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

Geetesh Bhardwaj† Rajdeep Sengupta‡§

March 31, 2009

Abstract

The dominant explanation for the meltdown in the US subprime mortgage market is that lending standards dramatically weakened after 2004. Using loan-level data, we examine underwriting standards on the subprime mortgage originations from 1998 to 2007. Contrary to popular belief, we find no 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 post-2004 would have resulted in a significantly lower ex post default if they were to be given a loan 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 post-2004 cannot be the explanation for collapse of subprime mortgage market.

JEL Codes: G21, D82, D86.

Keywords: mortgages, subprime, underwriting, crisis.

*Thanks to Mara Faccio, Kris Gerardi, Gary Gorton, Jim Hamilton, Bruce Mizrach, Sangsoo Park, Geert Rouwenhorst, Dan Thornton, Dave Wheelock and Paul Wilson for their comments and suggestions on earlier drafts of this paper.

†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

Existing wisdom on financial crises argues that the upward phase of the credit cycle is often associated with a weakening of lending standards. The hypothesis that "most bad loans are made in the good times" has been viewed, by policymakers and academics alike, as one of the principal features of credit crises.[1] While this is arguably true of most historical episodes of credit crises, the same reason has been put forward for having caused more recent events. Academic research and policy initiatives on the current crisis in subprime mortgages in the U.S. have argued that there was a significant decline in the underwriting standards adopted by subprime lenders. For example, the President's Working Group on Financial Markets (March, 2008) has concluded:

> 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.[2]

Much of the same sentiment is echoed in the popular press.[3] Despite the analysis of these events from newspapers and business journals, there has been little economic analysis of the proposition of examining underwriting standards in 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? To examine these questions, study loan-level data on over nine million subprime mortgages from the LoanPerformance (henceforth,

[1] There is a significant volume of theoretical as well as empirical studies supporting this hypothesis (see Gorton and He, 2008 and references therein). Using a large panel data of over two million loan contracts, Asea and Bloomberg (1998) find evidence that banks become tight during recessions and lax during upturns. Morgan and Lown (2001) find similar evidence using the Federal Reserve's Survey of Senior Loan Officers.

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

[3] 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. 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.

LP) database over the period 1998-2007. This is the largest available repository on 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 has 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 on the latter is largely determined by the former. We address the endogeneity problem that confronts the use of mortgage characteristics like 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 adopting techniques in Chiappori and Salanie (2000). Following this, we study the determinants of mortgage characteristics (like LTV and interest rate) and mortgage delinquencies in the subprime market by accounting for both features mentioned above. Finally, we devise a counterfactual technique to answer the question as to whether there was a decline in mortgage underwriting within the subprime market after 2004.

Our results indicate that, despite anecdotal evidence to the contrary, the hard information available on mortgage originations do not reveal a secular decline in underwriting standards for subprime originations, particularly after 2004. 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. The results show that while underwriting may have weakened along some dimensions (e.g. lower documentation), it also strengthened in others (e.g. higher FICO scores). Our test of endogeneity bias presents evidence of a strong correlation, conditional on observable characteristics, between the individual's choice of loan-to-value ratio (coverage), and the ex-post occurrence of the event of delinquency (risk). When we account for this endogeneity problem, our estimation results fail to document evidence that can suggest that underwriting standards within the subprime market deteriorated over this period. On the contrary, the hard data on mortgage originations provide evidence of credible underwriting over this period which attempted to adjust riskier borrower attributes with lower LTV ratios and higher FICO scores. Moreover, there is compelling evidence to suggest that this adjustment strengthened in some categories.

Using counterfactual analysis, we argue that one cannot reject the hypothesis that there was no weakening of underwriting within the subprime mortgage market for originations after 2004. Our results seem to indicate that (average) observable risk characteristics on loans underwritten post-2004 would have registered a significantly lower ex post default in 2001 and 2002 than (average) observable risk characteristics on loans underwritten in their current years (2001 or 2002). Stated differently, if loans underwritten in 2005 (or 2006 or 2007) were originated in 2001 or 2002, then they would have performed significantly better on average than loans underwritten in 2001 or 2002. In spite of the endogeneity problems of including mortgage terms, we show that the counterfactual results are robust to the inclusion of mortgage terms (like 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 default and delinquency in the subprime market. Of course, our analysis does not rule out the hypothesis that underwriting standards in this market were probably poor to begin with. At the very least, unobservable risk characteristics and market conditions (like house price appreciation) had a greater role to play than was earlier believed.

The subprime mortgage crisis has generated a substantial literature, not least because of the ensuing turmoil in financial markets. Mian and Sufi (2008) show that disintermediation-driven loan originations led to increased risk-taking on the part of lenders from 2001 to 2005. They argue that the rapid expansion in the supply of credit is responsible for the house price boom in the early 2000s and the subsequent mortgage defaults in the last couple of years. These findings are supported in other work on subprime mortgages (Gerardi et al., 2008; DellArricia et al., 2008). In a companion paper (Bhardwaj and Sengupta, 2008), we show that the boom in the subprime market was indeed sustained by prepayments during the period of appreciating house prices. However, the question of interest in this paper is whether there was a gradual shift over the years to a riskier consumer base *within* the subprime market. Undoubtedly, if one considers the entire mortgage market, the mere growth of a subprime segment makes a decline in underwriting standards a tautology. Moreover, our results are based on loan-level data available at the time of origination for first lien securitized subprime mortgages. Although, subsequent changes like originating a second lien on the property or a fall in the property price are important determinants of loan performance, they are hardly reflective of underwriting standards at origination.

Several studies have argued that the subprime mortgage market in the US witnessed a sharp decline in underwriting standards (Demyanyk and van Hemert, 2008). These studies have based their arguments on the originate-to-distribute hypothesis, implying that underwriting standards declined as mortgage originators could successfully pass on credit risk through the process of securitization (Keys et al. 2008). But, as Gorton (2008) points out, this assumption appears exceptionally simplistic in the face of detailed evidence on the securitization process. The analysis in this paper improves upon prior research in two ways. First, our study takes into account the multidimensional nature of credit risk arguing that any focus on a single borrower or mortgage characteristic is misleading. Second, 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 has largely been ignored by recent work (see Demyanyk and van Hemert, 2008).

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 concludes.

