ADVERSE SELECTION IN MORTGAGE SECURITIZATION[*]

Sumit Agarwal[1], Yan Chang[2], and Abdullah Yavas[3]

Abstract

We investigate lenders' choice of loans to securitize and whether the loans they sell into the secondary mortgage market are riskier than the loans they retain in their portfolios. Using a large dataset of mortgage loans originated between 2004 and 2008, we find that banks sold low-default risk loans into the secondary market while keeping higher-default risk loans in their portfolios. This result holds for both subprime and prime loans. We do find strong support for adverse selection with respect to prepayment risk; securitized loans had higher prepayment risk than portfolio loans. It appears that in return for selling loans with lower default risk, lenders retain loans with lower prepayment risk. Small lenders place more emphasis than large lenders on default risk versus prepayment risk of the loans they retain. Securitization strategies of lenders changed during the sample period as they became less willing to retain higher-default loans after the housing market reached its peak. There are also differences in the performance of loans sold to GSEs and loans sold to private issuers. Loans sold to private issuers have lower prepayment rates in each year while relative default rates vary across the years.

---

[*] We thank Jacqui Barrett for excellent research assistance. We are grateful to Brent Ambrose, Marsha Courchane, Tony Sanders, Peter Zorn, and the participants in Finance workshops at the University of South Carolina, National University of Singapore and the 2010 ASSA Meeting for their helpful comments.

[1] Senior Financial Economist, Research Department, Federal Reserve Bank of Chicago, email: ushakri@yahoo.com

[2] Credit and Prepayment Modeling Director, Freddie Mac, 8200 Jones Branch Drive, McLean, VA 22102, email: yan_chang@freddiemac.com

[3] UW Real Estate Program Distinguished Professor, School of Business, University of Wisconsin-Madison, 5251 Grainger Hall, 975 University Avenue, Madison, WI 53706, e-mail: ayavas@bus.wisc.edu

ADVERSE SELECTION IN MORTGAGE SECURITIZATION

Abstract

We investigate lenders' choice of loans to securitize and whether the loans they sell into the secondary mortgage market are riskier than the loans they retain in their portfolios. Using a large dataset of mortgage loans originated between 2004 and 2008, we find that banks sold low-default risk loans into the secondary market while keeping higher-default risk loans in their portfolios. This result holds for both subprime and prime loans. We do find strong support for adverse selection with respect to prepayment risk; securitized loans had higher prepayment risk than portfolio loans. It appears that in return for selling loans with lower default risk, lenders retain loans with lower prepayment risk. Small lenders place more emphasis than large lenders on default risk versus prepayment risk of the loans they retain. Securitization strategies of lenders changed during the sample period as they became less willing to retain higher-default loans after the housing market reached its peak. There are also differences in the performance of loans sold to GSEs and loans sold to private issuers. Loans sold to private issuers have lower prepayment rates in each year while relative default rates vary across the years.

1. Introduction

The U.S. economy is experiencing one of the worst financial and economic crises since the Great Depression. The crisis was triggered by collapse of the bubble in residential real estate markets. Many commentators cite the remarkable growth of securitization in recent years as a major contributor to the rise of the real estate bubble and the ensuing crisis. Part of the argument is that securitization creates additional layers of agency problems in loan origination, which leads to lax underwriting and thus to higher default rates.

In this paper, we empirically examine the potential adverse selection problems in mortgage securitization. In particular, we investigate whether loans into the secondary mortgage market are riskier than loans kept in lenders' portfolios. The conventional wisdom is that lenders may know more about the credit quality of a borrower than what is reflected in the hard information collected, such as the credit score, income and debt payments of the borrower. The lender may know more, for instance, about local real estate market trends, future prospects of the borrower's employer, and the borrower's job security. Lenders would have incentives to take advantage of their unobservable private information about borrowers and retain higher-quality loans on their balance sheets while selling inferior loans into the secondary mortgage market.

Investors in mortgage loans are concerned with three kinds of risk. Interest rate risk refers to the fact that a change in interest rates leads to an opposite change in the value of the mortgage. Interest rate risk is independent of the borrower's characteristics, and hence is not subject to potential adverse selection concerns. Prepayment risk refers to risk that mortgages may be repaid; prepayments often take place in the form of refinancing due to a decline in interest rate, which is precisely when prepayment is costly for the investor. Default risk refers to the likelihood that the borrower may stop making payments. Earlier studies of adverse selection in mortgage markets focus mostly on default risk. In this paper, we consider both prepayment risk and default risk, and show that both risks play a critical role in lenders' securitization strategies. The significance of the problem studied in the current paper is evident from the size of the market for mortgages and mortgage-related securities. Total volume of outstanding mortgage debt in the U.S., which has grown three times as fast as the gross domestic product in the last decade, is about $14.6 trillion. Total U.S. mortgage-related securities exceed $8.8 trillion. As the current crisis has illustrated, a jump in mortgage default rates can have dramatic consequences for the real economy.

Lenders typically sell mortgage loans in the secondary market to either Fannie Mae or Freddie Mac, which are government-sponsored enterprises (GSEs), or to a private sector financial institution, such as subsidiaries of investment banks, large banks, and homebuilders.[4] There are important differences between GSEs and private issuers that may impact their loan purchasing decisions. First, GSEs and private issuers differ with respect to default risk. GSEs offer investors guarantees against default risk, while private issuers often pass the default risk on to parties that are willing to bear it. Second, while mortgage securitizations by GSEs typically involve a single form of an investment bond called a Pass-Through Certificate, private placement of mortgage-backed securities involves multiple forms of investments created by splitting the principal and interest components of the mortgage pool into various tranches. A third difference is that GSEs have historically purchased only traditional fixed-rate mortgage products. They only recently began to purchase alternative mortgages such as adjustable-rate mortgages (ARMs), interest only ARMs, Alt-A, and subprime loans. Private label issuers have been purchasing these

[4] Secondary market institutions often create pools of loans and sell the payment rights of the loans in the pool to investors around the globe. Of the total volume of $7.6 trillion in pooled mortgages at the end of 2008, about $5 trillion is securitized or guaranteed by GSEs or government agencies. The remaining $2.6 trillion is pooled by private mortgage conduits (source: http://www.federalreserve.gov/econresdata/releases/mortoutstand/mortoutstand20090331.htm).

alternative mortgages on a much larger scale and for a longer time. Furthermore, GSE issuance of mortgage-backed securities is subject to detailed SEC filings and public reporting requirements, while private issuers are not. These differences give more flexibility to private label issuers, enabling them to create securities that better diversify the risks of individual loans. As a result of these differences, private label issuers might be willing to purchase some loans that GSEs would not. On the other hand, GSEs have implicit government guarantees, which might give them an incentive to be more aggressive in risk-taking than private issuers. As these differences are likely to impact the loan purchasing decisions of GSEs and private issuers, we distinguish between GSEs and private issuers in our analysis.

We use a large detailed dataset of residential mortgage loans from Lender Processing Services (LPS) to compare default and prepayment risks of loans retained in lenders' portfolios with those of loans sold to GSEs and private issuers for loans originated between 2004 and 2008. Results of our analysis of default outcomes defy the conventional wisdom. We show that originators chose to sell low-default risk, not high-default risk loans to GSEs and private issuers in all years studied.

What is interesting about these results is not only that lenders did not keep the low-default risk loans to themselves and sell the lemons in the secondary market, but also that they did not even randomize their choices of loans to securitize and loans to retain. Rather, they sold lower-default risk loans into the secondary market and retained higher-default risk loans in their portfolios. This result is surprising. Although lenders need to pay attention to their reputation with purchasers in the secondary market, reputational concerns alone cannot explain these results. Lenders could presumably maintain good reputations simply by randomizing their decisions of the loans to securitize.

One could argue that securitization often involves an exchange whereby a large originator assembles a pool of conforming mortgages and trades the pool with the GSEs or private issuers in return for securities backed by the pool of mortgages. Thus, one could argue that this practice significantly reduces the potential for any adverse selection problems. However, owning securities backed by a pool of mortgage loans is different from owning the loans directly. The reason is that the GSEs bear the default risk in the former case, but the originator bears the risk in the latter case. Furthermore, not all mortgage loan sales take place through a swap program; lenders sell a significant portion of their mortgage loans for cash. Thus, this argument cannot explain why banks would retain higher-default risk loans.

Another contributing factor could be regulatory capital requirements. The current risk-based capital rules require banks to have more capital reserves for higher-risk classes of loans. This gives banks incentives to retain riskier mortgage loans with higher expected return and securitize less risky mortgage loans, as long as both groups of loans have the same capital requirements when held on banks' balance sheets (Ambrose, LaCour-Little and Sanders, 2005). One compelling argument supporting our default results is that GSEs and private financial institutions that purchase mortgage loans from originators have very high underwriting standards. In order for an originator to be able to sell a loan to a GSE or private issuer, the loan needs to satisfy a certain set of criteria. If, for instance, the lender wishes to sell the loan to Fannie Mae, the lender needs to enter the loan and borrower data to the DeskTop Underwriting program of Fannie Mae and obtain approval. Freddie Mac has similar software called Loan Prospector. Although the specific formulae behind the approval decisions of the GSEs and private issuers are not known to lenders, they are believed to involve high standards. It is possible, therefore, that a subset of loans that fail to meet GSE and private issuer criteria are still acceptable to lenders. If there are fewer loans meeting secondary market criteria than the total demand by GSEs and private issuers, originators may choose to approve and retain some of the loans whose risk levels are below the standards of secondary market purchasers. Thus, even if the originators pick a higher-quality set of the conforming loans for their portfolios, the presence of nonconforming loans in their portfolios may push the average default rate of their portfolios above those of GSEs and private issuers. To capture this possibility, we use proxies to differentiate between conforming and non-conforming loans. Our default risk results remain unchanged when we control for whether or not a loan is conforming.

A related argument is that when a lender sells a loan into the secondary market, the lender reps and warrants to the purchasing institution that she has accurately followed the purchaser's underwriting requirements and that the loan meets the purchaser's standards. In order to implement this, GSEs selectively check loans that go into default, and if they discover that the lender's representation and warrants were violated, they can force the lender to purchase the loan back at par. GSEs also check a random sample of non-defaulted loans, and can force repurchase of all loans with any rep and warrant violations. GSEs keep track of the repurchase record of originators and impose higher fees and capital requirements on originators with high repurchase rates. These measures make it costly for lenders to have securitized loans go into default, and may induce them to be more conservative with loans that they sell into the secondary market than loans that they retain in their portfolios.

In this paper, we offer a more compelling explanation for our default risk results that defy the conventional wisdom. This explanation comes from our analysis of prepayment risk. We find that loans sold to GSEs and private issuers have higher prepayment risk than portfolio loans. It appears that in return for selling loans with lower or comparable default risks, lenders retain loans with lower prepayment risk. During the years of high refinancing and low default, retaining loans with lower prepayment rates was a much more profitable strategy than retaining loans with lower default rates. We find, for instance, that as we move from 2004 to 2008, and as default concerns in the market started growing, this trade-off changed. Lenders became less willing to retain higher default risk in return for lower prepayment risk, and there was less of a difference in the default probabilities of securitized and portfolio loans. Furthermore, trading low default risk for low prepayment risk helps originators maintain their reputations, minimize the probability and the cost of being required to repurchase loans, avoid being flagged as a high-repurchase originator, and satisfy the high underwriting standards of the secondary market, as these concerns all pertain to default risk, but not to prepayment risk.

