
| I(Pre-payment penalty) | 0.41 | 0.49 | 0.20 | 0.40 | 89.0 | 0.28 | 0.45 | 0.15 | 0.36 | 26.6 |
| I(No/Low documentation) | 0.82 | 0.39 | 0.67 | 0.47 | 64.0 | 0.80 | 0.40 | 0.72 | 0.45 | 15.3 |
| Fixed rate term (months) | 75.7 | 99.9 | 204.0 | 149.9 | −149.9 | 112.5 | 126.5 | 241.6 | 138.7 | −79.2 |
| Mortgage maturity (months) | 340.1 | 66.4 | 339.8 | 68.9 | 0.8 | 333.2 | 68.4 | 328.8 | 72.9 | 5.1 |
| I(Portfolio loan) | 0.01 | 0.11 | 0.17 | 0.38 | −109.9 | 0.04 | 0.20 | 0.16 | 0.36 | −31.6 |
| I(GSE securitization) | 0.29 | 0.45 | 0.38 | 0.48 | −36.6 | 0.44 | 0.50 | 0.54 | 0.50 | −16.4 |
| I(Private-label securitization) | 0.70 | 0.46 | 0.44 | 0.50 | 100.6 | 0.52 | 0.50 | 0.30 | 0.46 | 36.8 |
| Other covariates | | | | | | | | | | |
| ΔHPI 12-month pre-origination | 0.140 | 0.104 | 0.139 | 0.106 | 1.8 | 0.109 | 0.096 | 0.107 | 0.095 | 2.2 |
| I(African-American) | 0.06 | 0.23 | 0.06 | 0.23 | 1.2 | 0.06 | 0.24 | 0.06 | 0.25 | −1.7 |
| I(Hispanic) | 0.17 | 0.38 | 0.15 | 0.36 | 10.3 | 0.12 | 0.33 | 0.13 | 0.33 | −1.5 |
| I(Female) | 0.32 | 0.47 | 0.25 | 0.43 | 30.8 | 0.34 | 0.47 | 0.26 | 0.44 | 13.2 |
| I(No co-signer) | 0.68 | 0.47 | 0.57 | 0.50 | 43.7 | 0.69 | 0.46 | 0.57 | 0.50 | 20.9 |
| I(Low/Moderate Income) (tract) | 0.30 | 0.46 | 0.27 | 0.44 | 13.9 | 0.31 | 0.46 | 0.30 | 0.46 | 3.2 |

The summary statistics show that the propensity-score matching procedure matched steered and non-steered matched observations well. By construction, this sample is designed to minimize the joint differences on a limited set of observable characteristics. Yet, the summary statistics for the propensity-score matched sample displayed in the upper left-hand panel of Table 1 suggest that the means and standard deviations of each continuous variable used in PSM are very similar for the steered and non-steered samples. It is worth noting that the average FICO scores in our sample are around 710, and the average first-lien LTV ratios at origination are under 70 percent. In other words, the borrowers in our sample do not match the profile of a subprime borrower purchasing (or refinancing) their house with the minimum amount of equity possible. Over 80% of mortgages in the PSM sample are owner-occupied, and most (59%) are used for home purchases.

However, achieving a tight covariate balance in observables through matching still produces considerable variation in the means of the outcome variables, listed in the lower panel of Table 1. The steering hypothesis suggests that the 'steered' group is charged a higher interest rate and has better ex-post-credit quality than the control group. Indeed, we see that borrowers in this group have higher average interest rates (6.96% versus 6.59%), while experiencing lower unconditional average rates of default (6.3% versus 7.7%).[19] These differences are statistically as well as economically significant. Furthermore, we also observe sizable differences in propensities to originate loans with certain contract features between the two groups. A much higher fraction of the steered group loans are option ARM (38% versus 16%) or interest-only mortgages (32% versus 16%), and carry prepayment penalties (41% versus 20%).

For the strict-matching sample, the results are fairly similar, although the resulting sample is much smaller. As with the propensity-score matching sample, the key covariates are closely matched between the steered and non-steered samples. The comparison of outcome

[19]The initial or first observed APR is the interest rate reported six months after the loan was originated. This allows us to avoid capturing initial teaser rates commonly offered on certain loan contracts but typically lasting only one month.

variables between the groups is also similar to that in strict-matching sample. The steered group has higher average interest rate, lower realized delinquency rates, and higher rates of incidence of high-margin mortgage products (option ARMs, IO loans, and loans with prepayment penalties). It is worth noting that relying on the strict-matching procedure generates a sample that contains a smaller fraction of non-amortizing mortgage contracts, such as interest-only loans or option ARMs. Amromin, Huang, Sialm, and Zhong (2018) show that such contracts were common among relatively high-income borrowers purchasing more expensive homes that defaulted at high rates. The difference in the relative performance of such contracts between the two sample design approaches accounts for relative differences in income, loan amount, and default rates in the left and right panels of Table 1. In Section 4.4 we discuss performance patterns in greater detail.

# 4 Empirical Results

## 4.1 Regression Specification

Once we identify a sample of steered borrowers (i.e., those whose application is approved by the original lender or its affiliate), we conduct a cross-sectional regression analysis evaluating borrower and loan contract characteristics to determine whether that group of borrowers differs from the control group. In doing so, we control for various factors, including various fixed effects. The regressions used in most tables use the following specification:

$$\begin{aligned} Response_i &= \alpha + \beta I(SteeredBorrower)_i + \delta BorrowerControls_i \\ &\quad + \theta MortgageControls_i + \gamma FixedEffects_i + \varepsilon_i \end{aligned} \tag{1}$$

where *Response* is the loan-level response variable, such as the interest rate on mortgages, default status of loans, etc.; $I(SteeredBorrower)$ is an indicator variable for the whether a loans was steered (or, equivalently, whether the eventual lender is the same as the original one); $BorrowerControls$ are a set of borrower characteristics including logged borrower income, and the FICO credit score of the borrower (splined into the ranges: 621–660, 661–720, 721–760, and $> 760$). $MortgageControls$ are a set of loan-specific characteristics, which include the following variables: logged loan amount, LTV ratio at origination (splined into 80%–89%, 90%–99%, and $\geq$ 100%), binary indicators of various contract types (amortizing ARM, option ARM, IO), refi indicator, prepayment penalty indicator, owner-occupier indicator, conventional mortgage indicator, and no/low documentation indicator. In addition, we control for the 12-month change in the zip-level house price index. Appendix A provides detailed variable descriptions. $FixedEffects_i$ account for either fixed effects for the state interacted with calendar quarter, or fixed effects for each pair of matched steered and non-steered loans. We double-cluster standard errors in all regressions at the state and calendar quarter levels. In most tables, we present the propensity-score matched sample forms in Panel A, and the regression results for the strict matching approach in Panel B.

## 4.2 Who is Successfully Steered?

We begin the analysis by examining the demographic characteristics of borrowers from affiliated lenders relative to the unaffiliated group. We rely on (partial) demographic information and precise geographic location captured in HMDA to answer the question of which borrowers were more likely to be steered. In particular, we can use data on borrower's gender, identification as African-American or Hispanic, indicator of not having a co-applicant and indicator of a loan being secured by a property in a low or moderate-income census tract.[20] Since steering means that borrowers are taking an inferior product relative to what they can get otherwise, we expect that borrowers from affiliated lenders share characteristics linked to lower levels of financial sophistication.

[20]Before 2004, HMDA required respondents to choose among six racial or ethnic classifications. In 2004, the reporting rules separated questions on ethnicity (Hispanic or non-Hispanic) and race (white, African-American, Asian, American Indian and Alaska Native, Hawaiian or other Pacific Islander). This creates potential problems with making race and ethnicity classifications consistent over the two periods. A related problem arises with determining race and ethnicity in records where either of the two fields is missing. We follow the (Avery, Brevoort, and Canner, 2007, pp. 361–362) approach to addressing this issue.

We start with a set of steered and PSM-matched non-steered loans. By construction, this set is evenly split between steered and non-steered loans. More importantly, its construction ensures that each loan pair is closely matched on a set of key loan and borrower characteristics.[21] For this set of loans, we estimate the likelihood of being steered as a function of HMDA variables, absorbing a set of fixed effects as in the earlier tables. Our preferred method employs the linear probability model, given the large number of fixed effects in some specifications.

