Where's the Smoking Gun? A Study of Underwriting Standards for U.S. Subprime Mortgages*

Geetesh Bhardwaj[†] Rajdeep Sengupta[‡§]

September 2009

Abstract

The dominant explanation for the meltdown in the U.S. subprime mortgage market is that lending standards dramatically weakened after 2004. Using loan-level data, we examine underwriting standards on securitized subprime mortgage originations from 1998 to 2007. Contrary to popular belief, we find little evidence of a dramatic weakening of lending standards within the subprime market. We show that while underwriting may have weakened along some dimensions, it certainly strengthened along others. Our results indicate that (average) observable risk characteristics on mortgages underwritten after 2004 would have resulted in a significantly fewer ex post defaults if such mortgages had originated in 2001 or 2002. We show that while it is possible that underwriting standards in this market were poor to begin with, deterioration in underwriting after 2004 cannot be the dominant explanation for the collapse of the subprime mortgage market.

JEL Codes: G21, D82, D86.

Keywords: mortgages, subprime, underwriting, crisis.

---

*We are grateful to Franklin Allen, Eric Bond, Diana Bonfim, Gordon Dahl, John Duca, Ronel Elul, Maria Fabia-Penas, Mara Faccio, Kris Gerardi, Radha Gopalan, Gary Gorton, Jim Hamilton, Michael McCracken, Bruce Mizrach, Sangsoo Park, Geert Rouwenhorst, Rhiannon Sowerbutts, Dan Thornton, Nancy Wallace, Dave Wheelock, Paul Willen, Paul Wilson, and Thierry Tressel for helpful comments and suggestions and to Yu Man Tam for excellent research assistance. We would also like to thank the seminar participants of Rutgers University, the University of Arkansas at Fayetteville, and the conference participants at the 45th Annual Conference on Bank Structure and Competition at the Federal Reserve Bank of Chicago, the Banking and Financial Intermediation Conference at the European Banking Center in Tilburg University, the Financial Crisis Conference at the Yale International Center, the Summer Research Conference in Finance at the Indian School of Business, and the 9th Annual Bank Research Conference at the FDIC.

†Senior Economist, The Vanguard Group. The views expressed herein are those of the individual author and and do not necessarily reflect the official positions Vanguard Group Inc.

‡Economist, Federal Reserve Bank of St. Louis. The views expressed are those of the individual author and do not necessarily reflect official positions of the Federal Reserve Bank of St. Louis, the Federal Reserve System, or the Board of Governors.

§Correspondence: Research Division, Federal Reserve Bank of St. Louis, P.O. Box 442, St. Louis, MO 63166-0442. Phone: (314) 444-8819, Fax: (314) 444-8731, Email: rajdeep.sengupta@stls.frb.org.

1 Introduction

Conventional wisdom has argued that deterioration in underwriting standards was central to the collapse of the market for subprime mortgages. The hypothesis that "most bad loans are made in good times" has been viewed, by policymakers and academics alike, as one of the principal features of credit crises.[1] The current mortgage crisis in the United States is no exception. Indeed, the recent boom in the U.S. housing market witnessed a surge in nonprime mortgage originations from 2000 to 2006. Given the lower underwriting requirements for nonprime mortgages, this explosive growth naturally caused a decline in lending standards for the overall mortgage market.

Of greater interest, however, is the question of whether there was a decline in underwriting standards *within* the market for subprime mortgages. There is a remarkable increase in early default rates for post-2004 originations, especially during 2006 and 2007 (Figure 1).[2] Such high and early defaults on subprime mortgages led both policymakers and academics to believe that there was a significant deterioration in underwriting standards *within* the subprime mortgage market, particularly for these later vintages.[3] For example, the President's Working Group on Financial Markets (March, 2008) concluded that

> The turmoil in financial markets was triggered *by a dramatic weakening of underwriting standards for U.S. subprime mortgages, beginning in late 2004*, and extending into early 2007.[4]

Much of the same sentiment is echoed in the popular press.[5] Despite the analysis of these

[1] There is a significant volume of theoretical and empirical studies supporting this hypothesis (see, for example, Gorton and He, 2008, and references therein).

[2] Figure 1 illustrates the default probabilities by loan age for each year of origination (vintage) of subprime mortgages between 2000 and 2007. The default probabilities are calculated using the Kaplan-Meier product limit estimator (see appendix B and Section 3.2 for more details).

[3] Notably, high default rates for 2005-2007 vintages occur well before the loan age of 24 (and 36) months, typically the reset date on hybrid adjustable rate mortgage (ARM) products. Clearly, a jump in the payment obligations on hybrid-ARM resets would not explain these high early default rates. In appendix A, we argue that early defaults for post-2004 originations might be better explained if one studies early prepayment patterns on pre-2004 originations.

[4] Policy Statement on Financial Market Developments, March 2008 (emphasis in the original).

[5] Such examples are ubiquitous in newspaper reports. To cite a few examples: "Strange was becoming increasingly common: loans that required no documentation of a borrower's income. No proof of employment.

events in business and academic journals, there has been little economic analysis of the proposition of examining underwriting standards within the subprime mortgage market. This paper presents summary evidence on subprime mortgage underwriting standards. At the cost of parsing the policy statement above too literally, we examine two related questions. First, was there a dramatic weakening of underwriting standards *within* the subprime mortgage market? Second, did *this* "weakening" begin around late 2004 so as to trigger extensive defaults in subprime mortgages? To examine these questions, we study loan-level data on more than 9 million subprime mortgages from the LoanPerformance (LP) database over the period 1998-2007. This is the largest available repository on securitized subprime mortgages (see Section 2 for details). Our aim is to study the underlying distribution and evolution of borrower and mortgage (loan) characteristics in the subprime market with a view to identifying the deterioration in underwriting standards.

We argue that any study of underwriting standards in this environment needs to account for two important features of credit risk that have largely been ignored up to this point. The first takes into account the multidimensional nature of credit risk: It is often possible to compensate for the increase in the ex ante risk of one borrower attribute by raising the requirement standards along another dimension. The second involves the idea that while both borrower attributes and mortgage characteristics determine credit risk, the terms and conditions of the latter are largely determined by the former. We address the endogeneity problem that confronts the use of mortgage terms such as loan-to-value (LTV) ratio and mortgage interest rate as explanatory variables in determining loan performance. To this end, we first develop a test for endogeneity bias by adopting techniques in Chiappori and Salanie (2000). Following this, we study the determinants of mortgage characteristics (such as LTV ratio and interest rate) and mortgage delinquencies in the subprime market by accounting for both features mentioned above. Finally, we devise a counterfactual technique to determine whether there was a decline in underwriting standards within the subprime market after 2004.

No money down.I was truly amazed that we were able to place these loans" (The Bubble: How homeowners, speculators and Wall Street dealmakers rode a wave of easy money with crippling consequences. The Washington Post, June 15, 2008). "House prices levitated as mortgage underwriting standards collapsed. The credit markets went into speculative orbit, and an idea took hold. Risk, the bankers and brokers and professional investors decided, was yesteryear's problem." (Why no Outrage? Wall Street Journal, July 19, 2008).

Our results show that the hard information available on mortgage originations does not reveal deterioration in underwriting standards for securitized subprime originations, particularly after 2004.[6] Given the multidimensional nature of ex ante credit risk, it is difficult to emphasize weakening in terms of some attributes as a decline in overall underwriting standards. While underwriting may have weakened along some dimensions (e.g., lower documentation), it also strengthened in others (e.g., higher FICO[7] scores). Hard data provide evidence of credible underwriting over this period that attempted to adjust riskier borrower attributes with lower LTV ratios and higher FICO scores. Moreover, there is compelling evidence to suggest that lenders emphasized FICO scores not only as an adequate indicator of credit risk, but also as a means to adjust for other riskier attributes on the origination.

In addition, we present evidence showing that the effectiveness of FICO scores at origination in gauging default risk did not deteriorate over the years. To test the effectiveness of origination FICO scores, we examine the performance of a given change in origination FICO scores in terms of ex post default.[8] We find evidence that improvement in FICO score increases the ex post survival probability. Further, if one controls for other attributes on the loan origination, the improvement in FICO score significantly increases the ex-post survival probability, especially for later vintages. Moreover, this result is robust to the inclusion (or exclusion) of contract terms such as cumulative LTV (CLTV) ratio and mortgage rates as controls in this estimation procedure. The result is also robust to different specifications of FICO score groups and to variations in terms of transitions across these groups. In summary, these results seem to suggest that the lender emphasis on FICO scores at the time of origination was not misplaced.

Critical to this result is the evidence of endogeneity bias. Our test of endogeneity bias

[6] The results presented here are based on the hard information available on securitized subprime originations. The distinction between hard and soft information follows Stein (2002). Stein argues that the decision of whether to extend credit on a home mortgage loan application is typically based on hard information because it is readily verifiable and can be credibly transmitted. On the other hand, an unsecured "character loan" is based on soft information that cannot be verified by anyone other than the agent who produces it. This is not, however, to diminish the role of soft information for the subprime market.

[7] Borrower credit score at the time of loan origination is denoted by FICO™ (an industry standard developed by the Fair Isaac Corporation) with a number in the range 300 to 850.

[8] Using absolute measures to compare the performance of FICO scores across different default regimes can be misleading. For example, if defaults were to rise because of some exogenous shock, an increase in the default rates would occur across all FICO scores. A more relevant measure in this context would be one that compares loan performance for a given change in the FICO score across the different default regimes. Accordingly, we compare the difference (increase) in survival probabilities of an origination with a higher FICO score (or belonging to a higher FICO score group) relative to one with a lower FICO score (or belonging to a lower FICO score group).

presents evidence of a positive correlation, conditional on observable characteristics, between the individual's choices of LTV ratio (coverage[9]) and the ex post occurrence of default (risk). If we do not account for this endogeneity problem and include mortgage terms as explanatory variables in our default estimation, we introduce a positive bias on the explanatory variables such as FICO scores at origination. As a result, the positive bias reduces the magnitude of the negative relationship between FICO scores and ex post default.[10]

To address the multidimensional nature of underwriting, we use counterfactual analysis to help answer the following question: How would ex post default rates change if a "representative borrower" given a loan in 2005, for example, were to have been given a loan in, say, 2001? Our results indicate that (average) observable risk characteristics on loans underwritten after 2004 would have resulted in a significantly lower ex post default if they had been underwritten in 2001 and 2002 than (average) observable risk characteristics on loans underwritten in 2001 or 2002. Stated differently, if loans underwritten in 2005 (or 2006 or 2007) originated in 2001 or 2002, then they would have performed significantly better on average than loans underwritten in 2001 or 2002. Despite the endogeneity problems of including mortgage terms, we show that the counterfactual results are robust to the inclusion of mortgage terms (such as LTV ratio and interest rate) as explanatory variables of mortgage default. In light of this evidence, it is unclear how deterioration in underwriting since 2004 can be the dominant explanation of delinquencies in the subprime market. Of course, our analysis does not rule out the hypothesis that underwriting standards in the subprime market were probably poor to begin with. At the very least, unobservable risk characteristics and market conditions (such as house price appreciation) had a greater role than was earlier believed.

There is a large segment of literature analyzing different features of the subprime mortgage market. Earlier contributions include Cutts and van Order (2005) and Pennington Cross and Chomsisengphet (2007). More recent papers studying the subprime crisis include Dell'Arricia et al. (2008), Demyanyk and Van Hemert (2009), Elul (2009), Gerardi et al. (2009), Keys

[9] In the interest of familiarity, we retain the use of the term *coverage*, although semantically, it might be more appropriate to its original usage in the insurance market setting (Chiappori and Salanie, 2000). Throughout the paper, "higher coverage" implies a higher CLTV ratio on the mortgage.

[10] While the inclusion of mortgage terms certainly dampens the effect of FICO scores on default, it does not reverse the earlier result of improvement in the effectiveness of FICO scores.

et al. (2009), Mayer et al. (2008), and Mian and Sufi (2008). Our paper makes several new contributions to this literature. First, we show the change in mortgage underwriting does not explain the widespread defaults and subsequent collapse in the subprime mortgage market. Our results show that while underwriting declined on certain dimensions, a multidimensional study of underwriting fails to provide evidence of a secular decline in underwriting standards. In particular, we present evidence showing that lenders seem to have attempted to offset riskier attributes on originations by increasing the average quality of borrowers (as measured by their credit scores) to whom such loans were made. More important, this adjustment appears to have strengthened over the years in our sample period. Second, we provide evidence indicating that, especially on the basis of ex post loan performance for later vintages, lender emphasis on credit scores was not misplaced. Contrary to conventional wisdom, the effectiveness of an increase in credit scores on loan performance actually shows improvement over the years in our sample period. Third, we establish and account for the endogeneity problem that confronts the use of mortgage terms as explanatory variables in determining mortgage default. This endogeneity problem is important, because it introduces a positive bias that dampens the negative effect of (higher) credit scores on mortgage delinquencies. Fourth, we use counterfactual analysis to demonstrate that, at least in terms of underwriting standards, loans underwritten for later vintages would have performed no worse than mortgages of earlier vintages. This raises serious doubts on the conventional wisdom that a decline in underwriting for subprime mortgages is central to the collapse of this market. In contrast, it does not rule out the possibility that the design on subprime mortgage contracts has been flawed since its inception.

The rest of the paper is organized as follows. Section 2 presents the summary data on borrower and mortgage characteristics, while Section 3 provides summary evidence on mortgage underwriting. In Section 4, we present a brief discussion of limitations of characterizing underwriting standards and provide a theoretical framework for our analysis. Section 5 provides the evidence on endogeneity bias and estimation results on underwriting and loan performance in the subprime market. The counterfactual analysis is described in Section 6. Section 7 provides a discussion of the results.

2 Data and Summary Statistics

For the purposes of this study, we analyze loan-level data from the Asset-Backed Securities (ABS) database of the LP data repository.[11] Although this database contains both subprime and Alt-A pools, we restrict our analysis to subprime loans for the purposes of this study.[12] Following industry convention and standard practice in this field, we classify a loan as *subprime* if it belongs to a subprime pool in the ABS database.[13] LP data include only those loans that were securitized in the ABS market, as opposed to loans that were retained by originators in their portfolios. In addition to various borrower and mortgage characteristics, LP records repayment behavior on the loan. Mayer and Pence (2008) observe that LP captures around 90 percent of the subprime securitized market from 1999 to 2002 and nearly all of the market from 2003. In what follows, our analysis focuses on over 9 million first-lien subprime loans in the ABS database that originated between 1998 and 2007 and follows their repayment behavior up to December 2008.

We begin our discussion on summary statistics with a brief description of the trends in univariate data. Because the subprime market evolved fairly rapidly over the years in our sample period, we record changes in underwriting standards by vintage (year of mortgage origination) throughout this paper. Around 1998, more than half of subprime originations were fixed rate mortgages (FRMs). Over the years, there has been a clear shift toward the origination of more adjustable rate products (ARMs). Of these, the majority of originations were hybrid-ARM products (e.g., 2/28 and 3/27 mortgage products).[14] At their peak around 2005, hybrid-ARMs

[11] This is the largest database on nonprime loans with loan-level data on over 17 million nonprime (both subprime and Alt-A) mortgages originated in the United States. However, the dataset is not without its limitations: First, there is little information on the households that held these mortgages. For example, there are no data on household debt, income, employment, or demographics. Second, unlike other studies using mortgage data, the lack of identifiers in this database makes it difficult to match and combine these data with other databases to broaden the scope of analysis. Third, we do not have data on mortgage applications and are therefore unable to compare approvals to loan applications that were denied. Finally, even for loans in the database, we are unable to track multiple liens or mortgages on the same property.

[12] Loosely speaking, subprime pools include loans to borrowers with incomplete or impaired credit histories, whereas Alt-A pools include loans to borrowers of higher credit quality but who are unable or unwilling to provide documentation on the loan (Fabozzi, 2006).

[13] Other definitions involve identifying originations of lenders specializing in subprime originations, or using specific criteria such as lower credit scores and so on, to define subprime loans. It is important to understand that the guidelines for selection into subprime or Alt-A pools vary across originators or arranger of the securities. By our definition, we classify a mortgage as subprime if market participants labeled this mortgage as subprime at securitization.

[14] Hybrid-ARMs are specialized products that include an initial period over which the repayment schedule on

accounted for almost 80 percent of products in the subprime mortgage market. Contrary to conventional wisdom, teaser rates on hybrid-ARM products were not significantly lower compared to closing rates on other mortgage products in the subprime market. In fact, we do not find any significant difference between the unconditional means of closing rates on FRMs and hybrid-ARMs mortgage products. This is true for originations of all vintages in our sample period.

Perhaps a lesser-known fact about subprime mortgages is that a majority (around 60-70 percent) of subprime originations between 1998 and 2007 were refinances. More than half of the originations for every single year in this period were cash-out refinances. No cash-out refinances account for about 11-16 percent of originations between 1998 and 2003, but their proportion drops to 6-7 percent of total originations, once the Federal Reserve started raising interest rates in 2004. On an annual basis, roughly 90 percent of originations are on owner-occupied properties. Second homes account for a small proportion, about 1-2 percent of originations, while non-owner (i.e., investor) occupied properties account for 7-9 percent. Our data show little change in underwriting in terms of occupancy and purpose of the loan; proportions of the sample under different categories for either characteristic were fairly stable over the sample period.

We observe a trend toward riskier loans in terms of lower documentation and higher CLTV ratios.[15] From roughly 18-19 percent of originations in 1999-2000, the proportion of low-doc loans increased to 35-36 percent of originations for 2005-2006. However, no-doc loans remain an insignificant 0.4-0.7 percent of the total originations for all vintages. In addition, subprime lenders increasingly began to originate mortgages with high CLTV ratios. For example, originations with CLTV ratios in the (90, 100] range increased from 3-4 percent in 1998-1999 to 35-40 percent of total in 2005-2006. In contrast, average borrower FICO scores on originations increased over this period. For example, only 30 percent of the originations in 2000 had credit scores above 620, whereas the number was more than 50 percent in 2005. These trends in

the mortgage resembles that of a FRM and a subsequent period over which the mortgage product acts like an ARM. During the fixed-leg of the hybrid-ARM, the mortgagee pays a lower introductory closing rate called the *teaser rate*. The teaser rate remains in effect until the *reset date*, after which the repayment schedule on the hybrid-ARM resembles an ARM.

[15] We have used the CLTV ratios as they provide a better measure of home equity for the borrower.

univariate data do not reveal a secular decline in lending standards. While the trend shows increased risk-taking on the part of lenders in terms of documentation requirements and high CLTV loans, there is also a trend toward higher borrower quality, as summarized by average FICO scores. More important, these trends are discernible over the entire sample period and do not suggest anything particularly special about originations after 2004.

Turning our attention to multivariate analyses of underlying risk characteristics, we find that borrowers with lower documentation have, on average, higher FICO scores. Table 1 shows the distribution of FICO scores conditional on documentation level on originations of various vintages. The proportion of borrowers in the lowest FICO score group (less than 620) has declined over the years. At the same time, there has been an increase in the proportion of borrowers in the 620-659 score group and the 660-719 score group, especially for originations without full documentation. The distribution of FICO conditional on CLTV shows a similar pattern (Table 2). For all years, originations with higher CLTVs typically have higher FICO scores. As in the case of loan documentation, there has been a shift in population from the lowest FICO score group (less than 620) to the two intermediate FICO score groups (620-659 and 660-719), especially for originations with higher CLTVs.

3 Summary Evidence on Underwriting for Subprime Mortgages

3.1 FICO Score and Risk Characteristics

Based on the evidence presented in Tables 1 and 2, it is difficult to argue, as some have claimed, that there was a secular decline in lending standards in terms of a borrower's observable risk characteristics. Despite exposing themselves to more credit risk on some borrower attributes (for example, by lowering documentation requirements), lenders seem to have attempted to offset this by increasing the average quality of borrowers (as measured by their credit scores) to whom such loans were made.

For a more rigorous test of this hypothesis, we use regression techniques to determine equilibrium underwriting behavior. Borrower FICO scores are regressed on other borrower attributes and loan characteristics. The regression estimates in Table 3 summarize equilibrium underwriting for subprime mortgages originated between 2000 and 2007.[16] In addition to borrower characteristics used as regressors in Panel A, Panel B of Table 3 includes terms on the mortgage contract, such as the CLTV ratio and the *closing rate spread*. The *closing rate spread* is defined here as the difference between the closing rate on the origination (the teaser rate for hybrid-ARMs) and the 30-year conventional mortgage rate.[17] Regression coefficients indicate the presence of underwriting efforts to control for overall credit risk by varying credit score requirements on loan approvals. For example, a large negative and significant coefficient on the full-documentation dummy (both panels) indicates that, after controlling for other borrower attributes, a borrower with low or no documentation has a significantly higher FICO requirement than a similar borrower providing full documentation on the loan. As one would expect, the FICO requirement for loan approval on non-owner (investor) occupied homes is the highest, followed by that on second homes, whereas approvals for owner-occupied originations have the lowest required FICO scores. Not surprisingly, mortgages on properties with greater value have progressively higher required FICO scores. For loans of all vintages, property values in a lower quartile have, on average, a lower FICO score than those property values in the immediately higher quartile. Evidently, refinances have a lower FICO score, on average, than direct home purchases. The large negative coefficient on the *closing rate spread* variable in Panel B indicates that originations on low FICO scores in equilibrium have a higher mortgage rate. In addition, equilibrium FICO scores are higher on originations with higher CLTV ratios.

The regression coefficients indicate that underwriters attempted to adjust for borrowers' riskier attributes by requiring higher average FICO scores. Moreover, changes in the size of the coefficients over the years seem to suggest that the size of this adjustment appears to have increased over the years in our sample period. To test this hypothesis more formally, we use a fully interacted dummy variable model of the regression in Panel A of Table 3. The dummy

[16] In what follows, we report the regression estimates for all subprime mortgages that originated between 2000 and 2007. The results for the years of origination 1998 and 1999 are not given here but are available on request. Unless mentioned otherwise, regression estimates in the paper control for property type (dummies for single-family residence, condo, townhouse, cooperative, etc), property location (dummies for the state in which the property is located), and loan source (dummies for broker, realtor, wholesale, retail, etc).

[17] The 30-year *conventional mortgage rate* is the monthly average contract rate on commitments for prime FRMs, released by the Federal Home Loan Mortgage Corporation.

variable takes the value 1 for all originations after a given calendar year and 0 otherwise. We present the estimates on four specifications in Table 4, starting with an interacted dummy for post-2002 originations and ending with one for post-2005 originations. The estimated coefficient of 21.77 on the dummy variable for the post-2004 vintage shows that the improvement in FICO scores for originations between 2005 and 2007 was statistically as well as economically significant.

