RISKY BORROWERS OR RISKY MORTGAGES:
Disaggregating Effects Using Propensity Score Models

Working Paper: November 30, 2009

By
Lei Ding
Roberto G. Quercia
Janneke Ratcliffe
Center for Community Capital

Wei Lei
Center for Responsible Lending
Durham N.C.

Prepared with financial support from

Center *for* Community Capital
Research and analysis on the transformative power of capital

UNC
COLLEGE OF
ARTS & SCIENCES

Risky Borrowers or Risky Mortgages
Disaggregating Effects Using Propensity Score Models

Lei Ding, [a] * Roberto G. Quercia, [b] Wei Li, [c] Janneke Ratcliffe [b]

Draft on November 30, 2009

[a] Department of Urban Studies and Planning, Wayne State University, Detroit, MI
[b] Center for Community Capital, University of North Carolina, Chapel Hill, NC
[c] Center for Responsible Lending, Durham, NC

* Contact author: Telephone: 313-577-0543, E-mail: lei_ding@wayne.edu

Risky Borrowers or Risky Mortgages
Disaggregating Effects Using Propensity Score Models

Abstract:

In this research, we examine the relative risk of subprime mortgages and community reinvestment loans. Using the propensity score matching method, we construct a sample of comparable borrowers with similar risk characteristics but holding the two different loan products. We find that community reinvestment loans have a lower default risk than subprime loans, very likely because they are not originated by brokers and lack risky features such as adjustable rates and prepayment penalties. Our results suggest that similar borrowers holding community reinvestment loans exhibit significantly lower default risks.

Introduction

Explanations for the current foreclosure crisis abound. There are the obvious culprits: overextended borrowers, risky mortgages, reckless originators, and investors and other secondary market participants who failed to act with due diligence (e.g. Mian and Sufi, 2008; Quercia and Ratcliffe, 2008). Moreover, there are some who blame government regulation, such as the Community Reinvestment Act (CRA), designed to increase the credit supply to traditionally underserved, but creditworthy, population (Cravatts, 2008; Krauthammer, 2008). From this perspective, the CRA and similar regulation are said to have put pressure on lenders and the government sponsored enterprises (GSEs) to extent mortgages to over-leveraged, uncreditworthy, and/or irresponsible low-income and minority borrowers.

The debate over what caused the mortgage mess and how best to fix it has important policy implications. What is missing in the debate is an empirical examination of the relative performance of similar borrowers holding either a typical CRA loan or a subprime product. Such an analysis will help inform policy by answering the question of whether high default rates represent just the higher risk profile of borrowers holding subprime loans or the risky characteristics of subprime loans. Although borrowers holding subprime loans generally are weaker across key underwriting criteria, many borrowers holding subprime products actually qualify for a prime mortgage (Hudson and Reckard, 2005; Brooks and Simon, 2007). Some products or features that are more prevalent among subprime loans, such as prepayment penalties, adjustable rates, and balloon payments, have been found to be associated with elevated default risk (e.g. Ambrose, LaCour-Little and Huszar, 2005; Pennington-Cross and Ho, 2006; Quercia, Stegman and Davis, 2007). Are the higher default rates reported in the subprime sector mainly the result of risky loan products?

We address this issue by comparing the performance of subprime loans and CRA loans in a special lending program called Community Advantage Program (CAP). To solve the problem of selection bias since performance differences may be due to differences in the borrowers who receive each product type, we rely on propensity score matching methods to construct a sample of comparable borrowers. We find that for borrowers with similar risk characteristics, the estimated default risk is about 70 percent lower with a CRA loan than with a subprime mortgage. Broker-origination channel, adjustable rates, and prepayment penalties all contribute substantially to the elevated risk of default among subprime loans. When broker origination is combined with both adjustable rates and prepayment penalties, the borrower's default risk is four to five times higher than that of a comparable borrower with a prime-term CRA mortgage. Though CAP has some program specific characteristics, the results of this study clearly suggest that mortgage default risk cannot be attributed solely to borrower credit risk; the high default risk is significantly associated with the characteristics of loan products. Thus, the results are not consistent with the concerns of those blaming the borrowers likely to benefit from CRA and similar regulations. Done responsibly, targeted lending programs stimulated by the CRA can do a much better job in providing sustainable homeownership for the low- to moderate-income (LMI) population than subprime lending. The results have important policy implications on how to respond to the current housing crisis and how to meet the credit needs of all communities, especially the LMI borrowers, in the long run.

Compared with prior work, this study is characterized by several important differences. First, while most early studies focused on the performance of mortgages within different markets, the focus here is on similar LMI borrowers with different mortgages, allowing us to compare the relative risk of different mortgage products. Second, because of data constraints, research on the performance of CRA loans is scarce. With a unique dataset, this study examines the long term viability of the homeownership opportunities that CRA-type products provide, relative to that of subprime alternatives. Third, there have been few discussions and applications of the propensity score matching method in real estate research. This study uses propensity score models to explicitly address the selection bias issue and constructs a comparison group based on observational data. This method allows us to isolate the impact of loan product features and origination channel on the performance of mortgages. Finally, while the propensity score model cannot capture all the information for estimating the propensity of taking out a subprime loan, this study makes full use of the loan interest rate information to shed some light on the impact of the unobservable heterogeneity on the mortgage performance.

Literature Review

Risk of Subprime Mortgages

Subprime mortgages were originally designed as refinancing tools to help borrowers with impaired credit consolidate debt. With the reformed lending laws, the adoption of automated underwriting, risk-based pricing, as well as the persistent growth in house prices nationwide, the subprime lending channel soon expanded its credit to borrowers on other margins. The subprime surge was rapid and wide: between 1994 and 2006, the subprime share of all mortgage originations more than quadrupled, from 4.5 percent to 20.1 percent; and subprime loan originations increased more than seventeen fold, from $35 billion to about $600 billion. The surge was largely fueled by securitization (private Wall Street issuances) over the same period, the volume of securitized subprime mortgage loans increased over forty-four-fold, from $11 billion to more than $483 billion in 2006, accounting for more than 80 percent of all subprime lending (Inside Mortgage Finance, 2008).

Beginning in late 2006, a rapid rise in subprime mortgage delinquency and foreclosure caused a so-called meltdown of the subprime market. The Mortgage Bankers Association (MBA) reports that the serious delinquency rate for subprime loans in the second quarter of 2008 was 7.6 times higher than that for prime loans (17.9 percent versus 2.35 percent). Although subprime mortgages represented about 12 percent of the outstanding loans, they represented 48 percent of the foreclosures started during the same quarter (MBA, 2008). Delinquency and default rates for subprime loans typically are six times to more than 10 times higher than those of prime mortgages (Pennington-Cross, 2003; Gerardi, Shapiro and Willen, 2007; Immergluck, 2008).

It may be true that borrowers holding subprime loans are generally weaker across key underwriting criteria. A subprime borrower used to refer to an individual who had any of the following characteristics: 1) a FICO score below 620, 2) a delinquent debt repayment in the previous two years, 3) a bankruptcy filing in the previous five years (Gerardi et al. 2007). Recent subprime home-purchase loans became available to borrowers who may have had impaired credit history or were perceived to have elevated credit risks, such as “low-doc” or “no doc” borrowers, “low-down” or “zero-down” payment borrowers, or borrowers with high debt-to-income ratios (DTIs). All these risk characteristics are usually significantly associated with a higher default risk of the mortgages these borrowers hold.

At this point, it is important to make a distinction between borrowers and mortgage products. It can be said that there are two types of borrowers and two types of mortgage products: prime and subprime. Not all prime borrowers get prime mortgages and not all subprime borrowers get subprime mortgages. Borrowers who do not meet all the traditional underwriting guidelines can be considered subprime but these borrowers can receive prime-type mortgages as they may through CRA efforts. Similarly, borrowers with good credit can receive subprime products characterized by high debt to income and loan to value ratios, no or low documentation, teaser and adjustable rates and other such risky characteristics (the so called Alt-A market).

Some loan features and loan terms are more prevalent in the subprime sector than in other markets and are also associated with higher default risk. As summarized by Cutts and Van Order (2005) and Immergluck (2008), characteristics of subprime loans relative to prime loans include: 1) high interest rates, points, and fees, 2) prevalence of prepayment penalties, 3) prevalence of balloon payments, 4) prevalence of adjustable-rate mortgages (ARMs), 5) popularity of broker originations. After 2004, some “innovative” mortgage products, such as interest-only, payment option, negative amortization, hybrid ARMs, and piggy-back loans became more popular in the subprime sector (Immergluck, 2008). In the literature, Calhoun and Deng (2002) and Quercia et al. (2007) find that subprime ARMs have a higher risk of foreclosure because of the interest-rate risk. At the aggregate level, the share of ARMs appears to be positively associated with market risk as measured by the probability of the property value to decline in the next two years (Immergluck, 2008). Subprime hybrid ARMs, which usually have prepayment penalties, bear particularly high risk of default at the time the interest rate is reset (Ambrose et al. 2005; Pennington-Cross and Ho, 2006).

As to the feature of prepayment penalties and balloons, Quercia et al. (2007) find that refinanced loans with prepayment penalties are 20 percent more likely than loans without to experience a foreclosure while loans with balloon payments are about 50 percent more likely to experience a foreclosure than those without. Prepayment penalties also tend to reduce prepayments and increase the likelihood of delinquency and default among subprime loans (Danis and Pennington-Cross, 2005).

Recently, mortgage brokers have played a greater role in the subprime sector. In 2003 brokers originated about 48 percent of all subprime loans; in 2006 the share was estimated between 63 percent and 80 percent (Ernst et al. 2008), higher than the share of about 30 percent of broker-originated loans among all mortgages in recent years (Inside Mortgage Finance, 2008). Empirical evidence on the behavior of broker-originated mortgages is scarce. LaCour-Little and Chun (1999) find that for the four types of mortgages analyzed, loans originated by a third party (including broker and correspondence) were more likely to prepay than loans originated by a lender. Alexander, Grimshaw, McQueen and Slade (2002) find that third-party originated loans do not necessarily prepay faster but they default with greater frequency than similar retail loans, based on a sample of subprime loans originated from 1996 to 1998. They suggest that third-party originated mortgages have higher default risk than similar retail loans because brokers are rewarded for originating a loan but not held accountable for the loan’s subsequent performance.

Thus, the higher default rates reported in subprime lending may be because of risky borrowers, risky loan products, or a combination of both.

CRA Lending

The Community Reinvestment Act (CRA) of 1977 was created in response to charges that financial institutions were engaging in redlining and discrimination. The Act mandates that federally insured depository institutions help meet the credit needs of communities in which they operate in a manner consistent with safe and sound operation (Bernanke, 2007). Regulators assess each bank’s CRA record when evaluating these institutions’ applications for mergers, acquisitions, and branch openings. The performance of large institutions is measured under three categories of bank activities: lending, services, and investment, with the lending test carrying the most weight (at least 50 percent).[1] For the lending test, it examines the amount and proportion of lending activities made within an institution's assessment area.[2] Usually, loans are regarded as "CRA-related" if they are made by CRA-regulated institutions within their assessment areas to low-income borrowers (those with less than 80% area median income (AMI), regardless of neighborhood income) or in a low- income neighborhood (with less than 80% AMI, regardless of borrower income) (Avery, Bostic and Canner, 2000).

