Did Securitization Lead to Lax Screening? Evidence From Subprime Loans 2001-2006*

Benjamin J. Keys†

Tanmoy Mukherjee‡

Amit Seru§

Vikrant Vig¶

First Version: November 2007

This Version: January 2008

*Acknowledgements: We thank Viral Acharya, Patrick Bolton, Charles Calomiris, Douglas Diamond, John DiNardo, Charles Goodhart, Anil Kashyap, Adair Morse, Daniel Paravisini, Giallume Plantin, Manju Puri, Raghuram Rajan, Adriano Rampini, Steve Schaefer, Robert VanOrder, Annette Vissing-Jorgensen, Luigi Zingales, and seminar participants at Duke (Fuqua School of Business) and London Business School for useful discussions. The opinions expressed in the paper are those of the authors and do not reflect the views of Sorin Capital Management. All remaining errors are our responsibility.

†University of Michigan, e-mail: benkeys@umich.edu

‡Sorin Capital Management, e-mail: tmukherjee@sorincapital.com

§University of Chicago, GSB, e-mail: amit.seru@chicagogsb.edu

¶London Business School, e-mail: vvig@london.edu

Did Securitization Lead to Lax Screening? Evidence From Subprime Loans 2001-2006

Abstract

Theories of financial intermediation suggest that securitization, the act of converting illiquid loans into liquid securities, could reduce the incentives of financial intermediaries to screen borrowers. We empirically examine this question using a unique dataset on securitized subprime mortgage loan contracts in the United States. We exploit a specific *rule of thumb* in the lending market to generate an instrument for ease of securitization and compare the composition and performance of lenders' portfolios around the ad-hoc threshold. Conditional on being securitized, the portfolio that is more likely to be securitized defaults by around 20% more than a similar risk profile group with a lower probability of securitization. Crucially, these two portfolios have similar observable risk characteristics and loan terms. We use variation across lenders (banks vs. independents), state foreclosure laws, and the timing of passage of anti-predatory laws to rule out alternative explanations. Our results suggest that securitization *does* adversely affect the screening incentives of lenders.

I Introduction

Securitization, converting illiquid loans of banks into liquid securities held by a dispersed group of shareholders, has grown tremendously in recent years and has transformed the traditional role of financial intermediaries. By improving risk sharing, securitization can lower the cost of capital, expanding access to the market for agents who would previously have been excluded.[1] The dramatic expansion of the "subprime" housing market, where loans are made to riskier borrowers in exchange for higher payments, would not have been possible without lenders' improved ability to spread risk through securitizing mortgages into mortgage-backed securities (MBS), and is at least partially responsible for the increase in homeownership by 6% in the last ten years.

However, amidst a rising wave of subprime delinquencies and foreclosures, there is increasing concern that securitization also weakens the incentives of banks to screen and monitor borrowers (Stiglitz (2007)). By selling loans to investors and removing them from their balance sheets, banks no longer bear the final burden of a delinquent loan. At the same time, the complexity of the securitization process, which bundles and sells packages of loans called "tranches," greatly reduces the transparency of individual loans' quality to investors. Securitization changes the incentives of lenders to adequately evaluate borrowers' ability to repay, making this information asymmetry between lenders and investors potentially costly. Understanding these costs is particularly important as policymakers debate on the scope and structure of the securitization process.

This paper sheds light on this aspect by empirically investigating the relationship between securitization and screening standards in the context of subprime mortgage backed securities. We exploit a specific credit score *rule of thumb* in the lending market as an instrument for ease of securitization and find evidence that securitization does reduce lenders' incentives to screen borrowers. Around these cutoffs, loans have similar observable risk features, demographic characteristics, and loan terms. When loans are easier to securitize by being just above a credit score cutoff, we find these loans default more frequently, by 20%, than similar loans below the threshold.

The theoretical motivation of our analysis follows from the literature on financial intermediation which argues that banks alleviate the asymmetric information problem (screening) and the incentive problem (monitoring) associated with lending. However, for a bank to provide an efficient level of these services, it must also be given the incentives to do so (Diamond (1984); Holmstrom and Tirole (1997)). In particular, it is necessary that the bank hold or retain the

[1]Furthermore, as Loutskina and Strahan (2007) have shown, securitization mitigates the real effects of monetary policy on bank lending.

risk of loans that it creates (Gorton and Pennacchi (1995)). This notion derives from the classic liquidity-incentives trade-off at the core of the financial contracting literature.[2] Thus, by spreading the risk of an individual contract across many investors, securitization can potentially destroy the incentives of banks to carefully screen borrowers. Of course, since banks and investors are involved in repeated relationships, reputation concerns may prevent any moral hazard from lenders. What the effects of securitization on screening are, thus, remains an empirical research question.

There are two major challenges to empirical research in this area. The first obstacle is related to the endogeneity of securitization. Loans that are bundled and sold as securities differ on several dimensions from loans that are not securitized. For example, securitized loans tend to be safer and less informationally sensitive as compared to loans that are held by banks (Drucker and Puri (2007)). Therefore, comparing screening activity between securitized loans and those that are not is likely to capture the effect of these (observable and unobservable) differences rather than capturing the causal effect of securitization on screening behavior.

We overcome this obstacle by using a research design that randomly increases the securitization likelihood of a loan as compared to a loan with similar characteristics. The methodology can be best understood using the following example. Loans to individuals that have a credit score above x are admitted to a *program*. Our methodology documents the causal effect of the program by comparing loans of those individuals that just qualified for the program (treatment group) to that of individuals who just failed to do so (control group). The methodology compares the performance of loans made to individuals with scores of $x + \epsilon$ to those with scores of $x - \epsilon$, arguing that the assignment of treatment (admission to the program) is randomized i.e., it is not driven by any observable or unobservable characteristics of the borrower. More specific to our analysis, we exploit a rule of thumb in the subprime lending market which increases the probability of securitization (program) around a specific credit score cutoff.

The credit cutoff is based on the summary measure of borrower credit quality known as the FICO score. Since the mid-1990s, the FICO score has become the most recognizable credit indicator used by lenders, rating agencies, and investors. The credit score cutoff, a FICO score of 620, has followed from the lending guidelines by Fannie Mae and Freddie Mac at the advent of the subprime market. We argue that steadfast adherence to this (ad-hoc) cutoff by investors (investment banks, hedge funds) in their default models generates an increase in the ease of securitization for loans that fall just above this credit cutoff relative to loans below this cutoff. Therefore, by comparing the portfolio of loans on either side of the credit score threshold, we

[2]See Coffee (1991); Bhide (1993); Maug (1998); Diamond and Rajan (2003); Aghion et al. (2004); DeMarzo and Urosevic (2006) for more on the liquidity-incentives trade-off. This theoretical argument has been tested in Gorton and Pennacchi (1995) and Sufi (2006).

can assess whether differential access to securitization led to changes in the behavior of lenders who offered these loans to consumers with nearly identical risk profiles.

A second obstacle is related to data limitations. There is substantial variation in contractual terms even in the loans that are securitized, as they differ on several dimensions such as maturity, leverage, priority, documentation, and fixed vs. floating interest rate terms. To make any causal claims, it is necessary to isolate differences in loan outcomes independent of these contract characteristics. This places enormous demands on the data that is being analyzed. To overcome this second obstacle, we use detailed data on subprime loans contracts from the LoanPerformance database.

Using a sample of more than two million home purchase loans during the period 2001-2006, we empirically confirm that the number of loans securitized varies systematically around the credit cutoff. For loans with a potential for significant "soft" (unobservable) information – "low" documentation loans – we find that there are 110% more loans securitized above the credit threshold of 621 vs. 619. Importantly for our research design, we do not find any differences in any observable characteristics of borrowers or loan terms around this credit threshold. Strikingly, we find that low documentation loans that are originated above the credit threshold tend to default more by around 20% within two years of origination (relative to a mean of around 5%; this amounts to roughly a 1% increase in absolute delinquencies). We obtain this result after controlling for all of the observable characteristics of loans. Since the only difference between the loans around the credit threshold is the increased likelihood of securitization, the increased default probability of loans above the credit threshold must be due to a reduction in screening by lenders.

We conduct additional analyses to rule out alternative explanations for our results. First, differences in the performance of the loans around the credit threshold could be driven by borrower selection on observables. However, the observable loan characteristics are smooth through the discontinuity and even the distributions of interest rates and loan-to-value (LTV) ratios are identical in the portfolios on either side of the FICO score threshold. Second, an argument might be made that banks screen similarly around the credit threshold but are able to sell portfolio of loans above and below the threshold to investors with different risk tolerance. This argument is also unlikely to explain our results since the securitized bundles sold to investors contain loans across the FICO spectrum.

Another alternative explanation is that lenders with different monitoring technologies (banks /thrifts vs. independent lenders) operate around the credit threshold. Contrary to this explanation, we find that screening incentives are affected similarly for independent lenders as well as banks around the credit cutoffs. Similarly, lender incentives might vary depending on the state in which they originate the loan. For instance, screening incentives of lenders to screen should be weakest in states where they are likely to recover collateral relatively easily, should the borrower default. Therefore, our results might be driven by differences in the composition of loans from different states around the cutoff. We find no evidence for this alternative. We do, however, find that the screening incentives of lenders are weakened more in states where foreclosure laws are lender friendly.

Alternatively, if FICO scores could be manipulated, lower quality borrowers might appear at higher credit scores and that might explain our findings. Note that as per the rating agency (Fair Isaac), aside from outright fraud, it is difficult to strategically manipulate one's FICO score in a targeted manner. Nevertheless, to examine this alternative more closely, we exploit a natural experiment that relies on the fact that FICO scores tend to be quite sticky (www.myfico.com) – i.e., it takes long periods of time (more than 3 to 6 months) to improve credit scores. The natural experiment involves the passage of anti-predatory laws in New Jersey (2002) and Georgia (2003) that reduced securitization in the subprime market drastically. However, subsequent to protests by market participants, the laws were amended substantially and the market reverted to pre-predatory law levels. We are able to exploit the time series variation in securitization likelihood in the two states and show that the main effects show up relatively quickly (i.e., over short horizons). This rules out manipulation in credit scores as an alternative explanation for our results.

Finally, as an added test of the role of soft information on screening incentives of lenders, we investigate the "full" documentation loan lending market. These loans have potentially significant "hard" (observable) information since complete background information about the borrower's ability to repay is provided. In this market, we identify another credit cutoff (FICO of 600) based on the advice of the three credit repositories. We find that 100% more loans are securitized at credit threshold of 601 vs. 599 for loans with full documentation. Interestingly, however, we find no significant difference in default rates of full documentation loans originated around this credit threshold. As more hard information is provided for full documentation loans, this result is consistent with the notion that transparency reduces moral hazard in the subprime market.

Our paper is related to the literature that examines how loan contracts are designed to solve the moral hazard and adverse selection problems that accompany loans sales. Gorton and Pennacchi (1995) present a model of incentive compatible loan sales and empirically document that banks retain a larger share of the riskier loans to mitigate incentive problems. In a recent paper, Drucker and Puri (2007) investigate how banks reduce agency problems that arise from loan sales. Using data on syndicated loan sales they document that sold loans, for example, have more restrictive covenants when agency problems are more severe. Our research adds to this literature by clarifying that securitization can *directly* affect the risk composition of the portfolio originated by banks. The fact that we find similar contracts around the threshold suggests that from the lenders' perspective, differences in default risk are compensated for by reduced screening costs.

Our results suggest that lenders acted strategically in providing loans for those borrowers just above the credit threshold. Differential ease of securitization reduced the willingness of lenders to bear the cost of screening, as the dangers of default are no longer directly borne by the lender. This is consistent with the views of Mian and Sufi (2008), who argue that disintermediation was the driving factor behind the increase in mortgage defaults from 2005 to 2007. Our findings strongly imply a role for policymakers and market participants to improve screening mechanisms and encourage more transparency in the securitization process.

The rest of the paper is organized as follows. Section II provides an overview of lending in the subprime market. Section III describes the data and sample construction. Section IV discusses the empirical methodology used in the paper, while Section V presents the empirical results in the paper. Section VI presents evidence to address alternative explanations. Section VII concludes.

II Lending in Subprime Market

II.A Overview of History

Approximately 60% of outstanding U.S. mortgage debt is traded in mortgage-backed securities (MBS), making the U.S. secondary mortgage market the largest fixed-income market in the world (Chomsisengphet and Pennington-Cross (2006)). The bulk of this securitized universe ($3.6 trillion outstanding as of January 2006) is comprised of agency pass-through pools – those issued by Freddie Mac, Fannie Mae and Ginnie Mae. The remainder, approximately, $2.1 trillion as of January 2006 has been securitized in non-agency securities. The two markets are separated based on the eligibility criteria of loans that the government agencies have established. Broadly, agency eligibility is established on the basis of loan size, credit score, and underwriting standards. While the non-agency MBS market is relatively small as a percentage of all U.S. mortgage debt, it is nevertheless large on an absolute dollar basis. Unlike the agency market, the non-agency (refereed to as subprime in the paper) market was not always this size. This market gained momentum in the mid- to late-1990s. Inside B&C Lending – a publication which covers subprime mortgage lending extensively – reports that total subprime lending (B&C originations) has grown from $65 billion in 1995 to $500 billion in 2005.

According to Gramlich (2007), many factors led to the growth of subprime lending. First, it became legal to charge high rates and fees to borrowers based on risk. These high rates were not possible until the Depository Institutions Deregulation and Monetary Control Act (DIDMCA) was adopted in 1980, which preempted state usury laws in the form of interest rate caps. The Alternative Mortgage Transaction Parity Act (AMTPA) in 1982 permitted the use of variable interest rates and balloon payments. Second, the Tax Reform Act of 1986 (TRA) prohibited the deduction of interest on consumer loans, yet allowed interest deductions on mortgages for a primary residence as well as an additional home. This made even high-cost mortgage debt cheaper than consumer debt for many homeowners. Finally, market changes also contributed to the growth and maturation of subprime loans. In an environment of low and declining interest rates, such as during the late-1990s and early-2000s, mortgage brokers and mortgage companies switched from the prime to subprime market to maintain loan volume.

Growth in mortgage-backed securities led to an increase in securitization rates (the ratio of the dollar-value of loans securitized divided by the dollar-value of loans originated) from less than 30 percent in 1995 to over 80 percent in 2006. The market is largely accounted for by the top 50 lenders in terms of both volume and value. Inside B&C Lending reports that the market share of the top 50 lenders in the market has steadily grown from about 60 percent in 1995 to over 95 percent in 2006.

II.B Characteristics of Subprime Lending

From the borrower's perspective, the primary distinguishing feature between prime and subprime loans is that the up-front and continuing costs are higher for subprime loans. Up-front costs include application fees, appraisal fees, and other fees associated with originating a mortgage. The continuing costs include mortgage insurance payments, principle and interest payments, late fees for delinquent payments, and fees levied by a locality (such as property taxes and special assessments). The subprime mortgage market actively prices loans based on the risk associated with the borrower. Specifically, the interest rate on the loan depends on credit scores, debt-to-income ratios and the documentation level of the borrower. In addition, the exact pricing may depend on loan-to-value ratios (the amount of equity of the borrower), the length of the loan, the flexibility of the interest rate (adjustable, fixed, or hybrid), the lien position, the property type and whether stipulations are made for any prepayment penalties.[3]

In addition to being higher risk, subprime borrowers differ from prime borrowers on a number of different (albeit correlated) dimensions. As compared to the prime market, loans originated in the subprime market are predominantly in African-American census tracts (Canner and Passmore (1999); Immergluck and Wiles (1999); Calem, Hershaff, and Wachter (2004)). Moreover,

[3]For example, the rate and underwriting matrix of Countrywide Home Loans Inc., a leading lender of prime and subprime loans, shows how the credit score of the borrower and the loan-to-value ratio are used to determine the rate at which different documentation-level loans are made (www.countrywide.com).

borrowers who are credit-constrained are most likely to finance the purchase of a home by using a subprime mortgage (Bostic, Calem, and Wachter (2005); Nichols, Pennington-Cross, and Yezer (2005)). Fueling recent policy discussions surrounding "predatory" lending practices, surveys have found that subprime borrowers are less likely to search for the best loan terms and are less knowledgeable about the mortgage process (Courchane, Surette, and Zorn (2004)).

For investors who hold the eventual mortgage-backed security, credit risk in the agency sector is mitigated by an implicit or explicit government guarantee, but non-agency (subprime) securities have no such guarantee. Instead, credit enhancement for non-agency deals is in most cases provided internally by means of a deal structure which bundles loans into "tranches," or segments of the overall portfolio (Lucas, Goodman and Fabozzi (2006)). The vast majority of tranches in non-agency deals carry triple-A ratings, and credit risk (i.e., the risk that all principal will not be returned) is channeled to a small percentage of lower-rated tranches by cashflow rules that are designed to protect the "senior" higher rated bonds. Thus, in addition to time tranching deal cashflows in structures with various cashflow windows, non-agency deals are also credit tranched. This process mitigates the risk to triple-A bonds after the loans have been removed from banks' balance sheets.

III Data

Our primary data contain individual loan data leased from LoanPerformance. The database is the only source which provides a detailed perspective on the non-agency securities market. The data includes information on issuers, broker dealers/deal underwriters, servicers, master servicers, bond and trust administrators, trustees, and other third parties. As of December 2006, more than 8,000 home equity and nonprime loan pools (over 7,000 active) that include 16.5 million loans (more than seven million active) with over $1.6 trillion in outstanding balances were included. LoanPerformance estimates that as of 2006, the data covers over 90% of the subprime loans that are securitized (for more discussion on LoanPerformance also see Demyanyk and Van Hemert (2007) and recent papers by Pennington-Cross).[4] The dataset includes all standard loan application variables such as the loan amount, term, LTV ratio, credit score, and interest rate type – data elements that are typically disclosed and form the basis of contracts in non-agency deals. We now describe some of these variables in more detail.