The OLS results are shown in Table 2. Starting with Panel A, Column (1), we find that all else equal, Hispanic applicants had a somewhat higher propensity to be steered (no effect for African-American applicants). The results are weaker once we control for borrower and mortgage characteristics (Column (2)). Female applicants and applicants who did not have a co-borrower were much more likely to be steered towards more expensive loans. We also found applicants residing in LMI census tracts considerably more likely to be steered. The magnitudes of the estimated coefficients are in the order of 0.03–0.20, suggesting, for instance, that borrowers with no co-signers are up to 20% more likely to be steered. In Columns (3) and (4), we introduce dummies for each matched pair, thus exploiting differences in characteristics within matched pairs. All demographic variables appear as strong predictors of the likelihood of being steered. In Panel B, we present the results from the strict matching procedure. Here, the results on the dummies for female applicants, no co-signer, and low/moderate income are as before. The results, however, for minorities are different: we find no significant results for Hispanic or African-American applicants.

These results are broadly aligned with existing empirical evidence on which population subgroups display lowest levels of financial literacy. A literature survey by Lusardi and Mitchell (2014) highlights substantial shortfalls in financial literacy among certain groups.

[21] Recall that the PSM algorithm conditions on borrower income, loan amount, FICO score, and LTV at origination. It also requires an exact match loan purpose and type, occupancy status, state, and application date within 90 days of the steered loan.

### Table 2. Borrower Characteristics and Lender Affiliation

The table presents regressions of the lender affiliation indicator on borrower personal and area characteristics and various fixed effects described in the text. The sample is constructed using propensity-score matching (Panel A) and a strict matching algorithm (Panel B). All regressions are OLS regressions. Standard errors are double-clustered by calendar month and state of origination. Data sources: Home Mortgage Disclosure Act (HMDA), 1998–2006; U.S. Census. $t$-statistics are reported in brackets. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively. See Section 3 for details on data sources and sample construction. Variable definitions are provided in Appendix A.

**Panel A: Propensity-Score Matching Sample**

| Dependent variable: | I(SteeredBorrower) | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| I(African-American) | −0.013 | −0.012 | −0.020 | −0.008 |
| | [−0.77] | [−0.80] | [−0.43] | [−0.21] |
| I(Hispanic) | 0.036*** | 0.003 | 0.073** | 0.001 |
| | [3.04] | [0.22] | [2.08] | [0.05] |
| I(Female) | 0.062*** | 0.043*** | 0.121*** | 0.065*** |
| | [14.43] | [7.96] | [7.11] | [7.82] |
| I(No co-signer) | 0.101*** | 0.030*** | 0.205*** | 0.045*** |
| | [9.33] | [5.30] | [6.38] | [3.48] |
| I(Low/Moderate Income) | 0.048*** | 0.018 | 0.104*** | 0.035 |
| | [4.77] | [1.44] | [3.60] | [1.25] |
| I(Refinancing) | | 0.041*** | | |
| | | [3.05] | | |
| I(Owner-occupier) | | 0.066*** | | |
| | | [3.40] | | |
| I(Conventional (non-FHA)) | | −0.279*** | | |
| | | [−8.13] | | |
| Borrower and mortgage characteristics | No | Yes | No | Yes |
| State × Qtr fixed effects | Yes | Yes | No | No |
| Matched pair fixed effects | No | No | Yes | Yes |
| Observations | 133,011 | 124,011 | 133,011 | 124,011 |
| Adjusted $R^2$ | 0.026 | 0.271 | −0.928 | −0.171 |

In particular, the young and the old, women, minorities, the least educated and those with lower incomes all display markedly lower levels of financial sophistication. Consistently with these results, a recent study of the Urban Institute finds that mortgage applications by single women are more likely to be denied.[22] These results also resonate with Morton, Zettelmeyer, and Silva-Risso (2003) who document that women, minorities, and the elderly pay more, on average, for cars. By and large, these are the groups identified as more likely to be steered

[22]Goodman, Laurie, Jun Zhu, and Bing Bai, We're Still Shortchanging Women When It Comes to Mortgages, *Urban Wire*, 8 September 2016.

**Table 2. Borrower Characteristics and Lender Affiliation (Cont.)**

**Panel B: Strict-Matching Sample**

| Dependent variable: | I(SteeredBorrower) | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| I(African-American) | −0.054** | −0.049*** | −0.018 | −0.110* |
| | [−2.22] | [−2.80] | [−1.28] | [−1.70] |
| I(Hispanic) | −0.011 | −0.016 | 0.010 | −0.034 |
| | [−0.60] | [−0.64] | [0.81] | [−0.50] |
| I(Female) | 0.051*** | 0.031*** | 0.018*** | 0.041 |
| | [14.49] | [4.73] | [3.58] | [1.47] |
| I(No co-signer) | 0.122*** | 0.046*** | −0.007 | 0.083*** |
| | [8.86] | [5.26] | [−0.69] | [4.28] |
| I(Low/Moderate Income) | 0.043*** | 0.019* | −0.001 | 0.047 |
| | [4.27] | [1.84] | [−0.12] | [1.40] |
| Borrower characteristics | No | Yes | Yes | Yes |
| Mortgage characteristics | No | Yes | Yes | Yes |
| State × Bank × Qtr fixed effects | No | No | Yes | No |
| Matched pair fixed effects | No | No | No | Yes |
| Observations | 24,047 | 17,618 | 17,618 | 17,618 |
| Adjusted $R^2$ | 0.021 | 0.271 | 0.660 | −0.131 |

by their mortgage lender.

Overall, the results suggest that female borrowers, borrowers without a co-signer, and borrowers residing in low- and moderate-income areas were most likely steered towards borrowing from an affiliate lender. This result supports the mechanism we proposed earlier in Section 2.2. Specifically, lenders are more likely to steer applicants with lower levels of financial sophistication to minimize the risk that rejected but qualified borrowers shop around and end up with a different lender. Furthermore, existing research suggests that these populations might be less informed about credit markets in general and thus might be more likely to be vulnerable to lender steering practices (Berndt et al., 2016).

## 4.3 Characteristics of Steered Mortgages

To further understand whether the group of borrowers from affiliated lenders were steered to inferior products, we examine the characteristics of the mortgage products they took and compare them to those taken by non-borrowers from affiliated lenders.

### 4.3.1 Interest Rate

A central part of the steering hypothesis is that borrowers from affiliated lenders are led to take mortgage products that are more profitable to the originator. The most direct measure of loan profitability is risk-adjusted interest rate.

In Table 3, Panel A, we report the results of regressing the mortgage APR on the variable of interest—steered indicator—as well the other control variables and fixed effects as described in Section 4.1, using the propensity-score matched sample. The regressions show that borrowers from affiliated lenders pay interest rates up to 72 basis points higher than those of similar but non-borrowers from affiliated lenders.

The most parsimonious specification in Column (1) indicates an estimated interest rate differential of 39 basis points after soaking up the effects of loan origination date and property location (state). Since mortgages of different contractual forms have substantial variation in their interest rate—owing to the term premium and the frequency of interest rate resets—it is especially important to account for loan characteristics. When we add such controls in Column (2), the estimated interest rate differential nearly doubles to 72 basis points. The magnitude of the effect is large both in absolute terms and relative to the mean interest rate of 6.59% in the control group. Columns (3) and (4) provide a tighter specification that includes pairwise fixed effects and produce an estimated differential of 69 basis points. Panel B finds similar magnitude results (up to 54 basis points) using the strict-matched sample.

## Table 3. Interest Rate Paid, by Lender Affiliation

The table presents regressions of the initial interest rate on lender affiliation indicator, as well as a variety of fixed effects and borrower and mortgage characteristics. Borrower and mortgage controls include: logged borrower income, FICO credit score of the borrower (splined into the ranges: 621–660, 661–720, 721–760, and $> 760$), logged loan amount, LTV ratio at origination (splined into 80%–89%, 90%–99%, and $\geq 100\%$), amortizing ARM indicator, interest-only indicator, refi indicator, prepayment penalty indicator, owner-occupier indicator, conventional mortgage indicator, and no/low documentation indicator. All regressions are OLS regressions. Standard errors are double-clustered by calendar month and state of origination. Data sources: Home Mortgage Disclosure Act (HMDA), 1998–2006; U.S. Census. *t*-statistics are reported in brackets. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively. See Section 3 for details on data sources and sample construction. Variable definitions are provided in Appendix A.