The CRA lending test also examines the use of *innovative or flexible* lending practices to address the credit needs of LMI households and community. In response, many banks have developed "CRA Special Lending Programs" or have introduced mortgage products characterized by more flexible underwriting standards. Survey results suggest that most financial institutions offer these special programs, and that most of the programs relate to home mortgage lending, which typically feature some combination of special outreach, counseling and education, and underwriting flexibility (especially in terms of reduced cash to close, alternative credit verification and higher debt-to-income thresholds) (Avery et al. 2000). A review article by Apgar and Duda (2003) suggests the CRA has had a positive impact on underserved population by originating a higher proportion of loans to low-income borrowers and communities than they would have without CRA. At the same time, one study suggests that there is no evidence that CRA-affected lenders cut interest rates to CRA-eligible borrowers or that there is a regulation-driven subsidy for CRA loans (Canner, Laderman, Lehnert and Passmore, 2002).

CRA-type mortgages are different from subprime loans in that CRA products usually have prime-term characteristics. In general, they are believed to carry a higher risk because they are originated by liberalizing one or two underwriting criteria. Moreover, CRA products are originated by federally insured depository institutions covered by CRA while two out of three subprime lenders are independent mortgage companies not covered by CRA (Bernanke, 2007). A few studies investigating the delinquency behaviors among CRA borrowers suggest the delinquency rate of CRA mortgages is comparable to that of FHA loans after excluding loans with low loan-to-value ratios (LTV) (e.g., Quercia, Stegman, Davis and Stein, 2002). Because of data constraints, little is known about the long term viability of the homeownership opportunities that these products provide.

Why Different Markets Coexist

To increase the flow of funds into low-income population and neighborhoods, the CRA encourages lenders to meet credit needs within their service or catchment area, taking into account safety and soundness considerations. Liberalizing one or two traditional mortgage underwriting standards allows lenders to make loans to those who would otherwise not qualify for a prime mortgage (for instance, not requiring mortgage insurance when the downpayment is less than 20 percent makes loans more affordable for some borrowers).

In this sense, both CRA and subprime products may target many of the same borrowers. In fact, recent studies suggest there is a significant overlap between borrowers holding subprime mortgages and those holding prime loans, FHA loans, and other loan products, particularly among LMI borrowers with marginal credit quality. Freddie Mac, for example, finds that about 20 percent of subprime borrowers could have qualified for a prime rate mortgage (Hudson and Reckard, 2005). A Wall Street Journal report suggests 61 percent of subprime mortgages went to borrowers with credit that would have qualified them for conventional loans by 2006 (Brooks and Simon, 2007). Bocian, Ernst and Li (2007) suggest that a significant portion of subprime borrowers (estimates range from 10 percent to almost 40 percent) could have qualified for low-priced prime loans.

Why would many people who could qualify for low-cost prime-type loans take out subprime products? First of all, many borrowers, especially those with impaired credit history, are usually financially unsophisticated and may feel they have limited options. Courchane, Surette and Zorn (2004) indicate that subprime borrowers "are less knowledgeable about the mortgage process, are less likely to search for the best rates, and are less likely to be offered a choice among alternative mortgage terms and instruments" (p.365). Especially, for some nontraditional mortgages, including interest-only mortgages, negative amortization mortgages, and mortgages with teaser rates, they were apparently not well understood by many borrowers. When borrowers do not know the best price and are less likely to search for the best rates, it is likely that they cannot make the right decision when they shop for mortgage products. In fact, Courchane et al. (2004) find that search behavior as well as adverse life events, age, and Hispanic ethnicity contribute to explaining the choice of a subprime mortgage.

Second, *predatory lending* or abusive lending practices are concentrated in the subprime sector which may explain why some borrowers end up with certain loans. Unscrupulous lenders, or brokers as their agents, may take advantage of uninformed borrowers by charging fees and rates not reflected of the risk, by not informing borrowers of lower cost loan alternatives, and by offering products and services without full disclosure of terms and options. Renuart (2004) highlights the role of loan steering and abusive push-marketing of subprime lending practices, in which lenders steer borrowers to subprime products instead of low-cost prime alternatives. A major reason for this is that there are higher incentives from originating subprime mortgages than from low-cost alternatives. Compared to traditional prime mortgages, subprime mortgages generated much higher profit for originators before the bust – 3.6 percent versus 0.93 percent for Countrywide alone in 2004 (Morganson, 2008). For brokers, in addition to the standard origination fees, they are compensated by a yield-spread premium (YSM), which is an extra payment that brokers receive from lenders for delivering a mortgage with a higher interest rate than that for which the borrower may qualify (Ernst et al. 2008). Thus, brokers are usually more concerned about mortgage volume and features that generate fees and points from borrowers and commissions and premiums from lenders, instead of the loan's subsequent performance. Because the subprime market is characterized by complicated pricing tiers and product types that are not easy to understand, the steering problem is likely to be more pronounced in the subprime sector than in other markets in which products are generally standardized. Furthermore, the originators usually do not have to be held accountable for the loan's long term performance as most of subprime loans originated in recent years were securitized (80 percent in 2006). For brokers, broker fees and the yield spread premiums are paid upon settlement of the loan, at which point the broker would have no further stake in the performance of that loan. Of course, banks and investors, as well as brokers and banks, are involved in repeated relationships, reputation concerns may somewhat prevent the moral hazard of lenders. But the not well-designed compensation mechanism and the lack of responsibility for the long-term sustainability of mortgages provide the incentive for many lenders and brokers to originate subprime loans than other less profitable products to maximize their own profit.

In the literature, similar behaviors have been examined with the information asymmetry theory, moral hazard theory, and agency cost theory. For an originator to provide an efficient level of such services as marketing and underwriting mortgage products, it must be given the proper incentives to do so. But Alexander et al. (2002) suggest that third-party originators have the incentive to game with lenders and investors either passively or actively in the credit underwriting process: intentionally lacking rigor in the screening process, exaggerating measures of credit worthiness or property value, or targeting and putting borrowers with marginal quality to high-cost subprime with risky loan terms instead of lower cost alternatives.[3] Mian and Sufi (2008) blame the moral hazard on behalf of originators selling risky mortgages is the primary cause of the loose underwriting and the subsequent mortgage foreclosure crisis. Keys, Mukherjee, Seru and Vig (2008) also suggest that securitization leads to lax screening by adversely affecting the screening incentives of lenders.

In short, borrowers generally sort to prime/CRA, subprime or other mortgage markets based on their risk profile. However, the lack of financial sophistication of some borrowers, the poor alignment of incentives, and moral hazard considerations are some of the many reasons borrowers—especially marginally qualified borrowers—may receive less desirable mortgage products than they can be qualified for.

Data

Data for this study come from one LMI-targeted lending program, the Community Advantage Program (CAP), developed by Self-Help in partnership with a group of lenders, Fannie Mae, and the Ford Foundation. Participating lenders establish their own guidelines. The most common variants from typical conventional, prime standards are: reduced cash required to close (through lower down payment and/or lower cash reserve requirements);[4] alternative measures or lower standards of credit quality;[5] and flexibility in assessing repayment ability (through higher debt ratios and/or flexible requirements for employment history).[6] These guidelines variants could be combined or used to offset each other.[7] Nearly 90 percent of the programs feature exceptions in at least two of these areas, and more than half feature exceptions in all three. The majority of programs combine neighborhood and borrower targeting.

Under the LMI-targeted CAP lending program, participating lenders are able to sell these nonconforming mortgages to Self-Help, which then securitizes and sells them to Fannie Mae or other investors. Participating lenders originate and service the loans under contract with Self-Help. It should be emphasized that, while many of the borrowers are somewhat credit impaired, the program cannot be characterized as subprime. The vast majority of CAP loans are retail originated (in contrast to broker originated) and feature terms associated with the prime market: thirty-year fixed-rate loans amortizing with prime-level interest rates, no prepayment penalties, no balloons, with escrows for taxes and insurance, documented income, and standard prime-level fees. As a LMI-targeting program, CAP has some program-specific characteristics such as income and geographic limitations.[8]

The data of subprime loans come from a proprietary database from Lender Processing Services, Inc. (LPS, formerly McDash Analytics), which provides loan information collected from approximately 15 mortgage servicers. LPS' coverage in the subprime market by volume increased from 14 percent in 2004 to over 30 percent in 2006, based on our estimation using data from Inside Mortgage Finance. There is no universally accepted definition of *subprime mortgage*; the three most commonly used definitions are 1) those categorized as such by the secondary market, 2) those originated by a subprime lender as identified by HUD's annual list, and 3) those that meet HUD's definition of a "high-cost" mortgage (Gerardi et al. 2007). For the purposes of this paper we primarily follow the first definition, since we can identify those B&C loans in LPS but could not identify lenders' information and mortgages' APR. We further consider high-cost ARMs as subprime in this analysis. Less than 20% of loans in our LPS study sample are included solely because they are considered high-cost, defined as having a margin greater than 300 basis points (Poole, 2007). In addition, we appended to our data selected census and aggregated HMDA variables at a zip code level, including the Herfindahl-Hirschman Index ("HHI") calculated from HMDA, racial and educational distribution from census data, and area average FICO scores calculated from the LPS data.

We started from a sample of 9,221 CAP loans originated from 2003 to 2006. All are first-lien, owner-occupied, fixed-rate conforming home purchase loans with full or alternative documentation. National in scope, these loans were originated in 41 states, with about two-thirds concentrated in Ohio, North Carolina, Illinois, Georgia and Oklahoma. To make sure subprime loans are roughly comparable to CAP loans, as Exhibit 1 shows, we limited our analysis to subprime mortgages also characterized as first-lien, single-family, purchase-money, and conforming loans with full or alternative documentation that originated during the same period. We further excluded loans with missing values for some key underwriting variables (FICO score, LTV, DTI, and documentation status) and loans without complete payment history. Finally, because we want to compare CAP and subprime loans in the same market, we excluded those subprime loans in areas without CAP lending activities. This gave us a sample of 42,065 subprime loans. Table 2 summarizes some important characteristics of both CAP loans and subprime loans in this analysis. Significance tests show that almost all variables across the two groups differ significantly before matching, indicating that the covariate distributions are different between CAP and subprime loans in the original sample.

Though drawn from similar markets, the CAP borrowers (including all active loans originated as early as 1990s) are not experiencing the same mortgage woes as subprime borrowers. As Exhibit 2 shows, 3.21 percent of our sample of community lending borrowers were 90-days' delinquent or in foreclosure process in the second quarter of 2008. This was slightly higher than the 2.35 percent delinquency rate on prime loans but well below the 17.8 percent on subprime loans nationwide. Especially, over 27 percent of subprime ARMs were in foreclosure or serious delinquency, which was almost nine times that of community lending loans.