2 Data and Summary Statistics

We analyze loan-level mortgage data from the Asset Backed Securities (ABS) Database of the LoanPerformance (LP) data repository.[4] It is the largest database on non-prime loans with loan level data on over 17 million mortgages originated in the United States.[5] Although this database contains both subprime and Alt-A pools, we restrict our analysis to subprime loans for the purposes of this study.[6] Loosely speaking, subprime pools include loans to borrowers with incomplete or impaired credit histories while Alt-A pools include loans to borrowers who generally have high credit scores but who are unable or unwilling to document a stable income history or are buying second homes or investment properties (Fabozzi, 2000). 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. Apart from various borrower and mortgage characteristics, it records all activity on the loan since securitization including repayment behavior. In what follows, our analysis will focus on first lien subprime loans in the ABS database. 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 to 2005.[7]

We begin our discussion on summary statistics with a brief description of the trends in univariate data. Since the subprime market evolved fairly rapidly over the years in our sample period (1998-2007), we record changes in underwriting standards by vintage (year of mortgage origination) throughout this paper.

Table 1 summarizes first lien subprime mortgages by product type (fixed-rate mortgages

[4] Our data are current up to December 2008. LoanPerformance securities databases comprise the mortgage market's largest and most comprehensive mortgage securities data repository. See http://www.loanperformance.com/data-power/default.aspx

[5] However, the data set is not without its limitations: First, there is little information on the households that had subprime mortgages. For example, there are no data on household debt, income, employment and 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.

[6] We classify a loan as a subprime loan if it belongs to a subprime pool in the ABS database. The industry classification of subprime and Alt-A is at the pool level rather than on individual loan characteristics. Therefore, while subprime and Alt-A loans each have distinct loan credit and documentation characteristics, it is possible for a subprime pool to include a loan with characteristics more suitable for the Alt-A pool and vice versa.

[7] For details on the coverage of the LP subprime database and how it relates to other available databases on prime and subprime mortgages in the United States, see Meyer and Pence (2008).

(FRMs) and adjustable-rate mortgages (ARMs)) for every year of origination from 1998 to 2007.[8] The majority of ARMs are hybrid-ARMs; that is, lenders employ an initial (fixed) *teaser rate* that resets into a fully indexed rate to attract borrowers to the product. For a hybrid-ARM, this teaser rate is often fixed for longer periods of time such as two or three years. To simplify classification over a very broad range of product types in the market, we define these products as ARM2 and ARM3 respectively.[9] As seen from Table 1, the subprime mortgage market comprises of mainly three product types: FRM, ARM2 (which includes the hybrid 2/28 mortgage product) and ARM3 (which includes the hybrid 3/27). All other product types make up a smaller fraction of the subprime mortgage market and their market share was on the decline for most of the sample period. Clearly, there has been a shift over the years from FRMs to ARMs in the subprime market. The proportion of FRMs declined from more than half of total originations in 1998 to less than a fifth in 2006. At the same time there have been dramatic increases in ARM2 and ARM3 product types.

Contrary to conventional wisdom, teaser rates on hybrid-ARM products were not significantly lower when 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 (1998-2007). Moreover, performance data of subprime mortgage products shows that originations of 2006 and 2007 vintages were the worst performing mortgages. Among them, 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. These results suggest that loan performance on subprime mortgages can hardly be explained by variations in the distribution of product types. Therefore, in what follows, the results presented in the paper are for data pooled over all mortgage products. Results on in-

[8]Fixed rate mortgages (FRMs) have an interest rate that is set (or locked) at the closing of the loan and does not change over the life of the loan. However, rates are subject to change for adjustable rate mortgages (ARMs). ARMs typically *reset* annually and the periodic contractual rate is based on the *index* (an underlying reference rate like the LIBOR or COFI) and the *margin* (spread over the index).

[9]Not all ARM2 and ARM3 mortgages have a thirty-year maturity period. Therefore, while 2/28s and 3/27s make up the majority of loans in these two categories, they do not constitute all such loans.

dividual product type (ARMs and FRMs) are qualitatively similar and available upon request. For the purposes of this study, we focus our attention to the *closing rate spread*, 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*.[10]

Our data on securitized first-lien subprime originations show that there is little change in underwriting in terms of occupancy and purpose of the loan; the proportions of the sample under different categories for both characteristics were fairly stable over the 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 to 16 percent of total originations between 1998 and 2003, but their proportion drops to 6-7 percent of the total, 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 the total originations, while non-owner (i.e., investor) occupied properties account for 7-9 percent of the total.

However, we do observe a trend towards riskier loans in terms of lower documentation and high loan-to value (LTV) ratios. From roughly 18-19 percent of total originations in 1999-2000, the proportion of low-doc loans increased to 35-36 percent of total originations for 2005-2006. However, no-doc loans remain an insignificant 0.4-0.7 percent of the total originations for all years in our sample period. In addition, subprime lenders increasingly began to originate mortgages with high LTV ratios. For example, originations with LTV ratios in the (90, 100] range increased from 3-4 percent of the sample in 1998-1999 to 35-40 percent of the total in 2005-2006. Remarkably, however, 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.

The trends in 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

[10] 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 (FHLMC).

and high LTV loans, there is also a trend towards a 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 2 shows the distribution of FICO scores conditional on documentation level on originations of various vintages. The proportion of borrowers in the lowest FICO-score category (< 620) has declined over the years. At the same time, there has been an increase in the proportion of borrowers in the 620-659 and the 660-719 range, especially for originations without full documentation. The distribution of FICO conditional on LTV shows a similar pattern (Table 3). For all years, originations with higher LTVs typically have a higher FICO scores on average. Just like in the case of loan documentation, there has been a shift of population from the lowest FICO group (< 620) to the two intermediate FICO score groups (620-659 and 660-719) across the three LTV ranges.