**Panel A: Propensity-Score Matching Sample**

| Dependent variable: | Initial interest rate (%) | | | |
|---|---|---|---|---|
| Mean of control sample: | 6.79 | | | |
| | (1) | (2) | (3) | (4) |
| I(Steered Borrower) | 0.387*** | 0.721*** | 0.376* | 0.692*** |
| | [2.60] | [5.07] | [1.84] | [3.47] |
| Borrower characteristics | No | Yes | No | Yes |
| Mortgage characteristics | No | Yes | No | Yes |
| State × Qtr fixed effects | Yes | Yes | No | No |
| Matched pair fixed effects | No | No | Yes | Yes |
| Observations | 143,364 | 140,072 | 143,364 | 140,072 |
| Adjusted $R^2$ | 0.165 | 0.460 | 0.152 | 0.447 |

**Panel B: Strict-Matching Sample**

| Dependent variable: | Initial interest rate (%) | | | |
|---|---|---|---|---|
| Mean of control sample: | 6.44 | | | |
| | (1) | (2) | (3) | (4) |
| I(Steered Borrower) | 0.288*** | 0.540*** | 0.288** | 0.496** |
| | [2.89] | [4.29] | [2.17] | [2.32] |
| Borrower characteristics | No | Yes | No | Yes |
| Mortgage characteristics | No | Yes | No | Yes |
| State × Qtr fixed effects | Yes | Yes | No | No |
| Matched pair fixed effects | No | No | Yes | Yes |
| Observations | 26,503 | 19,758 | 26,503 | 19,758 |
| Adjusted $R^2$ | 0.198 | 0.428 | 0.405 | 0.452 |

### 4.3.2 Product Type

Next, we examine the type of mortgages and mortgage characteristics taken by borrowers indicatorged as steered, compared to similar borrowers in the non-steered group.

During the market run-up period of the early 2000s, many banks moved from a model of originate-to-hold to originate-to-distribute (Purnanandam, 2011; Bord and Santos, 2012). Thus, lenders were strongly incentivized to convince borrowers to take mortgages that generate high fees (Kolb, 2011). Thus, lenders were interested in originating loans that they could sell to Wall Street firms for a fee instead of quality loans that could be held on lenders' balance sheets (Keys et al., 2010).

In the analysis of product types, we select several mortgage types that are considered to have high-profit margins in the residential mortgage industry. The mortgage types that we study are interest-only mortgages, option ARMs (adjustable rate mortgages), mortgages with prepayment penalties, and no/low documentation mortgages. These features are not mutually exclusive except for interest-only mortgages and option ARMs. The choice of these variables is motivated by Engel and McCoy (2001) and by the available variables in the data. These mortgage products were suited to an economic environment in which the common wisdom is that property prices always increase (Shiller, 2017). Therefore, it made sense to take mortgages that offered minimal payment through flexible or no amortization or mortgages that offered teaser rates in the first few years (for instance, more than 80% of option ARMs in our sample came with very low initial teaser rates). The common strategy to avoid the reset of mortgage rates following the expiration of the teaser rates or onset of regular amortization was to refinance the mortgage,[23] which is, of course, advantageous to lenders who collect origination fees.

Interest-only loans are loans in which the borrower does not repay any of the principal amount for several years, thus lowering the monthly payment for a certain period. Option ARM mortgages are mortgages in which the borrower can decide about the monthly payment,

[23] E.g., Damon Darlin, Keep Eyes Fixed on Your Variable-Rate Mortgage, *New York Times*, July 15, 2006.

as long as it is equal to or above the minimum payment. The minimum payment is typically set below the interest servicing requirements, leading to negative amortization, i.e., borrowers accruing principal instead of repaying it. Lenders usually discontinue the optionality of the mortgage after a pre-specified number of periods, typically five years or less. Payments are re-calculated to allow full amortization over the remaining term (25 years $= 30 - 5$ in the example above). The optionality may also be terminated when the principal reaches a certain level, typically 110% or 125% of the original loan amount.[24] Mortgages with prepayment penalty are mortgages in which borrowers pay a penalty if they refinance the loan (repay the principal) earlier than scheduled. When they exist, prepayment penalties are typically set between 1 and 5 years. No/low documentation mortgages (also called stated-income mortgages) are mortgages in which borrowers need either none or limited documentation for their income.

We learn about the profitability of loan products from conversations with lenders in the industry. The information that these loan types are profitable also appears in written sources. Bowen, Jollineau, and Lougee (2014) cites the comments of the CEO of Washington Mutual (the largest mortgage originator at the time) from the 2004/Q3 conference call, where he says that the company focuses on high-margin mortgage products such as option ARM mortgages. A similar message is echoed in an article about the competition in the mortgage market.[25] Mortgages with prepayment penalties were Countrywide's favorite product since "...investors who bought securities backed by the mortgages were willing to pay more for loans with prepayment penalties...".[26] Steven Krystofiak, President of the Mortgage Brokers Association for Responsible Lending (MBARL), an advocacy group protecting consumers and the loan industry from outlandish and counterproductive loan programs, testified in 2006 in front of the Federal Reserve Board. He argued that banks originated increasing

[24]See detailed explanation of the mortgage types at `https://www.fdic.gov/consumers/consumer/interest-only/`.

[25]Ruth Simon and James R. Hagerty, Countrywide's New Scare, *Wall Street Journal*, October 24, 2007. Available at: `http://www.wsj.com/articles/SB119318489086669202`.

[26]Gretchen Morgenson, Inside the Countrywide Lending Spree, *New York Times*, August 26, 2007. Available at: `http://www.nytimes.com/2007/08/26/business/yourmoney/26country.html`.

amounts of stated-income (i.e., no/low doc) mortgages because they were selling them to securitizers for sizeable profit given the strong demand from Wall Street.[27]

The tests for the mortgage types are provided in Table 4. In both Panels A and B, there are eight regressions, where the dependent variables are indicators of whether the type of the mortgage is interest-only (Columns (1)–(2)), option ARM (Columns (3)–(4)), have a prepayment penalty (Columns (5)–(6)), or no/low documentation (Columns (7)–(8)). As in the previous tables, the specifications vary in their configuration of fixed effects. All specifications include controls for borrower and mortgage characteristics. Panel A presents results from the propensity-score matching sample and Panel B presents results from the strict-matching sample.

The results uniformly show that borrowers from affiliated lenders are more likely to take mortgages with features considered highly profitable in the mortgage industry. The magnitudes of the effect are very large. When considering the odd-numbered columns in Panel A, the results show that borrowers from affiliated lenders are 61% more likely to take interest-only loans (= 0.266/0.165) than borrowers in the non-steered group, 80% (= 0.129/0.161) more likely to take an option ARM mortgage, 71% (= 0.141/0.198) more likely to take a prepayment penalty mortgage, and 33% (= 0.219/0.671) more likely to take a no/low-documentation loan. The results in the even columns show almost identical results. The results in Panel B are even stronger due to the lower base rate.

#### 4.3.3 Securitization

Most mortgage loans in our sample (99%) were originated between 2003 and 2006. During this period, lenders increasingly originated mortgages to sell them to investment banks, which, in turn, packaged them into private-label mortgage-backed securities (PLS MBS) for capital-market investors (Mayer et al., 2009; Nadauld and Sherlund, 2013). According to the sources cited in Section 4.3.2, lenders originated mortgages with exotic features to satisfy

[27] Available at: `http://www.federalreserve.gov/secrs/2006/august/20060801/op-1253/op-1253_3_1.pdf`.

**Table 4. Mortgage Products, by Lender Affiliation**

The table presents regressions of indicators of mortgage type (interest-only, option ARM, prepayment penalty, and no/low documentation) on lender affiliation indicator, as well as a variety of fixed effects and borrower and mortgage characteristics. Borrower and mortgage controls include: logged borrower income, FICO credit score of the borrower (splined into the ranges: 621–660, 661–720, 721–760, and > 760), logged loan amount, LTV ratio at origination (splined into 80%–89%, 90%–99%, and ≥ 100%), amortizing ARM indicator, interest-only indicator, refi indicator, prepayment penalty indicator, owner-occupier indicator, conventional mortgage indicator, and no/low documentation indicator. All regressions are OLS regressions. Standard errors are double-clustered by calendar month and state of origination. Data sources: Home Mortgage Disclosure Act (HMDA), 1998–2006; U.S. Census. $t$-statistics are reported in brackets. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively. See Section 3 for details on data sources and sample construction. Variable definitions are provided in Appendix A.