Our explanation is supported by another result of our study. We find that, compared to large lenders, small lenders place more emphasis on default risk than prepayment risk. That is, they are not as likely to retain higher-default risk loans. In fact, unlike large lenders, small lenders sold *higher*-default risk loans into the secondary market in the last two years of our sample period. We attribute this to the fact that small lenders have less to lose from risking their reputations in the secondary market,[5] and they have less ability to diversify their default risk, as they have a smaller and geographically more concentrated portfolio of loans.

In 2006, the loans small lenders sold into the secondary market had both higher default risk and higher prepayment risk. This is in contrast with loans sold by large lenders, which had higher prepayment risk but lower default risk in each year. We interpret this difference between small and large lenders as evidence of asymmetric information in the market. While large banks handle more loan originations, and often evaluate them at their regional underwriting centers,

---

[5] During our sample period, small lenders not only sold fewer loans into the secondary market, but they also sold a smaller portion of their loan originations.

small banks are more likely to be small town or neighborhood banks. Hence, they are more likely to know the borrower personally and to have private (soft) information about the borrower. The asymmetric information that small lenders enjoy enables them to identify the lemons, loans that are riskier on both default and prepayment fronts, and sell them into the secondary market.

There is a small but growing literature on the role of securitization in loan markets. Of particular relevance to us are three studies that address adverse selection problems in mortgage loan securitization. Elul (2009) finds that securitized prime loans have higher default rates than portfolio loans, but securitized subprime loans do not perform worse than portfolio loans. Keys, Mukherjee, Seru and Vig (2009) compare the performance of subprime loans and find that mortgage lenders will apply weaker screening standards for loans that they are likely to sell in the secondary market. Ambrose, Lacour-Little and Sanders (2005), however, report that securitized loans perform better than loans retained in bank's portfolio.

We use the loan-level dataset from LPS Analytics that includes around two-thirds of the mortgage market in the United States. Our results regarding default risk support the results of Ambrose, et al. (2005) that banks sell low-risk loans into the secondary market and retain higher-risk ones in their portfolios. One limitation of the Ambrose et al. (2005) study is that the loans in their sample were originated by a single lender. Our sample includes loans originated by more than 4,500 lenders.

Our sample includes prime and subprime loans. The results with respect to default risk and prepayment risk hold for subprime as well as prime loans, but there is less of a difference in default and prepayment risks of portfolio and securitized loans for subprime loans. In fact, the difference for subprime loans becomes insignificant in 2006 and 2007, the height of the real estate bubble. The difference between subprime and prime loans is likely due to the fact that subprime loans are subject to more scrutiny by investors than prime loans.

Comparison of the performance of loans sold to GSEs and those sold to private issuers shows that loans sold to private issuers have lower prepayment rates than loans sold to GSEs for each of the four years studied. The comparisons with respect to default risk show variations across years.

To analyze lenders' choices of loans to securitize, we need to estimate the expected default and prepayment probabilities of the loans from the lender's perspective at the time of origination. For this purpose, we conduct estimations for lenders' expectations using two types of expectation models. In the first case, we base lenders' estimates on a rational expectations model of default and prepayment risk. In the second case, we use an adaptive expectations model where lenders' estimates are based on observed default and prepayment probabilities in the previous two years. Our results prove to be robust to the choice of the expectations model.

We review the literature in the next section. Section 3 discusses the data. Section 4 presents the methodology and discusses the results. Section 5 offers concluding remarks.

2. Literature Review

Following Akerlof's seminal work (1970), a large body of literature has addressed asymmetric information and adverse selection issues in various fields. The applications in financial markets, however, have been largely theoretical. Examples include Leland and Pyle (1977), Stiglitz and Weiss (1981), Myers and Majluf (1984), Dunn and Spatt (1985), John and Williams (1985), Chari and Jagannathan (1989), Brueckner (1994), Carlstrom and Samolyk (1995), Posey and Yavas (2001), and DeMarzo (2005). A number of other papers provide theoretical explanations for creating asset-backed securities under alternative information structures (e.g., Gorton and Pennachi, 1995; Glaeser and Kallal, 1997; Riddiough, 1997; and DeMarzo, 2005).

There is also a small but growing empirical literature on adverse selection in mortgage financing. The dramatic expansion of secondary loan markets and the current financial crisis have led to a number of recent studies on the topic. A recent study by An, Deng and Gabriel (2009) tests for adverse selection problems in the market for commercial mortgage loans. They compare portfolio lenders that sell some loans into the secondary market and keep some in their portfolios and conduit lenders that sell all their loans into the secondary markets. Since conduit loans impose no adverse selection problem for purchasers in the secondary market, these loans should be priced higher than portfolio loans. This prediction is supported by their empirical findings. In another recent study, Downing, Jaffee and Wallace (2009) show that Freddie Mac utilizes its private information and sells more lower-credit quality residential mortgage-backed securities to bankruptcy remote special purpose securitization vehicles than the securities it retains in its portfolio.

Recent studies of residential mortgages have focused on the impact of securitization on the quality of loan screening and servicing. Keys, Mukherjee, Seru and Vig (2009) argue that mortgage lenders will apply weaker screening standards for loans that they are likely to sell in the secondary market. To test this argument, they compare the performance of subprime loans with credit scores just above and just below the cutoff point of the purchasers in the secondary market (i.e., loans with credit scores just above and just below 620). Lenders will apply weaker screening standards to applicants with credit scores just above the cutoff point because they know they can sell these loans in the secondary market. Similarly, lenders will screen borrowers with credit scores just below the cutoff point more carefully, as these loans are harder to securitize, and the lenders are more likely to keep these loans in their portfolios. As a result, one would expect loans with credit scores just below the cutoff point to perform better than loans with credit scores just above the cutoff point. They find that loans with credit scores just below the cutoff point are indeed less likely to default.

Bubb and Kaufman (2009) offer an alternative theory to the finding in Keys, et. al. (2009). They claim lenders collect more information about borrowers with just below the cutoff credit score because the benefits to lenders of collecting additional information are higher for higher-default risk borrowers. They test their theory against the securitization-based argument of Keys, et al. and argue that their theory is more consistent with the empirical evidence.

The objective in Piskorski, Seru, and Vig (2009) is to study the moral hazard problem created by securitization. The authors examine the impact of securitization on loan servicing and whether securitization inhibits modifications of loans for distressed borrowers. Studying the loans that were seriously delinquent, they find significantly lower foreclosure rates for portfolio loans than for securitized loans. They conclude that, relative to servicers of securitized loans, servicers of portfolio loans undertake actions that lead to more renegotiations and lower foreclosures. Adelino, Gerardi and Willen (2009) also study the impact of securitization on the servicing of seriously delinquent loans. They challenge the conclusions of Piskorski, Seru, and Vig (2009) and offer evidence to show that servicers renegotiate similar and low fractions of loans in their portfolio and securitized loans. They argue that the reluctance of servicers to renegotiate loans is not attributable to a moral hazard problem created by securitization, but rather to the fact that a delinquent borrower is likely to default again despite costly renegotiation, and that about one-third of seriously delinquent borrowers cure without receiving any modification.

Rajan, Seru and Vig (2008) show that as there is a higher degree of securitization in the market, mortgage rates rely increasingly on hard information about borrowers. They offer a Lucas critique on the use of statistical default models, warning researchers that models fitted in a low-securitization period may break down in a high-securitization period.

Mian and Sufi (2009) argue that the increase in securitization of subprime mortgages is closely correlated with the expansion in subprime mortgage credit and its divergence from the growth in income. It is interesting to note that their findings hold even in markets with very elastic housing supply that experienced relatively low house price growth during the credit expansion years.

Berndt and Gupta (2009) have a different focus. They examine how securitization impacts the long-run performance of the borrowing firm. Using data from the secondary market for syndicated loans, they show that borrowing firms whose loans are sold in the secondary market underperform other borrowers. The effect is more severe for small and high leverage firms.

The research most closely related to ours includes Elul (2009) and Ambrose, Lacour-Little and Sanders (2005). Using a loan-level dataset from LPS Analytics, Elul (2009) concludes that securitized prime loans have higher default rates than portfolio loans, and the relative performance of securitized loans worsens as the origination year moves from 2003 to 2007. However, securitized subprime loans do not perform worse than portfolio loans. Like Elul (2009), we also use the LPS Analytics dataset, but our findings for prime loans are contrary to his findings. We find that default rates of securitized loans are lower than portfolio loans for each of the years 2004 through 2008, though the relative performance of securitized loans deteriorates as the origination year moves from 2004 to 2008. The reason for the difference in our results is likely due to differences in methodologies. Unlike Elul (2009), we construct a model of lenders' expectations of default and prepayment probabilities for the loans they originate, and use these expectations to estimate lenders' choice of whether to securitize a loan. We also control for the spread that the loan enjoyed over the ten-year treasury rate and the pricing of risk in the market at the time, as the pricing of the loan's risk as well as the riskiness of the loan are likely to impact the lender's securitization decision.

Our methodology and some of our results are similar to those of Ambrose, LaCour-Little and Sanders (2005). Using data from a single lender, they find that its securitized loans performed better than the loans retained. The authors attribute their result to two factors: reputation concerns, and regulatory capital requirements. Our study differs from Ambrose et al. (2005) in three respects. First, our data set involves more than 4500 lenders, and our results point to significant differences across them. In particular, we find significant differences between securitization strategies of large versus small lenders. Second, we differentiate between GSEs and private labels as loan purchasers when we study the lender's decision to sell or retain that loan. Third, we examine the securitization decision under an adaptive expectations as well as a rational expectations model of default and prepayment risk.

A somewhat related line of literature examines the role of GSEs. Given the differences in interest rates of conforming loans (loans that qualify for a purchase by GSEs)[6] and non-conforming loans, it is widely accepted that GSEs reduce interest rates by expanding funds available to lenders. There is, however, disagreement on the extent to which GSEs reduce

[6] Conforming mortgages have a maximum loan size that varies over the years and can differ by geographic areas (for 2009 the limit is $417,000 for most counties) and have to satisfy certain borrower quality and loan characteristics, such as credit score, payment-to-income ratio, and loan-to-value ratio. The exact set of combinations of borrower and loan characteristics for conforming loans is not known. Originators use Freddie Mac's Loan Prospector and Fannie Mae's Desktop Underwriting programs to determine if a loan is conforming.

mortgage interest rates. Ambrose, Buttimer and Thibodeau (2001) argue, for instance, that a significant portion of the mortgage rate differential between conforming and non-conforming loans can be explained by the higher house price volatility associated with non-conforming loans. Ambrose, LaCour-Little and Sanders (2004) report that the interest rate differential also narrows if one corrects for endogeneity and sample selection bias problems in the data. Nothaft, Pearce and Sevanovic (2002) consider the role of the funding advantage that GSEs enjoy, as they benefit from federal guaranty and exemption from certain taxes. They find that about 27 to 30 basis points of the mortgage rate difference can be attributed to the funding advantage of GSEs. Our results contribute to this discussion by showing that part of the interest rate differential can be explained by the fact that loans purchased by GSEs have different prepayment rates as well as different default rates than from portfolio loans and private issue loans.

3. Data

Our data is provided by LPS Applied Analytics, Inc., and includes loan-level information collected from residential mortgage servicers.[7] As of July 2008, the dataset included loans from nine of the top ten servicers, and represented around two-thirds of the mortgage market in the United States, or more than 39 million active mortgage loans.[8] As the information is collected from mortgage servicers rather than from investors, agency and non-agency mortgage-backed securities as well as portfolio loans are included in the dataset.