For our purpose, the most important piece of information about a particular loan is the credit

[4]Note that only loans that are securitized are reported in the LoanPerformance database. Moreover, communication with the database provider suggests that the 10% of loans that are not reported are for privacy concerns from lenders. Importantly for our purpose, the exclusion is not based on any selection criteria that the vendor follows (e.g., loan characteristics or borrower characteristics).

worthiness of the borrower. The borrower's credit quality is captured by a summary measure called the FICO score. The FICO score has increasingly become the most recognizable measure used by lenders, rating agencies, and investors to assess borrower quality (Gramlich (2007)). The software used to generate the score from individual credit reports is licensed by the Fair Isaac Corporation (FICO) to the three major credit repositories – TransUnion, Experian, and Equifax. These repositories, in turn, sell FICO scores and credit reports to lenders and consumers. FICO scores provide a ranking of potential borrowers by the probability of having some negative credit event in the next *two years*. Probabilities are rescaled into a range of 400-900, though nearly all scores are between 550 and 800, with a higher score implying a lower probability of a negative event. The negative credit events foreshadowed by the FICO score can be as small as one missed payment or as large as bankruptcy. Borrowers with lower scores are proportionally more likely to have all types of negative credit events than are borrowers with higher scores.

FICO scores have been found to be accurate even for low-income and minority populations (for more information see www.myfico.com; also see Chomsisengphet and Pennington-Cross (2006)). More importantly, the applicability of scores available at loan origination extends reliably up to two years. By design, FICO measures the probability of a negative credit event over a two-year horizon. Mortgage lenders, on the other hand, are interested in credit risk over a much longer period of time. The continued acceptance of FICO scores in automated underwriting systems indicates that there is a level of comfort with their value in determining lifetime default probability differences.[5] Keeping this as a backdrop, most of our tests of borrower default will examine the default rates up to 24 months from the time the loan is originated.

Borrower quality can also be gauged by the level of documentation collected by the lender when taking the loan. The documents collected provide historical and current information about the income and assets of the borrower. Documentation in the market (and reported in the database) is categorized as full, limited or no documentation. Borrowers with full documentation provide verification of income as well as assets. Borrowers with limited documentation provide no information about their income but do provide some information about their assets. No documentation borrowers provide no information about income or assets, which is a very rare degree of screening lenience on the part of lenders. In our analysis, we combine limited and no documentation borrowers and call them low documentation borrowers. Our results are unchanged if we remove the very small portion of loans which are no documentation.

The data also provide information on the features of the loan contracts. Specifically, we have

[5] An econometric study by Freddie Mac researchers showed that the predictive power of FICO scores drops by about 25 percent once one moves to a three-to-five year performance window (Holloway, MacDonald and Straka, 1993). FICO scores are still predictive, but do not contribute as much to the default rate probability equation after the first two years.

information on the type of mortgage loan (fixed rate, adjustable rate, balloon or hybrid), and the loan-to-value ratio (LTV) of the loan, which measures the amount of the loan expressed as a percentage of the value of the home. Finally, there is also information about the property being financed by the borrower. There is also information on the purpose of the loan. Typically loans are classified as either for purchase or for refinance (with or without cashout), and we focus exclusively on loans for home purchases. Information about the geography where the loan is located (zipcode) is also available in the database.

Most of the loans in our sample are for the owner-occupied single-family residences, townhouses, or condominiums. Therefore, to ensure reasonable comparisons we restrict the loans in our sample to these groups. We also drop non-conventional properties, such as those that are FHA or VA insured or pledged properties, and also exclude buy down mortgages. We also exclude Alt-A loans, since the coverage for these loans in the database is limited.[6] Only those loans with valid FICO scores are used in our sample. While we have information on loans starting 1997, we conduct our analysis for the period January 2001 to December 2006, since the securitization market in the subprime market grew to a meaningful size post-2000 (Gramlich (2007)). However, we do examine the period 1997 to 2000 when discussing how the market has evolved over time in Table I.

IV Empirical Methodology

IV.A Credit Thresholds as Instruments

We exploit a specific *rule of thumb* in the lending market which makes securitization of loans more likely if a certain FICO score threshold is attained. This generates an instrument for the ease of securitization of a loan. The idea then is to compare the characteristics of the loan market above and below the ad-hoc credit threshold.

The cutoff that we exploit is at the FICO score of 620. Historically, this score was established as a minimum threshold in the mid-1990's by Fannie Mae and Freddie Mac in their guidelines on loan eligibility. According to Fair Isaac, "...those agencies [Fannie Mae and Freddie Mac], which buy mortgages from banks and resell them to investors, have indicated to lenders that any consumer with a FICO score above 620 is good, while consumers below 620 should result in further inquiry from the lender...."[7] Similarly, guidelines by Freddie Mac suggest that FICO

[6]These borrowers are generally considered to be less risky – i.e., these borrowers on average have higher FICO scores.

[7]This was reported by Craig Watts, a spokesperson for Fair, Isaac and Company in an interview to Detroit Free Press. Similarly, Charles Capone, Jr., a senior Analyst with Microeconomic and Financial Studies Division U.S. Congressional Budget Office Washington, DC, wrote in "Research Into Mortgage Default and Affordable

scores below 620 are placed in the *Cautious Review Category*, and Freddie Mac considers a score below 620 "as a strong indication that the borrower's credit reputation is not acceptable."[8] There is also evidence that rating agencies (Fitch and Standard and Poor's) use this cutoff to determine default probabilities of loans when rating mortgage backed securities with subprime collateral (Temkin, Johnson and Levy (2002)).

We argue that adherence to this (ad-hoc) cutoff by investors (investment banks, hedge funds) in their default models, following the advice of GSEs, Fair Isaac, and rating agencies, generates an increase in demand for securitized loans which are just above the credit cutoff relative to loans below this cutoff. Our empirical methodology then is to examine how the loans that are originated around this credit threshold perform. In other words, we use this threshold to instrument for the ease of securitization – the notion being that loans of borrowers with creditworthiness just above the cutoff, 620^{+}, are randomly assigned a higher probability of securitization relative to loans made to borrowers with creditworthiness just below the cutoff, 620^{-}, because of the rule of thumb elaborated earlier. We discuss this in greater detail in the next subsection.

IV.B Methodology

Our empirical strategy uses the ad-hoc thresholds in the FICO score (at 620) to causally link securitization to screening performed by lenders. By focusing on the lender as a unit of observation we attempt to learn about the differential impact securitization has on behavior of lenders around the cutoff.

When a borrower approaches a lender for a mortgage loan, the lender asks the borrower to fill out a credit application. In addition, the lender also obtains the borrower's credit report from the three credit bureaus. Part of the background information on the application and report could be considered "hard" information (e.g., the FICO score of the borrower), while the rest is "soft" (e.g., a previous relationship with the lender, which appraiser was used to value the house, how many years of documentation were provided by the borrower, joint income status) in the sense that it is less easy to summarize on a legal contract. The lender expends effort to process the soft and hard information about the borrower and, based on this assessment, offers a menu of contracts to the borrower. Subsequently, borrowers decide to accept or decline the

Housing: A Primer" that for most of the 1990s, the mortgage market viewed a FICO score of 620 as the bottom cutoff of loans that could be sold to Fannie Mae or Freddie Mac. Popular press has also noted frequently that borrowers above 620 are considered to be of the good kind and that a score of 620 is the line between good and bad borrowers (for e.g., see www.money.cnn.com/2003/02/17/pf/banking/chatzky/ or more recently http://online.wsj.com/article/SB119662974358911035.html.

[8] Freddie Mac, Single-Family Seller/Servicer Guide, Chapter 37, Section 37.6: Using FICO Scores in Underwriting (03/07/01).

loan contract offered by the lender.

Once a loan contract has been accepted, the loan can be sold as part of a securitized pool to investors. Notably, only the hard information about the borrower (FICO score) and the contractual terms (e.g., LTV ratio, interest rate) are used by investors when buying these loans as a part of securitized pool.[9] In fact, the variables about the borrowers and the loan terms in the LoanPerformance database are identical to those used by investors and rating agencies to rate tranches of the securitized pool. Therefore, while lenders are compensated for the hard information about the borrower, the incentive for lenders to process soft information critically depends on whether they have to bear the risk of loans they originate (Gorton and Pennacchi (1995) and Parlour and Plantin (2007)).

The central claim in the paper is that lenders are less likely to expend effort to process soft information as the ease of securitization increases. Specifically, consider a situation in which an adhoc cutoff rule in the lending market exists around a credit score x. The lender knows that it is much easier to securitize the loan of a x^+ applicant – and therefore less likely to stay on its balance sheet if the loan were to be originated – as compared to the x^- applicant, even though they are nearly identical in terms of their *observable* risk profiles. In other words, loans at x^+ are more liquid than loans at x^-. Now suppose that two borrowers – one with credit score of $x + \epsilon$ (x^+) and the other with a credit score of $x - \epsilon$ (x^-) – approach the lender for a loan. Given the adhoc rule, lenders are likely to do a more careful review of both soft and hard information for borrowers with credit score of x^-. There is widespread evidence of this behavior; for instance, Advantage Mortgage's website claims that "...all loans with credit scores below 620 require a second level review....There are no exceptions, regardless of the strengths of the collateral or capacity components of the loan."[10]

To begin with, our tests empirically establish a statistical discontinuity in the distribution of loans securitized around the credit threshold of 620. In order to do so, we first show that the number of loans securitized dramatically increases when we move from 620^- to 620^+. We measure the extent of the jump by using techniques which are commonly used in the literature on regression discontinuity (e.g., see DiNardo and Lee (2004) and Card et al. (2007)). Specifically, we collapse the data on each FICO score (500-800) i, and estimate equations of the form:

$$Y_i = \Big(\alpha + \beta T_i + \theta f(FICO(i)) + \delta T_i * f(FICO(i)) + \epsilon_i \Big), \quad (1)$$

[9]See Testimony of Warren Kornfeld, Managing Director of Moodys Investors Service before the subcommittee on Financial Institutions and Consumer Credit U.S. House of Representatives May 8, 2007.

[10]See www.advantagemtg.com. This position for loans below 620 is reflected in lending guidelines of numerous other lenders. We also conducted a survey of origination matrices used by the top 50 originators in the subprime market (from a list obtained from Inside B&C Lending). We obtained origination matrices from the websites of many of these originators. These credit thresholds are being used by nearly all the lenders.

where Y_i is the number of loans at FICO score i, T_i is an indicator which takes a value of 1 at FICO ≥ 620 and a value of 0 if FICO < 620 and ϵ_i is a mean-zero error term. $f(FICO)$ and $T * f(FICO)$ are flexible seventh-order polynomials, with the goal of these functions being to fit the smoothed curves on either side of the cutoff as closely to the data presented in the figures as possible.[11] $f(FICO)$ is estimated from 620^- to the left, and $T * f(FICO)$ is estimated from 620^+ to the right. The magnitude of the discontinuity, β, is estimated by the difference in these two smoothed functions evaluated at the cutoff. We perform these regressions using FICO scores between 500 and 800 – the range of nearly all credit scores observed in the data – for each year after 2000. In this framework, the coefficient on the indicator T_i, β, yields a consistent estimate of the extent of the jump in the number of loans in the local vicinity of the discontinuity.

After documenting a large jump at the ad-hoc credit thresholds, we focus on the performance of the loans around these thresholds. We evaluate the performance of the loans by examining the default probability of loans – i.e., whether or not the loan defaulted t months after it was originated. If lenders screen similarly for the loan of credit quality 620+ and the loan of 620- credit quality, there should not be any discernable differences in default rates of these loans. Since we are comparing the composition of loans around the cutoff, and controlling for any observable loan characteristics, our maintained claim is that any differences in default rates on either side of the cutoff should be only due to impact that securitization has on lenders' screening standards. The underlying assumption is that as we approach the cutoff from either side, any unobserved differences in the features of the loans (not correlated with securitization) should disappear in the limit.[12] Thus any comparison around the discontinuity will be a consistent estimate of the impact of securitization.

This claim relies on several identification assumptions which drive our empirical estimation. First, at a given price, we assume that the number of mortgage loans demanded among potential buyers with a credit score of 620^+ is same as the number of loans that demanded at 620^-, and the discontinuity is only in the number of loans that are securitized at these thresholds. This is a plausible assumption since there is no reason to believe that housing demand also jumps discontinuously at these credit thresholds. This is also confirmed in Figure 1A, where we report the distribution of FICO scores of borrowers. There is no discontinuity in the number of borrowers at each score – i.e., as explained in Section III, the distribution of these scores across the population is smooth. Second, we assume that screening is costly for the lender. The notion is that collection of information – hard systematic data (e.g., FICO score) as well as

[11] We have also estimated these functions of the FICO score using 3rd order and 5th order polynomials in FICO, as well as relaxing parametric assumptions and estimating using local linear regression. The estimates throughout are not sensitive to the specification of these functions.

[12] Formally, $\lim_{FICO \to 620-}\{E[\epsilon_i|X_i, T_i = 0]\} = \lim_{FICO \to 620+}\{E[\epsilon_i|X_i, T_i = 1]\}$, where X_i are any other observable loan characteristics.

soft information (e.g., joint income status) about the creditworthiness of the borrower – would require time and effort by loan officers. If lenders did not have to expend resources to collect information, it would be tough to argue that the differences in performance we estimate are a result of ease of securitization around the credit threshold affecting banks incentives to screen and monitor. Again, this seems a reasonable assumption (see Gorton and Pennacchi (1995)).

Third, we assume that there is no explicit manipulation of FICO scores by the lenders or borrowers. If untrue, some risky borrowers at 620^- might be incentivized to increase their credit scores to 620^+. We provide several pieces of evidence that support this assumption.[13] First, the independent rating agency (Fair Isaac) responsible for generating these scores argues that it is difficult to strategically manipulate FICO scores in a narrowly targeted manner (see www.myfico.com). Taking steps to improve one's credit, such as paying off credit card debt, would improve the credit score but not in a precise (or predictable) way to move a borrower from one side of the threshold to the other. Second, the credit scores reported in LoanPerformance are the median of the credit scores reported by the three repositories – TransUnion, Experian, and Equifax.[14] Since each of these has a different set of creditworthiness criteria, this calculation makes it even harder to manipulate all three scores. Finally, FICO scores tend to be quite sticky – i.e., it takes long periods of time to substantially improve credit scores, generally in the range of at least 3-6 months (www.myfico.com). Our analysis in Section VI.D focuses on a natural experiment and shows that the effects of securitization show up relatively quickly (i.e., over short horizons).

V Main Empirical Results

V.A Descriptive Statistics

As noted earlier, the non-agency market differs from the agency market on three dimensions: FICO scores, loan-to-value ratios and the amount of documentation asked of the borrower. We next look at the descriptive statistics of our sample with special emphasis on these dimensions. Even though the coverage of the database is thin before 2000, we use data in this section from 1997 to 2006 to give some sense of how the market has evolved. Our main analysis will, however, only include years of data since 2000.

Our analysis uses more than two million loans (2,942,934) across years. As mentioned earlier,

[13]It is not clear for whom the manipulation is easy. If for instance it is easier for the good quality borrowers, then it would bias us against finding any increased defaults at higher credit scores (relative to lower credit threshold), since good quality borrowers would leave lower thresholds and move to higher ones.

[14]Technically these "FICO" scores are called EMPIRICA, Experian/Fair Isaac Risk Model, and BEACON score for TransUnion, Experian, and Equifax, respectively.

the non-agency securitization market experienced a dramatic increase in the late-1990s, which is apparent in Panel A of Table I, which shows the number of loans securitized across years. These patterns are similar to those in Demyanyk and Van Hemert (2007). The rapid appreciation of house prices (Figure 1B) and the interest rate policy of the Fed following the internet stock market run up from 2000 onwards certainly contributed to the enormous growth of the subprime market (Chomsisengphet and Pennington-Cross (2006)). Note that, based on estimates provided by LoanPerformance, the total number of non-agency loans originated has hovered between 1.1 (2004 onwards) to 1.5 (early 2000) times the number of loans that are securitized.

Panel A also shows that average FICO scores of individuals who access the subprime market has been increasing over time. For instance, the mean FICO score in our subprime borrower database increased from 611 in 1997 to roughly 640 in 2006. Part of the reason behind this increase in average FICO scores is the ease of access that borrowers with higher credit scores had once securitization took off in the late-1990s. In addition, this increase in FICO scores is consistent with a secular trend of increasing average creditworthiness among all credit-scored accounts. As more households have obtained a credit score by having sufficient credit history, these households with thin-but-positive credit files have pushed up the average credit score nationwide.

We can also see that LTV ratios have gone up over time, as borrowers have put in less and less equity into their homes when financing loans. This increase is consistent with a better appetite of market participants to absorb risk. In fact, this is often considered the bright side of securitization – borrowers are able to borrow at better credit terms since risk is being borne by investors who can bear more risk than individual banks. Finally, consistent with arguments above, the market has seen an increase in number of loans with reduced "hard" information in the form of limited or no documentation. Note that while limited documentation provides no information about income but does provide some information about assets, a no documentation loan provides no information about income or assets. In our analysis we combine both types of limited-documentation loans and denote them as *low* documentation loans.

