ISSN 1936-5330

THE FEDERAL RESERVE BANK *of* KANSAS CITY
RESEARCH WORKING PAPERS

Credit Scoring and Loan Default

Geetesh Bhardwaj and Rajdeep Sengupta

February 2015
RWP 15-02

RESEARCH WORKING PAPERS

Credit Scoring and Loan Default

Geetesh Bhardwaj and Rajdeep Sengupta

February, 2015

A metric of credit score performance is developed to study the usage and performance of credit scoring in the loan origination process. We examine the performance of origination FICO scores as measures of *ex ante* borrower creditworthiness using loan-level data on *ex post* performance of subprime mortgages. Parametric and nonparametric estimates of credit score performance reveal different trends, especially on originations with low credit scores. The data suggest a trend of increased emphasis on higher credit scores accompanying a trend of increased riskiness in other origination attributes. Over time, this increased emphasis on credit scoring coincided with deterioration in FICO performance largely due to the fact that higher credit score originations of later cohorts were more likely to have riskier attributes. However, controlling for other attributes on originations and changes in economic conditions, we find that, as measures of borrower ranking, FICO performance on subprime loans over the years remains fairly stable.

JEL Codes: G21
Keywords: credit score, information sharing, subprime, performance

1. Introduction

Over the last couple of decades, technological advances and private arrangements of information sharing have increased the use of credit scoring in almost all forms of loan origination (Altman and Saunders, 1998; Berger et al. 2005). However, the use of credit scoring is not without its limitations (Mester, 1997; Avery et al. 2000). Still, “a good model should be able to accurately predict the average performance of loans made to groups of individuals who share similar values of the factors identified as being relevant to credit quality” (Mester 1997, p.11). Despite the limitations of credit scoring, most approval processes continue to use credit scores as a measure of borrower creditworthiness at the time of loan origination (Avery et al. 2003; Brown et al. 2009). The continued importance of credit scoring in loan approvals merits careful study of the use and performance of such metrics.

We undertake an investigation at understanding the usage and performance of credit scoring in the U.S. subprime mortgage market. Using loan-level data available on subprime mortgages, we identify how credit scores were used in conjunction with other observable characteristics at the time of origination. We study how well credit scores measure ex ante credit risk using data on ex post performance of the origination. To this end, we introduce simple measures of credit score performance that help determine the impact of credit scoring and its usage in terms of observed loan performance.

The industry standard for measuring consumer credit risk in the U.S. is the FICO score.[1] There are important reasons to focus on FICO performance in the subprime mortgage market. While FICO scores have been an integral part of the prime mortgage approval process, they

[1] FICO score refers to the credit score developed by Fair Issac Corporation, (now known as FICO). Throughout this paper the terms “FICO score” and “credit score” are used interchangeably. See section 2 for more details on the FICO score.

played an essential role in extending credit beyond the prime segment to some the riskiest consumer loans in recent times, namely, U.S. subprime mortgages (Fishelson-Holstine, 2005).[2] Subsequently, high default rates of subprime mortgages have raised important questions about the efficacy and usage of credit scoring in loan origination (Demyanyk, 2008).[3] At the same time, examining FICO performance for the subprime segment allows us to study credit scoring performance at the lower-end of the credit score range. Stated differently, it helps determine whether FICO scores can be a stable metric for credit risk in lending to credit-impaired borrowers. In this sense, it would also help increase our understanding of the challenges faced in lending to credit-impaired borrowers.[4]

Our results reveal that, for the most part, the performance of credit scoring as a measure of (relative) credit risk remains fairly stable. The results provide little evidence of deterioration in the performance of FICO scores as rankings of borrower ex ante credit risk. However, they also suggest a pattern in which credit scoring was likely used to offset other riskier attributes on the origination—leading to an unconditionally higher rate of default, especially on originations with low credit scores.

Before discussing credit score performance, some features of the data are worth emphasizing. First, we find little evidence of credit score inflation in the U.S. over this period: there is hardly any difference between the credit score distribution of the credit eligible U.S.

---

[2] FICO scores were first recommended for use in mortgage lending by Fannie Mae and Freddie Mac back in 1995. In the subsequent years, they became a significant factor in the development of the U.S. subprime mortgage market (Fishelson-Holstine, 2005).

[3] Foust and Pressman (2008) document the industry view on how the FICO scores were "being blamed for failing to flag risky home-loan borrowers." From a policy research perspective, Demyanyk (2008) uses the performance of subprime mortgages to express doubts about the effectiveness of FICO scores.

[4] To be sure, we do not address the question whether lending to credit-impaired borrowers is a desirable objective from the standpoint of social welfare (see Bolton and Rosenthal, 2005 for a discussion). Our objective is somewhat modest: If financial inclusion were to be the desired policy objective, we examine whether FICO scores could be reliably used a metric of credit risk in lending to credit-impaired borrowers.

population for the period 2000-2002 and that for the period 2004-2006.[5] In contrast, we observe a clear improvement in the distribution of credit scores on subprime originations between the 2000-2002 cohorts and 2004-2006 cohorts.[6] Second, as has been well documented, there was an increase in the proportion of subprime originations with riskier attributes over this period (Mayer et al. 2009). Interestingly, we find that an origination with a riskier attribute (such as lower documentation and higher LTV) was more likely to have a higher FICO score in the 2004-2006 cohorts than a similar origination in 2000-2002 cohorts. Stated differently, we find that, among the various attributes on the origination, the apparent tradeoff between the credit score and other attributes (e.g., LTV) grew larger over time. Finally, we show that there was an overall increase in FICO scores from the earlier cohorts to the later cohorts even after adjusting for other origination attributes. This result indicates that the increase in credit scores across cohorts might be interpreted both in terms of adjustment for the increased riskiness in other origination attributes and the increased strength of adjustment to such riskier attributes. In sum, this pattern is suggestive of an increased emphasis on credit scoring to offset other attributes on the origination.[7]

We develop a simple measure of FICO score performance in terms of ex post loan performance. It is important to mention here that both academics and practitioners alike view FICO scores in terms of rankings of borrowers rather than absolute metrics that provide time-

[5]Anecdotal evidence has been provided showing that credit scoring itself is subject to manipulation (Foust and Pressman, 2008). In such cases, increases in a borrower's credit score occur without any increase in the borrower's creditworthiness. We discuss this issue in greater detail in Section 6.
[6]This result may seem odd at first, but it is useful to remember that on an annual basis over 60 (50) percent of subprime originations were (cash-out) refinances. To the extent these refinances included borrowers refinancing from prime segments, the improvement in credit scores should not appear anomalous.
[7]Needless to say, this is merely suggestive. As with most of this literature, the analysis here abstracts from a structural framework that models the incentives of agents involved in the origination process. For a heuristic discussion of the various agents involved and their incentives, see Ashcraft and Schuermann (2008).

invariant estimates of the probability of default.[8] Continuing in this vein, we propose measures of FICO scores performance that determine whether such credit score "rankings" are maintained in terms of observed loan performance. Any metric of FICO score performance should ideally account for factors that confound the effect of credit scores on loan performance. This becomes especially relevant in comparing loans that perform under different macroeconomic conditions. In short, while determining the effect of FICO scores on loan performance, we should ideally control for other risk attributes on the origination as well as the environment in which these loans perform.[9]

Our measure of FICO score performance is the difference between the survival probabilities of originations in a higher FICO score group and the survival probabilities for originations in the next lower FICO score group *in the same cohort*.[10] Since the FICO score groups are of the same cohort, they are each subject to the same underwriting trends and macro shocks (events), allowing us to net out the effects of the macroeconomic environment in our measure of FICO score performance.

Moreover, we derive nonparametric as well as parametric measures of this difference in probabilities. The nonparametric measure uses the unconditional survival probabilities. On the other hand, the parametric measure is derived by using estimates from a competing-risk hazard regression that allows us to control for other origination characteristics and local economic conditions such as unemployment and home prices (negative equity). In this manner, our

[8]See Section 2 for a discussion of FICO scores as rankings of borrowers. For reasons as to why absolute measures of credit scores performance are misleading, please refer to the discussion in Appendix C.
[9]Ceteris paribus, a loan in the 2002-2004 cohort is likely to have a lower default rate than an origination with a similar credit score in 2004-2006, simply because the latter performed in an environment of rapidly increasing home prices. Rapidly increasing home prices allow the borrower to avoid default by selling the property or refinancing the loan—an option that is no longer available under stagnant or declining prices.
[10]While our choice of credit score intervals as groups is ad hoc, the use of credit score groups is motivated by institutional factors (see Sections 2 and 3 for details). Our results are robust to variations in the choice of credit score groups.

approach attempts to mitigate the effect of factors that might confound measures determining the performance of credit scoring in terms of observed loan performance.