The preceding analysis indicates the presence of credible underwriting (i.e., the appropriate sign on the coefficient). However, we cannot comment on whether such underwriting was adequate in terms of the marginal rates of adjustment across different borrower attributes (i.e., the magnitude of the coefficient). Stated differently, we observe that the FICO scores on low documentation loans for all the vintages were, on average, higher than those on full-documentation loans. However, we do not know if the difference in FICO score of 19.26 points (as recorded on loans of 2006 vintage in Panel A, Table 3) as opposed to that of 15.14 points (as recorded on loans of 2000 vintage) is sufficient to offset the increase in the borrower risk profile (i.e., the low documentation on loans). Still, the evidence presented above indicates that lenders increasingly relied on FICO scores to offset other riskier attributes of borrowers.

3.2 FICO Scores and Default Risk

We conclude this section with some evidence on FICO scores and default behavior on subprime mortgages. In doing so, we provide some preliminary evidence that might help explain the increasing reliance on FICO scores. Our data allow for tracking mortgage repayment behavior on a monthly basis, thereby allowing us to determine the current status on the loan in terms of prepayments, delinquencies, and foreclosures. We can also distinguish among a 30-day, a 60-day, or a 90-day delinquency status on the loan. Following industry conventions, we define a mortgage to be in *default* (or in serious delinquency) if it records a 90-day delinquency event at any point in its repayment history.[18] Default rates and the probability of surviving a delinquency are calculated by using the Kaplan and Meier (1958) product limit estimator.

[18] Although we define default as a 90-day delinquency throughout the paper, the results are qualitatively similar for alternative definitions using a 60-day delinquency or a foreclosure as default.

Appendix B provides a formal treatment of this non-parametric approach in the context of mortgage repayment behavior.

As seen in Figure 1, it is clear that defaults started to rise sharply in 2006 and 2007, primarily for mortgages that originated between 2004 and 2007. As an example, about 28 percent of mortgages that originated in 2001 were in serious delinquency by the fourth calendar year (end of 2004), whereas the same proportion of defaults for 2005 originations occurred in just over two calendar years (by the first quarter of 2006). The numbers are even more striking when one considers that around 35 percent of mortgages that originated in 2006 were in default by the end of 2007. For post-2004 originations, most serious (90-day) delinquencies occur well before the reset dates on hybrid-ARM products. For example, 26 percent of originations of 2006 vintage and 32 percent of originations of 2007 vintage were seriously delinquent within the first 18 months. The corresponding numbers on originations of 2001 and 2002 vintage were 7.9 and 7.6 percent, respectively.[19] Most of the commentary on subprime mortgages has sought to explain this significant increase in default probabilities by a weakening in lending standards for originations after 2004.

At this point, it is important to recall several results from our analysis above. First, our analysis of summary data seems to indicate a trend toward higher FICO scores alongside lower documentation and higher CLTV ratios. Second, we observed that the average FICO score is significantly higher for originations whose other attributes (such as lower documentation or higher LTV ratios) are arguably riskier. Third, we present evidence to suggest that this adjustment strengthened over the years in our sample period. These underwriting patterns suggest that lenders placed emphasis on the FICO score not only as an adequate indicator of credit risk, but also as a means to adjust for other riskier attributes on the origination. With the benefit of hindsight, some industry experts have faulted originators on this account:

> ... [T]he crucial mistake many lenders made was relying on FICO credit scores to gauge default risk, regardless of the size of the downpayment or the type of loan.[20]

[19] These results suggest that loan performance on subprime mortgages can hardly be explained by variations in the distribution of product types (Mayer et al. 2008). For that reason, the results presented here are for data pooled over all mortgage products. Results on individual product types (ARMs and FRMs) are qualitatively similar and are available on request.

[20] "The woman who called Wall Street's meltdown." (*Fortune Magazine*, August 4, 2008). However, this is not

Anecdotal evidence has also been provided in support of the hypothesis that FICO scores failed as predictors of default.

However, one needs to approach this argument with caution. For instance, if some exogenous factor increases the default rate on mortgages for later vintages (post-2004 originations), it is likely to show worsening performance across all FICO scores for the later vintages. This is precisely what we observe in the data; there has been a significant increase in defaults for post-2004 originations for all FICO scores. The increase in defaults rates for a given FICO score is reflective of the increase in overall default rates for the later vintages. This can hardly be viewed as evidence of a decline in the effectiveness of FICO scores. Therefore, to test for the effectiveness of FICO, we demonstrate that the increase in the probability of survival for a given improvement in FICO scores does not deteriorate across the vintages.

To address this issue, we develop a metric of performance for origination FICO of a given vintage in terms of ex post loan performance of that vintage. Our metric is the difference in survival probabilities for an origination with a higher FICO score relative to one with a lower FICO score of the same vintage. This measure of performance relative to other originations in the same vintage is uncorrelated with exogenous factors determining default. For the ease of exposition, we split our sample into originations within different FICO score groups. Next, we calculate as a first pass the non-parametric estimates of the (unconditional) survival probabilities for originations within each FICO score group. In Section 5, we provide the parametric estimates of the survival probabilities for the group after controlling for other attributes on the origination.

Table 5 reports the difference (increase) in the probability of a loan surviving a 90-day delinquency event after two calendar years for originations. For the purposes of this analysis, we split the sample into various FICO score groups at intervals of 40 points, starting at a FICO score of 540. The rows in Table 5 show the percentage point increases in survival probabilities for originations in a higher FICO score group relative to those in its immediately lower FICO score group. Rows 1 and 2 in Table 5 find that such increases in survival probabilities among the lowest FICO score groups show deterioration in performance of origination FICO across

borne out in terms of the evidence in our data (see Table 8).

the vintages. In contrast, rows 4 through 6 show that the highest FICO score groups show improvement in origination FICO performance across the vintages. This contrasting pattern could have motivated the underwriting to seek higher FICO scores on riskier originations. In Section 5, we perform a rigorous test of this hypothesis by controlling for other attributes on the origination. However, it is important to point out that the overall trend is not driven solely by the highest FICO originations. Even without the two highest group "transitions" that show maximum improvements, we do not observe deterioration in FICO performance. This is shown by the average computed in the last row of Table 5. In appendix C, we confirm the robustness of this result for other specifications of the groups using different interval widths and starting FICO scores to demarcate these groups.[21]

These findings are important in our context for two reasons. First, as already discussed, more recent originations with higher FICO scores tend to be riskier in terms of other attributes (i.e., entail a greater likelihood that the origination has a lower documentation or a higher CLTV ratio). Second, there is anecdotal evidence suggesting that FICO scores at later vintages did not necessarily reflect the "true" creditworthiness of the borrower.[22] Naturally, one would expect the relative performance of higher FICO scores to be significantly worse than those of earlier vintages. However, we do not find evidence to support these hypotheses. In summary, the evidence from our non-parametric tests suggests that lender emphasis on FICO scores was not misplaced.

4 Mortgage Underwriting, Asymmetric Information, and Endogeneity Bias

The importance of information problems in any borrower-lender scenario cannot be overemphasized, especially when it pertains to a market for borrowers who would not otherwise qualify

[21] Other specifications include FICO score groups chosen at different intervals (such as 20 and 40) and at different starting FICO scores (such as 520, 521, 540, and 541). See appendix C for details.

[22] Some observers claim that a low-interest rate environment, as prevailed over the early part of this decade, enabled borrowers to improve creditworthiness and inflate their credit scores. Others include the possibility of "doctoring" a person's credit score to increase it (for anecdotal evidence, see Foust and Pressman, 2008). In either case, this would imply that the effectiveness of higher FICO scores should decline, especially on later originations.

for more conventional sources of financing. In this section, we emphasize the role of information asymmetries in the loan underwriting process. However, it is important to first list the limitations of our study in examining underwriting standards for subprime mortgages.

4.1 Limitations of Characterizing Underwriting Standards

First, approving loan applications of borrowers who would previously be considered uncreditworthy can be viewed as a weakening of underwriting standards. The subprime market extends credit to borrowers who would otherwise be denied loans in the prime market. Taken to its logical conclusion, one could view the emergence of subprime lending as a weakening of underwriting standards for the U.S. housing market as a whole. Significantly, for loans older than 60 months in our sample, default probabilities on subprime mortgages have never been lower than 0.28. These facts raise important questions about the viability of the subprime market as a whole. However, such questions are beyond the scope of this paper. For our purposes, it is important to keep in mind that our examination of a weakening in underwriting standards is relative to subprime mortgages of earlier vintages and not vis-à-vis mortgages in other segments of the market (prime, jumbo, and Alt-A).

Second, several characteristics of the borrower are summarized to determine overall credit risk. Lenders are known to compensate for the increase in the ex ante risk of one borrower attribute by raising the requirement standards along another dimension. Stated differently, borrower credit risk is multidimensional. This study takes into account the multidimensional nature of credit risk, arguing that any focus on a single borrower or mortgage characteristic is misleading. Accordingly, defining a decline in underwriting standards requires aggregating each borrower characteristic to build a summary measure that fulfills a variety of desirable conditions. Needless to say, the solution to this aggregation problem has proved elusive. To the best of our knowledge, we are not aware of a single metric that adequately summarizes a variety of borrower characteristics. Therefore, in Section 6, we adopt a counterfactual technique to cope with this problem.

Third, mortgage underwriting refers to the process used by a mortgagee (lender) to assess the credit risk of the mortgagor (borrower). The process involves summarizing the ex ante risk of default from a profile of borrower attributes with the purpose of approving or denying the borrower's loan application. Therefore, underwriting is based on the borrower's *observable* characteristics *at the time of origination.*

A final caveat relates to the determinants of ex post default on subprime mortgages as a testament to declining underwriting standards. Mortgage characteristics are themselves outcomes of the underwriting process. Cutts and Van Order (2005) show that, in the case of the subprime market, terms of the mortgage contract are determined by variations in borrower attributes. Consequently, treating mortgage terms as exogenous to the likelihood of mortgage default leads to endogeneity bias. The rest of this section discusses this endogeneity problem and the underlying theory in greater detail.

4.2 Theoretical Framework and Endogeneity Bias

Theoretical research has long emphasized the potential importance of asymmetric information in impairing the efficient operation of credit markets. There is strong evidence to suggest that loan markets, especially those marked as "nonprime", do not function according to the competitive ideal. For example, Adams et al.(2008) show how moral hazard and adverse selection in the subprime auto-loan market can significantly affect market outcomes, especially since subprime borrowers not only have imperfect or impaired credit histories but also tend to be more liquidity constrained. In this context, theoretical studies on the effect of asymmetric information in the mortgage market assume greater importance (Brueckner, 2000; Cutts and van Order, 2005). For the purposes of this paper, we draw on such theoretical work and recent empirical studies (Chiappori and Salanie, 2000; Chiappori et al. 2006) that establish the importance of asymmetric information to financial market settings.

Chiappori and Salanie (2000) show that under both adverse selection and moral hazard, one should observe a positive correlation (conditional on observables) between *risk* and *coverage*.[23] If different mortgage contracts are actually sold to observationally identical borrowers, then the

[23] Alternative approaches to testing for asymmetric information in insurance markets have been suggested in recent work (see, for example, Finkelstien and McGarry, 2006 and references therein).

frequency of default among the subscribers to a contract should increase with the LTV ratio on the mortgage. In a model of lender competition under adverse selection, where riskiness is an exogenous and unobservable characteristic of an agent, the correlation stems from the fact that high-risk agents are more likely to opt for the mortgage contract with the lower downpayment but a higher interest rate (Brueckner, 2000). Under moral hazard, the reverse causality would generate the same correlation: Borrowers buying into mortgages with higher LTV ratios for any unspecified or exogenous reasons are likely to exert less effort to repay the loan and therefore become riskier.

These theoretical results lead to the following two predictions. First, higher-risk subprime borrowers self-select into mortgage contracts that offer features (such as low downpayment), that at a given price, are more valuable to them than to lower-risk individuals. Second, equilibrium pricing on underwriting contracts reflects variation in the risk pool across different contracts. In particular, features of mortgage contracts that are selected by high-risk types should be priced more highly than those purchased by low-risk types.

Table 6 reports actual interest rates on offer for 30-year FRMs in the subprime market by Option One Mortgage Corporation in November 2007.[24] This table summarizes the actual origination process in the subprime market. Note that for a given borrower type—characterized by the borrower's credit grade and FICO score—the interest rates on offer vary with the down-payment on the loan. In other words, observably riskier borrowers are required to put up more equity to qualify for the same interest rate. Based on this outline, we can make the following inferences about the process of mortgage origination.

First, conditional on observable risk, borrowers are offered menus of contracts varying in their interest rate and LTV requirements as given in Table 6. Borrower characteristics define borrower credit grade, which together with borrower credit score, determines the menus of contracts available to the borrower. In terms of actual mortgage originations, this means that a borrower can choose among the contract terms given along a row in Table 6. Second, within the menu of contracts on offer, contracts with a higher LTV ratio typically come with a higher

[24] This table is similar to Table 4 in Cutts and Van Order (2005), which was prepared from Option One Mortgage Corporation rates effective in September 2002. Not surprisingly, differences in the two tables illustrate how mortgage originators cut back on loan offers after the downturn in this market.

rate of interest. This feature is critical to our understanding of the underwriting process. The borrower's downpayment on the mortgage determines the interest rate on the loan and vice versa. Stated differently, we can use this feature to model the determinants of a mortgage contract on either of these mortgage terms, but not both.

4.3 Estimation Strategy

Determinants of loan terms Subprime mortgage contracts are essentially summarized by the following three attributes: (1) product type (FRM or ARM), (2) LTV ratio, and (3) the interest rate (spread over prime rate) on the loan. Evidently, predictions of empirical contract theory are corroborated in terms of common practice (see Table 6): A given borrower can choose two but not all of the three terms of the mortgage contract on offer. Conditional on observable risk (as summarized from credit grade and scores), a borrower's choice of LTV ratio (and product type) determines the rate (spread) on his or her mortgage. Alternatively, the borrower's choice of monthly payment (mortgage rate) and product type, from among the menu of contracts on offer, determines the downpayment requirement (LTV ratio). Accordingly, we can focus our attention to the determinants of the mortgage contract as follows:

$$Type^* = \mathbf{X}\boldsymbol{\delta} + \delta_Z Z + v \tag{1}$$

$$Type = \text{FRM}\left[Type^* > 0\right], \tag{2}$$

$$Z = \mathbf{X}\boldsymbol{\gamma} + u, \tag{3}$$

where $\mathbf{X}$ is a vector of borrower attributes and Z is either the LTV ratio on the mortgage or the interest rate, but not both. It is important to mention here that the first and second equations are structural equations that determine product type, but the third equation is a reduced form equation for LTV ratio or interest rate.[25]

[25] See Maddala (1983, Chapter 7) and Wooldridge (2002, Chapter 15) for a discussion of discrete response models with continuous endogenous explanatory variables.

Determinants of default and delinquency To derive testable predictions about the ex post occurrence of default, we estimate the semiparametric hazard rate regression for the 90-day delinquency event. The hazard function $h(t)$ is the instantaneous probability of delinquency at age t, and is given by

$$h(t) = \lim_{\Delta t \to 0} \frac{\Pr(t \leq T < t + \Delta t | T \geq t)}{\Delta t}. \tag{4}$$

Following Cox (1972), the semiparametric representation that we estimate takes the form

$$h(t) = h_0(t) \exp(\mathbf{X}\beta), \tag{5}$$

where $h_0(t)$ is baseline hazard function.

Testing endogeneity bias For mortgages of every vintage, we set up a two-equation model, similar to the approach in Chiappori and Salanie (2000):

$$Z_i = X_i \gamma + u_i \tag{6}$$

$$h_i(t) = h_0(t) \exp(X_i \beta). \tag{7}$$

The first equation, identical to equation (3), is an ordinary least squares regression with LTV ratio (or interest rate spread) as the dependent variable. The second equation, identical to equation (5), is a Cox proportional hazard rate regression model.[26]

[26] The object of interest in a Cox proportional hazard rate regression model is hazard ratio (HR), which has the interpretation of a multiplicative change in the instantaneous probability of delinquency for a marginal change in a particular risk characteristic. HR is analogous to the odds ratio in logistic regressions. Let $h(t|X)$ be the instantaneous probability of delinquency at age t conditional on other characteristics given by vector X. We can define the estimated HR for marginal change in risk characteristic x_i as

$$\begin{aligned} \widehat{HR}(t \,|x_i = x_i + \Delta x_i\,) &= \frac{h_0(t) \exp(x_1 \widehat{\beta}_1 + x_2 \widehat{\beta}_2 + \cdots + (x_i + \Delta x_i)\, \widehat{\beta}_i + \cdots)}{h_0(t) \exp(x_1 \widehat{\beta}_1 + x_2 \widehat{\beta}_2 + \cdots + x_i \widehat{\beta}_i + \cdots)} \\ &= \exp(\Delta x_i \widehat{\beta}_i). \end{aligned} \tag{8}$$

$$h(t|X, x_i = x_i + \Delta x_i) = h(t|X) * \widehat{HR}(t \,|x_i = x_i + \Delta x_i\,)$$

The martingale residuals of the Cox model are calculated as

$$\hat{\eta}_i = \delta_i - \hat{H}_0(t)\exp(X_i\hat{\beta}), \tag{9}$$

where $\hat{H}_0(t)$ is the estimated cumulative baseline hazard rate and δ_i is an indicator that takes the value 1 when a delinquency is recorded at loan age t for mortgage i and 0 otherwise.

We estimate the two equations independently and compute the residuals $\widehat{u}_i$ and $\hat{\eta}_i$. Following Chiappori and Salanie (2000), the test statistic for the null of conditional independence $cov(\varepsilon_i, \eta_i) = 0$ is defined by

$$W = \frac{(\sum_{i=1}^{n} \widehat{u}_i \hat{\eta}_i)^2}{\sum_{i=1}^{n} \widehat{u}_i^2 \hat{\eta}_i^2}, \tag{10}$$

where W is distributed asymptotically as a $\chi^2(1)$.[27]

5 Results

5.1 The Evidence on Endogeneity Bias

The test of endogeneity bias is based on the conditional independence between the individual's choice of LTV ratio (coverage) and the ex post occurrence of the event of delinquency (risk). Table 7 shows the conditional correlation between risk and coverage under various specifications. The first specification uses the *closing rate spread* as the dependent variable in equation (8), while the second specification uses the CLTV ratio. Both specifications yield similar results: The conditional correlations for all vintages are positive and significant. The Chiappori and Salanie (2000) test statistic in equation (7) confirms the statistical significance of the results. In addition to the Chiappori and Salanie (2000) test statistic, we construct a bootstrap confidence

[27] Chiappori and Salanie (2000) use a probit equation to estimate the probability of accident in insurance markets and their test statistic is calculated by weighting each individual by days under insurance. In this case, we use the hazard rate regression for calculating the probability of default, which explicitly takes the age of the mortgage into account. Furthermore, we estimate the probit model on the event of default and the test by weighting each mortgage by the age (in months) at the time of delinquency event. The results are qualitatively similar.

interval for testing the significance of correlation (conditional on observables) between risk and coverage.[28] The bootstrap exercise further confirms that the estimated conditional correlation between risk and coverage in this market is significant and positive.

The importance of asymmetric information for subprime credit markets is corroborated by other studies (see, for example, Adams et al., 2008, on the subprime auto loan market). However, most empirical work on credit markets, like Chiappori and Salanie (2000), cannot reject the null of zero correlation between risk and coverage. It appears that for most conventional credit markets, there is little correlation between the coverage of a contract and the ex post riskiness of its subscribers (see references in Chiappori et al., 2006).[29] Therefore, it is perhaps likely that the strong endogeneity bias in subprime markets is sufficiently weaker when it comes to other mortgage markets (like that for prime mortgages). However, these results confirm the endogeneity problem that confronts the use of mortgage characteristics such as the CLTV ratio (and interest rate spread) as explanatory variables in determining loan performance.

In our regression on mortgage defaults given below, we show that ignoring this endogeneity bias leads to faulty inferences. The inclusion of endogenous variables such as the CLTV ratio (or the *closing rate spread*) as explanatory variables in a default regression introduces a positive bias on estimated coefficients. We can comment on the direction of this bias since the estimated conditional correlations are significantly positive. For explanatory variables such as the FICO score and the full-documentation dummy, one expects a negative coefficient in the hazard rate regression. Consequently, the positive bias introduced by including endogenous variables such as the CLTV ratio reduces the true impact of such explanatory variables on the probability of default.

[28]The bootstrap methodology can be described as follows. Borrower characteristics on mortgage i with LTV of z_i are denoted by X_i. Also, the age in months at which mortgage i faces the 90-day delinquency event is denoted by y_i. Constructing the bootstrap confidence interval involves the following steps:

Step 1: We draw a bootstrap sample $(z^*, y^*, X^*) = \{(z_1^*, y_1^*, X_1^*), (z_2^*, y_2^*, X_2^*), \ldots, (z_n^*, y_n^*, X_n^*)\}$ with replacement from $(z_1, y_1, X_1), (z_2, y_2, X_2), \ldots, (z_n, y_n, X_n)$.

Step 2: From the bootstrap sample estimate equations (6) and (7), recover the OLS residuals on equation (6), and the martingale residuals in equation (9); and calculate the correlation between the two estimated residuals.

Step 3: Repeat the process B times to obtain the distribution of estimated correlation between risk and coverage.

[29]The absence of a positive correlation does not necessarily imply that such markets do not suffer from asymmetric information. As Finkelstien and McGarry (2006) demonstrate, alternative tests can reveal the existence of asymmetric information along multiple dimensions.

5.2 Determinants of Mortgage Terms

Table 8 reports estimates of equation (6) for first-lien subprime originations between 2000 and 2007, with CLTV ratio as the dependent variable in Panel A and *closing rate spread* as the dependent variable in Panel B. Following our discussion in the previous section, all borrower attributes including FICO score (scaled here by a factor of 100) are included as explanatory variables but mortgage characteristics are excluded. In addition, we control for property type, property location, and lender type. The estimation results can be summarized as follows:

(1) We observe a scale effect in subprime underwriting. For higher-valued properties, borrowers have lower CLTV ratios on average, presumably because doing so lowers the exposure for lenders. This is reflected in the progressively lower coefficients for properties in higher-valued quartiles, showing that mortgages on properties with higher values have, on average, a lower CLTV ratio. Not surprisingly, originations on lower-valued properties with consequently higher CLTV ratios, have higher interest rates.

(2) Owner-occupied homes have significantly higher CLTV ratios and lower rates than non-owner occupied homes. Here, too, underwriting seems to have succeeded in getting non-owners (i.e., investors) to make greater downpayments on loans of identical size.

(3) Mortgages with full documentation have significantly higher CLTV ratios and lower rates than low- or no-doc loans. But the size of the CLTV coefficients in Panel A declines over the sample period. Evidently, underwriters' effort at tempering low-documentation loans with lower CLTVs, on average, was probably weakening over the years. However, originations with lower documentation required higher mortgage rates over the years, as seen from the size of the interest rate coefficients in Panel B.

(4) Borrowers with higher FICO scores are also the ones with higher CLTV ratios. But here the trend of adjustment of FICO scores with lower CLTV ratios seems to have grown stronger over the years. Also, equilibrium mortgage rates are lower for borrowers with higher FICO scores.