A significant part of the empirical analysis will compare the performance of full documentation and low documentation loans separately. To understand them better, we split the sample into these two categories in Panels B and C of Table I. We find similar trends for number of loans, LTV ratios and FICO scores in the two documentation groups. The low documentation loan market does see a larger increase in the number of loans, growing 40 times over by 2006 relative to 1997; in comparison, the number of full documentation loans grew 20 times during the same time period. Average LTV ratios are lower and FICO scores higher for low documentation as compared to the full documentation sample. This possibly reflects the additional uncertainty lenders have about the quality of low documentation borrowers. In the analysis which follows, we present results *within* a given category, either full or low documentation.

In Panels D and E, we provide general trends and summary statistics of key variables used in the analysis. The number of originated loans post-2000 has increased by about eight times relative to the market before that period (up from an average of 305,000 to 2,637,000). In terms of dollars the increase has been roughly seven-fold (from about $100 billion to $700 billion). The average loan size has also increased by roughly 45% (from $113,000 to $165,000). Note that the average loan size is significantly smaller than the average loan size in the agency market – which is around $300,000. Loans that are classified as first-time purchases have gone up over time (38% post-2000 vs. 28% pre-2000). Though we have information on purchases and refinancing contracts (cash outs and without cash outs), our analysis focuses on loans for home purchases. We find qualitatively similar results for other loan purpose classes as well. Notably, the initial interest rate on subprime mortgages has fallen since 2000. This is consistent with sharp declines in both the 1-year and 10-year Treasury yields post-2000. Finally, we also observe that there has been an increase in interest-only and balloon loans since 2000 (30% vs. 10%). A large part of the reason for increase in these loans has to do with the good market conditions that prevailed since 2000, allowing for these exotic and potentially risky loan structures to develop and gain in popularity (Gramlich (2007)).

V.B Estimating Discontinuity in Number of Loans Around Credit Threshold

We first present results that show that large differences exist in the number of loans that are securitized around the credit threshold we described earlier. We then examine whether this jump in securitization has any consequences on the subsequent performance of the loans above and below this credit threshold.

As mentioned in Section IV, the rules of thumb in the lending market impact the ease of securitization around the credit score of 620. We therefore expect to see a substantial increase in the number of loans just above this credit threshold as compared to number of loans just below this threshold. In order to examine this, we start by plotting the number of loans at each FICO score in the two documentation categories around the credit cutoff of 620 across years starting with 2001 and ending in 2006. As can be seen from Figure 2, there is a marked increase in number of low documentation loans around the credit score of 620 – that is, at 620^{+} relative to number of loans at 620^{-}. We do not find any such jump for full documentation loans at FICO of 620.[15] Given this evidence, we focus on the 620 credit threshold for low documentation loans.

From Figure 2, it is clear that the number of loans see roughly a 100% jump in 2004 for low documentation loans around the credit score of 620 – i.e., the number of loans securitized

[15] We will elaborate more on full documentation loans in Section VI.

at 620^+ are 100% higher in number as compared to loans securitized at 620^-. Clearly, this is consistent with the hypothesis that the ease of securitization is higher at 620^+ than at scores just below this credit cutoff.

To estimate the jumps in the number of loans, we use the methods described above in Section IV.B using the specification provided in equation (1). As reported in Table II, we find that low documentation loans see a dramatic increase above the credit threshold of 620 after 2000. In particular, the coefficient estimate (β) is significant at the 1% level and is on average around 110% (from 73 to 193%) higher for 620^+ as compared to 620^- for post-2000 loans.[16] These jumps are plainly visible from the yearly graphs in Figure 2. As an additional check, we conducted permutation tests (or "randomization" tests), where we varied the location of the discontinuity (T_i) across the range of all possible FICO scores and re-estimated equation (1). Although there are other gaps in the distribution in other locations in various years, the estimates at 620 for low documentation are strong outliers relative to the estimated jumps at other locations in the distribution (results not shown).

V.C Contract Terms and Borrower Demographic Profile Around Credit Threshold

Before examining the subsequent performance of loans around the credit threshold, we first check if there are any differences in hard information – either in terms of contract terms or other borrower characteristics – around this threshold. Though we control for these differences when we evaluate the performance of loans, it is insightful to examine whether borrower and contract terms also systematically differ around the credit threshold. If they were different, the patterns in the number of loans could be explained by differential composition effects on either side of the cutoff. We start by examining the contract terms – LTV and interest rates – around the credit threshold. Figures 3 and 4 show the distribution of LTV and interest rates on loan terms offered on low documentation loans across the FICO spectrum. As is apparent we find these loan terms to be very similar – i.e., we find no differences in contract terms for low documentation loans above and below the 620 credit score.

We test this formally using an approach equivalent to equation (1), replacing the dependent variable Y_i in the regression framework with contract terms (loan-to-value ratios and interest rates) and present the results in Panels A and B of Table III. Our results suggest that there is no difference in loan terms around the credit threshold. For instance, for low-documentation loans originated in 2006, the average loan-to-value ratio across the collapsed FICO spectrum is

[16] For instance, in 2001, the estimated discontinuity in Panel A is 85. The mean average number of low documentation loans at a FICO score for 2001 is 117. The ratio is around 73%.

85%, whereas our estimated discontinuity is only -1.05%, a 1.2% difference. Similarly for the interest rate, for low-documentation loans originated in 2005, the average interest rate is 8.2%, and the difference on either side of the credit score cutoff is only about -0.091%, a 1% difference. We repeated similar tests for whether or not the loan is ARM, FRM or interest only/balloon and find similar results.[17] These differences are well within the range of sampling variation. Permutation tests, which allow for the location of the discontinuity T_i to occur at each possible FICO score, confirmed that the estimates at 620 for low documentation are within the range of other jump estimates across the spectrum of FICO scores (results not shown).

Next, we examine whether the characteristics of borrowers differ systematically around the credit threshold. In order to evaluate this, we look at the distribution of the population of borrowers across the FICO spectrum for low documentation loans. The data on borrower demographics comes from Census 2000 and is at the zip code level. As can be seen from Figure 5, median household income of the zip codes of borrowers around the credit thresholds look very similar for low documentation loans. We plotted similar distributions for average percent minorities residing in the zip code, and average house value in the zip code across the FICO spectrum (unreported) and again find no differences around the credit threshold.[18]

We use the same specification as equation (1) for the number of loans, this time with the borrower demographic characteristics as dependent variables and present the results formally in Table IV (Panels A, B and C). Consistent with the patterns in the figures, we find no differences in borrower demographic characteristics around the credit score threshold. For low-documentation loans originated in 2005, for example, the median household income across the FICO spectrum is $47,390, and the estimated difference on either side of the cutoff is $197. These differences are also small for average percent minority, with the average percentage being 13.1% for low-documentation loans in 2005 and the estimated discontinuity around the cutoff of 0.3%, and for median household value, with an average across the FICO scores of $143,499 and an estimated difference of $1,215 (0.9%). Overall, our results indicate that observable characteristics of loans and borrowers are not different around the credit threshold.

V.D Performance of Loans Around Credit Threshold

We now focus on the performance of the loans that are originated close to the credit score threshold. Note that our analysis above suggests that there is no difference in terms of observable hard information about contract terms or about borrower demographic characteristics around

[17]These results are available upon request from the authors.

[18]Of course, since the census data is at the zip code level, we are to some extent smoothing our distributions. We note, however, that when we conduct our analysis on differences in number of loans (from Section V.B), aggregated at the zip code level, we still find jumps around the credit threshold within each individual zip code.

the credit score thresholds. Nevertheless, we will control for these differences when evaluating subsequent performance of loan. The notion is that if there is any difference in the performance of the loans above and below the credit threshold, it can be attributed to differences in unobservable soft information about the loans. In our analysis, we measure performance of a loan by tracking whether or not it became delinquent. We call a loan delinquent if it is in any of the following states: (a) payments on the loan are 60+ days late; (b) the loan is in foreclosure; or (c) the loan is real estate owned (REO), i.e. the bank has re-taken possession of the home. Accordingly, we code a dummy variable *Delinquency* that takes a value 1 if the loan is in any of the states listed above.[19]

We estimate the differences in default rates on either side of the cutoff using the same framework as equation (1), using the dollar-weighted fraction of loans defaulted within 10-15 months as the dependent variable. We collapse the data into two-point FICO bins instead of individual bins for precision and estimate third-order polynomials on either side of the threshold for each year. By estimating the magnitude of β in each year separately, we ensure that no one cohort (or "vintage") of loans is driving our results. As shown in Figure 6, the low documentation loans exhibit discontinuities in default rates at the FICO score of 620. In particular for loans originated in 2004, the estimate of β is .01 (t-stat=2.57), and the mean delinquency rate is .05, suggesting a 20% increase in defaults to the right of the credit score cutoff. Similarly, in 2005, the estimated size of the jump is .016 (t-stat=2.82), the mean delinquency rate for all FICO bins is .082, which is again a 20% increase in defaults around the FICO score threshold.

To show how delinquency rates evolve over the age of the loan, in Figure 7 we plot the delinquency rates of 620^+ and 620^- for low documentation loans (dollar weighted) by loan age for time periods after 2000. As discussed earlier, we restrict our analysis to about two years after the loan has been originated. As can be seen from the figure, the differences in the delinquency rates are stark. Contrary to what one might expect, around the credit threshold we find that loans of higher credit scores actually default *more often* than lower credit loans in the post-2000 period. The differences begin around four months after the loans have been originated and persist up to two years.[20] Differences in default rates also seem quite large in terms of magnitudes. Those with a credit score of 620^- are about 20% less likely to default after a year as compared to loans of credit score 620^+ for the post-2000 period.

[19] Estimates from various industry reports (e.g., Deutsche Bank report in November 2007) suggest that this is a sensible measure. Using data from LoanPerformance, these reports find that about 80% of the 60+ loans roll over to 90+ and another 90% roll over from 90+ to foreclosure in the subprime market. Our results are invariant to using other definitions of delinquency.

[20] The fact that we find no delinquencies early on in the duration of the loan is not surprising, given that originators are required to take back loans on their books if the loans default within three months.

To examine this more precisely, we estimate logit regressions of the following form:

$$Y_i = \Phi\Big(\; \alpha + \beta T_i + \gamma_1 X_i + +\delta_1 T_i * X_i + \mu_t + \epsilon_i \; \Big). \tag{2}$$

The dependent variable is the dummy variable (*Delinquency*) for loan i that takes a value of 1 or 0 as defined above. T takes the value 1 if FICO is between 621 and 625, and 0 if it is between 615 and 619 for low documentation loans. This restricts the analysis to the immediate vicinity of the cutoffs. Controls include FICO scores, the interest rate on the loan, loan-to-value ratio, borrower demographic variables, squares and cubic polynomials of these variables as well as interaction of these variables with T. We also include a dummy variable for the type of loan (adjustable or fixed rate mortgage). We control for age of the loan by including three dummy variables – that take a value of 1 if the month since origination is between 0-10, 11-20 and more than 20 months respectively. Year of origination fixed effects are included in the estimation and standard errors are clustered at the loan level.

As can be seen from the logit coefficients in Table V, results from this regression are qualitatively similar to those reported in the figures. In particular, we find that β is positive when we estimate the regressions for low documentation loans in the post-2000 period. The economic magnitudes are similar to those in the figures as well. For instance, keeping all other variables at their mean level, low documentation loans with credit score of 620^- are about 20-25% less likely to default after a year as compared to low documentation loans of credit score 620^+ for post-2000 period. These are large magnitudes – for instance, note that the mean delinquency rate for low documentation loans post-2000 is around 4.45%; the economic magnitude of the effects in Column (2) suggest that the difference in the absolute delinquency rate between loans around the credit threshold is around 1% for low documentation loans.

Overall, we find that even after controlling for all observable characteristics of the loan contracts or borrowers, loans of higher FICO scores perform *worse* in the future around the credit threshold. We have argued so far that ease of securitization at higher credit scores around the thresholds leads to lax screening by lenders. We now discuss more evidence in both the cross-section and time-series that is consistent with this explanation.

VI Testing Alternative Explanations

VI.A Selection and Cost of Capital

So far we have not considered if our results can be explained by arguments that involve borrower selection on observables. In particular, contract terms offered to borrowers above the credit threshold might attract riskier pool of borrowers. If this were the case, it would not be surprising if the loans above the credit threshold perform worse than those below it. In order to examine this hypothesis, we first examine the distribution of interest rates and loan-to-value ratios of contracts offered around 620 for low documentation loans. Figure 8A depicts the Epanechnikov kernel density of the interest rate on low documentation loans in the year 2004 for two FICO groups – 620^- (615-619) and 620^+ (621-625).[21] The distribution of interest rates observed in the two groups lie directly on top of one another. A Kolmogorov-Smirnov test for equality of distribution functions cannot be rejected at the 1% level. Similarly, Figure 8B depicts density of LTV ratios on low documentation loans in the year 2004 for 620^- and 620^+ groups. Again, a Kolmogorov-Smirnov test for equality of distribution functions cannot be rejected at the 1% level.[22] This suggests that selection based on observables is unlikely to explain our results.

Second, an argument might be made that banks screen similarly around the credit threshold but are able to sell portfolio of loans above and below the threshold to investors with different risk tolerance. If this were the case, it could potentially explain our results in Section V.D. Two facts however suggest that this cannot be the case. First, since all the loans in our sample are securitized, our results on performance on loans around the credit threshold are *conditional* on securitization. Second, securitized loans are sold to investors in pools which contains a mix of loans from the entire credit score spectrum. As a result, it is difficult to argue that loans of 620^- are purchased by different investors as compared to loans of 620^+. Thus a cost of capital argument cannot explain our results.

VI.B Composition of Loans From Different Types of Lenders

Lenders in the subprime market (e.g., banks/thrifts such as Countrywide vs. independents like Ameriquest) differ in the amount of supervision they face. In particular, banks/thrifts and their subsidiaries undergo rigorous examinations from their regulators (Office of Thrift Supervision, Federal Deposit Insurance Corporation and the Fed). One might therefore think that banks/thrifts would screen more actively than independents. Could it be that our results are an artifact of banks operating primarily below the credit threshold while independents operating primarily above the threshold? If this is the case, we should find that the performance results

[21] The bandwidth for the density estimation is selected using the plug-in formula of Sheather and Jones (1991).

[22] We also examine prices and terms (subordination levels) at which the pool of loans are sold to eventual investors (e.g., hedge funds). In particular, if adverse selection is severe above credit thresholds, we should find that the pool of loans sold above the credit threshold should have different pricing terms compared to pool of loans just below the credit threshold. However, we find no differences in prices and subordination levels of a pool of primarily low documentation loans around credit threshold of 620. However, we should interpret these findings with caution since by pooling loans one biases against finding any differences. In other words, since loans of exact 620^- and 620^+ form small parts of the sample, we are less likely to see differences in prices of the pools even if investors did price loans of exact 620^- and 620^+ differently.

disappear once we restrict our sample only to loans originated by either banks or independents. In this subsection, we examine if this is the case.

To conduct this test, we need to classify lenders in our sample into banks, thrifts, subsidiaries of banks/thrifts, and independent lenders. It is tough to exactly know the identity of all the lenders in the database since many of the names are abbreviated. In order to ensure that we are able to cover a majority of our sample, we focus on classifying the top 50 lenders (by origination volume) across the years in our sample period. We obtain the list from the publication 'Inside B&C mortgage' and matched the names in the list of top lenders to names of lenders in our sample (including abbreviations). Subsequently, we classify the lenders into two categories – banks (banks, subsidiaries, thrifts) and independents. Note that with this match we are able to cover about 75% of the entire sample.

In Panel A of Table VI, we estimate equation (1) separately for two types of lenders to quantify discontinuities in number of loans. We estimate the regressions after pooling data from 2001 to 2006 and use time fixed effects. As is shown in Columns (1) and (3), we find an increase in the number of low documentation loans – i.e., the coefficient on the indicator (T) is significant. Both banks and independents experience about a 90% increase in number of low documentation loans securitized above the credit threshold of 620 as compared to below it.

Next, we estimate differential delinquencies for loans securitized by different types of lenders and present the results in Column (2) and Column (4) of the two panels. Specifically, we use equation (2) for estimating the probabilities of delinquency separately for two types of lenders. Our results suggest that loans securitized by both banks and independents experience more delinquencies at the higher credit threshold. Keeping all other variables at their mean level, low documentation loans securitized by banks with credit score of 620^{-} are about 20% less likely to default after a year as compared to low documentation loans securitized by banks with credit score 620^{+}. Interestingly, this effect is similar for independent firms as well. This suggests that it is unlikely that our results are driven by differences in the composition of lenders around the thresholds.

VI.C Composition of Loans From Different Types of States

Lenders' screening incentives likely vary depending on the state in which they originate the loan, and would be weakest in states where they are most likely to recover collateral relatively easily, should the borrower default. Following the intuition of the "lazy banks" model of Manove, Padilla and Pagano (2001), screening and collateral promised to the lender can serve as substitutes. In the U.S., many states protect borrowers by imposing restrictions on the foreclosure process. This imposes costs on lenders and therefore should ex-ante lead to more screening by the lenders. If, on the other hand, certain states make collateral easy to possess – through creditor-friendly foreclosure policy for instance – the screening incentives of lenders in these states may be affected more. Therefore, our results might be driven by differences in the composition of loans from different states around the cutoff. Accordingly, we follow Pence (2003) and classify states as being lender friendly or not. Subsequently, we estimate the main regressions for the two groups.

In Panel B of Table VI, we estimate equation (1) separately for two types of states to quantify discontinuities in the number of loans. We estimate the regressions after pooling data from 2001 to 2006 and use time fixed effects. As is shown in Columns (1) and (3), we find an increase in the number of low documentation loans for both types of states. For instance, both types of states experience a 110% increase in number of low documentation loans securitized above the credit threshold of 620 as compared to below it.

