WORKING PAPER NO. 10-8
"CREAM-SKIMMING" IN SUBPRIME MORTGAGE SECURITIZATIONS: WHICH SUBPRIME MORTGAGE LOANS WERE SOLD BY DEPOSITORY INSTITUTIONS PRIOR TO THE CRISIS OF 2007?

Paul Calem
Federal Reserve Board of Governors

Christopher Henderson
Federal Reserve Bank of Philadelphia

Jonathan Liles
Freddie Mac

March 2010

"Cream-Skimming" in Subprime Mortgage Securitizations:

Which Subprime Mortgage Loans Were Sold by Depository Institutions Prior to the Crisis of 2007?

Paul Calem

Division of Banking Supervision and Regulation, Federal Reserve Board

Christopher Henderson

Department of Supervision, Regulation and Credit, Federal Reserve Bank of Philadelphia

and Jonathan Liles

Mission Division, Freddie Mac

March 2010

The views expressed in this paper are those of the authors and do not necessarily reflect the position of the Federal Reserve Board, the Federal Reserve Bank of Philadelphia, or Freddie Mac. We thank Robert Avery, Michael Gordy, Paul Kupiec, Kate Samolyk, Robert Tsien, George Walker, and Peter Zorn for helpful comments. We also thank Vidya Shenoy and Ali Cannoni for excellent research assistance. This paper is available free of charge at www.philadelphiafed.org/research-and-data/publicationsK.

“Cream-Skimming” in Subprime Mortgage Securitization:

Which Subprime Mortgage Loans Were Sold by Depository Institutions Prior to the Crisis of 2007?

Abstract

Depository institutions may use information advantages along dimensions not observed or considered by outside parties to “cream-skim,” meaning to transfer risk to naïve, uninformed, or unconcerned investors through the sale or securitization process. This paper examines whether “cream-skimming” behavior was common practice in the subprime mortgage securitization market prior to its collapse in 2007. Using Home Mortgage Disclosure Act data merged with data on subprime loan delinquency by ZIP code, we examine the bank decision to sell (securitize) subprime mortgages originated in 2005 and 2006. We find that the likelihood of sale increases with risk along dimensions observable to banks but not likely observed or considered by investors. Thus, in the context of the subprime lending boom, the evidence supports the cream-skimming view.

“Cream-Skimming” in Subprime Mortgage Securitization:

Which Subprime Mortgage Loans Were Sold by Depository Institutions Prior to the Crisis of 2007?

1. Introduction

Financial innovation has made transferring risk from primary lending markets through securitization a common practice. Banks may securitize a portion of their loan portfolios to minimize regulatory capital requirements or to diversify funding opportunities and risk. Investors seek to purchase asset-backed securities because they are relatively liquid and, in principle, allow the investor ownership claim on a particular bank asset with quantifiable default and prepayment risk. The financial crisis of 2007, however, curtailed much securitization activity. We focus our attention on information problems associated with securitization as a factor leading to the crisis.

Do depository institutions sometimes enjoy information advantages for securitized assets along dimensions not observed or considered by outside parties? And do they use these advantages to transfer risk to uninformed or unconcerned investors through the sale or securitization process? Was such “cream-skimming” behavior common practice in the subprime mortgage securitization market prior to its collapse in 2007? These questions are of interest not only in relation to subprime mortgages originated and sold by depository institutions but also in regard to the role played by myopic or ill-informed investors in the subprime mortgage market generally.

This concept of “cream-skimming” is similar to but distinct from the traditional concept of a “lemons market” as applied to loan securitization. When investors cannot fully verify the quality of loans being securitized, the lender potentially can keep the good (low-risk) loans and sell the bad. In the traditional lemons market, as in Akerlof (1970), buyers are aware of their informational disadvantage. Various outcomes are possible, including breakdown of the market, pooling and pricing of risk, or the development of mechanisms to separate the good from the bad.[1] “Cream-skimming” as defined here, however, occurs when investors are unaware of or indifferent to the information asymmetry. Conversely, “risk detection” among investors, ratings agencies, and market participants through higher

[1] In the literature, these typically arise as pooling or separating equilibriums in models of contracting between parties with asymmetric information.

asset yields, lower ratings, or greater credit enhancements implies the absence of cream-skimming. Moreover, "cream-skimming" does not imply malfeasance on the part of depository institutions but might result from investors and ratings agencies being overconfident in their ability to quantify default risk or the likelihood of events that cause defaults to be correlated.

Figures 1 and 2 provide evidence consistent with the view that mortgage securitization markets during the house price boom of 2005 through 2006 were characterized by possibly overly optimistic views of mortgage credit risk and, thus, were potentially conducive to cream-skimming. Figure 1 indicates that investors in mortgage-backed securities did not demand higher yields for holding these securities relative to the risk-free Treasury yield. Additionally, Figure 2 shows that the level of credit enhancements for senior tranches of prime, nonagency deals remained relatively consistent, demonstrating that ratings agencies and investors were either unaware of the increasing credit risk of the underlying mortgages or the ratings agencies made errors in assessing the risk, possibly opening the door for originators to "cream-skim."[2]

A competing view is that securitization would be favored for relatively transparent, lower-risk assets, whereas banks would tend to retain opaque, higher-risk loans because of the higher transactions costs of securitization or market discipline (for example, Hill 1996). In the case of mortgages, such a tendency may be reinforced by incentives arising from regulatory capital requirements (Calem and LaCour-Little 2004; Calem and Follain 2007). Consistent with this view, for instance, Ambrose, LaCour-Little, and Sanders (2005) find evidence in data from a single large lender that risky mortgages are retained on balance sheet, while lower-risk mortgages are sold into secondary markets.

The recent experience in the subprime mortgage market suggests, however, that investors may have become less responsive to risk, increasing the potential for cream-skimming. Beginning in 2004, growing demand for subprime RMBS among investors was satisfied increasingly through the origination and sale of nontraditional loans, including hybrid adjustable rate mortgages (hybrid ARMs) with artificially low initial interest rates. Moreover, many of the subprime loans originated and sold in 2005 and 2006 involved high loan-to-value ratios or were associated with a "piggyback" second lien, and high post-reset debt-payment-to-income ratios, often exceeding 55 percent (Cagan 2007).

[2] We thank Rita Csejtey and Lili Corson at the Federal Reserve Bank of New York for graciously sharing securities data using Intex.

This paper provides evidence that banks leveraged asymmetric information (or an excessive optimism or breakdown of due diligence among investors) regarding the credit risk of subprime home purchase mortgages, to shift risk onto investors in the RMBS market. Using Home Mortgage Disclosure Act (HMDA) data merged with data on subprime loan delinquency by ZIP code, we examine the bank decision to sell (securitize) subprime mortgages originated in 2005 and 2006. We find that the likelihood of sale increases with risk along dimensions observable to banks but not likely observed or considered by investors. For instance, the likelihood of sale increases with the ratio of loan amount to borrower income (a proxy for the post-reset debt-payment-to-income ratio) and is positively related to the ex-post (January 2008) rate of serious delinquency among subprime loans in the ZIP code where the property is located. While these findings are consistent with either cream-skimming- or lemons market-type explanations, we believe (for reasons to be discussed) that overall the empirical results favor the cream-skimming view.

At the outset, we emphasize that the focus here is on the banking side of depository institutions that originated subprime loans and on overall industry patterns, rather than the behavior of any individual institution. In particular, we do not address the intriguing question of why some institutions may have assumed risks on the trading side (through subprime RMBS purchases and related activities) that they had previously avoided by securitizing subprime loans originated on the banking side.

The paper is organized as follows. Section 2, which immediately follows, briefly reviews a sampling of related literature on bank loan securitization and the subprime mortgage market. Section 3 presents a stylized, theoretical framework that will help motivate the empirical analysis. Section 4 presents an overview of the data used for the analysis and a variety of descriptive statistics constructed from these data that relate to the structure of the subprime mortgage market and characteristics of subprime lending during 2005 and 2006. Section 5 develops our empirical analysis of the loan sale decision of depository institutions, involving the estimation of logistic regression equations by year and institution size. Section 6 presents the estimation results and Section 7 concludes.

2. **Literature Review**

Our paper contributes to the traditional literature on bank loan sale and securitization in relation to credit risk. It also adds to the burgeoning new literature on the growth of the subprime mortgage market and the ensuing foreclosure crisis and collapse of the market.

Banks have an incentive to sell or securitize loans to diversify funding opportunities and generate liquidity, diversify or manage credit and interest rate risk, and reduce regulatory capital requirements. In examining the relationship between securitization and credit risk, the literature focuses on the role of asymmetric information and on incentives deriving from regulation.

In the context of the traditional lemons market, it is clear how the presence of information asymmetry could encourage securitization of riskier loans. If lenders have more information about borrower credit quality than the purchasers of the securitized assets, then the lenders would take advantage of this information asymmetry and sell all risky loans. This is the classic lemons argument as presented by Akerlof (1970).

It is equally clear, however, that alternative outcomes are possible depending on the structure of the market and strategies available to buyers and sellers. Contracts may evolve that give lenders incentives to either pool or separate risks. In the latter case, lower-risk loans may be selected for securitization and the higher-risk loans retained, or the opposite outcome could prevail.

For instance, Hill (1996) develops a model in which the securitized assets are lower risk and their credit quality more transparent. Securitization reduces financing costs for firms that face limited financing opportunities due to limited information concerning the credit quality of their assets. According to Hill, securitization “divides the firm into slices which permit more specialized appraisal”; with selection of lower-risk assets for securitization, “security investors needn't appraise the particularly costly-to-appraise residual risks and prospects of such firms.”

Dewatripont and Tirole (1994, Chapter 10) also suggest that securitized loans are likely to be of higher credit quality. They argue that banks are uniquely equipped to add value to loans made to higher-risk borrowers by managing the future relationship with the borrower and that securitizing such loans may weaken this ability or the incentive to apply it.

In contrast, DeMarzo and Duffie (1999) propose a model in which asymmetric information drives securitization of riskier loans. In particular, a lender receives a higher price for a lower-risk security (where the credit quality differential is observable only to the lender) only by keeping a fraction of the security issue to signal the value of the lender's private information.[3] This signaling through repurchase of the security entails a cost and, therefore, provides a disincentive to securitize higher-quality loans. Moreover, the signaling enables the market to distinguish higher- from lower-risk security issues and to price accordingly.

Although information asymmetries are one factor in how securitization transactions are structured, other factors such as the level of regulatory capital and reputational risk help explain a bank's incentive to securitize a pool of receivables. For instance, Calem and Follain (2007) maintain that for lower-risk mortgage loans, existing regulatory capital levels (established under the original Basel Accord) are too high, creating an incentive to securitize these loans. As noted in the introduction, Ambrose, LaCour-Little, and Sanders (2005) examine data from a single large lender and find that risky mortgages are more likely to be retained on balance sheet. They point to incentives arising from capital regulation and reputational risk as likely explanations for their findings.

The collapse of the subprime market and beginning of the foreclosure crisis in 2007 and subsequent turmoil in mortgage and housing markets have spurred a variety of research on problems in the subprime or broader mortgage market that precipitated the crisis. Much of this research has focused on the deterioration of underwriting standards and house-price depreciation as primary factors (Smith 2007; Demyanyk and van Hemert (2009); Gerardi, Shapiro, and Willen 2008; Hahn and Passell 2008; Sherlund 2008). Haughwout, Peach, and Tracy (2008) focus on early payment default and emphasize that only part of the increase in default during 2007 is attributable to these factors. Demyanyk and van Hemert (2009) argue that the decline in underwriting standards prior to the crisis could have been detected but was masked by rapid house-price appreciation. Coleman, LaCour-Little, and Vandell (2008) present evidence that the expansion of credit resulting from looser underwriting standards contributed to the rise in house prices.

The issue broadly related to the subject of our paper—the role of securitization and associated moral hazard problems—has also garnered attention, with several researchers pointing to securitization as a principal culprit in the crisis. Ashcraft and Schuermann (2008) identify a number of market frictions affecting the subprime mortgage origination and securitization process and argue that the associated

[3] DeMarzo (2005) goes further by modeling how informed market participants create derivative securities by tranching pools of loans using specific risk characteristics.

misaligned incentives and adverse selection were largely responsible for the market's collapse. A partial list includes agency problems associated with brokers, such as incentives to misrepresent borrower credit quality; cream-skimming by portfolio lenders (the focus of our study); and rating agency conflicts of interest.[4] Golding, Green, and McManus (2008) and Hull (2009) also focus on misaligned incentives of market participants, in particular, compensation of loan originators and security traders disassociated from subsequent credit performance of the loans, and ratings agencies being paid by the issuers of the securities being rated. They put forth recommendations aimed at increasing transparency and reducing moral hazard in both the primary and secondary mortgage markets.

Rajan, Seru, and Vig (2009) draw a distinction between the "hard information" relied on by investors to value securitized loans and "soft information" accessible to originators but not verifiable by a third party. They argue that securitization of subprime mortgages reduced the incentive to collect soft information, resulting in less effective credit screening. In contrast, we focus on the cream-skimming that may result from such an information asymmetry. Note that for an individual firm, these are mutually exclusive outcomes, but that is not the case for the market as a whole; some institutions may engage in cream-skimming, while others may specialize in collecting only hard information and securitize all of the loans they originate.

Ben-David (2007) focuses on the propensity to overstate collateral values by borrowers, intermediaries, and originators when it is advantageous to do so in the presence of asymmetric information. In particular, originators are able to expand their business by securitizing more loans as house prices rise. White (2009) emphasizes the role of overly optimistic evaluations of the credit risk of mortgage-backed securities, in part due to agency problems and in part to inadequate information and "carelessness." Coval, Jurek, and Stafford (2009) point to the amplification of errors in evaluating the risk of the underlying securities of structured finance products and in the mispricing of these products.

Mian and Sufi (2009) present evidence that mortgage securitization expanded the supply of credit to riskier segments of borrowers. They show that the massive expansion of credit was disassociated from income growth in areas where it occurred and find an association with subsequent deterioration in credit performance. Their finding of an association between increased securitization and higher default rates at the neighborhood level is similar to that observed in our study, although the empirical

[4] Ernst, Bocian, and Li (2008) argue that mortgage brokers also often exploit an information advantage relative to the borrower to engage in predatory lending.

approaches differ.[5] Their analysis suggests a direct link between securitization and the recent mortgage crisis.