In summary, the CAP and subprime samples have identical characteristics for the following important underwriting variables: lien status, amortization period, loan purpose, occupancy status, and documentation type. They were originated during the same time period and roughly in the same geographic areas. But the two samples differ in other underwriting factors, including DTI, LTV, and FICO score, and in loan amount and some loan features that are more common only for subprime loans. In the next section, we use the propensity score matching (PSM) method to develop a new sample by matching CAP loans with comparable subprime loans.

Methodology

The PSM method has been widely used to reduce selection biases in recent program evaluation studies. PSM was first developed by Rosenbaum and Rubin (1983) as an effort to more rigorously estimate causal effects from observational data. Basically, PSM accounts for observable heterogeneity by pairing participants with nonparticipants on the basis of the conditional probability of participation, given the observable characteristics. The PSM approach has gained increasing popularity among researchers from a variety of disciplines, including biomedical research, epidemiology, education, sociology, psychology, and social welfare (see review in

Guo, et al., 2006). There is some evidence that nonparametric PSM methods can produce impact estimates that are closer to the experimental benchmark than the parametric approach (Essama-Nssah, 2006).

There are three basic steps involved in implementing PSM. First, a set of covariates is used to estimate the propensity scores using *probit* or *logit*, and the predicted values are retrieved. Then each participant is paired with a comparable nonparticipant based on propensity scores. In the last step, regression models or other methods can be applied to the matched group to compare the outcomes of participants and nonparticipants. Here we describe these steps in our analysis in more details.

In this case, because receiving a subprime is a choice/assignment process rather than randomly assigned we used the PSM method to adjust this selection bias. In the first step, we employed logistic regression models to predict the propensity ($e(x_i)$) for borrower i (i= 1,…,N) of receiving subprime loans (S_i= 1) using a set of conditioning variables (x_i).

$$e(x_i)=pr(S_i=1|X_\mathrm{i}= x_i) \tag{1}$$

In the second step, we used the nearest-neighbor with caliper method to match CAP borrowers with borrowers holding subprime loans based on the estimated propensity scores from the first step. The method of nearest-neighbor with caliper is a combination of two approaches: traditional nearest-neighbor matching and caliper matching. [9] This method begins with a randomly sort of the participants and nonparticipants, then selecting the first participant and finding the nonparticipant subject with the closest propensity score within a predetermined common-support region called caliper (δ). The approach imposes a tolerance level on the distance between the propensity score of participant i and that of nonparticipant j. Formally, assuming $c(p_i)$ as the set of the neighbors of i in the comparison group, the corresponding neighborhood can be stated as follows.

$$c(p_i) = \{ j \mid \delta > \|p_i - p_j\| \} \tag{2}$$

If there is no member of the comparison group within the caliper for the treated unit i, then the participant is left unmatched and dropped from the analysis. Thus, caliper is a way of imposing a common support restriction. Naturally, there is uncertainty about the choice of a tolerance level since a wider caliper can increase the matching rate but it also increase the likelihood of producing inexact matching. A more restrictive caliper increases the accuracy but may significantly reduce the size of the matched sample.

In the third step, we employed a multinomial regression model (MNL) to further control factors that may influence the performance of the new sample after loan origination, many of which are time-varying. In each month the loan can be in only one state or outcome (active, default, or prepaid). Since the sum of the probabilities of each outcome must equal to one, the increase in the probability of one outcome necessitates a decrease in the probability of at least one competing outcome. Thus the multinomial logit model is a competing risk model.

We can think of mortgage borrowers as having three options each month:

- DEFAULT: This study treats the incidence of the first 90-day delinquency as a proxy of default.
- PREPAID: If a loan was prepaid before it is seriously delinquent, it is considered a prepayment.
- ACTIVE: Active and not default (not seriously delinquent in some models)

The probability of observing a particular loan outcome is given by:

$$\begin{aligned}
\Pr(y_{it} = j) &= \frac{e^{\beta_j Z_{it} + \gamma_j S_i}}{1 + \sum_{k=1}^{2} e^{\beta_k Z_{it} + \gamma_k S_i}} \quad for\ j = 1,2 \\
\Pr(y_{it} = j) &= \frac{1}{1 + \sum_{k=1}^{2} e^{\beta_k Z_{it} + \gamma_k S_i}} \quad for\ j = 0 \\
\ln L &= \sum_{t=1}^{T} \sum_{i=1}^{N} \sum_{j=0}^{2} d_{ijt} \ln(\Pr(y_{it} = j))
\end{aligned} \tag{3}$$

where $j=0,1,2$ represents the three possible outcomes of a loan and the omitted category ($j=0$) remains active and not seriously delinquent (ACTIVE). d_{ijt} is an indicator variable taking on the value 1 if outcome j occurs to loan i at time t, and zero otherwise. Z contains a set of explanatory variables and β is the coefficient. To identify the difference between the performance of CAP loans and subprime loans, S contains a subprime dummy variable or indicators of subprime loan characteristics. Specifically, we considered the impact one origination channel and two loan characteristics: the prepayment penalty, the adjustable rate, and the broker origination channel. We constructed six mutually exclusive dummy variables for the combinations of these three characteristics,[10] such as *sub_bro&arm&ppp* for "broker-originated subprime loans with adjustable rates and prepayment penalties" and *sub_arm* for "retail-originated subprime loans with adjustable interest rates and no prepayment penalties." None of the CAP loans have these features, and they are set as the reference group in both models.

In the context of observational studies, the PSM methods seek to mimic conditions similar to an experiment so that the assessment of the impact of the program can be based on a comparison of outcomes for a group of participants (i.e. those with $S_i = 1$) with those drawn from a comparison group of non-participants ($S_i = 0$). We need to check whether our observational data meet the two primary assumptions underlying the PSM methods: the *conditional independence* assumption and the *overlap* assumption.

Conditional Independence Assumption: [11]

To yield consistent estimates of program impact, matching methods rely on a fundamental assumption known as "*conditional independence,*" which can be formally stated as:

$$(y_0, y_1) \perp w | x \tag{4}$$

This expression states that potential outcomes are orthogonal to treatment status, given the observable covariates. In other words, conditional on observable characteristics, participation is independent of potential outcomes and unobservable heterogeneity is assumed to play no role in participation (Dehejia and Wahba, 2002). Assuming that there are no unobservable differences between the two groups after conditioning on x_i, any systematic differences in outcomes between participants and nonparticipants are due to participation. So the plausibility of an evaluation method depends largely on the correctness of the propensity score model underlying program design and implementation.

Our first strategy is to use a well specified logit regression to estimate the probability of taking out a subprime mortgage for each cohort, grounded on a sound understanding of the subprime market. We determined the conditional variables that are associated with the use of subprime loans based on a review of subprime lending and mortgage choice literature, as discussed in the next section. Second, it is possible that lenders have access to more information about the borrower and local market than the information in our dataset and the unobservable lender information would influence the estimation results. Our strategy is to rerun the multinomial regression model by including the unobservable borrower heterogeneity as an independent variable, which is proxied by interest rate variables if the mortgage note rate can be assumed to an effective predictor of the level of credit risk.

Overlap assumption:

For matching to be feasible, there must be individuals in the comparison group with the same or similar propensity as the participant of interest. This requires an overlap in the distribution of observables between the treated and the comparison groups.

The overlap assumption is usually stated as:

$$0 < pr(w = 1 | x) < 1 \tag{5}$$

This implies the possible existence of a nonparticipant analogue for each participant. When this condition is not met, then it would be impossible to find matches for a fraction of program participants.

In this case, as we discussed in the literature review, it is highly likely that there is significant overlap between the CRA-type CAP loans and the subprime sample since both of them focus on households with marginal credit quality and have identical loan characteristics such as lien status, loan purpose, occupancy status, and documentation type. As shown in Exhibit 3, the distribution of credit scores for the CAP and subprime borrowers, subprime borrowers tend to have lower FICO scores than CAP borrowers, but there is a significant overlap in these distributions. This overlap allows us to conduct a meaningful analysis of the performance of different loan products.

Empirical Analysis

Propensity Score Matching

Recent empirical studies suggest that borrowers take out subprime mortgages based on their credit score, income, payment history, level of down payment, debt ratios, and loan size limits; there is mixed evidence on the effect of demographics (Courchane et al. 2004; Cutts and Van Order, 2005; Chomsisengphet and Pennington-Cross, 2006; LaCour-Little, 2007). Based on the literature review, we included the key underwriting factors of FICO score and DTI in our analysis. These variables are assumed to directly affect credit risk and therefore affect mortgage choice/assignment, since higher credit risk is hypothesized to be associated with a greater probability of taking out a subprime mortgage. For example, lower FICO scores are assumed to be associated with higher credit risk, so we expect subprime loans to capture the majority of the borrowers with lower FICO scores. LTV, another important underwriting variable, is also generally considered to raise endogeniety concerns (LaCour-Little, 2007). In this case, higher LTV is one distinct characteristic of most CAP loans, with over 82 percent of CAP loans having an LTV equal to or higher than 97 percent. By contrast, most subprime loans have an LTV of less than 90 percent. Courchane et al. (2004) also suggest that high LTV may be associated with higher risk but is not necessarily associated with getting a subprime mortgage. Because our focus is the impact of borrower and neighborhood characteristics on borrowers' choice/assignment of mortgages, we decided not to include LTV variables in the model.[12]

In addition to the underwriting variables, we included loan amount as an explanatory variable since fixed costs are usually a large component of loan originations. We further included several factors measuring local market dynamics and credit risk. We constructed a zip-code-level credit risk measure: the mean FICO score for mortgages originated in the preceding year from the LPS data. Our hypothesis is that subprime lenders tend to market in neighborhoods or areas with a larger share of potential borrowers who have impaired credit history. The zip-code educational distribution was included as a proxy of residents' financial knowledge and literacy. Because some literature suggests that subprime lending is more likely to be concentrated in minority neighborhoods (Calem et al. 2004), we included the share of minority in the zip code in the models. Furthermore, we constructed a zip-code-level HHI using HMDA data to measure the extent of competition in the market in which borrowers' properties are located.[13] The HHI measure also partially represents the volume of transactions in the area, since more transactions in a hot market could, though not necessarily would, attract more lenders to the market. In addition, we included quarterly calendar dummy variables to account for fluctuations in the yield curve that could affect market dynamics.

Exhibit 5 presents the results from logistic regression models for different vintages. Across different years, credit risk measures are highly predictive: borrower FICO score, coded into buckets with above 720 as the holdout category, is highly predictive of the use of subprime loans; coefficients are relatively large and decrease monotonically as credit score categories increase. In other words, as expected, the higher the FICO score, the lower the probability of taking out a subprime mortgage. Compared to those with very high DTI (>42 percent), borrowers with lower DTIs are generally less likely to receive subprime loans; exceptions are the buckets with low DTI (<28 percent) for the 2005 and 2006 samples. While it seems CAP borrowers had very high DTIs in 2006, the results generally suggest that borrowers with very high DTIs are more likely to receive subprime loans. In all the models, loan amount is positive for the use of subprime loans, consistent with the hypothesis that subprime borrowing involves higher costs, with costs being driven by large fixed components.