We find that nonparametric and parametric estimates reveal different patterns in FICO score performance depending on the level of origination FICO scores. At low FICO scores, nonparametric estimates typically show deterioration in performance for later cohorts. In contrast, this pattern of deterioration is reversed when we control for other attributes on the origination—FICO score performance shows no deterioration in terms of our parametric measure. Further, for high FICO scores, neither the nonparametric nor the parametric measures find deterioration in FICO score performance.[11]

The results can be explained in terms of the patterns of FICO score usage described above. Significantly, the usage patterns also vary with the FICO score level. For low credit-score levels, the higher of two adjacent FICO score groups is more likely to have riskier attributes on the origination in later cohorts than in earlier cohorts. Consequently, this lowers the unconditional performance of FICO improvement for the later cohorts as obtained from our nonparametric measure. For the same reason, this decline is reversed if we control for other origination attributes: our metric shows improvement in FICO score performance for later cohorts in terms of our parametric measure. The pattern of FICO score usage is different at higher levels of origination FICO score. Adjustment of riskier attributes with higher FICO scores is not significantly large for earlier cohorts and remains roughly unchanged over the cohorts. Accordingly, FICO score performance at high FICO score levels shows improvement over the cohorts in terms of both parametric and nonparametric measures.

Why do parametric measures of FICO score performance show improvement at later

[11] Given the selection issues for ARM and FRM products in the subprime segment, as shown in Pennington-Cross and Ho (2010), we demonstrate that our results also hold in the subsample of Subprime-ARM loans and Subprime-FRM loans.

cohorts compared with earlier cohorts? Our interpretation is that later cohorts (2004-2006) faced a high-default environment, and while originations of all FICO scores did poorly, the performance of low FICO score originations was significantly poorer.[12] Consequently, improvements in FICO scores demonstrate significantly higher improvements in survival probabilities for later cohorts. In summary, our results show a deteriorating pattern of FICO score usage—evidence suggesting the allowance of higher ex ante riskier characteristics on originations with higher scores. This implies higher unconditional default rates over time, but if one controls for such riskier attributes, our results provide little evidence of deterioration in the performance of FICO scores as rankings of borrower ex ante credit risk.

This study adds to the literature on information sharing, credit scoring, and loan default. Several studies point to the fact that information sharing, credit scoring increases lending volumes, especially to borrowers of low credit quality (Berger et al. 2005). As a result aggregate default rates may increase because of an increase in the proportion of lower-grade borrowers in the credit-eligible pool (Brown et al. 2009). To the best of our knowledge, we do not know of any study examining the performance of credit scoring for such low-grade borrowers. This study allows us to examine how credit scoring is used and how it performs within this lower-grade of credit impaired borrowers.

Our results find broad support in the literature on credit scoring for credit-impaired (subprime) borrowers. For example, emphasis on credit scoring has been shown to increase default rates by reducing the incentives for screening (Keys et al. 2010). This leads to a reduction in borrower quality across the spectrum (generated due to a low screening intensity) which is more severe for borrowers on whom such screening was more valuable, namely low FICO score

[12]We attribute this non-linearity to the interaction effect of non-increasing home prices with FICO scores on default. See Table 11 Bhardwaj and Sengupta (2012) for estimates of this interaction effect (Appendix C).

borrowers. While their study points to the lack of screening intensity, we demonstrate greater risk-taking in terms of other observable characteristics. In essence, both studies point to different outcomes that result from an emphasis on FICO scores in loan origination. Therefore, while we believe that the (lack of) screening intensity had a very important role, it is not the entire story (see Appendix C). In this sense, our study may be viewed as a complementary explanation of the data that has received scant attention in the literature. For low-FICO score borrowers, the increased risk in other origination attributes was accompanied by increasing credit scores over time. As a result, unconditional metrics show a fall in credit score performance because of other attributes on the origination. These results suggest that increased emphasis on credit scoring to offset other riskier attributes can have harmful consequences on market outcomes, especially for low-FICO score borrowers.

We begin by providing a brief primer on FICO scoring in Section 2. Section 3 discusses the patterns of credit scoring use over the various cohorts of subprime originations. The parametric and nonparametric measures of FICO performance are explained in Section 4, and the results are provided in Section 5. Section 6 concludes.

2. A Short Primer on FICO scores

FICO scores are the best known and most widely used consumer credit scores in the United States. They are published by Fair Isaac Corporation, which maintains that FICO is a statistical summary measure of credit risk based on information from a consumer's credit files.[13] The measure ranges between 300 and 850 with higher scores going to the more creditworthy borrowers. It is important to mention here that the statistical models used to determine FICO

[13]Fair Isaac (2007) argues, “A credit score is a number that summarizes your credit risk, based on a snapshot of your credit report at a particular point in time. A credit score helps lenders evaluate your credit report and estimate your credit risk.”

scores are proprietary and therefore not public information.[14]

Two features of FICO scores require elaboration. First, it is important to note that FICO scores do not provide for an absolute quantification of risk. Rather, they present a relative ranking of borrowers in terms of their credit risk. This feature of FICO has been confirmed in academic studies as well as industry releases. Prominently, Keys et al. (2010, p. 316, emphasis added) note that, "FICO scores provide *a ranking of potential borrowers*…." In the realm of practitioners, this feature of credit scoring is common knowledge: In their public release statements, Transunion, the third largest credit bureau in the U.S., claim that:

> *Credit scores are not an absolute statement of risk for an individual consumer, rather they state a consumers' risk in relation to other consumers.*[15]

Therefore, any assessment of the performance of FICO would have to demonstrate that the rank ordering of FICO performance was preserved at all FICO scores. Stated differently, determining the performance of a given FICO score should be made relative to a higher or lower FICO score as will be demonstrated below.

Second, most observed FICO scores have an inherent "lumpiness" to them. Individual FICO scores are obtained from each of three credit bureaus based on the information that each credit bureau has on the borrower.[16] To the extent that information on an individual borrower's credit report differs across the credit bureaus (either because of differences in their data collection methods and/or errors and omissions), the FICO algorithm consequently runs on different information sets. Therefore, it is not unlikely to obtain different FICO scores from

[14]While the detailed algorithm is not available, Fair Isaac provides us with summary details about the information that is used in calculating FICO scores. See http://www.myfico.com/crediteducation/whatsinyourscore.aspx for details.
[15]http://www.transunion.com/docs/rev/business/financialservices/VantageScore_CreditScoreBasics-Part1.pdf
[16] Notably, Fair Isaac Corporation hosts the proprietary formulas used to calculate FICO scores—but, the company does not collect information on each individual borrower's credit history. Credit information on the borrower is collected independently by three major credit bureaus in the U.S., namely, Equifax, Experian and Transunion.

different credit bureaus for the same borrower.[17] At the same time, however, the information sets (at each bureau) on each individual borrower are not likely to be widely different from each other. Therefore, if an individual borrower were to obtain his FICO score from all three bureaus simultaneously, he or she might obtain three different FICO scores but these FICO scores are likely to be "lumped" together.[18] To summarize, the relative quantification of risk and lumpiness are two important properties of FICO.

3. Data and Summary Trends

3.1 Subprime Mortgages and Origination FICO

We use loan-level data on over seven million first-lien mortgages originated during 2000-2006 from the subprime database in the CoreLogic (formerly, Loan Performance) data repository.[19] This is widely regarded as the most comprehensive database on subprime mortgages and captures over 90 percent of the mortgages that have been securitized as subprime. For the purposes of this study we include all loans securitized under a subprime pool in the data repository as subprime.[20] The database includes detailed information about mortgage and property characteristics at the time of origination.[21] In addition, it includes certain borrower level data, such as the borrower's FICO score at the time of origination. The origination FICO score in our data is typically the one obtained from one of three credit bureaus chosen by the mortgage

[17] For more details on this, see http://www.myfico.com/crediteducation/questions/Why-Three-Scores.aspx
[18] We use the term "lumpiness" to reflect the fact that observed FICO scores have this margin of error. Just to be sure, we are not claiming that there is an error in the process with which FICO evaluates borrower credit history. The error is largely due to the fact that FICO may process different information on the same borrower because the source of information is a different credit bureau.
[19] See http://www.loanperformance.com/data-power/default.aspx. While the database contains Alt-A mortgages, our focus is restricted to subprime mortgages only.
[20] This is the most commonly used definition on subprime mortgages, especially for empirical work on databases that include only securitized mortgages such as the one used here. Foote et al. (2008) observe that other definitions of subprime mortgages tend to exclude large segments of borrowers that used subprime mortgages.
[21] Details on this proprietary database, including its evolution, coverage, and comparison with other mortgage databases, are available in GAO (2010).

originator at the time of origination. Therefore, in examining the performance of FICO, we have to allow for a degree of lumpiness in the FICO score.[22]

In what follows, we describe the summary data patterns for origination FICO on subprime mortgages and how they compare over time with (i) FICO scores in the general population and (ii) other origination attributes on subprime mortgages. Much of these data patterns are fairly well-documented in the literature on subprime mortgages. Therefore, in the interest of brevity, we point out only the sailent features of these patterns here. A full description of these patterns is presented in a web appendix (see Web Appendix B).