We also estimate differential delinquencies for loans securitized by different state groups and present the results in Columns (2) and (4). Specifically, we use equation (2) for estimating the probabilities of delinquency separately for the two groups of states. Our results suggest that loans securitized by lenders in both categories experience more delinquencies at the higher credit threshold. However, the effects are stronger in creditor-friendly states. Keeping all other variables at their mean level, low documentation loans securitized by lenders in creditor-friendly states with credit score of 620^{-} are about 20% less likely to default after a year as compared to loans with credit score 620^{+}. However, loans securitized in less-friendly states have a smaller difference in the probability of default around the credit thresholds (and statistically insignificant). Overall, the results in this section suggest that our results are not driven by differences in the composition of loans from different states around the cutoff. We do, however, find support for the argument that screening incentives of lenders are weakened more in states where foreclosure laws are lender friendly.

VI.D Manipulation of Credit Scores: Evidence From a Natural Experiment

If FICO scores could be manipulated, lower quality borrowers might artificially appear at higher credit scores and that might explain our findings. Note that as per the rating agency (Fair Isaac), it is difficult to strategically manipulate one's FICO score in a targeted manner. Nevertheless, to examine this alternative more closely, we exploit a natural experiment that relies on the argument that FICO scores tend to be quite sticky (www.myfico.com) – i.e., it takes relatively long periods of time (more than 3 to 6 months) to improve credit scores. The natural experiment involves passing of anti-predatory laws in two states which reduced the securitization in the subprime market drastically. Additionally, subsequent to protests by market participants, the laws were amended substantially and the market reverted to pre-predatory law levels. We exploit the time series variation in securitization likelihood in the two states to examine how long it takes for the main effects to appear.

In October 2002, the Georgia Fair Lending Act (GFLA) went into effect, imposing anti-predatory lending restrictions which at the time were considered the toughest in the United States. The law allowed for unlimited punitive damages when lenders did not comply with the provisions and that liability extended to holders in due course. Once GFLA was enacted, the market response was swift. Fitch, Moodys, and S&P refused to rate securities that included Georgia loans. Fannie Mae and Freddie Mac announced that in January 2003 it would no longer purchase high-cost home loans made in Georgia. All of this led many lenders to pull out of the market. In effect, the demand for securitization of mortgage loans from Georgia also fell drastically during the same period. In response to these actions, the Georgia Legislature amended GLFA in early 2003. The amendments removed many of the GFLAs ambiguities and eliminated covered loans. Subsequent to April 2003, the market revived in Georgia. Similarly, New Jersey enacted its law, the New Jersey Homeownership Security Act of 2002. Many of the provisions were similar to the Georgia law. As in Georgia, lenders and ratings agencies expressed concerns when the law was passed and decided to substantially reduce the number of loans that were securitized in these markets. The Act was later amended in June 2004 in a way that relaxed requirements and eased lenders' concerns.

Our experimental design is to look at the number of loans securitized and the performance of loans above and below the credit threshold in both Georgia and New Jersey during the period when the securitization market was affected and compare it with period before the law was passed as well as with the period after the law was amended. We also use time fixed effects to control for any macro factors besides the law. Our empirical strategy uses equations (1) and (2) with an additional dummy variable that captures whether or not the law is in effect (*NoLaw*). The expectation is that passage of the anti-predatory law and its subsequent impact on the demand for securitization would have more of an effect at 620^{+} for low documentation loans relative to loans at 620^{-}. Specifically, we should find little differences in origination at 620^{+} relative to 620^{-} and between performance of loans during the period when the laws are in effect. However, we expect our results of Section V to appear in the period before the law was passed as well as in the period after the law was amended.

Our results are striking. Panel A of Table VII suggests that the number of loans securitized around the credit thresholds fell by around 80% at 620^{+} relative to 620^{-} during the period when the law was passed in Georgia and New Jersey. This effectively nullifies any jumps at 620 for low documentation loans that we documented earlier. We find results similar to those reported in the main tests during the period before the law was passed as well as during the period when the law was amended. Columns (1) and (2) of Panel B show that the default rates for loans above the credit threshold are higher than that of loans below the credit threshold in both Georgia and New Jersey *only* when the law was either not passed or was amended. In absolute terms, the magnitudes are large and are similar to those reported earlier (around 1% in absolute delinquency rates for low documentation loans). In addition, during the period when the law is in effect, we find that default rates for loans above the credit threshold is *lower* than loans below the credit threshold. Restricting our analysis to loans originated within six months after the laws were reversed, Columns (3) and (4) show that the reversal has immediate effects on the performance of the loans that are securitized. Overall, this evidence rules out manipulation in credit scores as an alternative explanation for our results.

VI.E Additional Robustness

VI.E.1 Full Documentation Credit Threshold

The results presented above are for low documentation loans, which necessarily have an unobserved component of borrowers' creditworthiness. In the full documentation loan market, on the other hand, there is no omission of hard information on the borrower's ability to repay. In this market, we identify a credit threshold at the FICO score of 600, the score that Fair Isaac (and the three credit repositories) advises lenders as a bottom cutoff for low risk borrowers. They note "...anything below 600 is considered someone who probably has credit problems that need to be addressed..."(see www.myfico.com). Similarly Fannie Mae in its guidelines notes "...a borrower with credit score of 600 or less has a high primary risk..." (see www.allregs.com/efnma/doc/). The Consumer Federation of America along with Fair Isaac (survey report in March 2005) suggests that "...FICO credit scores range from 300-850, and a score above 700 indicates relatively low credit risk, while scores below 600 indicate relatively high risk which could make it harder to get credit or lead to higher loan rates." Einav, Jenkins and Levin (2007) make a similar observation when they note that "...a FICO score above 600 [is] a typical cut-off for obtaining a standard bank loan."

Figure 9 reveals that there is a substantial increase in the number of full documentation loans above the credit threshold of 600. This pattern is consistent with the notion that lenders are more willing to securitize at a lower credit threshold (600 vs. 620) for full documentation loans since there is less uncertainty about these borrowers relative to those who provide less documentation. The magnitudes are again large – around 100% higher at 600^+ than at 600^- in 2004 – for full documentation loans. In Panel A of Table VIII, we estimate regressions similar to equation (1) and find the coefficient estimate is also significant at 1% and is on average around 100% (from 80 to 141%) higher for 600^+ as compared to 600^- for post-2000 loans. We repeated a similar analysis for loan characteristics (LTV and interest rates) and borrower demographics and find no differences for full documentation loans above and below the credit score of 600. Panels B and C present the estimates from the regressions.

Interestingly, we find that for full documentation loans, those with credit scores of 600^{-} (FICO between 595 and 599) are about as likely to default after a year as compared to loans of credit score 600^{+} (FICO between 601 and 605) for the post-2000 period. Both Figures 10 and 11 and results in Panel D support this conjecture. Following the methodology used in Figures 6 and 7, we show the default rates annually across the FICO distribution (Figure 10) and across the age of the loans (Figure 11). The estimated effects of the ad-hoc rule on defaults are negligible in all specifications. The absence of differences in default rates around the credit threshold, while maintaining the same magnitude of the jump in the number of loans, is consistent with the notion that the pattern of delinquencies around the low-documentation threshold are primarily due to soft information of the borrower. With so much information collected by the lender for full documentation loans, there is less soft information that can be gathered. Consequently, for full documentation loans there is no difference in how the loans perform subsequently after hard information has been controlled for. Put another way, these results show that transparency (i.e., more hard information with full documentation) reduces moral hazard in the subprime market.

VI.E.2 Other Checks

Another potential concern is related to whether there is indeed differential securitization at the credit threshold. Our results in Section VI.D alleviate these concerns since they show that our results disappear when ease of securitization above the credit threshold goes down during a period of strict enforcement of predatory lending laws. Additionally, we conduct several falsification tests, repeating our analysis at other credit scores where there is no jump in securitization. As an example, we plot the delinquency rates of 620^{+} and 620^{-} for full documentation loans (2001-2006) in Figure 12. In sharp contrast to results reported in Section V.D, the higher credit score bucket defaults *less* than the lower credit score bucket.

Finally, we also observe jumps in other parts of the distribution as other ad-hoc cutoffs have appeared in the market in the past three years (e.g., 600 for low documentation and 580 for full documentation loans in 2005 and 2006). A possible reason for this is the dramatic house price appreciation from 2001 to 2003 which made refinancing easy and kept defaults in the subprime market relatively low. Consequently, investor demand for loans with lower credit scores might have increased, despite the written guidelines of agencies and lenders' statements. The exact reasons for the appearance of these alternative standards is worthy of further investigation. For the purpose of this paper, we remain agnostic about why these other cutoffs have appeared.

Nevertheless, we conducted our analysis at these thresholds and find results for delinquencies that are smaller in magnitude but consistent with those reported for the predominant cutoff.

VII Conclusion

The goal of this paper is to empirically investigate whether securitization had an adverse effect on the ex-ante screening activity of banks. We exploit a specific "rule of thumb" in the subprime lending market to generate an instrument for ease of securitization. Comparing characteristics of the loan market above and below the ad-hoc credit threshold, we show that securitization does indeed weaken the screening incentives of financial intermediaries. These results should be interpreted in a local average treatment on treated framework, as the estimated coefficients are only valid in the immediate vicinity of the credit score threshold. An 80% increase in securitization volume is on average associated with about a 20% increase in defaults. These defaults are being driven by characteristics of the loan or the borrower that are unobservable to both the researchers and the securities market. That we find any effect on default behavior in one portfolio compared to another with virtually identical risk profiles, demographic characteristics, and loan terms suggests that the ease of securitization may have a direct impact on incentives elsewhere in the subprime housing market, as well as in other securitized markets.

There are several broad implications of our paper. First, we empirically demonstrate the economic trade-off between liquidity and incentives, a core feature of an extensive theoretical literature in financial contracting and corporate governance. The results underscore the role of illiquidity in preserving banks' willingness to adequately assess borrowers' creditworthiness (see also Mian and Sufi (2008)).

Second, it is important to note that we refrain from making any welfare claims. We believe securitization is an important innovation and has several merits. It is often asserted that securitization improves the efficiency of credit markets. This is true only to the extent that diversifying risk outweighs the dangers in increasing the distance between borrowers and investors in a particularly non-transparent manner. The benefits of securitization are limited by information loss, and in particular the costs we document in the paper. More generally, what types of credit products should be securitized? Our conjecture is that the answer depends crucially on the information structure of the particular loans. Loans with more "hard" information are likely to benefit from securitization as compared to loans that involve "soft" information. A careful investigation of this question is a promising area for future research.

Finally, our findings caution against policy that emphasizes excessive reliance on default models. The use of default models to predict and manage risk has become widespread in recent years and is also one of the key features of the Basel II Accord that is slated for implementation soon. The recent subprime crisis has demonstrated that these default models have mispriced risk and therefore implementation of Basel II may need to re-examined. Our research suggests that these pricing models put excessive weight on an individual credit score, and ignore essential elements of strategic behavior on the part of lenders that are likely to be important. Incorporating these strategic elements into default models, although challenging, is another important direction for future research.

References

1. Aghion, Philippe ; Bolton, Patrick and Tirole, Jean (2004), "Exit Options in Corporate Finance: Liquidity versus Incentives", *Review of Finance*, 8, 327-353.

2. Bhide, Amar (1993), "The Hidden Costs of Stock Market Liquidity", *Journal of Financial Economics*, 34, 31-51.

3. Bostic, Raphael; Calem, Paul, and Wachter, Susan, (2005) "Hitting the Wall: Credit as an Impediment to Homeownership", *Retsinas and Belsky, eds., Building Assets, Building Credit: Creating Wealth in Low-Income Communities*, Brookings Institution Press.

4. Calem, Paul; Hershaff, Jonathan and Wachter, Susan (2004), "Neighborhood Patterns of Subprime Lending: Evidence from Disparate Cities", *Working paper.*

5. Canner, Glenn and Passmore, Wayne (1999), "The Role of Specialized Lenders in Extending Mortgages to Lower-Income and Minority Homebuyers", *Federal Reserve Bulletin*, 709-723.

6. Card, David; Mas, Alexandre and Rothstein, Jesse (2007), "Tipping and the Dynamics of Segregation in Neighborhoods and Schools", *Working Paper, UC Berkeley Center for Labor Economics.*

7. Chomsisengphet, Souphala and Pennington-Cross, Anthony (2006), "The Evolution of the Subprime Mortgage Market", *Federal Reserve Bank of St. Louis Review*, 88:1, 31-56.

8. Coffee, John, (1991), "Liquidity Versus Control: The Institutional Investor as Corporate Monitor", *Columbia Law Review*, 91, 1277-1368.

9. DeMarzo, Peter and Urosevic, Branko (2006), "Optimal Trading and Asset Pricing With A Large Shareholder", *Journal of Political Economy*, 114, 774-815.

10. Demyanyk, Yuliya and Van Hemert, Otto (2007), "Understanding the Subprime Mortgage Crisis", *Working paper.*

11. Diamond, Douglas (1984), "Financial Intermediation And Delegated Monitoring", *Review of Economic Studies*, 51, 393-414.

12. Diamond, Douglas and Rajan, Raghuram (2003), "Liquidity Risk, Liquidity Creation and Financial Fragility: A Theory Of Banking", *Journal of Political Economy*, 109, 287-327.

13. DiNardo, John and Lee, David S. (2004), "Economic Impacts of New Unionization On Private Sector Employers: 1984-2001", *Quarterly Journal of Economics*, 119, 1383-1441.

14. Drucker, Steven and Puri, Manju (2007), "On Loan Sales, Loan Contracting, and Lending Relationships", *Working Paper.*

15. Einav, Liran; Jenkins, Mark and Levin, Jonathan (2007), "Contract Pricing in Consumer Credit Markets", *Working Paper.*

16. Gramlich, Edward (2007), "Subprime Mortgages: America's Latest Boom and Bust", *The Urban Institute Press.*

17. Gorton, Gary and Pennacchi, George (1995), "Banks and Loan Sales: Marketing Nonmarketable Assets", *Journal of Monetary Economics*, 35, 389-411.

18. Holloway, Thomas; MacDonald, Gregor and Straka, John, (1993), "Credit Scores, Early-Payment Mortgage Defaults, and Mortgage Loan Performance", *Freddie Mac Working Paper*.

19. Holmstrom, Bengt and Tirole, Jean (1997), "Financial Intermediation, Loanable funds, and The Real Sector", *Quarterly Journal of Economics*, 52, 663-692.

20. Immergluck, Daniel and Wiles, Marti (1999), "Two Steps Back: The Dual Mortgage Market, Predatory Lending, and the Undoing of Community Development", *Chicago: The Woodstock Institute*.

21. Loutskina, Elena and Strahan, Philip (2007), "Securitization And The Declining Impact Of Bank Finance On Loan Supply: Evidence From Mortgage Acceptance Rates", *Working Paper*.

22. Lucas, Douglas; Goodman, Laurie and Fabozzi, Frank (2006), "Collateralized Debt Obligations: Structures and Analysis", *Wiley Finance*, Hoboken, New Jersey.

23. Manove, Michael; Padilla, Jorge and Pagano, Marco (2001), "Collateral vs. Project Screening: a Model of Lazy Banks", *RAND Journal of Economics*, 32, 726-744.

24. Maug, Ernst, (1998), "Large Shareholders as Monitors: Is There a Tradeoff between Liquidity and Control?", *Journal of Finance*, 53, 65-98.

25. Mian, Atif and Sufi, Amir (2008), "The Consequences of Mortgage Credit Expansion: Evidence from the 2007 Mortgage Default Crisis", *Working paper*.

26. Nichols, Joseph; Pennington-Cross, Anthony and Yezer, Anthony (2005), "Borrower Self-Selection, Underwriting Costs, and Subprime Mortgage Credit Supply", *Journal of Real Estate Finance and Economics*,30, 197-219.

27. Parlour, Christine and Plantin, Guillaume, (2007), "Loan Sales and Relationship Banking", *Journal of Finance*, forthcoming.

28. Pence, Karen (2003), "Foreclosing on Opportunity: State Laws and Mortgage Credit", *Federal Reserve Working Paper*.

29. Sheather, Simon and Jones, Chris, (1991), "A Reliable Data-Based Bandwidth Selection Method for Kernel Density Estimation", *Journal of the Royal Statistical Society*, 53, 683690.

30. Stiglitz, Joseph, (2007), "Houses of Cards", *The Guardian*, October 9, 2007.

31. Sufi, Amir (2006), "Information Asymmetry and Financing Arrangements: Evidence from Syndicated Loans", *Journal of Finance*, 62, 629-668.

32. Temkin, Kenneth; Johnson, Jennifer and Levy, Diane (2002), "Subprime Markets, The Role of GSEs and Risk-Based Pricing", Prepared for U.S. Department of Housing and Urban Development Office of Policy Development and Research.

Table I
Summary Statistics

This table reports the summary statistics of the key variables used in the analysis. Information on loans comes from LoanPerformance database. The LoanPerformance database captures all the data elements that are typically disclosed in non-agency deals. We use information on FICO scores, LTV (loan-to-value) ratio of the loan, information on the documentation reported by the borrower when taking the loan. Documentation is categorized as full, limited and no. Full documentation loans provide verification of income as well as assets of the borrower. Limited documentation provides no information about the income but does provide some information about the assets. No documentation loans provide no information about income or assets. We combine limited and no documentation loans and call them 'low documentation' loans. Using the information in the database, we restrict the loans in our sample to purchases of owner-occupied single-family residences, townhouses, or condominiums. We drop non-conventional properties, such as those that are FHA or VA insured or pledged properties, and also exclude buydown mortgages. We also restrict our sample to those loans with valid FICO scores. We report summary statistics for the period 1997 to 2006. Panel A reports the general statistics on number of loans, documentation, LTV and FICO scores across these years for the entire sample. Panels B and C report similar statistics as in Panel A (except documentation) for low and full documentation loans respectively. Panels D and E provide descriptive statistics on key variables used in the analysis.