3 Evidence on Underwriting in the Subprime Mortgage Market

3.1 FICO score and Risk Characteristics

Based on the evidence presented in the summary Tables 2-3, 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 employ regression techniques to determine equilibrium underwriting behavior. Borrower FICO scores are regressed on other borrower attributes and loan characteristics. The regression estimates in Table 4 summarize equilibrium underwriting for subprime mortgages.[11] In addition to borrower characteristics used as regressors in Panel A, Panel B of Table 4 includes the *closing rate spread*, defined above. 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-doc dummy (both panels) indicates that after controlling for other borrower attributes, a borrower with low or no documentation has a significantly higher FICO than a similar borrower providing full documentation on the loan. As one would expect, FICO requirement for loan approval on non-owner (investor) occupied homes is the highest, followed by that on second homes, while approvals for owner occupied originations have the lowest FICO scores. Not surprisingly, mortgages on properties with greater value have progressively higher FICO scores. For loans of all vintages, property values in a lower quartile have on average a lower FICO than those property values in the immediately higher quartile. Evidently, refinances have a lower FICO 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. Although not reported here, including the LTV ratio as one of the regressors yields similar results: originations with higher LTV are seen to have greater FICO scores in equilibrium.

The regression coefficients indicate that underwriting sought to adjust for borrower's riskier attributes with 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 employ a fully interacted dummy variable model of the regression in Panel A of Table 4. The dummy variable takes the value one for all originations after a given calendar year, and zero otherwise. We present the estimates on four specifications in Table 5, 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 post-2004 vintage shows that the improvement in FICO scores for originations between 2005 and 2007 was statistically as well as economically significant.

[11] We control for property type (dummies for single-family residence, condo, townhouse, co-operative, etc), property location (dummies for the state in which the property is located) and loan source (dummies for broker, realtor, wholesale, retail etc).

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 that on full documentation loans (Panel A, Table 4). However, we do not know if the difference in FICO of 19.26 points (as recorded on loans of 2006 vintage) 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 due to 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. Our data allows for tracking repayment behavior on subprime mortgages on a monthly basis. Therefore, we can determine the nature (30-day, 60-day or 90-day) and timing (month) of the delinquency event. 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.[12] Default rates and the probability of surviving a delinquency are calculated by using the Kaplan and Meier (1958) product limit estimator. We begin this non-parametric approach to survival and hazard function estimation by formalizing it in the current context of mortgage repayment behavior.

Following Kaplan and Meier (1958), the delinquency rate $D(t)$ at month t (the age of the mortgage in months) is defined as

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

where T is the age in months for the delinquency event (60 day, 90 day, or foreclosure) of a randomly selected mortgage and $S(t) \equiv P(T > t)$ is the survivor function or the probability of

[12] Although we use 90-day delinquencies throughout the paper,the results for 60-day delinquencies and foreclosures are qualitatively similar and are available on request.

surviving the delinquency event beyond age t. Let $t_{(1)} < t_{(2)} < ... < t_{(k)}$ represent the ordered age in months at the time of delinquency event. For all these months, let n_i be the number of surviving mortgages just prior to month $t_{(i)}$. Surviving mortgages not only exclude the ones that have been delinquent, but also the ones that have been refinanced prior to age $t_{(i)}$. If d_i is the number of mortgages that go delinquent at age $t_{(i)}$, then the Kaplan-Meier estimator of surviving the event of delinquency is defined as

$$\hat{P}(T > t) = \prod_{i=1}^{k}(1 - \frac{d_i}{n_i}) \tag{2}$$

From our results, it is clear that defaults started to rise sharply in 2006 and 2007, primarily for originations between 2004 and 2007. To give an example, about 21 percent of mortgages originated in 1998 were in serious delinquency by the fifth year (end of calendar year 2002) whereas the same proportion of defaults for 2004 originations occurred within three years (end of calendar year 2006). The numbers are even more striking when one considers that around 35 percent of mortgages originated in 2006 had defaulted by the end of 2007. 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.

Table 6 reports the distribution of default (90-day delinquency) probabilities conditional on borrower FICO scores. While there has been a significant increase in defaults over the years within a given FICO-score group, the trend is hardly monotonic. Almost always, mortgages of 2003 vintage show anomalous behavior that breaks away from this trend. However, in one aspect the numbers are just as one would expect: delinquency probabilities of lower FICO-score groups for a given vintage are greater than that for higher FICO-score groups of that vintage. This result is robust for loans of all vintages and across all four FICO-score groups. The consistency of the result over the age of the mortgage and across the different vintages is remarkable. Clearly, these summary results indicate that lenders' assessment of FICO scores being good predictors of loan performance was not incorrect.

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 towards higher FICO scores alongside lower documentation and higher LTVs. Second, we observed that average FICO score is significantly higher for originations whose other attributes (like lower documentation, or higher loan-to-value ratios) are arguably riskier. Third, we present evidence to suggest that this adjustment strengthened over the years in our sample period. This feature of underwriting suggests lenders' emphasis on 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 even faulted originators on this account:

> ... the crucial mistake many lenders made was relying on FICO credit scores to gauge default risk, regardless of the size of the down payment or the type of loan.[13]

However, if the summary evidence presented on ex post (realized) default in Table 6 provides any indication of (ex-ante) credit risk in the subprime market, it appears that the lenders were justified in their assessment. In Sections 5 and 6, we present a more rigorous analysis in support of this conclusion.

4 Mortgage Underwriting, Asymmetric Information and Endogeneity Bias

The importance of information problems in any borrower-lender scenario can hardly be overemphasized, especially when it pertains to a market for borrowers who would not otherwise qualify for more conventional sources of financing. In this section, we emphasize the role of information asymmetries in the loan underwriting process. But before this, it is important to list the limitations of our study in examining underwriting standards for subprime mortgages.

4.1 Limitations of Characterizing Underwriting Standards

To begin with, approving loan applications of borrowers that would previously be considered uncreditworthy can be viewed as a weakening of underwriting standards. Interestingly, the

[13] "The woman who called Wall Street's meltdown" - *Fortune Magazine*, Aug. 4, 2008.

subprime market was primarily conceived to supply 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 US 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 28 percent. 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 the prime market.

Secondly, 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.

Thirdly, 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. Earlier studies have used both borrower attributes and loan characteristics as determinants of ex post default (Demyanyk and van Hemert, 2008). However, it is important to remember that mortgage characteristics are themselves outcome 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 "non-prime" 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*.[14] If different mortgage contracts are actually sold to observationally identical borrowers, then the frequency of default among the subscribers to a contract should increase with the loan-to-value 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

[14] 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).

LTV 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 (like low downpayment), that at a given price, are more valuable to them than to lower-risk individuals. Second, equilibrium pricing of underwriting policies reflects variation in the risk pool across different policies. 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 7 reports actual interest rates on offer for 30-year fixed rate mortgages in the subprime market by Option One Mortgage Corporation in November 2007.[15] This table summarizes the 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 downpayment on the loan. In other words, borrowers of riskier type have 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.