The LPS dataset provides extensive information about the loan, property and borrower characteristics at the time of origination as well as dynamically updated loan information subsequent to origination. Property-related variables include appraisal amount, geographic

[7] This dataset is generally known as the "McDash"dataset. McDash Analytics, the company that originally created the dataset was acquired by LPS in 2008.
[8] http://www.lpsvcs.com/NewsRoom/Pages/20080722.aspx

location, and property type (single-family residence, condo or other type of property). Loan characteristics available to us include origination amount, term to maturity, lien position, whether or not the loan is conventional, loan purpose (purchase or refinance), and lender-defined subprime flag, as well as coupon rate on the mortgage. Credit risk-related variables include debt-to-income ratio, FICO score, loan-to-value (LTV) ratio of the borrower at origination and level of documentation provided.

Beyond the data that are available at origination, dynamically updated variables capture changes made to the loans since origination as well as their performance at a monthly frequency. Variables of interest include coupon rates (which change for ARMs and have the potential to change for loan modifications), delinquency status (current, 31-60 days delinquent, 61-90 days delinquent, over 91 days delinquent, foreclosure, REO, or paid off), investor type (held in portfolio, private securitization, GNMA, FNMA, FHLMC, GNMA buyout loans, Local Housing Authority, or Federal Home Loan Bank), current FICO score, and principal balance, as well as scheduled principal balance if the borrower pays according to the original terms of the loan. Most critical to this research, the investor type variable tracks securitization decisions regarding the loan made over time, and the delinquency variable provides information on the loan's default and prepayment events.

We also have access to variables through HMDA (Home Mortgage Disclosure Act) data. The merging of the LPS dataset with HMDA data gives us access to additional information on the borrower and the lender. These data include socioeconomic and demographic information on the borrower such as borrower income. We are also able to use the HMDA data to control for lender differences (e.g., the number of loans originated by a lender in a given year).

We focus on conventional, fixed-rate mortgages for single-family residences and condos originated between January 2004 and June 2008. Second mortgages, HELOCs, and loans above $650,000 are excluded. Although we allow both prime and subprime loans to enter the dataset, we impose additional restrictions on the prime loans. We confine the analysis to prime loans with FICO scores above 620 and loan-to-value ratios below 95 percent. These criteria enable us to reduce selectivity bias by restricting the sample to only those loans that are qualified for securitization through both GSE and private channel. None of the subprime loans in our dataset were purchased by the GSEs, so additional constraints are not required for this group of loans.

Over time, the McDash dataset has grown dramatically with the addition of new reporting servicers. The addition of these servicers to the dataset means that both seasoned loans and new originations are included, but only information available after servicers sign on to the dataset is included in the dataset. This could potentially left-censor the data because earlier loans that have defaulted or prepaid prior to the servicer beginning reporting will not be included, while loans that have remained current will. To reduce the extent of left-censoring in the data, we eliminated loans that entered McDash a year or more after origination.

We then categorize the loans as being held in portfolio, sold to the GSEs, or sold to private label securities. Loans held by the Local Housing Authority and the Federal Home Loan Bank are grouped with the portfolio loans, as the institutions that originated the loans still bear the credit risk for these loans. GSE loans include Government National Mortgage Association (Ginnie Mae), Federal National Mortgage Association (Fannie Mae), Federal Home Loan Mortgage Corporation (Freddie Mac), and Government National Mortgage Association buyout loans. We define investor type as the most common type of investor within the 12-month period after origination. In our data sample, 78.3% of the loans with an observable investor type are classified as GSE loans, 7.4% as held in portfolio, and 14.3% as privately securitized.

We define a loan in default if it is over 61 days delinquent, or foreclosed, or has experienced an REO sale.[9] A loan is considered prepaid if it has been paid in full in a month when the scheduled principal balance amount is greater than $500 and the prepayment is not preceded by delinquency events.[10] We also create a dummy variable to indicate whether a loan conformed to GSE standards at the time of origination. Since we do not have an official GSE credit standard, we follow the definition proposed by Ambrose, LaCour-Little and Sanders (2005). We label a loan as conforming if it was held by one of the agencies above at some point during the 12 months after origination. If the loan was not held by a GSE, we label it as conforming if the FICO score was higher than 660, the origination amount was below the conforming limit for that geographic area and time, and the loan has private mortgage insurance if the LTV ratio is above 80. Overall, 92.2% of the loans in our sample are identified as conforming according to this definition.

Table 1a provides the descriptive statistics for the prime loan sample, broken out by origination year and investor type. The average FICO score for loans sold to the GSEs shows a slight decline over the years, while the LTV ratio increases every year. Loans considered low- or no-documentation as a share of all loans sold to the GSEs also increased from 18% in 2004 to 28% in 2007. The quality of loans held in portfolio, however, appears to improve slightly over time. Although the LTV ratio hovers at around 70% for every year in the sample, the average

[9] According to the Office of Thrift Supervision (OTS), a loan is in delinquency if a monthly payment is not received by the loan's due date. This is a slightly less strict definition of delinquency than the Mortgage Bankers Association's definition. An REO (Real Estate Owned) sale follows an unsuccessful foreclosure when a buyer for the property cannot be found, and the mortgage lender repossesses the property to sell separately.
[10] The minimum principal balance of $500 is used to differentiate a prepayment from a scheduled final month's payment of a loan.

FICO score shows a slight increase, and the proportion of conforming loans increases over time while the share of low/no documentation and jumbo loans decreases over time. The credit quality of the loans that are privately securitized remains more or less constant over time. The shares of jumbo loans, conforming loans and low/no-documentation loans hardly change over the sample period. Overall there is also a trend of lenders retaining a larger share of the prime loans that they originate. While in 2004, only 3.6% of loans originated were held in portfolio, in 2007 the rate increased to 16.6%.

Table 1b reports the summary statistics for the subprime loans. None of the subprime loans in our dataset were purchased by the GSEs, and the great majority of them were privately securitized. As expected, the average FICO score for the subprime loans is about 100 points fewer than the average FICO score for the prime loans, and the average LTV ratio is about 10 percentage points higher for the subprime loans. Compared to the subprime loans held in portfolio, the subprime loans that are privately securitized tend to have slightly higher average FICO scores for every year in the sample and the origination UPB (unpaid principal balance) tend to be lower. For every other variable of interest, subprime loans that are kept in portfolio seem comparable to the subprime loans that are privately securitized. Contrary to the prime market, 2006 saw the largest portion of subprime fixed-rate loans securitized: when in 2005 lenders kept roughly 30% of the subprime loans originated in portfolio, in 2006 the proportion drastically dropped to only 3%.

4. Methodology and Results

To determine the relationship between lenders’ securitization choice and expected loan performance, our approach consists of four steps. For prime and subprime loans and for each year of origination, we divide the sample population into a random 75% estimation sample and a 25% holdout sample. First, based on the 75% estimation sample, we construct a hazard model using the observed default and prepayment outcomes in the next 24 months. In the second step, we apply the coefficients obtained from the first step to the holdout sample consisting of the other 25% of the population, and calculate their expected default and prepayment probabilities. In the third step, we further account for the degree of over- or under-pricing for the loans in the holdout sample. In the last step, we regress the observed securitization outcome on the loans in the holdout sample on their expected default probabilities, prepayment probabilities, over-pricing and under-pricing indicators obtained from the previous two steps, and other variables controlling for the market environment at the time of origination. We then estimate the relationship between a loan's expected performance and the lender's securitization choice.

To account for variations in lenders' formation of expectations regarding loan performance, we use two alternative ways to apply the estimation parameters. First, under the rational expectations approach, we assume that the lender has perfect foresight regarding the contribution of loan characteristics to the outcome probabilities, and the expectations for loans in the holdout sample are formed the same way as those in the estimation sample. This way, we use parameters estimated from loans originated in the same year to apply to the holdout sample. For example, to form expected prepayment and default probabilities in the next 24 months for loans originated in 2006, we use parameters estimated from a different sample, also originated in 2006, observed through 2008.

Our second approach to modeling lenders' expectations is the adaptive expectations approach; lenders form their expectations based on their experiences up to the time of loan origination. In other words, they draw conclusions for the 2006 loans by learning from the performance of loans originated before 2006. For this case, we lag the estimation sample by two years compared to the holdout sample, so that parameters estimated will be from a 24-month period before origination of the new loan, For prime loans, we use parameters estimated from the 2004, 2005 and 2006 full populations to the holdout sample for 2006, 2007 and 2008, respectively. For subprime loans, as only few were originated in 2008, we estimate the expected probabilities for 2006 and 2007 loans using parameters from 2004 and 2005. This approach not only allows us an ex ante view of lender expectations but also gives us a view of the securitization strategy in the post-boom period, when it is still too early to draw conclusions from actual loan performance.

This approach is similar to that of Ambrose et al. (2005) in the sense that the expected performance, rather than the realized performance (as in Elul, 2009), is used to explain the securitization decision. This approach controls for potential performance differences due to factors post-securitization, such as the moral hazard issues pointed out in Piskorski, Seru and Vig (2009), and approximates lenders' ex ante information at the time a securitization decision is made.

Default and Prepayment Estimation

We model the loans' default and prepayment probabilities in a competing risk hazard framework. At each point in time, the borrower may decide to terminate the loan by refinancing or moving and prepaying the balance owed, or to default and put the house back to the lender. If neither of these events occurs at that point, the loan survives for another period, and the observation is considered censored. We implement the hazard model as a multinomial logit model with quadratic baseline function for timing of an event.

We control for borrower and mortgage characteristics in the default and prepayment estimation. To ensure consistency between estimation and forecast, the explanatory variables are taken as of the time of origination. These variables include: borrower credit score (FICO), borrower income (Income), origination loan-to-value ratio (LTV ratio), whether the loan was considered conforming by GSE standards (Conform), whether the loan amount was above conforming loan limits (Jumbo), and loan underwriting documentation level (Low Documentation). We also include time since origination and its squared term to control for loan age effects.

The estimation coefficients are reported in Tables 2a and 2b for prime and subprime loans. Borrower credit score (FICO) is negatively correlated with default probability for both the prime and subprime sector and nearly across all origination years 2004 through 2007, indicating that borrowers with better credit standing are less likely to default. The only exception is subprime loans originated in 2006; mid-year 2006 is considered the starting point of the housing downturn. This may indicate that for these loans, the overriding default factor is the house price collapse, rather than borrower credit quality. We also find a reduced probability to prepayment for borrowers with higher FICO scores, except for subprime loans originated in 2004 and 2005, where the effect is insignificant.

Higher-income borrowers tend to have reduced probability of default for loans originated in 2004 through 2006, and an increased probability of default for loans originated in 2007. This may be due to the fact that particular geographic areas with the most severe delinquencies, namely, California, Florida, Nevada and Arizona, were also areas where housing prices had risen the most, requiring higher income to qualify. Higher income is also associated with higher prepayment propensity among prime borrowers for most of the years, while among subprime borrowers the income effect on prepayment is insignificant.

Higher loan-to-value (LTV) ratio measured as of origination is found to contribute positively to borrower default probabilities for both prime and subprime loans in all years. A high LTV increases the probability that the house will go underwater, adding to borrower incentive to default. A high LTV also restricts the borrower's ability to refinance in cases of financial distress, as we estimate that higher LTV loans are less likely to prepay in nearly all cases.