Panel A: Entire Sample

Year	Number of Loans	% Low Documentation	Mean Loan-To-Value	Mean FICO
1997	24,067	24.9%	80.5	611
1998	60,094	23.0%	81.5	605
1999	104,847	19.2%	82.2	610
2000	116,778	23.5%	82.3	603
2001	136,483	26.0%	84.6	611
2002	162,501	32.8%	85.6	624
2003	318,866	38.9%	87.0	637
2004	610,753	40.8%	86.6	639
2005	793,725	43.4%	86.3	639
2006	614,820	44.0%	87.0	636

Panel B: Low Documentation Sample

Year	Number of Loans	Mean Loan-To-Value	Mean FICO
1997	5,990	76.2	632
1998	13,808	77.6	621
1999	20,167	79.3	627
2000	27,413	79.5	622
2001	35,427	81.4	630
2002	53,275	83.9	646
2003	124,039	85.2	657
2004	249,298	86.0	658
2005	344,308	85.5	659
2006	270,751	86.3	655

Panel C: Full Documentation Sample

Year	Number of Loans	Mean Loan-To-Value	Mean FICO
1997	18,077	81.9	604
1998	46,286	82.6	600
1999	84,680	82.9	605
2000	89,365	83.2	597
2001	101,056	85.7	604
2002	109,226	86.4	613
2003	194,827	88.1	624
2004	361,455	87.0	626
2005	449,417	86.9	623
2006	344,069	87.5	621

Panel D: Summary Statistics - Key Variables

	2000 and earlier			2001 and later		
	Mean	Median	Std. Dev.	Mean	Median	Std. Dev.
Average loan size ($000)	113.4	93	94.6	165.2	130.5	125.2
FICO score	606.14	605	61.77	635.63	633	53.85
Loan-to-Value ratio	81.95	80	11.2	86.47	80	9.87
Debt-to-Income ratio	19	10	20.3	29.6	38	19.8
Initial Interest Rate	10.19	10.28	2.03	8.27	7.95	1.84

Panel E: Summary Statistics - Other Variables

	2000 and earlier	2001 and later
FRM (%)	29	18.5
ARM (%)	61	51
Interest-only (%)	0.2	16.6
Balloon (%)	9.8	13.6
No documentation (%)	0.6	0.7
Limited documentation (%)	22.3	40.2
Full documentation (%)	77.2	59.1
Prepayment penalty (%)	72.7	73.6
First Time Purchase (%)	28	38
Refinance with no cash out (%)	58	53
Refinance with cash out (%)	14	8
Number of loans	305,786	2,637,148

Table II
Discontinuity in Number of Low Documentation Loans

This table reports estimates from a regression which uses the number of loans at each FICO score as the dependent variable. In order to estimate the discontinuity for each year, we collapse the number of loans at each FICO score and use an equation of type:

$$Y_i = \left(\alpha + \beta T_i + \theta f(FICO(i)) + \delta T_i * f(FICO(i)) + \epsilon_i \right),$$

where Y_i is the average number of loans at FICO score i and T_i is an indicator variable that takes a value 1 if the FICO score is more than 620 for low documentation loans and 0 otherwise. $f(FICO)$ and $T * f(FICO)$ are estimated as flexible seventh-order polynomials. In the case of low-documentation loans, one line is estimated from 620^- (T) to the left, and one line is estimated from 620^+ to the right. Information on loans comes from LoanPerformance database. We combine limited and no documentation loans and call them 'low documentation' loans. We restrict the loans in our sample to purchases of owner-occupied single-family residences, townhouses, or condominiums. We drop non-conventional properties, such as those that are FHA or VA insured or pledged properties, and also exclude buydown mortgages. We also restrict our sample to those loans with valid FICO scores. We estimate these regressions for FICO scores between 500 and 800 since that is where the bulk of the data is. Data is for the period 2001 to 2006. We report t-statistics in parentheses. ***, ** and * denote significance at 1%, 5% and 10% respectively.

Low Documentation Loans

Year	β	t-stat	Observations	R^2	Mean
2001	36.83**	(2.10)	299	0.96	117
2002	124.41***	(6.31)	299	0.98	177
2003	354.75***	(8.61)	299	0.98	413
2004	737.01***	(7.30)	299	0.98	831
2005	1,721.64***	(11.78)	299	0.99	1,148
2006	1,716.49***	(6.69)	299	0.97	903

Table III

Loans Characteristics around Discontinuity in Low Documentation Loans

This table reports the estimates of the regression that uses loan-to-value and interest rates as the dependent variable. In order to estimate the discontinuity, for each year, we calculate the mean interest rate and LTV ratio at each FICO score and use the equation of type:

$$Y_i = \Big(\ \alpha + \beta T_i + \theta f(FICO(i)) + \delta T_i * f(FICO(i)) + \epsilon_i \ \Big),$$

where Y_i is either mean loan-to-value or interest rates at FICO score i and T_i is an indicator variable that takes a value 1 if the FICO score is more than 620 for low documentation loans and 0 otherwise. $f(FICO)$ and $T * f(FICO)$ are estimated as flexible seventh-order polynomials. In the case of low-documentation loans, one line is estimated from 620^- (T) to the left, and one line is estimated from 620^+ to the right. Because the measures of the interest rate and LTV are estimated means, we weight each observation by the inverse of the variance of the estimate. Information on loans comes from LoanPerformance database. We combine limited and no documentation loans and call them 'low documentation' loans. We restrict the loans in our sample to purchases of owner-occupied single-family residences, townhouses, or condominiums. We drop non-conventional properties, such as those that are FHA or VA insured or pledged properties, and also exclude buydown mortgages. We also restrict our sample to those loans with valid FICO scores. We estimate these regressions for FICO scores between 500 and 800 since that is where the bulk of the data is. Data is for the period 2001 to 2006. We report t-statistics in parentheses. ***, ** and * denote significance at 1%, 5% and 10% respectively.

Low Documentation Loans

	Loan To Value					Interest Rate				
Year	β	t-stat	Obs.	R^2	Mean (%)	β	t-stat	Obs.	R^2	Mean (%)
2001	0.67	(0.93)	296	0.76	80.3	0.06	(0.59)	298	0.92	9.4
2002	1.53^{**}	(2.37)	299	0.91	82.6	0.15	(1.05)	299	0.89	8.9
2003	2.44^{***}	(4.27)	299	0.96	83.4	0.10	(1.50)	299	0.97	7.9
2004	0.30	(0.62)	299	0.96	84.5	0.03	(0.39)	299	0.97	7.8
2005	-0.33	(0.96)	299	0.95	84.1	-0.09	(1.74)	299	0.98	8.2
2006	-1.06^{***}	(2.53)	299	0.96	84.8	-0.21^{**}	(2.35)	299	0.98	9.2

Table IV
Borrower Demographics around Discontinuity in Low Documentation Loans

This table reports the estimates of the regression that uses demographic characteristics of borrowers as the dependent variable. In order to estimate the discontinuity, for each year, we calculate the mean demographic characteristics at each FICO score and use the equation of type:

$$Y_i = \Big(\alpha + \beta T_i + \theta f(FICO(i)) + \delta T_i * f(FICO(i)) + \epsilon_i \Big),$$

where Y_i is various borrower demographic characteristics at FICO score i and T_i is an indicator variable that takes a value 1 if the FICO score is more than 620 for low documentation loans and 0 otherwise. $f(FICO)$ and $T * f(FICO)$ are estimated as flexible seventh-order polynomials. In the case of low-documentation loans, one line is estimated from 620^- (T) to the left, and one line is estimated from 620^+ to the right. Because the measures of demographics are estimated means, we weight each observation by the inverse of the variance of the estimate. Information on loans comes from LoanPerformance database. We combine limited and no documentation loans and call them 'low documentation' loans. We restrict the loans in our sample to purchases of owner-occupied single-family residences, townhouses, or condominiums. We drop non-conventional properties, such as those that are FHA or VA insured or pledged properties, and also exclude buydown mortgages. We also restrict our sample to those loans with valid FICO scores. We estimate these regressions for FICO scores between 500 and 800 since that is where the bulk of the data is. We obtain the demographic variables from Census 2000, matched using the zip code of each loan. Panel A represents results using %black in the neighborhood of the borrower's zip code as the dependent variable, Panel B represents results using income in the neighborhood of the borrower's zip code as the dependent variable and Panel C represents results using median house size in the neighborhood of the borrower's zip code as the dependent variable. Data is for the period 2001 to 2006. We report t-statistics in parentheses. ***, ** and * denote significance at 1%, 5% and 10% respectively.

Panel A: Percent Black in Zip Code

Year	β	t-stat	Observations	R^2	Mean (%)
2001	1.54	(1.16)	297	0.79	11.2
2002	0.32	(0.28)	299	0.63	10.6
2003	1.70**	(2.54)	299	0.70	11.1
2004	0.42	(0.53)	299	0.72	12.2
2005	-0.50	(0.75)	299	0.69	13.1
2006	0.25	(0.26)	299	0.59	14.7

Panel B: Median Income in Zip Code

Year	β	t-stat	Observations	R^2	Mean (%)
2001	1,963.23**	(2.04)	297	0.33	49,873
2002	-197.21	(0.13)	299	0.35	50,109
2003	154.93	(0.23)	299	0.50	49,242
2004	699.90	(1.51)	299	0.46	48,221
2005	662.71	(1.08)	299	0.64	47,390
2006	-303.54	(0.34)	299	0.68	46,396

Panel C: Median House Value in Zip Code

Year	β	t-stat	Observations	R^2	Mean (%)
2001	3,943.30	(0.44)	297	0.66	163,151
2002	-599.72	(0.11)	299	0.79	165,049
2003	-1,594.51	(0.36)	299	0.89	160,592
2004	-2,420.01	(1.03)	299	0.91	150,679
2005	-342.04	(0.14)	299	0.93	143,499
2006	-3,446.06	(1.26)	299	0.92	138,556

Table V

Delinquencies in Low Documentation Loans around the Credit Threshold

This table reports the estimates of the regression that uses delinquency status of a loan in a given month from origination as the dependent variable. More specifically, we estimate the logit regression of the type:

$$Y_i = \Phi\Big(\alpha + \beta T_i + \gamma_1 X_i + +\delta_1 T_i * X_i + \mu_t + \epsilon_i \Big).$$

The dependent variable is a dummy variable (*Delinquency*) for loan i that takes a value 1 if the loan is in any of the following states – 60 days unpaid or more, real estate owned or in foreclosure and 0 otherwise. T for each loan takes a value 1 if the FICO of the borrower is between 621 and 625 and 0 if it between 615 and 619 for low documentation loans (reported as FICO≥620). Controls include FICO scores, interest rate on the loan, loan-to-value ratio, borrower demographic variables as well as interaction of these variables with T. We also include a dummy variable for the type of loan (adjustable or fixed rate mortgage). We control for age by including three dummy variables – τ^{0-10}, τ^{10-20} and τ^{20+} – that take a value 1 if the month since origination is between 0-10, 11-20 and more than 20 months but less than 25 months respectively, i.e., $j \in \{0-10, 11-20, 20+\}$. We obtain the demographic variables from Census 2000, matched using the zip code of each loan. Information on loans comes from LoanPerformance database. Limited documentation provides no information about the income but does provide some information about the assets. No documentation loans provide no information about income or assets. We combine limited and no documentation loans and call them 'low documentation' loans. We restrict the loans in our sample to purchases of owner-occupied single-family residences, townhouses, or condominiums. We drop non-conventional properties, such as those that are FHA or VA insured or pledged properties, and also exclude buydown mortgages. We also restrict our sample to those loans with valid FICO scores. Time fixed effects (μ_t) are used in all the regressions. We estimate these regressions for FICO scores between 500 and 800 since that is where the bulk of the data is. Data is for the period 2001 to 2006. We report t-statistics in parentheses. Standard errors in the regression are clustered at the loan level. ***, ** and * denote significance at 1%, 5% and 10% respectively.

Low Documentation Loans

	Pr(Delinquency)=1			
	(1)	(2)	(3)	(4)
FICO≥620	0.12***	1.07***	0.08**	0.48**
	(3.42)	(5.45)	(2.17)	(2.46)
τ^{0-10}	-0.45***	-0.44***	-0.55***	-0.55***
	(17.54)	(17.35)	(21.79)	(21.57)
τ^{10-20}	1.48***	1.48***	1.52***	1.52***
	(59.99)	(59.93)	(60.91)	(60.86)
τ^{20+}	1.81***	1.79***	2.04***	2.03***
	(53.09)	(52.84)	(59.22)	(58.82)
FICO≥620*τ^{0-10}	0.01	-0.01	0.01	-0.01
	(0.13)	(0.27)	(0.09)	(0.35)
FICO≥620*τ^{10-20}	-0.04	-0.04	-0.04	-0.04
	(1.32)	(1.29)	(1.42)	(1.40)
FICO≥620*τ^{20+}	-0.10**	-0.09**	-0.10**	-0.09**
	(2.51)	(2.35)	(2.53)	(2.26)
Observations	1,393,655	1,393,655	1,393,655	1,393,655
Pseudo R^2	0.088	0.091	0.109	0.116
Other Controls	Yes	Yes	Yes	Yes
T*Other Controls	No	Yes	No	Yes
Time Fixed Effects	No	No	Yes	Yes
Mean Delinquency (%)	4.45			

Table VI

Number of Loans and Delinquencies in Low Documentation Loans around the Credit Threshold: Cross-Sectional Evidence

This table reports the estimates of the regressions on the differences in number of loans and performance of loans around the credit thresholds. We use specifications similar to (1) to estimate the number of loans regressions and (2) to estimate delinquency regressions. In Panel A, we estimate the regressions separately for banks and independents. We focus on classifying the top 50 lenders (by origination volume) across the years in our sample period. We obtain the list from the 'Inside B&C mortgage' and matched the names in the list of top lenders to names of lenders in our sample (including abbreviations). We classify the lenders into two categories – banks (banks, subsidiaries, thrifts) and independents – and estimate regressions for each of these groups. In Panel B, we estimate regressions separately for states that are creditor friendly (in terms of foreclosure laws) and for those states that are not. We follow Pence (2003) and classify states as being lender friendly or not. Subsequently, we estimate the main regressions for the two groups. We obtain the demographic variables from Census 2000, matched using the zip code of each loan. Information on loans comes from LoanPerformance database. Limited documentation provides no information about the income but does provide some information about the assets. No documentation loans provide no information about income or assets. We combine limited and no documentation loans and call them 'low documentation' loans. We restrict the loans in our sample to purchases of owner-occupied single-family residences, townhouses, or condominiums. We drop non-conventional properties, such as those that are FHA or VA insured or pledged properties, and also exclude buydown mortgages. We also restrict our sample to those loans with valid FICO scores. Time fixed effects are used in all the regressions. We estimate these regressions for FICO scores between 500 and 800 since that is where the bulk of the data is. We conduct this analysis for the period 2001 to 2006. We report t-statistics in parentheses. Standard errors in the delinquency regression are clustered at the loan level. ***, ** and * denote significance at 1%, 5% and 10% respectively.

Panel A: Banks and Independents

	Banks		Independents	
	Number of Loans	Pr(Delinquency)=1	Number of Loans	Pr(Delinquency)=1
	(1)	(2)	(3)	(4)
FICO≥620	159.49***	.82**	387.74***	.67***
	(3.42)	(2.19)	(4.01)	(3.21)
Observations	1,793	225,455	1,793	797,472
R^2/ Pseudo R^2	0.73	0.11	0.74	0.11
Other Controls	Yes	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes
Mean	136.9	4.44%	336.8	4.10%

Panel B: Creditor Friendly and Unfriendly States

	Friendly States		Unfriendly States	
	Number of Loans	Pr(Delinquency)=1	Number of Loans	Pr(Delinquency)=1
	(1)	(2)	(3)	(4)
FICO≥620	545.03***	.56**	265.72***	.44
	(4.44)	(2.25)	(4.24)	(1.33)
Observations	1,793	848,597	1,794	456,525
R^2/ Pseudo R^2	0.74	0.11	0.73	0.11
Other Controls	Yes	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes
Mean	403.5	4.30%	196.1	4.20%

Table VII

Number of Loans and Delinquencies in Low Documentation Loans around the Credit Threshold: Evidence From A Natural Experiment

This table reports the estimates of the regressions on differences in number of loans and performance of loans around the credit thresholds. We use specifications similar to (1) in Panel A to estimate the number of loans regressions and (2) in Panel B to estimate delinquency regressions. We restrict our analysis to loans made in Georgia and New Jersey. We define a dummy variable (*NoLaw*) that takes a value 1 if the anti-predatory law was not passed in a given year or was amended and a value 0 during the time period when then the law was passed. Subsequently, we modify (1) and (2) and include this additional dummy variable. We obtain the demographic variables from Census 2000, matched using the zip code of each loan. Information on loans comes from LoanPerformance database. We restrict the loans in our sample to purchases of owner-occupied single-family residences, townhouses, or condominiums. We drop non-conventional properties, such as those that are FHA or VA insured or pledged properties, and also exclude buydown mortgages. We also restrict our sample to those loans with valid FICO scores. Also, we estimate these regressions for FICO scores between 500 and 800 – since that is where the bulk of the data is. We conduct this analysis for the period 2001 to 2006. We report t-statistics in parentheses. Standard errors in the delinquency regression are clustered at the loan level. *** and ** denote significance at 1% and 5% respectively.