Further, zip-code-level average credit score is statistically significant and negatively related to the probability of taking out a subprime mortgage, suggesting that borrowers in areas with a higher share of low-score population are more likely to receive subprime loans. Zip-code-level education performs about as expected, with higher educational attainment roughly associated with a reduced probability of receiving a subprime mortgage. Borrowers in areas with a higher share of minorities are more likely to use subprime mortgages. Finally, higher HHIs are associated with a lower probability of taking out a subprime mortgage—suggesting that, at least in the period from 2003-2006, subprime loans were more likely to be in the markets with more intensive competition and/or more transactions.

In this analysis, we defined the logit rather than the predicted probability as the propensity score, because the logit is approximately normally distributed. For the one-to-one nearest neighbor with caliper match, we selected the subprime loan with the closest propensity score within a caliper for the first CAP loan after the subprime and randomly ordered CAP loans. We then removed both cases from further consideration and continued to select the subprime loan to match the next CAP loan. For the one-to-many match, we matched subprime loans with CAP loans with the closest propensity score within a caliper after all the loans were randomly sorted. Instead of removing the matched cases after matching, as in the one-to-one match, we kept the matched CAP loans in the sample and continued to find the matching CAP loans for the next subprime loan. This allows us to match as many subprime loans as possible for each CAP loan. We tried two different calipers, 0.1 and 0.25 times of standard error as suggested by Rosenbaum and Rubin (1985). In other words, we tried two matching algorithms, allowing us to match one CAP loan with one or multiple subprime loans, and two caliper sizes, allowing us to test the sensitivity of the findings to varying sizes. For the one-to-many matched sample, to ensure that our analysis is representative of the matched set, we apply a system of weights, where the weight is the inverse of the number of subprime loans that matched to one single CAP loan.

Exhibit 6 describes the four matching schemes and numbers of loans for the resamples: Match 1 and Match 2 are based on the one-to-one match; Match 3 and Match 4 are based on the one-to-many match. Match 1 and Match 3 use nearest neighbor matching within a more restrictive caliper of 0.1, while other matching schemes employ a wider caliper (0.25 times of the standard deviation of the propensity scores). The results show that the more restrictive caliper does not dramatically reduce the sample size; we lost about 791 cases (12 percent) from Match 2 to Match 1 and only one CAP loan from Match 4 to Match 3. Because the qualitative results do not change and a restrictive caliper can lower the likelihood of producing inexact matching, we focused on the schemes using the more restrictive caliper size of 0.1 (Matches 1 and 3) in our analysis of loan performance. For the one-to-one match (Match 1), we ended up with a sample of 5,558 CAP loans and 5,558 matching subprime loans. For the one-to-many match, the sample was 35,971 subprime loans matched to 3,943 CAP loans (Match 3).

We checked covariate distributions after matching. Both Match 1 and Match 3 remove all significant differences, except LTV variables, between groups. For the matched groups, as Exhibit 7 shows, borrowers are remarkably similar across all groups except for LTV ratios, and we got a reduced but more balanced sample of CAP and subprime borrowers. Compared to CAP loans, which are usually fixed-rate retail loans with no prepayment penalty, subprime loans have distinctive features and terms. A vast majority (86 percent) of subprime loans are adjustable rate mortgages; most (70 percent) were obtained through brokers; and many (41 percent) have prepayment penalties.

Performance of the Matched Sample

We turn now to the comparison of CAP loans and subprime loans with similar characteristics. For the matched sample, we observed the payment history during the period from loan origination to March 2008. During this period, CAP loans had a lower serious delinquency rate: only 9.0 percent had ever experienced 90-day delinquencies before March 2008, compared to 19.8 percent of comparable subprime loans (Exhibit 8). Subprime loans also had a higher prepayment rate, 38 percent compared to about 18 percent for the matched CAP loans.

In addition to the subprime variables, we considered in the MNL model important underwriting variables, including borrower DTI ratio, credit history, loan age, and loan amount, as well as the put option. According to the option-based theory, home equity plays a central role in determining the probability of foreclosure (Quercia and Stegman 1992). The value of the put option is proxied by the ratio of negative equity (unpaid mortgage balance minus estimated house price based on the house price index of the Office of Federal Housing Enterprise Oversight) to the original house price. We recognize that the inclusion of the put option may overestimate the risk of subprime loans since, as suggested in Zelman, McGill, Speer and Ratner (2007), some subprime loans may have second mortgages that were not captured here. We tried the same models without the put option variable; although the estimated default rate for the subprime loans is smaller, the qualitative results are fairly consistent with those in Exhibit 9 and Exhibit 10.

Falling interest rates may lead to faster prepayments and drive down delinquency rates as borrowers refinance their way out of potential problems. Rising interest rates can cause payment shocks at the reset date for adjustable-rate mortgages and reduce the ability of borrowers to afford a fixed-rate refinance. To capture the change in interest rate environment, we used the difference between the prevailing interest rates, which is proxied by the average interest rate of 30-year fixed-rate mortgages from the Freddie Mac Primary Mortgage Market Survey (PMMS), and the temporal average of the prevailing interest rates during the study period (Q1 2003 to Q1 2008).

Consistent with prior work, we further separated the matched sample into two cohorts based on years of origination. Subprime loans that originated in 2003 and 2004 were underwritten during a time of historically low interest rates and a strong economy, leading to a relatively good performance with very low default rates (Cutts and Merrill, 2008). Many borrowers were able to refinance their mortgages or sell their houses because of lax underwriting and high house price appreciation before 2007, which extinguished the default option. Instead, subprime loans that originated in 2005 and 2006, especially subprime ARMs, have not performed as well. These two cohorts capture some unobservable heterogeneity characterizing mortgages that originated in a booming housing market and those that originated in a softening housing market.

The results from the MNL regressions based on different matching samples are listed in Exhibit 9 (one-to-one match) and Exhibit 10 (one-to-many match). Model 1 considers the subprime dummy variable only, while Model 2 helps us explain the difference in performance between CAP and subprime loans. The results-based samples using varying algorithms are quite consistent; estimated coefficients for the explanatory variables are of the same sign and similar size, so Exhibit 10 only lists results for the subprime variables. Except for a few insignificant coefficients for the prepayment outcome, the subprime variables are significant and have expected signs. It is not easy to interpret the results based on the coefficients from the MNL regressions directly. We estimated the cumulative default and prepayment rates in the first 24 months after origination for borrowers with impaired credit score (FICO score 580-620) and with mean value of other regressors, except loan age and loan characteristics, based on the MNL regression results. The estimation results discussed below are listed in Exhibit 11, where we consider a 90-day delinquency as termination of a loan, although it may still be active after the delinquency.

Summary of Primary Findings

First of all, there is consistent evidence that subprime loans have a higher default risk and a higher prepayment probability than CAP loans. The estimated cumulative default rate for a 2004 subprime loan is 16.3 percent, about four times that of CAP loans (4.1 percent). For a 2006 subprime loan, the cumulative default rate is over 47.0 percent, about 3.5 times that of comparable CAP loans (13.3 percent). In other words, CAP loans are over 70 percent *less* likely to default than a comparable subprime loan across different vintages. We also notice that the default rate of the 2005-2006 cohort is significantly higher than that of the 2003-2004 cohort for loans with same loan features. Very likely this is because of changes in the underwriting standard and in economic conditions, as well as other unobservable heterogeneity.

We also found that subprime loans with adjustable rates have a significantly higher default rate than comparable CAP loans. And when the adjustable rate term is combined with the prepayment-penalty feature, the default risk of subprime loans becomes even higher. For a 2004 *sub_arm* loan (retail-originated subprime ARM without prepayment penalty), the estimated cumulative default rate would be 6.5 percent, slightly higher than that of CAP loans (4.1 percent). But if the adjustable rate subprime mortgage has a prepayment penalty, the estimated default rate increases to 13.5 percent for a 2004 *sub_arm&ppp* loan (retail-originated subprime ARM with prepayment penalty), over 100 percent higher than that of *sub_arm*. The same pattern also holds for the 2006 originations.

Finally, we found that the broker-origination channel is significantly associated with an increased level of default. For example, the estimated cumulative default rate for a 2004 *sub_bro&arm* loan (broker-originated adjustable-rate subprime loan without prepayment penalty) is 17.3 percent, significantly higher than the 6.5 percent of the *sub_arm* loans. For a 2006 *sub_bro&arm* loan, the estimated cumulative default rate is as high as 51 percent, much higher than the 16.8 percent of the *sub_arm* loans. The same pattern can also identified for adjustable-rate subprime loans with prepayment penalties. When a broker-originated subprime ARM has the term of prepayment penalty, the default risk for 2004 originations is 5.1 times as high as that of CAP loans (21.8 percent vs. 4.1 percent) and for 2006 originations 4.0 times as high (53.8 percent vs. 13.3 percent).

The results suggest that, all other characteristics being equal, borrowers are three to five times more likely to default if they obtained their mortgages through brokers. When this feature is combined with the adjustable rate and/or prepayment penalty, the default risk is even higher. One possible explanation is that, as suggested in Ernst et al. (2008) and Woodward (2008), loans originated through brokers have significantly higher closing costs and prices, which increases borrowers' costs and can lead to elevated default risk. It is also possible that borrowers obtaining loans through brokers are more likely to receive products with features that may increase the default risk. Finally, it is very likely that the broker-origination channel has a looser underwriting standard that has not been fully captured by the model, which allows unqualified borrowers to receive unsustainable risky products. All these contentions are consistent with the results, and additional research is needed to examine this issue in more detail.

As to the outcome of prepayment, we observed two obvious trends. The first is that subprime loans, especially subprime ARMs, have a significantly higher prepayment rate than CAP loans (Exhibit 11). Second, for recent originations (2005-2006), subprime loans with prepayment penalties are less likely to prepay than loans with similar terms but without prepayment penalties. But for early originations (2003-2004), the pattern is reversed: subprime loans with prepayment penalties have a higher prepayment rate, probably because they are more likely to be prepaid after the prepayment penalty period has expired. Although we were not able to determine the prepayment penalty clauses for all subprime loans because of missing values, for those loans with complete information prepayment penalties were most frequently levied within the first two to three years of loan origination. As of March 2008, then, most prepayment penalties for 2003-2004 originations have expired. But prepayment may also be part of the problem if the borrower prepaid the loans by refinancing into another subprime product.