Loans with reduced documentation require less paperwork in the underwriting process, and are generally issued to borrowers with variable or unverifiable income, such as borrowers who are self-employed or citizens of another country. LaCour-Little and Yang (2009) find that such loans issued during the most recent housing boom have higher propensity of default. Our results confirm that low- or no-documentation loans have a higher default probability, and a stronger effect is observed in the subprime sector. These loans are also more likely to prepay.

Cumulative Default and Prepayment Probability

We use the estimated coefficients to calculate the expected cumulative 12-month and 24-month prepayment and default probabilities for each loan in the holdout sample. The general form of the calculation of probability of outcome for loan *i* for each of the three outcomes at each point in time is:

$$P_i(outcome = j) = \left(\frac{\exp(\alpha_j + x'_{1,2,..,n}\beta_j)}{1 + \sum_{j=1}^{2} \exp(\alpha_j + x'_{1,2,..,n}\beta_j)} \right)$$

where j=1,2 (default, prepay), and to ensure that probabilities sum up to 1:

$$P_i(outcome = survive) = \left(\frac{1}{1 + \sum_{j=1}^{2} \exp(\alpha_j + x'_{1,2,..,n}\beta_j)} \right)$$

where α is the constant, X is the value of the independent variables, and β is the vector of coefficient estimates.

Tables 3a and 3b present the cumulative expected probabilities calculated assuming lenders have rational expectations. Compared to loans held in portfolio, loans sold to GSEs consistently have lower expected default probabilities. Loans privately securitized have the lowest default probabilities of the three types in 2004 and 2005, but then the highest default probabilities in 2006 and 2007. Loans sold to GSEs have the highest prepayment probabilities in all years, while loans held in portfolio and privately securitized have lower prepayment probabilities. For the subprime sector, privately securitized loans have lower expected default probabilities in all years.

Expected cumulative default and prepayment probabilities using the adaptive expectations approach are reported in Tables 3c and 3d. Under this approach, the private label security loans have the lowest expected default probabilities in 2006 and 2007. Loans sold to the GSEs have lower expected default rates than those held in portfolio in 2006, but higher rates in

2007, although the differences are minor in both cases. In 2008 loans sold to the GSEs have the lowest default probabilities, and the portfolio loans have significantly higher expected default probability than those sold for securitization. This may be a result of tightened securitization policies during the housing downturn, and lenders are forced to keep the worse performing loans in their own portfolios. The GSEs also have the highest expected prepayment probabilities in all cases. In the subprime sector, privately securitized loans show lower expected default probabilities than portfolio loans in both 2006 and 2007, although the sample size is small.

Yield Spread

Besides termination risks, we need to take into consideration the yield on individual loans in determining lenders' decisions to securitize. A lender may choose to keep a loan in its portfolio if the loan is over-priced relative to its risk. We measure a loan's pricing by the difference between a loan's coupon rate at origination and the contemporaneous yield on the ten-year Treasury bond. This yield spread is then regressed on the mortgage risk-related characteristics and market condition variables for the holdout sample. The residual between the actual yield and the predicted yield from the model gives an indication of the pricing of the mortgage. We assign a value of 1 to this indicator of over-pricing (high_spd=1) if the residual is positive and in the top quartile of sample distribution, and a value of 1 to the indicator of under-pricing (low_spd=1) if the residual is negative and in the bottom quartile of distribution.

Explanatory variables in the yield spread regression include loan characteristics that directly affect a loan's pricing such as whether the loan conforms to GSE standards (conform), whether it is above the conforming loan limits (jumbo), its loan-to-value ratio (ltv-ratio), logarithm of credit score of the borrower (log_FICO), and logarithm of the origination amount of the loan (log_UPB). We also consider market risk pricing indicators at the time of origination such as local house price volatility (log_sigma_hpi), credit spread premium (log_credit_spd), interest rate volatility (log_sigma_int), and shape of the yield curve (log_yield_curve). Differences between geographic locations are factored in through creation of Census region indicators. The steepness of the yield curve is measured by the ratio of the ten-year Treasury bond rate to the one-year Treasury note rate. The credit spread premium is defined as the difference between the 'AAA' bond index and the 'Baa' bond index. We calculate housing price volatility as the standard deviation in the OFHEO purchase-only non-seasonally adjusted house price index in the eight quarters prior to origination at the state level. We proxy for interest rate volatility by the standard deviation in the one-year Treasury bond rate 15 months before the origination of a loan.

The yield spread model parameters reported in Appendix 1 largely conform to our expectations. Higher yields are associated with loans with higher loan-to-value ratios and borrowers with lower credit scores. There is a positive correlation between a loan's yield premium and market credit spread premium, except for subprime loans in 2004. There is an interesting contrast between prime and subprime loans in their pricing of house price volatility. For every year in the prime sample, loans in areas with higher house price volatility are asked to pay higher yield premiums. Yet the opposite is true in the subprime sample where higher house price volatility areas are paying reduced yield premiums. This may be evidence of differences in practice between the two sectors, considering that for the years in the sample period, areas with higher volatilities generally are the high growth areas. While the prime market demands higher return to compensate for the higher risk associated with volatility, subprime market reduces their prices in these areas.

Decision to Securitize

In the last step, we use the observed securitization choice from the holdout sample to model the adverse selection. Our base model includes the expected default probability, expected prepayment probability, and the over- and under-pricing indicators calculated in the previous steps. In addition, we control for the mortgage's yield spread, whether the loan is above the conforming loan limit, the market credit spread premium, the shape of the yield curve and interest rate volatility. This is estimated in a multinomial logit equation that takes the general form as follows:

$$\log\left(\frac{\pi^j}{\pi^r}\right) = \alpha_j + x'_{1,2,\ldots,n}\,\beta_j$$

In our estimation π^j (j=1,2) represents the probability of the two outcomes of interest: sold to GSEs, and securitized through private label securities; and π^r is the residual probability of being kept in portfolio. The model reduces to a logit model for the subprime sector when there are only two outcomes, sold to private label securities or kept in portfolio. For primary loans, this equation is run with lender fixed effects where loans with the same holding company are given the same identifier.[11] To test for a small lender effect, we combine all lenders that issued fewer than 20 loans in that particular year and assign them one identifier. We also simulate both

[11] Loans from originating lenders with the same holding company (e.g., Wells Fargo California and Wells Fargo Washington) are grouped together.

models of expectations, where lenders are assumed to have rational expectations and where lenders can only learn from the past. Under the rational expectations assumption, the cumulative default and prepayment are estimated using parameters from observed performance on loans originated in the same year. The results are summarized in Tables 4a and 4b. Under the adaptive expectations assumption, the cumulative default and prepayment probabilities are estimated using parameters from loans originated two years prior. These results are presented in Tables 4c and 4d.

Our results show that lenders are less likely to sell a loan to a GSE or private label if the loan is expected to have a higher default probability. The coefficients on the expected cumulative default probability (cumdefault_24) are negative and significant for each year. The relationship between default probability and probability of securitization is weakens through the years examined. Under adaptive expectations, the effect is stronger for the GSEs than for private label securities; under rational expectations, the comparison between GSEs and private label securities varies by year. Lenders are also more likely to securitize loans with higher expected prepayment probabilities, and more so with GSEs than private label securities. This is seen in the positive and significant coefficients on the expected cumulative prepayment probability (cumprepay_24) variable. In the subprime sector, lenders are less likely to securitize higher-default risk loans in 2004 and 2005, but are indifferent to the default probabilities after that. As Elul (2009) points out, the difference is likely due to the fact that subprime loans are subject to greater investor scrutiny than prime loans. Expectations about prepayment probabilities are not a significant factor in securitization of subprime loans under either expectations model in any of the years except for 2004.

To examine the adverse selection behavior of different types of lenders further, we segment the data by lender size, proxied by their origination volume in a given year. Those originating 20 or fewer loans are grouped as small lenders, and the analysis is repeated on each group's securitization choice. Lender fixed effects are included in the large lender group. We do not include lender fixed effects in the small lender group as each lender contributes 20 loans or fewer to the sample, and lender-specific variations are not considered significant enough to bias estimation results. The results under rational expectations are shown in Appendix 2.[12] In Table A2a, the coefficients on the default probability variable are positive and significant for securitization with GSEs in years 2006 and 2007 and for private label securities in 2006. This is contrary to the results for the large lender group, presented in Table A2b, which are similar to results obtained from the overall sample reported in Table 4a. We attribute the difference to two factors. One is that smaller lenders are more likely to be small town lenders and have private information about the borrower to adversely select loans for securitization. The other is that small lenders are less likely to retain higher-default risk loans even if the loans come with lower prepayment probability because of their less diversified borrower base and their smaller cost of having damaged reputation in the secondary market. In 2006, loans sold by small lenders to GSEs have both higher default risk and higher prepayment risk than the loans held in portfolio. The loans sold to private labels have higher default risk but not significantly different prepayment risk from loans sold in portfolio. In 2007, prepayment probability does not play a significant role in small lenders' securitization decisions.

As a robustness check, we also model the ex post realized performance of the holdout sample on the securitization choice variables and other loan characteristics variables. In addition to adverse selection, this approach captures any moral hazard issues related to servicing of the loan

[12] The results under adaptive expectations are qualitatively similar.

after origination. Complete results are available upon request. We find that the GSE-securitized indicator is positively associated with prepayment probability but negatively associated with default probability, consistent with our ex ante findings. Loans securitized through private channels are more likely to default than GSE loans or loans held in lender portfolios. Consistent with the findings reported so far, loans sold to GSEs and private labels are more likely to prepay than those held in portfolio.

5. Conclusion

Are loans sold into the secondary mortgage market riskier than the loans that lenders retain in their portfolios? Analysis of a large dataset of mortgage loans originated between 2004 and 2008 reveals strong evidence that banks sold low-default risk loans into the secondary market and retained higher-default risk loans in their portfolios. This result holds for prime as well as subprime loans, although the difference is smaller for subprime loans. We also compare the performance of securitized loans and portfolio loans with respect to prepayment risk. We find support for adverse selection with respect to prepayment risk; securitized loans entailed a higher prepayment risk than portfolio loans. It appears that in return for selling loans with lower default risk, lenders retained loans with lower prepayment risk. This would be a profitable strategy in the early years of our sample period when prepayment risk driven by high refinancing activity was a bigger concern for lenders than default risk. As an additional support for the rationality of lenders' strategies, we find that as the bursting of the bubble approached and default concerns in the market started growing, lenders became less willing to retain higher default risk in return for lower prepayment risk.

Our results also identify differences in the performance of the loans purchased by GSEs and by private issuers. Loans sold to private issuers have lower prepayment rates, while relative default rates show variations across the years.

In addition to a trade-off between default risk and prepayment risk in lenders' securitization strategies, we also find evidence of asymmetric information between originators and purchasers of loans in the secondary mortgage market. In 2006, loans sold by small lenders into the secondary market had higher default risk and higher prepayment risk. This is in contrast with loans sold by large lenders, which had higher prepayment risk but lower default risk in each year. We interpret this difference as evidence of asymmetric information in the market. Small banks are more likely to be small town or neighborhood banks that are more likely to know borrowers personally and have some soft information about them. Large banks, on the other hand, handle a much larger number of originations, often evaluate them at the regional underwriting centers, and thus rely mostly on hard data about borrowers. Asymmetric information that small lenders enjoy enables them to identify the lemons, loans that have both higher default risk and higher prepayment risk, and sell them into the secondary market.