3.2 Subprime and Population FICO scores

Origination FICO of subprime mortgages improved significantly over time. The probability that a subprime borrower has a lower credit score is significantly higher on originations of earlier cohorts (2000-2002) than on originations of later cohorts (2004-2006). At the same time, the data also reveal a marginal improvement in the credit score distribution from 2000-2002 to 2004-2006 for the population as well. However, a cohort by cohort comparison between the periods 2000-2002 and 2004-2006 reveals that the improvement in credit scores is statistically significant for subprime originations but not for the entire U.S. consumer population. This would suggest that, while there is an improvement over the cohorts in the origination FICO for subprime mortgages, such an improvement cannot be attributed to trends in the overall population.

3.3 Subprime FICO scores and Other Origination Attributes.

[22] The data does not include the name of the credit bureau that was chosen by the originator. Therefore, we are unable to control for any systematic variation in the collection of consumer data across the three bureaus either.

As has been well documented in this literature, there was an increase in the proportion of subprime originations with higher LTV and originations lacking full documentation over the cohorts (see, for example, Mayer et al. 2009). Bhardwaj and Sengupta (2014) present evidence showing that this deterioration in other origination attributes was accompanied by increases in FICO scores on the origination. Stated differently, an origination with a riskier attribute (such as lower documentation and higher LTV) is more likely to have a higher FICO score in the 2004-2006 cohorts than in the 2000-2002 cohorts. In addition, they find that the tradeoff or adjustment between riskier attributes on the origination and higher credit scores on the same origination grew stronger over time. Below, we use regressions to examine these trends further.

Following Bhardwaj and Sengupta (2014), we run cohort-by-cohort least squares regressions of origination FICO on other non-score origination attributes. The estimated coefficients are reported in panel A of Table 1. Clearly the regression estimates is not meant to demonstrate any causal relationship. Rather, the coefficients show the equilibrium relationship between credit scores and other origination characteristics. Not surprisingly, most of the coefficients have the expected sign: For example, origination credit scores increase with the LTV and the lack of full documentation on the origination. Moreover, an increase in the magnitude (absolute values) of these coefficients over the cohorts appears to suggest that the strength of adjustment of increased FICO for riskier attributes increased over this period. Therefore, not only did FICO scores increase unconditionally over the cohorts, but also the strength of adjustment to offset other riskier attributes on the origination increased over the cohorts in our sample.

Panel B of Table 1 reports the estimates for the full sample (all cohorts) of the regression presented in Panel A. In addition to the regressors in Panel A, Columns (1) and (2) include dummy variables that take the value 1 for the later cohorts 2003-2006 and 2004-2006, respectively, and zero otherwise. A positive value of the coefficient on the dummy for later cohorts indicates an overall increase in FICO scores from the earlier cohorts to the later cohorts even after adjusting for other attributes on the origination.

These results indicate that the increase in credit scores over the cohorts can be explained both in terms of adjustment for the increased riskiness in other attributes on the originations and the increased strength of adjustment to such riskier attributes. To conclude, the trend of increasing credit scores accompanying the increase in credit risk exposure in terms of certain origination attributes (such as lack of full documentation and high LTV) suggest an attempt to offset this higher credit risk by increasing the average quality of borrowers (as measured by their credit scores) to whom such loans were made.

3.4 Understanding the Trends in Subprime Origination FICO

Importantly, the trends described above pertain to changes in hard information on the origination. This is different from much of the literature that has emphasized the collection of soft information, or lack thereof, on subprime originations. Needless to say, further research is needed to investigate how such trends (in both hard and soft information) were determined by the incentives of agents involved in the origination process.[23]

The evidence of an overall increase in credit scores prompted assertions that unsuspecting prime borrowers may have been misled by originators into subprime products (Brooks and

[23]The incentives of agents including, but not limited to, the mortgagor, originator, arranger, servicer, investor, warehouse lender, and the credit rating agency played an important role in the origination process (Ashcraft and Schuermann, 2008). Unfortunately, our data put limitations on investigating the role of agents' incentives on trends in subprime originations. Therefore, as with much of this literature that examines these data, our treatment of the origination process is akin to that of a neoclassical production function of mortgages. Stated differently, the analysis abstracts from the incentives of the various agents involved in the origination (production) process and examines the characteristics and performance of subprime mortgages as (final) products of the origination (production) process.

Simon, 2007). However, consistent with the evidence presented above, Foote et al. (2008) argue that originations to borrowers with such high credit scores were marked as subprime because of other attributes on the origination. In order for borrowers with such high credit scores to originate prime mortgages, the risk on other attributes such as documentation and LTV of the origination would have to be lower as well. In summary, originations of later cohorts include a significantly large proportion of borrowers with higher credit scores. However, as noted previously, these high credit scores were accompanied by other riskier attributes on the origination and therefore not classified as prime.

As mentioned earlier, we are not the first to observe this trend over the subprime cohorts toward higher FICO scores. Practitioners have often referred this phenomenon as a movement in the subprime segment toward the Alt-A market segment (Bhattacharya et al. 2006). Some observers have even described this as a creation of the Alt-B market segment (Zimmerman, 2006).[24] However, as argued previously, simply raising FICO scores on the origination would not necessarily make them Alt-A loans. This leads to some obvious questions: Could higher origination credit scores on later cohorts succeed in offsetting the increased riskiness in terms of other attributes on the origination? Did the performance of FICO scores withstand post-origination factors, such as the stagnant or deteriorating home prices around 2006-2007? We attempt to answer these questions in the next two sections, where we first devise a metric for FICO score performance and then examine how such scores performed over the cohorts of subprime mortgages.

4. Measuring FICO performance

[24]For example, Zimmerman (2006, p. 106) observes, "... FICOs in subprime at 624 in 2004 are at a record high level. In part, the increase in subprime FICOs reflects the rapid move by subprime issuers into the lower end of the Alt-A market, sometimes referred to as the Alt-B or the gap part of the non-agency market."

In this section, we develop our measure of origination credit score performance. We consider two features of the measurement problem. First, drawing from earlier assertions of FICO scores as a ranking of borrowers, we evaluate the performance of a given FICO score relative to that of a lower or higher FICO score. In this sense our measure determines whether the rank ordering in FICO score is reflected in terms of loan performance. Second, we take into account the "lumpiness" of FICO scores as discussed above. Therefore, we evaluate the performance of FICO score groups as opposed to the performance of individual FICO scores. Given that FICO scores from different credit bureaus are likely to lie in the neighborhood of the observed FICO score in our data, it makes sense to evaluate performance of FICO score groups rather than individual scores.

Our metric for credit score performance is the difference in survival probabilities for an origination with a higher FICO in comparison to one with a lower FICO *in the same cohort*. First, we split our sample into originations belonging to different FICO score groups as discussed above. Next, we calculate the (unconditional) survival probabilities for originations within each FICO score group. Our nonparamertic measure of origination FICO performance is the difference in the survival probabilities of a given origination FICO score group to that of its immediately lower FICO score group of the same cohort. Since both (higher and lower) credit score groups are of the same cohort, they are each subject to the same underwriting trends and macro shocks (events), allowing us to net out the effects of these factors in our measure of credit score performance.

For the parametric estimates of our performance measure, we use estimates from a competing-risk proportional hazard model.[25] Prepayment and default on the mortgages are

[25]It is important to point out that the distinction here between our non-parametric and parametric measure is that, while the former is the difference between the unconditional probabilities, the latter measure controls for other

modeled as competing risks. The FICO score groups enter the regressions as explanatory (dummy) variables. The estimated hazard ratio for a given FICO score group are then used to derive a parametric measure of survival probabilities for that group. Our parametric measure of origination FICO performance is the difference in the *estimated* survival probabilities of a given origination FICO score group to that of its immediately lower FICO score group of the same cohort. A formal description of the methodology used to derive our parametric estimates is presented in Appendix A.

5. Results

Following standard conventions in this literature, we record the first serious delinquency (90-day delinquency) event as an indicator of default on the loan. Therefore, the survival probabilities discussed below are the probabilities of surviving a 90-day delinquency event two years from the month of origination. Alternative definitions of default yield qualitatively similar results. The use of two calendar years as the size of the interval is motivated by the fact that most originations were hybrid-ARM mortgages designed in theory to ensure a refinance in two years (Gorton, 2008).[26] Indeed, the large increase in defaults on subprime originations during 2006-2007 occurs well before the two-year period.

Pennington-Cross and Ho (2010) point to selection issues based on the choice of mortgage product, especially ARM loans as opposed to FRM mortgages. Therefore, in the interest of completeness, the entire analysis is conducted first for the full sample of loans and then separately for subsamples of ARM and FRM loans respectively. As shown below, all of the

attributes on the origination while calculating this difference. Of course, one can control for other origination attributes using non-parametric estimation techniques. Strictly speaking, therefore, some inaccuracies may creep into terminology of this nature. Our aim here is to make a semantic distinction between the two measures in a simple way.