Elul (2009) compares the repayment performance of securitized residential mortgages to those retained in portfolio, for loans originated during 2003 through 2007. The analysis controls for various loan-level factors, including the initial loan-to-value ratio and the borrower's credit score at origination, and for local economic conditions, including house-price appreciation. The analysis indicates that nonagency prime, adjustable rate securitized loans originated after 2003 (and fixed rate loans originated after 2005) performed significantly worse than similar, nonsecuritized, loans. According to Elul, the findings "suggest that adverse selection may have been present in the prime nonagency mortgage market and may have contributed to deterioration in underwriting standards."[6]

3. A Framework for Cream Skimming – Data Simulation

In this section we construct a simple data simulation that represents securitization by banks and highlights two contrasting (but not mutually exclusive) hypotheses: the "cream-skimming" view that banks sometimes can exploit informational advantages along opaque dimensions of credit risk by securitizing riskier loans, and the alternative view that securitization is favored for loans of higher credit quality. This stylized framework helps put into focus the issues motivating the empirical analysis to follow.

Essentially, we represent the choice of share of loans securitized as a decision under uncertainty, where excess securitization (relative to the ex-post realized target share) is more costly than retaining too large a share. This cost asymmetry reflects the simple intuition that it is easier to sell surplus balance-sheet assets at a later date than it is to unwind a securitization or replace assets on balance sheet. Initially, we assume that the bank does not have any private information (either ex-ante or ex-post) concerning the value of securitization. We show that under plausible conditions, the proportion securitized will decrease as the credit quality of the loans increases. While new, this approach is broadly

---

[5] Mian and Sufi (2009) conduct a panel analysis of mortgage lending patterns at the ZIP code level; our analysis is cross-sectional and at the loan level and focuses specifically on the subprime loan sale decision along dimensions likely associated with bank information advantages.

[6] Elul (2009) uses a loan-level data set from LPS Analytics. The study also examines the subprime market and does not find a difference in performance between securitized and retained loans; however, the data set's coverage of the subprime market is less representative than for prime and contains relatively few retained subprime loans.

similar to other arguments that have emphasized higher costs or reduced benefits from securitization of riskier or more opaque assets.

To motivate this framework, we posit that the benefits and costs of securitization depend on macroeconomic, sector-specific, and institution-specific factors that are not fully observed ex-ante. For example, changing macroeconomic conditions might affect the bank's need for liquidity or the value it adds to retained loans through its risk-mitigation activities. For a given product and risk category of loans, let X denote the ex-post target share, that is, the share that would have been securitized conditional on a realization of these random variables. Ex-ante, X is a random variable that we assume has a beta distribution, denoted F(X, *a, b*), with mean μ and standard deviation σ given by:

$$(1)\quad \mu = a/(a+b);\quad \sigma^2 = ab/(a+b)^2(a+b+1) = \mu(1-\mu)/(a+b+1);\ \ a > 0;\ \ b > 0$$

The beta distribution has advantages of being bounded between 0 and 1 and encompassing a variety of possibilities, such as uniform, unimodal and bimodal symmetric distributions, and unimodal, asymmetric long-tailed distributions.

Let S denote the rate of securitization chosen by the bank ex-ante. We assume that the cost of over- or under-securitizing conditional on X, denoted c(X), is proportional to X – S, specifically:

$$(2)\quad c(x) = X - S \text{ if } S > X;\ \ c(x) = \alpha(X - S) \text{ if } X \geq S, \text{ where } \alpha < 1$$

We first establish a general result concerning the relationship between ex-ante uncertainty and percent securitized. For this purpose, it is useful to characterize a beta distribution in terms of its median and interquartile range. Thus, between two beta distributions with the same median, the one with a larger interquartile range incorporates more uncertainty. Also, between two distributions both having infinite density at zero and zero density at one (or vice versa), the one with the larger interquartile range (or, equivalently, the one closer to uniform) represents greater uncertainty. We rely on the following results, derived in the Appendix:

Proposition 1: Let F(x, a_1, b_1) and F(x, a_2, b_2) be two beta distributions with identical medians x^m. If F(x, a_1, b_1) has a wider interquartile range, then F(x, a_1, b_1) > F(x, a_2, b_2) for all x < x^m.

Proposition 2: Let F(x, a_1, b_1) and F(x, a_2, b_2) be two beta distributions such that, either $a_i < 1$ and $b_i \geq 1$ for i=1,2, or $a_i \geq 1$ and $b_i < 1$ for i=1,2. If F(x, a_1, b_1) has a wider interquartile range, then F(x, a_1, b_1) > F(x, a_2, b_2) for all 0 < x < 1.

The bank chooses a rate of securitization S to minimize expected cost:

$$(3)\quad \min_{S} \int c(x) f(x) dx$$

It is straightforward to verify that the solution S* satisfies:

$$(4)\quad F(S^*, a, b) = \alpha/(1+\alpha)$$

Let F^{-1} denote the inverse of the beta distribution F(X, *a, b)*; then:

$$(5)\quad S^* = F^{-1}(\alpha/(1+\alpha), a, b)$$

Since $\alpha < 1$, we have $S^* < x^m$, where x^m denotes the median. Therefore, it follows from Proposition 1 that across beta distributions having the same median, S* declines as the interquartile range widens. Likewise, across beta distributions as characterized in Proposition 2, S* declines as the interquartile range widens. Thus, in each case the rate of securitization declines as the degree of ex-ante uncertainty, represented by the interquartile range of the distribution, increases.[7] If we further assume that lower-quality (riskier or more opaque) assets are characterized by greater uncertainty regarding the ex-post target rate of securitization, then the ex-ante rate of securitization will decline with credit quality. Such an assumption seems reasonable; because the repayment performance of these loans is more uncertain, the costs and benefits associated with securitizing them are apt to be more uncertain as well. Note that S* will also depend on the median of the distribution x^m, which could reflect, for instance, the transactions costs of securitization and the marginal contribution of a loan sale to diversifying portfolio risk.

These ideas are illustrated in Figure 3 for the case of symmetric beta distributions for which $a=b$ and $x^m=\mu=½$. We assume that $a=b=q+0.05$, where q is an index of credit quality that ranges from 0 (C-rated) to 1 (A-rated), with larger q corresponding to higher credit quality. The solution S* in relation to the quality index q is indicated by the blue line in Figure 3 for the case $\alpha = 0.75$. S* is upward sloping in relation to q, reflecting the reduced variance (and interquartile range) associated with larger q.[8]

Further, it seems plausible that the marginal cost of excess securitization may be larger for lower-quality assets. For instance, it is reasonable to expect that the value added to retained loans through

---

[7] Letting δ denote the interquartile range, we have $dS^*/d\delta = -(\partial F/\partial\delta)/(\partial F/\partial x) < 0$.

[8] From (1) μ=½, it follows that $\sigma^2=1/(4+8a) = 1/4(1.01+2q)$.

risk-mitigation efforts is greater for loans that are riskier or more opaque.[9] We can represent such a relationship by replacing the parameter α, which quantifies the cost differential between over- and under-securitization, with a relationship $\alpha(q)$, where $d\alpha/dq \geq 0$. Not surprisingly, as shown by the magenta line in Figure 3 for the case $\alpha(q) = \min(q+0.20, 0.75)$ and $a=b=q+0.05$, this assumption results in a more rapid decline in the securitized share S* as credit quality decreases.

Informational advantages of banks. Thus far, we have assumed that the bank does not have private information concerning the return to securitization. We now consider that possibility. We consider only the possibility of cream-skimming as defined previously, as opposed to the traditional lemons market situation. That is, the dimensions along which the bank has private information are either not observed or not considered by the securities market.

Ties to local markets, for instance, may generate information advantages for depository institutions, including knowledge of local or neighborhood conditions influencing mortgage or small business credit risk.[10] Prior or ongoing lending relationships may be a source of additional private information. Another possibility is that banks benefit from superior controls over information processes because of better aligned incentives and regulatory supervision. This possibility may have been applicable particularly to the subprime mortgage market, where RMBS investors depended on rating agencies or other third parties that may have faced conflicting incentives (Ashcraft and Schuermann 2008).

Such information advantages create opportunities for banks to "cherry pick" among loans indistinguishable to the market, retaining those with lower expected loss conditional on information available only to the bank. Of course, in a context of repeat relationships, investors could punish banks that on previous occasions sold loans that underperformed, thereby deterring cream-skimming activity. Suppose, to the contrary, that investors are unresponsive to cream-skimming. This might be the case, for instance, in an environment where investors are unaware of certain risks, mistakenly relax their due diligence activity, or are overly optimistic. Now let q index a credit quality dimension observed only (or with greater accuracy) by the bank. In this case, the bank's marginal benefit from securitization will increase as credit quality decreases, because of the ability to cream-skim lower-quality loans. We again can illustrate this effect by replacing the parameter α with a relationship $\alpha(q)$, where now $d\alpha/dq < 0$ for

[9] This value added is mostly realized under circumstances (realizations of X) that generate higher default rates.
[10] Community banks by definition have ties to local markets. Larger banking organizations may also develop awareness of local and neighborhood conditions by monitoring the profitability of branch networks or broker and correspondent relationships, or through community reinvestment activities.

a range of q. The yellow line in Figure 3 represents the case $\alpha(q) = 1\text{-}q\text{+}3q^2/4$, again with $a=b=q+0.05$. Not surprisingly, the possibility of cream-skimming leads to reversal of the relationship between (asymmetrically observed) credit quality q and the rate of securitization; for $q < 0.44$, securitized share increases as credit quality declines. We question whether this latter scenario had some role in the crisis of 2007.

4. Data

Our empirical analysis focuses on the loan sale (securitization) decision of depository institutions regarding subprime home purchase mortgages originated in 2005 and 2006. The analysis relies on loan-level HMDA data, which indicate whether a loan was sold during the year when it was originated, and the type of institution purchasing the loan.[11]

In addition to indicating whether a loan is sold or retained, the HMDA data provide a variety of loan, borrower, and property characteristics that we use to develop our empirical analysis. These include the loan amount; loan purpose (home purchase or refinance) and type (conventional or government insured); identity of the institution that originated the loan; income of the borrower, state, county, and census tract location of the property being financed; and ownership status of the property (owner-occupied primary residence or not). For high-cost loans, defined in the HMDA regulation as first-lien loans with APR spread greater than or equal to 300 basis points above the applicable Treasury yield (500 basis points for junior liens), the APR spread is also provided. We limit the population for our analysis to high-cost loans plus loans originated by HUD-identified subprime specialists (whether high-cost or not), defining this to be the subprime population in the HDMA data.[12] We further restrict the sample to conventional, first-lien, single-family home purchase loans.[13]

---

[11] Purchaser categories include Fannie Mae; Freddie Mac; Ginnie Mae, Farmer Mac; commercial or savings bank or savings and loan association; affiliated institution; mortgage, finance, or life insurance company or credit union; other type of purchaser. Note that the disposition of a loan originated late in the year generally is not known, since a sale of the loan would typically not occur until the following year.

[12] We also exclude loans larger than $1,000,000.

[13] The restriction to home purchase loans is motivated in part for the sake of brevity and in part by the lack of distinction in HMDA data between two important categories of refinance loans: cash-out refinancing used to extract accumulated home equity, and rate-refinancing used to obtain a lower interest rate or to reduce (at least temporarily) the monthly mortgage payment. Moreover, a prior relationship with the borrower is more likely to be a factor in refinance lending. Inability to distinguish such factors complicates interpretation of empirical results based on refinance loans. Results obtained with refinance loans are similar to those reported below for home purchase loans.

We supplement the HMDA data with institution-specific information: primary regulator; total assets; and whether the institution is a subprime lending specialist as defined by HUD. For depository institutions, we also add an indicator for thrift institutions (saving and loan institutions and savings banks) and an indicator for whether the institution has a branch in the county where the loan was originated.[14]

We also combine the loan-level data with a number of local and neighborhood-level housing and mortgage market variables. Annual house-price appreciation rates by metropolitan statistical area (MSA) and state are calculated using the Federal Housing Finance Agency's ((FHFA), formerly the Office of Federal Housing Enterprise Oversight (OFHEO)) weighted repeat-sales price index. Percent change in housing starts from the previous year, by MSA and state, are obtained from Economy.com. Neighborhood-level variables include number of owner-occupied units in the census tract where the property is located, from the 2000 U.S. census, and information from HMDA aggregated to the tract level. The latter include, by year, the proportion of mortgages that are high-cost loans; the fraction of subprime mortgages that are originated by HUD-identified subprime lenders; the proportion of home purchase loans that are for non-owner-occupied properties; and the fraction of high-cost home purchase loans that are second lien.[15]

HMDA does not provide detailed information on loan terms, loan-to-value ratio (LTV), or borrower credit information such as FICO score. We accessed such information, along with data on subsequent payment performance, at the aggregate ZIP code level using the LoanPerformance TrueStandings Servicing™ online data analytics tool. This platform provides aggregated ZIP-code-level information on the current payment status of active mortgages and original loan terms and LTVs and FICO score ranges for both paid-off and active mortgages serviced by the top mortgage servicing institutions.[16] In contrast to the more commonly used loan-level subprime securities database of LoanPerformance, the subprime servicing database includes subprime loans retained in bank portfolios as well as those in securities.[17]

---

[14] We thank Robert Avery of the Federal Reserve Board of Governors for supplying the institution-level data.

[15] The aggregated tract-level measures from HMDA are defined with respect to conventional, single-family, home purchase, and refinance loans originated by the HMDA-reporting institution. Other than the fraction of second liens, they are calculated with respect to first liens only.

[16] This online business intelligence platform accesses the subprime mortgage database of LoanPerformance, a division of FirstAmerican CoreLogic. Information about TrueStandings Servicing® is available at www.loanperformance.com.

[17] Loans assigned to the subprime database are serviced by institutions that specialize in servicing subprime loans or are identified as subprime by the servicing institution. Despite the recent demise of most subprime specializing institutions, the subprime database continues to track the performance of active subprime loans,

Specifically, we obtain estimates of first-lien subprime mortgage delinquency and foreclosure rates as of October 2008 by ZIP code and of the proportion of subprime mortgages originated during 2005 and 2006 in each ZIP code by range of origination LTV ratio, FICO score range, and interest rate type. We merge these ZIP-code-level data into our loan-level HMDA data. Since HMDA data indicate the state, county, and census tract associated with a mortgage, not the ZIP code, we first map each state, county, and census tract into one or more ZIP codes.[18]

Overview of the subprime market in 2005 and 2006. Figures 3 through 12 present descriptive information on the composition of the subprime home purchase loan market in 2005 and 2006, providing background for our analysis. Figures 3 through 8 describe subprime lending activity by type of institution, based on the subprime HMDA data. Figure 9 provides summary statistics from the HMDA data merged with the data on MSA house-price appreciation and housing starts, with 2004 data appended for comparison. Figures 10 and 11 provide summary statistics on product mix and loan characteristics from the HMDA data merged with the data from LoanPerformance TrueStandings Servicing™, again with 2004 appended for comparison. Figure 12 presents information from HMDA on proportion of subprime, first-lien home purchase loans originated with "piggyback" second liens in 2004 through 2006.