The Impact of Unobservable Heterogeneity

To check how unobservable borrower risk characteristics impact the results, we treated unobservable heterogeneity as an omitted variable, and solved this problem by including a proxy of the omitted variable as a regressor in the outcome equation along with the subprime dummy and other controls. Our first proxy of borrower unobservable heterogeneity is the risk premium (*rate_sp*), which is the mortgage interest rate minus the national average rate of 30-year fixed-rate mortgages from the PMMS. Of course, the risk premium variable may be an endogenous variable here, because if priced properly mortgage interest rates are determined by an assessment of a borrower's risk profile and some mortgage characteristics. To address the endogeneity issue, we used the residue of the risk premium (*rate_resid*) as a proxy of the unobservable lender/borrower risk characteristics based on an OLS model using observable information to predict mortgage risk premium.[14]

The qualitative results generally do not change when the proxies of unobservable heterogeneity are considered (Model 3 and Model 4 in Exhibit 12). The inclusion of the risk premium variables seems help explain the borrowers' prepayment behavior but not the default behavior. The coefficients of the subprime variables for the default option vary only slightly and have the same significance in different models. The noticeable difference is that for prepayment option once the risk premium variables are controlled, the coefficients of the subprime variables become much smaller for the 2005-2006 cohort but the signs and significance are the same. The coefficients of the risk premium variables (*rate_sp* and *rate_resid*) are generally insignificant for the default option (with only one exception of the 2003-2004 cohort which is slightly significant). As to the prepayment option, risk premium variables have a positive impact on the probability of prepayment for the 2005-2006 cohort but have a negative impact, though with a magnitude close to zero, for the 2003-2004 cohort, possibly because of changes in some uncaptured market condition information.

In summary, we demonstrate that the results we obtained earlier are robust enough even after controlling for proxies of the unobservable heterogeneity among borrowers. As a result, we are more confident about the conclusions about the relative risk of different loan products.

Empirical Results of Other Controls

Because the results for most of the variables are generally consistent across different models, discussion of other control variables is based primarily on Model 1, as summarized in Exhibit 9. For other controlled variables, the results suggest:

Other risk variables

- Put option: Borrowers with less or negative equity in their homes (larger value of *put*) are more likely to default and less likely to prepay. The results confirm the common wisdom that the level of equity in a home is a strong predictor for prepayment and default.
- Credit history: As expected, there is consistent evidence that borrowers with lower credit scores are more likely to experience serious delinquency.
- Debt-to-income ratio: Higher debt-to-income ratios are associated with a higher default risk for the 2003-2004 cohort, but the coefficients are insignificant for the 2005-2006 sample.

Loan characteristics

- Size of unpaid balance: Larger loan size is generally associated with lower default risk. Larger loan size is also associated with higher prepayment probability for the 2003-2004 cohort.

Area and neighborhood controls

- Area credit risk: Average credit score in the zip code is significantly and negatively associated with default risk. There is also some evidence that zip code average credit score is positively associated with prepayment probability (for the 2005-2006 vintage).
- Interest rate dynamics: For different cohorts, the impact of interest rate environment is different. For the 2003-2004 cohort, the increase in average interest rate decreases the prepayment probability but for the recent cohort, the increase in average interest rate increases the default risk and has no significant impact on the prepayment probability.
- County unemployment rate: Average county unemployment rate is generally insignificant in explaining the default and prepayment behaviors across different models.

Time dummies

- Dummies of 2003 and 2005 originations: The 2005 originations are significantly less likely to default, compared to the 2006 cohort.

Conclusions

As the current economic crisis worsens, the debate continues as to what cause the initial foreclosure crisis in the mortgage markets. In this study, we examine the relative default risk of two of the suspects: subprime mortgages and community reinvestment loans. Using propensity matching methods, we constructed a sample of comparable borrowers with similar risk characteristics but holding the two different loan products. We found that, for comparable borrowers, the estimated default risk is much lower with a CRA loan than with a subprime mortgage. More narrowly, we found that the broker-origination channel, an adjustable rate, and a prepayment penalty, all contribute substantially to the elevated risk of default among subprime loans. In the worst scenario, when broker origination is combined with the features of adjustable rate and prepayment penalty, the default risk of a borrower is four to five times as high as that of a comparable borrower holding a CRA-type product. Though CAP has some program-specific features, the results clearly suggest that the relative higher default risk of subprime loans may not be solely attributed to borrower credit risk, instead it is significantly associated with the characteristics of the products and the origination channel in the subprime market. Thus, the results suggest that when done right and responsibly, lending to LMI borrowers is viable proposition. Borrowers and responsible CRA lending should not be blamed for the current housing crisis.

Our results are consistent with recent regulatory action.[15] Key features of subprime loans—underwriting that ignores ability to pay, the inclusion of prepayment penalties, escalating interest rates and hidden fees--make it difficult for families to stay current on their mortgage payments. Federal Reserve rules issued in 2006 and recent amendments to the Truth in Lending Act (Regulation Z) have banned negative amortization for high-priced loans and most prepayment penalties. They have also banned underwriting loans without regard to a borrower's ability to pay. Unfortunately, broker origination also significantly increases default risk. However, there is no Federal law and only a few states have sufficiently regulated the incentive structure of the broker origination channel, especially the yield spread premium which many have argued may lead brokers to originate loans that may not be in the best interest of the borrower.[16]

In the current economic situation, many borrowers holding subprime mortgages with risky loan features are having difficulty making their current payments and many have already been seriously delinquent or in default. One proposed solution has been to modify troubled owner-occupied subprime loans with FHA-insured loans or more sustainable fixed-rate products at a significant discount (Inside B&C Lending, 2008). This research demonstrates that if subprime-like borrowers receive loans with prime rather than subprime terms and conditions, their default rate would be much lower.

Because the mortgage industry was originally criticized for failing to serve lower-income and minority households and more recently for flooding the market with unsustainable mortgages with risky features, our findings are important for policymakers. This research suggests that loans with prime terms and conditions offered through special CRA lending programs provide LMI and minority households, even those with somewhat imperfect credit histories, more sustainable homeownership options than subprime loans.

While our results are interesting for understanding the performance difference between subprime and CRA loans, we would like to emphasize that CAP has some program specific characteristics. Though national in scope, CAP is geographically concentrated in certain markets. In addition, this analysis focuses solely on home purchase lending activities and borrowers with full or alternative documentation only. As such, it is unclear whether or not our findings for the CAP program are applicable to national population of CRA loans and the entire subprime market. However, CAP borrowers are matched with subprime borrowers with similar risk profiles, focusing in this way on the less risky portion of the subprime market. We have also excluded from the analysis investor loans and low- or no-doc subprime mortgages, all of which are generally associated with a higher credit risk. Further, if borrowers are indeed steered to low- and no-doc loans in the subprime market even when they could have documented their income, as has been asserted by some observers, this would suggest that the increased risk of having one's mortgage originate in the subprime market is even greater than captured in this paper. As such, this research provides more convincing evidence of the relative risk of the CRA-type loans and the impact of loan features and origination channels on loan performance.

Endnotes:

[1] For more complete details of CRA regulations, see http://www.ffiec.gov/cra/default.html.

[2] The CRA assessment area for a retail-oriented banking institution must include "the areas in which the institution operates branches and deposit-taking automated teller machines and any surrounding areas in which it originated or purchased a substantial portion of its loans" (Avery et al. 2000, p. 712).

[3] As Alexander et al. (2002) suggest that some practices of possible gaming of brokers with lenders include at least reporting the highest FICO score from the three bureaus, pulling a FICO score after challenging a derogatory, and shopping for cooperative appraisers.

[4] Examples of guidelines that reduced cash required to close include: Lesser of $500 or 1percent from borrower's own funds; Maximum LTV of 98percent and maximum combined LTV (including soft seconds) of 103percent; No reserves required.

[5] Examples of guideline flexibility with respect to credit history include: Demonstrate 6-month satisfactory payment history with four sources of credit, either traditional or non-traditional; FICO scores thresholds below 620 accepted in certain programs.

[6] Examples of underwriting flexibility in assessing the ability to repay include: Maximum total ratio of debt payments to income ratio of 43 percent, or up to 45 percent if new housing payment is not more than 25 percent higher than prior housing payment.

[7] Examples of offsetting or combined guideline flexibilities include: Maximum total ratio of debt payments to income varies from 38percent to 48percent with borrowers with higher credit scores allowed higher ratios; Higher downpayments or reserve requirements for borrowers with FICO below 620.

[8] To qualify for the CAP program, borrowers must meet one of three criteria: (1) have income under 80 percent of the area median income (AMI) for the metropolitan area; (2) be a minority with income below 115 percent of AMI; (3) or purchase a home in a high-minority (>30%) or low-income (<80% AMI) census tract and have an income below 115 percent AMI.

[9] Other common matching algorithms include: nearest-neighbor matching, kernel matching, local linear matching, Mahalanobis metric matching, Mahalanobis metric matching including the propensity score, and difference in differences methods (see review in Guo et al. 2006 and Essama-Nssah, 2006).

[10] Unfortunately, there are too few loans in the matched sample for retail-originated fixed-rate mortgages (less than 20 for the one-to-one match for each category), which does not allow us to conduct meaningful analysis, and so they were dropped from further analysis.

[11] This assumption is also known as the *exogeneity,* or *unconfoundedness,* or *ignorable treatment assignment,* or *conditional homogeneity, or selection on observables* assumption *(*Essama-Nssah, 2006).

[12] To empirically test the impact on results of including/excluding LTV variables, we tried logistic regression models with LTV variables. As expected, LTV ratio is highly significant in predicting the use of subprime loans, with lower LTVs consistently and monotonically related to the use of subprime loans. The match rate is lower than those reported in Exhibit 6, but the qualitative results on the performance of mortgages do not change.

[13] The HHI is constructed as the sum of squared market shares of firms in a zip code. Based on HMDA data, we got the market share of firms in a census tract and then matched to corresponding zip codes. When a census tract overlaps multiple zip codes, we assume the share of loans for the particular firm is the same as the share of house units of the tract in this zip code. As such, the index ranges from 10,000 in the case of 100% market concentration to near zero in the case of many firms with equally small market shares.

[14] We assume mortgage risk premium is determined by a set of borrower, neighborhood characteristics in the propensity score estimation and loan characteristics that may influence pricing including LTV, adjustable rates, and prepayment penalties. We ran OLS regressions for different cohorts and the R squares of the four regressions range from 0.4 for the 2004 cohort to 0.61 for the 2003 cohort. The regression results are available upon request.

[15] Home Ownership and Equity Protection Act bans balloon payments, negative amortization, most prepayment penalties for high-rate/high-fee loans. The Revision of Regulation Z of Truth in Lending Act in July 2008 further bans any prepayment penalties if the payment can change in the initial four years and for high-priced loans prepayment penalties cannot last for more than two years.

[16] Effective on October 1, the House Bill 2188 in North Carolina bans rate or yield spread premiums.

Reference:

Alexander, W.P., S.D. Grimshaw, G.R. McQueen and B.A. Slade. Some Loans Are More Equal than Others: Third-Party Originations and Defaults in the Subprime Mortgage Industry. *Real Estate Economics,* 2002, 30(4): 667-697.

Ambrose, B.W., M. LaCour-Little and Z.R. Huszar. A Note on Hybrid Mortgages. *Real Estate Economics*, 2005, 33(4): 765-782.