[26]This is particularly true for 2/28 hybrid ARM products--the most popular product in the subprime mortgage market.

results on FICO performance hold for all three samples. However, these results are more prominent in the full sample and the Subprime-ARM subsample but less so in the Subprime-FRM subsample. It bears to keep in mind that, on an annual basis, FRMs are never more than one-third of all subprime originations and for the peak years (2004-2006) less than a quarter of all subprime loans.

5.1 Non-parametric measure of FICO performance

Tables 2.1-2.3 report the difference (increase) in probability of surviving a 90-day delinquency event two calendar years after origination. The results for the full sample of subprime loans are reported in Table 2.1, while that for subsamples of Subprime-ARM and Subprime-FRM are reported in Table 2.2 and Table 2.3 respectively. As mentioned above, we split the sample into various FICO score groups for the purposes of this analysis. The results for two such groupings starting at a FICO of 540 are recorded here: the first at intervals of 40 points (Panel A) and the second at intervals of 20 points (Panel B).[27] The rows in Table 2.1-2.3 show the percentage-point increases in survival probabilities for originations in a higher FICO score group relative to those in its immediately lower FICO score group. Needless to say, this is not the only way to compare performance between two FICO score groups. An alternative way of implementing this measure would involve measuring the increase in survival probabilities for a given increase in credit score and then averaging across all credit scores in the FICO score group. This second method yields materially similar results that are available on request.

Three features of the non-parametric measure of differences in survival probabilities are noteworthy. First, the metric is typically smaller at later cohorts than earlier cohorts at low levels

[27] We run robustness checks by varying both the cutoffs and the starting FICO of the groups. The results are qualitatively similar and available upon request.

of FICO (top rows in both panels of Table 2.1-2.3). Consequently, changes in this metric suggest a deterioration of credit score performance for the later cohorts at low levels of FICO score. Second, at high levels of FICO scores (bottom rows in both panels of Table 2.1-2.3), the metric is typically greater at later cohorts than earlier cohorts. Likewise, this suggests an improvement in FICO performance for later cohorts in terms of our nonparametric measure. Third, these changes are robust if one considers bigger FICO score groups, and therefore bigger transitions as shown in Panel A. In contrast, the metric is significantly noisier for smaller FICO groups as shown in Panel B. This feature of the data appears to lend support to the notion of a degree of lumpiness to FICO scores, as discussed above. As mentioned above, these results are more prominent in the full sample and the ARM subsample but less so in the FRM subsample. Having derived the nonparametric (unconditional) measures of credit score performance on loan default, we turn to parametric measures wherein we derive the measure of credit score performance after controlling for other origination characteristics.

5.2 Parametric Measure of FICO performance

To derive this measure, we estimate the competing risk proportional hazard model for prepayment and default. The estimated hazard ratios for the three sets of regressions (full sample, Subprime-ARM and Subprime-FRM) are given in Tables 3.1-3.3 respectively. The regressions are estimated using dummy variables for individual FICO score groups. The FICO score groups selected for the regressions in Tables 3.1-3.3 are the same as those given in Panel A of Table 2.1-2.3 respecitvely.[28] The estimated hazard ratios are provided in Table 3.1-3.3 with the dummy for the lowest FICO score group (less than 540) chosen as the omitted dummy variable. The results

[28] In the interest of brevity, the results for regressions using the groups of Panel B are not reported here, but are available on request.

appear robust across all cohorts; the estimated hazard ratios are highly significant in all specifications. A hazard ratio greater than 1 indicates that the increase in the relevant variable is associated with an increase in the default hazard—the converse is true for a hazard ratio that is less than 1. Therefore, the hazard ratio of 0.604 on the FICO: 580-619 dummy variable for the 2003 cohort implies that the default hazard on 2003 originations with FICO scores in the interval 580-619 is 0.604 times the default hazard on 2003 originations with FICO scores less than 540.

As determinants of loan performance, we consider the principal origination characteristics such as the origination FICO score, CLTV, and loan documentation. *Full Documentation* is a dummy variable equal to 1 if the mortgage has full documentation and zero otherwise. *Fees and Points* are the fees and discount points charged by the lender at settlement on a 30-yr FRM prime mortgage, taken from the Freddie Mac PMMS Survey. *Present Value Annualized Ratio (PVAR)* measures the ratio of the present value of the payments on mortgage principal outstanding using the existing mortgage rate to that using the current rate available on refinance. *Interest Volatility* is the standard deviation of the six-month LIBOR for the previous 24 months. *House Price Volatility* is the standard deviation of house price change in the MSA-level FHFA house price index for the previous 24 months. *PosUnempG* is a dummy variable equal to one if the MSA records and increase in the unemployment rate over the previous year and zero otherwise. *Negative Equity* is a dummy variable equal to 1 if the current LTV exceeds 100% and zero otherwise.[29]

Tables 4.1-4.3 report the increases in probability of surviving a 90-day delinquency after two years for originations in a higher FICO score group relative to those in its immediate lower

[29] In addition, we also control for loan characteristics such as loan type (conventional, VA, FHA, government, etc.), loan purpose (purchase, cash-out refinance, no cash-out refinance, etc.) and property characteristics such as property type (condo, townhouse, etc.) , property location (dummies for the state in which the property is located), loan source (dummies for broker, realtor, wholesale, retail etc.), occupancy status (dummy variables for owner-occupied, investor-owned, or second home) and a dummy for mortgages with prepayment penalties.

FICO score group, after controlling for other attributes on the origination. The results for the full sample of subprime loans are reported in Table 4.1, while that for subsamples of Subprime-ARM and Subprime-FRM are reported in Table 4.2 and Table 4.3 respectively. For example, panel A of Table 4.1 shows parametric estimates for the difference in survival probabilities of adjacent FICO score groups such as those given in panel A of Table 2.1.[30]

Importantly, after controlling for other attributes on the origination (as given by the regressions in Table 3.1-3.3), the increases in survival probabilities do not show deterioration over the cohorts. Notably, this result holds true for both high and low FICO score levels. At low levels of FICO scores, the deterioration in performance as seen in the nonparametric estimates (top rows in Table 2.1-2.3) above is now reversed (top rows in Table 4.1-4.3). Controlling for other attributes on the origination, we find that the performance metric at low FICO levels for later cohorts is typically greater than those at earlier cohorts, thereby suggesting an improvement in FICO performance for later cohorts (top rows of both panels in Table 4.1-4.3). Again this result holds at high levels of FICO, which means that results using nonparametric measures continue to hold in the parametric case. Importantly, these results hold for the full sample, Subprime-ARM and Subprime-FRM; however, the results are stronger in the full sample and the Subprime-ARM subsample but less so in the Subprime-FRM subsample.

5.3 Low FICO scores versus High FICO scores

Comparing the values of our metric obtained in Table 2.1 with those obtained in Table 4.1 reveals an interesting pattern for high and low FICO scores on the origination. In the nonparametric case, we documented deterioration in FICO performance for later cohorts at low

[30]Similarly, Panel B of Tables 4.1-4.3 show the same groups as Panel B of Tables 2.1-2.3 respectively. In the interest of brevity, we do not provide the corresponding regression estimates for the groups in Panel B of Table 3. The results are available on request.

FICO levels (top rows in both panels of Table 2.1). However, after controlling for other attributes, parametric estimates in Table 4.1 show that this pattern is reversed. Stated differently, the performance metric is typically lower for later cohorts in terms of our nonparametric measure but higher for later cohorts in terms of our parametric measure. In contrast, the FICO performance metric at high FICO score levels is higher for later cohorts than earlier cohorts under both measures. That is, we observe an improvement in FICO performance at high FICO levels (bottom rows in both panels of Tables 2.1 and 4.1) in terms of both nonparametric and parametric measures.

To explain these patterns in our results, we recall our summary results on origination FICO trends in relation to other origination attributes. For the sake of exposition, we focus our attention on one such attribute, namely, LTV. Using the adjacent FICO score groups listed in Panel A of Table 4.1, we plot the (kernel) density functions for LTV for each adjacent group-pair in Figure 1. The plots in the left column show the distribution of LTV for the 2000-2002 cohorts, whereas plots in the right column show the distribution of LTV for the 2004-2006 cohorts, respectively. The top, middle, and bottom rows show the LTV distribution for adjacent FICO score group-pairs of less than 540 and 540-579, 580-619 and 620-659, and 660-699 and 700-739, respectively. These group pairs correspond to the first, third and fifth rows in Table 2.1.