The distribution of high-cost home purchase loans by type of institution is presented in Figure 3. In 2005, a slim majority of the loans were originated by subprime specialists; nonspecialist nondepository institutions had the next highest share, nearly 30 percent, followed by large (more than $10 billion in total assets) and small depository institutions, respectively. In 2006, the share of subprime specialists declined to about 36 percent, while the share of nonspecialist nondepository institutions increased to match that of the specialists. The shares of large and small depositories increased as well, from around 20 to about 30 percent combined share.

Figure 4 reports the disposition of subprime home purchase loans (whether sold or retained, and type of purchaser) in 2005 and 2006, by type of institution originating the loan. Similar patterns are observed in each year. Nearly 80 percent of the originations of subprime specialist institutions are reported sold to nonaffiliated nondepository institutions. For convenience, we shall refer to this loan sale category as securitization, since it is likely that these loans were packaged in securities and sold to

because the servicing of these loans has largely been transferred to other institutions that contribute to the database.

18 Where a census tract traverses more than one ZIP code, we allocate the tract across the ZIP codes in proportion to loan counts observed in Freddie Mac internal data.

investors. Small quantities are sold by subprime specialists to affiliated institutions or nonaffiliated depository institutions, while a little less than 20 percent are recorded as not sold. It is likely that most of the latter are in the securitization "pipeline" or are "warehouse" loans being held for sale, including loans originated near the end of the HMDA-reporting year. At nonspecialist nondepository institutions (including independent mortgage companies as well as mortgage subsidiaries of depository institutions or of bank or thrift holding companies), the share sold directly to the securities market was somewhat smaller, closer to 70 percent, reflecting a somewhat larger share sold to affiliates and depositories, while the share recorded as not sold was again not quite 20 percent.   About half of the loans of large depository institutions in 2005, rising to 55 percent in 2006, and less than 20 percent of the originations of small depository institutions, are sold directly to the securities market. Close to 20 percent of the originations of large depository institutions are sold to affiliates, while 80 percent of the originations of small depository institutions are recorded as not sold.

Figure 5 indicates the average ratio of loan amount to income by type of institution in 2005 and 2006. This ratio is highest at subprime specialists, close to 2.5 in each year, suggesting relatively high risk exposure with respect to the repayment capacity of borrowers. It is marginally smaller at large depository institutions and nonspecialist nondepository institutions, in the range of 2.2 to 2.4, and much lower, about 1.5, at small depository institutions.[19]   Average loan sizes show a similar pattern by institution category, as indicated in Figure 6.

Average APR spread in 2005 and 2006 by type of institution is shown in Figure 7. Subprime specialists and nonspecialist nondepository institutions have larger average APR spreads than depository institutions. In 2006, spreads for subprime specialists rose dramatically, while spreads for small depository institutions declined, relative to the other institution categories.

Figure 8 indicates percent of loans associated with a nonprimary residence (investment property or second home), by type of institution in 2005 and 2006. Percent of loans in this occupancy category is greatest for small depositories, about 25 percent, and smallest for subprime specialists, only 10 percent, in both years. The share of loans associated with nonprimary residence increased substantially between 2005 and 2006 for nonspecialist nondepository institutions and for large depositories.

Housing market conditions in 2004 through 2006 are summarized in Figure 9. Sample averages of annual percent change in the local area FHFA house price index and annual percent change in local area

[19] Note that the decline in the average loan-amount-to-income ratio between 2005 and 2006 does not necessarily imply declines in payment-to-income ratios, since subprime product mix changed, and shorter-term interest rates rose.

housing starts are shown, along with annual percent change in the FHFA national house-price index.[20] House-price appreciation and growth in housing starts declined in 2006 from the elevated levels observed in 2004 and 2005. The sample average rate of house-price appreciation exceeds the percent change in the FHFA national index each year, indicating that subprime home purchase loans were disproportionately originated in markets with high rates of house-price appreciation.

Figure 10 indicates that about half of the dollar volume of subprime home purchase mortgages originated in 2005 was in hybrid ARMs (mostly 2-28 and 3-27); about 30 percent was standard ARMs; about 10 percent were fixed-rate (FRMs); and the remainder (less than 10 percent) was a mix of nontraditional products (balloon and others, including interest-only mortgages). From 2005 to 2006, substantial shifts occurred from hybrid ARMs to the nontraditional products and from FRMs to standard ARMs, a continuation of a trend observed between 2004 and 2005.

As shown in Figure 11, around a third of subprime home purchase mortgages in 2005 and 2006 were categorized as low documentation. A slim majority had FICO scores greater than 620, indicating that borrower credit rating was only one of a number of factors distinguishing subprime from prime loans. Fewer than 20 percent have an LTV of 90 percent or greater. These proportions were similar in 2004.

Note, however, that these data indicate only the LTV associated with the first lien. Figure 12 provides estimates of the proportion of first-lien, subprime home purchase loans associated with "piggyback" second liens in 2004, 2005, and 2006.[21] About half of first-lien subprime loans originated in 2005 and 2006 had a "piggyback" second, a considerable increase compared with 2004 and even more so compared with previous years.[22]

5. Empirical Approach

The primary objective of our empirical analysis is to explore the subprime loan sale decisions of depository institutions in 2005 and 2006 for evidence of cream-skimming behavior. A finding that

---

[20] Sample averages are obtained by weighting the local values by the subprime loan counts in the sample.

[21] Our estimate of proportion of first liens with a piggyback second is based on HMDA data with matched first and second liens (generously provided by Robert Avery of the Federal Reserve Board) and is calculated as follows. Let N = total number of matched seconds (subprime or prime) in these data; M = total number of subprime firsts; n1= number matched to a subprime first; n2 = number matched to a prime first; and Y= total no. of unmatched seconds (subprime or prime). We impute y1=Y*(n1/N) to be the number of unmatched seconds that are piggybacks to a subprime first. Then the total matched to subprime first-lien (actual + imputed) = n1 + y1, and the proportion of firsts with a piggyback second is (n1 + y1)/M.

[22] Lien status is not provided in HMDA data prior to 2004. However, monthly 2004 HMDA data indicate rapid growth in second-lien loan originations during the year.

likelihood of sale is inversely related to credit quality along dimensions associated with bank informational advantages would be consistent with cream skimming.

We estimate logit regression models relating the disposition of subprime loans originated by depository institutions (sold or not) to factors likely to be associated with information advantages; other indicators of credit quality; and other control variables. We estimate separate equations for large and small institutions (greater or less than $10 billion in assets) and by year (2005 and 2006). We include institution-specific fixed effects in the equation for large institutions and a vector of institution-specific characteristics in the equation for small institutions.

Most of the original HMDA sample (75 percent of the 2005 sample and 70 percent for 2006) is excluded by the restriction to loans originated by depository institutions, as shown in Table 1, lines 1 and 2a. We further restrict the sample by excluding loans sold to affiliates and or to nonaffiliate depository institutions, since the meaning of the sale is ambiguous and the ultimate disposition of the loan is not known (they may or may not be resold in securities.) In addition, we exclude the depository subsidiary of AIG and large institutions that originate fewer than 100 loans.[23] The effect of these exclusions on sample size is shown in Table 1, lines 2b and 2c.

As shown in line 3, about 15 percent of the remaining HMDA sample for each year cannot be merged with the LoanPerformance data, because information on loans originated in the ZIP code is lacking in the LoanPerformance database. The post-merge sample sizes are shown in line 4.

Most of the loans excluded on the basis of sale to affiliate or to nonaffiliate depository institutions were sales to affiliates by large depository institutions. Although in the aggregate, about 20 percent of the subprime loans originated by large depository institutions in both 2005 and 2006 were sold to affiliates, most of these sales were by institutions that sell all or nearly all of their loans. In estimating the fixed effects model, where institution-specific effects are accounted for, such institutions are necessarily excluded from the sample. Thus, we need not be concerned about censoring bias arising from these exclusions.[24]

---

[23] These institutions are viewed as nonrepresentative of large depository institutions that are active in the mortgage market. The AIG subsidiary is excluded because its parent is an insurance company. The large institutions that originate fewer than 100 loans are not active mortgage lenders and their loans represent a negligible share of the large institution sample. Moreover, it is convenient to exclude them for the purpose of estimating a fixed effects model. Note that depository subsidiaries of investment banks also end up excluded from the final sample, on the basis of criteria introduced below, because they sell more than 90 percent of the loans they originate.

[24] With exclusion of the institutions that sell all or nearly all of their loans, the share of the large bank sample excluded on the basis of sale-to-affiliate is 7 percent in 2005 and 8 percent in 2006.

Specifically, because we allow for systematic differences across institutions through inclusion of fixed effects, we exclude from the large bank sample institutions that sell more than 90 percent or fewer than 5 percent of the loans they originate. Row 5 of Table 1 reports the impact of these exclusions on sample size. [25]

Bank informational advantages. Banks are likely to have an information advantage along dimensions associated with neighborhood-specific credit factors, reflecting their ties to local markets. We employ two alternative approaches to test for cream-skimming related to neighborhood-specific information. The first approach, termed Model 1, includes several neighborhood risk factors in the vector $\mathbf{X}_i$ in equation (7):

- Percent of home purchase loans originated in census tract that are subprime
- Percent of census tract's subprime loans that are originated by subprime specialists
- Housing market depth (log of the number of owner-occupied units in the census tract)

The second approach, Model 2, includes a single neighborhood variable:

- The future (January 2008) subprime mortgage default rate in the census tract[26]

Neighborhood concentrations of subprime loans can generate elevated neighborhood default rates with adverse spillover effects on collateral values.[27] Concentrations of loans originated by subprime specialists can have similar effects.[28] Neighborhoods with fewer home sales will tend to have fewer informed appraisals and, therefore, a higher default risk associated with a particular measured LTV

[25] We apply a weaker criterion at the upper end to account for the fact that sales of loans originated near the end of the year are underreported. Three institutions in 2005 and four in 2006 were excluded because they had fewer than 5 percent sold loans, and ten institutions in 2005 and eight in 2006 were excluded because they had more than 90 percent sold loans. In addition, for consistency, for both years we exclude Fremont Mortgage, which straddled the 90 percent threshold, meeting it in 2005 and falling just short of it in 2006.

The final large bank samples are composed of 24 and 31 institutions, respectively, for 2005 and 2006. A list of these institutions is available from the authors on request.

[26] The default rate is measured as number of loans 90-days or more delinquent or in foreclosure, divided by number of active loans.

[27] See Lee (2008) for a literature review of price-related spillover effects.

[28] About 10 percent of the loans originated by subprime specialists each year in 2005 and 2006 were not high cost as defined in HMDA data but would have been higher risk than loans originated by prime lenders. Subprime specialists may also have tended to serve higher-risk segments among borrowers obtaining high-cost loans, as suggested by the data in Figures 5 through 7.

ratio.[29] If depository institutions are better informed regarding these risk factors and use their advantage to cream-skim, then the likelihood of loan sale would increase in relation to the first two measures and decrease in relation to the third.[30] Similarly, if they are cream-skimming, then the likelihood of sale would be positively related to the ex-post neighborhood default rate.

In both Models 1 and 2, we include two additional variables that we associate with likely information advantages of depository institutions:

- An indicator for whether the institution has a branch in the county where the property is located
- The ratio of loan amount to borrower income

Loans originated "out-of-market" (where the institution has no branch presence) more likely were obtained through mortgage brokers compared to loans in areas covered by the bank's branch network. Broker originations have been associated with agency problems and elevated credit risk (Jiang, Nelson, and Vytlacil 2009). If origination channel is not consistently reported to investors, then the possibility of cream-skimming based on origination channel arises.[31] Such cream-skimming would be reflected in the likelihood of sale being positively related to branch presence.

The ratio of loan amount to income is a proxy for debt-repayment capacity, where a larger loan amount relative to income indicates a greater payment burden and increased credit risk. Underwriting in the subprime RMBS market generally relied on standard payment-to-income ratios based on initial loan terms and did not account for potential post-reset payment shocks associated with hybrid ARMs and nontraditional loan products. Depository institutions may have gained an information advantage by applying more robust measures of repayment capacity of subprime borrowers than those used by ratings agencies or investors. Note that the advantage here relates to how information is used rather

---

[29] It is commonly argued that housing market depth is associated with the accuracy of appraised values and the resulting LTV ratios (Nakamura 2010). Because of potential endogeneity of number of home sales, we employ the log of the number of owner-occupied units as an instrument.

[30] Alternatively, subprime lenders may have relied disproportionately on brokers, in which case subprime share or share originated by subprime specialists may be correlated with share originated by brokers. As discussed below, broker channel may be a source of risk concerning which investors may have incomplete information, creating an opportunity for cream skimming.

[31] Origination channel is a requested but not a required field for rating agency RMBS models. Even when loans originated through the wholesale channel are identified, the investor or rating agency likely would not be able to distinguish between third-party originators who have long-term relationships with the bank from those that do not. The latter would be more strongly associated with agency problems and increased credit risk (Jiang, Nelson, and Vytlacil 2009).

than its availability, and as such could reflect superior controls over information processes. Cream-skimming in this case would imply that the likelihood of sale increases with the loan-to-income ratio.

Other credit-quality variables. Models 1 and 2 include several loan-level variables that are also indicators of credit quality, though not along dimensions where it seems particularly likely that banks would enjoy an information advantage. These variables are:

- APR spread—a set of dummy variables indicating the ranges [3.0, 3.25), [3.25, 5.25), [5.25, 7.0), or > 7.0
- An indicator for nonprimary residence (non-owner-occupied or second home)

Generally, larger APR spreads are associated with borrowers who are perceived (by both the originator and the market) to have greater risk of default, although APR spreads as measured in HMDA data are based on nominal maturities and therefore may not be comparable across fixed-rate and variable-rate mortgage products.[32] Very high APR loans also involve legal risks arising from predatory lending laws and regulations. For a given APR spread, the relationship between ownership type and credit risk is ambiguous, depending on how factors such as choice of loan product, prepayment speed, financial sophistication, and nature of the banking relationship differ by ownership type.[33]

Models 1 and 2 also include two measures of local area housing market conditions:

- MSA or (for non-MSA areas) state annual house-price appreciation
- MSA or state annual percent change in housing starts

Possibly, banks have information advantages along these local area dimensions as well, although not likely to the same degree as at the neighborhood level. Their relationship to perceived credit risk during 2005 and 2006 is somewhat ambiguous. While rising house prices and growth in housing starts usually imply reduced risk of default, they can also be associated with overheated or volatile markets.