Apgar, W.C. and M. Duda. The Twenty-Fifth Anniversary of the Community Reinvestment Act: Past Accomplishments and Future Regulatory Challenges. *FRBNY Economic Policy Review,* 2003, June: 169-191.

Avery, R.B., R.W. Bostic and G.B. Canner. CRA Special Lending Programs, *Federal Reserve Bulletin*, 2000, 86: 711-731.

Bernanke, B.S. The Community Reinvestment Act: Its Evolution and New Challenges. Speech, The Community Affairs Research Conference, Washington, D.C. March 30, 2007, http://www.federalreserve.gov/newsevents/speech/Bernanke20070330a.htm.

Brooks, R. and R. Simon. Subprime Debacle Traps Even Very Credit-Worthy As Housing Boomed, Industry Pushed Loans To a Broader Market, *Wall Street Journal,* 2007, Dec. 3, A1.

Bocian, D., K. Ernst and W. Li. Unfair lending: The effect of race and ethnicity on the price of subprime mortgages. Federal Reserve System Community Affairs Research Conference, Washington, DC, 2007.

Calem, P.S., K. Gillen and S.M. Wachter. The Neighborhood Distribution of Subprime Mortgage Lending. *Journal of Real Estate Finance and Economics,* 2004, 29 (4): 393-410.

Calhoun, C.A. and Y. Deng. A Dynamic Analysis of Fixed- and Adjustable-Rate Mortgage Terminations. *Journal of Real Estate Finance and Economics*, 2002, 24(1/2): 9-33.

Chomsisengphet, S. and A. Penning-Cross. The Evolution of the Subprime Mortgage Market. *Federal Reserve Bank of St. Louis Review*, 2006, January/February: 31-56.

Cravatts, R.L. Unintended fallout of loan crisis, *The Boston Globe*, 2008, July 26.

Courchane, M.J. The Pricing of Home Mortgage Loans to Minority Borrowers, How Much of the APR Differential Can We Explain? *Journal of Real Estate Research,* 2007. 29(4): 399-440.

Courchane, M.J., B.J. Surette and P.M. Zorn. Subprime Borrowers, Mortgage Transitions and Outcomes. *Journal of Real Estate Finance and Economics*, 2004, 29(4): 365-392.

Cutts, A.C. and Merrill, W.A. Interventions in Mortgage Default: Policies and Practices to Prevent Home Loss and Lower Costs. in Retsinas, N.P. and E.S. Belsky, eds. *Borrowing to Live: Consumer and Mortgage Credit Revisited*, 2008, pp. 203-254, Brookings Institution Press. Washington, DC

Cutts, A.C. and R. Van Order. 2005. On the Economics of Subprime Lending, *Journal of Real Estate Finance and Economics*, 2005, 30(2): 167–196.

Danis, M.A. and A. Pennington-Cross. The Delinquency of Subprime Mortgages. Federal Reserve Bank of St. Louis, Research Division Working Paper, 2005-022A, 2005.

Dehejia, R.H. and W. Sadek. Propensity Score-Matching Methods for Nonexperimental Causal Studies. *The Review of Economics and Statistics*, 2002, 84(1): 151-161.

Ernst, K., D. Bocian and W. Li. Steered Wrong, Brokers, Borrowers, and Subprime Loans. Center for Responsible Lending, 2008, http://www.responsiblelending.org/research/steered-wrong-brokers-borrowers-and-subprime-l-oans.html.

Essama-Nssah, B. Propensity Score Matching and Policy Impact Analysis: A Demonstration in EViews**,** *World Bank Policy Research Working Paper*, No. 3877, 2006.

Gerardi, K., A.H. Shapiro and P. Willen. 2007. Subprime Outcomes, Risky Mortgages, Homeownership Experiences, and Foreclosures. *Federal Reserve Bank of Boston Working Papers* 07-15, 2007.

Glenn B.C., E.Laderman, A. Lehnert and W. Passmore. Does the Community Reinvestment Act (CRA) cause banks to provide a subsidy to some mortgage borrowers? *Finance and Economics Discussion Series* 2002-19, 2002.

Guo, S., R.P. Barth and C. Gibbons. Propensity Score Matching Strategies for Evaluating Substance Abuse Services for Child Welfare Clients. *Children and Youth Services Review*, 2006, 28(4): 357-383.

Hudson, M. and E.S. Reckard. More Homeowners with Good Credit Getting Stuck with Higher-Rate Loans. *Los Angeles Times*, 2005, October 24.

Immergluck, D. From the Subprime to the Exotic, Excessive Mortgage Market Risk and Foreclosures. *Journal of the American Planning Association*, 2008, 74 (1): 59-76.

Inside B&C Lending. FDIC’s Mod Plan a Possible Model for Subprime Servicing, *Inside B&C Lending*, 2008, August 29: 4-5.

Inside Mortgage Finance. *The 2008 Mortgage Market Statistical Annual*. Inside Mortgage Finance Publications: Bethesda, MD, 2008.

Keys, B.J., T.K. Mukherjee, A. Seru and V. Vig. Did Securitization Lead to Lax Screening? Evidence from Subprime Loans (April 2008). EFA 2008 Athens Meetings Paper. 2008. Available at SSRN: http://ssrn.com/abstract=1093137.

Krauthammer, C. Catharsis, Then Common Sense, *Washington Post,* 2008, September 26, A23.

LaCour-Little, M. and Chun, G.H. Third Party Originators and Mortgage Prepayment Risk: An Agency Problem *Journal of Real Estate Research*, 1999, 17(1/2): 55-70.

LaCour-Little, M. The Home Mortgage Preferences of Low-and-Moderate Income Households. *Real Estate Economics*, 2007. 35(3): 265-290.

Mian, A. and A. Sufi. The Consequences of Mortgage Credit Expansion: Evidence from the 2007 Mortgage Default Crisis, 2008. Available at SSRN: http://ssrn.com/abstract=1072304.

Morganson, G. *Inside the Countrywide Spending Spree*, *New York Times*, August 26, 2008.

Mortgage Bankers Association (MBA). National Delinquency Survey, Second Quarter 2008, 2008.

Pennington-Cross, A. Credit History and the Performance of Prime and Nonprime Mortgages. *Journal of Real Estate Finance and Economics,* 2003, 27(3): 279–301.

Pennington-Cross, A. and G. Ho. The Termination of Subprime Hybrid and Fixed-Rate Mortgages. *Federal Reserve Bank of St. Louis, Research Division Working Paper* 2006-042A, July 2006.

Poole, W. Reputation and the Non-Prime Mortgage Market. Speech at St. Louis Association of Real Estate Professionals, Saint Louis, MO, July 20, 2007.

Quercia, R.G. and M.A. Stegman. Residential mortgage default: A review of the literature. *Journal of Housing Research*, 1992, 3(2): 341-376.

Quercia, R.G., M.A. Stegman, W.R. Davis and E Stein. The performance of community reinvestment loans: Implications for secondary market purchases. In *Low-income homeownership: Examining the unexamined goal*, edited by N. P. Retsinas and E. S. Belsky. Washington, DC: The Brookings Institution, 2002.

Quercia, R.G., M.A. Stegman and W.R. Davis. The Impact of Predatory Loan Terms on Subprime Foreclosures, The Special Case of Prepayment Penalties and Balloon Payments. *Housing Policy Debate*, 2007, 18(2): 311-346.

Quercia, R.G. and J. Ratcliffe. The Preventable Foreclosure Crisis. *Housing Policy Debate*, 2008, 19(4): 775-783.

Renuart, E. An Overview of the Predatory Mortgage Lending Process. *Housing Policy Debate*, 2004, 15 (3): 467-502.

Rosenbaum, P.R. and D.B. Rubin. The Central Role of the Propensity Score in Observational Studies for Causal Effects. *Biometrika*, 1983, 70: 41-55.

Rosenbaum, P.R. and D.B. Rubin. Constructing a Control Group Using Multivatiae Matched Sampling Methods that Incorporate the Propensity Score. *American Statistician*, 1985, 39: 33-38.

Woodward, S.E. A Study of Closing Costs for FHA Mortgages, U.S. Department of Housing and Urban Development, May 2008.

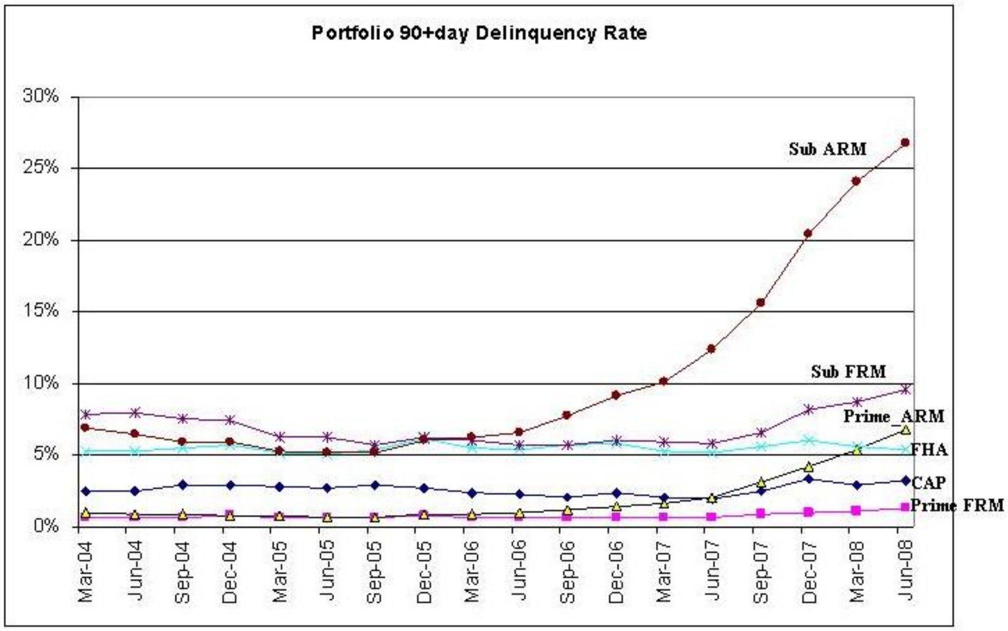
Zelman, I.L., D. McGill, J. Speer and A. Ratner. Mortgage Liquidity du Jour, Underestimated No More. *Credit Suisse,* March 12, 2007.