Firstly, conditional on observable risk, borrowers are offered menus of contracts varying in their interest rate and LTV requirements as given in Table 7. 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 7.

Secondly, within the menu of contracts on offer, contracts with a higher LTV typically come with a higher 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 terms, but not both.

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

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) loan-to-value ratio (LTV) 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 7): 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 (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, will determine his downpayment requirement (LTV). Accordingly, we can focus our attention to the determinants of the mortgage contract as follows:

$$Type^{*} = \mathbf{X}\boldsymbol{\delta}+\delta_{Z}Z + v \tag{3}$$

$$Type = \mathrm{FRM}\,[Type^{*} > 0] \tag{4}$$

$$Z = \mathbf{X}\boldsymbol{\gamma} + u \tag{5}$$

where $\mathbf{X}$ is a vector of borrower attributes and Z is either the LTV on the mortgage or the mortgage 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 or mortgage rate.[16]

Determinants of default and delinquency To derive testable predictions about the ex-post occurrence of default, we estimate the semi-parametric 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

$$h(t) = \lim_{\Delta t \to 0} \frac{\Pr(t \leq T < t + \Delta t | T \geq t)}{\Delta t} \tag{6}$$

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

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

$$h(t) = h_0(t)\exp(\mathbf{X}\beta) \tag{7}$$

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).

$$\begin{aligned} Z_i &= X_i\gamma + u_i & (8) \\ h_i(t) &= h_0(t)\exp(X_i\beta) & (9) \end{aligned}$$

The first equation, identical to equation (5), is an ordinary least squares regression with LTV ratio (or interest rate spread) as the dependent variable. The second equation, identical to equation (7), is a Cox proportional hazard rate regression model.[17]

The martingale residuals of the Cox model are calculated as

$$\hat{\eta}_i = \delta_i - \hat{H}_0(t)\exp(X_i\hat{\beta}) \tag{11}$$

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 zero 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

[17] The object of interest in a Cox proportional hazard rate regression model is hazard ratio, that has the interpretation of a multiplicative change in the instantaneous probability of delinquency for a marginal change in a particular risk characteristic. Hazard ratio 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 hazard ratio (HR) for marginal change in risk characteristic x_i as

$$\begin{aligned} \widehat{HR}(t\,|x_i = x_i + \Delta x_i\,) &= \frac{h_0(t)\exp(x_1\widehat{\beta}_1 + x_2\widehat{\beta}_2 + \cdots + (x_i + \Delta x_i)\,\widehat{\beta}_i + \cdots)}{h_0(t)\exp(x_1\widehat{\beta}_1 + x_2\widehat{\beta}_2 + \cdots + x_i\widehat{\beta}_i + \cdots)} \\ &= \exp(\Delta x_i\widehat{\beta}_i). \end{aligned} \tag{10}$$

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

$cov(\varepsilon_i, \eta_i) = 0$ is defined by:

$$W = \frac{(\sum_{i=1}^{n} \widehat{u}_i \hat{\eta}_i)^2}{\sum_{i=1}^{n} \widehat{u}_i^2 \hat{\eta}_i^2} \tag{12}$$

where W is distributed asymptotically as a $\chi^2(1)$.[18]

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 loan-to-value ratio (coverage), and the ex-post occurrence of the event of delinquency (risk). Table 8 shows the conditional correlation between risk and coverage under various specifications. The first specification uses LTV ratio as the dependent variable in (8), while the second specification uses the *closing rate spread*. Both specifications yield similar results: the conditional correlations for all vintages are positive and significant. Chiappori and Salanie (2000) test statistic in (7) confirms the statistical significance of the results. In addition to the Chiappori and Salanie (2000) test statistic, we construct a bootstrap confidence interval for testing the significance of correlation (conditional on observables) between risk and coverage.[19] 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

[18] Chiappori and Salanie (2000) estimate a probit equation for 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.

[19] 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 (8) and (9), recover the OLS residuals on equation (8) and the martingale residuals in (11); 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.

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).[20] 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 like LTV ratio (and interest rate spread) as explanatory variables in determining of loan performance.

In our regression results below, we show that ignoring this endogeneity bias leads to faulty inferences. The inclusion of endogenous variables like LTV 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 like the FICO score and the full documentation dummy, one expects a expects a negative coefficient in the hazard rate regression. Consequently, the positive bias introduced by including endogenous variables like LTV ratio reduces the true impact of such explanatory variables on the probability of default.

In the next section, we report the OLS estimates of LTV and *closing rate spread* on borrower characteristics. These regressions show a high degree of explanatory power of the borrower characteristics for both LTV and the *closing rate spread*. Consequently, borrower characteristics can be used as instruments for LTV ratios and interest rate spreads. This helps in coping with the problem of endogeneity bias for the default regression.

5.2 Determinants of Mortgage Terms

Table 9 reports estimates of equation (8) for first lien subprime originations between 2000 and 2007, with LTV ratio as the dependent variable in Panel A and *closing rate spread* as the

[20] 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.

dependent variable in Panel B.[21] 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. Borrowers in higher valued properties have lower LTVs 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 higher property values have on average a lower loan-to-value ratio. Not surprisingly, originations on lower property values, and consequently higher LTVs, have lower interest rates.

(2) Owner-occupied homes have significantly higher LTVs 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 downpayment on loans of identical size.

(3) Mortgages with full-documentation have significantly higher LTVs and lower rates than low or no documentation loans. But, the size of the LTV coefficients in Panel A decline over the sample period. Evidently, underwriting attempts at tempering low-documentation loans with lower LTVs on average was probably getting weaker 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 LTVs, and consequently higher mortgage rates. But here the trend of adjustment of FICO scores with lower LTVs seems to have gotten stronger over the years.

(5) No cash-out refinances have lower LTVs 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 LTV 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 LTV regression (Table 9) to

[21] The results for the years of origination 1998 and 1999 are not given here, but are available upon request.

those in the FICO regression (Table 4). 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. While refinances have lower FICO scores on average, borrowers refinancing loans also have lower LTVs. 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 LTV 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 to the full extent of the appreciated value.[22] 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 LTVs on average than purchases.