[32] This problem is mitigated in 2005 and 2006 by the relatively flat yield curve environment. See Avery, Brevoort, and Canner (2006) for discussion of this issue.

[33] For example, if investors are more financially sophisticated, have other relationships with the bank, or have a slower prepayment speed, they may obtain more favorable credit terms than owner-occupants of similar credit risk, in which case investors will be riskier than owner-occupants with the same APR spread.

Model 1 incorporates some additional neighborhood (census tract) variables such that, although they may capture neighborhood risk factors, their expected relationship to cream-skimming is somewhat ambiguous. (We include them in Model 2 only when checking robustness.) They are:

- Percent of subprime home purchase loans that are junior lien
- Percent of subprime, first- lien home purchase loans tract that are high LTV (LTV $\geq$ 90)
- Percent of subprime, first- lien home purchase loans that are low FICO (FICO < 620)

On the one hand, neighborhood concentrations of junior-lien, high-LTV, or other higher-risk subprime loans can generate elevated neighborhood default rates that, as noted earlier, may have adverse spillover effects on collateral values. Asymmetric information regarding these neighborhood risk factors could motivate cream-skimming.[34] On the other hand, these variables might proxy for the loan-level risk characteristics (piggyback status, LTV, and FICO score) of the subprime loans originated by depository institutions, whereby the relationship of these variables to the likelihood of sale would reflect that of the underlying loan-level factors.[35] Since these risk characteristics are observable to investors in securities, there would be no role for cream-skimming. Moreover, an observed, positive relationship between the borrower's equity stake in the home (as represented by piggyback status and LTV) and the likelihood of sale, particularly at smaller banks, might reflect a risk diversification motive rather than cream-skimming activity. Banks seeking to eliminate geographic concentrations of loans in their portfolios in order to mitigate collateral risk tied to house-price volatility would tend to focus on loans with higher combined LTV, for which such risks are magnified.

Bank characteristics. As noted, the model estimated for large banks incorporates institution-specific fixed effects. For small banks, the following institution-specific characteristics are controlled for:

- Dummy variable identifying subprime specialist institutions[36]
- Set of dummy variables indicating primary regulator (OCC, Federal Reserve, FDIC, or OTS)
- Dummy variable for thrift institutions[37]

---

[34] Further, there are some indications that rating agency models underestimated the credit risk of loans originated with a piggyback second lien; if so, an opportunity to "cream-skim" based on piggyback status could arise. See http://online.wsj.com/article/SB118714461352698015.html

[35] Piggyback status is potentially endogenous, because the decision to slice the loan into senior and junior pieces may be associated with a decision to sell one piece and retain the other. The results are robust to dropping this variable.

[36] There are seven such institutions in the small bank sample in both 2005 and 2006.

- Institution size (log of total assets)

In particular, we expect that smaller institutions will face higher transactions costs associated with loan sale or securitization, so that the likelihood of sale will increase with size of the institution.

Other control variables. Additional control variables included in both Models 1 and 2 are:

- Loan size—a set of dummy variables indicating the ranges (in $1,000) < 55, [55,155), [155,255), and >255
- A dummy variable indicating whether the subject property is located in a metropolitan area
- A dummy variable identifying loans originated by subprime specialists with APR spread < 3.0

In addition, when estimating Model 1, for census tracts with missing values of the two variables from the LoanPerformance data (percent high LTV and percent low FICO), we set these value equal to zero. We then include two dummy variables identifying these cases. Similarly, when estimating Model 2, if the future subprime mortgage default rate in the census tract is missing, we set it equal to zero, and we include a dummy variable for these cases.

Row 6 of Table 1 indicates the number of observations excluded from the estimation samples due to missing data on borrower income, and Rows 7 and 8 provide the final sample sizes for Models 1 and 2, respectively. Table 2 provides sample mean and median values of the continuous variables included in the estimated logit equations, by loan disposition (retained or sold). Summary statistics for the large bank sample are shown in panel A, and for the small bank sample in panel B. Table 3, panels A and B, provides frequencies for the categorical variables (other than institution fixed effects) by value of the variable and loan disposition, for large and small banks, respectively.

Limitations of the analysis. An important limitation of the analysis is our inability to control at the loan level for loan and borrower characteristics related to credit risk or loan pricing but not reported in HMDA data. These include interest rate type or loan product category, borrower FICO score, and LTV or (for loans with piggyback seconds) combined LTV ratio.

---

[37] Thrift institutions are self-identified in the HMDA data; they may have savings and loan or savings bank charters and be OTS or FDIC insured.

Another limitation is that we cannot observe the extent to which banks may have provided credit enhancements or retained first loss or residual risk positions when selling or securitizing loans, although, anecdotally, sale and securitization of subprime mortgages by depository institutions generally involved significant risk transfer.[38] Moreover, we address only the loan sale or securitization decision of the banking side of the organization—many of the same subprime loans that were packaged into ABS may have cycled back to the trading or investment portfolios of these organizations.

6. Results

Estimation results for Model 1 are presented in Table 4. Results for the small bank sample are shown in panel A, and for the large bank sample in panel B. Estimation results for Model 2 are presented in Table 5 panels A and B for small and large banks, respectively. Overall, the results are consistent with cream-skimming, indicating a strong tendency for depository institutions to retain higher-quality loans and sell the lower-quality loans, along dimensions associated with an information advantage.

Looking first at the loan-to-income ratio, in all cases the estimated coefficient is statistically significant, with similar magnitude for large and small banks and between Models 1 and 2. A full unit increase in loan-to-income ratio (say, from 2 to 3) is associated with roughly a 20 percent increase in the likelihood of sale in 2005 and around 10 percent in 2006.

All three neighborhood risk factors in Model 1are statistically significant and show the likelihood of sale increasing with risk, both in 2005 and 2006. The relationship of log number of owner-occupied units to the likelihood of sale is much stronger for small compared to large institutions. Possibly, large banks are achieving geographic diversification benefits from retaining loans originated in smaller metropolitan areas, and these benefits partly offset the aforementioned neighborhood effects.

In Model 2, for both the large and small bank samples in both years, the likelihood of sale exhibits a statistically significant, positive relation to the future subprime mortgage default rate in the neighborhood. In 2005, for instance, large banks were 18 percent more likely to securitize and small banks 32 percent more likely to securitize per 10-percentage-point increment in future delinquency rate

---

[38] Typically, the investment bank arranging the sale of the security or a third party would retain first loss or residual risk positions. Covenants and warrants were generally uncommon for most mortgage ABS tranches.

of the census tract. The relationship between loan sale and the expected future default rate suggests that banks anticipated credit deterioration at the neighborhood level and used this information to select which loans to retain and which to securitize.

Results for out-of-market originations are mixed. For small banks, the estimated coefficient on the out-of-market dummy variable is statistically significant in both years and both models and indicates roughly a 30-percentage-point lower likelihood of sale for loans originated out-of-market. For large banks, there appears to be no relationship. One possible explanation of the different findings for large and small banks is that large banks may more consistently record origination channel. In that case, large banks would more consistently identify loans originated by brokers, and thus preclude cream-skimming. Another possible explanation is that large banks achieve greater geographic diversification benefit from retaining loans originated outside of the counties where they have branches.[39]

Other credit-quality variables. For both models in both years, for both large and small banks, the estimation results indicate that the likelihood of sale declines with APR spread. This relationship conforms to the view that banks tend to retain loans that securities market participants can observe to be higher risk. These results are consistent with those of Ambrose, LaCour-Little, and Sanders (2005), who find that banks hold higher-risk loans on book for purposes that include reputational risk and capital arbitrage.

Estimation results for local area housing market conditions and for the additional neighborhood risk variables included in Model 1 are mixed. Most notably, for the small bank sample, a larger share of high-LTV and piggyback subprime home purchase loans in a neighborhood is strongly associated with increased likelihood of sale. This result is consistent with the risk diversification motive posited above.

Other control variables. The estimated relationship of loan size to the likelihood of sale differs between small banks, which are most likely to retain the smallest loans, and large banks, which are most likely to retain the largest loans. One possible explanation is that small and large banks face differing types of transactions costs because they access the securities market in different ways. Small banks typically access the market indirectly by selling loans to aggregators who combine the loans of a number of banks into a security. This process may involve a relatively high transactions cost per loan, generating a preference toward selling larger loans. Moreover, selling larger loans may provide greater risk

[39] Out-of-market originations by large banks are likely to be more wide-ranging and distant than those of small banks.

diversification benefit, which may be more important for small than for large banks. Larger banks typically access the securities market directly by packaging their loans into pools, with pool-level transactions costs likely to predominate. This process may generate a preference toward retaining larger loans in order to reduce the per-loan cost of issuing a security.

For both models in both years, for both large and small banks, estimation results show that the likelihood of sale is lower for loans originated outside of metropolitan areas. Among small banks, this relationship may reflect higher transactions costs of securitization for banks located in rural areas. Among larger banks, it may reflect geographic diversification benefits associated with retaining loans originated in nonmetropolitan areas. Alternatively, it may reflect cream-skimming, since loans originated outside of metropolitan areas have been better performing.[40]

A number of control variables for type of institution included in the estimated equations for small banks are each statistically significant. Institution size exhibits a positive relation to the likelihood of sale, suggesting lower transactions costs of securitization for larger institutions. FDIC-regulated institutions (which are state-chartered and not members of the Federal Reserve System) have the lowest likelihood of loan sale, possibly tied to higher transactions costs. OTS-regulated institutions and thrift institutions have the highest likelihood of loan sale, which probably reflects a stronger risk-diversification motive due to the typically narrower product and geographic scope of their lending activities.

Robustness. We explored robustness of the estimation results in various ways. Including observations with missing income (by setting the loan-to-income ratio to zero for these observations, and identifying them with a dummy variable in the estimated equations) has no appreciable impact on any of the Model 1 or Model 2 estimation results. Likewise, the results are robust to excluding them from the sample loans sold to another depository institution.

Dropping the potentially endogenous variable percent of census tract subprime home purchase loans that are junior lien has little impact on the estimated coefficients of other variables in Model 1. Estimating Model 1 using the full HMDA sample, after omitting the two neighborhood variables derived from LoanPerformance data (percent of first-lien subprime mortgages that are high LTV and percent low

[40] Mean values for the January 2008 subprime default rate are 19.4 and 14.7 in the 2006 large bank sample and 16.8 and 13.9 in the 2006 small bank sample, for metropolitan and nonmetropolitan areas, respectively, and are nearly the same in the 2005 samples.

FICO) also has little impact on the estimated coefficients of other variables. Model 2 results are robust to including these two neighborhood variables.

All of the estimation results are robust to inclusion of neighborhood measures of subprime product mix (shares of standard ARMs, hybrid ARMs, and nontraditional products) derived from LoanPerformance data. They have no appreciable impact on sign, magnitude, or statistical significance of the other variables, nor do they exhibit consistent relationships to loan disposition.

7. Conclusion

Depository institutions may use information advantages along dimensions not observed or considered by outside parties to "cream-skim," meaning to transfer risk to naïve, uninformed, or unconcerned investors through the sale or securitization process. This paper provides evidence of cream-skimming by depository institutions in the subprime mortgage securitization market. Using Home Mortgage Disclosure Act (HMDA) data merged with data on subprime loan delinquency by ZIP code, we examine the bank decision to sell (securitize) subprime mortgages originated in 2005 and 2006. We find that the likelihood of sale increases with risk along dimensions observable to banks but not likely observed or considered by investors.

In particular, the likelihood of sale increases with the ratio of loan amount to borrower income, which suggests that depository institutions applied more robust measures of borrower repayment capacity than those used by ratings agencies or investors. The likelihood of sale also increases along neighborhood dimensions of credit risk, including the proportion of loans that are high cost and the proportion originated by subprime specialists in the ZIP code where the property is located. In addition, among small depository institutions, the likelihood of sale is greater for loans originated out-of-market and it is inversely related to the depth of the neighborhood housing market. For both large and small institutions, the likelihood of sale is positively related to the ex-post (January 2008) rate of serious delinquency among subprime loans in the ZIP code where the property is located, when this is substituted for individual neighborhood risk factors in an alternative model specification. To the extent that depository institutions are better informed regarding neighborhood risk factors, these results provide evidence of cream-skimming behavior.

While these findings are consistent with either cream-skimming or lemons market-type explanations, we believe that overall the empirical results favor the cream-skimming view. First, the cream-skimming explanation is consistent with the ex-post outcome in this market, where credit losses have greatly exceeded expectations. Second, in a lemons market equilibrium in which the loan sale decision separates risky from safe loans, investors would be aware of their disadvantage and the higher risk characterizing the sold loans would be reflected in pricing of the loans. Our empirical results, however, indicate an inverse relationship between the APR spread and the likelihood of sale, contrary to the prediction of the lemons market model. This evidence also aligns with research that shows gaps in risk detection among ratings agencies and investors, which could have allowed "cream-skimming" to go unchecked until the financial crisis of 2007 had become fully recognized.

Our findings, on the one hand, suggest that depository institutions had information advantages that enabled them to cream-skim in deciding which subprime mortgages to securitize and sell to investors during 2005 and 2006. Depository institutions' share of the subprime loan origination market was relatively small, however, so that cream-skimming by itself cannot be considered a major factor in precipitating the subsequent breakdown of the nonagency mortgage securitization market. The findings, on the other hand, suggest that investors had been inattentive to risk or ill-informed along critical dimensions, including the repayment capacity of the borrower and neighborhood concentrations of credit risk. Alternatively, they had less effective controls around information processes than depository institution originators, because of a breakdown of due diligence or misaligned incentives. From this perspective, the findings contribute to an understanding of the dynamics leading to the market collapse, and they are consistent with and complementary to other recent research highlighting the role of agency problems in the subprime mortgage securitization market.

The several caveats noted earlier bear repeating. An important limitation of the analysis is our inability to control at the loan level for loan and borrower characteristics related to credit risk or loan pricing but not reported in HMDA data. Also, we cannot observe the extent to which banks may have provided credit enhancements or retained first loss or residual risk positions when selling or securitizing loans, and we address only the loan sale or securitization decision of the banking side of the organization. Many of the same subprime loans that were packaged into ABS may have cycled back to the trading or investment portfolios of these organizations.