Almost always, the LTV kernel density plot for the higher FICO score group (in gray) lies to the right of the kernel density plot of its immediately lower FICO score group (in black-dashed). At high FICO score groups this difference is marginal but increases progressively as one moves to lower FICO score groups. Moreover, for the low FICO score groups this difference appears to have increased significantly over the cohorts from 2000-2002 (top row, left column in Figure 1) to 2004-2006 (top row, right column in Figure 1). In contrast, the difference is marginal at high FICO score levels to begin with. In addition, there is hardly any change in these differences from the 2000-2002 cohorts (bottom row, left column in Figure 1) to the 2004-2006 cohorts (bottom row, right column in Figure 1) at high FICO levels.

Formally, we conduct Anderson's (1996) test for stochastic dominance described above for each of the six plots in Figure 1. The results are not stated here but are available on request. We are able to establish that in three of the six plots, the LTV distribution of the higher FICO score group stochastically dominates that in its immediately lower FICO group. In particular, the LTV distribution for the FICO group 540-579 stochastically dominates that for the less than 540 FICO score group for both 2000-2002 and 2004-2006 cohorts. In addition, the LTV distribution for the FICO group 620-659 stochastically dominates that for the 580-619 score group for 2004-2006 cohorts but not for the 2000-2002 cohorts. Lastly, we fail to establish stochastic dominance of the LTV distribution for the FICO group 700-739 over that for the 660-699 FICO score group for both 2000-2002 and 2004-2006 cohorts. In summary, these results formalize what was described above: The difference between LTV on higher FICO score originations and that on the lower FICO score originations increased from earlier cohorts to later cohorts. Significantly, these differences are more discernable at low levels of FICO scores than at high FICO score levels.

The contrast in the LTV distribution between high and low FICO score originations can help explain the anomalous trends in our performance metric over the cohorts. At low FICO score levels, the higher of the two adjacent FICO score groups is more likely to have riskier attributes on later cohorts than earlier cohorts. First, this lowers the unconditional performance of FICO improvement for the later cohorts as shown by our nonparametric measure in Table 4. Once we control for these riskier attributes, the pattern is reversed—we observe an improvement in FICO performance for later cohorts. In contrast, there is a marginal difference in LTV between adjacent FICO score groups at high FICO score levels. Moreover, this difference does not show any discernible change for the later cohorts in our sample. Therefore at high FICO levels, our performance metric shows improvement for later cohorts in terms of our nonparametric measure; and, even after controlling for other origination attributes, FICO performance at high score levels shows improvement for the later cohorts in terms of our parametric measure.

While our results do not establish deterioration in performance of credit scores, the methods by which these scores have been implemented remain questionable. As demonstrated above, higher credit score groups were more likely to include other riskier attributes on the origination than their immediately lower ones. This trade-off grew stronger for later cohorts especially among low levels of FICO. Naturally, for originations at low levels of FICO, we witness deterioration for later cohorts in terms of the unconditional metric of FICO performance. However, controlling for such riskier attributes implies that FICO performance actually improved for later cohorts even for originations with low FICO scores. In contrast, the trade-off between riskier attributes and higher credit scores was less pronounced at high FICO score levels. Accordingly, our results show improvement in FICO performance in terms of our nonparametric as well as parametric measures.

6. Conclusion

Anecdotal evidence on credit scoring has pointed to possible manipulation that may increase the credit scores of borrowers without any real improvements in their creditworthiness (see, for example, Foust and Pressman, 2008). In theory, score manipulation has minimum impact in terms of our metric if its occurrence were to be uniformly distributed. However, this is unlikely: A more probable scenario is one in which manipulation is more likely to occur at low levels of credit scores. Moreover, most anecdotal accounts argue that such manipulation increased over time. Therefore, if credit score manipulation affects default rates, it is most likely to be reflected in our results at low levels of FICO and for later cohorts.

More important, evidence of manipulation of credit scores should be reflected in anomalous behavior in terms of our parametric measure—a measure that controls for other characteristics on the origination. However, the evidence shows the opposite: parametric measures of FICO performance show improvement at all levels of FICO. This result is fairly robust and holds true for multiple variations of credit score groupings. In light of this, we conclude that the evidence from our data does not reflect any anomalous behavior that would suggest that such manipulation was widespread. That is not to say that such instances of manipulation did not occur, but simply that given our large sample size, score manipulation would have to be fairly widespread to affect our results.

We embarked on this study with a view to determine the efficacy and usage of credit scoring among some of the riskiest loans in recent history. To this end, this paper has introduced a simple yet effective measure for evaluating the performance of credit scoring. As mentioned earlier, the advantage of using such a measure is twofold. First, it lends itself to both non-parametric and parametric measurement. Second, it minimizes the impact of situational factors. Using this measure, we find that credit score performance is robust to both high- and low-default environments. However, evidence suggests that some of the increase in credit scores over the cohorts can be explained as adjustment for the increased riskiness in other attributes on the originations. This was particularly true for low levels of credit scores—resulting in a sharp deterioration of credit score performance in terms of our nonparametric measure. Significantly, once we control for other (riskier) attributes in the origination, our parametric measure of credit score performance shows improvement over the cohorts. This would suggest an increased emphasis on credit scoring—not only as a measure of credit risk but to offset risk on other origination attributes. In part, this emphasis led to deterioration in loan performance even though average credit quality—as measured in terms of credit scores—actually improved over time.

References

Altman, Edward I. and Anthony Saunders (1997). Credit Risk Measurement: Developments over the last 20 Years. Journal of Banking & Finance, 21(11-12), pp. 1721-1742

Anderson, Gordon (1996). Nonparametric Tests of Stochastic Dominance in Income Distributions. Econometrica, 64(5), pp. 1183-93.

Ashcraft Adam B. and Til Schuermann, "Understanding the Securitization of Subprime Mortgage Credit," Foundations and Trends in Finance 2, no. 3 (July 2008): 191-309.

Avery, Robert B.; Raphael W. Bostic; Paul S. Calem; and Glenn B. Canner (2000). Credit Scoring: Statistical Issues and Evidence from Credit-Bureau Files. Real Estate Economics, 28(3), pp. 523-547.

Avery, Robert B.; Paul S. Calem; and Glenn B. Canner (2003). An Overview of Consumer Data and Credit Reporting, Board of Governors Federal Reserve Bulletin, February, pp. 47-73.

Berger, Allen N & Frame, W Scott & Miller, Nathan H, 2005. "Credit Scoring and the Availability, Price, and Risk of Small Business Credit," Journal of Money, Credit and Banking, vol. 37(2), pages 191-222, April.

Bhardwaj G. and R. Sengupta, (2012). Subprime mortgage design, Journal of Banking & Finance, Volume 36, Issue 5, May, Pages 1503-1519.

Bhardwaj G. and R. Sengupta, (2014). "Subprime cohorts and loan performance," Journal of Banking & Finance, Volume 41, Issue 3, March, pages 236-252.

Bhattacharya; A., W. Berliner; and J. Lieber, (2006) "Alt-A Mortgages and MBS." In: Frank Fabozzi, Ed. The Handbook of Mortgage-Backed Securities. McGraw-Hill: New York. pp. 187-206.

Bolton, P. and H. Rosenthal, eds., (2005) Credit Markets for the Poor, New York: Russell Sage Foundation.

Brooks, Rick and Ruth Simon "Subprime Debacle Traps Even Very Credit-Worthy. As Housing Boomed, Industry Pushed Loans to a Broader Market". Wall Street Journal, Leader, December 3, 2007.

Brown, Martin; Tullio Jappelli; and Marco Pagano (2009). Information Sharing and Credit: Firm-Level Evidence from Transition Countries. Journal of Financial Intermediation, 18(2), pp. 151-172.

Demyanyk, Y. (2008) "Did credit scores predict the subprime crisis?," The Regional Economist, Federal Reserve Bank of St. Louis, issue Oct, pages 12-13.

Fair Isaac Corporation (2007). Understanding your FICO Score. UYFS Booklet 100807.

Fishelson-Holstein, Hollis, (2005) "Credit Scoring Role in Increasing Homeownership for Underserved Populations," in Building Assets, Building Credit: Creating Wealth in Low-Income Communities, Nicolas P. Retsinas and Eric S. Belsky, eds. (Washington, DC: Brookings Institution Press, 2005).

Foote, Christopher L.; Kristopher Gerardi; Lorenz Goette; and Paul S. Willen (2008). Just the Facts: An Initial Analysis of Subprime's Role in the Housing Crisis. Journal of Housing Economics, 17(4), pp. 291-305.

Foust, Dean and Aaron Pressman Credit Scores: Not-So-Magic Numbers. BusinessWeek, February 7, 2008.

Gary B. Gorton, 2008. "The Panic of 2007," NBER Working Papers 14358, National Bureau of Economic Research, Inc.

Government Accountability Office (2010) Nonprime Mortgages: Data on Loan Performance by Cohort Year, Product Type, and Location GAO-10-805, August 24, 2010. http://www.gao.gov/products/GAO-10-805

Kaplan, E. L. and Paul Meier (1958). Nonparametric Estimation From Incomplete Observations. Journal of the American Statistical Association, 53, pp. 457-481.