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

5.3 Determinants of Mortgage Default

Table 10 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 type, and property location. Confirming our earlier results in Section 3, these estimates show that a higher FICO score (scaled here by dividing the borrower's credit score by 100) significantly lowers the probability of default. Therefore, for originations in 2002, loans to borrowers with a 100 point higher FICO score reduces the probability of default by 59 percent. Likewise, requiring full documentation on loans originated in 2002 reduces the probability of default by 18 percent.

[22] The lender now faces a less risky borrower who has built up equity in the house. Gorton (2008) argues that 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.

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. Clearly, a priori beliefs about the effect of individual borrower characteristics on credit risk are validated: for example, originations with full documentation have a significantly lower probability of default than low- or no-doc loans.

Viewed independently, the evidence in Table 10 tells us little about underwriting standards. On the other hand, when these regression results are examined in conjunction with previous regression results in Tables 4 and 9, we are able to get a clearer picture of underwriting standards. Earlier, we showed evidence to suggest that the underwriting process attempted to adjust riskier borrower characteristics with higher FICO (Section 3.1) and lower LTVs (Section 5.2). These results also suggest that lenders adjusted higher LTVs 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 LTV and lower documentation by raising FICO scores 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 (like LTV 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 regression in Panel B in Table 10. Including the *closing rate spread* introduces a positive bias on the estimates of explanatory variables like FICO scores, thereby reducing the impact of FICO scores as a determinant of ex post default. This is clearly evident from the higher hazard ratios (lower impact on the probability of default) obtained for originations of all vintages in Panel B of Table 10. A similar effect is observed if we include LTV ratio as a regressor in the hazard estimation.

6 Counterfactual Analysis

So far, we have failed to find the smoking gun. Both from the point of view of mortgage and borrower characteristics and from the standpoint of ex post default, observable underwriting trends do not provide evidence of a secular decline in lending standards. Moreover, there is no discernable 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 the strengthening of this adjustment over the years. However, we are yet to determine if the adjustment was "adequate" in terms of its magnitude. At the heart of this 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 us determine if underwriting 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 originated to a "representative borrower" in 2005 were to be given a loan 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 $S_{v,0}(t)$.

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

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.[23] In this specification, let $\mathbf{X}_1$ and $\mathbf{X}_2$ denote the "representative borrowers" of vintage v_1 and v_2 respectively. If unobservable characteristics and market conditions captured by the baseline survivor function are applied on the different borrower characteristics, we can identify the effect of $\mathbf{X}_2$ on the survivor function in v_1 as follows:

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

Such a counterfactual exercise helps us in testing 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 (7) for a given vintage v. Next, we calculate the estimated survivor function for the representative borrower of vintage v. Finally, we calculate the counterfactual survivor function for the representative borrower of a different vintage, say $\widetilde{v}$. Since 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 of summarizing the distribution of borrower characteristics.[24]

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 likely to survive

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

[24] 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.

an event of default than those of earlier vintages, namely 2001, 2002 or 2003, in the environment of these vintages. The results of counterfactual analysis are summarized in Table 11. Table 11 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% 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 performed significantly better 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 the loan performances between the representative borrowers of 2005 or 2007 vintages and that of the 2003 vintage.

These results are best illustrated in terms of the survival plots in the upper panel of Figure 1. As discussed above, we can reject the null hypothesis in favour of the alternative that the underwriting standards actually improved in the latter vintages when compared to 2001 and 2002 vintages. Originations of 2003 vintage perform significantly better than originations of 2006 but not for 2005 and 2007.

To check the robustness of our results, we conduct a similar counterfactual analysis, this time including LTV 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 1. 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 high LTV mortgages 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 default in subprime mortgages.

7 Conclusion

We begin by pointing out some of the limitations in our study. Firstly, 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, they would hardly provide a basis for determining a decline in underwriting at origination. Secondly, as with any empirical study, there is of course the possibility that there were borrower attributes observed by the lender, but that are not reported in the FALP data. Lack of data often makes it difficult to make a conclusive argument on some important characteristics, like for example, the debt to income ratio. Using HMDA 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 FALP database on individual mortgages creates significant problems. First, there is almost no 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 FALP data. For the data that is 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. Thirdly, some observers could raise the doubts about the veracity of the data. There is some anecdotal evidence that points to poor reporting, false documentation and outright fraud.[25] However, it is difficult to make this case for a significant proportion of a repository of nine million loan observations. Finally it needs to be mentioned that our examination of the underwriting standards is at level of the individual borrower and not at the level of the lending institution. We do not examine the hypothesis if, for example, originations of high-LTV mortgages were disproportionately high for a particular lending institution.

Nevertheless, this paper presents a contrarian perspective on underwriting standards in the subprime market. Our examination of the LP data shows scant evidence of a decline in underwriting standards. Moreover, our counterfactual analysis demonstrates that, at least on average,

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

we can reject the hypothesis of no decline in underwriting standards in favour 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, this begs the obvious questions as to what sustained the phenomenal growth in the subprime market for nearly a decade. And, of course, why did the subprime market collapse?

In a companion paper, Bhardwaj and Sengupta (2008), we attempt to answer these questions in sufficient detail. Following Gorton (2008), we argue that the subprime mortgage contracts were designed as "bridge finance" to give temporary credit relief to borrowers in anticipation of future earnings growth or build-up of borrower equity through rise in house prices, or both. Bhardwaj and Sengupta (2008) find 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. Evidently, refinancing mortgages during a period of rising house prices is easier than doing the same when the opposite is true.

References

Adam, W., Einav, L., Levin, J. (2007) , Liquidity Constraints and Imperfect Information in Subprime Lending. American Economic Review. 99(1), forthcoming.

Asea, P. and Blomberg, B. (1998) Lending Cycles. Journal of Econometrics. 83(1) 89-128.

Bhardwaj, G. and Sengupta, R. (2008). Did Prepayments Sustain the Subprime Mortgage market? Working Paper.

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., and B. Jullien, B. Salanié and F. Salanié, (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 te Subprime Mortgage Market. IMF working paper: 08/106.

Demyanyk, Y. and Van Hemert, O. (2007). Understanding the subprime mortgage Crisis. Working Paper.

Fabozzi, F.J. ed. (2006). The Handbook of Mortgage-Backed Securities. 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.

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 E.J. Vyltacil (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.