It should be noted that observing securitized loans to have lower default risk than portfolio loans does not necessarily mean securitization did not play a role in the rising default probabilities that triggered the current financial crisis. Rather, securitization has led to a greater supply of funds for mortgage lending, which in turn might have contributed to deterioration in underwriting standards (Greenspan, 2010). Investigating the impact of securitization on default probabilities would require a study of the impact of securitization on overall underwriting standards in the industry.

References

Adelino, Manuel, Kristopher Gerardi, and Paul S. Willen (2009), "Why Don't Lenders Renegotiate More Home Mortgages? Redefaults, Self-Cures, and Securitization," Federal Reserve Bank of Boston Public Policy Paper 09-4.

Akerlof, G. A. (1970), "The Market for "Lemons": Quality Uncertainty and the Market Mechanism," *Quarterly Journal of Economics* 84 (3), 488–500.

Ambrose, Brent W., Richard J. Buttimer, and Thomas Thibodeau (2001), "A New Spin on the Jumbo/Conforming Loan Rate Differential," *Journal of Real Estate Finance and Economics*, 23: 309-335.

Ambrose, Brent W., Michael Lacour-Little, and Anthony Sanders (2004),"The Effect of Conforming Loan Status on Mortgage Yield Spreads: A Loan Level Analysis," *Real Estate Economics*, 32, 541-569.

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.

An, Xudong, Deng, Yongheng and Stuart A. Gabriel (2009), "Asymmetric information, adverse selection, and the pricing of CMBS," UCLA Working paper.

Berndt, Antje and Anurag Gupta (2009), "Moral Hazard and Adverse Selection in the Originate-to-Distribute Model of Bank Credit," *Journal of Monetary Economics* 56, 725-743.

Brueckner, J. K. (1994), "Borrower Mobility, Adverse Selection, and Mortgage Points," Journal of Financial Intermediation, 3, 416-441.

Bubb, Ryan and Alex Kaufman (2009), "Securitization and Moral Hazard: Evidence from a Lender Cutoff Rule," Federal Reserve Bank of Boston Public Policy Paper 09-5.

Carlstrom, Charles. T. and Ketherine A. Samolyk (1995), "Loan sales as a response to market-based capital Constraints," *Journal of Banking and Finance* 19, 627–646.

Chari, V. V., and R. Jagannathan (1989), "Adverse Selection in a Model of Real Estate Lending," *Journal of Finance*, 44, 499-508.

DeMarzo, Peter M. (2005), "The Pooling and Tranching of Securities: A Model of Informed Intermediation," *Review of Financial Studies*, 18: 1-35.

Downing, C., D. Jaffee, and N. Wallace (2009), "Is the Market for Mortgage-Backed Securities a Market for Lemons?" *Review of Financial Studies* 22(7), 2457-2494.

Duffee, G. R. and C. Zhou (2001), "Credit derivatives in banking: Useful tools for managing risk?" *Journal of Monetary Economics* 48, 25–54.

Dunn, K. B., and C. S. Spatt (1985), “Prepayment Penalties and the Due-On-Sale Clause,” *Journal of Finance*, 40, 293-308.

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

Glaeser, E. and H. Kallal (1997), “Thin Markets, Asymmetric Information, and Mortgage-Backed Securities,” *Journal of Financial Intermediation*, 6: 64-86.

Gorton, Gary B. and George G. Pennacchi (1995), “Banks and loan sales: Marketing nonmarketable assets,” *Journal of Monetary Economics*, 35, 389-411.

Greenspan, Alan (2010), “The Crisis,” paper presented at the April 2010 conference of the Brookings Papers on Economic Activity.

John, K., and J. Williams (1985), "Dividends, Dilution and Taxes: A Signaling Equilibrium," *Journal of Finance* 40(4), 1053-1070.

Keys, Benjamin, Tanmoy Mukherjee, Amit Seru, and Vikrant Vig (2009), “Did Securitization Lead to Lax Screening? Evidence from Subprime Loans,” forthcoming *Quarterly Journal of Economics*.

LaCour-Little, Michael, and Jing Yang (2009), “Taking the Lie out of Liar Loans,” working paper, presented at the FDIC symposium, *Improving Assessment of the Default Risk of Single-Family Mortgage*.

Leland, H. E., and D. H. Pyle (1977), "Informational Asymmetries, Financial Structure and Financial Intermediation," *Journal of Finance* 32(2), 371-387.

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

Myers, S. C., and N. S. Majluf (1984), "Corporate Financing and Investment Decisions when Firms Have Information that Investors Do Not Have," *Journal of Financial Economics* 13(2), 187-221.

Nothaft, Frank E., James E. Pearce and Stevan Stevanovic (2002), “Debt Spreads Between GSEs and Other Corporations,” *Journal of Real Estate Finance and Economics* 25(2/3), 151-172

Piskorski, Tomasz, Amit Seru and Vikrant Vig (2009), “Securitization and Distressed Loan Renegotiation: Evidence from the Subprime Mortgage Crisis,” forthcoming, *Journal of Financial Economics*

Posey, L. and A. Yavas (2001), "Adjustable and Fixed Rate Mortgages as a Screening Mechanism for Default Risk," *Journal of Urban Economics*, 49, 54-79.

Rajan, Uday; Seru, Amit and Vig, Vikrant (2008), “The Failure of Models That Predict Failure: Distance, Incentives and Defaults," *Working paper*, University of Chicago.

Riddiough, T. J. (1997), “Optimal Design and Governance of Asset-Backed Securities,” *Journal of Financial Intermediation,* 6: 121-152.

Stiglitz, J. E., and A. Weiss (1981), "Credit Rationing in Markets with Imperfect Information," *American Economic Review* 71(3), 393-410.

Table 1a. Descriptive Statistics for Prime Loans

	Kept in Portfolio				Sold to GSEs				Sold to Private Label			
	2004	**2005**	**2006**	**2007**	**2004**	**2005**	**2006**	**2007**	**2004**	**2005**	**2006**	**2007**
	n=13,031	n=22,007	n=44,947	n=51,128	n=276,807	n=411,422	n=391,674	n=226,341	n=69,024	n=92,850	n=51,031	n=30,770
FICO	728 (50.83)	731 (51.40)	731 (52.83)	734 (52.49)	731 (50.02)	727 (51.92)	726 (52.65)	724 (53.11)	733 (45.36)	728 (48.67)	720 (50.66)	726 (50.56)
Income (in $1,000s)	105.62 (238.29)	93.95 (138.77)	96.61 (160.75)	112.68 (145.61)	87.67 (107.71)	86.12 (83.26)	92.29 (94.82)	92.92 (88.17)	129.70 (159.93)	119.98 (131.16)	125.96 (157.07)	140.71 (179.09)
LTV ratio	69.34 (18.28)	69.99 (17.49)	69.76 (18.25)	70.19 (18.18)	65.16 (18.79)	67.28 (18.06)	68.46 (17.65)	69.94 (17.22)	68.92 (15.81)	70.21 (15.04)	70.49 (16.01)	69.88 (16.24)
Origination Amount ($)	208,112 (139,622)	207,230 (125,109)	199,959 (120,293)	219,857 (134,412)	159,229 (79,426)	174,503 (86,724)	184,475 (95,904)	190,115 (96,190)	235,151 (164.569)	253,475 (164,981)	263,432 (174,449)	289,785 (171,537)
Conform	.65 (.48)	.74 (.44)	.79 (.41)	.77 (.42)	1.00 (0)	1.00 (0)	1.00 (0)	1.00 (0)	.65 (.48)	.62 (.49)	.61 (.49)	.63 (.48)
Jumbo	.17 (.38)	.11 (.31)	.05 (.22)	.09 (.28)	.01 (.072)	.01 (.11)	.00 (.05)	.00 (.05)	.29 (.45)	.28 (.45)	.26 (.438)	.29 (.45)
Low/No Documentation	.13 (.33)	.07 (.254)	.06 (.24)	.05 (.21)	.18 (.38)	.25 (.43)	.30 (.46)	.28 (.45)	.21 (.41)	.22 (.41)	.26 (.438)	.24 (.43)
Coupon Rate	5.59 (.53)	5.68 (.50)	6.24 (.48)	6.19 (.48)	5.66 (.50)	5.83 (.43)	6.47 (.42)	6.31 (.42)	5.90 (.47)	6.00 (.43)	6.64 (.62)	6.42 (.48)

This table reports the descriptive statistics for single-family prime loans originated between January 2004 and June 2007. Means are reported. Standard deviations are in parentheses. Only conventional, fixed-rate mortgages are included. Second liens and loans above $650,000 are excluded.

Table 1b. Descriptive Statistics for Subprime Loans

	Kept in Portfolio				Sold to GSEs				Sold to Private Label			
	2004	**2005**	**2006**	**2007**	**2004**	**2005**	**2006**	**2007**	**2004**	**2005**	**2006**	**2007**
	n=430	n=2,991	n=302	n=483	n=0	n=0	n=0	n=0	n=3,841	n=7,354	n=9,902	n=1,804
FICO	607 (59.71)	616 (54.32)	623 (62.86)	611 (53.04)					641 (62.86)	620 (55.90)	625 (56.22)	619 (54.74)
Income (in $1,000s)	81.45 (77.05)	83.52 (139.50)	92.13 (93.96)	103.81 (207.26)					118.56 (240.89)	83.80 (129.42)	90.69 (178.69)	87.03 (126.38)
LTV ratio	78.86 (15.84)	80.35 (15.94)	84.11 (22.02)	77.38 (18.56)					77.64 (16.69)	80.33 (17.14)	78.99 (18.070)	77.24 (17.82)
Origination Amount ($)	175,399 (120,001)	186,206 (117,585)	223,425 (147,018)	215,912 (131,885)					154,679 (113,229)	165,339 (117,521)	190,715 (128,598)	202,955 (133,972)
Conform	.11 (.32)	.14 (.34)	.12 (.32)	.09 (.28)					.29 (.45)	.12 (.32)	.13 (.340)	.10 (.31)
Jumbo	.11 (.32)	.10 (.30)	.13 (.34)	.11 (.31)					.08 (.27)	.08 (.27)	.08 (.27)	.09 (.29)
Low/No Documentation	.21 (.41)	.24 (.43)	.24 (.43)	.16 (.37)					.32 (.47)	.28 (.45)	.21 (.41)	.17 (.37)
Coupon Rate	7.46 (1.25)	7.42 (1.09)	8.69 (1.56)	8.62 (1.50)					7.48 (1.20)	7.92 (1.19)	8.28 (1.40)	8.26 (1.36)

This table reports the descriptive statistics for single-family subprime loans originated between January 2004 and June 2007. Means are reported. Standard deviations are in parentheses. None of the subprime loans in our dataset were sold to the GSEs. Only conventional, fixed-rate mortgages are included. Second liens and loans above $650,000 are excluded.