(5) No cash-out refinances have lower CLTV ratios than purchases. This is hardly surprising

given the property price appreciation for most of our sample period. However, refinances (both cash-out and no cash-out) have lower CLTV ratios and lower mortgage rates than purchase originations. This result is explained below.

It is interesting to compare the signs on the coefficients in the CLTV regression (Table 8) with those in the FICO regression (see table 3). Given our a priori judgment of risk characteristics, the signs on the coefficients seem to indicate evidence of credible underwriting. For example, note that while full-documentation is associated with a lower FICO score, borrowers providing full documentation on loans are allowed to make a lower downpayment. The important exception is the sign of coefficients on loan purpose. Although borrowers' FICO scores are lower on average for refinances, these refinances also have lower CLTV ratios. Typically, loans are refinanced with the original lender, and, because of a recorded payment history, mortgage refinances are considered to be less risky a priori. This could explain the lower FICO score on refinances. Explaining the CLTV result requires a more nuanced view of subprime originations: Gorton (2008) shows that in the event of house price appreciation lenders can benefit even from a refinancing option, so long as the borrower does not extract the full amount of the appreciated value.[30] This implies that lenders try to ensure that the borrowers retain sufficient equity in the property on a refinance, which could explain why refinances have lower CLTV ratios, on average, than purchases.

In summary, our results indicate that the underwriting process attempted to adjust riskier borrower characteristics with lower CLTV ratios (and higher mortgage rates). Again, there is little evidence to suggest any significant deterioration in underwriting standards after 2004.

5.3 Determinants of Mortgage Default

Table 9 reports the estimated hazard ratios for the Cox proportional hazard rate regressions in equation (7). Here too, we control for borrower attributes, lender characteristics, property

[30] The lender now faces a less-risky borrower who has built up equity in the house. Gorton (2008, p.150) argues that subprime mortgages, the majority of which were hybrid ARMs, were designed "to provide an implicit embedded option on house prices for the lender." Unwilling to speculate on house prices and borrower repayment behavior for long periods, lenders treated subprime mortgages as bridge financing and sought the option to end the mortgage early. As a result, the fully-indexed rate is designed to be prohibitively high once it resets from the teaser rate, thereby essentially forcing a refinancing.

type, and property location. Panel A reports the hazard ratios for borrower characteristics excluding mortgage terms. Panel B includes mortgage terms such as CLTV ratio and the *closing rate spread* as explanatory variables. Clearly, a priori beliefs about the effect of individual borrower characteristics on credit risk are validated; originations with full documentation have a significantly lower probability of default than low-doc or no-doc loans. For example, the hazard ratio on the full-doc dummy variable for the 2003 vintage (Panel A) indicates that the default probability of a loan with full documentation is 0.7451 times the default probability on the origination with low-documentation for the same vintage. Likewise, a higher CLTV ratio increases the probability of default: We estimate a 1.0194 times increase in the probability of default for an increase in one percentage point of the CLTV in 2003 (Panel B).[31] In the same manner, the likelihood of default on the mortgage is reduced if the property is owner-occupied rather than for investment purposes and if the loan originated is a refinance as opposed to a direct purchase. Finally, within originations of the same vintage, those with higher FICO scores have a significantly lower probability of default than those with lower FICO scores. The model provides a good fit of the data. In appendix C, we report the Kaplan-Meier survivor function and the model-implied survivor function for the vintages 2005-2007.

To confirm our earlier summary results in Section 3, we estimate the same regression by using dummy variables for each of seven different FICO score groups. The groups selected for this regression are the same as those given in Table 5. The hazard ratios are provided in Table 10 with the lowest FICO score group (< 540) chosen as the base group. This procedure enables us to assign default probabilities across the various FICO groups and helps answer questions about the effectiveness of FICO scores across the various vintages. We estimate the probabilities of default for a FICO group as the product of the (actual) probability of default for the base group (for each vintage) times the hazard ratio for the FICO score group (see appendix B for details).

Table 11 reports the increases in probability of surviving a 90-day delinquency for originations in a higher FICO score group relative to those in its immediate lower FICO score group, after controlling for other attributes on the origination. The results show that after controlling

[31] This implies on average, the probability of default of a 2003 origination with a CLTV ratio of 85 percent is ($1.0194^5 =$) 1.1008 times the probability of default of a 2003 origination with a CLTV ratio of 80 percent.

for other attributes on the origination (as given by the regressions in Table 10), the increases in survival probabilities show a significant improvement over the vintages. First, this result holds across all transitions between adjacent FICO score groups, irrespective of whether they are low and high. Second, the results are robust even if mortgage terms such as CLTV ratio or *closing rate spread* are included as explanatory variables (Panel B). However, while the inclusion of mortgage terms dampens the effectiveness of FICO as evidenced by the lower increases in Panel B, the trend is not reversed.

Comparing the survival probabilities in Table 5 with those in Table 11 reveals an interesting trend. In Section 3.2, we documented deterioration in performance over the vintages for lower FICO originations (rows 1 and 2 in Table 5). However, after controlling for other attributes, Table 11 shows that this trend is reversed for lower FICO originations. Similarly, we recorded a sharp improvement in performance of FICO for higher FICO originations (rows 4 through 6 in Table 5). Whereas controlling for other attributes on the origination dampens this improvement in Table 11, the trend is not reversed. These trends can be explained in terms of our earlier results that there was an attempt to adjust riskier attributes with higher FICO scores and that this adjustment strengthened for the later vintages. Stated differently, originations of later vintage with higher FICO scores are more likely to have riskier attributes on average. Controlling for these riskier attributes would dampen the trend of improvement in FICO performance as seen from Table 11. Conversely, lower FICO originations on more recent vintages are less likely to have riskier attributes on average. Therefore, in controlling for these attributes, their improvement in performance is sufficiently large to reverse the earlier trend of deterioration in performance.

Viewed independently, the evidence in Tables 10 and 11 reveals little about underwriting standards. On the other hand, when these regression results are examined in conjunction with the other results in Tables 3, 8, and 11, a clearer picture of underwriting standards emerges. Earlier, we showed evidence to suggest that the underwriting process attempted to adjust riskier borrower characteristics with higher FICO scores (Section 3.1) and lower CLTVs (Section 5.2). These results also suggest that lenders adjusted higher CLTV ratios with higher FICO scores and that the strength of adjustment increased over the years. In this section, the hazard rate estimation shows that, ceteris paribus, FICO scores are an important determinant of ex post default. Taken together, there is significant evidence of credible mortgage underwriting on the basis of hard data available: Lenders tried to offset greater risk in terms of higher CLTV ratios and lower documentation by raising FICO score requirements at the time of loan origination because FICO scores are an important determinant of ex post default.

A final comment involves the use of mortgage terms (such as CLTV ratio or mortgage rate) in mortgage default estimation. Our results on endogeneity argue that the inclusion of such terms as explanatory variables would lead to biased estimates. This is best illustrated in terms of the differences in the estimates between Panels A and B for Tables 9 and 11. Including mortgage terms such as CLTV ratio and *closing rate spread* introduces a positive bias on the estimates of explanatory variables such as FICO scores (and the dummy variable for full documentation), thereby reducing the impact of FICO scores as a determinant of ex post default. This is clearly evident from the higher hazard ratios (Panel B of Tables 9 and 10) and consequently lower improvements in default probabilities (Panel B of Table 11).

6 Counterfactual Analysis

From the standpoint of mortgage and borrower characteristics as well as ex post default, observable underwriting trends do not provide evidence of a secular decline in lending standards. Moreover, there is no discernible change for post-2004 originations. On the contrary, we find evidence of credible underwriting in terms of the right direction of adjustment (higher FICO scores on low-doc originations) and some evidence to suggest this adjustment was strengthened over the years. However, we have yet to determine whether the adjustment was "adequate" in terms of its magnitude. At the heart of this analysis is the problem of aggregating a multidimensional profile of borrower attributes to a single metric that could summarize the overall credit risk of the borrower. Although this would help determine whether underwriting standards declined over this period, we are not aware of a direct solution to this problem.

In this section, we attempt to cope with this problem by using a counterfactual exercise. In so doing, we answer the following question: How would ex post default rates change if a mortgage that originated to a representative borrower in 2005 had originated in 2001? To this end, we estimate the proportional hazard rate model for a particular vintage and then use the estimated relationship to evaluate the *estimated proportional hazard survivorship function* for a representative borrower from a different vintage (see Cameron and Trivedi, 2006, for further details).

Let v be the index of vintage, $S_{v,0}(t)$ be the baseline survivor function, and $\mathbf{X}$ be the observable characteristic of the "representative borrower" of vintage v. The survivor function $S_v(t)$, for any vintage v and age of mortgage t, is the outcome of a mapping of observable borrower characteristics, $\mathbf{X}$, and unobservable characteristics and market conditions captured by baseline survivor function, $S_{v,0}(t)$:

$$S_v(t) = f\left(S_{v,0}(t), \mathbf{X}\right),$$

where function f maps $\left(S_{v,0}(t), \mathbf{X}\right)$ into the range of $S_v(t)$.

For our purposes, the objective is to forecast the impact on the survivor function of vintage v_2 in the environment of vintage v_1.[32] In this specification, let $\mathbf{X}_1$ and $\mathbf{X}_2$ denote the representative borrowers of vintages v_1 and v_2, respectively. If unobservable characteristics and market conditions captured by the baseline survivor function are applied to the different borrower characteristics, we can identify the effect of $\mathbf{X}_2$ on the survivor function in v_1 as follows:

$$S_{v_1}^{v_2}(t) = f\left(S_{v_1,0}(t), \mathbf{X}_2\right).$$

Such a counterfactual exercise helps us test the following hypothesis:

Null Hypothesis: Let $S_v(t)$ be the survivor function for vintage v and age of mortgage t; and $S_v^{\widetilde{v}}(t)$ be the counterfactual survivor function, which is the result of the forecasting problem described above, then $S_v(t) \approx S_v^{\widetilde{v}}(t)$, for all t.

We proceed as follows: First, we estimate the Cox proportional hazard model in equation (7) for a given vintage v. Next, we calculate the estimated survivor function for the represen-

[32] This problem is similar to **P-2** on program evaluation in Heckman and Vyltacil (2007).

tative borrower of vintage v. Finally, we calculate the counterfactual survivor function for the representative borrower of a different vintage, say $\widetilde{v}$. Because our representative borrower is constructed to best reflect borrower characteristics of a particular vintage, we define characteristics of this representative borrower as follows. Any attribute of the representative borrower of vintage v is calculated as the average of the values of the attribute of all borrowers who originated loans in year v. Therefore, if 28.6 percent of the sample had low- or no-documentation loans in 2002, the value of the "dummy" variable on documentation for 2002 vintage would be 0.286. Clearly, this is an oddity, but it is a simple way to summarize the distribution of borrower characteristics.[33]

With these tools in place, we can now use our counterfactual analysis to test the null hypothesis that there was no dramatic weakening of underwriting standards beginning around late 2004. The null hypothesis is that mortgages approved after 2004 are equally as likely to survive an event of default as those of earlier vintages–namely, 2001, 2002, and 2003–in the environment of these vintages.[34] The results of counterfactual analysis are summarized in Table 12. Table 12 has three panels corresponding to the counterfactual exercises using survivor function estimates based on 2001, 2002, and 2003 data. The numbers in parentheses are the 95 percent confidence intervals for the estimated survivor function. The results show that if a representative borrower in 2006 (likewise for 2005 and 2007) had originated mortgages in 2001 and 2002, she would have significantly better loan performance than representative borrowers of vintages 2001 and 2002 respectively. The counterfactual survival function using 2003 estimates shows that the loan performance of the representative borrower of 2006 vintage would have been worse than that of the representative borrower of the current (2003) vintage. However, there are no statistically significant differences in loan performance between the representative borrowers of 2005 or 2007 vintages and that of the 2003 vintage.

[33] Needless to say, the results of this counterfactual analysis are sensitive to the definition of the "representative borrower" of a particular vintage. To test the robustness of our results, we adopt an alternative procedure. We adopt the first step as before. In the second step, we recover the estimated survivor function for all borrowers in year v. In the third step, we calculate the counterfactual survivor function for all borrowers who originated loans in year $\tilde{v}$. A final step involves averaging across all borrowers of a given vintage to obtain the actual and the counterfactual survivor functions for years v and $\tilde{v}$, respectively. The results are qualitatively similar.

[34] Our choice of years on the counterfactual is motivated by the fact that the information set of the lender for post-2004 originations should arguably include the repayment behavior on 2001 and 2002 originations. Moreover, we conduct a reverse counterfactual analysis by examining survival functions for originations of 2001 and 2002 for vintages 2005-2007. The results are presented in appendix C.

These results are best illustrated in terms of the survival plots in the upper panel of Figure 2. As discussed above, we can reject the null hypothesis in favor of the alternative that the underwriting standards actually improved in the latter vintages compared with 2001 and 2002 vintages. Originations of 2003 vintage perform significantly better than originations of 2006 but not better than 2005 and 2007.

To check the robustness of our results, we conduct a similar counterfactual analysis, this time including CLTV ratio and *closing rate spread* as explanatory variables for the counterfactual estimates. As mentioned earlier, doing so introduces an endogeneity bias to our estimates. But we proceed nevertheless and the survival plots are shown in the lower panel of Figure 2. In comparison, the second counterfactual exercise reduces the differences in loan performance across the vintages. However, even with the inclusion of loan characteristics, the results of the counterfactual exercise remain robust. Evidently, the origination of mortgages with high CLTV ratios in later vintages did not have a significant impact in terms of ex post default. In summary, the counterfactual analysis is strong evidence against the hypotheses that a weakening of underwriting standards can explain recent defaults in subprime mortgages.

7 Discussion and Assessment

We fail to find evidence of deterioration in underwriting standards for later vintages of securitized subprime mortgages. Moreover, in light of the evidence, it is difficult to conclude that underwriting was central to the collapse of the subprime mortgage market. This non-result is a significant departure from conventional wisdom on the subprime crisis. However, it is not difficult to see why a discerning reader may not find this result implausible. The argument that a significant deterioration in underwriting after 2004 triggered the collapse of the subprime market implicitly suggests that originations of earlier vintages had relatively robust underwriting. Taken to its logical conclusion, it could also suggest that the underwriting framework for earlier vintages could help provide a sustainable framework for subprime originations for the future. In contrast, our results do not rule out the possibility that the design of subprime contracts could be flawed since the inception of this market.

There is sufficient evidence to suggest that this might indeed be true. Gorton (2008) enumerates the reasons underwriting to subprime borrowers would require a fundamental change to underwriting standards compared with other prime markets. Moreover, as he argues, if the interest rate on the mortgage is set to price the risk, such a rate is not likely to be affordable to these borrowers. Adams et al.(2008) show that the interest rates on subprime auto loans are significantly higher than those on subprime mortgages. As Gorton (2008) demonstrates, the subprime mortgage design embedded a price appreciation that made this market extremely dependent on home price appreciation. Bhardwaj and Sengupta (2009) show how prepayments were integral to the design on subprime contracts and how the subprime boom was sustained by high and early prepayments during a period of considerable house price appreciation (also see appendix A).

As with any empirical study of this kind, there are limitations in our study. First, it is extremely important to state that our conclusions are drawn from data available at the time of loan origination. Subsequent behavior of the borrower (e.g., originating a second lien on the property) is undeniably important in determining ex post delinquency and default. However, this would hardly provide a basis for determining a decline in underwriting at origination. Second, it is possible that there were borrower attributes observed by the lender but not reported in the LP data. Lack of data often hinders a conclusive argument on some important characteristics, for example, the debt-to-income ratio. Using different data, Mian and Sufi (2008) report that aggregate mortgage debt-to-income ratios for entire zip codes have increased significantly in the borrower population. However, using the debt-to-income ratios in the LP database on individual mortgages creates significant problems. First, there are very few data on the front-end debt-to-income ratio. Second, even for the back-end ratio, the field is sparsely populated for earlier vintages in the LP data. For the data that are available, we observe a trend of increasing (back-end) debt-to-income ratios. Again, our regression results show attempts to control for this increase by increasing other borrower attributes, namely, the FICO score. Appendix C presents the evidence on debt-to-income ratios.

Third, some observers may doubt the veracity of the data. Some anecdotal evidence points to poor reporting, false documentation, and outright fraud.[35] However, it is difficult to make this case for a significant proportion of a repository of more than 9 million loan observations. Fourth, it needs to be mentioned that our examination of the underwriting standards is at the level of the individual borrower and not at the level of the lending institution. We do not examine the hypothesis if, for example, the fraction of originations with high CLTV ratios were disproportionately high for a particular lending institution. Fifth, it is important to note that our sample includes mortgages that have been securitized as subprime. Our sample does not include loans that were classified as subprime but retained by originators in their portfolios. Therefore, loans that default even before they can be securitised are not part of our dataset. Since our data do not cover these very early defaults before securitization, we have a possibile selection bias in our results.

Finally, as is well known, the guidelines for classification into the subprime and Alt-A categories vary by arranger. There is significant evidence that points to a deterioration of underwriting standards in Alt-A mortgages (Sengupta, 2009). Because both Alt-A and subprime mortgages are likely to have the same originators, this result at first pass may seem implausible. However, a plausible explanation to our findings on securitized subprime mortgages might be explained in terms of anecdotal evidence on subprime. In their handbook chapter on Alt-A mortgages, Bhattacharya et al. (2006, p. 189) remark that "the demarcation between Alt-A and subprime loans has been blurred. Over time Alt-A has expanded to include loans with progressively less documentation and lower borrower credit scores. At the same time, subprime loans have, on average experienced a slow but steady rise in average credit scores. A result of this convergence has been the creation of the so-called alt-B sector, where loans using this nomenclature were securitized in 2004."

[35] Federal investigators are probing into allegations of fraud and misrepresentations by mortgage companies such as Countrywide Financial Corp. See, for example, "Loan Data Focus of Probe" (Wall Street Journal, March, 11 2008).

8 Conclusion

This paper presents a contrarian perspective on underwriting standards in the subprime market. Our examination of the LP data on securitized subprime originations shows scant evidence of a decline in underwriting standards. Moreover, our counterfactual analysis demonstrates that, at least on average, we can reject the hypothesis of no decline in underwriting standards in favor of improvement in underwriting standards. Of course, we cannot reject the premise that underwriting standards in the subprime market were poor to begin with. However, the question remains: What sustained the phenomenal growth in the subprime market for nearly a decade. And, of course, why did the subprime market collapse?

Gorton (2008) argues that the subprime mortgage contracts were designed as bridge financing to give temporary credit accommodation to borrowers in anticipation of future earnings growth, buildup of borrower equity through a rise in house prices, or both. In a similar vein, Bhardwaj and Sengupta (2009) show that, for early vintages, a significantly high proportion of subprime borrowers used early prepayments as an exit option from mortgage obligations. These early prepayments were largely sustained by the boom in house prices in the United States from 1995 to 2006. However, a sudden reversal in house price appreciation increased default in this market because it made this prepayment exit option cost prohibitive. Most important, high early defaults on post-2004 originations can be explained when one takes into account the high early prepayment rates for the pre-2004 vintages.

References

Adam, W., Einav, L., and Levin, J. (2009). Liquidity Constraints and Imperfect Information in Subprime Lending. American Economic Review. 99(1), 49-84.

Bhardwaj, G. and Sengupta, R. (2009). Did Prepayments Sustain the Subprime Mortgage market? Federal Reserve Bank of St. Louis Working Paper.

Bhattacharya, A., Berliner, W., and Lieber, J. (2006). Alt-A Mortgages and MBS. In: Fabozzi, F. (Ed.), The Handbook of Mortgage-Backed Securities. McGraw-Hill: New York.

Brueckner, J. K. (2000). Mortgage default with asymmetric information. Journal of Real Estate Finance and Economics, Springer, 20(3), 251-74.

Cameron, C. and Trivedi, P. (2005) Microeconometrics: Methods and Applications. Cambridge University Press, New York.

Chiappori, P., Jullien, B. , Salanie B., and Salanie, F. (2006). Asymmetric Information in Insurance: General Testable Implications. Rand Journal of Economics, 37(4), 2006.

Chiappori, P., and Salanie, B. (2000). Testing for Asymmetric Information in Insurance Markets. Journal of Political Economy, 108(1), 56-78.

Cox, D.R. (1972). Regression Models and Life-Tables (with Discussion). Journal of the Royal Statistical Society, Series B, 34, 187–220.

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

Dell'Arricia, G., Igan, D., and Laeven, L.(2008) Credit Booms and Lending Standards: Evidence from the Subprime Mortgage Market. IMF working paper: 08/106.

Demyanyk, Y. and Van Hemert, O. (2009). Understanding the subprime mortgage Crisis. Review of Financial Studies, forthcoming.

Elul, R. (2009), Securitization and Mortgage Default: Reputation vs. Adverse Selection. Federal Reserve Bank of Philadelphia, Working Paper.

Fabozzi, F.J. ed. (2006). The Handbook of Mortgage-Backed Securities. 6th edition. New York: McGraw-Hill.

Finkelstein, A. and McGarry, K. (2006) Multiple Dimensions of Private Information: Evidence from the Long-Term Care Insurance Market. American Economic Review 96(4) 938-958.

Foust, D. and Pressman, A. (2008). Credit Scores: Not-So-Magic Numbers. BusinessWeek, Issue 4701, 38-43.

Gerardi, K., Lehnert, A., Sherland, S., and Willen, P. (2009) Making Sense of the Subprime Crisis . Brookings Papers on Economic Activity, forthcoming.

Gorton, G. (2008). The Panic of 2007. Manuscript Prepared for the Federal Reserve Bank of Kansas City, Jackson Hole Conference, August 2008 .

Gorton, G. and He, P. (2008) Bank Credit Cycles. Review of Economic Studies, 75(4) 1181-1214.

Heckman, J.J. and Vyltacil, E.J. (2007). Econometric Evaluation of Social Programs, Part I: Causal Models, Structual Models and Econometric Policy Evaluation", Handbook of Econometrics, Vol 6B, Elsevier.

Kaplan, E. L., and Meier, P. (1958). Nonparametric Estimation from Incomplete Observations. Journal of the American Statistical Association, 53, 457-48.

Keys, B., Mukherjee, T., Seru, A. and Vig, V. (2009). Securitization and Screening: Evidence from Subprime Mortgage Backed Securities, mimeo, London Business School.

Maddala, G.S. (1983). Limited dependent and quantitative variables in econometrics. Cambridge: Cambridge University Press.

Mayer, C., and Pence K. (2008) Subprime Mortgages: What, Where and to Whom?, NBER Working Paper No. W14083.

Mayer, C., Pence, K., and Sherlund, S. (2009). The Rise in Mortgage Defaults. Journal of Economic Perspectives, 23(1), 27-50.

Mian, A., and Sufi, A. (2008). The Consequences of Mortgage Credit Expansion: Evidence from the 2007 Mortgage Default Crisis. Unpublished manuscript, University of Chicago Graduate School of Business.