Panel A: Number of Loans

	During Law (*NoLaw=0*)	Pre & Post Law (*NoLaw=1*)
	(1)	(2)
FICO≥620	10.71**	211.5***
	(2.30)	(5.29)
Observations	294	299
Other Controls	Yes	Yes
R^2	0.90	0.96
Mean	16	150

Panel B: Delinquencies

	Pr(Delinquency)=1			
	Entire Period 2001-2006		During Law and Six months After	
	(1)	(2)	(3)	(4)
FICO≥620	-.94**	-.91**	-1.04**	-1.02**
	(2.08)	(2.00)	(2.23)	(2.12)
FICO≥620*NoLaw	.91**	.88*	1.14**	1.13*
	(1.98)	(1.94)	(1.97)	(1.93)
Observations	109,536	109,536	14,883	14,883
Other Controls	Yes	Yes	Yes	Yes
Time Fixed Effects	No	Yes	No	Yes
Pseudo R^2	0.05	0.06	0.94	0.05
Mean Delinquency (%)	6.1		4.2	

Table VIII

Number of Loans, Loan Characteristics, Borrower Demographics and Delinquencies around the Credit Threshold for Full Documentation Loans

This table reports the estimates of the regressions on differences in number of loans, loan characteristics, borrower demographics and performance of loans around the credit threshold of 600 for full documentation loans. We use specifications similar to (1) in Panels A, B and C to estimate the number of loans, loan characteristics and borrower demographic regressions and (2) in Panel D to estimate delinquency regressions. Information on loans comes from LoanPerformance database. Data construction follows description in previous tables. Data is for the period 2001 to 2006. We report t-statistics in parentheses. ***, ** and * denote significance at 1%, 5% and 10% respectively.

Panel A: Full Documentation Loans

Year	β	t-stat	Observations	R^2	Mean
2001	306.85***	(5.70)	299	0.99	330
2002	378.49***	(9.33)	299	0.99	360
2003	780.72***	(11.73)	299	0.99	648
2004	1,629.82***	(8.91)	299	0.99	1205
2005	1,956.69***	(4.72)	299	0.98	1499
2006	2,399.48***	(6.97)	299	0.98	1148

Panel B: Loan Characteristics

	Loan To Value					Interest Rate				
Year	β	t-stat	Obs.	R^2	Mean (%)	β	t-stat	Obs.	R^2	Mean (%)
2001	0.820**	(2.09)	299	0.73	85.1	-0.097	(0.87)	299	0.97	9.5
2002	-0.203	(0.65)	299	0.86	85.8	-0.279***	(3.96)	299	0.97	8.6
2003	1.012***	(3.45)	299	0.95	86.9	-0.189***	(3.42)	299	0.99	7.7
2004	0.755**	(2.00)	299	0.96	86	-0.244***	(6.44)	299	0.99	7.3
2005	0.354	(1.82)	299	0.93	86.2	-0.308***	(5.72)	299	0.99	7.7
2006	-0.454	(1.96)	299	0.94	86.7	-0.437***	(9.93)	299	0.99	8.6

Panel C: Percent Black in Zip Code

Year	β	t-stat	Observations	R^2	Mean (%)
2001	2.32**	(2.03)	299	0.86	13.6
2002	-0.79	(1.00)	299	0.82	12.5
2003	0.40	(0.48)	299	0.87	12.5
2004	0.54	(0.96)	299	0.92	12.9
2005	-0.38	(0.85)	299	0.86	13.4
2006	-0.86	(1.40)	299	0.81	14.3

Panel D: Delinquencies

	Pr(Delinquency)=1			
	(1)	(2)	(3)	(4)
FICO≥600	-.06**	-.04	-.04*	-.02
	(2.30)	(0.28)	(1.65)	(0.15)
Observations	3,125,818	3,125,818	3,125,818	3,125,818
Pseudo R^2	0.073	0.075	0.081	0.084
Other Controls	Yes	Yes	Yes	Yes
T*Other Controls	No	Yes	No	Yes
Time Fixed Effects	No	No	Yes	Yes
Mean Delinquency (%)	4.54			

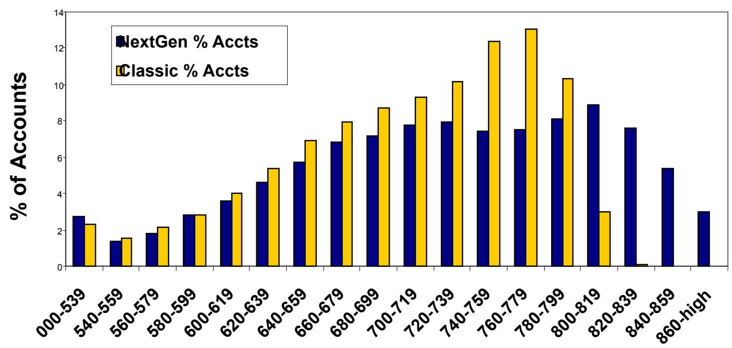

Figure 1A: National FICO Distribution

Figure 1A presents the FICO score distribution reported by Fair Isaac in the report published in 2005. Plots for other years (unreported) also look qualitatively similar to the distribution reported above. As can be seen from the histogram, there is no break in the distribution of accounts reported by Fair Isaac around the 620 credit threshold. This is true for both Fair Isaac scoring models (NextGen and Classic).

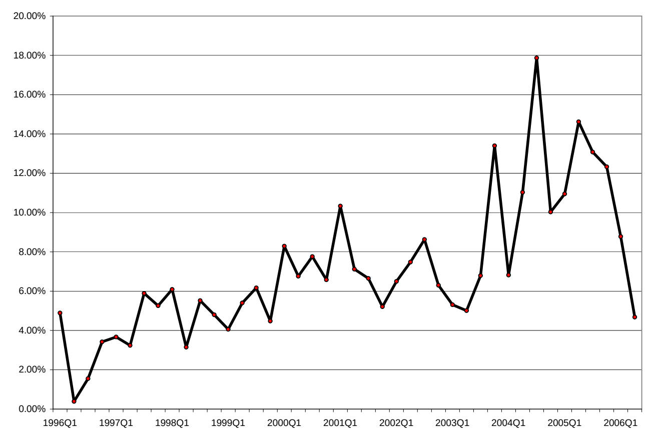

Figure 1B: House Price Appreciation

Figure 1B presents house price appreciation in the U.S during the period 1996 to 2006. The data comes from Office of Federal Housing Enterprise Oversight (OFHEO). The house price index tracked in the graph is based on average house price changes in repeat sales or refinancings of the same single-family properties and tracked on a quarterly basis. The data used by OFHEO is based on analysis of data obtained from Fannie Mae and Freddie Mac from more than 31 million repeat transactions over the past 31 years.

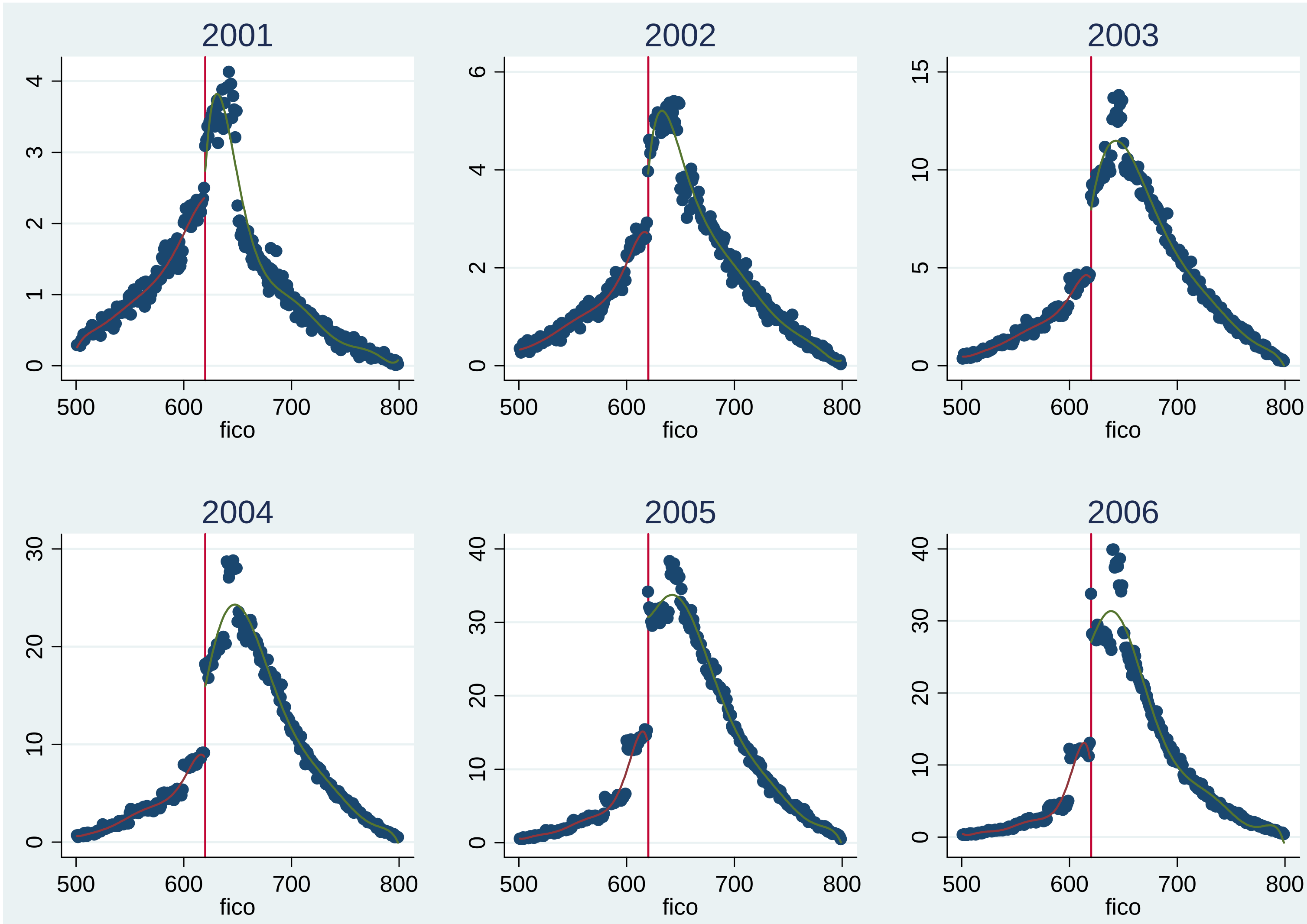

Figure 2: Number of Loans (Low Documentation)

Figure 2 presents the data for number of loans (in '00s) for low documentation loans. We plot the average number of loans at each FICO score between 500 and 800. We combine limited and no documentation loans and call them 'low documentation' loans. As can be seen from the graphs, there is a large increase in number of loans around the 620 credit threshold (i.e., more loans at 620^+ as compared to 620^-) from 2001 onwards. Data is for the period 2001 to 2006.

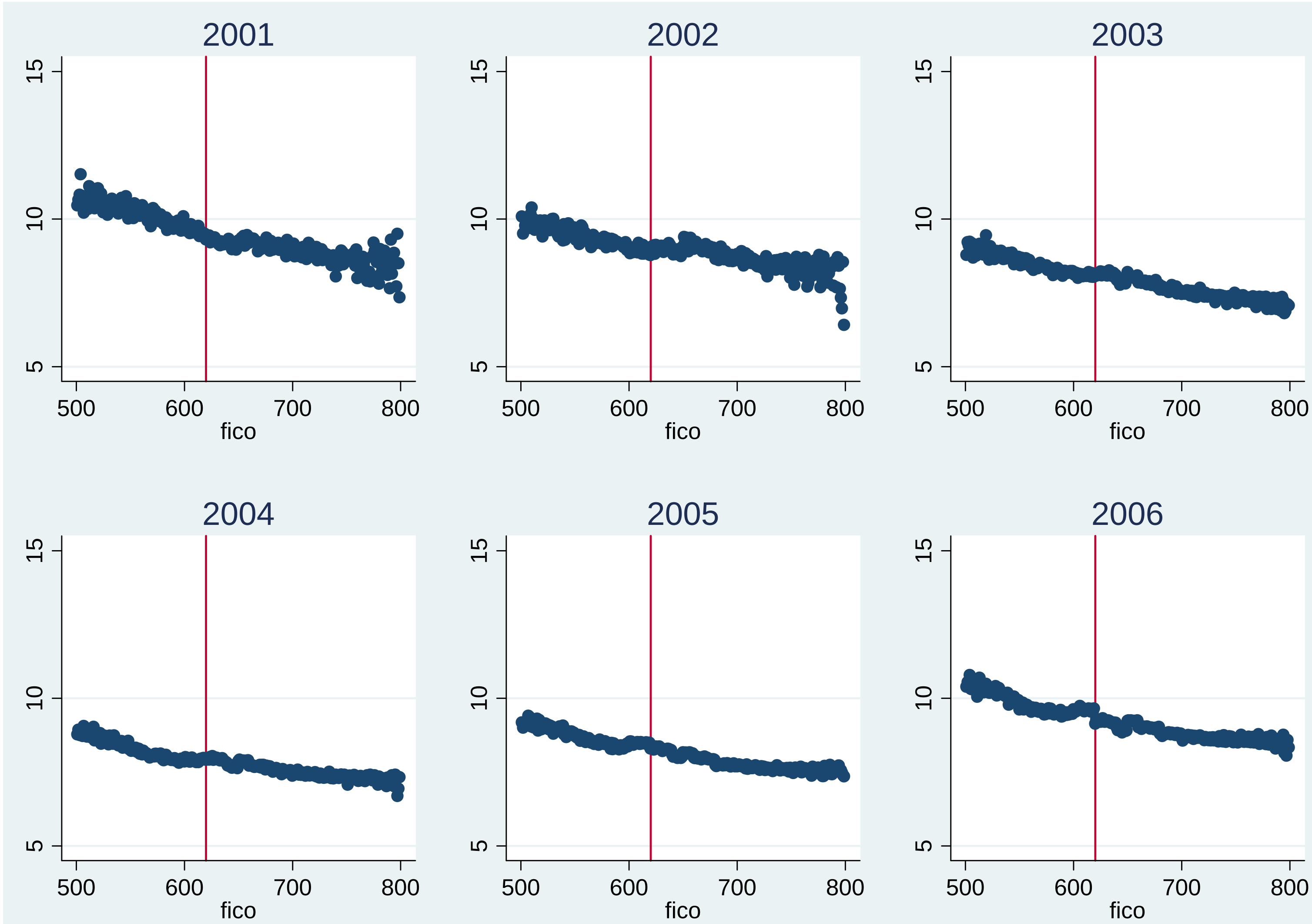

Figure 3: Interest Rates (Low Documentation)

Figure 3 presents the data for interest rate (in %) on low documentation loans. We plot average interest rates on loans at each FICO score between 500 and 800. We combine low and no documentation loans and call them 'low documentation' loans. As can be seen from the graphs, there is no change in interest rates around the 620 credit threshold (i.e., more loans at 620^{+} as compared to 620^{-}) from 2001 onwards. Data is for the period 2001 to 2006.

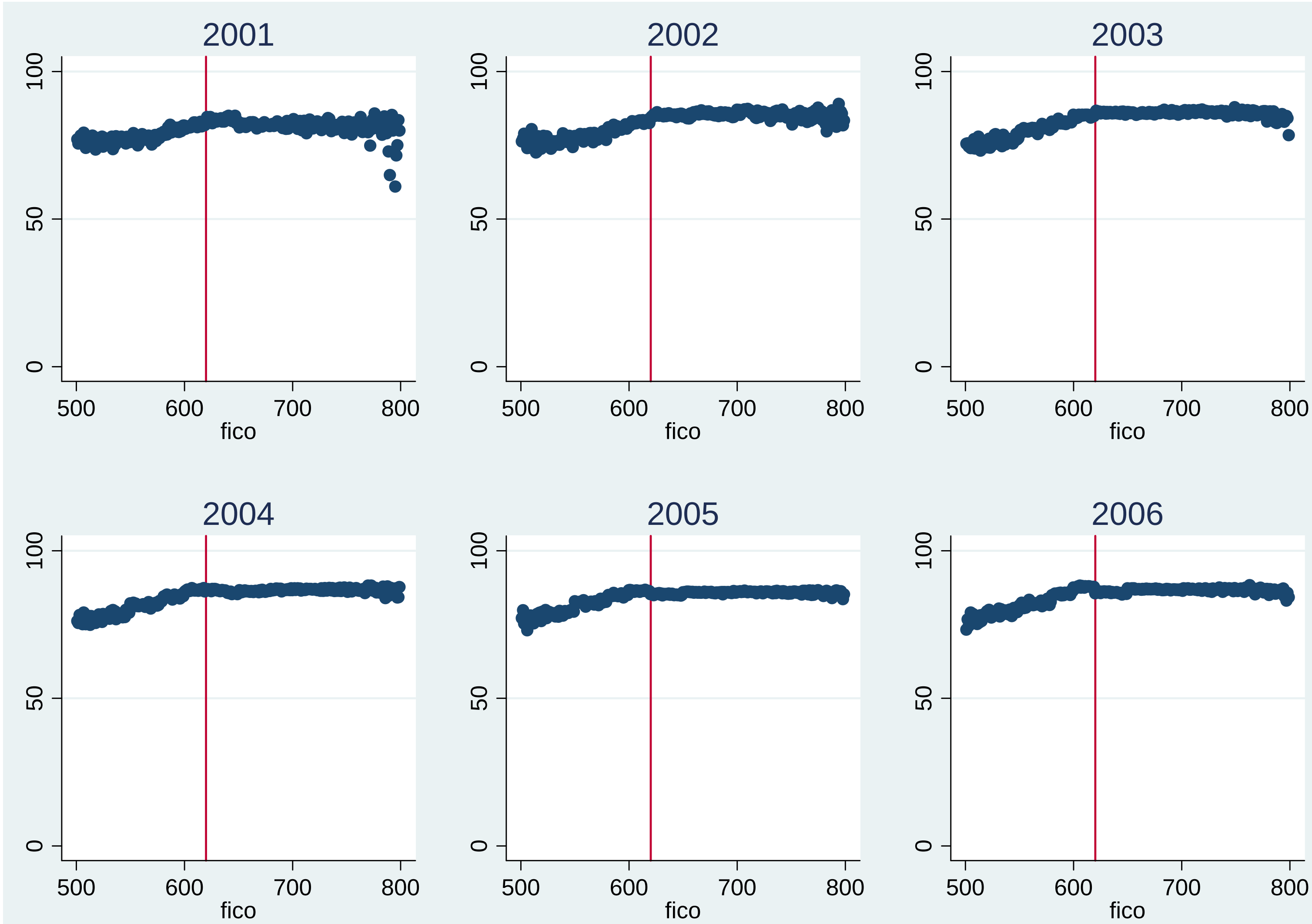

Figure 4: Loan-to-Value (Low Documentation)

Figure 4 presents the data for loan-to-value ratio (in %) on low documentation loans. We plot average loan-to-value ratios on loans at each FICO score between 500 and 800. We combine low and no documentation loans and call them 'low documentation' loans. As can be seen from the graphs, there is no change in loan-to-value around the 620 credit threshold (i.e., more loans at 620^{+} as compared to 620^{-}) from 2001 onwards. Data is for the period 2001 to 2006.