**Panel A: Propensity-Score Matching Sample**

| Dependent variable: | I(Interest Only) | | I(Option ARM) | | I(Prepayment penalty) | | I(No/Low doc) | |
|---|---|---|---|---|---|---|---|---|
| Mean of control sample: | 0.165 | | 0.161 | | 0.198 | | 0.671 | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| I(Steered Borrower) | 0.266*** | 0.262*** | 0.129*** | 0.125*** | 0.141*** | 0.136*** | 0.219*** | 0.221*** |
| | [5.60] | [4.03] | [8.70] | [6.15] | [6.13] | [4.11] | [5.30] | [3.99] |
| Borrower and mortgage characteristics | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| State × Qtr fixed effects | Yes | No | Yes | No | Yes | No | Yes | No |
| Matched pair fixed effects | No | Yes | No | Yes | No | Yes | No | Yes |
| Observations | 143,364 | 143,364 | 143,364 | 143,364 | 143,364 | 143,364 | 143,364 | 143,364 |
| Adjusted $R^2$ | 0.158 | 0.144 | 0.241 | 0.204 | 0.158 | 0.144 | 0.241 | 0.204 |

**Panel B: Strict-Matching Sample**

| Dependent variable: | I(Interest Only) | | I(Option ARM) | | I(Prepayment penalty) | | I(No/Low doc) | |
|---|---|---|---|---|---|---|---|---|
| Mean of control sample: | 0.093 | | 0.082 | | 0.148 | | 0.719 | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| I(Steered Borrower) | 0.190*** | 0.197*** | 0.195*** | 0.195*** | 0.174*** | 0.184** | 0.135*** | 0.139*** |
| | [8.14] | [6.05] | [4.36] | [2.87] | [3.82] | [2.55] | [7.41] | [5.39] |
| Borrower and mortgage characteristics | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| State × Qtr fixed effects | Yes | No | Yes | No | Yes | No | Yes | No |
| Matched pair fixed effects | No | Yes | No | Yes | No | Yes | No | Yes |
| Observations | 20,164 | 20,164 | 20,164 | 20,164 | 20,164 | 20,164 | 20,164 | 20,164 |
| Adjusted $R^2$ | 0.163 | 0.150 | 0.208 | 0.170 | 0.235 | 0.182 | 0.068 | 0.037 |

the demand from Wall Street: both the investment banks and the ultimate investors. This section explores whether steered mortgages were more likely to be sold to private market securitizers.

In Table 5, Panel A, we regress indicators for whether a mortgage was kept as a portfolio loan, securitized by a private market organization, or securitized by one of the governmentsponsored entities (GSEs). Our results strongly indicate that the steered loans were much more likely to be funded through private-label securitizations than held on bank portfolios. The point estimates in Columns (1) and (2) show that steered loans are 47% (= 0.207/0.440) more likely to be sold into a private-label MBS pool relative to being held in a bank's own portfolio. Columns (5) and (6) suggest that steered loans were equally likely to be sold to GSEs as mortgages in the control sample. (Note that the three funding outlets are mutually exclusive alternatives, summing up to 1.) The results in Panel B (strict-matching sample) show about half the magnitude compared with Panel A.

These results corroborate our conjecture that generating origination fees from selling mortgages to securitizers was the prime motivation for lenders to steer borrowers into exotic products.

## 4.4 The Cost of Predatory Lending: An Interplay between Refinancing and Defaulting

To examine the behavior of borrowers from affiliated lenders post-origination, we focus on two actions: refinancing and default. As shown earlier, many of the mortgage types that were sold to borrowers from affiliated lenders provided affordability in the short term through teaser rates, flexible payments, and zero or even negative amortization. This, however, came at the potential expense of high-interest rates later, a jump in payments when accelerated amortization kicked in, and higher loan-to-value ratios. The common wisdom during the years we study was that one could always refinance their mortgages and thus avoid the cost of non-traditional contracts later.[28] This was true as long as home prices continued to increase (2002–2006), but became more difficult in 2007 and almost impossible in 2008. On the flip side, once refinancing was not possible during a period of falling house prices, borrowers from affiliated lenders with mortgages experiencing rising payments (e.g., due to

[28]Demyanyk and Van Hemert (2011) document that 30% of alternative loan products were paid in full within the first year, and almost all were repaid in 3 years after origination.

## Table 5. Mortgage Allocation, by Lender Affiliation

The table presents regressions of indicators for the allocations of mortgages to banks' portfolios, private securitizations, and public (GSE) securitizations on lender affiliation indicator, as well as various fixed effects and borrower and mortgage characteristics. Borrower and mortgage controls include: logged borrower income, FICO credit score of the borrower (splined into the ranges: 621–660, 661–720, 721–760, and > 760), logged loan amount, LTV ratio at origination (splined into 80%–89%, 90%–99%, and ≥ 100%), amortizing ARM indicator, interest-only indicator, refi indicator, prepayment penalty indicator, owner-occupier indicator, conventional mortgage indicator, and no/low documentation indicator. All regressions are OLS regressions. Standard errors are double-clustered by calendar month and state of origination. Data sources: Home Mortgage Disclosure Act (HMDA), 1998–2006; U.S. Census. *t*-statistics are reported in brackets. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively. See Section 3 for details on data sources and sample construction. Variable definitions are provided in Appendix A.

**Panel A: Propensity-Score Matching Sample**

| Dependent variable: | I(Portfolio loan) | | I(Private-label securitization) | | I(GSE securitization) | |
|---|---|---|---|---|---|---|
| Mean in the control sample: | 0.17 | | 0.44 | | 0.38 | |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| I(Steered Borrower) | −0.231*** | −0.230*** | 0.207*** | 0.203*** | 0.025 | 0.028 |
| | [−12.32] | [−8.12] | [6.13] | [4.16] | [0.91] | [0.76] |
| Borrower and mortgage characteristics | Yes | Yes | Yes | Yes | Yes | Yes |
| State × Qtr fixed effects | Yes | No | Yes | No | Yes | No |
| Matched pair fixed effects | No | Yes | No | Yes | No | Yes |
| Observations | 134,083 | 134,083 | 134,083 | 134,083 | 134,083 | 134,083 |
| Adjusted $R^2$ | 0.172 | 0.139 | 0.314 | 0.300 | 0.372 | 0.376 |

**Panel B: Strict-Matching Sample**

| Dependent variable: | I(Portfolio loan) | | I(Private-label securitization) | | I(GSE securitization) | |
|---|---|---|---|---|---|---|
| Mean in the control sample: | 0.16 | | 0.54 | | 0.30 | |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| I(Steered Borrower) | −0.161*** | −0.177*** | 0.123*** | 0.117** | 0.040 | 0.063 |
| | [−13.77] | [−7.60] | [4.02] | [2.44] | [1.36] | [1.39] |
| Borrower and mortgage characteristics | Yes | Yes | Yes | Yes | Yes | Yes |
| State × Qtr fixed effects | Yes | No | Yes | No | Yes | No |
| Matched pair fixed effects | No | Yes | No | Yes | No | Yes |
| Observations | 19,199 | 19,199 | 19,199 | 19,199 | 19,199 | 19,199 |
| Adjusted $R^2$ | 0.140 | 0.031 | 0.322 | 0.320 | 0.350 | 0.386 |

the expiration of teaser rates and the onset of amortization) and larger balances relative to house value would be subject to high risk of default.

To summarize, we anticipate that the steered and non-steered groups would behave differently in terms of their refinancing and default propensities during the boom and bust phases of the housing cycle. In boom times, we expect refinancing activity to be higher for borrowers from affiliated lenders, a difference which would diminish once the crisis hit. The default rate should exhibit the opposite pattern: the likelihood of default should be similar between steered and non-steered during boom time but should be materially higher during crisis. The increased default rate at bad times manifests the costs of predatory lending.