Pennington Cross, A., and Chomsisengphet, S. (2007), Subprime Refinancing: Equity Extraction and Mortgage Termination. Real Estate Economics. 35(2), 233-263.

President's Working Group on Financial Markets (2008). Policy Statement on Financial Market Developments.

Sengupta R. (2009). A Brief Overview of the Alt-A Mortgage Market. Working Paper. Forthcoming.

Stein, J. (2002) Information Production and Capital Allocation: Decentralized versus Hierarchical Firms. The Journal of Finance, 57 (5). 1891-1921

Wooldridge, J. (2002) Econometric Analysis of Cross Section and Panel Data. Cambridge and London: MIT Press.

Figure 1: Probability of Default by Loan Age (in months)
The plot shows the Kaplan-Meier default probabilities by loan age for securitized first-lien subprime mortgages. The graph presents the default probabilities by years of originations (vintage) from 2000 through 2007. Each line shows the performance of originations of the same vintage. The details of the methodology for calculating the default probabilities are provided in appendix B.

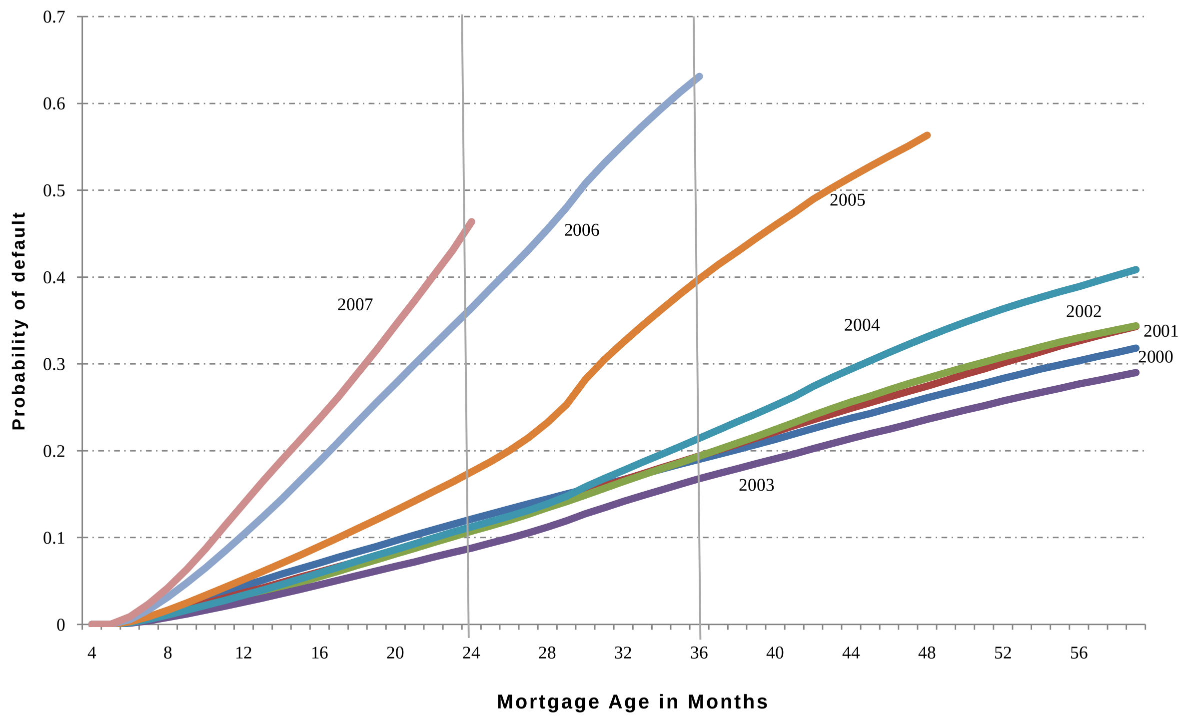


Table 1: FICO Distribution Conditional on Documentation Level on Loan by Vintage

Borrower credit score at the time of loan origination is denoted by FICO™ (an industry standard developed by the *Fair Isaac Corporation*) with a number in the range 300 to 850. Loans coded by the source with a non-blank documentation code are classified as *Full doc*, whereas those under a no documentation program or prospectus are classified as *No doc*. Others are classified as *Low doc*.

	Full-doc loans				Low-doc or No-doc loans			
	FICO Scores				FICO scores			
Vintage	< 620	620-659	660-719	≥ 720	< 620	620-659	660-719	≥ 720
1998	65.6	18.9	11.5	4.0	56.7	21.7	16.2	5.4
1999	67.4	18.4	10.9	3.3	53.3	22.1	18.2	6.4
2000	72.1	16.9	8.6	3.3	59.1	21.3	15.0	4.6
2001	67.8	18.8	10.0	3.3	50.2	25.2	18.7	5.8
2002	64.4	20.2	11.4	4.0	42.1	27.2	23.2	7.5
2003	58.4	22.2	13.9	5.4	37.3	27.4	26.2	9.1
2004	58.8	22.5	13.7	5.0	38.0	27.8	26.1	8.1
2005	58.8	23.2	13.6	4.5	34.5	30.1	26.9	8.6
2006	61.3	23.7	11.5	3.4	35.7	32.3	24.9	7.1
2007	60.5	24.7	11.9	2.9	42.3	30.2	22.1	5.4

Table 2: Distribution of FICO Scores Conditional on CLTV by Vintage

Borrower credit score at the time of loan origination is denoted by FICO score with a number in the range 300 to 850. The Cumulative Loan to Value (CLTV) Ratio is the proportion of loans (secured by the property) on all liens in relation to the property's value.

	CLTV ≤ 80				80 < CLTV ≤ 90				90 < CLTV ≤ 100			
Vintage	< 620	620-659	660-719	≥ 720	< 620	620-659	660-719	≥ 720	< 620	620-659	660-719	≥ 720
1998	63.2	18.4	12.5	5.9	61.9	21.1	12.5	4.4	52.2	22.2	17.2	8.4
1999	65.1	18.0	12.1	4.8	63.9	20.6	11.8	3.7	44.2	23.5	23.1	9.2
2000	70.4	16.5	9.9	3.2	71.1	18.1	8.6	2.2	48.1	29.3	17.1	5.5
2001	66.0	18.1	11.7	4.2	65.8	21.0	10.6	2.6	44.0	30.8	18.9	6.3
2002	62.0	19.3	13.6	5.2	61.8	21.9	12.9	3.4	30.2	36.1	25.2	8.5
2003	59.2	19.4	15.1	6.3	55.8	23.7	15.7	4.7	30.2	33.6	26.5	9.7
2004	61.9	19.2	13.7	5.2	57.5	23.2	15.0	4.3	31.0	32.9	27.0	9.0
2005	60.7	20.6	13.8	5.0	55.9	23.6	15.8	4.7	32.7	33.2	25.9	8.2
2006	65.1	19.7	11.4	3.9	60.4	23.3	12.9	3.4	34.8	35.4	23.3	6.4
2007	68.3	19.2	9.8	2.7	57.7	26.3	13.2	2.7	31.4	37.3	25.1	6.3

Table 3: Credit Score (FICO) Regression

This table reports OLS estimates with borrower FICO score as the left-hand side variable and other borrower characteristics as regressors. We control for property type (dummies for single-family residence, condo, townhouse, cooperative, etc.), property location (dummies for the state in which the property is located), loan source (dummies for broker, realtor, wholesale, retail, etc.) and number of units in the property. *Home Value n*th *Quartile* is a dummy that equals 1 if the value of the property lies in the *n*-th quartile of all property values in the data and 0 otherwise. The results for the years of origination 1998 and 1999 are not reported here but are available on request.

A. Controlling for All Attributes on the Origination (excluding Mortgage Terms)

Variable	**2000**	**2001**	**2002**	**2003**	**2004**	**2005**	**2006**	**2007**
Intercept	644.57***	667.17***	697.95***	716.68***	682.69***	702.73***	704.42***	704.73***
Full- Documentation	-15.14***	-18.49***	-22.07***	-19.44***	-17.74***	-18.87***	-19.26***	-17.25***
Owner-Occupied	-26.88***	-24.34***	-27.6***	-32.46***	-33.76***	-32.11***	-31.48***	-32.79***
Second Home	-3.71***	-3.28***	-8.51***	-12.86***	-14.46***	-7.58***	-8.26***	-15.37***
Refinance (Cash-Out)	-16.93***	-16.77***	-28***	-34.38***	-37.17***	-34.44***	-33.26***	-31.71***
Refinance (No Cash-Out)	-19.12***	-17.8***	-20.23***	-22.11***	-22.37***	-19.62***	-18.64***	-23.8***
Home Value First Quartile	-7.29***	-13.36***	-11.25***	-13.56***	-13.18***	-14.11***	-13.99***	-12.55***
Home Value Second Quartile	-5.38***	-9.2***	-7.35***	-8.87***	-8.25***	-8.25***	-8.96***	-8.56***
Home Value Third Quartile	-3.63***	-5.47***	-5.76***	-7.27***	-6.7***	-6.31***	-6.71***	-5.48***
Adjusted R^2	0.0766	0.0877	0.1336	0.1529	0.1684	0.1698	0.1766	0.1486

B. Including the Mortgage Terms: CLTV Ratio and Closing Rate Spread

Variable	**2000**	**2001**	**2002**	**2003**	**2004**	**2005**	**2006**	**2007**
Intercept	811.27***	851.48***	869.49***	882.52***	817.88***	793.96***	793.01***	830.85***
Closing Rate Spread	-17.69***	-22.61***	-24.23***	-27.54***	-25.72***	-24.04***	-17.7***	-20.48***
CLTV Ratio	0.32***	0.46***	0.72***	0.85***	0.97***	0.95***	0.95***	1.08***
Full- Documentation	-18.49***	-20.7***	-25.96***	-26.16***	-27.17***	-29.13***	-28.9***	-28.11***
Owner-Occupied	-29.37***	-29.6***	-33.71***	-38.5***	-42.24***	-46.18***	-42.45***	-44.59***
Second Home	-3.13***	-6.83***	-15.35***	-17.72***	-17.34***	-13.27***	-12.51***	-18.21***
Refinance (Cash-Out)	-16.05***	-16.88***	-22.91***	-23.63***	-22.18***	-18.26***	-17.77***	-18.96***
Refinance (No Cash-Out)	-18.42***	-18.37***	-19.8***	-17.91***	-15.68***	-13.07***	-10.38***	-16.83***
Home Value First Quartile	10.83***	14.15***	12.3***	10.65***	7.37***	3.37***	-0.042	4.01***
Home Value Second Quartile	3.58***	5.76***	4.93***	3.91***	1.04***	-1.61***	-2.96***	-1.81***
Home Value Third Quartile	1.27***	3.2***	1.01***	-0.73***	-2.12***	-2.77***	-3.59***	-2.85***
Adjusted R^2	0.2211	0.3294	0.3944	0.4320	0.4180	0.4106	0.3878	0.4124

The symbols ***, **, and * denote statistical significance at 1 percent, 5 percent, and 10 percent levels respectively.

Table 4: Fully Interacted Dummy Variable Regression of Credit Score (FICO) on Other Borrower Characteristics

This table reports OLS estimates of a fully interacted dummy variable regression of borrower FICO scores on other borrower attributes, for all the vintages (1998 onwards) pooled together; the dummy variable is turned on for latter vintages. We report four versions of this equation where dummy variable is turned on for post-2002 to post-2005 vintages.

	Dummy = 1 if vintage is …			
Variable	Post-2002	Post-2003	Post-2004	Post-2005
Intercept	671.18***	679.97***	680.24***	685.41***
Dummy	21.59***	9.90***	21.77***	19.15***
Full-Documentation	-19.52***	-20.77***	-20.03***	-20.01***
Full-Documentation x Dummy	0.89***	2.33***	1.15***	1.05***
Owner-Occupied	-26.39***	-28.46***	-30.48***	-30.83***
Owner-Occupied x Dummy	-6.20***	-4.12***	-1.40***	-0.79***
Second Home	-8.59***	-10.73***	-12.30***	-10.15***
Second Home x Dummy	-2.07***	0.46	3.90***	0.78
Refinance (Cash-Out)	-18.34***	-24.33***	-29.66***	-31.35***
Refinance (Cash-Out) x Dummy	-16.53***	-10.61***	-4.20***	-1.77***
Refinance (No Cash-Out)	-16.60***	-19.47***	-22.36***	-22.80***
Refinance (No Cash-Out) x Dummy	-4.19***	-1.18***	2.49***	2.76***
Home Value First Quartile	-8.2***	-7.48***	-8.19***	-9.36***
Home Value First Quartile x Dummy	-5.32***	-6.08***	-5.68***	-4.35***
Home Value Second Quartile	-5.08***	-4.48***	-4.88***	-5.48***
Home Value Second Quartile x Dummy	-3.26***	-3.83***	-3.58***	-3.39***
Home Value Third Quartile	-3.82***	-4.30***	-4.86***	-5.13***
Home Value Third Quartile x Dummy	-2.76***	-2.17***	-1.52***	-1.39***
Adjusted R^2	0.1549	0.1482	0.1440	0.1422

The symbols ***, **, and * denote statistical significance at 1 percent, 5 percent, and 10 percent levels respectively.

Table 5: Increase in Survival Probabilities for Improvements in FICO Score (Groups)
The numbers show percentage point increases in the Kaplan-Meier survival probabilities (for the first two years after origination) of originations in the higher FICO score group relative to those in the lower FICO score group. The Kaplan-Meier probabilities are calculated as discussed in appendix B. The FICO score groups used below are "< 540", "540-579", "580-619" … "700-739" and "≥740".

	Vintage							
FICO score	**2000**	**2001**	**2002**	**2003**	**2004**	**2005**	**2006**	**2007**
[< 540] to [540 – 579]	8.17	7.46	5.75	4.17	4.73	4.95	5.52	6.12
[540 – 579] to [580 – 619]	4.45	4.24	3.57	3.38	3.88	3.04	1.68	3.51
[580 – 619] to [620 – 659]	3.35	2.87	2.91	3.24	4.48	4.33	2.10	3.72
[620 – 659] to [660 – 699]	1.95	2.37	2.54	2.43	2.79	4.59	4.64	5.05
[660 – 699] to [700 – 739]	1.41	1.44	1.96	1.52	1.50	2.56	4.14	3.38
[700 – 739] to [≥740]	0.91	1.10	0.81	0.84	1.30	2.57	7.84	10.13
Average all	3.37	3.25	2.92	2.60	3.11	3.68	4.32	5.32
Average first five	3.87	3.68	3.35	2.95	3.48	3.90	3.62	4.35

Table 6: Mortgage Pricing Sheet, Option One Mortgage Corporation

The rate sheet is for fixed rate mortgages with two-year prepayment charge. The worksheet assumes full documentation, one unit house, and loan amount in the range $200,000 to $417,000. In case of secondary financing (CLTV > LTV) and credit score less than 660 (or ≥660) rate is adjusted upward by 155 basis points (or 90 basis points). For a similar table showing the rates available in 2002, see Table 4 in Cutts and van Order (2005).

		LTV			
Grade	**FICO Score**	**65**	**70**	**75**	**80**
	700+	8.65	8.70	8.80	8.90
	660	8.75	8.80	8.90	9.00
AA+	620	9.00	9.05	9.15	9.25
	580	9.55	9.60	9.90	10.05
	540	10.45	10.70	10.90	11.15
	700+	9.35	9.40	9.50	9.60
	660	9.45	9.50	9.60	9.70
AA	620	9.70	9.75	9.85	9.95
	580	10.15	10.20	10.35	10.50
	540	10.70	10.95	11.00	11.25
	700+	9.45	9.50	9.60	9.70
	660	9.55	9.60	9.70	9.80
A	620	9.80	9.85	9.95	10.05
	580	10.25	10.30	10.45	10.60
	540	10.80	11.05	11.10	11.35
	700+	9.85	9.95	10.10	10.25
	660	10.05	10.15	10.35	10.45
B	620	10.40	10.55	10.75	10.80
	580	10.95	11.00	11.25	11.35
	540	11.55	11.7	11.95	

Source: Option One Mortgage Corporation, west area rate sheet, effective 11/09/2007, downloaded on 07/03/2008, http://www.oomc.com/broker/broker_rateguide.asp.

Table 7: Test of Endogeneity Bias
Tabulated entries are the estimated correlation coefficients (conditional on observables) between risk and coverage as described in Chiappori and Salanie (2000). We conduct tests on two specifications, one with CLTV and the other with *Closing Rate Spread* as the dependent variable in equation (8).

	Dependent Variable in Equation (8)	
Vintage	**Closing Rate Spread**	**CLTV**
2000	0.13	0.04
2001	0.10	0.06
2002	0.10	0.05
2003	0.09	0.07
2004	0.08	0.09
2005	0.10	0.14
2006	0.11	0.19
2007	0.19	0.19

Table 8: Determinants of Mortgage Terms

The table reports OLS estimates for different mortgage terms as the dependent variable. FICO scores are scaled by a factor of 100. We control for property type (dummies for single-family residence, condo, townhouse, cooperative, etc.), property location (dummies for the state in which the property is located), and loan source (dummies for broker, realtor, wholesale, retail, etc.). *Home Value n*th *Quartile* is a dummy that equals 1 if the value of the property lies in the *n*-th quartile of all property values in the data and 0 otherwise. The results for the years of origination 1998 and 1999 are not reported here but are available on request.

A. Dependent Variable: CLTV

Variable	**2000**	**2001**	**2002**	**2003**	**2004**	**2005**	**2006**	**2007**
FICO (scaled)	1.73***	2.01***	3.02***	3.49***	4.41***	4.86***	5.21***	6.16***
Full- Documentation	5.6***	4.49***	3.34***	2.84***	1.74***	1.32***	0.9***	1.47***
Owner Occupied	4.14***	4.52***	4.77***	5.67***	5.48***	5.09***	5.49***	6.06***
Second Home	-0.5**	-1.48***	-0.31*	-0.81***	-0.78***	0.08	0.13	0.65**
Refinance (Cash Out)	-8.06***	-8.28***	-7.54***	-10.35***	-11.24***	-12.27***	-13.87***	-13.64***
Refinance (No Cash Out)	-6***	-6.04***	-5.11***	-8.25***	-9.39***	-9.1***	-9.88***	-10.73***
Home Value First Quartile	0.03	2.35***	3.4***	4.63***	4.18***	3.58***	2.93***	4.03***
Home Value Second Quartile	0.75***	2.5***	3.24***	4.07***	3.62***	2.92***	2.33***	2.82***
Home Value Third Quartile	0.68***	2.27***	2.91***	3.06***	2.46***	1.48***	1.25***	1.82***
Adjusted R^2	0.1409	0.1473	0.1595	0.2423	0.2901	0.3077	0.3357	0.3108

B. Dependent Variable: Closing Rate Spread

Variable	**2000**	**2001**	**2002**	**2003**	**2004**	**2005**	**2006**	**2007**
FICO (scaled)	-0.88***	-1.22***	-1.16***	-1.14***	-1.05***	-1.04***	-1.11***	-1.11***
Full-Documentation	-0.3***	-0.31***	-0.29***	-0.36***	-0.43***	-0.51***	-0.68***	-0.69***
Owner-Occupied	-0.4***	-0.43***	-0.4***	-0.43***	-0.51***	-0.67***	-0.73***	-0.75***
Second Home	0.1***	0.01	-0.2***	-0.21***	-0.27***	-0.29***	-0.41***	-0.34***
Refinance (Cash-Out)	-0.16***	-0.32***	-0.28***	-0.26***	-0.18***	-0.11***	-0.18***	-0.46***
Refinance (No Cash-Out)	-0.12***	-0.26***	-0.24***	-0.21***	-0.21***	-0.17***	-0.16***	-0.49***
Home Value First Quartile	0.98***	1.07***	0.83***	0.83***	0.79***	0.75***	0.73***	0.82***
Home Value Second Quartile	0.46***	0.56***	0.44***	0.45***	0.39***	0.35***	0.36***	0.38***
Home Value Third Quartile	0.23***	0.32***	0.25***	0.22***	0.18***	0.15***	0.16***	0.17***
Adjusted R^2	0.2800	0.4121	0.4194	0.4306	0.4055	0.4256	0.3741	0.3980

The symbols ***, ** and * denote statistical significance at 1-percent, 5-percent, and 10-percent levels respectively.

Table 9: Cox Proportional Hazard Rate Regression: Hazard Ratio for 90-day Delinquency Event
This table reports the estimated hazard ratios for the Cox proportional hazard rate regressions conducted for all loans originated in a given calendar year. We control for property type (dummies for single-family residence, condo, townhouse, cooperative, etc.), property location (dummies for the state in which the property is located), and loan source (dummies for broker, realtor, wholesale, retail, etc.). *Home Value n*th *Quartile* is a dummy that equals 1 if the value of the property belongs to the *n*-th quartile of all property values in the data and 0 otherwise. The results for the years of origination 1998 and 1999 are not reported here but are available on request.

A. Controlling for All Attributes on the Origination (excluding Mortgage Terms)

Variable	**2000**	**2001**	**2002**	**2003**	**2004**	**2005**	**2006**	**2007**
FICO score	0.9920***	0.9918***	0.9912***	0.9900***	0.9909***	0.9928***	0.9940***	0.9942***
Full-Documentation	0.8754***	0.8624***	0.8219***	0.7451***	0.7518***	0.6922***	0.6517***	0.6657***
Owner-Occupied	0.8076***	0.8022***	0.8127***	0.7825***	0.7493***	0.7725***	0.7729***	0.7611***
Second Home	0.6302***	0.5463***	0.5738***	0.6072***	0.5989***	0.7045***	0.6922***	0.6896***
Refinance (Cash-Out)	0.7625***	0.6605***	0.6414***	0.5419***	0.5164***	0.5015***	0.5558***	0.5738***
Refinance (No Cash-Out)	0.919***	0.7927***	0.7477***	0.5829***	0.5341***	0.539***	0.5975***	0.5831***
Home Value First Quartile	0.9015***	0.9301***	0.9493***	1.0221**	0.871***	0.7009***	0.6303***	0.6466***
Home Value Second Quartile	0.9321***	0.9054***	0.9142***	0.9663***	0.8432***	0.7029***	0.6761***	0.7003***
Home Value Third Quartile	0.9351***	0.911***	0.8966***	0.9359***	0.8689***	0.8512***	0.8362***	0.852***
LR test H_0: $\beta = 0$	22077	27461	44013	83203	123586	155671	157703	23435
(p-value)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)

B. Including the Mortgage Terms: CLTV Ratio and Closing Rate Spread

Variable	**2000**	**2001**	**2002**	**2003**	**2004**	**2005**	**2006**	**2007**
FICO score	0.9934***	0.9937***	0.9932***	0.9917***	0.9917***	0.9937***	0.9943***	0.9955***
CLTV ratio	1.0099***	1.0148***	1.0137***	1.0194***	1.024***	1.0266***	1.0293***	1.0265***
Full-Documentation	0.8513***	0.8191***	0.8183***	0.7487***	0.7799***	0.7517***	0.6971***	0.7411***
Closing Rate Spread	1.2381***	1.2268***	1.2428***	1.218***	1.2049***	1.2214***	1.1562***	1.284***
Owner-Occupied	0.8341***	0.8078***	0.827***	0.7717***	0.7301***	0.7797***	0.7266***	0.7496***
Second Home	0.5991***	0.5814***	0.6356***	0.6371***	0.6573***	0.7467***	0.7128***	0.7071***
Refinance (Cash-Out)	0.8439***	0.8161***	0.7698***	0.7013***	0.6942***	0.7284***	0.8624***	0.9063***
Refinance (No Cash-Out)	1.0294***	0.9495***	0.8756***	0.7261***	0.6962***	0.743***	0.8394***	0.8858***
Home Value First Quartile	0.7236***	0.7194***	0.7373***	0.7556***	0.6505***	0.5377***	0.5236***	0.4686***
Home Value Second Quartile	0.8395***	0.7841***	0.7822***	0.793***	0.702***	0.6077***	0.6038***	0.5919***
Home Value Third Quartile	0.8763***	0.8335***	0.8112***	0.8297***	0.7789***	0.7911***	0.7886***	0.7796***
LR test H_0: $\beta = 0$	35418	45374	72023	129707	208878	307473	333416	55519
(p-value)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)

The symbols ***, **, and * denote statistical significance at 1 percent, 5 percent, and 10 percent levels, respectively.