Lown, C. and Morgan, D. (2006) The Credit Cycle and the Business Cycle: New Findings Using the

Loan Officer Opinion Survey. Journal of Money, Credit & Banking. 38(6) 1575-1597

Keys, B., Mukherjee, T., Seru, A. and Vig, V. (2008). 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.

Mian, A., and Sufi, A. (2007). The consequences of mortgage credit expansion: evidence from the 2007 mortgage default crisis, unpublished manuscript, University of Chicago Graduate School of Business.

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

Wooldridge, J. (2002) Econometric analysis of cross section and panel data. Cambridge and London: MIT Press.

Table 1: Evolution of the Subprime Market (market share by product type)

Table summarizes first lien subprime mortgages by product type as fixed or adjustable rate mortgages (FRM and ARM) for every year of origination from 1998 to 2007. The numbers give us the market share for a particular product type. ARM2 and ARM3 denote hybrid-ARM products where the *teaser rate* is fixed for two and three years respectively. *Other* product types include ARM-other, Balloon, Two-Step, GPM, GEM and GPARM. The total number denotes the number of originations in each category.

Vintage	FRM	ARM2	ARM3	Other	Total Number
1998	51.33	26.53	4.52	17.62	253264
1999	38.88	29.34	19.21	12.57	369424
2000	32.58	43.29	14.78	9.35	399368
2001	31.70	48.69	12.44	7.17	498494
2002	28.37	54.84	12.62	4.17	755578
2003	33.57	52.60	11.37	2.46	1265769
2004	23.81	59.73	14.64	1.81	1922451
2005	18.66	65.48	13.22	2.64	2266502
2006	19.98	62.56	10.86	6.61	1776422
2007	27.59	50.23	9.92	12.26	330901
Total Number	2527945	5581132	1249659	479437	9838173

Table 2: 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 300-850. Loans coded by the source as with a non-blank documentation code are classified as *Full-doc* whereas those under a No doc program or prospectus are classified as *No doc*. Others are classified as *Low doc*.

	Full doc loans				Low-doc or No-doc loans			
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 3: Distribution of FICO scores conditional on CLTV 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 300-850. Cumulative Loan to value Ratio (CLTV) is the proportion of loans (secured by the property) on all liens in relation to its value.

	CLTV $\leq$ 80				80 < CLTV $\leq$ 90				90 < CLTV $\leq$ 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 4: Credit Score (FICO) Regression

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, co-operative, etc), property location (dummies for the state in which the property is located) and 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 one if the value of the property lies in the *n*-th quartile of all property values in the data and zero otherwise. The results for the years of origination 1998 and 1999 are not reported here, but are available upon request.

Panel A. All Borrower Characteristics

	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-Square	0.0766	0.0877	0.1336	0.1529	0.1684	0.1698	0.1766	0.1486

Panel B. All Borrower Characteristics plus Closing Rate Spread

	2000	**2001**	**2002**	**2003**	**2004**	**2005**	**2006**	**2007**
Intercept	837.59***	886.21***	918.64***	943.31***	895.65***	885.58***	874.07***	916.35***
Full- Documentation	-17.93***	-22.79***	-24.5***	-27.85***	-26.29***	-24.65***	-17.99***	-20.34***
Closing Rate Spread	-16.8***	-18.79***	-24.07***	-24.44***	-26.38***	-28.91***	-28.99***	-27.52***
Owner Occupied	-28.16***	-27.74***	-30.96***	-34.69***	-38.48***	-43.04***	-38.93***	-40.22***
Second Home	-3.21***	-7.49***	-15.51***	-18.65***	-18.42***	-13.41***	-12.74***	-18.28***
Refinance (Cash Out)	-18.64***	-20.72***	-28.54***	-33.17***	-34.39***	-31.24***	-32.47***	-35.68***
Refinance (No Cash Out)	-20.41***	-21.23***	-23.46***	-25.24***	-25.46***	-22.48***	-20.39***	-29.86***
Home Value First Quartile	11***	15.27***	14.64***	14***	10.9***	6.25***	1.9***	6.95***
Home Value Second Quartile	3.89***	6.89***	7.07***	6.88***	4.06***	0.75***	-1.38***	0.383
Home Value Third Quartile	1.53***	4.21***	2.91***	1.44***	-0.18**	-1.73***	-2.89***	-1.38***
Adjusted R-Square	0.2171	0.3215	0.3762	0.4057	0.3823	0.3738	0.3455	0.3506

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

Table 5: Fully Interacted dummy variable Regression of Credit Score (FICO) on Other Borrower Characteristics

Table reports OLS estimates of a filly 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			
Variable	2003-2007	2004-2007	2005-2007	2006-2007
Intercept	671.18***	679.97***	680.24***	685.41***
Dummy	21.59***	9.9***	21.77***	19.15***
Full- Documentation	-19.52***	-20.77***	-20.03***	-20.01***
Full- Doc x Dummy	0.89***	2.33***	1.15***	1.05***
Owner Occupied	-26.39***	-28.46***	-30.48***	-30.83***
Owner Occupied x Dummy	-6.2***	-4.12***	-1.4***	-0.79***
Second Home	-8.59***	-10.73***	-12.3***	-10.15***
Second Home x Dummy	-2.07***	0.46	3.9***	0.78
Refinance (Cash Out)	-18.34***	-24.33***	-29.66***	-31.35***
Refinance (Cash Out) x Dummy	-16.53***	-10.61***	-4.2***	-1.77***
Refinance (No Cash Out)	-16.6***	-19.47***	-22.36***	-22.8***
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***	-4.3***	-4.86***	-5.13***
Home Value Third Quartile x Dummy	-2.76***	-2.17***	-1.52***	-1.39***
Adj R-Sq	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 6: Probability of a 90 day delinquency conditional on FICO

Table reports the delinquency rate for all the vintages, for loans grouped by their FICO score. Delinquency rate is defined in section 5.2 as one minus Kaplan and Meier (1958) survivor function. Delinquency rate is thus one minus the probability of surviving the delinquency event beyond the given age in months.