Table 1: Sample Sizes (Loan Counts) at Various Stages of Sample Development

Stages of Sample Development	2005	2006
1. Initial HMDA sample[1]	1,404,749	1,289,036
2. HMDA Exclusions		
(a) Nondepository institution[2]	(1,063,743)	(891,255)
(b) Sold to affiliate or depository	(43,453)	(48,114)
(c) Other exclusions[3]	(17,160)	(2,984)
(c) Remaining	280,393	346,683
3. Merge to LoanPerformance TrueStandings Data[4]		
Missing data		
(a) Large institutions	(22,853)	(31,426)
(b) Small institutions	(18,588)	(21,590)
4. Remaining	238,952	293,667
(a) Large institutions	181,101	230,090
(b) Small institutions	57,851	63,577
5. Percent securitized exclusions		
Large institutions only		
(a) Sold < 5% or Sold >90%	(121,888)	(120,311)
(b) Remaining	59,213	109,779
6. Missing income		
(a) Large institutions	(4,366)	(7,365)
(b) Small institutions	(4,792)	(5,763)
7. Final Sample: Model 1 with Fixed Effects[5]		
Large institutions[6]		
(a) Observations read	54,847	102,414
(b) Observations used	54,798	102,239
(c) Total securitized	25,007	58,474
Small institutions		
(d) Observations read	53,059	57,814
(e) Observations used	53,026	57,631
(f) Total securitized	10,128	11,495
8. Final Sample: Model 2 without Fixed Effects[5]		
Large institutions[6]		
(a) Observations read	54,847	102,414
(b) Observations used	54,847	102,414
(c) Total securitized	25,024	58,587
Small institutions		
(a) Observations read	53,059	57,814
(b) Observations used	53,051	57,687
(c) Total securitized	10,141	11,513

[1] The HMDA base population consists of subprime, home purchase, first-lien originations, excluding loans in U.S. territories, missing geography, and outliers for loan amount, rate spread, and loan-to-income ratio.

[2] Depository institutions include commercial banks, savings banks, and credit unions. Nondepository institutions refer to commercial bank subsidiaries, subsidiaries of commercial bank holding companies, thrift institution subsidiaries, subsidiaries of a thrift holding company, liquidated commercial bank or thrift institutions, subsidiaries of a credit union, or independent mortgage banks.

[3] Other exclusions were large institutions with less than 100 subprime loans and subprime loans originated by the depository subsidiary of AIG.

[4] The number reported here is the number dropped from the estimation sample for Model 1, which is slightly different from that for Model 2.

[5] An additional number of observations were excluded due to missing information on number of owner-occupied units.

Table 2: Mean and Median Values of Continuous Variables

Panel A: Banks with assets < 10 billion

Variable	Retained/ Securitized	2005 mean	2005 median		2006 mean	2006 median
Loan amount	Retained	90.99	56.00		109.72	70.00
	Securitized	150.66	122.00		172.16	138.00
APR spread	Retained	4.42	4.01		4.33	3.91
	Securitized	4.22	3.86		4.04	3.57
Future delinquency rate in census tract	Retained	15.4%	14.9%		15.8%	15.0%
	Securitized	17.4%	16.6%		17.2%	16.4%
Tract pct second lien	Retained	20.7%	20.0%		20.7%	20.0%
	Securitized	30.0%	32.0%		30.1%	31.3%
Tract pct FICO< 620	Retained	53.0%	53.1%		54.6%	55.3%
	Securitized	47.8%	48.1%		49.5%	50.3%
Tract pct LTV > 90	Retained	19.2%	18.2%		22.7%	22.2%
	Securitized	17.1%	16.1%		20.4%	19.1%
Tract pct subprime specialist	Retained	43.7%	44.9%		32.7%	33.3%
	Securitized	53.5%	54.8%		38.5%	39.0%
Tract pct high cost	Retained	32.7%	31.0%		34.6%	33.2%
	Securitized	31.7%	29.2%		33.8%	31.8%
Loan-to-income ratio	Retained	1.42	1.09		1.52	1.22
	Securitized	2.19	2.17		2.13	2.12
Log of census tract owner-occupied units	Retained	7.24	7.29		7.23	7.29
	Securitized	7.16	7.21		7.16	7.22
Local area house-price appreciation rate	Retained	9.9%	7.4%		6.5%	6.0%
	Securitized	12.8%	8.7%		6.6%	6.4%
Local area pct change in housing starts	Retained	7.8%	5.7%		-8.9%	-9.3%
	Securitized	5.5%	4.4%		-12.3%	-13.9%

Table 2, cont'd.

Panel B: Banks with assets > 10 billion

Variable	Retained/ Securitized	2005 mean	2005 median		2006 mean	2006 median
Loan amount	Retained	167.61	129.00		204.94	159.00
	Securitized	203.74	159.00		203.43	165.00
APR spread	Retained	4.44	4.17		4.49	4.04
	Securitized	4.81	4.84		5.06	5.11
Future delinquency rate in census tract	Retained	17.9%	17.3%		18.2%	17.5%
	Securitized	19.4%	19.0%		19.2%	18.6%
Tract pct second lien	Retained	28.5%	30.8%		29.6%	31.3%
	Securitized	33.2%	36.0%		32.9%	35.2%
Tract pct FICO< 620	Retained					
	Securitized					
Tract pct LTV > 90	Retained	16.6%	15.5%		19.0%	17.4%
	Securitized	14.8%	13.7%		18.5%	16.8%
Tract pct subprime specialist	Retained	53.9%	55.4%		39.3%	39.8%
	Securitized	59.3%	60.6%		43.4%	43.8%
Tract pct high cost	Retained	31.9%	29.3%		34.5%	32.3%
	Securitized	32.8%	30.6%		37.5%	35.7%
Loan-to-income ratio	Retained	2.18	2.15		2.16	2.13
	Securitized	2.72	2.72		2.43	2.47
Log of census tract owner-occupied units	Retained	7.14	7.21		7.15	7.23
	Securitized	7.10	7.17		7.07	7.15
Local area house-price appreciation rate	Retained	13.6%	10.4%		6.7%	6.4%
	Securitized	14.0%	12.1%		6.2%	6.1%
Local area pct change in housing starts	Retained	6.5%	4.8%		-12.2%	-11.7%
	Securitized	5.5%	4.4%		-13.7%	-13.5%

Table 3: Mean Values of Categorical Variables

Panel A: Banks with assets < 10 billion

Variable	value	Frequencies		% Sold	
		y2005	y2006	y2005	y2006
Indicator for loan sale	0	76.7%	76.2%	0.0%	0.0%
	1	23.3%	23.8%	100.0%	100.0%
Insufficient data for measuring delinquency rate	0	97.8%	98.5%	23.6%	24.0%
	1	2.2%	1.5%	10.8%	13.1%
Bank has branch in county of origination	0	12.7%	10.8%	33.1%	35.6%
	1	87.3%	89.2%	21.9%	22.4%
Non-owner-occupied property	0	73.8%	72.5%	25.7%	26.0%
	1	26.2%	27.5%	16.7%	18.0%
Insufficient data for measuring LTV	0	99.8%	99.7%	23.4%	23.8%
	1	0.2%	0.3%	11.7%	15.4%
Metropolitan area indicator	0	23.2%	23.2%	13.4%	14.0%
	1	76.8%	76.8%	26.4%	26.8%
HUD subprime specialist	0	96.9%	97.7%	22.4%	23.3%
	1	3.1%	2.3%	54.4%	46.4%
OCC-regulated banks	1	22.9%	24.5%	25.7%	27.8%
FRB	2	11.8%	9.7%	34.2%	29.3%
FDIC	3	41.8%	43.1%	9.1%	13.6%
OTS	4	18.4%	18.4%	46.9%	39.9%
Thrift institution	0	77.8%	77.2%	18.1%	19.9%
	1	22.2%	22.8%	41.7%	37.0%

Table 3, cont'd.

Panel B: Banks with assets > 10 billion

Variable	value	Frequencies		% Sold	
		y2005	y2006	y2005	y2006
Indicator for loan sale	0	19.8%	30.6%	0.0%	0.0%
	1	80.2%	69.4%	100.0%	100.0%
Insufficient data for measuring delinquency rate	0	271.7%	214.5%	80.2%	69.4%
	1	2.0%	1.1%	70.5%	56.7%
Bank has branch in county of origination	0	77.0%	49.9%	79.7%	68.2%
	1	196.7%	165.7%	80.3%	69.7%
Non-owner-occupied property	0	239.0%	171.1%	82.3%	71.2%
	1	34.7%	44.5%	65.4%	62.4%
Insufficient data for measuring LTV	0	273.6%	215.4%	80.2%	69.4%
	1	0.2%	0.3%	69.5%	56.5%
Metropolitan area indicator	0	24.2%	21.2%	67.2%	58.0%
	1	249.6%	194.4%	81.4%	70.6%
HUD subprime specialist	0	271.8%	214.2%	80.1%	69.3%
	1	1.9%	1.4%	97.0%	82.4%
OCC-regulated banks	1	142.4%	77.8%	77.3%	64.0%
FRB	2	5.1%	7.8%	31.9%	39.5%
FDIC	3	99.4%	66.7%	90.5%	84.3%
OTS	4	26.8%	63.2%	66.3%	63.9%
Thrift institution	0	149.4%	88.8%	74.9%	60.0%
	1	124.3%	126.8%	86.5%	75.9%

Table 4: Model 1 Estimation Results

Panel A: Depository institutions with assets less than $10 billion
Dependent Variable: Subprime loan was sold (1,0)

Variable	Description	2005 Odds Ratio	2005 Chi Square	2006 Odds Ratio	2006 Chi Square
Specialist	Lender is subprime specialist	0.29	61.7*	0.05	67.7*
OCC1	Lender is a national bank	0.86	0.6	0.84	7.3*
FRB1	Lender is state-chartered bank and member of the Federal Reserve System	0.93	2.7	0.76	0.7
FDIC1	Lender is a state-chartered nonmember bank	0.42	910.8*	0.47	532.8*
OTS1	Lender is an S&L	1.54	103.3*	0.96	20.2*
Thrift	Thrift institution indicator	1.41	27.0*	1.58	71.7*
Tract % high cost	Fraction of tract loans that are high cost	2.70	96.3*	3.70	185.8*
Specialist rate0	Lender is subprime specialist; loan not high cost	8.35	35.1*	44.28	47.9*
Rate1	APR spread in [3.0, 3.25)	1.72	48.3*	2.34	170.5*
Rate2	APR spread in (3.25,5.25]	1.39	20.0*	1.67	67.2*
Rate3	APR spread in (5.25,7.0]	1.27	9.3*	1.75	63.7*
Investor	Loan for nonprimary residence	0.79	52.0*	0.71	138.7*
Loan-to-income ratio	Ratio of loan amount to borrower income	1.17	145.9*	1.09	47.9*
Log owner units	Log of census tract owner-occupied units	0.71	218.7*	0.79	134.5*
Lnsize1	Loan amount < $55,000	0.45	188.0*	0.40	327.6*
Lnsize2	Loan amount in [$55,000, $155,000)	1.10	4.0**	1.00	0.0
Lnsize3	Loan amount in [$155,000, 255,000)	1.21	13.7*	1.07	2.8
Metro	Property in an MSA	1.08	4.7**	1.14	15.9*
Branch	Indicator for loan originated in county where bank has a branch	0.70	99.2*	0.66	145.8*
Log assets	Log of institution total assets	1.24	540.5*	1.24	518.2*
Tract % second	Fraction of tract's high cost home purchase loans that are 2nd lien	18.56	489.1*	32.21	834.7*
Tract % Specialist	Subprime specialist share of subprime loans in tract	5.07	204.1*	2.79	84.4*
% LTV GT 90	Fraction of census tract's subprime home purchase loans with LTV ³ 90	3.99	78.8*	4.42	133.8*
LTV Dummy	Insufficient data for measuring LTV distribution	1.32	0.5	1.87	7.1*
% Low FICO	Fraction of census tract's subprime home purchase loans with FICO < 620	0.92	0.5	0.82	3.6
HPI change	Annual rate of change in local area HPI in year	1.81	11.7*	8.47	46.6*
MSASTARTS	Annual rate of change in local area housing starts in year	0.33	170.6*	1.09	1.2
Sample Size			52,781		56,021
C statistics			0.819		0.783

* Statistically significant at the 1 percent level

** Statistically significant at the 5 percent level

Table 4, cont'd.