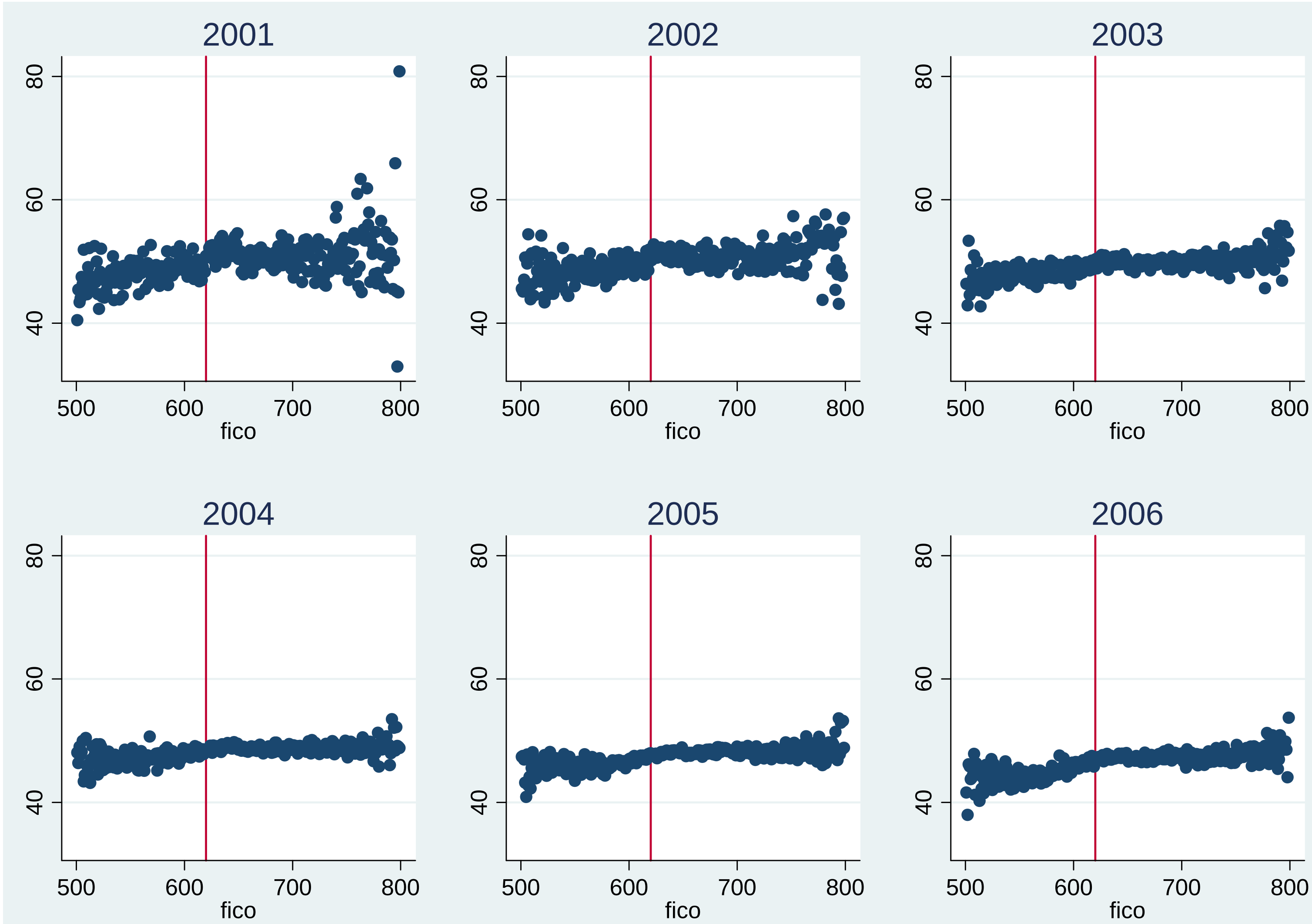

Figure 5: Median Household Income (Low Documentation)

Figure 5 presents median household income (in '000s) of zip codes in which loans are made at each FICO score between 500 and 800. We combine low and no documentation loans and call them 'low documentation' loans. As can be seen from the graphs, there is no change in median household income around the 620 credit threshold (i.e., more loans at 620^{+} as compared to 620^{-}) from 2001 onwards. We plotted similar distributions for average percent minorities taking loans, and average house size and find no differences around the credit thresholds. Data is for the period 2001 to 2006.

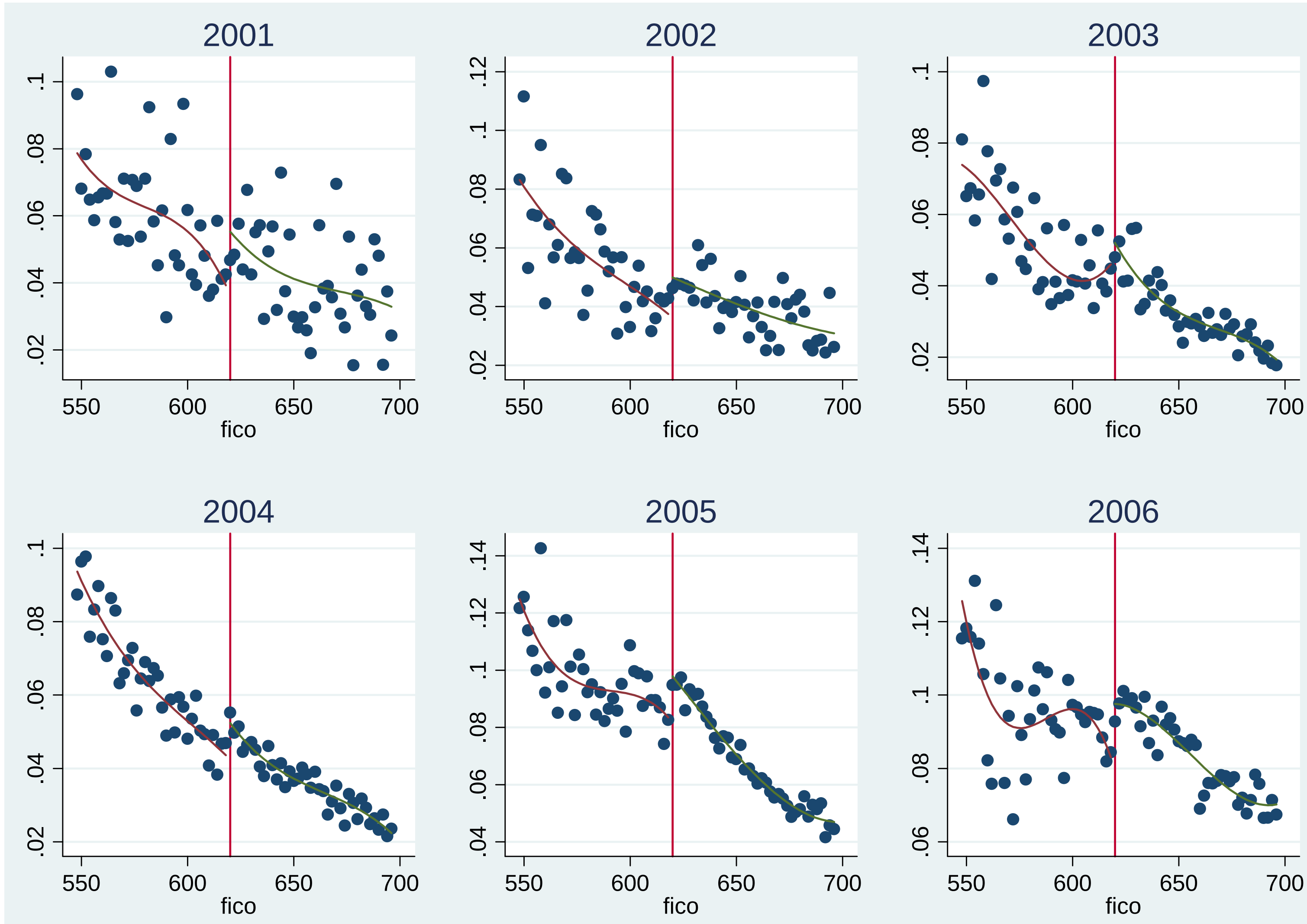

Figure 6: Annual Delinquencies for Low Documentation Loans

Figure 6 presents the data for actual percent of low documentation loans that became delinquent for 2001 to 2006. We plot the dollar weighted fraction of the pool that becomes delinquent for 2-point FICO bins between score of 550 and 700. The vertical line denotes the 620 cutoff, and a third order polynomial is fit to the data on either side of the threshold. Delinquencies are reported between 10-15 months for loans originated in all years except 2006. For loans originated in 2006, we take delinquencies between 5-10 months.

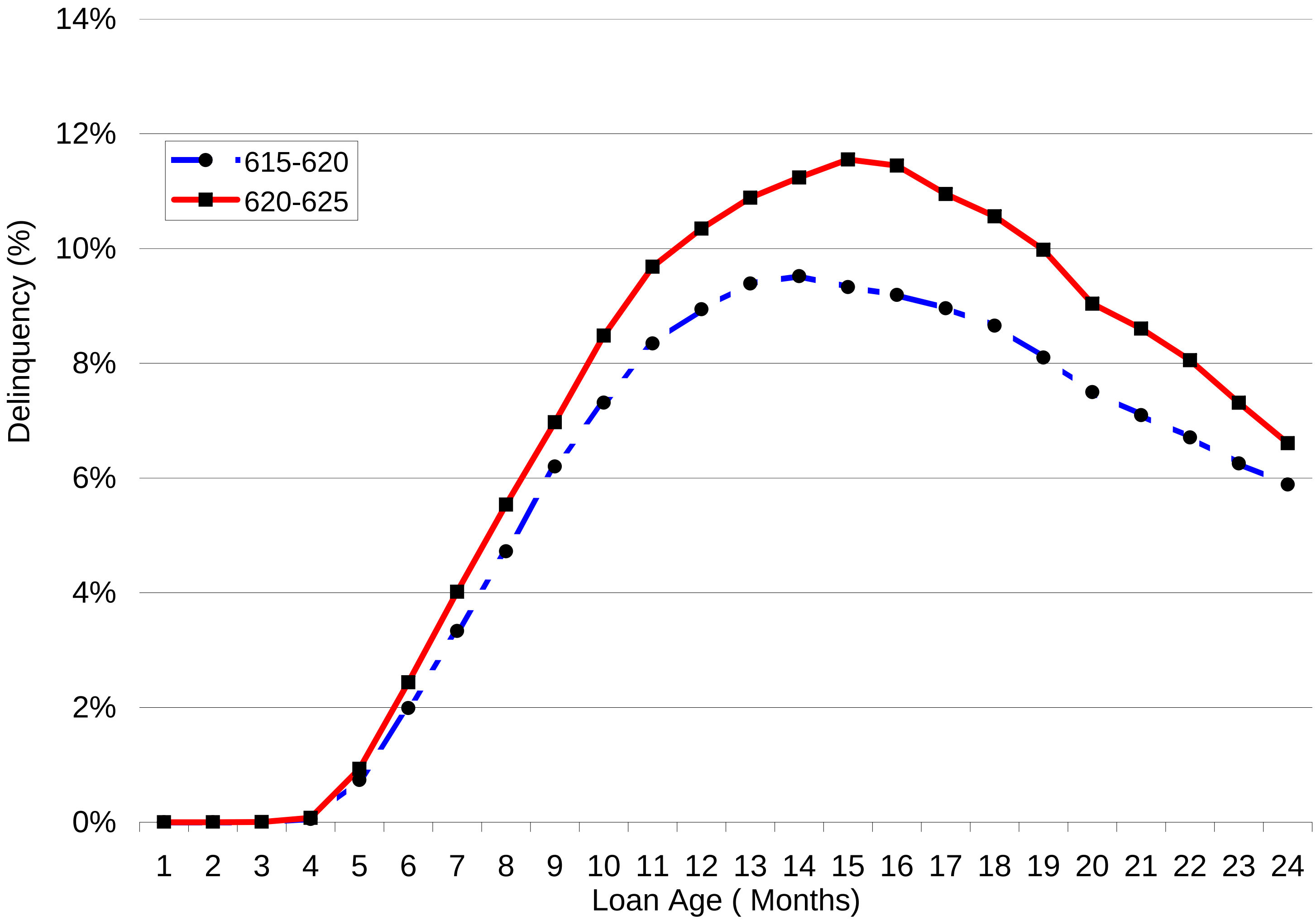

Figure 7: Delinquencies for Low Documentation Loans (2001-2006)

Figure 7 presents the data for actual percent of low documentation loans (dollar weighted) that became delinquent for 2001 to 2006. We track loans in two FICO buckets – 615-619 (620^-) in dotted blue and 621-625 (620^+) in red – from their origination date and plot the average loans that become delinquent each month after the origination date. As can be seen, the higher credit score bucket defaults *more* than the lower credit score bucket for post 2000 period. For brevity, we do not report plots separately for each year. The effects shown here in the pooled 2001-2006 plot are apparent in every year.

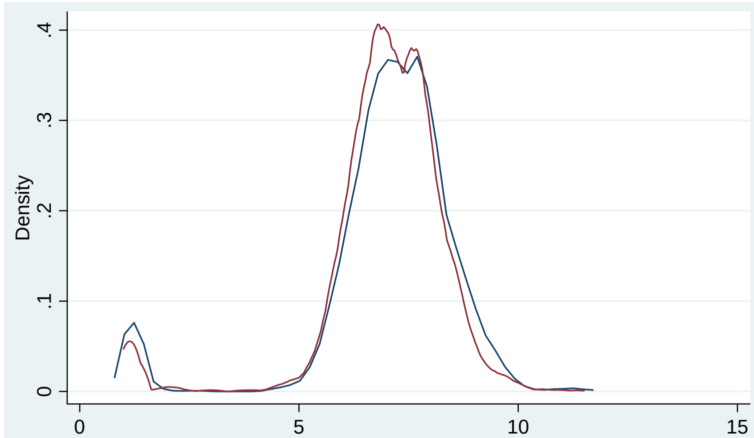

Figure 8A: Dispersion of Interest Rates (Low Documentation)

Figure 8A depicts the Epanechnikov kernel density of interest rate for two FICO groups for low documentation loans – 620^{-} (615-619) in blue and 620^{+} (621-625) in red. The bandwidth for the density estimation is selected using the plug-in formula of Sheather and Jones (1991). The figure shows that the density of interest rates on loans are similar for both the groups. A Kolmogorov-Smirnov test for equality of distribution functions cannot be rejected at the 1% level. Data for 2004 is reported here. We find similar patterns for 2001-2006. We do not report those graphs for brevity.

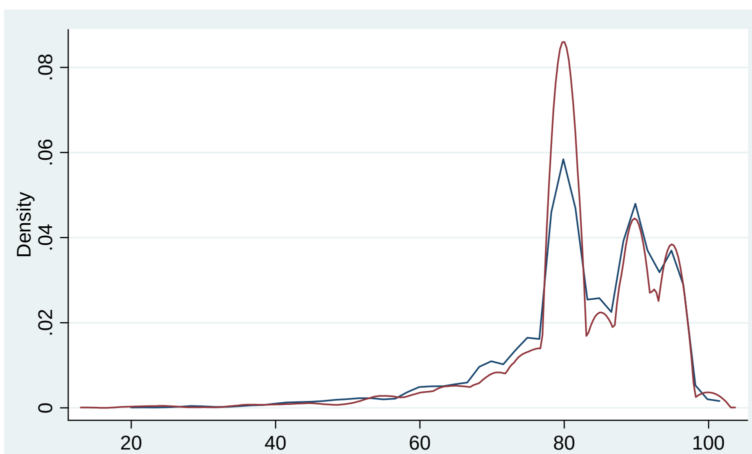

Figure 8B: Dispersion of Loan-to-Value (Low Documentation)

Figure 8B depicts the Epanechnikov kernel density of loan-to-value ratio for two FICO groups for low documentation loans – 620^{-} (615-619) in blue and 620^{+} (621-625) in red. The bandwidth for the density estimation is selected using the plug-in formula of Sheather and Jones (1991). The figure shows that the density of interest rates on loans are similar for both the groups. A Kolmogorov-Smirnov test for equality of distribution functions cannot be rejected at the 1% level. Data for 2004 is reported here. We find similar patterns for 2001-2006. We do not report those graphs for brevity.