Exhibit 1 Construction of Subprime Study Sample

	# of Observations
	Subprime
Step 1 Subprime Loans meeting the following criteria: home purchase loans, first-lien; single family house, 30-year amortization, conforming loans with a minimum loan amount of $10,000 only	544,849
Step 2 Exclude loans with no or limited documentation or missing information for the following variables: LTV, Fico score, DTI**,** documentation	86,697
Step 3 Exclude loans not in zip codes with CAP activities and loans without complete payment history	42,065

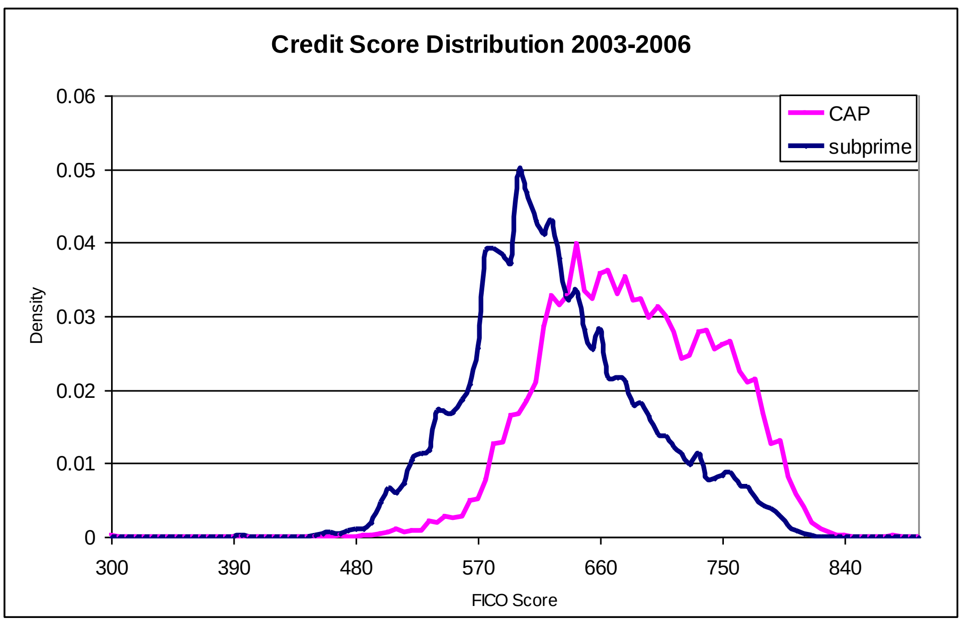
Note: based on authors' calculation from LPS. Subprime loans here include B&C loans and high-cost ARMs (with a margin greater than 300 basis points).

Exhibit 2 90-day Delinquency Rate by Loan Types



Source: Mortgage Banker Association (2008) and Self-Help

Exhibit 3 CAP and Subprime FICO Score Distribution (2003-2006)



Source: Lender Processing Services, Inc. (LPS) and Self-Help

Exhibit 4 Descriptive Statistics (Mean or Percentage)

Variable	CAP	Subprime
Debt-to-income ratio*		
DTI<28%	0.126	0.163
DTI 28-36%	0.278	0.158
DTI 36-42%	0.315	0.178
DTI>42%	0.281	0.501
FICO score*		
<580	0.031	0.213
580-620	0.109	0.263
620-660	0.224	0.225
660-720	0.324	0.192
>=720	0.312	0.107
LTV*		
<80%	0.037	0.369
80-90%	0.050	0.381
90-97%	0.090	0.167
>=97%	0.823	0.083
Loan characteristics		
Loan_amt*	100.86	148.1
ARMs*	-	0.903
Broker*	-	0.808
Prepayment penalty*	-	0.495
Note Rate*	6.66%	7.87%
Neighborhood/Local characteristics		
HHI index (in 10,000, 2005)*	0.051	0.036
Mean area FICO Score (2005)*	688.6	685.2
Share of minority *	0.293	0.482
Education distribution*		
Share of less high school	0.199	0.239
Share of high school	0.318	0.283
Share of some college	0.272	0.292
Share of college and above	0.211	0.186
Geography: top 5 states		
	OH (22.3%)	CA (19.2%)
	NC (14.6%)	TX (11.0%)
	IL (12.6%)	FL (10.1%)
	GA (11.4%)	IL (9.1%)
	OK (5.8%)	GA (5.3%)
Origination Year		
2003	2,670	4,680
2004	2,581	18,380
2005	2,251	11,703
2006	1,719	7,302
N	9,221	42,065

Note: * Bivariate χ^2 test or t test significant at the 0.01 level.

Exhibit 5 Logistic regression models predicting propensity scores

	2003		2004		2005		2006	
	Coef.	P-value	Coef.	P-value	Coef.	P-value	Coef.	P-value
dti<28	-0.172	0.088	0.006	0.941	0.616	0.000	1.324	0.000
dti 28-36	-1.369	0.000	-1.252	0.000	-0.603	0.000	0.216	0.018
dti 36-42	-1.411	0.000	-1.486	0.000	-0.837	0.000	-0.160	0.060
dti>42								
cscore<580	4.632	0.000	3.943	0.000	4.182	0.000	1.900	0.000
cscore 580-620	2.040	0.000	2.237	0.000	2.846	0.000	1.245	0.000
cscore 620-660	1.431	0.000	1.121	0.000	1.438	0.000	1.021	0.000
cscore 660-720	0.850	0.000	0.550	0.000	0.632	0.000	0.483	0.000
cscore >=720								
loan_amt	0.012	0.000	0.013	0.000	0.011	0.000	0.010	0.000
qtr1	0.055	0.585	-0.553	0.000	0.606	0.000	1.137	0.000
qtr2	-0.019	0.843	-0.062	0.407	0.315	0.000	0.891	0.000
qtr3	-0.545	0.000	0.070	0.342	0.073	0.372	0.601	0.000
qtr4								
HHI (in 10,000)	-14.763	0.000	-18.747	0.000	-21.058	0.000	-23.296	0.000
area credit score	-0.004	0.046	-0.004	0.053	-0.002	0.438	0.000	0.937
pctmin	-0.007	0.001	0.006	0.001	0.017	0.000	0.014	0.000
pct_less_high								
pct_high	-0.124	0.000	-0.077	0.000	-0.057	0.000	-0.144	0.000
pct_somecoll	0.062	0.000	0.049	0.000	0.054	0.000	0.015	0.037
pct_coll	-0.082	0.000	-0.067	0.000	-0.058	0.000	-0.092	0.000
_cons	6.015	0.000	5.411	0.000	2.164	0.177	6.127	0.001
Pseudo R^2	0.42		0.36		0.38		0.35	
	N=7,350		N=20,961		N=13,954			N=9,021

Exhibit 6 Description of matching schemes and resample sizes

Scheme	Description of matching method	N of original sample CAP	N of the new sample CAP	N of the new sample Subprime
Match1	Nearest 1-to-1 using caliper=0.1	9,221	5,558	5,558
Match2	Nearest 1-to-1 using caliper=0.25σ	9,221	6,349	6,349
Match3	Nearest 1-to-many using caliper=0.1	9,221	3,943	35,971
Match4	Nearest 1-to-many using caliper=0.25σ	9,221	3,944	36,236

Note: For the one-to-one nearest neighbor with caliper match, the subprime loan with the closest propensity score within a caliper for the first CAP loan was selected after the sample was randomly ordered. We then removed both cases from further consideration and continue to select the subprime loan to match the next CAP loan. For the one-to-many match, subprime loans were matched with CAP loans with the closest propensity score within a caliper after all the loans were randomly sorted. Instead of removing the matched cases after matching as in the one-to-one match, we kept the matched CAP loans in the sample and continued to find the matching CAP loan for the next subprime loan.

Exhibit 7 Significance tests of the resamples

Variable	Match 1		Match3	
Debt-to-income ratio	CAP	Subprime	CAP	Subprime
DTI<28%	0.229	0.221	0.223	0.218
DTI 28-36%	0.261	0.249	0.242	0.233
DTI 36-42%	0.375	0.391	0.397	0.403
DTI>42%	0.135	0.139	0.138	0.146
FICO score				
<580	0.047	0.049	0.165	0.164
580-620	0.15	0.155	0.251	0.241
620-660	0.256	0.241	0.296	0.292
660-720	0.305	0.305	0.165	0.164
>=720	0.242	0.25	0.123	0.139
LTV (* for match 1)				
<80%	0.042	0.314	0.044	0.305
80-90%	0.062	0.276	0.066	0.282
90-97%	0.11	0.209	0.117	0.208
>=97%	0.786	0.201	0.773	0.204
Loan characteristics				
loan_amt*	109.4	109.7	112.0	113.2
ARMs*		0.864		0.880
Broker*		0.696		0.682
Prepayment penalty*		0.413		0.422
Note Rate*	0.066	0.078	0.066	0.078
N	5,558	5,558	3,943	35,971**

Note: * Bivariate χ^2 test or t test significant at 0.01 level. **Statistics based on Match 3 are weighted average and the weight is the inverse of number of subprime loans that matched to one CAP loan.

Exhibit 8 Performance measures of the new samples

	Whole sample		2003-2004 Sample		2005-2006 Sample	
	% of 90-day	% prepayment	% of 90-day	% prepayment	% of 90-day	% prepayment
CAP	8.98	18.46	7.64	25.73	10.94	7.84
Subprime	19.81	38.27	12.97	50.06	29.81	21.04
N	11,116		6,600		4,516	

Note: Observation period is from origination to March 2008; if a loan was 90-day delinquent and then prepaid, it is considered as a 90-day delinquency only.

Exhibit 9 MNL regression results of default and prepayment (Match 1 in Exhibit 6)