Table 2a. Competing Risks Model of Mortgage Outcome for Prime Loans

	2004		2005		2006		2007	
Default Outcome	**Coefficient**	**P > |z|**	**Coefficient**	**P > |z|**	**Coefficient**	**P > |z|**	**Coefficient**	**P > |z|**
Constant	0.040	0.923	1.068**	0.001	-0.329	0.497	-0.615**	0.000
FICO	-0.018**	0.000	-0.018**	0.000	-0.016**	0.000	-0.015**	0.000
Income	-0.007**	0.000	-0.005**	0.000	-0.001**	0.000	0.000*	0.024
LTV ratio	0.034**	0.000	0.034**	0.000	0.034**	0.000	0.031**	0.000
Conform	-0.101	0.263	-0.207**	0.001	-0.330**	0.000	-0.342**	0.000
Time (in months)	0.342**	0.000	0.184**	0.000	0.252**	0.000	0.262**	0.000
$Time^2$	-0.009**	0.000	-0.004**	0.000	-0.005**	0.000	-0.004**	0.000
Jumbo	-0.213	0.217	-0.190	0.066	0.070	0.168	0.136**	0.007
Low Documentation	0.070	0.228	0.222**	0.000	0.125**	0.001	0.206**	0.000
Prepay Outcome	**Coefficient**	**P > |z|**	**Coefficient**	**P > |z|**	**Coefficient**	**P > |z|**	**Coefficient**	**P > |z|**
Constant	-4.996**	0.000	-3.236**	0.000	-5.447**	0.000	-7.865**	0.000
FICO	-0.005**	0.000	-0.004**	0.000	-0.001**	0.000	0.003**	0.000
Income	0.000	0.007	0.000**	0.000	0.000**	0.000	0.000**	0.000
LTV ratio	0.005**	0.000	-0.001**	0.000	-0.005**	0.000	-0.009**	0.000
Conform	0.205**	0.000	0.115**	0.000	0.388**	0.000	0.409**	0.000
Time (in months)	0.377**	0.000	0.171**	0.000	0.146**	0.000	0.083**	0.000
$Time^2$	-0.011**	0.000	-0.005**	0.000	-0.004**	0.000	0.000**	0.000
Jumbo	0.162**	0.000	-0.136**	0.000	0.257**	0.000	0.119*	0.014
Low Documentation	0.217**	0.000	0.055**	0.000	0.059**	0.000	0.051**	0.000
	Observations	6,110,843	Observations	8,883,551	Observations	8,106,087	Observations	5,179,805
	Pseudo R2	0.041	Pseudo R2	0.020	Pseudo R2	0.024	Pseudo R2	0.047

This table states the results from a competing risks model of the outcome to prepay, default or remain current on a given mortgage as estimated by a multinomial logit model. These results are estimated from a 75% sample of prime loans taken for each year of data. The dependent variable is whether a loan experienced default, prepayment or remained current within 24 months of origination. The independent variables are information available to lenders at the time of underwriting and include the borrower's FICO score (FICO), the borrower's income (Income), the loan-to-value ratio for the mortgage (LTV-ratio), whether or not the loan conforms to GSE standards (Conform), whether or not the loan amount exceeds GSE limits (Jumbo), and whether the loan application was low- or no-documentation (Low Documentation). ** significant at 1% level. * significant at 5% level.

Table 2b: Competing Risks Model of Mortgage Outcome for Subprime Loans

	2004		2005		2006		2007	
Default Outcome	**Coefficient**	**P > |z|**	**Coefficient**	**P > |z|**	**Coefficient**	**P > |z|**	**Coefficient**	**P > |z|**
Constant	-2.766**	0.000	-3.835**	0.000	-3.931**	0.000	-3.450**	0.000
FICO	-0.011**	0.000	-0.008**	0.000	-0.008	0.414	-0.009**	0.000
Income	-0.003**	0.004	0.000	0.229	0.000**	0.000	0.001*	0.020
LTV ratio	0.027**	0.000	0.021**	0.000	0.029**	0.000	0.028**	0.000
Conform	-0.705**	0.010	-0.690**	0.000	-0.628**	0.000	-0.389	0.128
Time (in months)	0.262**	0.000	0.279**	0.000	0.300**	0.000	0.257**	0.000
Time ^{2}	-0.007**	0.000	-0.008**	0.000	-0.009**	0.000	-0.007**	0.000
Jumbo	0.296	0.177	0.067	0.491	0.375**	0.000	0.258	0.074
Low Documentation	0.454**	0.001	0.479**	0.000	0.823**	0.000	0.678**	0.000
Prepay Outcome	**Coefficient**	**P > |z|**	**Coefficient**	**P > |z|**	**Coefficient**	**P > |z|**	**Coefficient**	**P > |z|**
Constant	-6.100**	0.000	-4.588**	0.000	-2.124**	0.000	-2.639*	0.034
FICO	-0.001	0.169	-0.001	0.241	-0.004**	0.000	-0.006**	0.002
Income	0.000	0.603	0.000	0.917	0.000	0.747	0.000	0.588
LTV ratio	-0.008**	0.001	-0.019**	0.000	-0.022**	0.000	-0.012*	0.013
Conform	-0.117	0.303	0.023	0.809	0.044	0.707	0.175	0.617
Time (in months)	0.436**	0.000	0.300**	0.000	0.250**	0.000	0.283**	0.000
Time ^{2}	-0.013**	0.000	-0.010**	0.000	-0.009**	0.000	-0.010**	0.000
Jumbo	-0.152	0.315	-0.240*	0.034	-0.064	0.656	-0.522	0.227
Low Documentation	0.104	0.200	0.179**	0.004	0.262**	0.001	-0.008	0.975
	Observations	68,291	Observations	158,897	Observations	152,185	Observations	36,147
	Pseudo R2	0.062	Pseudo R2	0.044	Pseudo R2	0.068	Pseudo R2	0.066

This table states the results from a competing risks model of the outcome to prepay, default or remain current on a given mortgage as estimated by a multinomial logit model. These results are estimation from a 75% sample of subprime loans taken for each year of data. The dependent variable is whether a loan experienced default, prepayment or remained current within 24 months of origination. The independent variables are information available to lenders at the time of underwriting and include the borrower's FICO score (FICO), the borrower's income (Income), the loan-to-value ratio for the mortgage (LTV-ratio), whether the loan conforms to GSE standards (Conform), whether or not the loan amount exceeds GSE limits (Jumbo), and whether the loan application was low- or no-documentation (Low Documentation). ** significant at 1% level. * significant at 5% level.

Table 3a. Predicted Cumulative Default and Prepayment Probabilities for Prime Loans under Rational Expectations

	2004			2005			2006			2007		
	Held in Portfolio	**Sold to GSEs**	**Sold as Private Label**	**Held in Portfolio**	**Sold to GSEs**	**Sold as Private Label**	**Held in Portfolio**	**Sold to GSEs**	**Sold as Private Label**	**Held in Portfolio**	**Sold to GSEs**	**Sold as Private Label**
	n=3,263	n=69,330	n=17,114	n=5,608	n=103,025	n=23,206	n=11,290	n=98,148	n=12,742	n=12,836	n= 56,662	n=7,611
Mean Cumulative Default Probabilities												
Month 12	0.26%	0.21%	0.17%	0.34%	0.31%	0.29%	0.51%	0.45%	0.56%	0.92%	0.89%	1.03%
Month 24	1.01%	0.81%	0.64%	1.17%	1.08%	0.99%	2.78%	2.44%	3.04%	5.53%	5.37%	6.20%
Mean Cumulative Prepayment Probabilities												
Month 12	4.11%	4.30%	4.17%	5.43%	5.83%	5.30%	6.42%	7.00%	6.48%	4.75%	4.93%	4.55%
Month 24	15.12%	15.79%	15.32%	14.33%	15.40%	14.01%	17.17%	18.71%	17.33%	15.75%	16.34%	15.08%

This Table reports the predicted cumulative default and prepayment probabilities for prime loans both 12 and 24 months after origination. The probabilities are calculated for each loan in the holdout sample using the coefficients estimated from the same year estimation sample as reported in Table 2a.

Table 3b. Predicted Cumulative Default and Prepayment Probabilities for Subprime Loans under Rational Expectations

	2004			2005			2006			2007		
	Held in Portfolio	**Sold to GSEs**	**Sold as Private Label**	**Held in Portfolio**	**Sold to GSEs**	**Sold as Private Label**	**Held in Portfolio**	**Sold to GSEs**	**Sold as Private Label**	**Held in Portfolio**	**Sold to GSEs**	**Sold as Private Label**
	n=96	n=0	n=903	n= 722	n=0	n=1,862	n=82	n= 0	n= 2,468	n=117	n=0	n=406
Mean Cumulative Default Probabilities												
Month 12	4.62%		3.61%	6.93%		6.58%	16.85%		11.22%	11.57%		10.94%
Month 24	16.66%		12.92%	22.77%		21.64%	52.32%		35.36%	41.31%		39.15%
Mean Cumulative Prepayment Probabilities												
Month 12	8.60%		8.30%	8.17%		8.41%	6.96%		7.68%	3.74%		3.61%
Month 24	30.31%		29.26%	21.03%		21.66%	15.05%		16.62%	9.38%		9.07%

This Table reports the predicted cumulative default and prepayment probabilities for subprime loans both 12 and 24 months after origination. The probabilities are calculated for each loan in the holdout sample using the coefficients estimated from the same year estimation sample as reported in Table 2b.

Table 3c. Predicted Cumulative Default and Prepayment Probabilities for Prime Loans under Adaptive Expectations

	2006			2007			2008		
	Held in Portfolio	**Sold to GSEs**	**Sold as Private Label**	**Held in Portfolio**	**Sold to GSEs**	**Sold as Private Label**	**Held in Portfolio**	**Sold to GSEs**	**Sold as Private Label**
	n=11,213	n=97,803	n=12,751	n=12,791	n=56,386	n= 7,702	n= 1,132	n=44,790	n=1,171
Mean Cumulative Default Probabilities									
Month 12	0.25%	0.23%	0.22%	0.34%	0.35%	0.29%	0.50%	0.38%	0.43%
Month 24	0.99%	0.91%	0.87%	1.15%	1.18%	0.98%	2.75%	2.07%	2.36%
Mean Cumulative Prepayment Probabilities									
Month 12	3.04%	3.44%	3.35%	5.23%	5.73%	5.21%	6.35%	6.89%	6.63%
Month 24	10.67%	12.05%	11.73%	13.75%	15.05%	13.68%	16.97%	18.40%	17.70%

This Table reports the predicted cumulative default and prepayment probabilities for prime loans both 12 and 24 months after origination. The probabilities are calculated for each loan in the holdout sample using the coefficients estimated from a sample with origination years that occurred two years earlier (these estimates are not reported but are very similar to those found in Table 2a).

Table 3d. Predicted Cumulative Default and Prepayment Probabilities for Subprime Loans under Adaptive Expectations

	2006			2007		
	Held in Portfolio	**Sold to GSEs**	**Sold as Private Label**	**Held in Portfolio**	**Sold to GSEs**	**Sold as Private Label**
	n=302	n=0	n=9902	n=483	n=0	n=1804
Mean Cumulative Default Probabilities						
Month 12	5.30%		4.06%	6.57%		5.95%
Month 24	18.85%		14.49%	21.60%		19.60%
Mean Cumulative Prepayment Probabilities						
Month 12	8.47%		8.81%	8.54%		8.56%
Month 24	29.61%		30.81%	22.75%		22.79%

This Table reports the predicted cumulative default and prepayment probabilities for subprime loans for both 12 and 24 months after origination. The probabilities are calculated for each loan in the holdout sample using the coefficients estimated from a sample with origination years that occurred two years earlier (these estimates are not reported but are very similar to those found in Table 2b).