	FICO: < 620								
	Calendar Year Ending								
Vintage	2000	2001	2002	2003	2004	2005	2006	2007	2008
2000	5.8%	15.3%	23.2%	31.2%	38.0%	44.2%	49.7%	55.3%	61.8%
2001		4.9%	14.3%	24.3%	33.6%	42.0%	48.5%	54.3%	60.1%
2002			4.2%	14.2%	25.1%	36.3%	44.2%	50.8%	57.3%
2003				3.7%	12.7%	23.8%	33.6%	41.6%	48.6%
2004					4.9%	16.0%	28.9%	42.8%	52.9%
2005						7.2%	22.6%	45.0%	62.3%
2006							12.1%	40.0%	66.4%
2007								16.0%	51.1%

	FICO: 620-659								
	Calendar Year Ending								
Vintage	2000	2001	2002	2003	2004	2005	2006	2007	2008
2000	1.8%	6.7%	12.5%	18.6%	24.2%	29.7%	34.3%	38.6%	44.1%
2001		2.0%	7.0%	14.3%	21.8%	28.7%	34.2%	39.0%	44.0%
2002			1.9%	7.8%	15.3%	23.7%	30.5%	36.0%	41.9%
2003				1.7%	6.7%	14.1%	21.0%	27.4%	34.0%
2004					2.3%	8.5%	18.7%	31.3%	40.2%
2005						4.1%	16.0%	39.8%	57.5%
2006							9.8%	36.2%	63.5%
2007								12.4%	44.6%

	FICO: 660-719								
	Calendar Year Ending								
Vintage	2000	2001	2002	2003	2004	2005	2006	2007	2008
2000	1.2%	4.5%	8.2%	12.2%	16.4%	21.2%	24.8%	28.9%	35.0%
2001		1.2%	4.3%	9.0%	13.6%	18.6%	22.7%	26.2%	29.5%
2002			1.2%	4.9%	9.6%	14.8%	18.9%	22.8%	27.9%
2003				1.1%	4.0%	8.1%	11.9%	15.9%	20.1%
2004					1.5%	5.5%	12.4%	21.9%	28.6%
2005						2.6%	11.0%	33.5%	49.3%
2006							7.6%	31.0%	58.5%
2007								11.0%	39.1%

	FICO: >= 720								
	Calendar Year Ending								
Vintage	2000	2001	2002	2003	2004	2005	2006	2007	2008
2000	0.8%	2.8%	5.3%	7.5%	10.1%	12.7%	15.1%	17.7%	21.4%
2001		1.0%	2.4%	4.7%	7.1%	8.8%	10.3%	11.9%	14.1%
2002			0.9%	2.7%	4.8%	7.1%	8.8%	11.1%	13.1%
2003				0.7%	2.1%	4.2%	5.7%	7.3%	9.1%
2004					1.0%	3.3%	6.8%	11.4%	14.3%
2005						1.9%	7.0%	23.2%	34.7%
2006							5.8%	22.4%	45.4%
2007								8.3%	29.7%

Table 7: Mortgage Pricing Sheet, Option one Mortgage Corporation

Rate sheet is for five year fixed mortgage with two year prepays charge. The worksheet assumes full documentation, one unit house, and loan amount in the range $200,000 - $417,000. In case of secondary financing (CLTV > LTV) and credit score less than 660 (or >=660) rate is adjusted upwards by 155 basis points (or 90 basis points).

		LTV			
Grade	Credit 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	

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 8: 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 test 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

9: Determinants of Mortgage Terms

We control for property type (dummies for single-family residence, condo, townhouse, co-operative, 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 one if the value of the property lies in the *n*-the quartile of all property values in the data and zero otherwise. The results for the years of origination 1998 and 1999 are not reported here, but are available upon request.

Panel A. Dependent Variable: CLTV

	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-Square	0.1409	0.1473	0.1595	0.2423	0.2901	0.3077	0.3357	0.3108

Panel B. Dependent Variable: Closing Rate Spread

	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-Square	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 10: Estimated 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, co-operative, 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 one if the value of the property lies in the *n*-th quartile of all property values in the data and zero otherwise. The results for the years of origination 1998 and 1999 are not reported here, but are available upon request.

Panel A. All Borrower Characteristics

	2000	2001	2002	2003	2004	2005	2006	2007
FICO (scaled)	0.4495***	0.4395***	0.4129***	0.3666***	0.3992***	0.4852***	0.5502***	0.5562***
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)

Panel B. All Borrower Characteristics plus Closing Rate Spread

	2000	2001	2002	2003	2004	2005	2006	2007
FICO (scaled)	0.5262***	0.5529***	0.5415***	0.4889***	0.514***	0.6343***	0.7105***	0.826***
Full- Documentation	0.902***	0.8837***	0.8666***	0.806***	0.8348***	0.8023***	0.7541***	0.8162***
Closing Rate Spread	1.2269***	1.2175***	1.2453***	1.2341***	1.2261***	1.2481***	1.1923***	1.3413***
Owner Occupied	0.8739***	0.8702***	0.8889***	0.8721***	0.8447***	0.9252***	0.8958***	0.9336***
Second Home	0.5987***	0.5743***	0.6394***	0.6409***	0.6642***	0.77***	0.7479***	0.7637***
Refinance (Cash Out)	0.7821***	0.7221***	0.7047***	0.5934***	0.554***	0.5464***	0.6134***	0.6855***
Refinance (No Cash Out)	0.9664***	0.8657***	0.8286***	0.6428***	0.5852***	0.6023***	0.6635***	0.7166***
Home Value First Quartile	0.7268***	0.7432***	0.7621***	0.8072***	0.6988***	0.5661***	0.5446***	0.4894***
Home Value Second Quartile	0.8476***	0.8114***	0.8101***	0.8443***	0.7511***	0.64***	0.6326***	0.622***
Home Value Third Quartile	0.8843***	0.8575***	0.8365***	0.8705***	0.8193***	0.8145***	0.811***	0.8087***
LR test H_0: $\beta = 0$	34398	42843	68978	121636	185109	245904	257072	45442
(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 1: 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 like LTV ratio and the closing rate spread are added as regressors for the counterfactual results on display in the lower panel.