We present the results in two ways. First, we discuss the raw data presented in Figure 1, Panels (a) and (b). Figure 1, Panel (a), shows the quarterly refinancing rate from 2004 through 2011 for steered (solid line) and non-steered (dotted line) loans. The refinancing rates in each calendar quarter are computed relative to a set of mortgages that survived at the beginning of the quarter (i.e., were not refinanced and did not default). We observe a substantially higher propensity of steered loans to refinance from early 2005 through early 2007. Steered loans had about 50 percent higher quarterly refinancing rates during this period. However, this gap effectively evaporates with the housing price bubble burst.

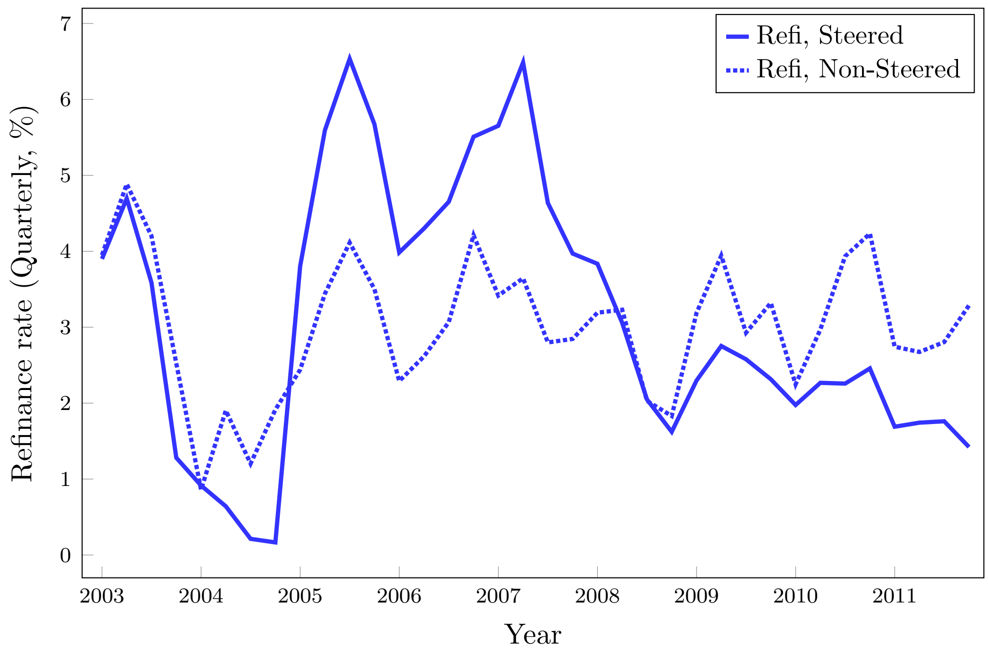
A similar analysis of default rates over time in Figure 1, Panel (b), shows them to be very similar for the two groups until early 2007. At that point—precisely when the propensity to refinance by the steered borrower slows down dramatically—their default rates accelerate well past the also rising rates of the borrowers from affiliated lenders. The inverse relationship between the series in these two figures illustrates the apparent tradeoff between the ability to refinance and mortgage default, which is pronounced more sharply among non-traditional mortgages of borrowers from affiliated lenders.

An alternative way to evaluate refinancing and default patterns is by focusing on origination cohorts, building on the insights of Demyanyk and Van Hemert (2011). While Panels (a) and (b) in Figure 1 display a clear pattern in calendar time, the sample comprises varying numbers of loans originating in different years under different market conditions. Hence, one might worry that secular changes in the prevalence of non-traditional mortgage contracts in different cohorts drive the difference between the two groups. To check this, we follow

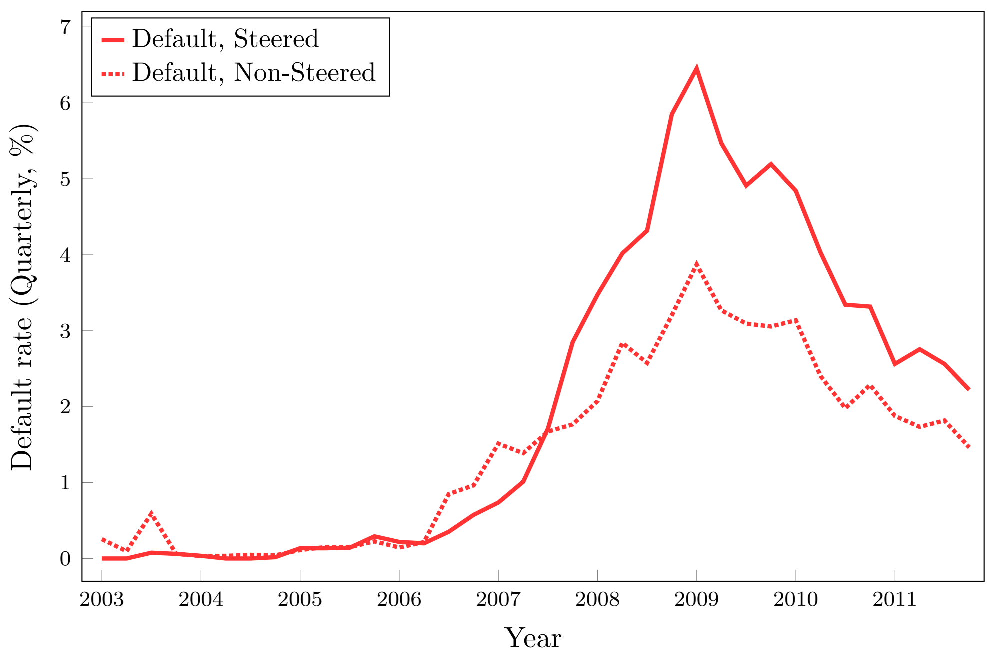
**Figure 1. Quarterly Refinancing and Default Rates**

This figure shows the quarterly refinancing and default rates from 2004 through 2011 for affiliated (solid line) and unaffiliated (dotted line) loans. The refinancing (default) rates in each calendar quarter are computed relative to a set of mortgages that survived at the beginning of the quarter (i.e., were not refinanced and did not default).

(a) Refinance rate




(b) Default rate



the Demyanyk and Van Hemert (2011) approach to show the refinancing and default rates over time for individual origination cohorts. These are broken into eight groups: steered mortgages that were originated from 2003 to 2006, and non-steered mortgages that were originated over the same periods. We plot the refinancing (default) rates in each origination cohort as shares of mortgages that survived at the beginning of the year (i.e., were not refinanced and did not default in prior years). Refinancing (default) rates of steered mortgages are presented with solid lines, while non-steered series are presented as dotted lines.

Figure 2, Panel (a), shows that for almost each origination cohort, refinancing activity is stronger for mortgages in the steered group. The gap in the refinancing rates between steered and non-steered groups is exceptionally high for 2005–2006. For instance, 27% of steered loans originated in 2005 that survived their first year were refinanced in 2006, compared to only 14% of non-steered loans. By 2007, however, the refinancing rates on remaining loans shrank dramatically for both groups—to 13% and 9%, respectively, nearly closing the gap.