Table 10: Cox Proportional Hazard Rate Regression: Hazard Ratio for 90-day Delinquency Event

This table reports the estimated hazard ratios for the Cox proportional hazard rate regressions conducted for all loans originated in a given calendar year. FICO scores below 540 are treated as the base group. We control for property type (dummies for single-family residence, condo, townhouse, cooperative, etc.), property location (dummies for the state in which the property is located), and loan source (dummies for broker, realtor, wholesale, retail, etc.). *Home Value n*th *Quartile* is a dummy that equals 1 if the value of the property belongs to the *n*-th quartile of all property values in the data and 0 otherwise. The results for the years of origination 1998 and 1999 are not reported here but are available on request.

A. Controlling for All Attributes on the Origination (excluding Mortgage Terms)

Variable	**2000**	**2001**	**2002**	**2003**	**2004**	**2005**	**2006**	**2007**
FICO: 540-579	0.6671***	0.6923***	0.6874***	0.7147***	0.7305***	0.7619***	0.7946***	0.7832***
FICO: 580-619	0.5071***	0.5258***	0.5125***	0.5215***	0.5491***	0.6092***	0.6719***	0.6437***
FICO: 620-659	0.368***	0.3822***	0.3712***	0.3459***	0.3713***	0.4521***	0.5229***	0.5122***
FICO: 660-699	0.2606***	0.2516***	0.2391***	0.2062***	0.2449***	0.3284***	0.3998***	0.3909***
FICO: 700-739	0.1841***	0.1666***	0.1506***	0.127***	0.1674***	0.2549***	0.3269***	0.3314***
FICO: $\geq$740	0.1459***	0.1012***	0.0997***	0.0784***	0.0996***	0.1755***	0.2322***	0.241***
Full-Documentation	0.8758***	0.8654***	0.8274***	0.7482***	0.7541***	0.6944***	0.653***	0.6685***
Owner-Occupied	0.8169***	0.806***	0.8192***	0.7827***	0.7523***	0.7782***	0.7772***	0.7667***
Second Home	0.6345***	0.5482***	0.5752***	0.6085***	0.6013***	0.7064***	0.6955***	0.6933***
Refinance (Cash-Out)	0.7641***	0.6603***	0.6435***	0.5483***	0.5256***	0.5088***	0.5644***	0.5789***
Refinance (No Cash-Out)	0.9241***	0.7979***	0.7519***	0.59***	0.5401***	0.543***	0.6027***	0.5867***
Home Value First Quartile	0.9129***	0.9397***	0.9569***	1.0244**	0.8755***	0.7044***	0.6329***	0.6485***
Home Value Second Quartile	0.9387***	0.908***	0.9191***	0.9674***	0.8445***	0.7046***	0.6774***	0.7013***
Home Value Third Quartile	0.9351***	0.911***	0.8966***	0.9359***	0.8689***	0.8512***	0.8362***	0.852***
LR test H_0: $\beta = 0$	21281	26772	42589	81926	121051	151229	155037	23036
(p-value)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)

B. Including the Mortgage Terms: CLTV Ratio and Closing Rate Spread

Variable	**2000**	**2001**	**2002**	**2003**	**2004**	**2005**	**2006**	**2007**
FICO: 540-579	0.71***	0.7568***	0.7628***	0.7631***	0.7617***	0.7853***	0.7924***	0.8385***
FICO: 580-619	0.5725***	0.6195***	0.6111***	0.5747***	0.5748***	0.6268***	0.6263***	0.6736***
FICO: 620-659	0.4494***	0.49***	0.4764***	0.409***	0.399***	0.464***	0.4799***	0.5529***
FICO: 660-699	0.3422***	0.3513***	0.3313***	0.2628***	0.2729***	0.3468***	0.3784***	0.4535***
FICO: 700-739	0.2579***	0.2536***	0.221***	0.1712***	0.1919***	0.2765***	0.3188***	0.4086***
FICO: ≥740	0.22***	0.1734***	0.1621***	0.1161***	0.1244***	0.2032***	0.2432***	0.3226***
CLTV Ratio	1.0106***	1.015***	1.0142***	1.0188***	1.0235***	1.0276***	1.03***	1.0265***
Full-Documentation	0.8528***	0.8342***	0.8171***	0.7475***	0.7764***	0.7345***	0.6889***	0.7309***
Closing Rate Spread	1.2468***	1.2414***	1.2601***	1.2408***	1.2325***	1.2405***	1.179***	1.3051***
Owner-Occupied	0.8339***	0.8056***	0.8266***	0.7642***	0.7331***	0.7722***	0.7337***	0.7644***
Second Home	0.6363***	0.5645***	0.6168***	0.6504***	0.6489***	0.7436***	0.7167***	0.7143***
Refinance (Cash-Out)	0.8389***	0.7763***	0.7473***	0.6916***	0.6856***	0.7111***	0.8394***	0.8904***
Refinance (No Cash-Out)	1.003	0.9151***	0.8479***	0.7124***	0.6813***	0.7139***	0.8109***	0.8538***
Home Value First Quartile	0.7238***	0.7148***	0.7351***	0.7593***	0.6521***	0.5401***	0.511***	0.4608***
Home Value Second Quartile	0.8356***	0.7746***	0.784***	0.7983***	0.7054***	0.6083***	0.5972***	0.5883***
Home Value Third Quartile	0.8771***	0.8224***	0.8119***	0.837***	0.7841***	0.7906***	0.7847***	0.7774***
LR test H_0: $\beta = 0$	26986	33339	52166	96343	150898	218605	226958	39158
(p-value)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)

The symbols ***, **, and * denote statistical significance at 1 percent, 5 percent, and 10 percent levels respectively.

Table 11: Parametric Estimates of Increase in Survival Probabilities for Transitions between Different FICO Score Groups

The numbers show parametric estimates of percentage point increases in the Kaplan-Meier survival probabilities (for the first two years after origination) of originations in the higher FICO score group relative to those in the lower FICO score group. The estimated probabilities are calculated after controlling for other attributes on the origination as given in appendix B. The FICO score groups used below are "< 540", "540-579", "580-619" … "700-739" and "≥740".

A. Controlling for All Attributes on the Origination (excluding Mortgage Terms)

FICO Score	**2000**	**2001**	**2002**	**2003**	**2004**	**2005**	**2006**	**2007**
[<540] to [540 – 579]	7.56	6.63	6.27	4.99	5.83	6.75	9.36	12.57
[540 – 579] to [580 – 619]	3.63	3.59	3.51	3.38	3.92	4.33	5.59	8.09
[580 – 619] to [620 – 659]	3.16	3.09	2.83	3.07	3.85	4.45	6.79	7.62
[620 – 659] to [660 – 699]	2.44	2.81	2.65	2.44	2.73	3.50	5.61	7.03
[660 – 699] to [700 – 739]	1.74	1.83	1.77	1.38	1.68	2.08	3.32	3.45
[700 – 739] to [≥740]	0.87	1.41	1.02	0.85	1.47	2.25	4.32	5.24
Average all	3.23	3.23	3.01	2.69	3.25	3.89	5.83	7.33
Average first five	3.70	3.59	3.41	3.05	3.60	4.22	6.13	7.75

B. Including the Mortgage Terms: CLTV Ratio and Closing Rate Spread

FICO Score	**2000**	**2001**	**2002**	**2003**	**2004**	**2005**	**2006**	**2007**
[<540] to [540 – 579]	6.58	5.24	4.76	4.14	5.15	6.08	9.45	9.36
[540 – 579] to [580 – 619]	3.12	2.96	3.04	3.29	4.04	4.49	7.57	9.56
[580 – 619] to [620 – 659]	2.79	2.79	2.70	2.90	3.80	4.61	6.67	6.99
[620 – 659] to [660 – 699]	2.43	2.99	2.91	2.56	2.73	3.32	4.62	5.76
[660 – 699] to [700 – 739]	1.91	2.10	2.21	1.60	1.75	1.99	2.71	2.60
[700 – 739] to [≥740]	0.86	1.73	1.18	0.96	1.46	2.08	3.44	4.99
Average all	2.95	2.97	2.80	2.58	3.16	3.76	5.75	6.54
Average first five	3.37	3.22	3.12	2.90	3.50	4.10	6.21	6.86

Table 12: Counterfactual Survival Analysis

Three panels report numbers corresponding to counterfactual exercise using survivor function estimates based on 2001, 2002, and 2003 data. The numbers in the parentheses are lower and upper confidence limits at 95 percent confidence interval for the estimated survivor function.

Panel A: Counterfactual Analysis 2001

Age of Loan (Months)	Survivor Function 2001	Counterfactual Survivor Function 2005	Counterfactual Survivor Function 2006	Counterfactual Survivor Function 2007
12	0.966	0.969	0.969	0.968
	(0.965,0.967)	(0.969,0.97)	(0.968,0.97)	(0.968,0.969)
24	0.903	0.913	0.911	0.91
	(0.902,0.905)	(0.912,0.914)	(0.91,0.913)	(0.908,0.911)
36	0.832	0.848	0.846	0.843
	(0.83,0.834)	(0.846,0.851)	(0.843,0.848)	(0.841,0.845)
48	0.764	0.786	0.782	0.779
	(0.761,0.767)	(0.783,0.789)	(0.779,0.785)	(0.776,0.782)
60	0.702	0.728	0.724	0.72
	(0.698,0.706)	(0.724,0.732)	(0.72,0.727)	(0.716,0.723)

Panel B: Counterfactual Analysis 2002

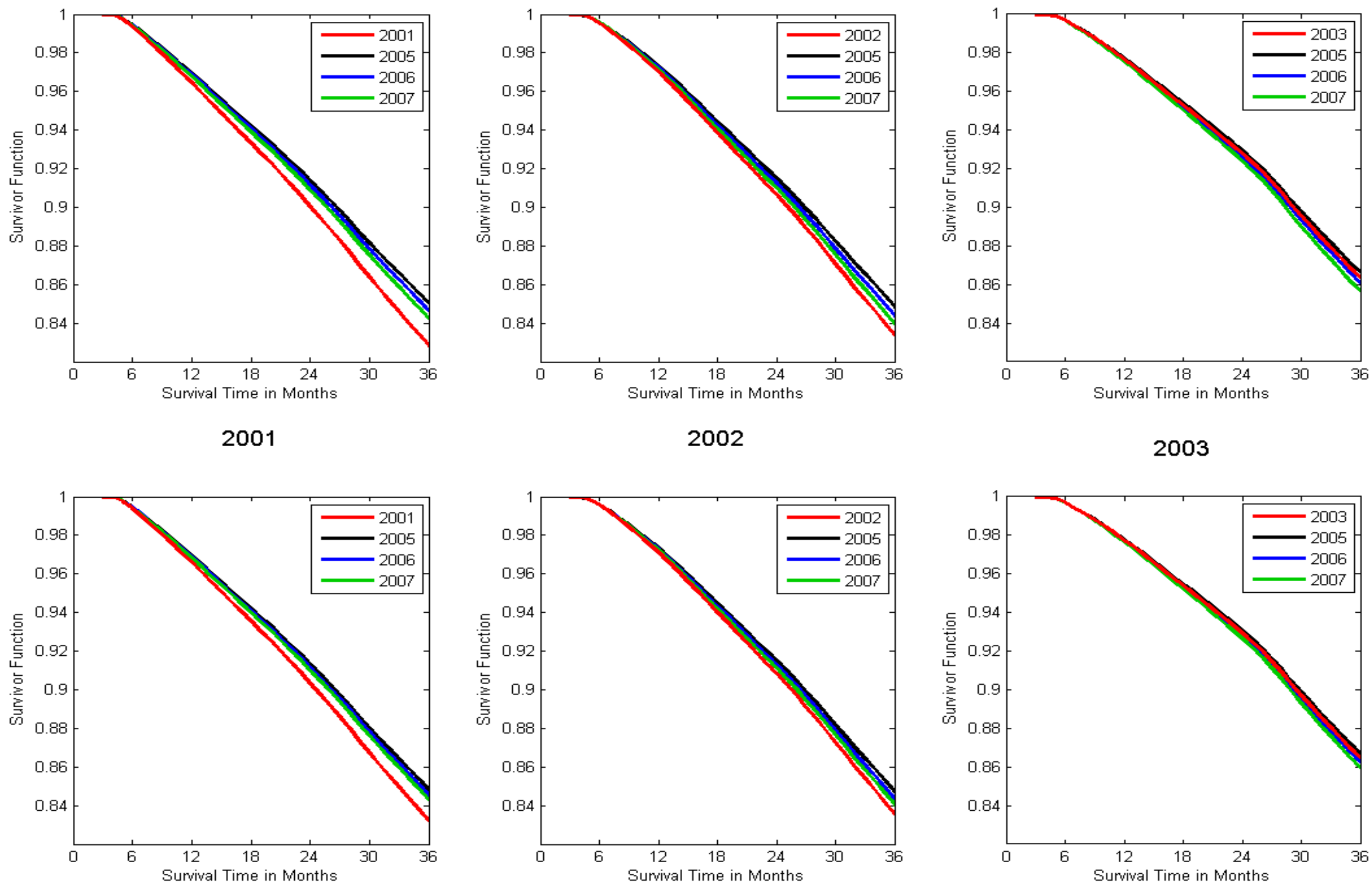
Age of Loan (Months)	Survivor Function 2002	Counterfactual Survivor Function 2005	Counterfactual Survivor Function 2006	Counterfactual Survivor Function 2007
12	0.971	0.973	0.972	0.972
	(0.97,0.971)	(0.973,0.974)	(0.972,0.973)	(0.971,0.972)
24	0.908	0.915	0.913	0.911
	(0.907,0.909)	(0.914,0.916)	(0.912,0.914)	(0.91,0.912)
36	0.835	0.847	0.843	0.84
	(0.834,0.837)	(0.845,0.849)	(0.842,0.845)	(0.839,0.842)
48	0.759	0.775	0.77	0.766
	(0.756,0.761)	(0.772,0.778)	(0.767,0.772)	(0.763,0.768)
60	0.7	0.72	0.713	0.708
	(0.697,0.703)	(0.716,0.723)	(0.71,0.716)	(0.705,0.711)

Panel C: Counterfactual Analysis 2003

Age of Loan (Months)	Survivor Function 2003	Counterfactual Survivor Function 2005	Counterfactual Survivor Function 2006	Counterfactual Survivor Function 2007
12	0.977	0.978	0.977	0.977
	(0.977,0.978)	(0.978,0.978)	(0.977,0.977)	(0.976,0.977)
24	0.929	0.93	0.928	0.926
	(0.928,0.93)	(0.93,0.931)	(0.927,0.929)	(0.926,0.927)
36	0.865	0.867	0.862	0.859
	(0.863,0.866)	(0.866,0.868)	(0.861,0.864)	(0.858,0.861)
48	0.804	0.807	0.801	0.796
	(0.802,0.806)	(0.805,0.809)	(0.799,0.803)	(0.794,0.798)
60	0.746	0.75	0.743	0.737
	(0.744,0.749)	(0.748,0.753)	(0.74,0.745)	(0.735,0.74)

Figure 2: Counterfactual Analysis for 2001-2003 vintage

The figures show the estimated proportional hazard survivorship function for representative borrowers from different vintages. The three columns correspond to the counterfactual exercises using survivor function estimates based on 2001, 2002, and 2003 data. The upper panel shows results for the counterfactual exercise with borrower characteristics only as regressors, whereas mortgage terms such as LTV ratio and the closing rate spread are added as regressors for the counterfactual results shown in the lower panel.



Appendix A: An alternative explanation for early defaults

This aim of this section is to show that high early defaults on post-2004 vintages might be better explained in terms of high early prepayment rates for the pre-2004 originations. Prepayments include either a refinancing or an outright sale of the property. In essence, a prepayment can be an exit option for distressed borrowers, and this option can be used either before the borrower becomes delinquent on the mortgage as a means to avoid delinquency or after delinquency as a means to avoid foreclosure.

To provide a more detailed picture of these trends, we track the following loan status variables over time: delinquency, prepayment, and foreclosure. Panel A of Table A.1 provides the percentage of loans that register a delinquency (as a fraction of total originations) within the first two calendar years after the year of origination for each vintage (year of origination). We record 30-day and a 60-day delinquencies over these periods separately. Next, we determine the percentage of delinquent loans that were either prepaid (Panel B) or went into foreclosure (Panel C) within the same period. Table A.1 presents these summary results for owner-occupied subprime originations.

The pattern that emerges is robust across occupancy categories, loan purpose, and product type and is summarized as follows:

1. Both 30-day and 60-day delinquencies are higher for loans that originated after 2004.

2. Loans that register a 30-day delinquency are more likely to be prepaid than loans that record a 60-day delinquency in the same period.

3. There is an increase in foreclosure rates on delinquent loans for originations after 2004, especially during 2006 and 2007.

4. Finally, and perhaps most important, there is a sharp decline in prepayment rates for post-2004 originations, especially during 2006 and 2007.

The patterns across the 30-day and 60-day delinquencies are similar. Interestingly, the total of post-delinquency early prepayments and foreclosures (obtained by adding the percentages in Panels B and C of Table A.1) for each vintage does not reveal an increasing trend. Stated differently, the percentages of delinquent loans that were either prepaid or went into foreclosure are not significantly different over our sample period. Remarkably, however, there is significant change in the composition. A sharp drop in the proportion of early prepayments is almost always accompanied by a sharp rise in the proportion of foreclosures for post-2004 originations, especially during 2006 and 2007.

However, this pattern still does not explain why delinquencies are marginally higher for post-2004 originations. To answer this query, we study pre-delinquency repayment behavior by adopting the same approach as above. Figures A.1 through A.3 show the percentage of mortgages for every vintage that either (1) registered a 30-day delinquency or (2) were prepaid *before* recording a 30-day delinquency within the first 18 months since origination. In addition, we calculate the total fraction of loans in either of these categories.

Note that the duration of study is different from Table A.1 in that we observe each loan for the first 18 months since origination rather than up to a specific calendar date. But just as observed previously, the total proportion of early prepayments and delinquencies is not significantly different over our sample period. Moreover, the pattern appears to be the same as that observed in Table A.1: A significantly large fraction of total originations is prepaid early in the subprime market, and this fraction drops for originations after 2004, especially during 2006 and 2007. Most important, these graphs show that these findings are robust across variations in product type (Figure A.1), occupancy (Figure A.2), and loan purpose (Figure A.3).

To summarize, we find a distinct pattern of high early prepayment behavior for earlier vintages followed by a sharp drop in early prepayment rates for later (post-2004) vintages, especially during 2006 and 2007. This result holds regardless of whether we track originations over a given loan age or particular calendar dates, whether we consider post-delinquency or pre-delinquency behavior, or whether we study repayment across the different categories of occupancy, product type, and loan purpose. After delinquency, the total fraction of loans that either go into foreclosure or are paid off remain roughly the same for all vintages. Before delinquency, the total fraction of loans that register a delinquency or get paid off before registering a delinquency is not significantly different for all vintages. Our data do not permit us to determine the cause of prepayment, but prepayment is an exit option available to a distressed borrower unable to make mortgage payments. Moreover, given that a significant proportion of subprime originations came with prepayment penalties for two or more years, the trend in early prepayments is indeed suggestive of prepayments under distress.

Why are early prepayments important to our analysis? Much of this relates to the dominant explanation behind subprime defaults. The dominant explanation for the subprime crisis has been that a severe weakening in underwriting standards occurred over the last few years, which eventually caused a downturn in this market. This section shows that examining high early prepayment rates for the pre-2004 originations seems a more fruitful line of research in explaining high early defaults on post-2004 vintages.

Appendix B: The product limit estimator for mortgage defaults

B1. Kaplan-Meier Survival Probabilities

The *survivor function,* or the probability of surviving default (a 90-day delinquency event), beyond loan age t is given by $S(t) \equiv P(T > t)$, where T denotes the duration in months from the month of origination. Following this, we define the default rate $D(t)$ at month t (the age of the mortgage in months) as

$$D(t) \equiv 1 - P(T > t). \tag{1}$$

Let $t_1 < t_2 < ... < t_k$ denote the observed age in months at the time of default (a 90-day delinquency event) in a sample size of N originations, $N \geq k$. Also, let n_j be the number of surviving mortgages just prior to month t_j. A surviving mortgage is defined as one that has neither defaulted nor been paid-off prior to age t_j. If we define d_j as the number of mortgages that default at age t_j, then the Kaplan-Meier estimator of the survivor function is

$$\hat{S}(t) \equiv \hat{P}(T > t) = \prod_{j|t_j \leq t} (1 - \frac{d_j}{n_j}). \tag{2}$$

B2. Parametric Estimates of Survival Probabilities

The object of interest in a Cox proportional hazard rate regression model is hazard ratio (HR), which is interpreted as a multiplicative change in the instantaneous probability of delinquency for a marginal change in a particular risk characteristic. Let $h(t|X)$ be the instantaneous probability of delinquency at age t conditional on other characteristics given by vector X. We can define the estimated HR for marginal change in risk characteristic x_i as

$$\begin{aligned} \widehat{HR}(t\,|x_i = x_i + \Delta x_i\,) &= \frac{h_0(t)\exp(x_1\widehat{\beta}_1 + x_2\widehat{\beta}_2 + \cdots + (x_i + \Delta x_i)\,\widehat{\beta}_i + \cdots)}{h_0(t)\exp(x_1\widehat{\beta}_1 + x_2\widehat{\beta}_2 + \cdots + x_i\widehat{\beta}_i + \cdots)} \\ &= \exp(\Delta x_i \widehat{\beta}_i) \end{aligned} \tag{3}$$

$$h(t|X, x_i = x_i + \Delta x_i) = h(t|X) * \widehat{HR}(t\,|x_i = x_i + \Delta x_i\,).$$

We begin by splitting our sample into various groups or intervals of FICO scores, $G_1, ..., G_n$. To do this we define n FICO group dummies, $FICOd_1, ..., FICOd_n$, such that $FICOd_k = 1$ if the origination FICO score lies in the interval G_k, $k = 1, ..., n$ and zero otherwise. To calculate the estimated survival probabilities for each FICO score group, we first estimate the HR for each group dummy, with the lowest group as the base group. The estimated HR for a given FICO score group, say G_k, is given by

$$\widehat{HR}(t|FICOd_k = 1, X) = \exp(\hat{\beta}_{FICOd_k=1})$$

where $\hat{\beta}_{FICOd_k=1}$ is the coefficient of the regression for the FICO score group G_k (or the FICO score dummy $FICOd_k = 1$).