Table 4a. Probability of Securitization for Prime Loans under Rational Expectations

	2004		2005		2006		2007	
GSE Outcome	**Coefficient**	**P > |z|**	**Coefficient**	**P > |z|**	**Coefficient**	**P > |z|**	**Coefficient**	**P > |z|**
Constant	2.931**	0.000	7.513**	0.000	11.754**	0.000	1.711	0.133
yield_spread	-0.255**	0.002	0.853**	0.000	1.270**	0.000	0.959**	0.000
Jumbo	-1.992**	0.000	-1.071**	0.000	-1.683**	0.000	-2.410**	0.000
credit_spread	-0.989*	0.024	-2.396**	0.000	-7.654**	0.000	-1.208	0.058
yield_curve	0.337**	0.000	-1.266**	0.000	-13.567**	0.000	4.390**	0.000
sigma_int	0.487	0.246	-1.898**	0.000	12.472**	0.000	-3.608**	0.000
cumprepayprob_24	13.053**	0.000	14.936**	0.000	32.441**	0.000	6.360**	0.000
cumdefaultprob_24	-58.106**	0.000	-45.471**	0.000	-4.463**	0.000	-0.718*	0.044
high_spd	0.154**	0.000	-0.095**	0.000	-0.249**	0.000	-0.090**	0.000
low_spd	-0.189**	0.000	-0.092**	0.000	-0.108**	0.000	-0.087**	0.000
Private Label Outcome	**Coefficient**	**P > |z|**	**Coefficient**	**P > |z|**	**Coefficient**	**P > |z|**	**Coefficient**	**P > |z|**
Constant	5.568**	0.000	10.268**	0.000	-0.803	0.616	21.138**	0.000
yield_spread	1.362**	0.000	1.922**	0.000	1.721**	0.000	0.966**	0.000
Jumbo	0.321**	0.000	0.748**	0.000	0.978**	0.000	0.415**	0.000
credit_spread	-5.533**	0.000	-4.754**	0.000	-4.111**	0.000	-5.331**	0.000
yield_curve	0.090	0.352	-1.529**	0.000	0.955	0.482	-12.071**	0.000
sigma_int	-1.701**	0.000	-3.969**	0.000	9.887**	0.000	-5.629**	0.000
cumprepayprob_24	9.992**	0.000	3.733**	0.000	3.565**	0.000	1.284*	0.026
cumdefaultprob_24	-71.992**	0.000	-33.662**	0.000	-4.940**	0.000	-2.505**	0.000
high_spd	0.107**	0.003	-0.107**	0.000	0.064**	0.010	0.230**	0.000
low_spd	-0.056	0.105	-0.135**	0.000	-0.114**	0.000	-0.268**	0.000
	Observations	89,707	Observations	131,836	Observations	122,180	Observations	77,109

This table reports the coefficients of a multinomial logit model which estimates the probability that a prime loan will be sold to the GSEs or privately securitized. This regression was estimated using the holdout sample that was created from the same origination years as the estimation samples. The independent variables are the yield spread at origination (yield_spread), whether or not the loan is above GSE loan limits (jumbo), the difference between the AAA bond index and the BBB bond index at the time of origination (credit_spread), the ratio between the 10-year Treasury rate and the 1-year Treasury rate (yield_curve), the interest rate volatility over the 24 months before origination (sigma_int), the cumulative 24 month prepayment and default probabilities (cumprepayprob_24 and cumdefaultprob_24) as well as the dummy variables to indicate whether the loans had a high yield spread (high_spd) or a low yield spread (low_spd). Lender fixed effects are also included but their coefficients are not reported here. ** significant at 1% level. * significant at 5% level.

Table 4b. Probability of Securitization for Subprime Loans under Rational Expectations

Private Label Outcome	2004 Coefficient	2004 P > |z|	2005 Coefficient	2005 P > |z|	2006 Coefficient	2006 P > |z|	2007 Coefficient	2007 P > |z|
constant	-6.826	0.080	20.454**	0.000	-32.904**	0.005	227.300**	0.003
yield_spread	0.172	0.222	0.347**	0.000	-0.106	0.346	0.007	0.975
Jumbo	-0.322	0.106	-0.051	0.558	-0.031	0.882	-0.245	0.570
credit_spread	7.778**	0.009	-14.768**	0.000	18.028**	0.001	-45.432	0.248
yield_curve	2.102	0.172	-10.215**	0.000	32.480**	0.001	-190.800**	0.001
sigma_int	0.197	0.961	13.935**	0.000	-9.141*	0.025	4.120	0.700
cumprepayprob_24	-5.926*	0.012	0.222	0.772	1.517	0.356	0.763	0.906
cumdefaultprob_24	-0.879*	0.033	-1.129**	0.001	-0.228	0.095	-0.205	0.757
high_spd	-0.134	0.447	0.110	0.151	-0.185	0.270	-0.071	0.813
low_spd	-0.388**	0.005	-0.050	0.402	-0.096	0.542	-0.289	0.262
	Observations	999	Observations	2,584	Observations	2,550	Observations	523

This table reports the coefficients of a multinomial logit model which estimates the probability that a subprime loan will be privately securitized. This regression is estimated using the holdout sample that was created from the same origination years as the estimation samples. The independent variables are the yield spread at origination (yield_spread), whether or not the loan is above GSE loan limits (jumbo), the difference between the AAA bond index and the BBB bond index at the time of origination (credit_spread), the ratio between the 10-year Treasury rate and the 1-year Treasury rate (yield_curve), the interest volatility over the 24 months before origination (sigma_int), the cumulative 24 month prepayment and default probabilities (cumprepayprob_24 and cumdefaultprob_24) as well as the dummy variables to indicate whether the loans had a high yield spread (high_spd) or a low yield spread (low_spd). Lender fixed effects are also included but their coefficients are not reported here. ** significant at 1% level. * significant at 5% level.

Table 4c. Probability of Securitization for Prime Loans under Adaptive Expectations

	2006		2007		2008	
GSE Outcome	**Coefficient**	**P > |z|**	**Coefficient**	**P > |z|**	**Coefficient**	**P > |z|**
Constant	15.602**	0.000	6.930**	0.000	45.970	0.711
yield_spread	0.666**	0.000	0.276**	0.000	0.409**	0.001
Jumbo	-1.859**	0.000	-2.292**	0.000	-2.760**	0.000
credit_spread	-6.191**	0.000	-1.849**	0.004	-49.204	0.707
yield_curve	-13.685**	0.000	0.183	0.805	4.192	0.701
sigma_int	12.218**	0.000	-5.302**	0.000	13.950	0.679
cumprepayprob_24	18.279**	0.000	17.306**	0.000	41.244**	0.000
cumdefaultprob_24	-60.664**	0.000	-40.426**	0.000	-5.475**	0.000
high_spd	-0.064**	0.002	0.088**	0.000	-0.199**	0.000
low_spd	-0.233**	0.000	-0.247**	0.000	-0.237**	0.000
Private Label Outcome	**Coefficient**	**P > |z|**	**Coefficient**	**P > |z|**	**Coefficient**	**P > |z|**
Constant	-1.104	0.486	23.642**	0.000	63.259	0.620
yield_spread	1.546**	0.000	0.637**	0.000	-0.211	0.159
jumbo	0.884**	0.000	0.471**	0.000	0.093	0.454
credit_spread	-4.185**	0.000	-5.502**	0.000	-65.600	0.627
yield_curve	1.404	0.297	-14.733**	0.000	5.627	0.617
sigma_int	9.349**	0.000	-6.947**	0.000	16.772	0.628
cumprepayprob_24	10.511**	0.000	13.427**	0.000	8.258**	0.001
cumdefaultprob_24	-48.050**	0.000	-40.116**	0.000	-4.462**	0.008
high_spd	0.144**	0.000	0.318**	0.000	0.092	0.217
low_spd	-0.139**	0.000	-0.332**	0.000	-0.118	0.081
	Observations	121,767	Observations	76,879	Observations	46,577

This table reports the coefficients of a multinomial logit model which estimates the probability that a prime loan will sold to the GSEs or privately securitized. This regression was estimated using the holdout sample that was created from loans that were originated two years after the estimation samples. The independent variables are the yield spread at origination (yield_spread), whether or not the loan is above GSE loan limits (jumbo), the difference between the AAA bond index and the BBB bond index at the time of origination (credit_spread), the ratio between the 10-year Treasury rate and the 1-year Treasury rate (yield_curve), the interest volatility over the 24 months before origination (sigma_int), the cumulative 24 month prepayment and default probabilities (cumprepayprob_24 and cumdefaultprob_24) as well as the dummy variables to indicate whether the loans had a high yield spread (high_spd) or a low yield spread (low_spd). Lender fixed effects are also employed but their coefficients are not reported here. ** significant at 1% level. * significant at 5% level.

Table 4d. Probability of Securitization for Subprime Loans under Adaptive Expectations

	2006		2007	
Private Label Outcome	**Coefficient**	**P > |z|**	**Coefficient**	**P > |z|**
Constant	-41.9895**	0.000	255.0**	0.000
yield_spread	-0.183**	0.002	-0.002	0.985
Jumbo	-0.257**	0.007	-0.075	0.675
credit_spread	20.234**	0.000	42.140*	0.030
yield_curve	39.960	0.293	-296.400**	0.000
sigma_int	-11.083**	0.000	-21.971**	0.007
cumprepayprob_24	2.140	0.080	1.284	0.318
cumdefaultprob_24	-0.304	0.189	-0.669	0.340
high_spd	-0.175*	0.044	0.043	0.768
low_spd	-0.072	0.397	0.045	0.736
	Observations	10,140	Observations	2,287

This table reports the coefficients of a multinomial logit model which estimates the probability that a subprime loan will be privately securitized. This regression was estimated using the holdout sample that was created from loans that were originated two years after the estimation samples. The independent variables are the yield spread at origination (yield_spread), whether or not the loan is above GSE loan limits (jumbo), the difference between the AAA bond index and the BBB bond index at the time of origination (credit_spread), the ratio between the 10-year Treasury rate and the 1-year Treasury rate (yield_curve), the interest volatility over the 24 months before origination (sigma_int), the cumulative 24 month prepayment and default probabilities (cumprepayprob_24 and cumdefaultprob_24) as well as the dummy variables to indicate whether the loans had a high yield spread (high_spd) or a low yield spread (low_spd). Lender fixed effects are also included but their coefficients are not reported here. ** significant at 1% level. * significant at 5% level.