Panel B: Depository institutions with assets greater than $10 billion Dependent Variable: Subprime loan was securitized (1,0)					
Variable	Description	2005		2006	
		Odds Ratio	Chi Square	Odds Ratio	Chi Square
Tract % high cost	Fraction of loans in census tract that are high cost	2.44	134.4*	1.83	100.0*
Rate1	APR spread in [3.0, 3.25)	7.57	486.3*	2.09	262.7*
Rate2	APR spread in [3.25, 5.25)	5.67	377.1*	1.59	120.3*
Rate3	APR spread in [5.25, 7.0)	3.24	166.7*	1.54	101.7*
Investor	Loan indicator for nonprimary residence	1.43	146.8*	1.33	158.5*
Loan-to-income ratio	Ratio (loan-level) created by dividing loan amount by income	1.18	210.2*	1.09	94.3*
Log owner units	Natural log of census tract owner-occupied units	0.94	13.1*	0.97	5.0**
Loan size1	Loan size (0, 55K]	2.49	333.5*	1.96	324.0*
Loan size2	Loan size (55, 155K]	2.78	744.6*	2.31	1100.1*
Loan size3	Loan size (155, 255K]	1.93	328.5*	1.70	540.9*
Metro	Metro area loan indicator	1.11	12.8*	1.19	52.4*
Branch	Indicator for county where bank has a branch	0.98	0.7	1.02	1.4
Tract % Second	Percent of 2nd liens in census tract	0.99	0.0	1.78	51.2*
Tract % Specialist	Percent of HUD subprime specialists in the census tract	1.48	17.5*	1.57	35.5*
% LTV GT 90	Fraction of subprime home purchase loans in the census tract with LTV greater than 90%	2.46	46.4*	2.17	68.3*
LTV dummy	Insufficient data for estimating LTV distribution	1.38	1.3	1.02	0.0
% Low Fico	Fraction of subprime home purchase loans in the census tract with FICO less than 620	1.24	4.9**	0.62	46.2*
HPI change	Annual house-price rate of change for given year	0.39	55.6*	0.38	24.2*
MSA Housing starts	Annual housing starts rate of change for the metropolitan area in given year	0.86	5.2**	1.23	14.6*
Sample Size			52,042		91,882
C statistics			0.754		0.714
* Statistically significant at the 1 percent level					
** Statistically significant at the 5 percent level					

Table 5: Model 2 Estimation Results

Panel A: Depository institutions with assets less than $10 billion
Dependent Variable: Subprime loan was sold (1,0)

Variable	Description	2005 Odds Ratio	2005 Chi Square	2006 Odds Ratio	2006 Chi Square
Specialist	Lender is subprime specialist	0.42	31.5*	0.07	56.9*
OCC1	Lender is a national bank	0.80	0.41	0.85	41.7*
FRB1	Lender is state-chartered bank and member of the Federal Reserve System	0.94	17.9*	0.66	6.2**
FDIC1	Lender is state-chartered nonmember bank	0.36	1117.4*	0.41	699.8*
OTS1	Lender is an S&L	1.30	74.7*	0.84	11.7*
Thrift	Thrift institution indicator	1.65	60.3*	1.73	109.2*
Branch	Indicator for loan originated in county where bank has a branch	0.71	101.0*	0.68	132.2*
Specialist rate0	Lender is subprime specialist; loan not high cost	8.81	38.8*	31.1	40.5*
Rate1	APR spread in [3.0, 3.25)	1.71	47.8*	1.83	91.8*
Rate2	APR spread in (3.25,5.25]	1.41	22.0*	1.31	19.6*
Rate3	APR spread in (5.25,7.0]	1.35	14.5*	1.47	31.6*
Log assets	Log of institution total assets	1.27	671.6*	1.26	614.3*
Investor	Loan for nonprimary residence	0.86	20.9*	0.78	71.7*
Loan-to-income ratio	Ratio of loan amount to borrower income	1.20	216.8*	1.11	78.5*
Lnsize1	Loan amount < $55,000	0.37	334.0*	0.32	572.3*
Lnsize2	Loan amount in [$55,000, $155,000)	1.02	0.2	0.89	9.6*
Lnsize3	Loan amount in [$155,000, 255,000)	1.18	11.5*	1.02	0.2
Metro	Property in an MSA	1.49	142.5*	1.53	190.4*
HPI change	Annual rate of change in local area HPI in year	2.76	43.8*	15.75	85.1*
MSASTARTS	Annual rate of change in local area housing starts in year	0.34	180.9*	0.78	10.1*
Tract % Bad	Subprime 90+ delinquency rate in census tract as of Jan. 2008	3.19	35.1*	1.84	10.6*
Bad rate dummy	Insufficient data for measuring tract delinquency rate	0.72	8.5*	0.83	2.5
Sample Size			52,806		56,077
C statistics			0.800		0.761

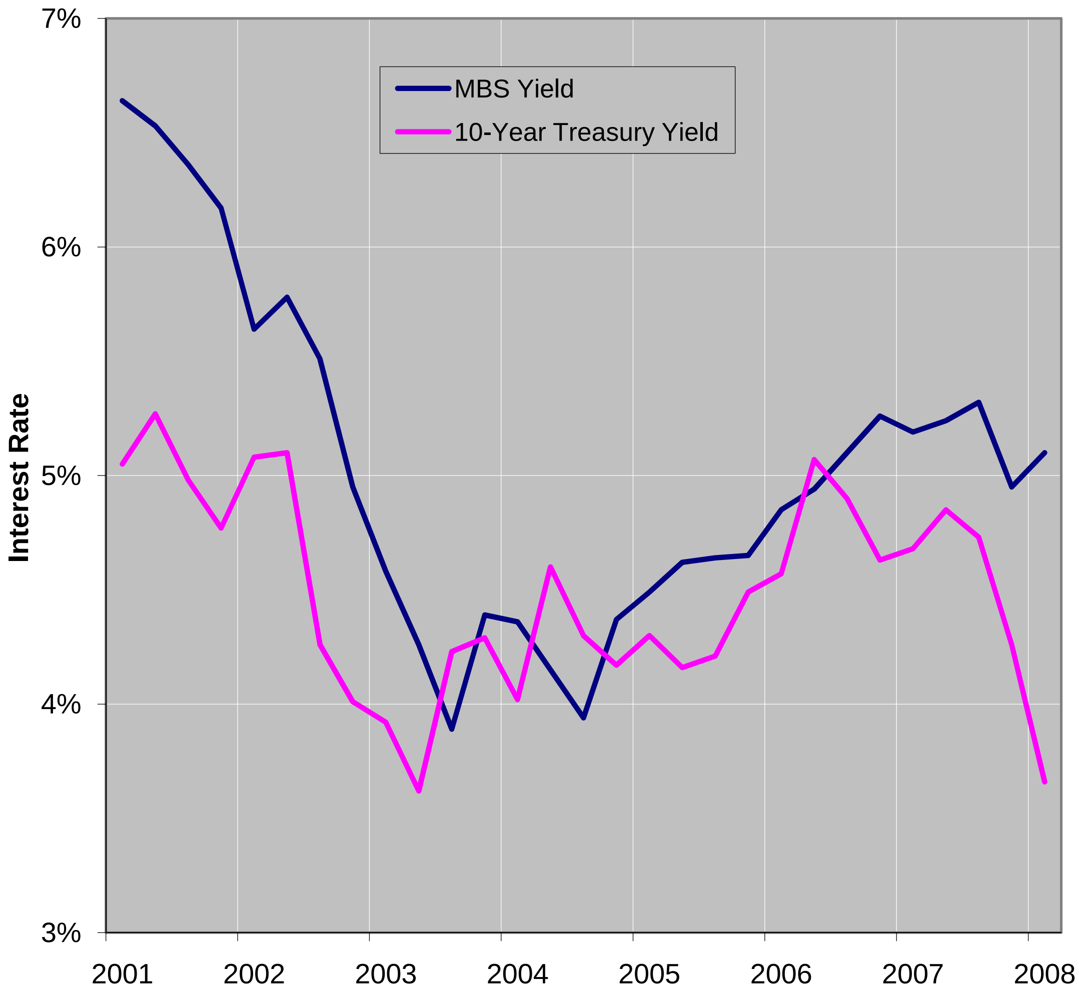
* Statistically significant at the 1 percent level

** Statistically significant at the 5 percent level

Table 5, cont'd.

Panel B: Depository institutions with assets greater than $10 billion Dependent Variable: Subprime loan was sold (1,0)					
Variable	Description	2005		2006	
		Odds Ratio	Chi Square	Odds Ratio	Chi Square
Branch	Bank has a branch in county were loan was originated	0.99	0.28	1.02	1.3
Rate1	APR spread in [3.0, 3.25)	7.18	464.1*	2.07	260.3*
Rate2	APR spread in [3.25, 5.25)	5.46	362.7*	1.58	120.1*
Rate3	APR spread in [5.25, 7.0)	3.17	161.6*	1.55	105.8*
Investor	Loan indicator for nonprimary residence	1.47	171.1*	1.34	173.3*
Loan-to-income ratio	Ratio (loan-level) created by dividing loan amount by income	1.19	228.3*	1.10	120.5*
Loan size1	Loan size (0, 55K]	3.33	705.4*	2.04	468.6*
Loan size2	Loan size (55, 155K]	3.36	1200.7*	2.36	1491.5*
Loan size3	Loan size (155, 255K]	2.09	432.1*	1.71	580.3*
Metro	Metro area loan indicator	1.15	22.1*	1.27	105.6*
HPI change	Annual house-price rate of change for given year	0.27	131.5*	0.50	12.5*
MSA Housing starts	Annual housing starts rate of change for the metropolitan area in given year	0.89	3.1	1.19	9.9*
Tract % bad	Subprime 90+ days delinquent in census tract as of Jan. 2008	1.76	13.0*	1.72	18.9*
Bad rate dummy	Insufficient data for measuring delinquency rate	0.88	1.6	0.86	2.7
Sample Size			52,090		92,052
C statistics			0.749		0.709
* Statistically significant at the 1 percent level					
** Statistically significant at the 5 percent level					

Figure 1: MBS Yield vs. 10-Year Treasury Yield



Source: Federal Reserve Board of Governors, Banking Conditions Report.

Figure 2: Average Credit Enhancement for Aaa Nonagency Prime Mortgage Securities

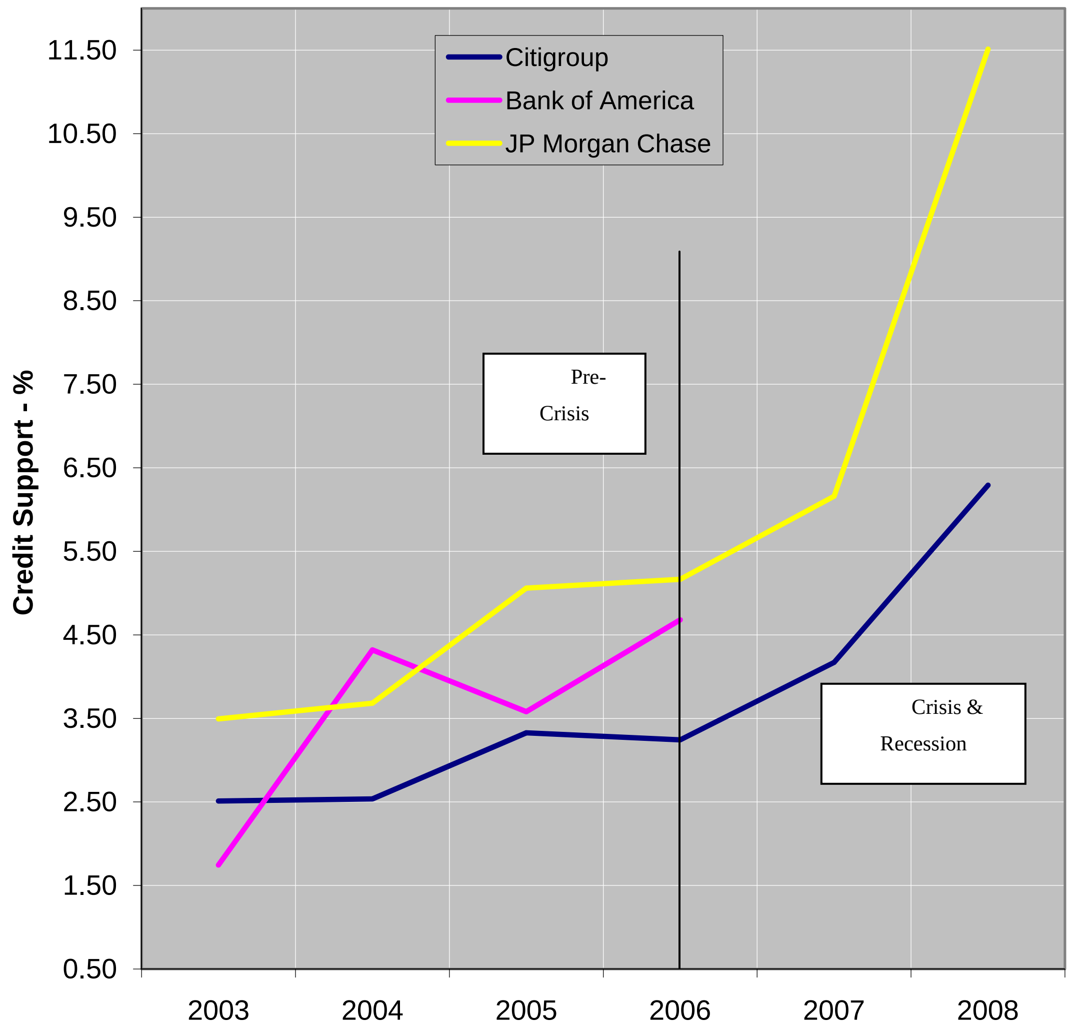
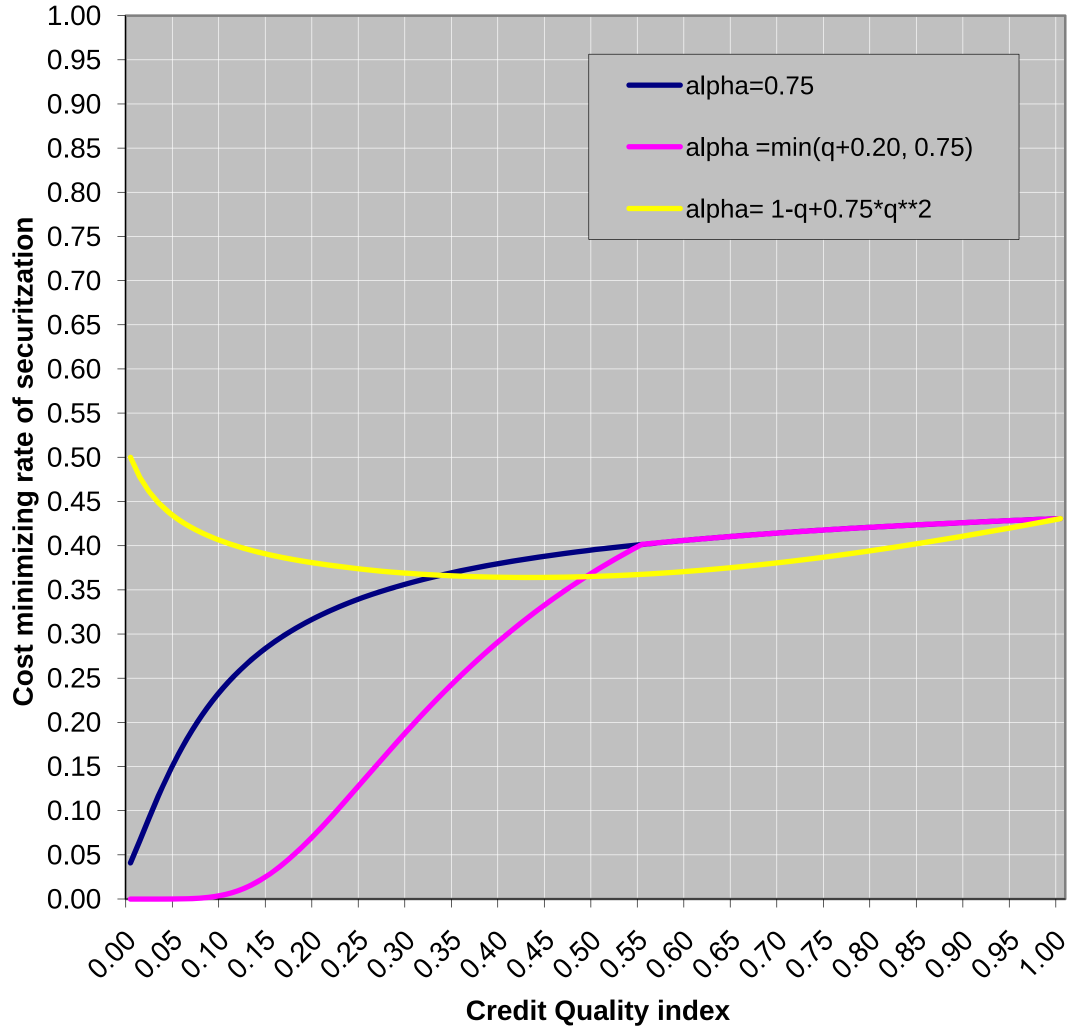