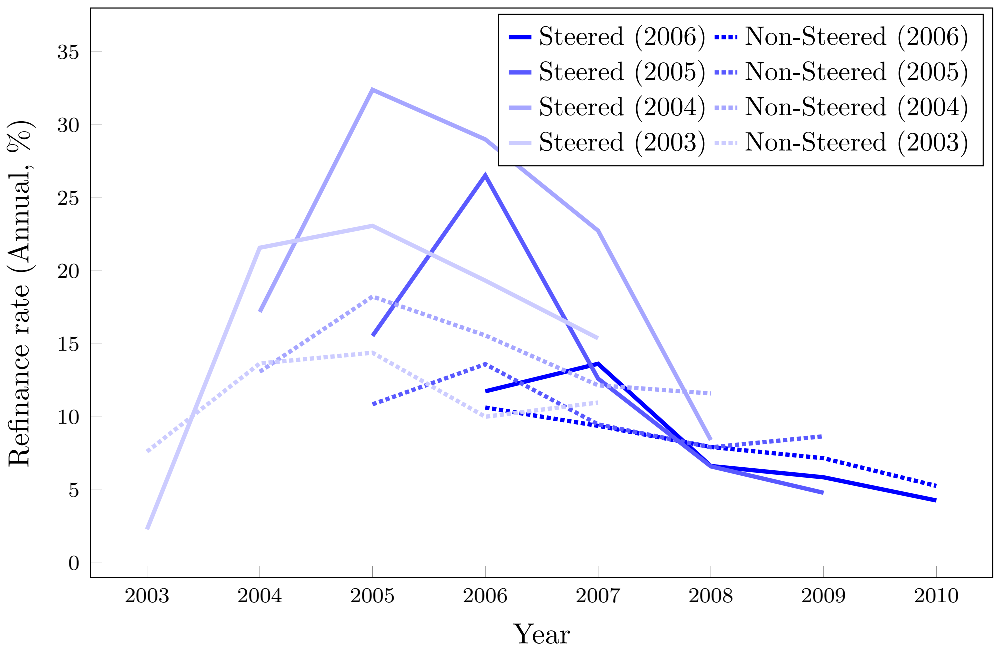
The cohort default rates in Figure 2, Panel (b), also display a familiar pattern. Default rates on surviving loans jump dramatically in 2007–2008 for all origination cohorts. Within each cohort, the difference in default rate between steered and non-steered, which was negligible in 2003–2005 starts opening up in 2006, and peaks in 2008, precisely at the same time when borrowers from affiliated lenders cannot refinance their mortgages. These figures underscore the differences between the steered- and non-steered borrowers from affiliated lenders within each cohort. Still, they became much more amplified during the time of falling home prices.

Finally, we present a similar analysis in a regression framework that accounts for time and cohort effects. We transform our data to a panel dataset where each observation reflects borrower–quarter. Borrowers can exit the sample if they defaulted (about 27% of the sample), refinanced (about 49% of the sample), or were transferred to a servicer that is not covered by McDash (only 2% of our sample). We are interested in exploring the timing of different exit events. We hypothesize that refinancing was the prevalent exit event when the real estate

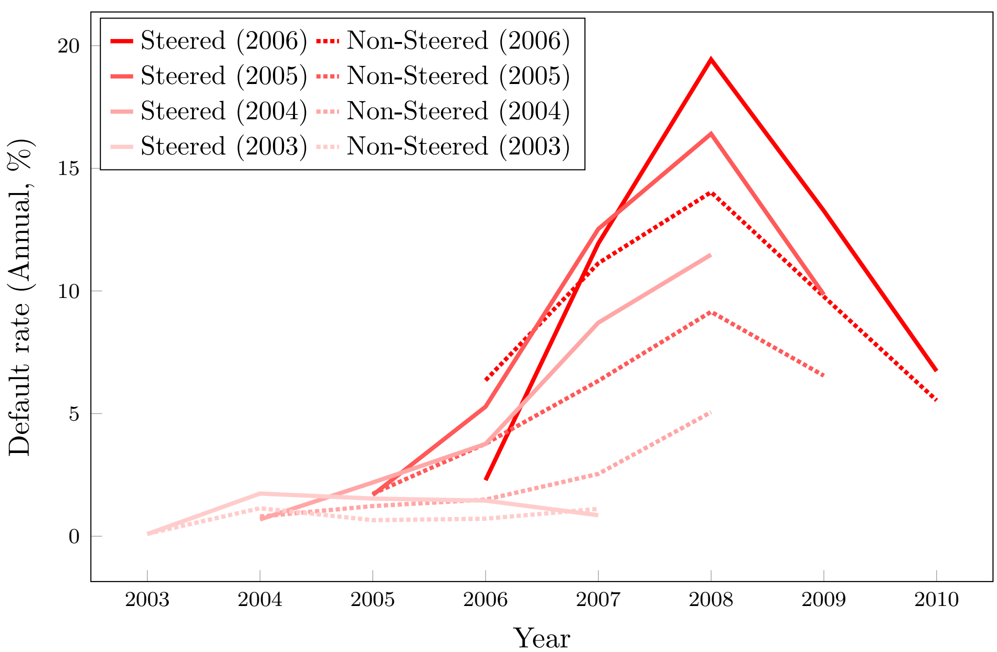
### Figure 2. Refinancing and Default Rate, by Origination Cohort

This figure shows the quarterly refinancing and default rates from 2003 through 2011 for affiliated (solid line) and unaffiliated (dotted line) loans stratified by their year of origination cohort. The refinancing (default) rates in each calendar quarter are computed relative to a set of mortgages that survived at the beginning of the quarter (i.e., were not refinanced and did not default).

(a) Refinance, by origination cohort




(b) Default, by origination cohort



market boomed, and default was prevalent after house prices stopped increasing. Ferreira and Gyourko (2012) show that housing prices peaked in most neighborhoods between the second half of 2005 and the second half of 2007, with the greatest concentration in 2006. The differential hazard rates of default and refinancing for steered and non-steered groups represent our estimates of the effects of predatory lending. Specifically, we estimate the following econometric model:

$$\begin{aligned} ExitEvent_{it} &= \alpha + \beta \sum_t CalendarQuarter_t \\ &\quad + \gamma I(SteeredBorrower)_i \cdot \sum_t CalendarQuarter_t \\ &\quad + \delta BorrowerControls_i + \theta MortgageControls_i + \varepsilon_{it} \end{aligned} \tag{2}$$

This specification controls for time and state fixed effects, annual origination cohort effects, and the same array of borrower and mortgage controls used in regressions in Tables 3–5. We are interested in the series of estimated coefficients $\beta + \gamma t$ for both refinancing and default hazards. These series, along with the estimated error bands, are plotted in Figure 3, Panels (a) and (b).

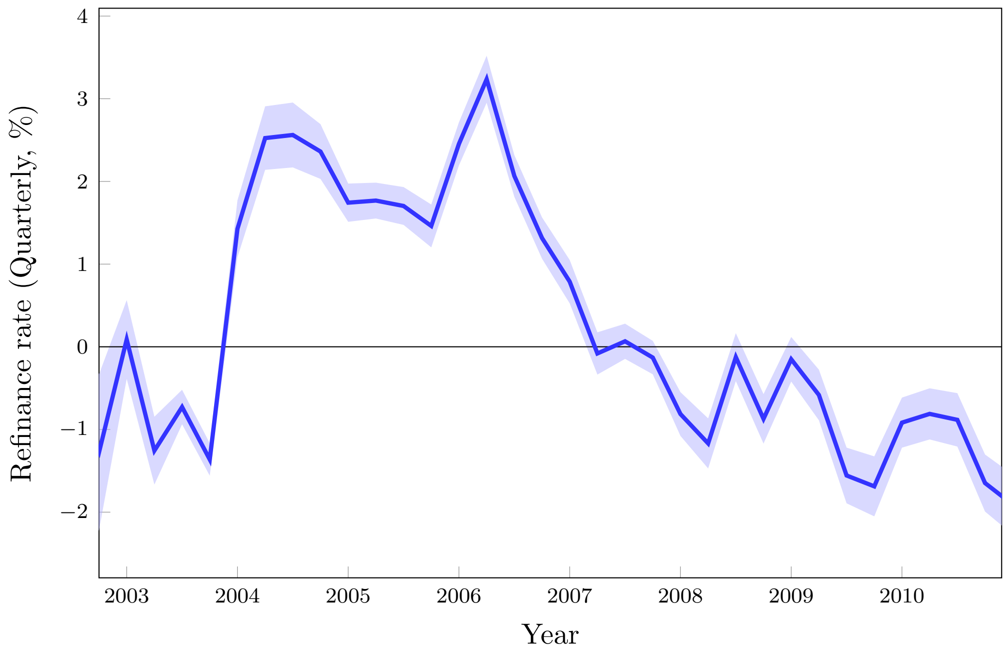
The regression results reaffirm observations from raw data and provide additional information about the magnitude of the effects. During the boom part of our sample, steered borrowers were more likely to refinance their mortgages, with conditional quarterly refinancing rates about two percentage points higher, on average (relative to the mean refinancing rate of 3.2% for non-steered borrowers). Once the real estate market stalls, the difference in the likelihood of refinancing collapses.