		2003-2004 Sample				2005-2006 Sample			
		Model 1		Model 2		Model 1		Model 2	
	Variable	Coef.	P>z	Coef.	P>z	Coef.	P>z	Coef.	P>z
Default	put	0.041	0.000	0.044	0.000	0.050	0.000	0.052	0.000
	dti 28-36	0.581	0.000	0.585	0.000	0.083	0.528	0.093	0.480
	dti 36-42	0.632	0.000	0.599	0.000	0.025	0.847	0.018	0.890
	dti>42	0.323	0.029	0.522	0.000	-0.241	0.065	0.015	0.907
	cscore<580	2.414	0.000	2.196	0.000	1.682	0.000	1.477	0.000
	cscore 580-620	1.991	0.000	1.790	0.000	1.278	0.000	1.057	0.000
	cscore 620-660	1.471	0.000	1.286	0.000	1.033	0.000	0.907	0.000
	cscore 660-720	0.634	0.000	0.512	0.001	0.448	0.004	0.388	0.011
	unpaid balance (in log)	-0.357	0.000	-0.266	0.008	-0.163	0.079	-0.066	0.482
	loan age (in log mon)	1.007	0.000	1.084	0.000	1.043	0.000	1.093	0.000
	area credit score	-0.010	0.000	-0.009	0.000	-0.012	0.000	-0.010	0.000
	average interest rate	-0.128	0.346	-0.142	0.299	0.522	0.000	0.507	0.000
	area unemp rate	0.044	0.120	0.045	0.106	0.045	0.120	0.025	0.393
	y2003 (y2005)	-0.078	0.389	-0.153	0.097	-0.607	0.000	-0.491	0.000
	subprime	1.592	0.000			1.596	0.000		
	sub_arm			0.540	0.004			0.361	0.033
	sub_arm&ppp			1.546	0.028			1.898	0.000
	sub_bro			1.945	0.000			1.446	0.000
	sub_bro&ppp			1.985	0.000			1.527	0.000
	sub_bro&arm			1.661	0.000			1.898	0.000
	sub_bro&arm&ppp			1.987	0.000			1.818	0.000
	cons	0.818	0.544	-0.963	0.482	1.291	0.347	-1.241	0.371
Prepay	put	-0.015	0.000	-0.013	0.000	-0.007	0.061	-0.006	0.185
	dti 28-36	0.289	0.000	0.301	0.000	-0.045	0.760	0.015	0.920
	dti 36-42	0.348	0.000	0.354	0.000	0.058	0.683	0.149	0.311
	dti>42	0.015	0.825	0.119	0.088	-0.300	0.030	-0.175	0.248
	cscore<580	0.142	0.322	-0.001	0.996	-0.090	0.663	-0.012	0.956
	cscore 580-620	0.080	0.321	-0.006	0.945	0.237	0.069	0.274	0.045
	cscore 620-660	0.323	0.000	0.262	0.000	-0.193	0.131	-0.140	0.285
	cscore 660-720	0.149	0.005	0.139	0.008	-0.076	0.521	-0.114	0.344
	unpaid balance (in log)	0.329	0.000	0.298	0.000	-0.055	0.537	-0.117	0.201
	loan age (in log mon)	0.459	0.000	0.512	0.000	0.697	0.000	0.699	0.000
	area credit score	0.001	0.381	0.002	0.091	0.007	0.001	0.008	0.001
	average interest rate	-0.197	0.007	-0.187	0.011	0.200	0.203	0.188	0.237
	area unemp rate	-0.016	0.338	-0.022	0.185	-0.029	0.409	-0.031	0.375
	y2003 (y2005)	-0.021	0.640	0.029	0.519	0.278	0.003	0.317	0.002
	subprime	0.922	0.000			1.238	0.000		
	sub_arm			0.611	0.000			1.132	0.000
	sub_arm&ppp			1.685	0.000			2.289	0.000
	sub_bro			0.437	0.000			1.207	0.001
	sub_bro&ppp			0.979	0.000			-0.241	0.510
	sub_bro&arm			1.080	0.000			1.660	0.000
	sub_bro&arm&ppp			1.340	0.000			0.947	0.000
	cons	-11.241	0.000	-11.612	0.000	-11.908	0.000	-11.495	0.000
	Log likelihood		-16790.3		-16683.1		-8262.9		-8157.0
	N		N=192,179 of 6,600 loans				N=93,646 of 4,516 loans		

Note: *sub_arm* represents subprime retail originated ARMs without prepayment penalty;
sub_ arm&ppp represents subprime retail originated ARMs with prepayment penalties;
sub_bro represents subprime broker originated fixed-rate mortgages without prepayment penalties; *sub_bro&ppp* represents subprime broker originated fixed-rate mortgages with prepayment penalties;
sub_bro&arm represents subprime broker originated ARMs without prepayment penalties; *sub_bro&arm&ppp* represents subprime broker originated ARMs with prepayment penalties.

Exhibit 10 MNL regression results of default and prepayment (Match 3 in Exhibit 6)

		2003-2004 Sample				2005-2006 Sample			
		Model 1		Model 2		Model 1		Model 2	
	Variable	Coef.	P>z	Coef.	P>z	Coef.	P>z	Coef.	P>z
Default	subprime	1.448	0.000			1.616	0.000		
	sub_arm			0.482	0.003			0.189	0.208
	sub_arm&ppp			1.658	0.000			2.073	0.000
	sub_bro			1.721	0.000			1.418	0.000
	sub_bro&ppp			1.770	0.000			1.581	0.000
	sub_bro&arm			1.638	0.000			1.906	0.000
	sub_bro&arm&ppp			1.843	0.000			1.833	0.000
	cap								
Prepay	subprime	0.940	0.000			1.308	0.000		
	sub_arm			0.666	0.000			1.192	0.000
	sub_arm&ppp			1.544	0.000			2.220	0.000
	sub_bro			0.510	0.000			1.235	0.000
	sub_bro&ppp			0.901	0.000			-0.451	0.111
	sub_bro&arm			1.052	0.000			1.751	0.000
	sub_bro&arm&ppp			1.385	0.000			1.073	0.000
	cap								
	N	N=341,367 of 16,604 loans				N= 528,292 of 23,310 loans			

Note: see note in Exhibit 9 for the definition of different loan products.
There should be 8 dummies for different combinations of loan features but the sample sizes of the buckets of retail-originated fixed-rate subprime with and without prepayments are too small, which does not allow us conduct meaningful analysis.

Exhibit 11 Estimated cumulative default and prepayment rate
(24 months after origination for a borrower with impaired credit score of 580-620)

	2004 Origination		2006 Origination	
	Default	prepayment	Default	prepayment
CAP	4.08%	10.34%	13.32%	7.47%
Subprime	16.28%	22.81%	47.04%	17.69%
sub_arm	6.53%	17.93%	16.82%	20.82%
sub_arm&ppp	13.48%	41.43%	43.30%	39.42%
sub_bro	24.15%	13.92%	40.61%	18.76%
sub_bro&ppp	23.33%	22.48%	47.84%	4.74%
sub_bro&arm	17.30%	25.37%	51.00%	24.27%
sub_bro&arm&ppp	21.82%	30.40%	53.78%	13.36%

Note: see note in Exhibit 9 for the definition of different loan products. The predicted cumulative default and prepayment rate is as of 24 months after origination for a borrower with a FICO score between 580-620 and holding a mortgage originated in 2004 or 2006, with the mean value of other regressors. The estimation is based on regression results in Exhibit 9.

Exhibit 12 MNL Regression Results of Default and Prepayment
(with proxy of unobservable heterogeneity)

		2003-2004 Sample						2005-2006 Sample					
		Model 1		Model 3		Model 4		Model 1		Model 3		Model 4	
	Variable	Coef.	P>z	Coef.	P>z	Coef.	P>z	Coef.	P>z	Coef.	P>z		
Default	put	0.041	0.000	0.041	0.000	0.041	0.000	0.050	0.000	0.049	0.000	0.049	0.000
	dti 28-36	0.581	0.000	0.571	0.000	0.582	0.000	0.083	0.528	0.081	0.543	0.078	0.560
	dti 36-42	0.632	0.000	0.606	0.000	0.632	0.000	0.025	0.847	0.024	0.859	0.018	0.893
	dti>42	0.323	0.029	0.349	0.019	0.323	0.030	-0.241	0.065	-0.232	0.077	-0.243	0.063
	cscore<580	2.414	0.000	2.271	0.000	2.413	0.000	1.682	0.000	1.628	0.000	1.690	0.000
	cscore 580-620	1.991	0.000	1.921	0.000	1.990	0.000	1.278	0.000	1.237	0.000	1.274	0.000
	cscore 620-660	1.471	0.000	1.422	0.000	1.471	0.000	1.033	0.000	1.015	0.000	1.032	0.000
	cscore 660-720	0.634	0.000	0.614	0.000	0.634	0.000	0.448	0.004	0.441	0.004	0.448	0.004
	unpaid balance (in	-0.357	0.000	-0.308	0.002	-0.357	0.000	-0.163	0.079	-0.122	0.240	-0.152	0.114
	loan age (in log mon)	1.007	0.000	0.996	0.000	1.007	0.000	1.043	0.000	1.040	0.000	1.042	0.000
	area credit score	-0.010	0.000	-0.010	0.000	-0.010	0.000	-0.012	0.000	-0.012	0.000	-0.012	0.000
	average interest rate	-0.128	0.346	-0.143	0.297	-0.128	0.348	0.522	0.000	0.518	0.000	0.522	0.000
	area unemp rate	0.044	0.120	0.038	0.186	0.044	0.121	0.045	0.120	0.044	0.121	0.045	0.120
	y2003 (y2005)	-0.078	0.389	-0.097	0.289	-0.077	0.393	-0.607	0.000	-0.602	0.000	-0.608	0.000
	rate_sp			0.075	0.033					0.038	0.274		
	rate_resid					-0.002	0.961					0.020	0.573
	subprime	1.592	0.000	1.446	0.000	1.594	0.000	1.596	0.000	1.515	0.000	1.559	0.000
	cons	0.818	0.544	0.268	0.846	0.814	0.546	1.291	0.347	0.719	0.629	1.237	0.371
Prepay	put	-0.015	0.000	-0.014	0.000	-0.014	0.000	-0.007	0.061	-0.018	0.000	-0.018	0.000
	dti 28-36	0.289	0.000	0.291	0.000	0.297	0.000	-0.045	0.760	-0.062	0.686	-0.139	0.356
	dti 36-42	0.348	0.000	0.356	0.000	0.360	0.000	0.058	0.683	0.082	0.579	-0.049	0.739
	dti>42	0.015	0.825	-0.008	0.906	0.002	0.975	-0.300	0.030	-0.123	0.386	-0.299	0.031
	cscore<580	0.142	0.322	0.264	0.068	0.117	0.405	-0.090	0.663	-0.734	0.001	-0.022	0.916
	cscore 580-620	0.080	0.321	0.136	0.099	0.072	0.376	0.237	0.069	-0.184	0.187	0.139	0.298
	cscore 620-660	0.323	0.000	0.361	0.000	0.324	0.000	-0.193	0.131	-0.373	0.004	-0.211	0.102
	cscore 660-720	0.149	0.005	0.158	0.003	0.143	0.007	-0.076	0.521	-0.159	0.181	-0.080	0.497
	unpaid balance (in	0.329	0.000	0.309	0.000	0.335	0.000	-0.055	0.537	0.338	0.002	0.129	0.157
	loan age (in log mon)	0.459	0.000	0.466	0.000	0.464	0.000	0.697	0.000	0.679	0.000	0.688	0.000
	area credit score	0.001	0.381	0.000	0.626	0.001	0.350	0.007	0.001	0.008	0.000	0.006	0.012
	average interest rate	-0.197	0.007	-0.184	0.012	-0.186	0.011	0.200	0.203	0.163	0.304	0.186	0.241
	area unemp rate	-0.016	0.338	-0.014	0.408	-0.014	0.385	-0.029	0.409	-0.022	0.528	-0.022	0.527
	y2003 (y2005)	-0.021	0.640	-0.023	0.613	-0.019	0.675	0.278	0.003	0.273	0.003	0.239	0.010
	rate_sp			-0.068	0.000					0.399	0.000		
	rate_resid					-0.059	0.002					0.367	0.000
	subprime	0.922	0.000	1.004	0.000	0.977	0.000	1.238	0.000	0.505	0.000	0.601	0.000
	cons	-11.24	0.00	-10.74	0.00	-11.39	0.00	-11.91	0.00	-17.37	0.00	-12.46	0.00
	Log likelihood	-16790.3		-16780.3		-16785.1		-8262.9		-8211.7		-8219.2	

Note: Model 1 is the same as the one in Exhibit 9. *rate_sp* represents the difference between the mortgage note rate and the average interest rate of 30-year fixed-rate mortgages from the Freddie Mac Primary Mortgage Market Survey in the same month. *rate_resid* represents the residue of the risk premium variable from OLS models of risk premium.