The instantaneous probability of delinquency at age t conditional on other attributes on the origination given by vector X is

$$\hat{h}(t|FICOd_k = 1, X) = h(t|FICOd_1 = 1)^* \widehat{HR}(t|FICOd_k = 1, X),$$

where $h(t|FICOd_1 = 1)$ is the instantaneous probability of delinquency at age t of the base group. We use the following estimate of this probability for the base group

$$\hat{h}(t|FICOd_1 = 1) = \frac{d_{FICOd_1=1}(t)}{n_{FICOd_1=1}(t)},$$

where $d_{FICOd_1=1}(t)$ is the number of delinquencies at age t with FICO scores in the interval G_1, and $n_{FICOd_1=1}(t)$ is the number of number of surviving mortgages (not in default or prepaid) at age t with FICO scores in the interval G_1.

Appendix C: Robustness checks

C1. FICO Scores, Risk Characteristics and Default Risk

We begin with robustness checks for Table 5 and Table 11. We use different specifications of FICO score groups with different "starting" FICO scores at 520, 521, 540 and 541. For each of last two "starting" FICO scores, we use a finer demarcation with groups at smaller intervals of 20 points each. The results are reported in Tables C1.5A-C1.5F, which replicate Table 5 for each different specification and Tables C1.11A-C1.11F, which replicate Table 11 for each different specification. Evidently the pattern outlined in the paper is robust for all these different specifications of FICO scores.

C2. The Back-End Debt-to-Income Ratios

As mentioned in Section 6, our results do not include the back-end debt-to-income ratios. There are very few data on the front-end debt-to-income ratio. Moreover, the back-end debt-to-income ratio field is sparingly populated for earlier vintages. Given the selection issues involved in reporting this field, the lack of these data for earlier vintages poses a problem for this particular study. Nevertheless, we present the results for the major tables in the paper. Table C2.3 presents the FICO regression with the debt-to-income ratio (just as Table 3 in the paper). Note that for earlier vintages the coefficient is negative but turns positive for the later vintages suggesting credible underwriting for the later years of origination. Table C2.8 presents the OLS estimates of origination attributes for the CLTV ratio and *closing rate spread* (just as Table 8 in the paper). Since we consider the back-end debt-to-income ratio there is a strong correlation between this variable and CLTVs in Panel A. Also, there is a positive sign on the spread coefficient, indicating that higher debt-to-income ratios are associated with higher spreads in Panel B. Finally, we replicate our results on the default regressions including the back-end ratios as Table C2.9. As expected, higher debt-to-income ratios increase the likelihood of default, but there is no discernible trend in this effect.

C3. Model Fit

Our test of model fit is provided as three plots for 2005-2007. For each year of origination, the plots include the Kaplan-Meier survivor function for the particular vintage as defined in Section 6. The other two plots are the estimated survivor functions corresponding to the two estimated models reported in Table 10, Panel A (including all borrower characteristics) and Panel B (including borrower characteristics plus CLTV ratio and *closing rate spread*).

C4. The Reverse Counterfactual

It is not implausible to assume that the repayment behavior on originations of earlier vintages was observable to the lender in making originations for later vintages. Naturally, our aim was to discover whether later originations would perform at least as well as earlier originations. This in turn motivated our choice of counterfactual to ascertain whether later originations would have performed at least as well as those of previous vintages. Here, we look at the reverse counterfactual. That is, we ask the question: How would ex post default rates change if a mortgage that originated to a "representative borrower" in 2001 had originated in 2005? We perform the same counterfactuals for 2005, 2006 and 2007. Again, as Figures C4.2A through C4.2C show, we find little evidence counter to our earlier assertion. We reject the null hypothesis that a representative borrower of 2001(2002) vintage would have performed significantly better than representative borrowers of 2005-2007 vintages if the mortgages were originated in 2005, 2006, and 2007.

Table A.1: Repayment Behavior of Owner-Occupied Households (up to two calendar years from year of origination).

The delinquency rate is based on the percentage of total loans in the sample. We consider loans that are both 30 days and 60 days delinquent. Among the loans that are delinquent, we select those that were prepaid and those that went into foreclosure. We do this separately for loans that were 30 days and 60 days delinquent. Panel B shows the prepayment rate for delinquent loans, the number of loans prepaid expressed as a percentage of loans that are delinquent in each category. Panel C shows the foreclosure rate on delinquent loans, the number of loans foreclosed expressed as a percentage of loans that are delinquent in each category.

Panel A. Delinquency Rate
(as a fraction of total loans)

	30-day delinquency			60-day delinquency		
Vintage	After 2 years	After 3 years	After 4 years	After 2 years	After 3 years	After 4 years
1998	0.21	0.28	0.32	0.08	0.12	0.14
1999	0.26	0.33	0.37	0.10	0.14	0.17
2000	0.31	0.37	0.40	0.13	0.17	0.20
2001	0.33	0.39	0.41	0.13	0.18	0.20
2002	0.33	0.37	0.39	0.13	0.17	0.19
2003	0.27	0.31	0.32	0.11	0.14	0.15
2004	0.29	0.33	0.35	0.13	0.17	0.19
2005	0.34	0.42		0.19	0.28	
2006	0.46			0.34		
2007*	0.40			0.30		

Panel B. Prepayment Rate for Delinquent Loans
(as a fraction of delinquent loans)

	30-day delinquency			60-day delinquency		
Vintage	After 2 years	After 3 years	After 4 years	After 2 years	After 3 years	After 4 years
1998	0.22	0.36	0.45	0.13	0.21	0.26
1999	0.22	0.36	0.46	0.13	0.20	0.26
2000	0.24	0.39	0.49	0.13	0.20	0.26
2001	0.26	0.44	0.54	0.15	0.24	0.30
2002	0.32	0.51	0.60	0.18	0.28	0.35
2003	0.35	0.53	0.60	0.20	0.30	0.35
2004	0.33	0.49	0.52	0.19	0.26	0.27
2005	0.22	0.27		0.10	0.11	
2006	0.09			0.03		
2007*	0.04			0.01		

Panel C. Foreclosure Rate for Delinquent Loans
(as a fraction of delinquent loans)

	30-day delinquency			60-day delinquency		
Vintage	After 2 years	After 3 years	After 4 years	After 2 years	After 3 years	After 4 years
1998	0.11	0.15	0.19	0.29	0.36	0.42
1999	0.11	0.16	0.19	0.29	0.39	0.42
2000	0.14	0.19	0.22	0.35	0.40	0.44
2001	0.12	0.17	0.20	0.29	0.37	0.41
2002	0.11	0.16	0.19	0.29	0.36	0.39
2003	0.11	0.15	0.17	0.27	0.33	0.37
2004	0.11	0.17	0.20	0.26	0.34	0.38
2005	0.17	0.28		0.30	0.42	
2006	0.27			0.36		
2007*	0.19			0.25		

* For 2007 vintage, the data are available for only one calendar year and five months. Source: FALP.

Figure A.1: Pre-delinquency Repayment Behavior by Product Type (up to loan age of 18 months)

The first bar for each vintage shows the proportion of total loans that register a 30-day delinquency within the first 18 months. The second bar for each vintage shows the percentage of loans that are paid off even before registering a 30-day delinquency. The third bar for each vintage show the total of the first two columns. The percentages are shown for fixed rate mortgages (FRM), adjustable rate mortgages with a two year teaser rate (ARM2) and adjustable rate mortgages with a three year teaser rate (ARM3).

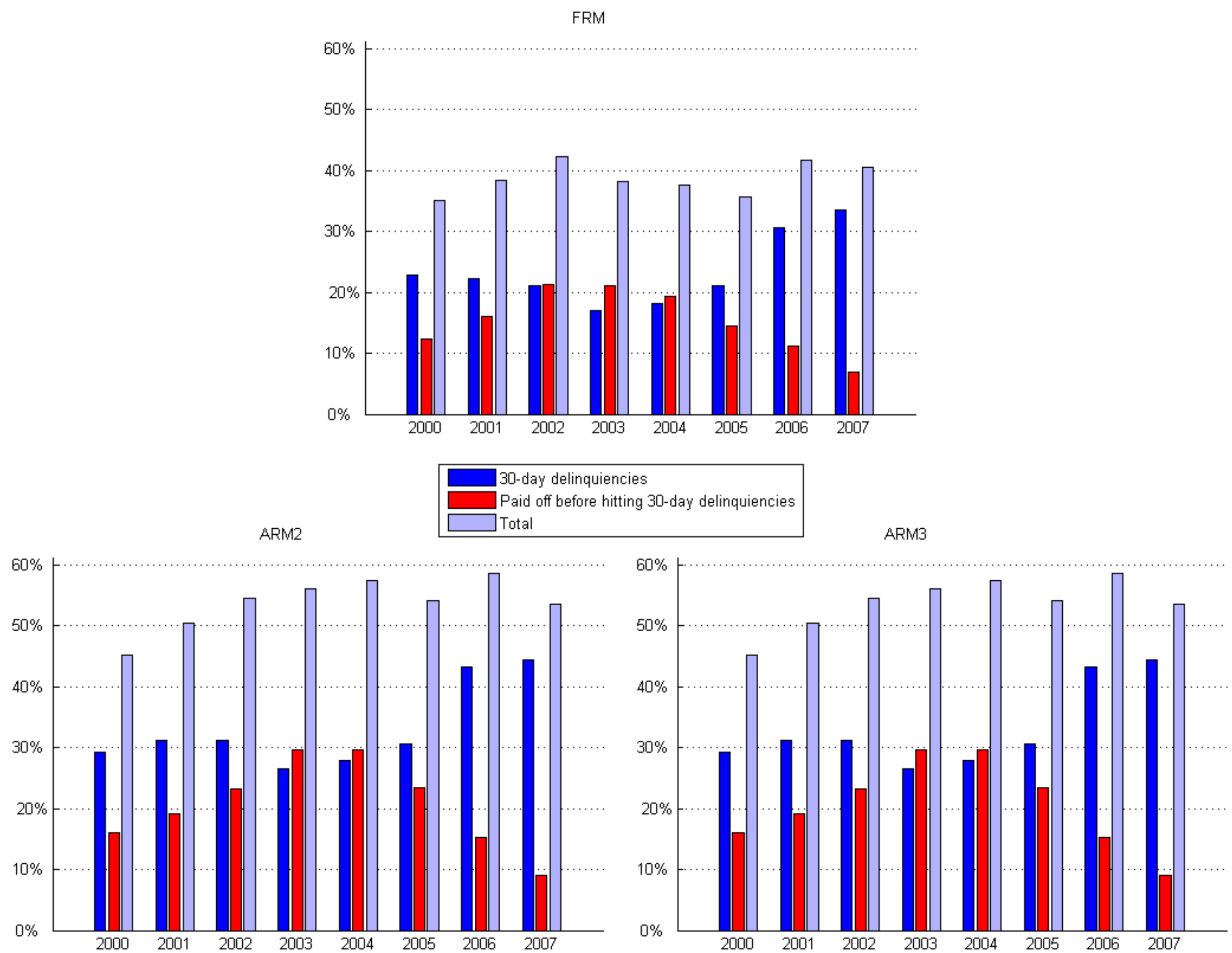


Figure A.2: Pre-delinquency Repayment Behavior by Occupancy (up to loan age of 18 months)

The first bar for each vintage shows the proportion of total loans that register a 30-day delinquency within the first 18 months. The second bar for each vintage shows the percentage of loans that are paid off even before registering a 30-day delinquency. The third bar for each vintage shows the total of the first two columns.

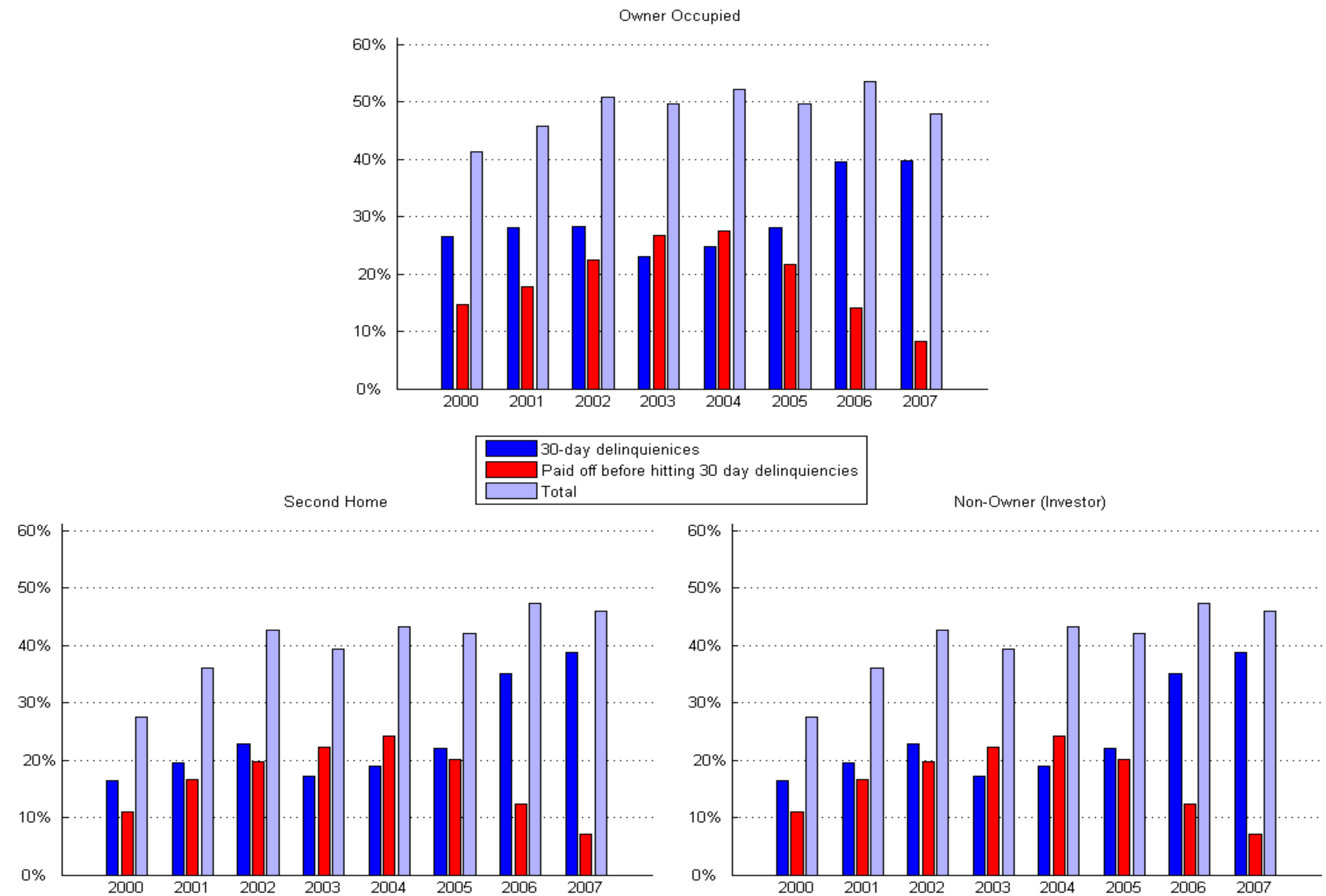


Figure A.3: Pre-delinquency Repayment Behavior by Purpose (up to loan age of 18 months)

The first bar for each vintage shows the proportion of total loans that register a 30-day delinquency within the first 18 months. The second bar for each vintage shows the percentage of loans that are paid off even before registering a 30-day delinquency. The third bar for each vintage shows the total of the first two columns

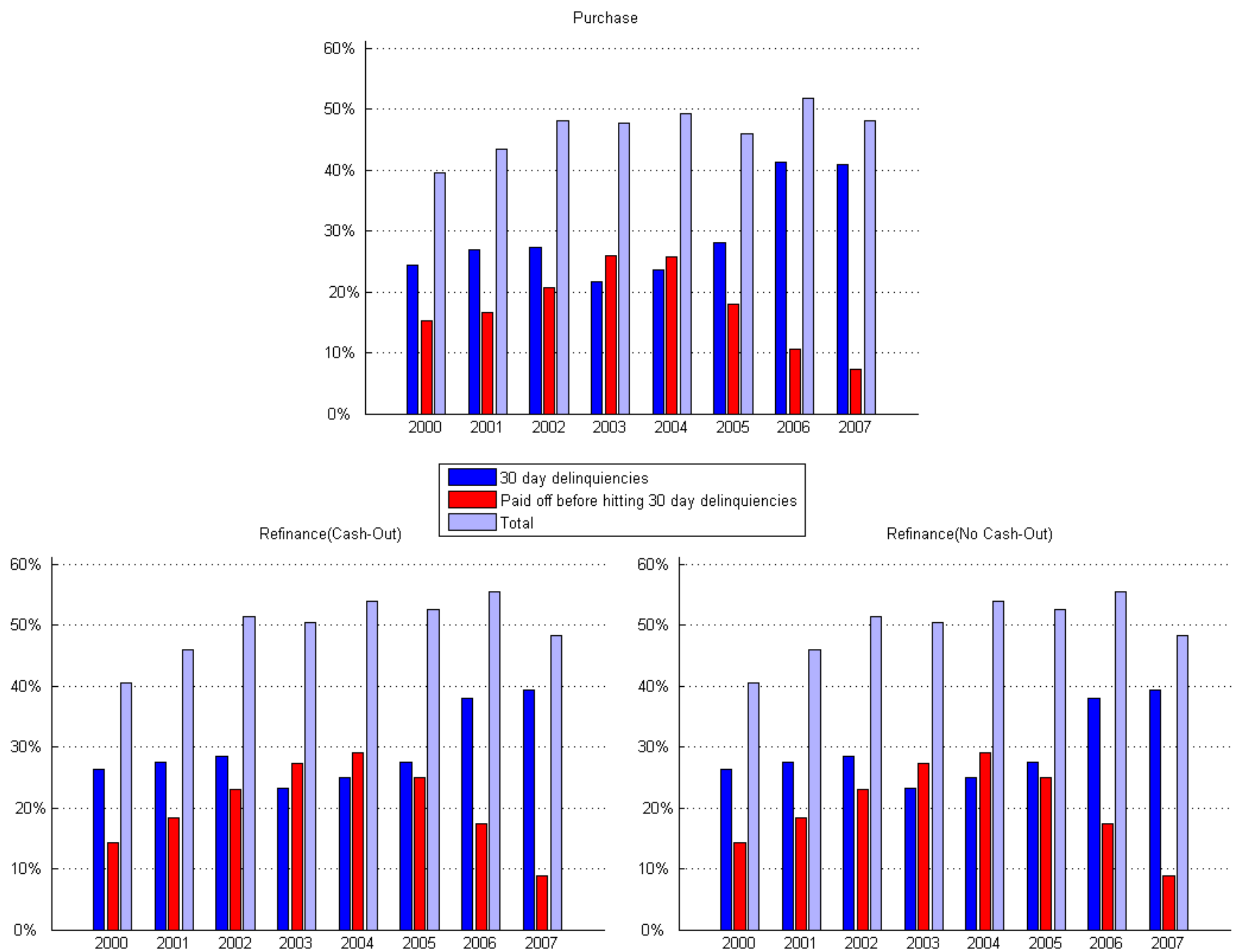


Table C1.5A: Increase in Survival Probabilities for Improvements in FICO Score (groups)

The numbers show a percentage point increase in the Kaplan-Meier survival probabilities (for the first two years after origination) of originations in the higher FICO score group relative to those in the lower FICO score group. The Kaplan Meier probabilities are calculated as discussed in appendix B. The FICO score groups used below are "< 520", "520-559", "560-599" … "680-719" and "≥ 720".

	Vintage							
FICO score	**2000**	**2001**	**2002**	**2003**	**2004**	**2005**	**2006**	**2007**
[< 520] to [520 – 559]	8.54	7.88	6.60	3.96	4.40	4.46	4.67	3.21
[520 – 559] to [560 – 599]	5.36	5.11	4.17	3.56	4.12	3.58	3.80	6.07
[560 – 599] to [600 – 639]	3.80	3.30	2.86	3.26	4.36	3.89	1.59	3.07
[600 – 639] to [640 – 679]	2.67	2.78	3.02	3.12	3.66	4.67	3.11	3.86
[640 – 679] to [680 – 719]	1.62	1.83	2.15	1.79	2.14	3.63	4.72	5.16
[680 – 719] to [≥720]	1.16	1.42	1.51	1.37	1.51	2.84	7.16	7.76
Average all	3.86	3.72	3.38	2.84	3.36	3.85	4.17	4.85
Average first five	4.40	4.18	3.76	3.14	3.74	4.05	3.58	4.27

Table C1.5B: Increase in Survival Probabilities for Improvements in FICO Score (groups)

The numbers show a percentage point increase in the Kaplan-Meier survival probabilities (for the first two years after origination) of originations in the higher FICO score group relative to those in the lower FICO score group. The Kaplan Meier probabilities are calculated as discussed in appendix B. The FICO score groups used below are "< 521", "521-560", "561-600" … "681-720" and "≥721".

	Vintage							
FICO score	**2000**	**2001**	**2002**	**2003**	**2004**	**2005**	**2006**	**2007**
[< 521] to [521 – 560]	8.45	7.90	6.54	3.83	4.38	4.49	4.62	3.19
[521 – 560] to [561 – 600]	5.34	5.10	4.16	3.62	4.11	3.49	3.67	6.05
[561 – 600] to [601 – 640]	3.71	3.13	2.77	3.15	4.28	3.88	1.63	2.88
[601 – 640] to [641 – 680]	2.66	2.88	3.06	3.19	3.71	4.74	3.24	4.10
[641 – 680] to [681 – 720]	1.59	1.82	2.14	1.75	2.08	3.60	4.63	5.12
[681 – 720] to [≥721]	1.17	1.32	1.45	1.38	1.53	2.82	7.21	7.92
Average all	3.82	3.69	3.35	2.82	3.35	3.83	4.17	4.88
Average first five	4.35	4.16	3.74	3.11	3.71	4.04	3.56	4.27

Table C1.5C: Increase in Survival Probabilities for Improvements in FICO Score (groups)

The numbers show a percentage point increase in the Kaplan-Meier survival probabilities (for the first two years after origination) of originations in the higher FICO score group relative to those in the lower FICO score group. The Kaplan Meier probabilities are calculated as discussed in appendix B. The FICO score groups used below are "< 540", "540-579", "580-619" … "700-739" and "≥740".