Keys, B., Mukherjee, T., Seru, A. and Vig, V., (2010), Did Securitization Lead to Lax Screening? Evidence from Subprime Loans. Quarterly Journal of Economics, 125, pp. 307--62.

Lee, Donghoon and Wilbert Van der Klaauw (2010). An Introduction to the FRBNY Consumer Credit Panel FRB of New York Staff Report No. 479.

Levitin, Adam J. and Susan M. Watcher (2010). Explaining the Housing Bubble. Georgetown University Law Center Working Paper No. 1669401.

Mayer, Christopher, Karen Pence; and Shane M. Sherlund (2009). The Rise in Mortgage Defaults. Journal of Economic Perspectives, 23(1), pp.27-50.

Mester, Loretta J. (1997). What's the Point of Credit Scoring? Federal Reserve Bank of Philadelphia Business Review, pp. 3-16.

Pennington-Cross, Anthony (2003). Credit History and the Performance of Prime and Nonprime Mortgages. The Journal of Real Estate Finance and Economics, 27(3), pp. 279-301.

Pennington-Cross, A. and Ho, G. (2010), The Termination of Subprime Hybrid and Fixed-Rate Mortgages, Real Estate Economics, American Real Estate and Urban Economics Association, vol. 38(3), pages 399-426.

Petersen, M.A. (2004) Information: Hard and Soft.Working paper,Northwestern University.

Rajan U., A. Seru and V. Vig, (2011). "The Failure of Models that Predict Failure: Distance, Incentives and Defaults." Chicago GSB Research Paper No. 08-19

Rothschild, Michael and Joseph E. Stiglitz (1970). Increasing Risk: I. A Definition. Journal of Economic Theory, 2(3), pp. 225-243.

Sengupta, R. (2010). "Alt-A: the forgotten segment of the mortgage market," Review, Federal Reserve Bank of St. Louis, issue Jan, pages 55-72.

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

Zimmerman, T. (2006) "Defining Nonagency MBS." In: Fabozzi, F. (Ed.), The Handbook of Mortgage-Backed Securities. McGraw-Hill: New York.

APPENDIX A:

A Formal Description of the FICO Performance Measures

Nonparametric Estimates

Formally, the *survival* probability of a 90-day delinquency event beyond loan age t is given by $S(t) = P(T > t)$, where T denotes the duration in months from the month of origination. Let $t_1 < t_2 < \cdots < t_m$ denote the observed age in months at the time of event in a sample size of Noriginations, $\geq m$. Also, let n_l be the number of surviving mortgages just prior to month t_l. A surviving mortgage is defined as one that has neither defaulted nor been paid-off prior to age t_l. If we define d_l as the number of mortgages that default at age t_l, then the Kaplan-Meier estimator of the survivor function is

$$\hat{S}_{NP}(t) = P(T > t) = \prod_{l|t_l \leq t}(1 - \frac{d_l}{n_l}) = \prod_{l|t_l \leq t}(1 - \widehat{\lambda}_l)$$

where, $\widehat{\lambda}_l = \frac{d_l}{n_l}$ is the nonparametric hazard estimate.

For ease of exposition, we split our sample into originations belonging to mutually exclusive and exhaustive FICO score groups, $\mathcal{F}_1, \mathcal{F}_2, \dots, \mathcal{F}_K$. To do this we define K FICO group dummies, $f_1, f_2, \dots, f_K$ such that $f_k = 1$ if FICO score the origination lies in the interval $\mathcal{F}_k$, $k = 1, \dots, K$ and zero otherwise.

Our measure of origination credit score performance $\hat{Q}^c_{k,k-1}$ is the difference in the survival probabilities of a given origination FICO score group to that of its immediately lower FICO score group of the same cohort, c , where $c = 2001, 2002, \dots, 2006$. For the nonparametric measure, we use the following

$$\hat{Q}^{c,NP}_{k,k-1}(t) = \hat{S}^c_{NP}(t, f_k = 1\ ) - \hat{S}^c_{NP}(t, f_{k-1} = 1\ )$$

where $\hat{S}^c_{NP}(t, f_k = 1\ )$ is the nonparametric estimate of the survivor function, $\hat{S}_{NP}(t)$, for cohort, c and FICO score group $\mathcal{F}_k$.

Parametric Estimates

We use the competing risk framework with proportional hazards to study the determinants of default and prepayment (Pennington-Cross and Ho, 2010). Default and prepayment are modeled as competing risks.[31] To formalize our argument, we split borrower repayment behavior into three possible outcomes: (i) the borrower defaults on the loan, (ii) the borrower prepays, and (iii) the loan is current or even 30-day or 60-day delinquent. We denote the exit routes by event j, where the two exit events are given by subscript $j = 1, 2$. Let T_{ij} be the age (in months) at which borrower i chooses event j.

The loan performance of borrower i is observed for $\min(T_{ij}\ )$ and the hazard function, $h_{ij}(t)$, specifying the instantaneous probability of occurrence of event $j(= 1, 2)$ for mortgage i, is given as

$$h_{ij}(t) = \lim_{\Delta t \to 0} \frac{\Pr(t \leq T_{ij} < t + \Delta t | T_{ij} \geq t)}{\Delta t} \quad (1)$$

[31]A mortgage is considered to be in default if it records a 90-day delinquency event (Cowan and Cowan, 2004). As mentioned earlier, prepayments include mortgages that are paid off either because the property is sold off and loan repaid or because the existing mortgage is refinanced.

Following Cox (1972), the semi-parametric representation that we estimate takes the form

$$h_{ij}(t) = h_{0j}(t) \exp(X_i(t)\beta_j) \tag{2}$$

where $h_{0j}(t)$ is the cumulative baseline hazard rate for event $j(= 1, 2)$ and $X_i(t)$ is the vector of covariates on mortgage i which includes both origination characteristics (such as FICO, CLTV, loan documentation, etc.) and time varying economic variables (such as house price volatility, PVAR, and negative equity) as described above.

Under the assumption of independence of the two outcomes (Pennington-Cross and Ho, 2010), the likelihood function for the competing risk hazard model is

$$\mathrm{L}(\beta_1, \beta_2) = \prod_{j=1}^{2} \prod_{i=1}^{n} \frac{\exp\left(X_i(t)\beta_j\right)}{\sum_{l \in R(t_j)} exp\left(X_l(t)\beta_j\right)}$$

where $R(t_j)$ is the risk set for event j at age t.

The estimated hazard ratio (HR) for marginal change in risk characteristic x_i is $\widehat{HR}(t, x_i + \Delta x_i) = \exp(\Delta x_i \widehat{\beta_i})$, whereas the estimated hazard ratio for a given FICO score group, say $\mathcal{F}_k$, is given by

$$\widehat{HR}(t, f_k = 1) = \exp\left(\hat{\beta}_{f_k=1}\right),$$

where $\hat{\beta}_{f_k=1}$ is the coefficient of the regression for the FICO score group $\mathcal{F}_k$ (or the FICO score dummy, $f_k = 1$).

For FICO score group k, the instantaneous probability of delinquency at age t is

$$\hat{h}(t, f_k = 1) = \hat{h}(t, f_1 = 1) \times \widehat{HR}(t, f_k = 1)$$

where $\hat{h}(t, f_k = 1)$ is a parametric estimate of the hazard rate at age t for the FICO group $\mathcal{F}_k$. We replace $\hat{h}(t, f_1 = 1)$ with its nonparametric equivalent, $\widehat{\lambda_j}(t, f_1 = 1)$, given as $\widehat{\lambda_j}(t, f_1 = 1) = \frac{d(t, f_1=1)}{n(t, f_1=1)}$, where $d(t, f_1 = 1)$ is the number of delinquencies at age t with FICO scores in the interval $\mathcal{F}_1$ and $n(t, f_1 = 1)$ is the number of number of surviving mortgages (not in default or prepaid) at age t with FICO scores in the interval $\mathcal{F}_1$. Finally, we obtain an estimate of $\hat{h}(t, f_k = 1)$. Accordingly, the parametric estimates of the survivor function is calculated as

$$\hat{S}_P^c(t, f_k = 1\,) = \prod_{j | t_j \leq t} (1 - \widehat{h_j}(t, f_k = 1))$$

Just as in the non-parametric case, we obtain estimates of the parametric measure of credit score performance as $\hat{Q}_{k,k-1}^{c,P}(t)$.

<u>Table 1: Credit Score (FICO) Regression</u>

OLS estimates with borrower FICO score as the left-hand side variable and other borrower characteristics as regressors. In addition to the variables shown here, 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. Jumbo is a dummy variable that equals one if the value of the property exceeds the conforming limit for that year and zero otherwise.