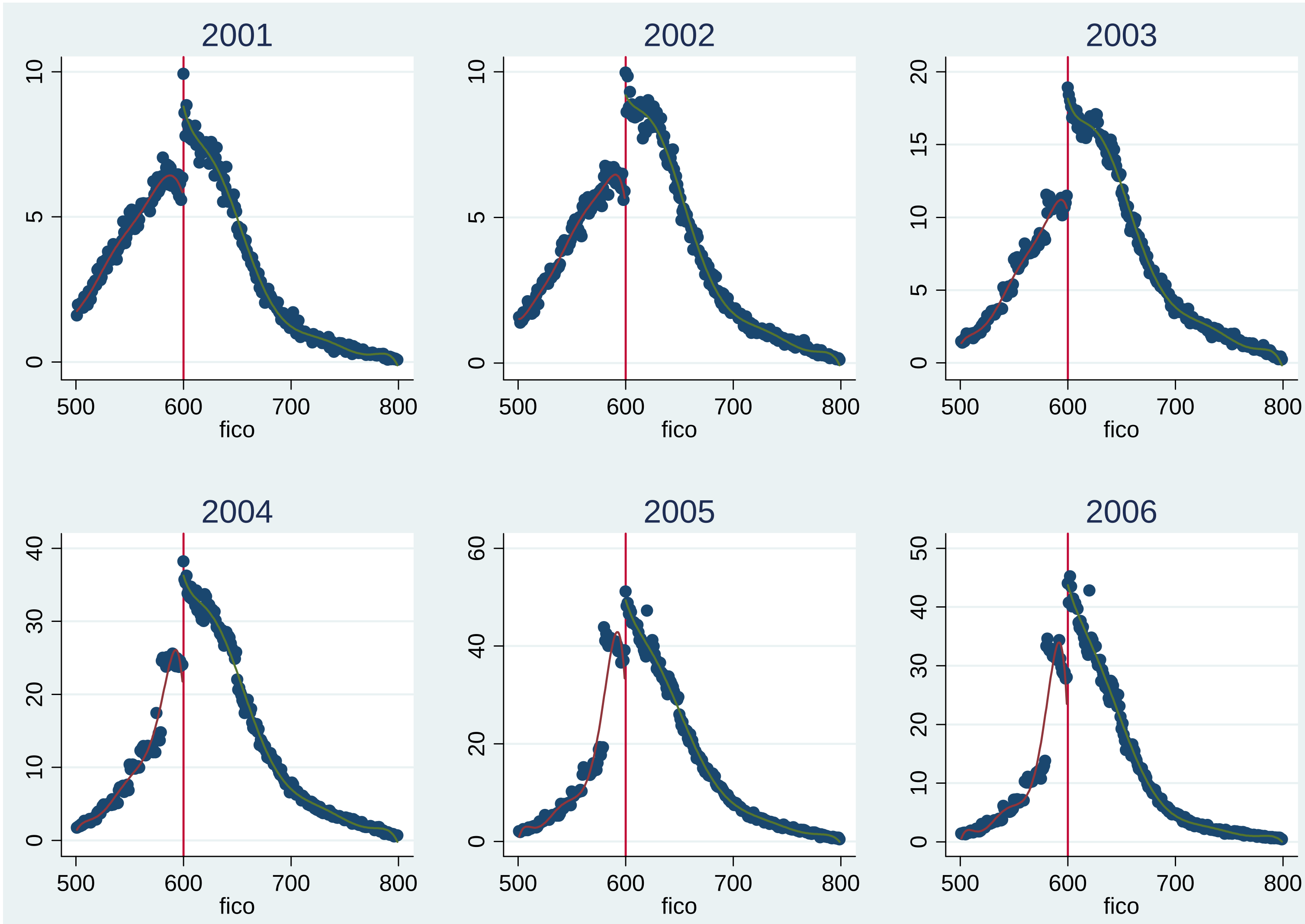

Figure 9: Number of Loans (Full Documentation)

Figure 9 presents the data for number of loans (in '00s) for full documentation loans. We plot the average number of loans at each FICO score between 500 and 800. As can be seen from the graphs, there is a large increase in number of loans around the 600 credit threshold (i.e., more loans at 600^{+} as compared to 600^{-}) from 2001 onwards. Data is for the period 2001 to 2006.

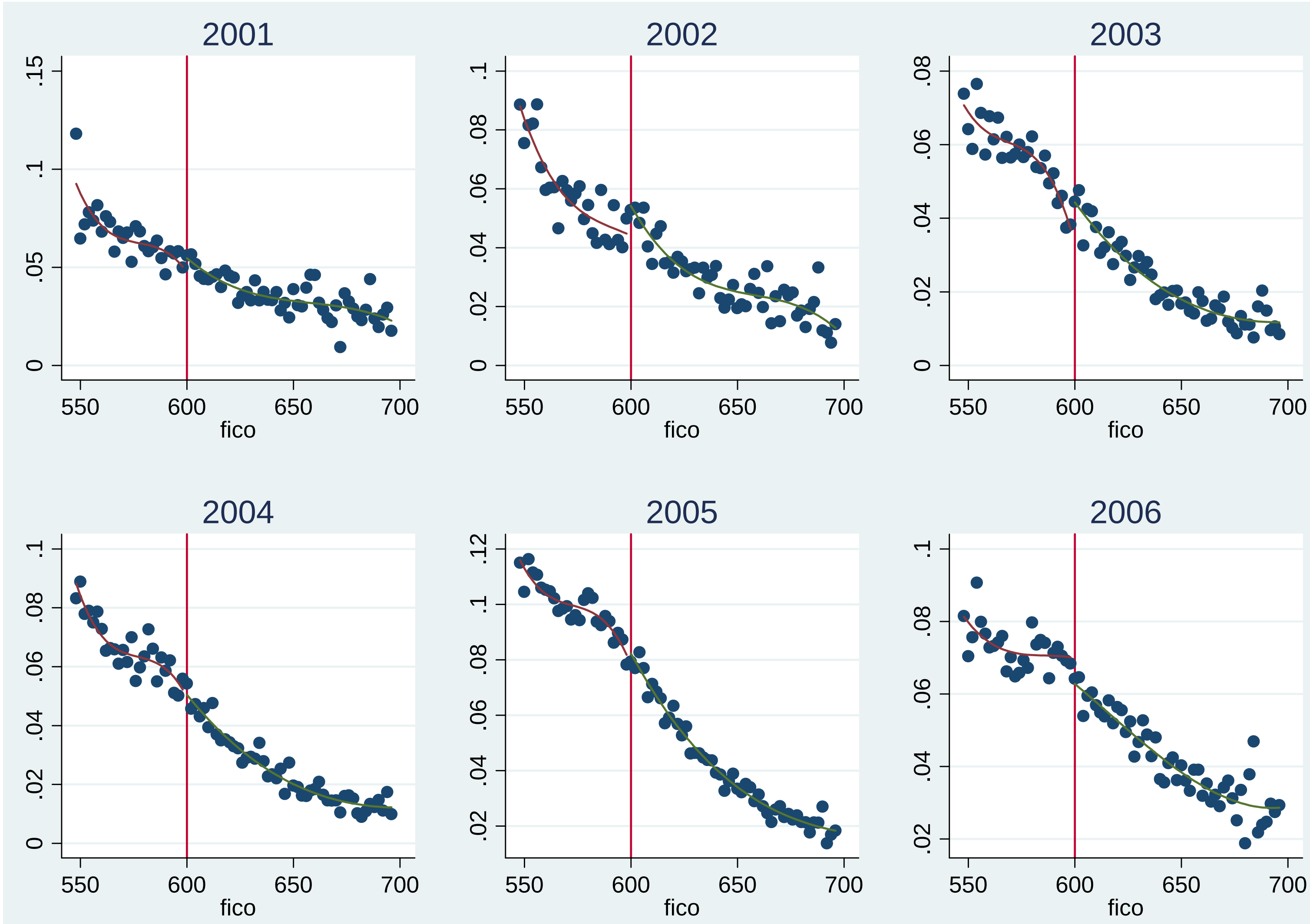

Figure 10: Annual Delinquencies for Full Documentation Loans

Figure 10 presents the data for actual percent of full documentation loans that became delinquent for 2001 to 2006. We plot the dollar weighted fraction of the pool that becomes delinquent for 2-point FICO bins between score of 550 and 700. The vertical line denotes the 600 cutoff, and a third order polynomial is fit to the data on either side of the threshold. Delinquencies are reported between 10-15 months for loans originated in all years except 2006. For loans originated in 2006, we take delinquencies between 5-10 months.

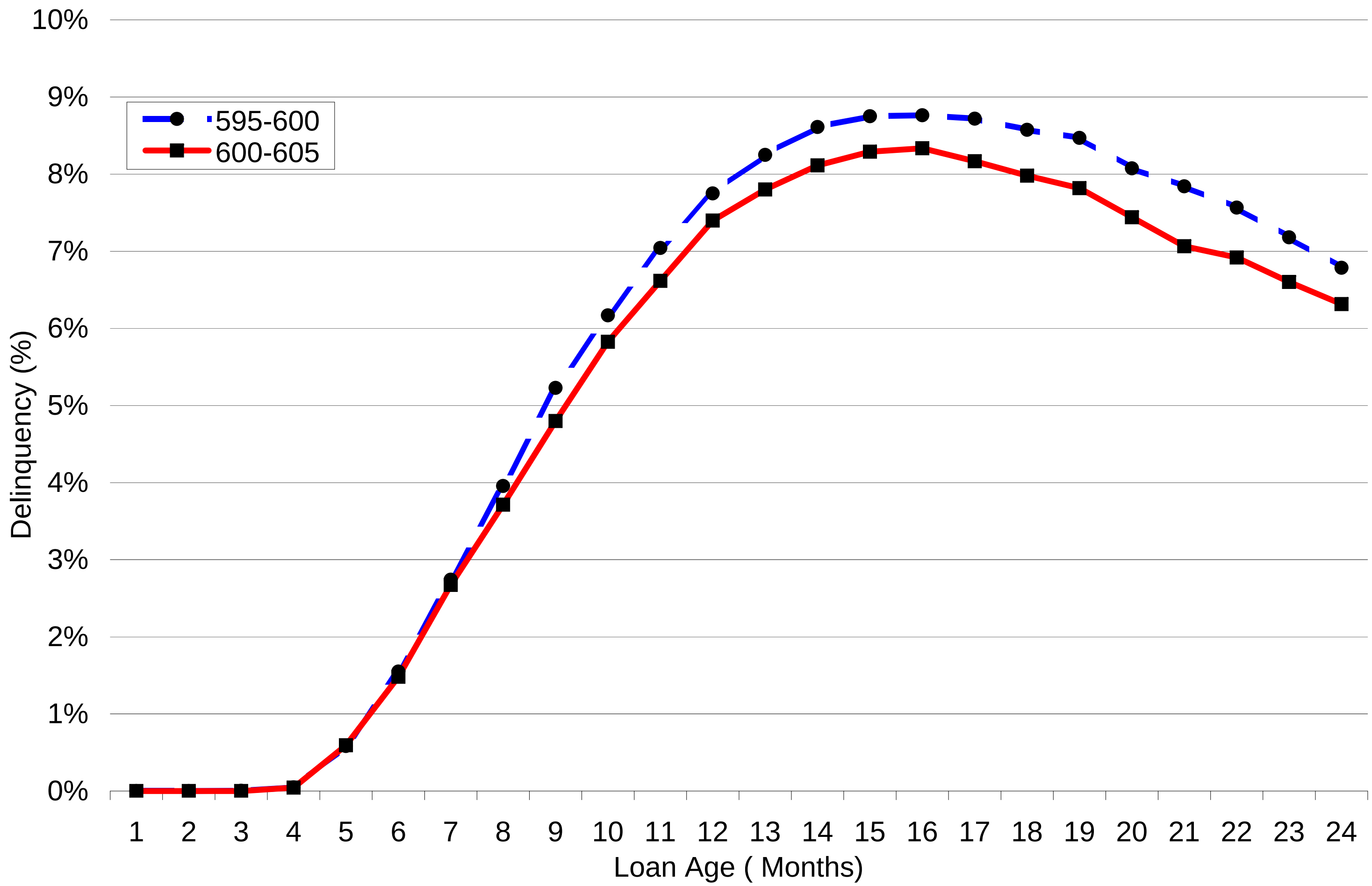

Figure 11: Delinquencies for Full Documentation Loans (2001-2006)

Figure 11 presents the data for actual percent of full documentation loans (dollar weighted) that became delinquent for 2001 to 2006. We track loans in two FICO buckets – 595-599 (600^{-}) in dotted blue and 601-605 (600^{+}) in red – from their origination date and plot the average loans that become delinquent each month after the origination date. As can be seen, the higher credit score bucket defaults *more* than the lower credit score bucket for post 2000 period. For brevity, we do not report plots separately for each year. The effects shown here in the pooled 2001-2006 plot show up for every year.

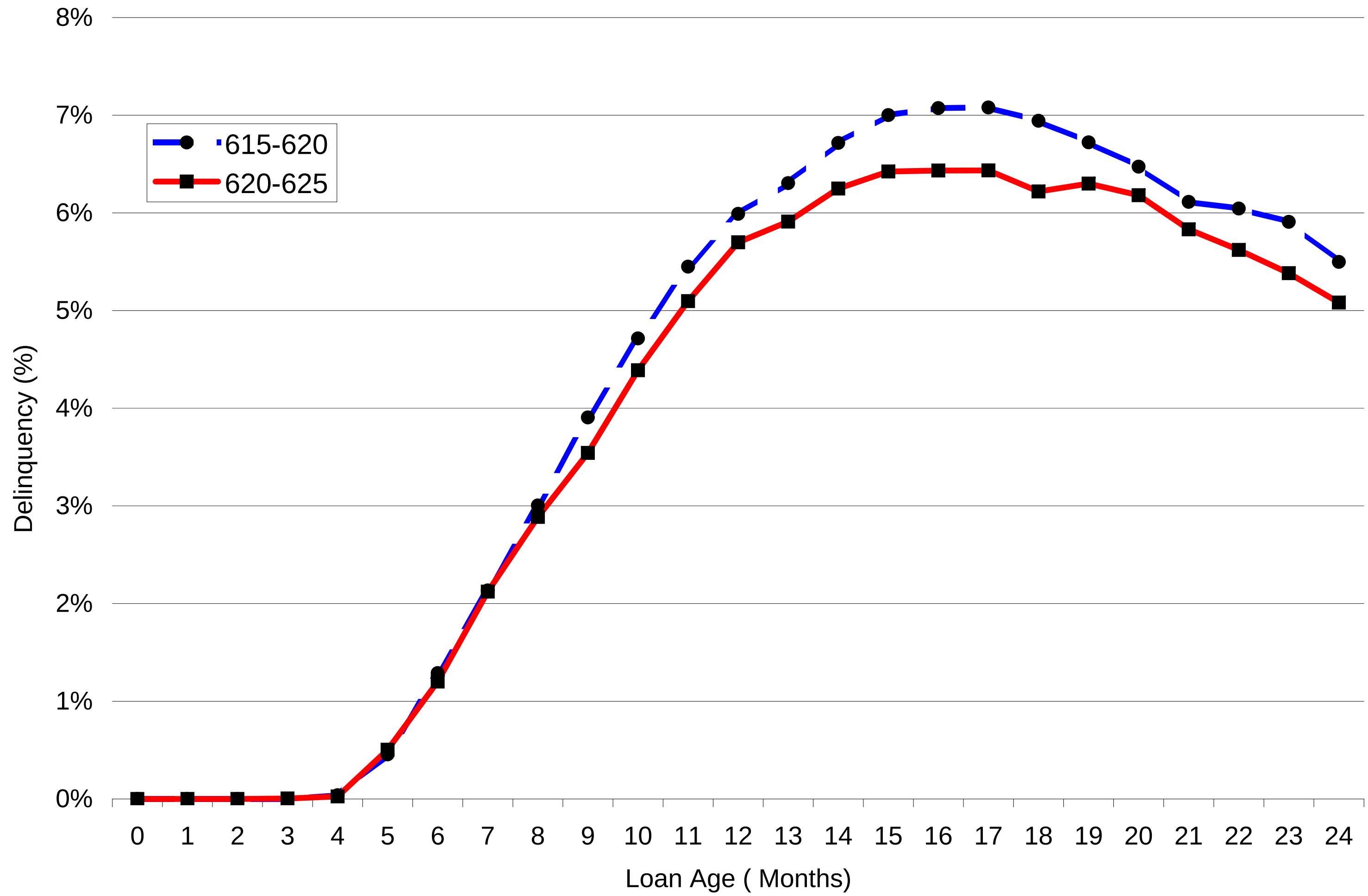

Figure 12: Falsification Test - Delinquencies for Full Documentation Loans Around FICO of 620

Figure 12 presents the falsification test by examining data for actual percent of full documentation loans (dollar weighted) that became delinquent for 2001 to 2006. We track loans in two FICO buckets – 615-619 (620^{-}) in dotted blue and 621-625 (620^{+}) in red – from their origination date and plot the average loans that become delinquent each month after the origination date. As can be seen, the higher credit score bucket defaults *less* than the lower credit score bucket for post 2000 period. For brevity, we do not report plots separately for each year. The effects shown here in the pooled 2001-2006 plot show up for every year.