Figure 3: Optimal Rate of Securitization



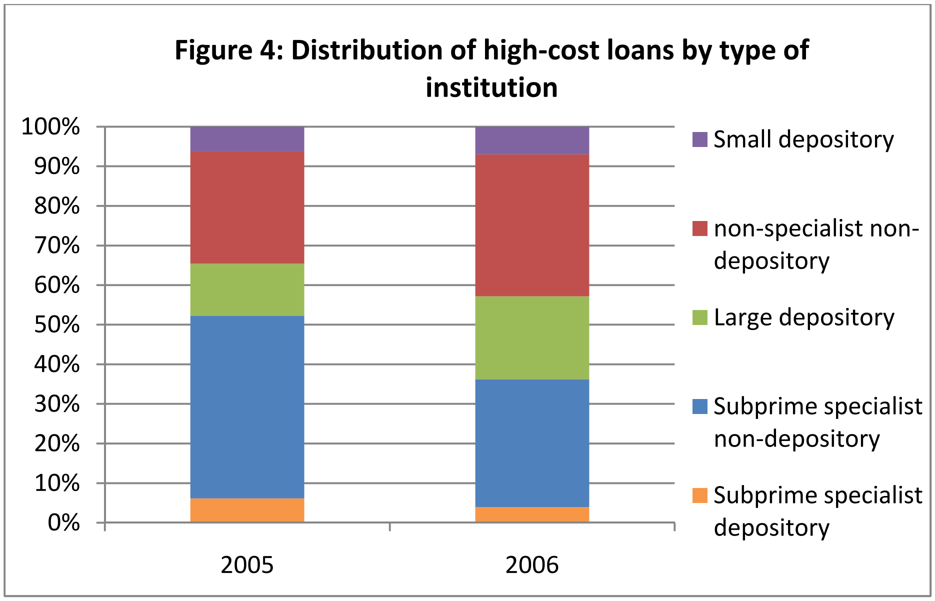
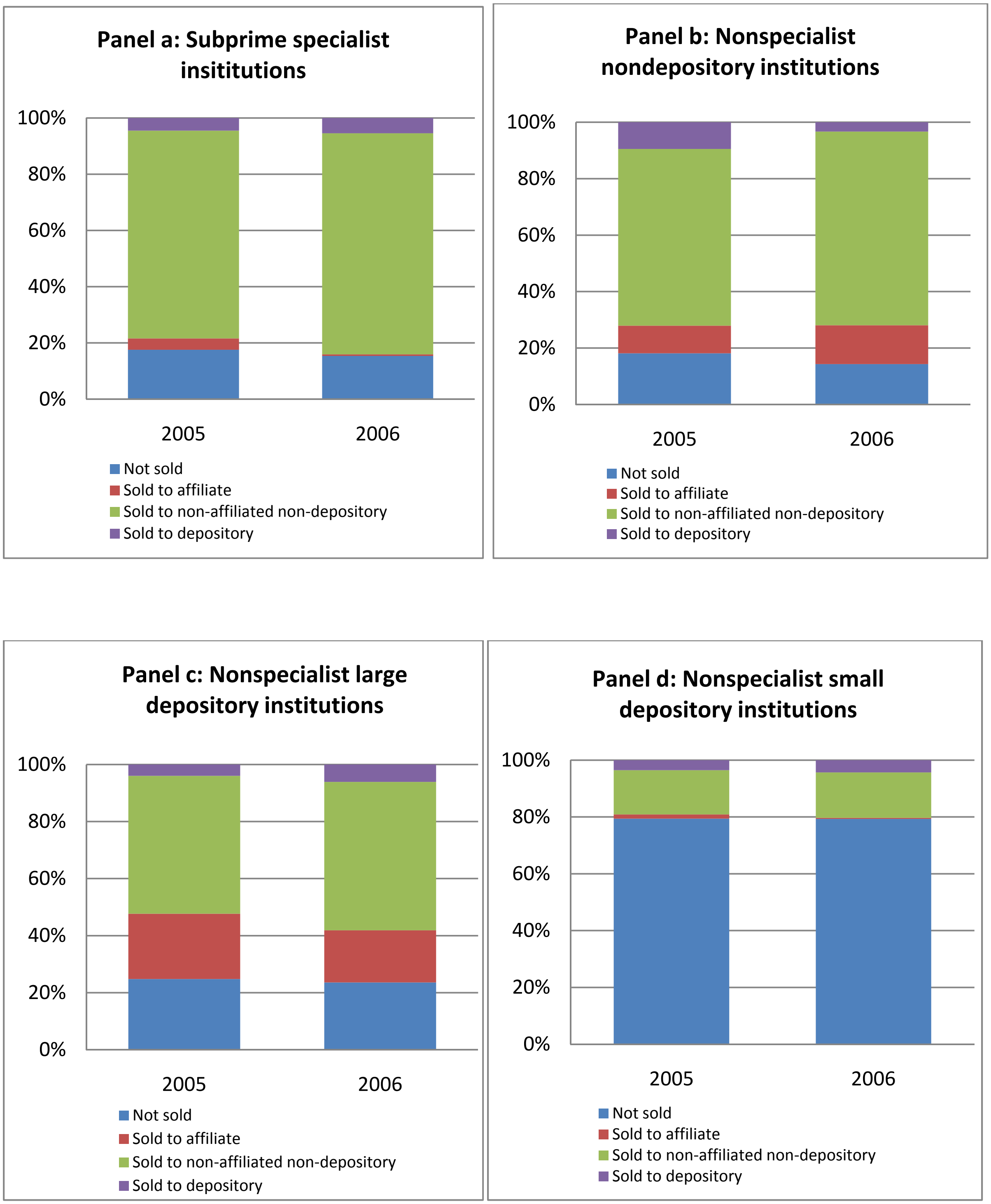


Figure 4: Distribution of high-cost loans by type of institution
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
2005
2006
Small depository
non-specialist non-depository
Large depository
Subprime specialist non-depository
Subprime specialist depository

Figure 5: Disposition of Subprime Loans (purchaser type)



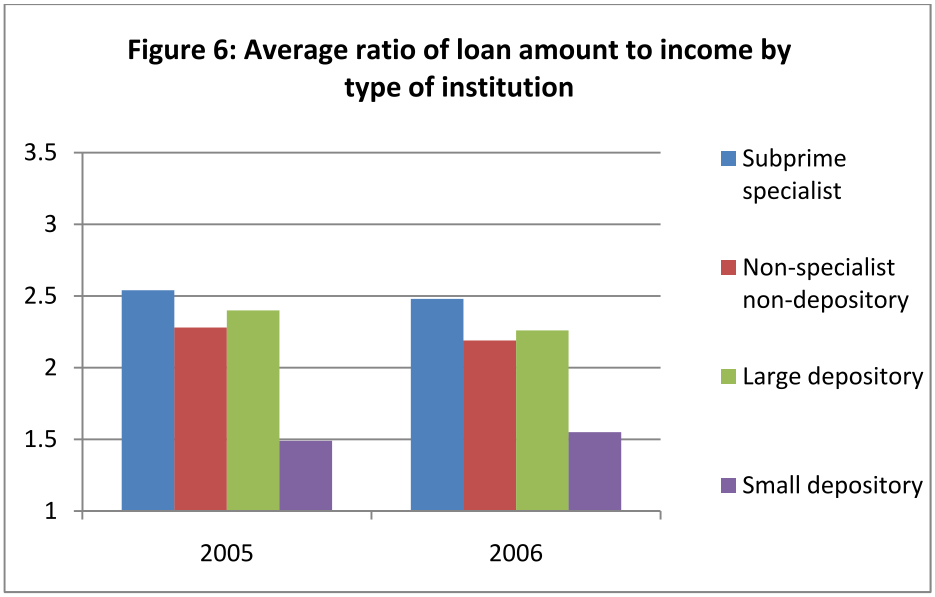


Figure 6: Average ratio of loan amount to income by type of institution
3.5
3
2.5
2
1.5
1
2005
2006
Subprime specialist
Non-specialist non-depository
Large depository
Small depository

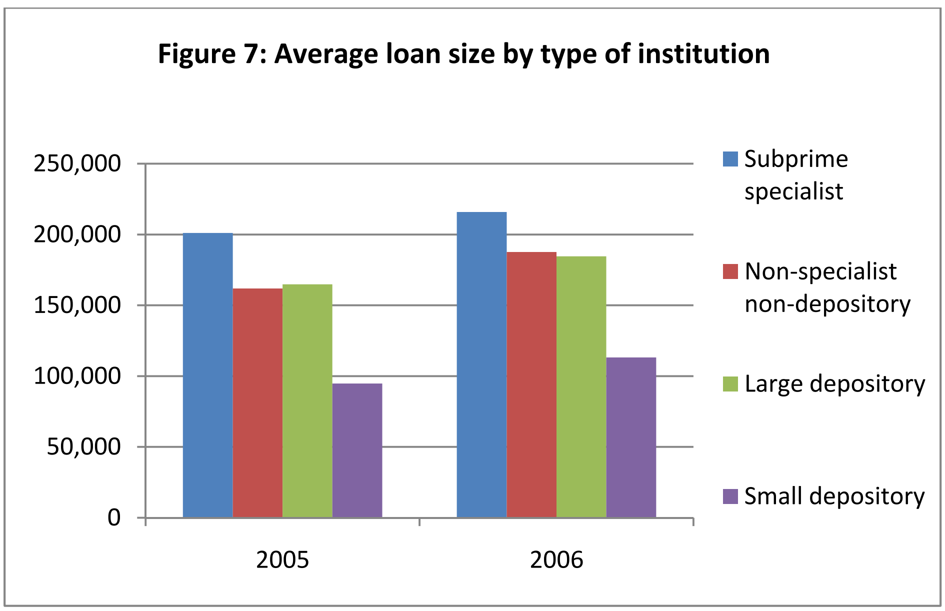


Figure 7: Average loan size by type of institution
250,000
200,000
150,000
100,000
50,000
0
2005
2006
Subprime specialist
Non-specialist non-depository
Large depository
Small depository

Figure 8: Mean APR spread for high-cost loans by type of institution

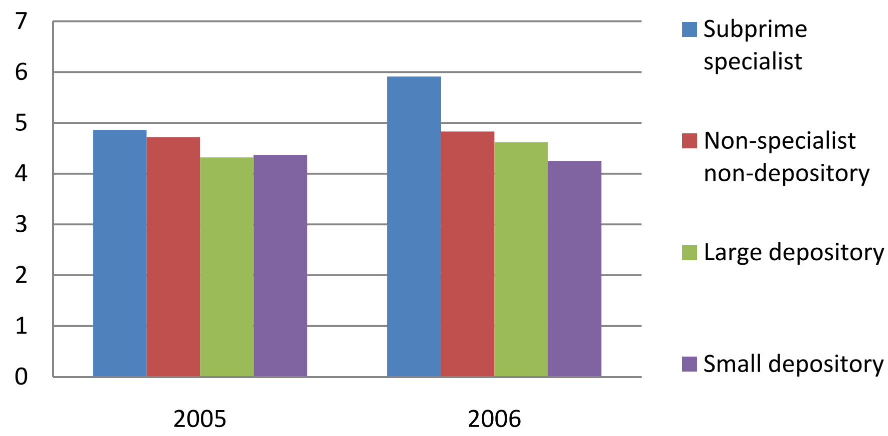
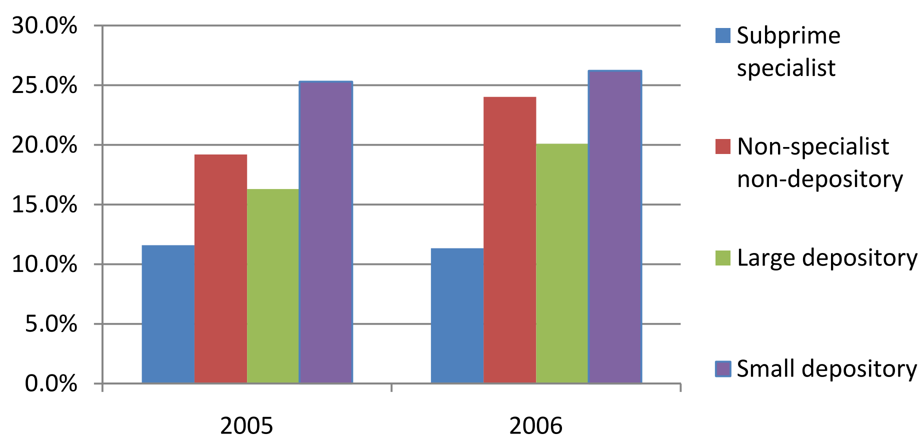


Figure 9: Percent of loans for nonprimary residence by type of institution



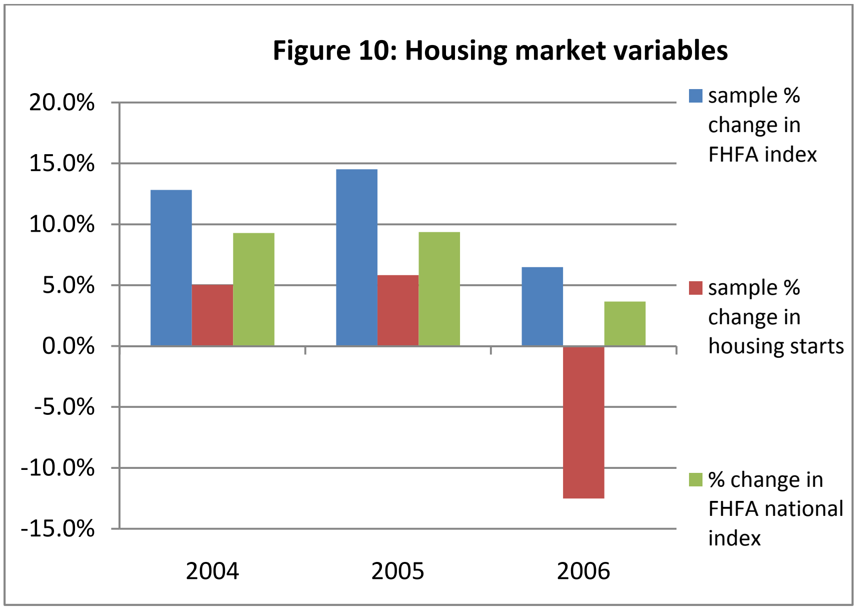


Figure 10: Housing market variables
20.0%
15.0%
10.0%
5.0%
0.0%
-5.0%
-10.0%
-15.0%
2004
2005
2006
sample % change in FHFA index
sample % change in housing starts
% change in FHFA national index