Default rates exhibit the reverse pattern. In the first half of the sample (2004 to early 2007), the default likelihood of steered borrowers is somewhat lower than that of the non-steered. In later years, the quarterly default rates of the steered were nearly two percentage points above that of the non-steered. This is similar in magnitude to the effect of predatory lending on default documented by Agarwal et al. (2014). This reversal in fortune coincides

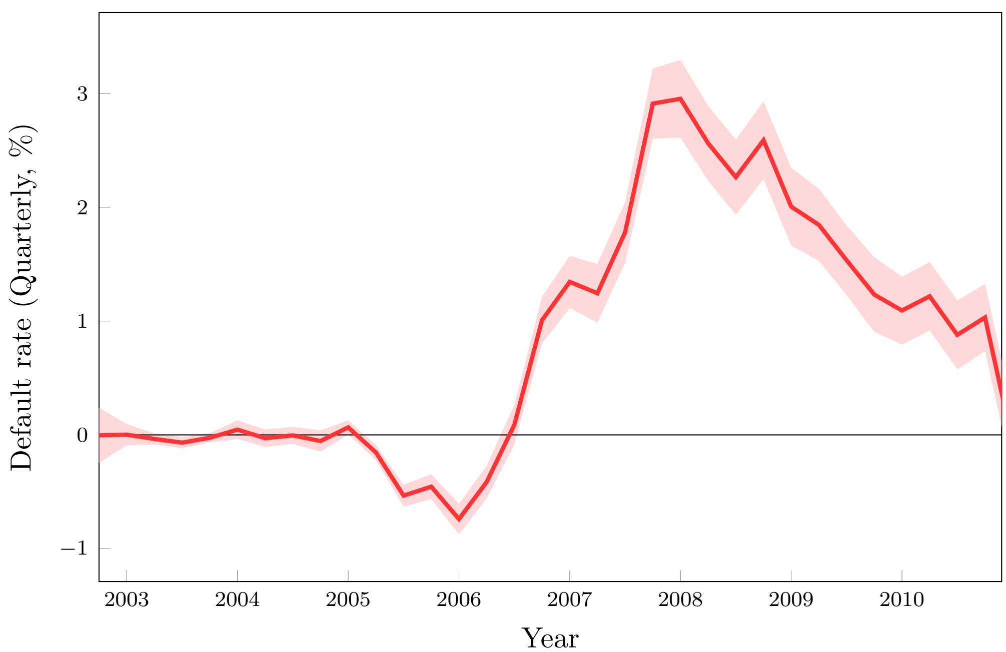
## Figure 3. Refinancing and Default Rate, by Origination Cohort

This figure shows the difference in quarterly refinancing and default rates of steered and non-steered loans obtained from regression analysis specified in Equation (2) in the text. The dashed lines show the 95 percent confidence interval.

(a) Refinance rate: Steered minus Non-Steered Borrowers




(b) Default rate: Steered minus Non-Steered Borrowers



almost perfectly with the disappearance of higher refinancing rates by steered borrowers.

# 5 Conclusion

During the housing boom of the 2000s, there were frequent accusations of predatory lending. In this study, we provide micro-level evidence about the process of predatory lending. The evidence is consistent with a narrative in which lenders engage in the origination-to-distribute model (Purnanandam, 2011) and thus attempt to steer borrowers into products that yield greater origination fees. While these products appear to be affordable in the short run, they expose the borrowers to risks in the long run.

Aggressive selling practices are generally unobservable to us as they occur at the point of sale: we typically do not know the consumer's true needs and which product the loan officer offers. We only observe the mortgage product that the consumer eventually sold. Therefore, in this study, our focus is on a very specific case in which steering activity is potentially visible—when applicants have already submitted their applications and then rejected by the lenders and consequently referred to an affiliate lender. We use borrowers who were rejected and eventually borrowed from a different lender as a control group.

Our results suggest that borrowers who were rejected by a lender and borrowed from an affiliate appear to be taking mortgages that have features of predatory loans (high interest, no/low documentation, and no/low amortization). Borrowers who were steered to expensive mortgage products tend to be of populations that are considered in the literature to be more vulnerable: females, single (no co-signers), and from low-income neighborhoods. We confirm that these loans are likely to be predatory as their characteristics and features match those described by the literature: high interest and exotic features (prepayment penalties, no/low amortization, and no/low documentation). Furthermore, we find that these mortgages were significantly more likely to end up securitized, giving rise to the claim of legal scholars that the demand from investors bolstered predatory lending.

When examining the behavior of borrowers and the performance of mortgages, we document that borrowers stayed for a relatively short time with predatory products. In good years (2003–2006), they were likely to refinance their mortgages (and therefore help lenders generate additional fees). In bad years (post-2006), their default rate was significantly higher, indicating that the mortgage products were not affordable for borrowers.

Overall, our results show that the true costs of predatory lending are revealed when a crisis hits. At that point, borrowers who carry expensive mortgages become fragile and default. The findings resonate with Zingales (2015), questioning the benefit that financial innovation that shrouds certain risks for financially unsophisticated borrowers brings to society.

# References


Agarwal, Sumit, Brent W Ambrose, and Vincent Yao, 2020a, Lender steering in residential mortgage markets, *Real Estate Economics* 48, 446–475.

Agarwal, Sumit, Gene Amromin, Itzhak Ben-David, Souphala Chomsisengphet, and Douglas D Evanoff, 2010, Learning to cope: Voluntary financial education and loan performance during a housing crisis, *American Economic Review: Papers & Proceedings* 100, 495–500.

Agarwal, Sumit, Gene Amromin, Itzhak Ben-David, Souphala Chomsisengphet, and Douglas D Evanoff, 2014, Predatory lending and the subprime crisis, *Journal of Financial Economics* 113, 29–52.

Agarwal, Sumit, Gene Amromin, Itzhak Ben-David, Souphala Chomsisengphet, and Douglas D Evanoff, 2020b, Financial education versus costly counseling: How to dissuade borrowers from choosing risky mortgages?, *American Economic Journal: Economic Policy* 12, 1–32.

Agarwal, Sumit, and Itzhak Ben-David, 2018, Loan prospecting and the loss of soft information, *Journal of Financial Economics* 129, 608–628.

Agarwal, Sumit, Itzhak Ben-David, and Vincent Yao, 2015, Collateral valuation and borrower financial constraints: Evidence from the residential real estate market, *Management Science* 61, 2220–2240.

Agarwal, Sumit, Itzhak Ben-David, and Vincent Yao, 2017, Systematic mistakes in the mortgage market and lack of financial sophistication, *Journal of Financial Economics* 123, 42–58.

Agarwal, Sumit, Richard J Rosen, and Vincent Yao, 2016, Why do borrowers make mortgage refinancing mistakes?, *Management Science* 62, 3494–3509.

Amromin, Gene, Jennifer Huang, Clemens Sialm, and Edward Zhong, 2018, Complex mortgages, *Review of Finance* 22, 1975–2007.

Avery, Robert, Kenneth Brevoort, and Glenn Canner, 2007, Opportunities and issues in using HMDA data, *Journal of Real Estate Research* 29, 351–380.

Azmy, Baher, 2005, Squaring the predatory lending circle, *Florida Law Review* 57, 295.

Barr, Michael S, 2005, Credit where it counts: The Community Reinvestment Act and its critics, *NYU Law Review* 80, 513.

Ben-David, Itzhak, 2011, Financial constraints and inflated home prices during the real estate boom, *American Economic Journal: Applied Economics* 3, 55–87.

Berndt, Antje, Burton Hollifield, and Patrik Sandås, 2016, How subprime borrowers and mortgage brokers shared the pie, *Real Estate Economics* 44, 87–154.

Bocian, Debbie Gruenstein, Keith S Ernst, and Wei Li, 2008, Race, ethnicity and subprime home loan pricing, *Journal of Economics and Business* 60, 110–124.

Bond, Philip, David K Musto, and Bilge Yilmaz, 2009, Predatory mortgage lending, *Journal of Financial Economics* 94, 412–427.

Bord, Vitaly, and João AC Santos, 2012, The rise of the originate-to-distribute model and the role of banks in financial intermediation, *Economic Policy Review* 18, 21–34.

Bostic, Raphael, Souphala Chomsisengphet, Kathleen C Engel, Patricia A McCoy, Anthony Pennington-Cross, and Susan Wachter, 2012, Mortgage product substitution and state anti-predatory lending laws: Better loans and better borrowers?, *Atlantic Economic Journal* 40, 273–294.