	Vintage							
FICO score	**2000**	**2001**	**2002**	**2003**	**2004**	**2005**	**2006**	**2007**
[< 540] to [540 – 579]	8.17	7.46	5.75	4.17	4.73	4.95	5.52	6.12
[540 – 579] to [580 – 619]	4.45	4.24	3.57	3.38	3.88	3.04	1.68	3.51
[580 – 619] to [620 – 659]	3.35	2.87	2.91	3.24	4.48	4.33	2.10	3.72
[620 – 659] to [660 – 699]	1.95	2.37	2.54	2.43	2.79	4.59	4.64	5.05
[660 – 699] to [700 – 739]	1.41	1.44	1.96	1.52	1.50	2.56	4.14	3.38
[700 – 739] to [≥740]	0.91	1.10	0.81	0.84	1.30	2.57	7.84	10.13
Average all	3.37	3.25	2.92	2.60	3.11	3.68	4.32	5.32
Average first five	3.87	3.68	3.35	2.95	3.48	3.90	3.62	4.35

Table C1.5D: Increase in Survival Probabilities for Improvements in FICO Score (groups)

The numbers show a percentage point increase in the Kaplan-Meier survival probabilities (for the first two years after origination) of originations in the higher FICO score group relative to those in the lower FICO score group. The Kaplan Meier probabilities are calculated as discussed in appendix B. The FICO score groups used below are "< 541", "541-580", "581-620" … "711-740" and "≥741".

	Vintage							
FICO score	**2000**	**2001**	**2002**	**2003**	**2004**	**2005**	**2006**	**2007**
[< 541] to [541 – 580]	8.13	7.47	5.77	4.15	4.71	4.82	5.45	6.10
[541 – 580] to [581 – 620]	4.42	4.19	3.54	3.40	3.97	3.18	1.70	3.59
[581 – 620] to [621 – 660]	3.33	2.87	2.89	3.21	4.41	4.38	2.20	3.60
[621 – 660] to [661 – 700]	1.94	2.31	2.55	2.42	2.75	4.51	4.60	5.15
[661 – 700] to [711 – 740]	1.29	1.47	1.89	1.45	1.48	2.56	4.14	3.37
[711 – 740] to [≥741]	1.06	1.05	0.81	0.88	1.29	2.50	7.89	10.20
Average all	3.36	3.22	2.91	2.59	3.10	3.66	4.33	5.34
Average first five	3.82	3.66	3.33	2.93	3.46	3.89	3.62	4.36

Table C1.5E: Increase in Survival Probabilities for Improvements in FICO Score (groups)

The numbers show a percentage point increase in the Kaplan-Meier survival probabilities (for the first two years after origination) of originations in the higher FICO score group relative to those in the lower FICO score group. The Kaplan Meier probabilities are calculated as discussed in appendix B. The FICO score groups used below are "< 540", "540-559", "560-579" … "720-739" and "≥ 740".

	Vintage							
FICO score	**2000**	**2001**	**2002**	**2003**	**2004**	**2005**	**2006**	**2007**
[< 540] to [540 – 559]	6.77	6.17	4.50	3.16	3.46	3.71	3.88	3.61
[540 – 559] to [560 – 579]	2.72	2.47	2.37	1.87	2.38	2.24	2.94	4.52
[560 – 579] to [580 – 599]	2.26	2.34	1.83	1.59	1.42	0.51	-0.35	0.56
[580 – 599] to [600 – 619]	1.78	1.40	1.16	1.66	2.45	2.83	1.33	1.69
[600 – 619] to [620 – 639]	1.90	1.56	1.60	1.64	2.52	1.67	0.82	2.33
[620 – 639] to [640 – 659]	1.26	1.46	1.71	1.84	1.80	2.91	1.46	1.32
[640 – 639] to [660 – 679]	1.01	1.25	1.22	1.09	1.40	2.22	3.06	3.33
[660 – 639] to [680 – 699]	1.01	1.08	1.43	1.17	1.33	2.36	2.89	3.16
[680 – 639] to [700 – 719]	0.63	0.47	0.72	0.43	0.40	0.57	1.30	0.50
[700 – 639] to [720 – 739]	0.00	0.45	0.48	0.66	0.53	1.17	1.94	2.00
[720 – 739] to [≥740]	0.91	0.83	0.52	0.46	0.99	1.87	6.66	8.88
Average all	1.84	1.77	1.60	1.42	1.70	2.00	2.36	2.90
Average first ten	1.93	1.87	1.70	1.51	1.77	2.02	1.93	2.30

Table C1.5F: Increase in Survival Probabilities for Improvements in FICO Score (groups)

The numbers show a percentage point increase in the Kaplan-Meier survival probabilities (for the first two years after origination) of originations in the higher FICO score group relative to those in the lower FICO score group. The Kaplan Meier probabilities are calculated as discussed in appendix B. The FICO score groups used below are "< 541", "541-560", "561-580" … "721-740" and "≥ to 741".

	Vintage							
FICO score	**2000**	**2001**	**2002**	**2003**	**2004**	**2005**	**2006**	**2007**
[< 541] to [541 – 560]	6.68	6.16	4.53	3.10	3.47	3.70	3.90	3.64
[541 – 560] to [561 – 580]	2.82	2.50	2.36	1.97	2.30	2.02	2.76	4.42
[561 – 580] to [581 – 600]	2.16	2.31	1.80	1.55	1.55	0.80	-0.27	0.67
[581 – 600] to [611 – 620]	1.78	1.22	1.10	1.56	2.29	2.73	1.40	1.51
[611 – 620] to [621 – 640]	1.94	1.67	1.69	1.73	2.62	1.83	0.84	2.28
[621 – 640] to [641 – 660]	1.16	1.52	1.62	1.83	1.78	2.78	1.53	1.61
[641 – 640] to [661 – 680]	1.07	1.16	1.30	1.10	1.38	2.23	3.01	3.25
[661 – 640] to [681 – 700]	0.58	0.78	0.90	0.84	1.05	1.94	1.99	2.96
[681 – 640] to [711 – 720]	0.94	0.89	1.19	0.67	0.62	0.94	2.17	0.55
[711 – 640] to [721 – 740]	0.00	0.25	0.41	0.71	0.60	1.19	1.91	2.45
[721 – 740] to [≥741]	1.06	0.90	0.56	0.47	0.93	1.79	6.71	8.68
Average all	1.83	1.76	1.59	1.41	1.69	2.00	2.36	2.91
Average first ten	1.91	1.84	1.69	1.50	1.77	2.02	1.93	2.33

Table C1.11A: Parametric Estimates of Increase in Survival Probabilities for Improvements in FICO Score (groups)
The numbers show parametric estimates of percentage point increases in the Kaplan-Meier survival probabilities (for the first two years after origination) of originations in the higher FICO score group relative to those in the lower FICO score group. The estimated probabilities are calculated after controlling for other attributes on the origination as given in appendix B. The FICO score groups used below are "< 520", "520-559", "560-599" … "680-719" and "≥ 720".

A. Controlling for All Attributes on the Origination (excluding Mortgage Terms)

	Vintage							
FICO score	**2000**	**2001**	**2002**	**2003**	**2004**	**2005**	**2006**	**2007**
[< 520] to [520 – 559]	8.07	6.73	6.63	4.59	5.48	5.92	7.79	8.41
[520 – 559] to [560 – 599]	4.83	4.69	4.47	3.93	4.54	5.12	7.23	10.62
[560 – 599] to [600 – 639]	3.08	3.19	2.97	3.16	3.85	4.40	5.94	7.13
[600 – 639] to [640 – 679]	2.96	3.15	2.97	2.84	3.25	4.06	6.17	7.22
[640 – 679] to [680 – 719]	1.87	2.36	2.11	1.85	2.19	2.63	4.33	5.56
[680 – 719] to [≥720]	1.66	1.94	1.61	1.23	1.69	2.43	4.28	4.64
Average all	3.74	3.68	3.46	2.93	3.50	4.09	5.96	7.26
Average first five	4.16	4.03	3.83	3.27	3.86	4.42	6.29	7.79

B. Including the Mortgage Terms: CLTV Ratio and Closing Rate Spread

	Vintage							
FICO score	**2000**	**2001**	**2002**	**2003**	**2004**	**2005**	**2006**	**2007**
[< 520] to [520 – 559]	7.39	5.26	4.93	3.44	4.97	5.25	7.52	5.23
[520 – 559] to [560 – 599]	4.10	3.98	3.93	3.89	4.38	5.09	8.79	10.83
[560 – 599] to [600 – 639]	2.65	2.74	2.67	3.06	3.91	4.64	6.65	7.27
[600 – 639] to [640 – 679]	2.82	3.19	3.22	2.90	3.21	4.00	5.57	6.46
[640 – 679] to [680 – 719]	1.92	2.70	2.56	2.12	2.28	2.51	3.53	4.70
[680 – 719] to [≥720]	1.87	2.40	2.03	1.44	1.70	2.29	3.42	4.01
Average all	3.46	3.38	3.22	2.81	3.41	3.96	5.91	6.42
Average first five	3.77	3.57	3.46	3.08	3.75	4.30	6.41	6.90

Table C1.11B: Parametric Estimates of Increase in Survival Probabilities for Improvements in FICO Score (groups)
The numbers show parametric estimates of percentage point increases in the Kaplan-Meier survival probabilities (for the first two years after origination) of originations in the higher FICO score group relative to those in the lower FICO score group. The estimated probabilities are calculated after controlling for other attributes on the origination as given in appendix B. The FICO score groups used below are "< 521", "521-560", "561-600" … "681-720" and "≥ 721".

A. Controlling for All Attributes on the Origination (excluding Mortgage Terms)

	Vintage							
FICO score	**2000**	**2001**	**2002**	**2003**	**2004**	**2005**	**2006**	**2007**
[< 521] to [521 – 560]	7.96	6.86	6.63	4.54	5.48	5.97	7.87	8.20
[521 – 560] to [561 – 600]	4.82	4.55	4.40	3.91	4.50	5.07	7.13	10.79
[561 – 600] to [601 – 640]	3.07	3.23	2.96	3.14	3.85	4.41	5.99	7.08
[601 – 640] to [641 – 680]	2.85	3.10	2.96	2.83	3.21	4.02	6.11	7.22
[641 – 680] to [681 – 720]	1.94	2.33	2.09	1.80	2.16	2.59	4.26	5.46
[681 – 720] to [≥721]	1.61	1.88	1.56	1.23	1.67	2.42	4.26	4.73
Average all	3.71	3.66	3.43	2.91	3.48	4.08	5.94	7.25
Average first five	4.13	4.01	3.81	3.24	3.84	4.41	6.27	7.75

B. Including Mortgage Terms: CLTV Ratio and Closing Rate Spread

	Vintage							
FICO score	**2000**	**2001**	**2002**	**2003**	**2004**	**2005**	**2006**	**2007**
[< 521] to [521 – 560]	7.27	5.42	4.98	3.41	4.91	5.27	7.56	4.93
[521 – 560] to [561 – 600]	4.11	3.83	3.85	3.87	4.37	5.07	8.73	11.16
[561 – 600] to [601 – 640]	2.65	2.79	2.67	3.02	3.92	4.65	6.64	7.14
[601 – 640] to [641 – 680]	2.69	3.14	3.20	2.90	3.18	3.95	5.49	6.42
[641 – 680] to [681 – 720]	2.02	2.65	2.54	2.06	2.26	2.48	3.48	4.60
[681 – 720] to [≥721]	1.81	2.31	1.95	1.44	1.69	2.27	3.40	4.16
Average all	3.43	3.36	3.20	2.78	3.39	3.95	5.88	6.40
Average first five	3.75	3.57	3.45	3.05	3.73	4.28	6.38	6.85

Table C1.11C: Parametric Estimates of Increase in Survival Probabilities for Improvements in FICO Score (groups)
The numbers show parametric estimates of percentage point increases in the Kaplan-Meier survival probabilities (for the first two years after origination) of originations in the higher FICO score group relative to those in the lower FICO score group. The estimated probabilities are calculated after controlling for other attributes on the origination as given in Appendix B. The FICO score groups used below are "< 540", "540-579", "580-619" … "700-739" and "≥740".

A. Controlling for All Attributes on the Origination (excluding Mortgage Terms)

	Vintage							
FICO score	**2000**	**2001**	**2002**	**2003**	**2004**	**2005**	**2006**	**2007**
[< 540] to [540 – 579]	7.56	6.63	6.27	4.99	5.83	6.75	9.36	12.57
[540 – 579] to [580 – 619]	3.63	3.59	3.51	3.38	3.92	4.33	5.59	8.09
[580 – 619] to [620 – 659]	3.16	3.09	2.83	3.07	3.85	4.45	6.79	7.62
[620 – 659] to [660 – 699]	2.44	2.81	2.65	2.44	2.73	3.50	5.61	7.03
[660 – 699] to [700 – 739]	1.74	1.83	1.77	1.38	1.68	2.08	3.32	3.45
[700 – 739] to [≥740]	0.87	1.41	1.02	0.85	1.47	2.25	4.32	5.24
Average all	3.23	3.23	3.01	2.69	3.25	3.89	5.83	7.33
Average first five	3.70	3.59	3.41	3.05	3.60	4.22	6.13	7.75

B. Including Mortgage Terms: CLTV Ratio and Closing Rate Spread

	Vintage							
FICO score	**2000**	**2001**	**2002**	**2003**	**2004**	**2005**	**2006**	**2007**
[< 540] to [540 – 579]	6.58	5.24	4.76	4.14	5.15	6.08	9.45	9.36
[540 – 579] to [580 – 619]	3.12	2.96	3.04	3.29	4.04	4.49	7.57	9.56
[580 – 619] to [620 – 659]	2.79	2.79	2.70	2.90	3.80	4.61	6.67	6.99
[620 – 659] to [660 – 699]	2.43	2.99	2.91	2.56	2.73	3.32	4.62	5.76
[660 – 699] to [700 – 739]	1.91	2.10	2.21	1.60	1.75	1.99	2.71	2.60
[700 – 739] to [≥740]	0.86	1.73	1.18	0.96	1.46	2.08	3.44	4.99
Average all	2.95	2.97	2.80	2.58	3.16	3.76	5.75	6.54
Average first five	3.37	3.22	3.12	2.90	3.50	4.10	6.21	6.86

Table C1.11D: Parametric Estimates of Increase in Survival Probabilities for Improvements in FICO Score (groups)
The numbers show parametric estimates of percentage point increases in the Kaplan-Meier survival probabilities (for the first two years after origination) of originations in the higher FICO score group relative to those in the lower FICO score group. The estimated probabilities are calculated after controlling for other attributes on the origination as given in appendix B. The FICO score groups used below are "< 541", "541-580", "581-620" … "711-740" and "≥741".

A. Controlling for All Attributes on the Origination (excluding Mortgage Terms)

	Vintage							
FICO score	**2000**	**2001**	**2002**	**2003**	**2004**	**2005**	**2006**	**2007**
[< 541] to [541 – 580]	7.56	6.63	6.25	4.99	5.82	6.68	9.39	12.49
[541 – 580] to [581 – 620]	3.55	3.52	3.50	3.37	3.95	4.40	5.65	8.16
[581 – 620] to [621 – 660]	3.18	3.13	2.83	3.05	3.80	4.43	6.77	7.62
[621 – 660] to [661 – 700]	2.39	2.75	2.64	2.43	2.72	3.46	5.50	7.01
[661 – 700] to [711 – 740]	1.71	1.85	1.74	1.36	1.64	2.06	3.33	3.39
[711 – 740] to [≥741]	0.95	1.33	1.02	0.85	1.46	2.24	4.29	5.30
Average all	3.22	3.20	2.99	2.67	3.23	3.88	5.82	7.33
Average first five	3.68	3.58	3.39	3.04	3.59	4.21	6.13	7.73

B. Including Mortgage Terms: CLTV Ratio and Closing Rate Spread

	Vintage							
FICO score	**2000**	**2001**	**2002**	**2003**	**2004**	**2005**	**2006**	**2007**
[< 541] to [541 – 580]	6.56	5.23	4.74	4.17	5.16	6.03	9.60	9.47
[541 – 580] to [581 – 620]	3.05	2.91	3.03	3.27	4.05	4.57	7.53	9.50
[581 – 620] to [621 – 660]	2.82	2.83	2.71	2.88	3.75	4.56	6.59	6.95
[621 – 660] to [661 – 700]	2.39	2.92	2.90	2.55	2.72	3.28	4.53	5.74
[661 – 700] to [711 – 740]	1.88	2.15	2.17	1.57	1.72	1.96	2.72	2.50
[711 – 740] to [≥741]	0.98	1.60	1.17	0.96	1.45	2.08	3.43	5.17
Average all	2.95	2.94	2.79	2.57	3.14	3.75	5.73	6.55
Average first five	3.34	3.21	3.11	2.89	3.48	4.08	6.19	6.83

Table C1.11E: Parametric Estimates of Increase in Survival Probabilities for Improvements in FICO Score (groups): The numbers show parametric estimates of percentage point increases in the Kaplan-Meier survival probabilities (for the first two years after origination) of originations in the higher FICO score group relative to those in the lower FICO score group. The estimated probabilities are calculated after controlling for other attributes on the origination as given in appendix B. The FICO score groups used below are "< 540", "540-559", "560-579" … "720-739" and "≥740".

A. Controlling for All Attributes on the Origination (excluding Mortgage Terms)

	Vintage							
FICO score	**2000**	**2001**	**2002**	**2003**	**2004**	**2005**	**2006**	**2007**
[< 540] to [540 – 559]	6.29	5.45	4.93	3.82	4.46	5.16	6.73	9.12
[540 – 559] to [560 – 579]	2.43	2.24	2.49	2.13	2.52	2.81	4.64	6.20
[560 – 579] to [580 – 599]	1.91	1.82	1.55	1.49	1.56	1.56	1.56	3.18
[580 – 599] to [600 – 619]	1.14	1.37	1.51	1.67	2.20	2.82	3.63	3.89
[600 – 619] to [620 – 639]	2.04	1.79	1.37	1.62	2.07	2.04	3.68	4.22
[620 – 639] to [640 – 659]	1.25	1.49	1.71	1.55	1.70	2.45	3.29	3.62
[640 – 639] to [660 – 679]	1.56	1.59	1.35	1.24	1.37	1.67	3.13	4.14
[660 – 639] to [680 – 699]	0.43	1.00	0.91	0.91	1.15	1.37	2.02	2.96
[680 – 639] to [700 – 719]	1.21	0.90	0.97	0.62	0.76	0.93	1.54	0.88
[700 – 639] to [720 – 739]	0.67	0.76	0.64	0.57	0.59	0.89	1.42	1.80
[720 – 739] to [≥740]	0.46	0.96	0.63	0.51	1.11	1.70	3.42	4.09
Average all	1.76	1.76	1.64	1.47	1.77	2.13	3.19	4.01
Average first ten	1.89	1.84	1.74	1.56	1.84	2.17	3.16	4.00

B. Including the Mortgage Terms: CLTV Ratio and Closing Rate Spread

	Vintage							
FICO score	**2000**	**2001**	**2002**	**2003**	**2004**	**2005**	**2006**	**2007**
[< 540] to [540 – 559]	5.57	4.30	3.66	3.03	3.94	4.59	6.73	6.49
[540 – 559] to [560 – 579]	1.99	1.87	2.19	2.18	2.43	2.85	5.15	5.64
[560 – 579] to [580 – 599]	1.75	1.54	1.43	1.50	1.78	1.80	3.45	5.47
[580 – 599] to [600 – 619]	0.87	1.08	1.19	1.55	2.18	2.84	3.69	3.10
[600 – 619] to [620 – 639]	1.88	1.65	1.36	1.50	2.05	2.28	3.74	4.23
[620 – 639] to [640 – 659]	1.13	1.46	1.77	1.54	1.65	2.28	2.92	3.15
[640 – 639] to [660 – 679]	1.67	1.72	1.49	1.26	1.32	1.58	2.46	3.23
[660 – 639] to [680 – 699]	0.32	1.14	1.05	1.04	1.23	1.28	1.66	2.67
[680 – 639] to [700 – 719]	1.42	0.98	1.25	0.71	0.76	0.89	1.22	0.33
[700 – 639] to [720 – 739]	0.77	1.00	0.77	0.65	0.61	0.85	1.16	1.51
[720 – 739] to [≥740]	0.39	1.12	0.69	0.56	1.06	1.51	2.63	3.93
Average all	1.61	1.62	1.53	1.41	1.73	2.07	3.17	3.61
Average first ten	1.73	1.67	1.62	1.50	1.80	2.12	3.22	3.58

Table C1.11F: Parametric Estimates of Increase in Survival Probabilities for Improvements in FICO Score (groups): The numbers show parametric estimates of percentage point increases in the Kaplan-Meier survival probabilities (for the first two years after origination) of originations in the higher FICO score group relative to those in the lower FICO score group. The estimated probabilities are calculated after controlling for other attributes on the origination as given in appendix B. The FICO score groups used below are "< 541", "541-560", "561-580" … "721-740" and "≥ to 741".