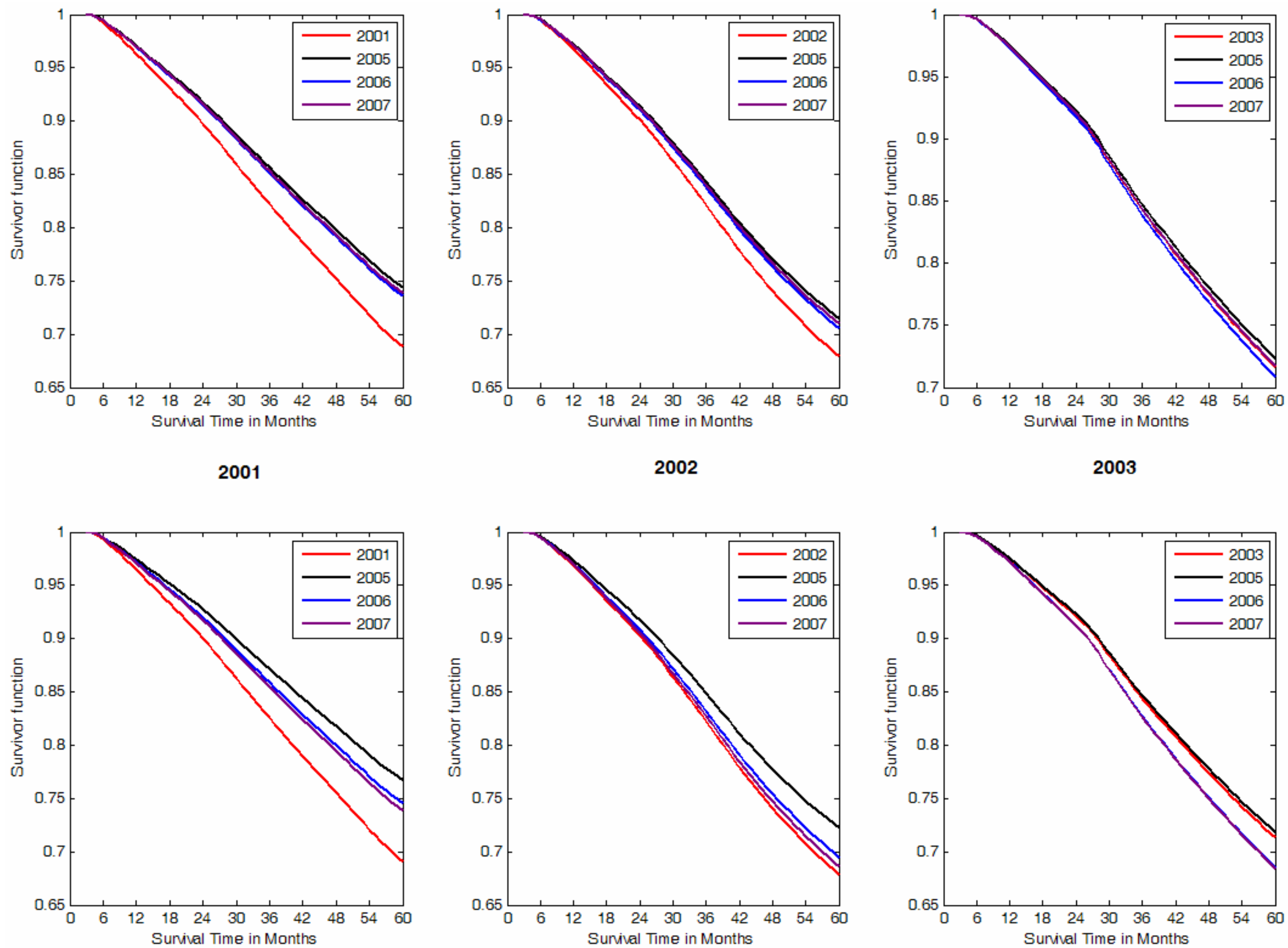

Table 11: Counterfactual Survival analysis

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

Panel 1: Counterfactual Analysis 2001				
Age of Loan (Months)	Survivor Function 2001	Counterfactual Survivor Function 2005	Counterfactual Survivor Function 2006	Counterfactual Survivor Function 2007
12	0.964	0.971	0.97	0.97
	(0.963,0.964)	(0.97,0.972)	(0.969,0.971)	(0.97,0.971)
24	0.897	0.918	0.915	0.916
	(0.896,0.898)	(0.917,0.919)	(0.914,0.916)	(0.915,0.917)
36	0.822	0.856	0.852	0.853
	(0.82,0.824)	(0.855,0.858)	(0.85,0.854)	(0.851,0.855)
48	0.752	0.798	0.792	0.794
	(0.75,0.755)	(0.795,0.801)	(0.789,0.794)	(0.791,0.796)
60	0.688	0.744	0.736	0.738
	(0.685,0.692)	(0.741,0.747)	(0.733,0.74)	(0.735,0.742)
Panel 2: Counterfactual Analysis 2002				
Age of Loan (Months)	Survivor Function 2002	Counterfactual Survivor Function 2005	Counterfactual Survivor Function 2006	Counterfactual Survivor Function 2007
12	0.967	0.972	0.971	0.971
	(0.967,0.968)	(0.971,0.972)	(0.97,0.971)	(0.971,0.972)
24	0.901	0.914	0.91	0.912
	(0.9,0.902)	(0.913,0.915)	(0.909,0.911)	(0.911,0.913)
36	0.821	0.843	0.838	0.84
	(0.82,0.823)	(0.841,0.845)	(0.836,0.839)	(0.838,0.842)
48	0.74	0.771	0.763	0.767
	(0.738,0.743)	(0.768,0.773)	(0.761,0.765)	(0.764,0.769)
60	0.679	0.715	0.706	0.71
	(0.676,0.682)	(0.712,0.718)	(0.703,0.709)	(0.707,0.713)
Panel 3: Counterfactual Analysis 2003				
Age of Loan (Months)	Survivor Function 2003	Counterfactual Survivor Function 2005	Counterfactual Survivor Function 2006	Counterfactual Survivor Function 2007
12	0.974	0.975	0.973	0.974
	(0.974,0.974)	(0.974,0.975)	(0.973,0.973)	(0.974,0.974)
24	0.92	0.922	0.917	0.92
	(0.919,0.92)	(0.921,0.923)	(0.917,0.918)	(0.919,0.921)
36	0.843	0.847	0.838	0.844
	(0.842,0.844)	(0.846,0.848)	(0.837,0.84)	(0.842,0.845)
48	0.775	0.781	0.769	0.776
	(0.773,0.777)	(0.779,0.783)	(0.767,0.771)	(0.774,0.778)
60	0.716	0.723	0.709	0.717
	(0.714,0.718)	(0.721,0.726)	(0.706,0.711)	(0.715,0.72)