Appendix 1. Yield Spread Regressions

Table A1a. Yield Spread Regression for Prime Loans

	2004		2005		2006		2007	
	Coefficient	**P> |t|**	**Coefficient**	**P> |t|**	**Coefficient**	**P> |t|**	**Coefficient**	**P> |t|**
constant	6.283**	0.000	8.249**	0.000	9.103**	0.000	9.463**	0.000
conform	-0.060**	0.000	-0.042**	0.000	0.057**	0.000	0.055**	0.002
jumbo	0.158**	0.000	0.184**	0.000	0.283**	0.000	0.304**	0.000
log_ltv	0.380**	0.000	0.283**	0.000	0.224**	0.000	0.226**	0.000
log_credit_spd	0.385**	0.000	1.006**	0.000	0.756**	0.000	0.418**	0.000
log_FICO	-0.716**	0.000	-0.837**	0.000	-0.938**	0.000	-1.131**	0.000
log_UPB	-0.134**	0.000	-0.190**	0.000	-0.172**	0.000	-0.189**	0.000
log_sigma_hpi	0.080**	0.000	0.064**	0.000	0.051**	0.000	0.055**	0.000
log_sigma_int	0.756**	0.000	0.447**	0.000	0.662**	0.000	-0.301**	0.000
log_yield_curve	1.298**	0.000	1.039**	0.000	-4.503**	0.000	-5.157**	0.000
south	-0.046**	0.000	0.003	0.429	-0.025**	0.000	-0.029**	0.000
midwest	0.014**	0.010	0.086**	0.000	0.087**	0.000	0.077**	0.000
west	-0.003	0.523	0.001	0.703	-0.022**	0.000	-0.022**	0.000
	Observations	89,707	Observations	131,836	Observations	122,180	Observations	77,109
	R-squared	0.2853	R-squared	0.1479	R-squared	0.1719	R-squared	0.1797

This table reports the coefficients from an Ordinary Least Squares regression of the mortgage yield spreads at the time of origination for prime mortgages using the holdout samples. The dependent variable (yield_spread) is the 10-year Treasury rate subtracted from the mortgage coupon rate. The independent variables are whether or not the loan conforms to GSE standards (conform), whether or not the loan amount is above GSE loan limits (jumbo), the logarithm of loan-to-value ratio (log_ltv), the logarithm of market credit risk premium (log_credit_spd), the logarithm of borrower's credit score (log_FICO), the logarithm of origination amount (log_UPB), the logarithm of house price volatility (log_sigma_hpi), the logarithm of interest rate volatility (log_sigma_int), the logarithm of the ratio of the 10-year Treasury rate to the 1-year Treasury rate (log_yield_curve) and dummy variables to indicate the geographic location of the property. ** significant at 1% level. * significant at 5% level.

Table A1b. Yield Spread Regression for Subprime Loans

	2004		2005		2006		2007	
	Coefficient	**P> |t|**	**Coefficient**	**P> |t|**	**Coefficient**	**P> |t|**	**Coefficient**	**P> |t|**
constant	39.435**	0.000	32.231**	0.000	45.129**	0.000	55.682**	0.000
conform	-0.109	0.177	0.126	0.052	0.054	0.467	0.221	0.175
jumbo	0.476**	0.000	0.484**	0.000	0.647**	0.000	0.551**	0.002
log_ltv	0.821*	0.011	1.216**	0.000	1.598**	0.000	1.649**	0.000
log_credit_spd	-1.888**	0.000	0.868*	0.020	3.745**	0.000	2.060	0.237
log_FICO	-4.610**	0.000	-3.502**	0.000	-5.509**	0.000	-8.085**	0.000
log_UPB	-0.852**	0.000	-0.931**	0.000	-1.039**	0.000	-0.737	0.000
log_sigma_hpi	-0.069	0.162	-0.110**	0.000	-0.172**	0.000	-0.248*	0.017
log_sigma_int	1.394**	0.000	0.230	0.490	-0.146	0.681	-1.016**	0.000
log_yield_curve	2.154**	0.008	0.091	0.840	2.414	0.144	-8.759**	0.001
south	-0.104	0.276	-0.090	0.123	0.052	0.448	0.168	0.229
midwest	-0.341**	0.003	-0.146*	0.041	0.080	0.368	0.052	0.753
west	-0.369**	0.000	-0.281**	0.000	0.114	0.144	0.086	0.574
	Observations	999	Observations	2,584	Observations	2,550	Observations	523
	R-squared	0.6126	R-squared	0.4288	R-squared	0.4727	R-squared	0.5286

This table reports the coefficients from an ordinary least squares regression of the mortgage yield spreads at the time of origination for subprime mortgages using the holdout samples. The dependent variable (yield_spread) is the 10-year Treasury rate subtracted from the interest rate. The independent variables are whether or not the loan conforms to GSE standards (conform), whether or not the loan amount is above GSE loan limits (jumbo), the logarithm of loan-to-value ratio (log_ltv), the logarithm of market credit risk premium (log_credit_spd), the logarithm of borrower's credit score (log_FICO), the logarithm of origination amount (log_UPB), the logarithm of house price volatility (log_sigma_hpi), the logarithm of interest rate volatility (log_sigma_int), the logarithm of the ratio of the 10-year Treasury rate to the 1-year Treasury rate (log_yield_curve) and dummy variables to indicate the geographic location of the property. ** significant at 1% level. * significant at 5% level.

Appendix 2. Lender Size and Securitization Choice

Table A2a. Probability of Securitization for Prime Loans under Rational Expectations – Small Lenders

	2004		2005		2006		2007	
GSE Outcome	**Coefficient**	**P > |z|**	**Coefficient**	**P > |z|**	**Coefficient**	**P > |z|**	**Coefficient**	**P > |z|**
Constant	5.175*	0.012	9.456*	0.031	-1.067	0.924	10.292	0.208
yield_spread	0.691	0.120	0.220	0.605	-0.745*	0.035	-0.704**	0.004
Jumbo	-1.565**	0.000	-1.011**	0.000	-1.360**	0.000	-1.694**	0.000
credit_spread	0.340	0.897	-3.950	0.157	2.685	0.586	0.607	0.904
yield_curve	-1.378*	0.020	-2.531	0.237	-0.756	0.936	-8.613	0.098
sigma_int	-4.380	0.102	-3.526	0.233	-2.058	0.567	-1.800	0.419
Cumprepayprob_24	1.862	0.589	13.703**	0.003	23.790**	0.000	5.305	0.200
cumdefaultprob_24	-20.771	0.091	-21.605*	0.041	14.605**	0.007	12.710**	0.003
high_spd	-0.327	0.061	-0.081	0.602	-0.023	0.892	0.250	0.141
low_spd	-0.072	0.701	0.210	0.173	-0.150	0.339	-0.020	0.881
Private Label Outcome	**Coefficient**	**P > |z|**	**Coefficient**	**P > |z|**	**Coefficient**	**P > |z|**	**Coefficient**	**P > |z|**
Constant	8.247**	0.000	13.379**	0.003	-3.808	0.751	15.888	0.088
yield_spread	2.312**	0.000	1.309**	0.003	-0.765*	0.037	-0.729*	0.022
Jumbo	0.872**	0.000	0.880**	0.000	1.236**	0.000	1.252**	0.000
credit_spread	-3.784	0.172	-6.507*	0.024	5.607	0.287	-1.180	0.833
yield_curve	-1.422*	0.023	-3.019	0.174	5.482	0.587	-10.577	0.079
sigma_int	-4.388	0.121	-6.901*	0.024	-1.500	0.698	-1.186	0.640
Cumprepayprob_24	-11.036**	0.002	5.687	0.240	-10.200	0.129	-7.697	0.110
Cumdefaultprob_24	2.581	0.841	-8.415	0.444	15.468**	0.006	5.977	0.181
high_spd	-0.378*	0.038	-0.072	0.659	0.482**	0.007	0.771**	0.000
low_spd	0.011	0.957	0.160	0.327	-0.337	0.051	-0.561**	0.001
	Observations	4,752	Observations	6,989	Observations	6,441	Observations	5,903

This table reports the coefficients of a multinomial logit model which estimates the probability that a prime loan originated by a small lender (defined as those originating 20 or fewer loans in a given year in our sample) will be bought by the GSEs or privately securitized. This regression was estimated using the holdout sample that was created from the same origination years as the estimation samples. The independent variables are the yield spread at origination (yield_spread), whether or not the loan amount is above GSE loan limits (jumbo), the difference between the AAA bond index and the BBB bond index at the time of origination (credit_spread), the ratio between the 10-year Treasury rate and the 1-year Treasury rate (yield_curve), the interest volatility over the 24 months before origination (sigma_int), the cumulative 24 month prepayment and default probabilities (cumprepayprob_24 and cumdefaultprob_24) as well as the dummy variables to indicate whether the loans had a high yield spread (high_spd) or a low yield spread (low_spd). ** significant at 1% level. * significant at 5% level.

Table A2b. Probability of Securitization for Prime Loans under Rational Expectations – Large Lenders

	2004		2005		2006		2007	
GSE Outcome	**Coefficient**	**P > |z|**	**Coefficient**	**P > |z|**	**Coefficient**	**P > |z|**	**Coefficient**	**P > |z|**
Constant	2.751**	0.000	7.413**	0.000	16.475**	0.000	1.413	0.221
yield_spread	-0.290**	0.001	0.868**	0.000	1.394**	0.000	0.994**	0.000
Jumbo	-2.001**	0.000	-1.072**	0.000	-1.700**	0.000	-2.435**	0.000
credit_spread	-0.925*	0.038	-2.360**	0.000	-7.936**	0.000	-1.224	0.058
yield_curve	0.382**	0.000	-1.235**	0.000	-13.466**	0.000	4.750**	0.000
sigma_int	0.609	0.153	-1.858**	0.000	12.640**	0.000	-3.623**	0.000
cumprepayprob_24	13.530**	0.000	14.965**	0.000	32.530**	0.000	6.294**	0.000
cumdefaultprob_24	-59.838**	0.000	-46.127**	0.000	-4.940**	0.000	-0.975**	0.007
high_spd	0.169**	0.000	-0.093**	0.000	-0.270**	0.000	-0.093**	0.000
low_spd	-0.195**	0.000	-0.097**	0.000	-0.087**	0.000	-0.091**	0.000
Private Label Outcome	**Coefficient**	**P > |z|**	**Coefficient**	**P > |z|**	**Coefficient**	**P > |z|**	**Coefficient**	**P > |z|**
constant	5.281**	0.000	10.122**	0.000	3.394*	0.043	21.918**	0.000
yield_spread	1.324**	0.000	1.933**	0.000	1.884**	0.000	1.001**	0.000
Jumbo	0.309**	0.000	0.743**	0.000	0.971**	0.000	0.389**	0.000
credit_spread	-5.414**	0.000	-4.703**	0.000	-4.358**	0.000	-5.527**	0.000
yield_curve	0.126	0.200	-1.484**	0.000	1.578	0.253	-12.703**	0.000
sigma_int	-1.684**	0.000	-3.855**	0.000	9.853**	0.000	-5.819**	0.000
cumprepayprob_24	11.339**	0.000	3.595**	0.000	3.883**	0.000	1.590**	0.007
cumdefaultprob_24	-77.474**	0.000	-34.518**	0.000	-5.485**	0.000	-2.473**	0.000
high_spd	0.122**	0.001	-0.106**	0.000	0.031	0.229	0.214**	0.000
low_spd	-0.061	0.080	-0.140**	0.000	-0.079**	0.002	-0.251**	0.000
	Observations	84,955	Observations	124,847	Observations	115,739	Observations	71,206

This table reports the coefficients of a multinomial logit model which estimates the probability that a prime loan originated by a large lender (defined as those originating more than 20 loans in a given year in our sample) will be bought by the GSEs or privately securitized. This regression was estimated using the holdout sample that was created from the same origination years as the estimation samples. The independent variables are the yield spread at origination (yield_spread), whether or not the loan is above GSE loan limits (jumbo), the difference between the AAA bond index and the BBB bond index at the time of origination (credit_spread), the ratio between the 10-year Treasury rate and the 1-year Treasury rate (yield_curve), the interest volatility over the 24 months before origination (sigma_int), the cumulative 24 month prepayment and default probabilities (cumprepayprob_24 and cumdefaultprob_24) as well as the dummy variables to indicate whether the loans had a high yield spread (high_spd) or a low yield spread (low_spd). Lender fixed effects are also included but their coefficients are not reported here. ** significant at 1% level. * significant at 5% level.