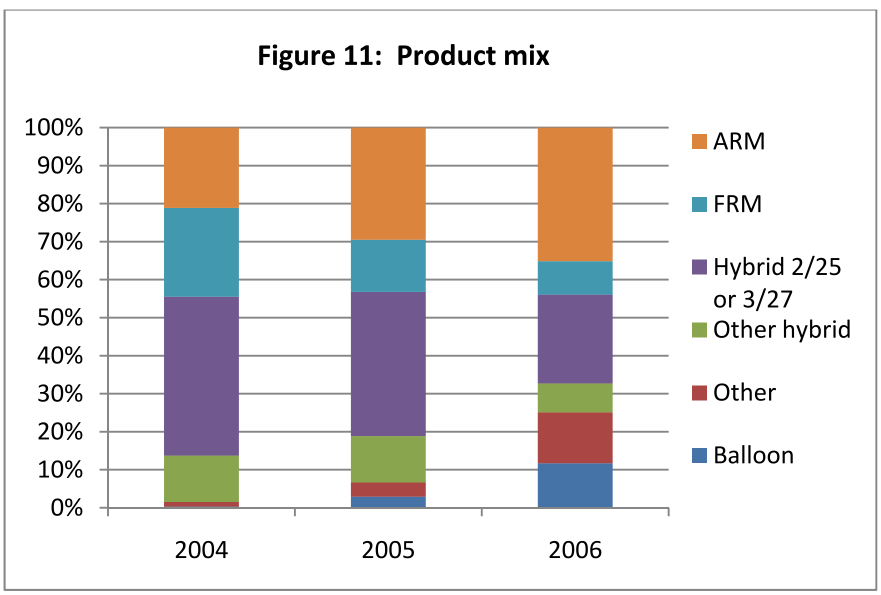


Figure 11: Product mix
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
2004
2005
2006
ARM
FRM
Hybrid 2/25 or 3/27
Other hybrid
Other
Balloon

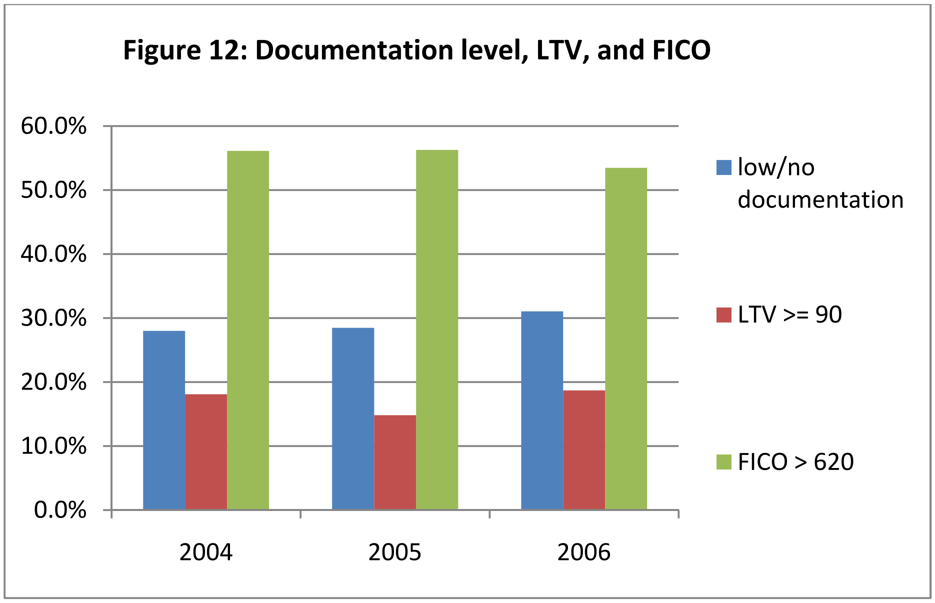


Figure 12: Documentation level, LTV, and FICO
60.0%
50.0%
40.0%
30.0%
20.0%
10.0%
0.0%
2004
2005
2006
low/no documentation
LTV >= 90
FICO > 620

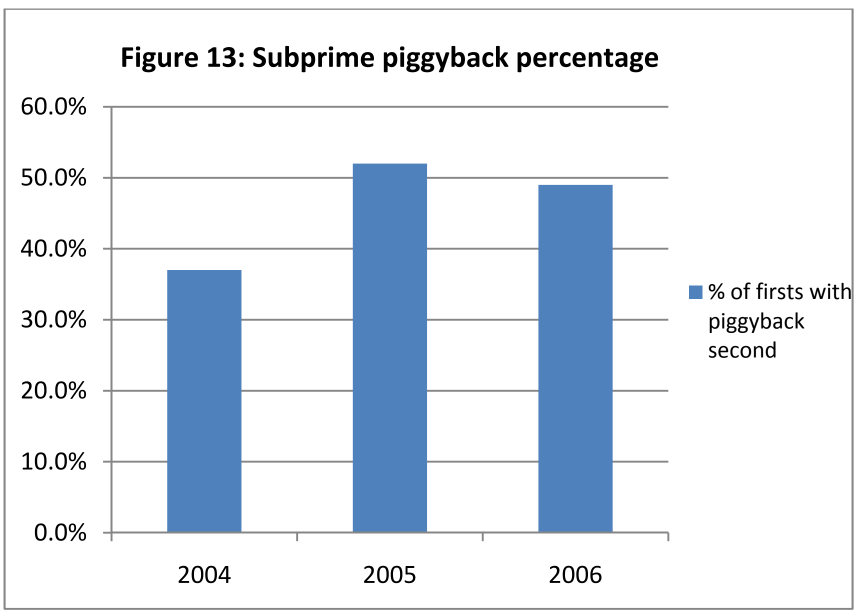


Figure 13: Subprime piggyback percentage
60.0%
50.0%
40.0%
30.0%
20.0%
10.0%
0.0%
2004
2005
2006
% of firsts with piggyback second

References

Akerlof, George A. (1970), “The Market for ’Lemons’: Quality Uncertainty and the Market Mechanism,” *Quarterly Journal of Economics* 84:3, pp. 488-500.

Ambrose, Brent, Michael LaCour-Little, and Anthony Sanders (2005), “Does Regulatory Capital Arbitrage, Reputation, or Asymmetric Information Drive Securitization?” *Journal of Financial Services Research*, 28:1, pp. 113-133.

Ashcraft, Adam, and Til Schuermann (2008), “Understanding the Securitization of Subprime Mortgage Credit,” Federal Reserve Bank of New York Staff Report no. 318.

Avery, Robert B., Kenneth P. Brevoort, and Glenn B. Canner (September 2006), “Higher-Priced Home Lending and the 2005 HMDA Data,” *Federal Reserve Bulletin* 92 pp. A123-A166.

Ben-David, Itzhak (2007), “Financial Constraints, Inflated Home Prices, and Borrower Default during the Real-Estate Boom,” Working Paper no. 2009-03-001, Fisher College of Business, Ohio State University.

Cagan, Christopher L. (March, 2007), “Mortgage Payment Reset the Issue and the Impact,” *First American* CoreLogic.

Calem, Paul S., and James R. Follain (2007), “Regulatory Capital Arbitrage and the Potential Competitive Impact of Basel II in the Market for Residential Mortgages,” *Journal of Real Estate Finance and Economics* 35 (2), pp. 197-219.

Calem, Paul S., and Michael LaCour-Little (2004), “Risk-Based Capital Requirements for Mortgage Loans,” *Journal of Banking & Finance* 28, pp. 647-672.

Coleman, Major, Michael LaCour-Little, and Kerry D. Vandell (2008), “Subprime Lending and the Housing Bubble: Tail Wags Dog?” *Journal of Housing Economics* 17, pp. 272-290.

Coval, Joshua, Jakub Jurek, and Erik Stafford (Winter 2009), “The Economics of Structured Finance,” *Journal of Economic Perspectives* 23(1), pp. 3-25.

DeMarzo, Peter (2005), “The Pooling and Tranching of Securities: A Model of Informed Intermediation,” *Review of Financial Studies Mini Issues* 18(1), pp. 1-35.

DeMarzo, Peter, and Darrell Duffie (January, 1999), “A Liquidity-Based Model of Security Design,”*Econometrica* pp. 65-99.

Demyanyk, Yuliya, and Otto Van Hemert (2009), “Understanding the Subprime Mortgage Crisis,” *Review of Financial Studies Mini Issues,* published online May 4, 2009.

Dewatripont, Mathias, and Jean Tirole (1994), *The Prudential Regulation of Banks.* Cambridge, Massachusetts: The MIT Press

Elul, Ronel (2009), “Securitization and Mortgage Default: Reputation vs. Adverse Selection,” Working Paper no. 09-21, Federal Reserve Bank of Philadelphia.

Ernst Keith, Debbie Bocian, and Wei Li (2008), “Steered Wrong: Brokers, Borrowers, and Subprime Loans,” Center for Responsible Lending.

Gerardi, Kristopher, Adam Hale Shapiro, and Paul S. Willen (2008), “Subprime Outcomes: Risky Mortgages, Homeownership Experiences, and Foreclosures,” *Federal Reserve Bank of Boston Working Papers No. 07-15.*

Golding, Edward, Richard K. Green and Douglas A. McManus (February 2008), “Imperfect Information and the Housing Finance Crisis,” *Joint Center for Housing Studies at Harvard University* UCC08-6.

Hahn, Robert and Peter Passell (February, 2008), “Better that the Fed Regulates Subprime Mortgages” *Economist Voice.*

Haughwout, Andrew, Richard Peach, and Joseph Tracy (August 2008), “Juvenile Delinquent Mortgages: Bad Credit or Bad Economy,” Federal Reserve Bank of New York Staff Report No. 341.

Hill, Claire A. (Winter, 1996), “Securitization: A Low-Cost Sweetener for Lemons,” Washington University *Law Quarterly* pp. 1061-1120.

Hull (2009), “The Credit Crunch of 2007: What Went Wrong? Why? What Lessons Can Be Learned?” Joesph L. Rotman School of Management, University of Toronto.

Jiang, Wei, Ashlyn Nelson, and Edward Vytlacil, (2009), “Liar’s Loan? Effects of Loan Origination Channel and Loan Sale on Delinquency,” manuscript, Columbia University.

Lee, Kai-yan (2008), **“**Foreclosure’s Price-Depressing Spillover Effects on Local Properties: A Literature Review,” Federal Reserve Bank of Boston Community Affairs Discussion Paper no. 2008-01.

Mian, Atif, and Amir Sufi (2009), “The Consequences of Mortgage Credit Expansion: Evidence from the U.S. Mortgage Default Crisis,” *Quarterly Journal of Economics,* 124:4, pp. 1449-1996.

Nakamura, Leonard (2010), “How Much Is That Home Really Worth? Appraisal Bias and House-Price Uncertainty,” *Federal Reserve Bank of Philadelphia Business Review* First Quarter.

Rajan, Uday, Amit Seru, and Vikrant Vig (2009), “The Failure of Models That Predict Failure: Distance, Incentives and Defaults,” University of Chicago Graduate School of Business Research Paper no. 08-19.

Sherlund, Shane M. (2008), “The Past, Present, and the Future of Subprime Mortgages,” Federal Reserve Board Finance and Economics Discussion Series 2008-63.

Smith, Brent C. (2007), “The Subprime Mortgage Market: A Review and Compilation of Research and Commentary,” *The Homer Hoyt Institute*.

White, Lawrence J. (2009), “The Credit Rating Agencies and the Subprime Debacle,” *Critical Review* 21 (2-3), pp. 389-399.

Appendix 1: Derivation of Propositions 1 and 2

To verify Proposition 1, let F(x, a_1, b_1) and F(x, a_2, b_2) be two beta distributions with identical medians x^m. Then F(x, a_1, b_1) and F(x, a_2, b_2) intersect at x=0, where F(x, a_1, b_1)=F(x, a_2, b_2)=0; at x=1, where F(x, a_1, b_1)=F(x, a_2, b_2)=1; and at x= x^m, where F(x, a_1, b_1)=F(x, a_2, b_2)=1/2. Next, we show that these are the only points at which the two distributions will intersect.

The beta distribution is defined as F(x, a, b)=B(x, a, b)/B(0, a, b), where B(x, a, b) is the beta function. Calculating the second derivative of the beta distribution, we obtain:

(A-1) $\partial^2F/\partial x^2 = [(a-1)x^{a-2}(1-x)^{b-1} - (b-1)x^{a-1}(1-x)^{b-2}]/B(x, a, b)$

It follows that:

(A-2) $\partial^2F/\partial x^2 > (=) (<) 0$ if and only if $[(a-1)(1-x) - (b-1)x] > (=) (<) 0$.

Hence, there can be at most one x_0 such that $\partial^2F/\partial x^2= 0$; i.e., at most one inflection point of the distribution, implying that any two beta distributions can intersect at most once in (0, 1). It follows that if F(x, a_1, b_1) and F(x, a_2, b_2) intersect at a common median value x^m, then either F(x, a_1, b_1) > F(x, a_2, b_2) for all x < x^m, or vice versa. It then follows immediately that F(x, a_1, b_1) > F(x, a_2, b_2) for all x < x^m if and only if F(x, a_1, b_1) has the wider interquartile range.

To verify Proposition 2, let F(x, a_1, b_1) and F(x, a_2, b_2) be two beta distributions. From (A-2), if $a_i < 1$ and $b_i \geq 1$ for i=1 and 2, then the distributions are strictly concave; if $a_i \geq 1$ and $b_i < 1$ for i=1 and 2 they are strictly convex. In either case, the distributions will intersect only at 0 and 1, from which it follows that F(x, a_1, b_1) > F(x, a_2, b_2) for all x in (0, 1) if and only if F(x, a_1, b_1) has the wider interquartile range.