A. Controlling for All Attributes on the Origination (excluding Mortgage Terms)

	Vintage							
FICO score	**2000**	**2001**	**2002**	**2003**	**2004**	**2005**	**2006**	**2007**
[< 541] to [541 – 560]	6.24	5.47	4.93	3.83	4.50	5.14	6.89	8.92
[541 – 560] to [561 – 580]	2.52	2.20	2.47	2.12	2.44	2.72	4.40	6.41
[561 – 580] to [581 – 600]	1.77	1.73	1.53	1.48	1.64	1.73	1.75	3.23
[581 – 600] to [611 – 620]	1.18	1.47	1.55	1.67	2.20	2.75	3.62	3.81
[611 – 620] to [621 – 640]	2.09	1.77	1.34	1.60	2.03	2.08	3.69	4.25
[621 – 640] to [641 – 660]	1.13	1.49	1.72	1.55	1.68	2.39	3.26	3.60
[641 – 640] to [661 – 680]	1.54	1.53	1.33	1.24	1.38	1.67	3.06	4.16
[661 – 640] to [681 – 700]	0.53	0.99	0.91	0.87	1.12	1.34	1.95	2.83
[681 – 640] to [711 – 720]	1.17	0.94	0.95	0.62	0.76	0.93	1.63	0.87
[711 – 640] to [721 – 740]	0.55	0.73	0.60	0.57	0.58	0.88	1.33	1.88
[721 – 740] to [≥741]	0.61	0.90	0.65	0.51	1.10	1.70	3.45	4.10
Average all	1.76	1.75	1.63	1.46	1.77	2.12	3.18	4.01
Average first ten	1.87	1.83	1.73	1.55	1.83	2.16	3.16	4.00

B. Including the Mortgage Terms: CLTV Ratio and Closing Rate Spread

	Vintage							
FICO score	**2000**	**2001**	**2002**	**2003**	**2004**	**2005**	**2006**	**2007**
[< 541] to [541 – 560]	5.49	4.29	3.67	3.07	3.98	4.57	6.93	6.33
[541 – 560] to [561 – 580]	2.13	1.87	2.17	2.15	2.38	2.80	5.03	6.10
[561 – 580] to [581 – 600]	1.58	1.42	1.39	1.48	1.81	1.92	3.48	5.27
[581 – 600] to [611 – 620]	0.92	1.21	1.24	1.54	2.19	2.81	3.69	3.03
[611 – 620] to [621 – 640]	1.94	1.62	1.34	1.48	2.01	2.26	3.68	4.26
[621 – 640] to [641 – 660]	0.97	1.48	1.79	1.55	1.62	2.23	2.88	3.05
[641 – 640] to [661 – 680]	1.67	1.64	1.46	1.27	1.35	1.59	2.40	3.29
[661 – 640] to [681 – 700]	0.43	1.11	1.07	0.98	1.19	1.24	1.61	2.52
[681 – 640] to [711 – 720]	1.38	1.07	1.22	0.71	0.75	0.89	1.30	0.28
[711 – 640] to [721 – 740]	0.61	0.93	0.71	0.66	0.61	0.82	1.06	1.61
[721 – 740] to [≥741]	0.59	1.04	0.72	0.55	1.04	1.53	2.68	4.05
Average all	1.61	1.61	1.53	1.40	1.72	2.06	3.16	3.62
Average first ten	1.71	1.66	1.61	1.49	1.79	2.11	3.21	3.57

Table C2.3: Credit Score (FICO) Regression (including Back-end Debt-to-Income Ratios)

Table reports OLS estimates with borrower FICO score as the left-hand-side variable and other borrower characteristics as regressors. We control for property type (dummies for single-family residence, condo, townhouse, cooperative, etc.), property location (dummies for the state in which the property is located), loan source (dummies for broker, realtor, wholesale, retail, etc.) and number of units in the property. *Home Value n*th *Quartile* is a dummy that equals 1 if the value of the property lies in the *n*-th quartile of all property values in the data and 0 otherwise. The results for the years of origination 1998 and 1999 are not reported here but are available on request.

Variable	**2000**	**2001**	**2002**	**2003**	**2004**	**2005**	**2006**	**2007**
Intercept	645.32***	670.6***	698.97***	718.64***	682.61***	702.26***	703.79***	699.44***
Debt-to-Income Ratio	-0.05***	-0.17***	-0.09***	-0.12***	0.03***	0.01***	0.02***	0.16***
Full Documentation	-14.83***	-18.01***	-21.92***	-19.32***	-17.78***	-18.86***	-19.26***	-17.33***
Owner-Occupied	-26.75***	-24.03***	-27.54***	-32.40***	-33.81***	-32.13***	-31.51***	-32.66***
Second Home	-3.56***	-3.25***	-8.70***	-12.84***	-14.5***	-7.61***	-8.31***	-15.10***
Refinance (Cash-Out)	-16.94***	-16.67***	-28.03***	-34.54***	-37.1***	-34.4***	-33.22***	-31.37***
Refinance (No Cash-Out)	-18.95***	-17.50***	-20.07***	-22.08***	-22.33***	-19.6***	-18.63***	-23.28***
Home Value First Quartile	-7.39***	-13.74***	-11.42***	-13.81***	-13.11***	-14.07***	-13.92***	-12.03***
Home Value Second Quartile	-5.41***	-9.31***	-7.42***	-9.03***	-8.20***	-8.22***	-8.92***	-8.28***
Home Value Third Quartile	-3.62***	-5.45***	-5.77***	-7.35***	-6.68***	-6.29***	-6.69***	-5.35***
Adjusted R^2	0.0768	0.0898	0.1342	0.1537	0.1685	0.1698	0.1766	0.1519

The symbols ***, **, and * denote statistical significance at 1percent, 5 percent, and 10 percent levels respectively.

Table C2.8: Determinants of Mortgage Terms (including Back-end Debt-to-Income Ratios)
We control for property type (dummies for single-family residence, condo, townhouse, cooperative, etc), property location (dummies for the state in which the property is located) and loan source (dummies for broker, realtor, wholesale, retail etc.). *Home Value n*th *Quartile* is a dummy that equals 1 if the value of the property lies in the *n*-the quartile of all property values in the data and 0 otherwise. The results for the years of origination 1998 and 1999 are not reported here but are available on request.

A. Dependent Variable: CLTV Ratio

Variable	**2000**	**2001**	**2002**	**2003**	**2004**	**2005**	**2006**	**2007**
FICO	2.02***	2.43***	3.39***	4.2***	5.09***	5.37***	5.8***	6.57***
Debt-To-Income Ratio	3.67***	5.04***	5.81***	9.53***	7.7***	4.74***	6.43***	5.64***
Full-Documentation	5.45***	4.51***	3.31***	2.79***	1.85***	1.54***	1.25***	1.67***
Owner-Occupied	4.32***	4.64***	4.78***	5.75***	5.46***	5.18***	5.51***	6.13***
Second Home	-0.26	-1.3***	0.29**	-0.51***	-0.49***	0.25***	0.1	0.9***
Refinance (Cash-Out)	-7.68***	-7.96***	-6.79***	-9.64***	-10.49***	-11.67***	-13.25***	-13.09***
Refinance (No Cash-Out)	-5.96***	-5.9***	-4.39***	-7.56***	-8.62***	-8.59***	-9.25***	-10.17***
Home Value First Quartile	-0.03	2.37***	3.33***	4.4***	3.99***	3.4***	2.8***	3.78***
Home Value Second Quartile	0.71***	2.42***	3.05***	3.8***	3.37***	2.78***	2.17***	2.69***
Home Value Third Quartile	0.65***	2.18***	2.71***	2.81***	2.24***	1.36***	1.13***	1.78***
Adjusted R^2	0.1534	0.1583	0.1681	0.2635	0.3101	0.3235	0.3558	0.3355

B. Dependent Variable: Closing Rate Spread

Variable	**2000**	**2001**	**2002**	**2003**	**2004**	**2005**	**2006**	**2007**
FICO	-0.89***	-1.21***	-1.15***	-1.14***	-1.05***	-1.04***	-1.11***	-1.12***
Debt-To-Income Ratio	-0.16***	0.28***	0.25***	0.16***	0.23***	0.11***	0.12***	0.45***
Full-Documentation	-0.29***	-0.31***	-0.29***	-0.36***	-0.43***	-0.51***	-0.68***	-0.69***
Owner Occupied	-0.4***	-0.44***	-0.41***	-0.43***	-0.51***	-0.67***	-0.73***	-0.75***
Second Home	0.11***	0	-0.2***	-0.21***	-0.27***	-0.29***	-0.41***	-0.33***
Refinance (Cash-Out)	-0.16***	-0.32***	-0.27***	-0.26***	-0.18***	-0.11***	-0.18***	-0.46***
Refinance (No Cash-Out)	-0.12***	-0.27***	-0.25***	-0.21***	-0.21***	-0.17***	-0.16***	-0.48***
Home Value First Quartile	0.97***	1.08***	0.83***	0.84***	0.8***	0.75***	0.73***	0.84***
Home Value Second Quartile	0.46***	0.56***	0.44***	0.46***	0.4***	0.35***	0.36***	0.39***
Home Value Third Quartile	0.23***	0.32***	0.25***	0.23***	0.18***	0.15***	0.16***	0.17***
Adjusted R^2	0.2804	0.4131	0.4203	0.4309	0.4063	0.4258	0.3744	0.4030

The symbols ***, **, and * denote statistical significance at 1 percent, 5 percent, and 10 percent levels respectively.

Table C2.9: Estimated Cox Proportional Hazard Rate Regression (including Debt-to-Income Ratio)

This table reports the estimated hazard ratios for the Cox proportional hazard rate regressions conducted for all loans originated in a given calendar year. We control for property type (dummies for single-family residence, condo, townhouse, cooperative, etc.), property location (dummies for the state in which the property is located) and loan source (dummies for broker, realtor, wholesale, retail etc.). *Home Value n*th *Quartile* is a dummy that equals 1 if the value of the property lies in the *n*-th quartile of all property values in the data and 0 otherwise. The results for the years of origination 1998 and 1999 are not reported here but are available on request.

Variable	**2000**	**2001**	**2002**	**2003**	**2004**	**2005**	**2006**	**2007**
FICO	0.9920***	0.9919***	0.9914***	0.9905***	0.9913***	0.9934***	0.9948***	0.9949***
Debt-To-Income Ratio	1.2134***	1.4736***	2.2747***	2.4973***	2.0873***	1.4623***	1.2711***	0.9654***
Full-Documentation	0.8633***	0.8430***	0.8130***	0.7430***	0.7556***	0.7145***	0.6713***	0.6844***
Owner-Occupied	0.8128***	0.8083***	0.8186***	0.7971***	0.7515***	0.7827***	0.7746***	0.7536***
Second Home	0.5907***	0.5635***	0.6078***	0.5984***	0.6124***	0.7048***	0.6880***	0.6859***
Refinance (Cash-Out)	0.7747***	0.6937***	0.6695***	0.5660***	0.5416***	0.5397***	0.5944***	0.6002***
Refinance (No Cash-Out)	0.9548***	0.8264***	0.7762***	0.6098***	0.5632***	0.5828***	0.6425***	0.6239***
Home Value First Quartile	0.9005***	0.935***	0.9482***	1.0054	0.8589***	0.6872***	0.6336***	0.6356***
Home Value Second Quartile	0.9345***	0.9127***	0.9096***	0.9519***	0.8300***	0.6959***	0.6763***	0.6933***
Home Value Third Quartile	0.9320***	0.9195***	0.8896***	0.9231***	0.8589***	0.8476***	0.8359***	0.8476***
LR test H0: β = 0	28703	37700	62477	114039	172908	217112	231396	33118
(p-value)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)

The symbols ***, **, and * denote statistical significance at 1 percent, 5 percent, and 10 percent levels respectively.

Figure C3.1: Actual and Fitted Survivor Functions

The plots include the Kaplan-Meier estimated survivor function for the particular vintage as defined in Section 6. The other two plots are the estimated survivor functions corresponding to the two estimated models reported in Table 10, Panel A (including all borrower characteristics) and Panel B (including borrower characteristics plus CLTV ratio and *Closing Rate Spread*).

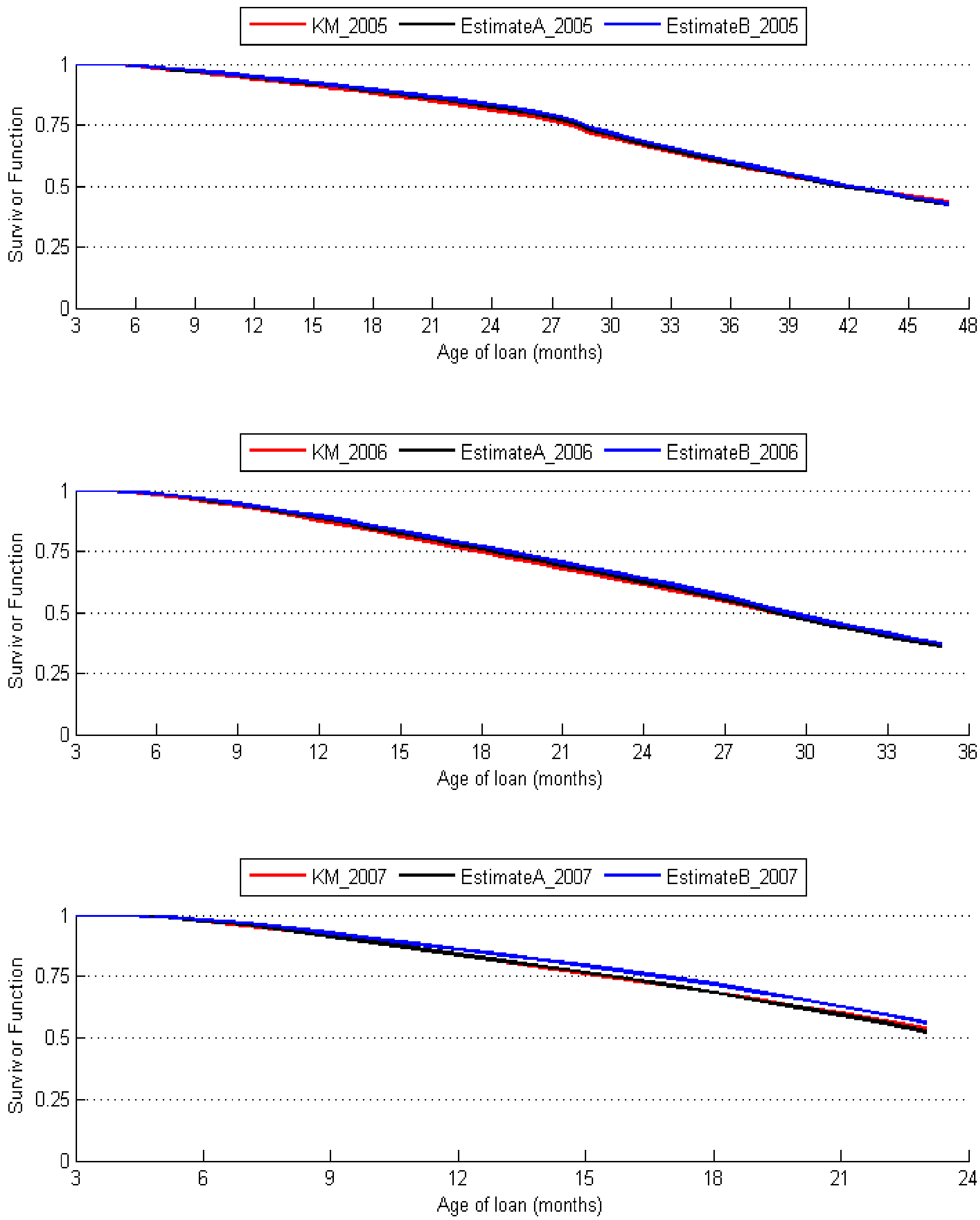


Figure C4.2A: Counterfactual Analysis for 2005 Vintage

The figures show the estimated proportional hazard survivorship function for representative borrowers from different vintages. The plots correspond to the counterfactual exercises using survivor function estimates, based on 2005 data, for loans originated in 2001 and 2002. The upper panel shows results for the counterfactual exercise with borrower characteristics only as regressors, whereas mortgage terms such as CLTV ratio and the *closing rate spread* are added as regressors for the counterfactual results on display in the lower panel.

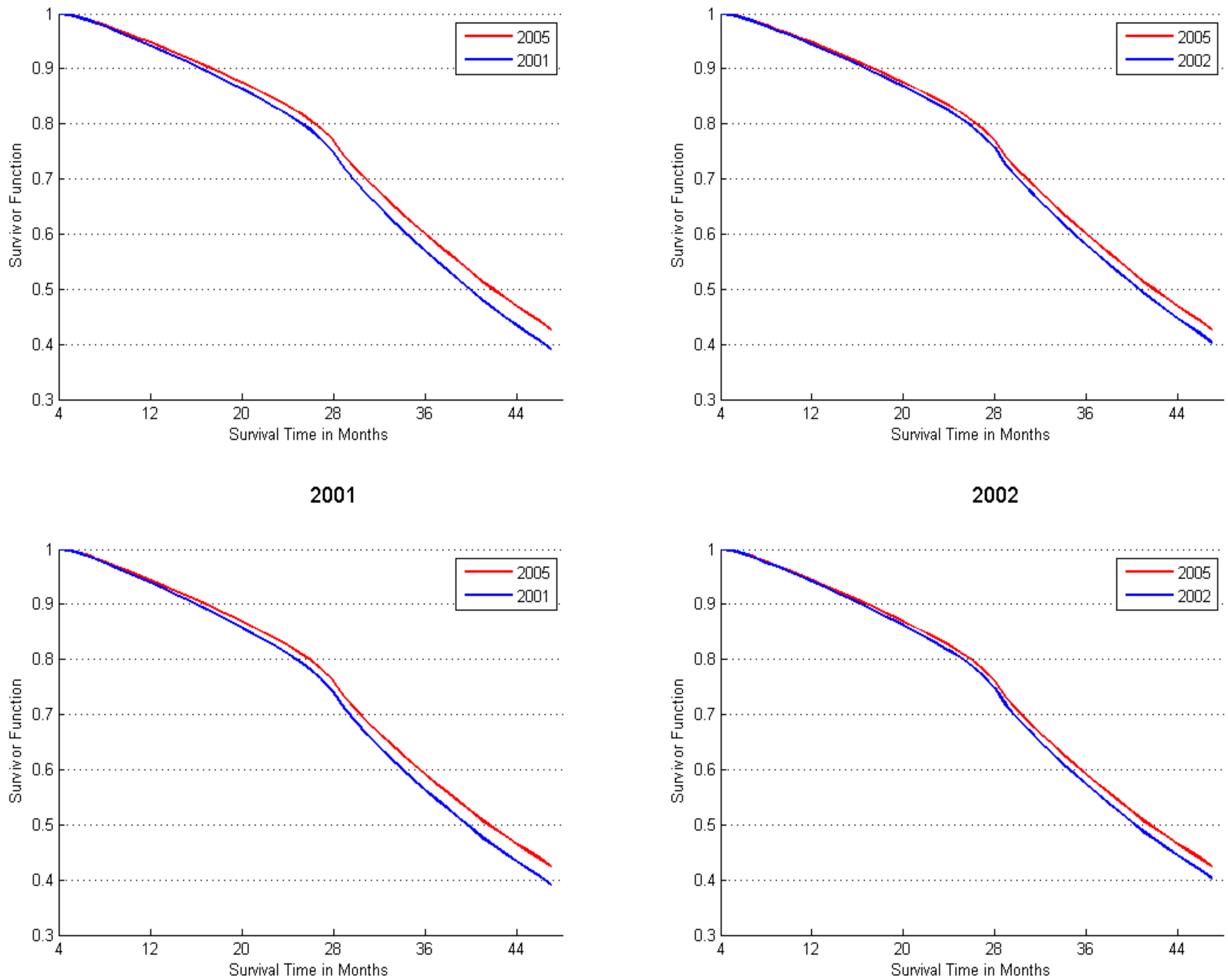


Figure C4.2B: Counterfactual Analysis for 2006 Vintage

The figures show the estimated proportional hazard survivorship function for representative borrowers from different vintages. The plots correspond to the counterfactual exercises using survivor function estimates, based on 2006 data, for loans originated in 2001 and 2002. The upper panel shows results for the counterfactual exercise with borrower characteristics only as regressors, whereas mortgage terms such as CLTV ratio and the *closing rate spread* are added as regressors for the counterfactual results on display in the lower panel.

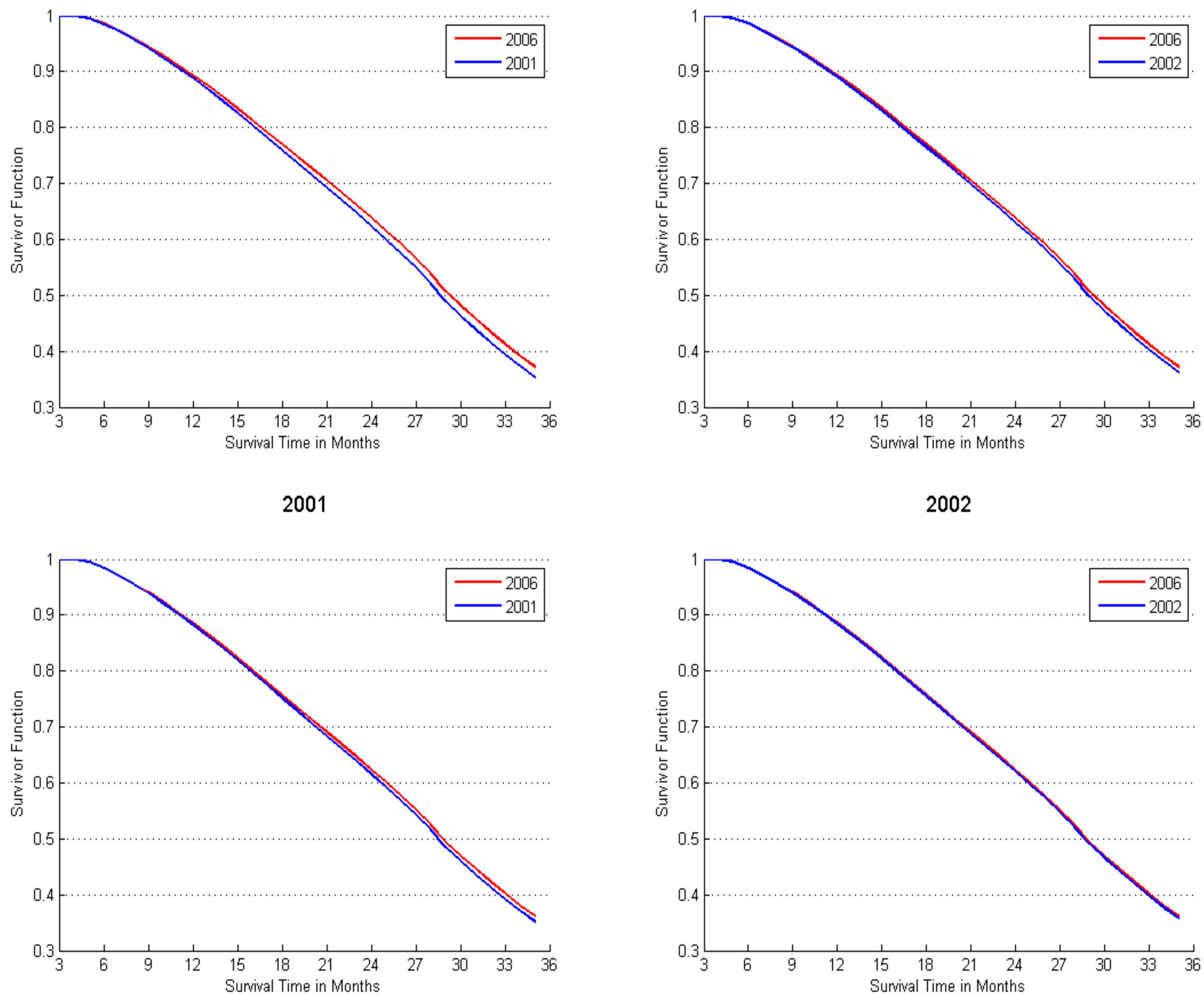


Figure C4.2C Counterfactual Analysis for 2007 Vintage

The figures show the estimated proportional hazard survivorship function for representative borrowers from different vintages. The plots correspond to the counterfactual exercises using survivor function estimates, based on 2007 data, for loans originated in 2001 and 2002. The upper panel shows results for the counterfactual exercise with borrower characteristics only as regressors, whereas mortgage terms such as CLTV ratio and the *closing rate spread* are added as regressors for the counterfactual results on display in the lower panel.