Bostic, Raphael W, Kathleen C Engel, Patricia A McCoy, Anthony Pennington-Cross, and Susan M Wachter, 2008, State and local anti-predatory lending laws: The effect of legal enforcement mechanisms, *Journal of Economics and Business* 60, 47–66.

Bowen, Robert M, S Jane Jollineau, and Barbara A Lougee, 2014, WaMu's option-ARM strategy, *Issues in Accounting Education* 29, 9–24.

Carr, James H, and Lopa Kolluri, 2001, Predatory lending: An overview, *Fannie Mae Foundation* 1–17.

Commission, Financial Crisis Inquiry, 2010, *The financial crisis inquiry report: Final report of the national commission on the causes of the financial and economic crisis in the United States* (Government Printing Office).

Demyanyk, Yuliya, and Elena Loutskina, 2016, Mortgage companies and regulatory arbitrage, *Journal of Financial Economics* 122, 328–351.

Demyanyk, Yuliya, and Otto Van Hemert, 2011, Understanding the subprime mortgage crisis, *Rview of Financial Studies* 24, 1848–1880.

Di Maggio, Marco, Amir Kermani, and Sanket Korgaonkar, 2019, Partial deregulation and competition: Effects on risky mortgage origination, *Management Science* 65, 4676–4711.

Egan, Mark, 2019, Brokers versus retail investors: Conflicting interests and dominated products, *Journal of Finance* 74, 1217–1260.

Eggert, Kurt, 2001, Held up in due course: Predatory lending, securization, and the holder in due course doctrine, *Creighton Law Review* 35, 503.

Elliehausen, Gregory, and Michael E Staten, 2004, Regulation of subprime mortgage products: An analysis of North Carolina's predatory lending law, *Journal of Real Estate Finance and Economics* 29, 411–433.

Engel, Kathleen C, and Patricia A McCoy, 2001, A tale of three markets: The law and economics of predatory lending, *Texas Law Review* 80, 1255.

Engel, Kathleen C, and Patricia A McCoy, 2006, Turning a blind eye: Wall Street finance of predatory lending, *Fordham Law Review* 75, 2039.

Ferreira, Fernando, and Joseph Gyourko, 2012, Heterogeneity in neighborhood-level price growth in the United States, 1993–2009, *American Economic Review: Papers & Proceedings* 102, 134–140.

Garmaise, Mark J, 2015, Borrower misreporting and loan performance, *Journal of Finance* 70, 449–484.

Gennaioli, Nicola, Andrei Shleifer, and Robert Vishny, 2012, Neglected risks, financial innovation, and financial fragility, *Journal of Financial Economics* 104, 452–468.

Griffin, John M, and Gonzalo Maturana, 2016, Who facilitated misreporting in securitized loans?, *Review of Financial Studies* 29, 384–419.

Gurun, Umit G, Gregor Matvos, and Amit Seru, 2016, Advertising expensive mortgages, *Journal of Finance* 71, 2371–2416.

Harvey, Keith D, and Peter J Nigro, 2004, Do predatory lending laws influence mortgage lending? An analysis of the North Carolina predatory lending law, *Journal of Real Estate Finance and Economics* 29, 435–456.

Ho, Giang, and Anthony Pennington-Cross, 2006, The impact of local predatory lending laws on the flow of subprime credit, *Journal of Urban Economics* 60, 210–228.

Immergluck, Dan, and Geoff Smith, 2003, Measuring neighborhood diversity and stability in home-buying: Examining patterns by race and income in a robust housing market, *Journal of Urban Affairs* 25, 473–491.

Jiang, Wei, Ashlyn Aiko Nelson, and Edward Vytlacil, 2014, Securitization and loan performance: Ex ante and ex post relations in the mortgage market, *Review of Financial Studies* 27, 454–483.

Keys, Benjamin J, Tanmoy Mukherjee, Amit Seru, and Vikrant Vig, 2010, Did securitization lead to lax screening? Evidence from subprime loans, *Quarterly Journal of Economics* 125, 307–362.

Kolb, Robert W, 2011, *The financial crisis of our time* (Oxford University Press).

Lusardi, Annamaria, and Olivia S Mitchell, 2014, The economic importance of financial literacy: Theory and evidence, *American Economic Journal: Journal of Economic Literature* 52, 5–44.

Mac, Freddie, 1996, Automated underwriting: Making mortgage lending simpler and fairer for america's families, Industry report, Freddie Mac.

Mayer, Christopher, Karen Pence, and Shane M Sherlund, 2009, The rise in mortgage defaults, *Journal of Economic Perspectives* 23, 27–50.

Morton, Fiona Scott, Florian Zettelmeyer, and Jorge Silva-Risso, 2003, Consumer information and discrimination: Does the internet affect the pricing of new cars to women and minorities?, *Quantitative Marketing and Economics* 1, 65–92.

Nadauld, Taylor D, and Shane M Sherlund, 2013, The impact of securitization on the expansion of subprime credit, *Journal of Financial Economics* 107, 454–476.

Ondrich, Jan, Stephen Ross, and John Yinger, 2003, Now you see it, now you don't: Why do real estate agents withhold available houses from black customers?, *Review of Economics and Statistics* 85, 854–873.

Piskorski, Tomasz, Amit Seru, and James Witkin, 2015, Asset quality misrepresentation by financial intermediaries: Evidence from the RMBS market, *Journal of Finance* 70, 2635–2678.

Purnanandam, Amiyatosh, 2011, Originate-to-distribute model and the subprime mortgage crisis, *Review of Financial Studies* 24, 1881–1915.

Reiss, David, 2005, Subprime standardization: How rating agencies allow predatory lending to flourish in the secondary mortgage market, *Florida Stare University Law Review* 33, 985.

Renuart, Elizabeth, 2004, An overview of the predatory mortgage lending process, *Housing Policy Debate* 15, 467–502.

Shiller, Robert J, 2017, Narrative economics, *American Economic Review* 107, 967–1004.

White, Alan M, 2008, Deleveraging the American homeowner: The failure of 2008 voluntary mortgage contract modifications, *Conneticut Law Review* 41, 1107.

Zingales, Luigi, 2015, Presidential address: Does finance benefit society?, *Journal of Finance* 70, 1327–1363.

# Appendix A  Variable Definitions

| Variable | Description | Source |
|---|---|---|
| I(Affiliated lender) | 1 if the rejected mortgage application is approved soon after by an affiliated lender; 0 if unaffiliated | HMDA, authors' calculations |
| FICO score | FICO score at origination | McDash |
| LTV ratio | First-lien loan-to-value ratio at origination | McDash |
| Borrower income | Borrower income at origination, as reported | HMDA |
| Loan amount | First-lien mortgage amount at origination | McDash |
| I(Refinancing) | 1 if a mortgage is identified as refinancing an existing mortgage | McDash |
| I(Owner-occupied) | 1 if a property is reported to be owner-occupied | McDash |
| I(Conventional (non-FHA)) | 1 if a mortgage originated outside of FHA/VA | McDash |
| Initial interest rate (%) | | McDash |
| 90-day delinquency w/ 2 years | | McDash |
| I(Interest Only) | 1 if a mortgage calls for interest-only payments for a pre-specified number of years, fixed amortization schedule thereafter | McDash |
| I(Option ARM) | 1 if a mortgage has an adjustable interest rate but required payments may be less than interest charges subject to time and LTV restrictions | McDash |
| I(Prepayment penalty) | 1 if a mortgage contract has a penalty for refinancing before a pre-specified time | McDash |
| I(No/Low documentation) | 1 for mortgages that are listed as not being underwritten based on fully documented income and assets | McDash |
| Fixed rate term | | McDash |
| Mortgage maturity | | McDash |
| I(Portfolio loan) | | McDash |
| I(GSE securitization) | | McDash |
| I(Private-label securitization) | | McDash |
| ΔHPI 12-months pre-origination | Annual change in ZIP or MSA home price index in the 12 months preceding mortgage origination | CoreLogic |
| I(African-American) | | HMDA |
| I(Hispanic) | | HMDA |
| I(Female) | | HMDA |
| I(No co-signer) | | HMDA |
| I(Low/Moderate Income tract) | | McDash |
| ???I(Fixed rate) | 1 if a mortgage is identified as having a fixed interest rate | McDash |
| ???I(Amortizing ARM) | 1 if a mortgage has an adjustable interest rate but amortizes over a pre-determined period | McDash |