Panel A: By Cohort

	Cohort						
	2000	2001	2002	2003	2004	2005	2006
Intercept	619.84***	617.46***	641.96***	648.6***	597.66***	607.79***	616.86***
CLTV	0.32***	0.51***	0.71***	0.76***	0.92***	0.92***	0.89***
Full-Doc	-15.01***	-20.11***	-23.89***	-21.39***	-19.02***	-20.08***	-20.09***
Owner Occupied	-26.22***	-24.65***	-29.05***	-34.25***	-35.64***	-34.21***	-33.46***
Second Home	-0.65	-0.81	-5.86***	-9.07***	-9.34***	-5.33***	-5.62***
Refinance (Cash Out)	-13.81***	-12.49***	-21.89***	-25.47***	-25.18***	-21.2***	-18.57***
Refinance (No Cash Out)	-16.44***	-14.4***	-15.54***	-14.66***	-12.6***	-9.46***	-8.15***
Jumbo	9.31***	13.78***	14.01***	14.66***	12.04***	10.16***	11.48***
Adjusted R-Squared	0.0773	0.0957	0.1527	0.1765	0.2025	0.2077	0.2161

Panel B: For all Cohorts (full sample)

	(1)	(2)
Intercept	604.35***	607.12***
Dummy for 2003-2006 cohorts	9.37***	
Dummy for 2004-2006 cohorts		3.45***
CLTV	0.83***	0.86***
Full-Doc	-20.37***	-20.65***
Owner Occupied	-33.26***	-33.36***
Second Home	-6.18***	-6.14***
Refinance (Cash Out)	-21.16***	-20.92***
Refinance (No Cash Out)	-12.14***	-12.18***
Jumbo	11.86***	11.98***
Adjusted R-Squared	0.188	0.1853

Table 2.1: Increase in (Non-parametric) Survival Probabilities for Improvements in FICO score (groups)—All Subprime Loans

The numbers show a percentage point increase in the Kaplan-Meier survival probabilities (for the first two years after origination) of originations in the higher FICO score group relative to those in the lower FICO score group.

Panel A.
The FICO score groups used below are “less than 540", “540-579", “580-619" … “700-739" and “greater than or equal to 740".

	Cohort						
Improvement in FICO score	2000	2001	2002	2003	2004	2005	2006
[< 540] to [540 – 579]	8.17	7.46	5.75	4.17	4.73	4.95	5.52
[540 – 579] to [580 – 619]	4.45	4.24	3.57	3.38	3.88	3.04	1.68
[580 – 619] to [620 – 659]	3.35	2.87	2.91	3.24	4.48	4.33	2.10
[620 – 659] to [660 – 699]	1.95	2.37	2.54	2.43	2.79	4.59	4.64
[660 – 699] to [700 – 739]	1.41	1.44	1.96	1.52	1.50	2.56	4.14
[700 – 739] to [≥740]	0.91	1.10	0.81	0.84	1.30	2.57	7.84
Average All	3.37	3.25	2.92	2.60	3.11	3.68	4.32

Panel B.
The FICO score groups used below are "less than 540", "540-559", "560-579" … "720-739" and "greater than or equal to 740".

	Cohort						
Improvement in FICO score	2000	2001	2002	2003	2004	2005	2006
[< 540] to [540 - 559]	6.68	6.16	4.53	3.10	3.47	3.70	3.90
[540 - 559] to [560 - 579]	2.82	2.50	2.36	1.97	2.30	2.02	2.76
[560 - 579] to [580 - 599]	2.16	2.31	1.80	1.55	1.55	0.80	-0.27
[580 - 599] to [600 - 619]	1.78	1.22	1.10	1.56	2.29	2.73	1.40
[600 - 619] to [620 - 639]	1.94	1.67	1.69	1.73	2.62	1.83	0.84
[620 - 639] to [640 - 659]	1.16	1.52	1.62	1.83	1.78	2.78	1.53
[640 - 659] to [660 - 679]	1.07	1.16	1.30	1.10	1.38	2.23	3.01
[660 - 679] to [680 - 699]	0.58	0.78	0.90	0.84	1.05	1.94	1.99
[680 - 699] to [700 - 719]	0.94	0.89	1.19	0.67	0.62	0.94	2.17
[700 - 719] to [720 - 739]	0.00	0.25	0.41	0.71	0.60	1.19	1.91
[720 - 739] to [≥740]	1.06	0.90	0.56	0.47	0.93	1.79	6.71
Average All	1.83	1.76	1.59	1.41	1.69	2.00	2.36

Table 2.2: Increase in (Non-parametric) Survival Probabilities for Improvements in FICO score (groups)—Subprime ARM Loans

The numbers show a percentage point increase in the Kaplan-Meier survival probabilities (for the first two years after origination) of originations in the higher FICO score group relative to those in the lower FICO score group.

Panel A.

The FICO score groups used below are “less than 540", “540-579", “580-619" … “700-739" and “greater than or equal to 740".

	Cohort						
Improvement in FICO score	2000	2001	2002	2003	2004	2005	2006
[< 540] to [540 – 579]	9.14	8.58	6.75	4.44	4.23	4.36	4.50
[540 – 579] to [580 – 619]	5.82	5.08	4.36	3.89	3.50	2.50	0.59
[580 – 619] to [620 – 659]	3.71	3.29	3.32	3.79	4.39	4.03	0.41
[620 – 659] to [660 – 699]	2.07	2.38	2.85	2.50	2.87	4.69	3.70
[660 – 699] to [700 – 739]	1.17	1.65	2.01	1.57	1.44	2.59	3.02
[700 – 739] to [≥740]	0.42	0.44	0.70	0.60	0.88	2.22	5.51
Average All	3.72	3.57	3.33	2.80	2.88	3.40	2.96

Panel B.

The FICO score groups used below are "less than 540", "540-559", "560-579" … "720-739" and "greater than or equal to 740".

	Cohort						
Improvement in FICO score	2000	2001	2002	2003	2004	2005	2006
[< 540] to [540 - 559]	7.83	7.07	5.32	3.78	3.16	3.21	3.22
[540 - 559] to [560 - 579]	2.55	2.91	2.69	1.22	2.01	2.09	2.31
[560 - 579] to [580 - 599]	3.50	2.65	2.23	2.26	1.23	0.11	-0.72
[580 - 599] to [600 - 619]	2.20	2.01	1.63	1.91	2.42	2.70	0.51
[600 - 619] to [620 - 639]	1.97	1.75	1.85	1.99	2.46	1.42	-0.29
[620 - 639] to [640 - 659]	1.44	1.35	1.67	2.11	1.78	2.92	1.02
[640 - 659] to [660 - 679]	1.40	1.31	1.60	1.02	1.45	2.26	2.68
[660 - 679] to [680 - 699]	-0.39	0.79	0.79	0.82	1.19	2.14	1.27
[680 - 699] to [700 - 719]	1.38	0.86	1.52	0.85	0.52	0.84	1.66
[700 - 719] to [720 - 739]	0.07	0.79	0.02	0.56	0.53	1.24	1.49
[720 - 739] to [≥740]	0.38	-0.06	0.69	0.26	0.56	1.47	4.59
Average All	2.03	1.95	1.82	1.53	1.57	1.85	1.61

Table 2.3: Increase in (Non-parametric) Survival Probabilities for Improvements in FICO score (groups)—Subprime FRM loans

The numbers show a percentage point increase in the Kaplan-Meier survival probabilities (for the first two years after origination) of originations in the higher FICO score group relative to those in the lower FICO score group.

Panel A.

The FICO score groups used below are “less than 540", “540-579", “580-619" … “700-739" and “greater than or equal to 740".

	Cohort						
Improvement in FICO score	2000	2001	2002	2003	2004	2005	2006
[< 540] to [540 – 579]	7.84	7.20	5.57	4.07	5.01	5.81	7.76
[540 – 579] to [580 – 619]	4.32	3.83	3.30	3.22	3.69	3.96	5.08
[580 – 619] to [620 – 659]	3.00	2.65	2.59	2.93	3.21	3.55	5.39
[620 – 659] to [660 – 699]	1.68	2.20	2.49	2.36	1.69	2.96	4.94
[660 – 699] to [700 – 739]	1.43	1.45	1.86	1.45	0.98	1.58	3.25
[700 – 739] to [≥740]	0.89	0.98	0.73	0.85	0.84	1.37	4.04
Average All	3.20	3.05	2.76	2.48	2.57	3.21	5.08

Panel B.

The FICO score groups used below are "less than 540", "540-559", "560-579" … "720-739" and "greater than or equal to 740".

	Cohort						
Improvement in FICO score	2000	2001	2002	2003	2004	2005	2006
[< 540] to [540 - 559]	6.49	5.83	4.45	3.04	3.49	4.37	5.40
[540 - 559] to [560 - 579]	2.57	2.61	2.11	1.91	2.71	2.54	4.15
[560 - 579] to [580 - 599]	2.34	2.08	1.80	1.49	1.42	1.64	0.97
[580 - 599] to [600 - 619]	1.53	0.99	0.93	1.51	1.89	2.17	4.17
[600 - 619] to [620 - 639]	1.83	1.55	1.39	1.45	1.75	1.58	2.01
[620 - 639] to [640 - 659]	0.93	1.43	1.70	1.73	1.33	2.11	3.22
[640 - 659] to [660 - 679]	0.77	1.13	1.14	1.10	0.85	1.42	2.45
[660 - 679] to [680 - 699]	1.01	0.72	1.08	0.85	0.38	1.04	2.03
[680 - 699] to [700 - 719]	0.85	0.87	1.03	0.73	0.67	0.73	1.28
[700 - 719] to [720 - 739]	-0.10	0.34	0.49	0.54	0.22	0.65	1.89
[720 - 739] to [≥740]	0.95	0.77	0.43	0.53	0.72	0.99	2.90
Average All	1.74	1.66	1.50	1.35	1.40	1.75	2.77

Table 3.1 : Estimated Cox proportional hazard rate regression (All Subprime Loans)

Table reports the hazard ratio estimates from a competing risk hazard model with loan default and prepayment as competing risks. In addition to the variables shown here, 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 and a dummy for mortgages with prepayment penalties. *Full Documentation* is a dummy variable equal to 1 if the mortgage has full documentation and zero otherwise. *Fees and Points* are the fees and discount points charged by the lender at settlement on a 30-yr FRM prime mortgage, taken from the Freddie Mac PMMS Survey. *Present Value Annualized Ratio (PVAR)* measures the ratio of the present value of the payments on mortgage principal outstanding using the existing mortgage rate to that using the current rate available on refinance. *Interest Volatility* is the standard deviation of the six-month LIBOR for the previous 24 months. *House Price Volatility* is the standard deviation of house price change in the MSA-level FHFA house price index for the previous 24 months. *PosUnempG* is a dummy variable equal to one if the MSA records and increase in the unemployment rate over the previous year and zero otherwise. *Negative Equity* is a dummy variable equal to 1 if the current LTV exceeds 100% and zero otherwise.

Panel A: Default hazard

Cohort	2000	2001	2002	2003	2004	2005	2006
FICO: 540-579	0.789***	0.794***	0.774***	0.788***	0.739***	0.762***	0.776***
FICO: 580-619	0.701***	0.644***	0.62***	0.604***	0.565***	0.618***	0.644***
FICO: 620-659	0.592***	0.538***	0.507***	0.461***	0.418***	0.49***	0.532***
FICO: 660-699	0.453***	0.408***	0.37***	0.335***	0.315***	0.385***	0.446***
FICO: 700-739	0.356***	0.285***	0.255***	0.245***	0.243***	0.316***	0.384***
FICO: >=740	0.29***	0.201***	0.212***	0.196***	0.194***	0.252***	0.328***
CLTV	1.008***	1.012***	1.013***	1.019***	1.017***	1.027***	1.04***
Full Documentation	0.869***	0.813***	0.807***	0.723***	0.77***	0.699***	0.68***
Fees and Points	1.149***	1.094***	1.044***	1.069***	1.14***	1.053***	0.959***
PVAR	1.001***	1.001***	1.02***	1.033***	1.09***	1.124***	1.058***
Interest Volatility	1.015***	1.022***	1.017***	0.996***	1.02***	1.103***	1.055***
House Price Volatility	0.644***	0.579***	0.617***	0.735***	0.779***	0.79***	0.697***
PosUnempG dummy	1.203***	1.655***	1.523***	1.278***	1.161***	1.25***	1.099***
Negative Equity Dummy	1.046**	1.017**	1.153***	1.141***	1.117***	1.029**	1.115***
LR test	136628	156116	199491	265265	379406	443464	126835
(p-value)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)

Panel B: Prepayment hazard

Cohort	2000	2001	2002	2003	2004	2005	2006
FICO: 540-579	1.291***	1.249***	1.136***	1.114***	1.043***	1.057***	1.057***
FICO: 580-619	1.512***	1.367***	1.206***	1.198***	1.085***	1.081***	1.062***
FICO: 620-659	1.657***	1.438***	1.282***	1.265***	1.148***	1.151***	1.075***
FICO: 660-699	1.76***	1.507***	1.335***	1.307***	1.189***	1.257***	1.207***
FICO: 700-739	1.823***	1.551***	1.408***	1.321***	1.234***	1.348***	1.288***
FICO: >=740	1.996***	1.501***	1.401***	1.329***	1.26***	1.438***	1.461***
CLTV	0.994***	0.995***	0.997***	0.999***	0.996***	0.992***	0.991***
Full Documentation	1.068***	1.057***	1.039***	1.03***	1.037***	1.087***	1.068***
Fees and Points	1.037***	1.025***	1.011***	1.027***	1.04***	0.996***	0.978***
PVAR	1.001***	1.002***	1.015***	1.025***	1.061***	1.065***	1.043***
Interest Volatility	1.007***	1.01***	1.003***	0.994***	1.004***	1.019***	1.02***
House Price Volatility	1.001***	1.055***	1.087***	1.12***	1.162***	1.155***	1.199***
PosUnempG dummy	0.994**	0.943***	1.046***	0.936***	0.912***	0.904***	1.089***
Negative Equity Dummy	0.509***	0.505***	0.514***	0.667***	0.346***	0.586***	0.54***
LR test	90573	124004	242996	685001	1198543	833056	125112
(p-value)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)

Table 3.2 : Estimated Cox proportional hazard rate regression (Subprime ARMs)

Table reports the hazard ratio estimates from a competing risk hazard model with loan default and prepayment as competing risks. In addition to the variables shown here, 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 and a dummy for mortgages with prepayment penalties. *Full Documentation* is a dummy variable equal to 1 if the mortgage has full documentation and zero otherwise. *Fees and Points* are the fees and discount points charged by the lender at settlement on a 30-yr FRM prime mortgage, taken from the Freddie Mac PMMS Survey. *Present Value Annualized Ratio (PVAR)* measures the ratio of the present value of the payments on mortgage principal outstanding using the existing mortgage rate to that using the current rate available on refinance. *Interest Volatility* is the standard deviation of the six-month LIBOR for the previous 24 months. *House Price Volatility* is the standard deviation of house price change in the MSA-level FHFA house price index for the previous 24 months. *PosUnempG* is a dummy variable equal to one if the MSA records and increase in the unemployment rate over the previous year and zero otherwise. *Negative Equity* is a dummy variable equal to 1 if the current LTV exceeds 100% and zero otherwise.

Panel A: Default hazard

Cohort	2000	2001	2002	2003	2004	2005	2006
FICO: 540-579	0.869***	0.846***	0.839***	0.854***	0.744***	0.764***	0.783***
FICO: 580-619	0.778***	0.703***	0.707***	0.718***	0.575***	0.626***	0.666***
FICO: 620-659	0.687***	0.599***	0.611***	0.562***	0.426***	0.499***	0.562***
FICO: 660-699	0.571***	0.454***	0.467***	0.435***	0.319***	0.393***	0.474***
FICO: 700-739	0.478***	0.332***	0.351***	0.336***	0.252***	0.324***	0.414***
FICO: >=740	0.401***	0.249***	0.282***	0.268***	0.205***	0.26***	0.355***
CLTV	1.005***	1.008***	1.008***	1.014***	1.016***	1.027***	1.039***
Full Documentation	0.933***	0.843***	0.887***	0.782***	0.778***	0.703***	0.69***
Fees and Points	1.148***	1.089***	1.042***	1.084***	1.141***	1.054***	0.959***
PVAR	1.002***	1.002***	1.024***	1.044***	1.092***	1.124***	1.059***
Interest Volatility	1.018***	1.024***	1.012***	0.995***	1.019***	1.103***	1.054***
House Price Volatility	0.541***	0.599***	0.637***	0.758***	0.781***	0.791***	0.701***
PosUnempG dummy	1.2***	1.51***	1.494***	1.163***	1.188***	1.272***	1.108***
Negative Equity Dummy	1.289**	1.212***	1.216***	1.392***	1.133***	1.031**	1.106***
LR test	94139	89060	128690	160883	326689	405776	111593
(p-value)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)

Panel B: Prepayment hazard

Cohort	2000	2001	2002	2003	2004	2005	2006
FICO: 540-579	1.171***	1.138***	1.13***	1.101***	1.046***	1.055***	1.067***
FICO: 580-619	1.278***	1.19***	1.214***	1.173***	1.093***	1.078***	1.071***
FICO: 620-659	1.366***	1.266***	1.321***	1.282***	1.166***	1.149***	1.082***
FICO: 660-699	1.396***	1.337***	1.452***	1.392***	1.212***	1.26***	1.226***
FICO: 700-739	1.376***	1.435***	1.568***	1.479***	1.256***	1.353***	1.303***
FICO: >=740	1.528***	1.508***	1.638***	1.564***	1.28***	1.448***	1.516***
CLTV	0.996***	0.999***	0.999**	1.000**	0.995***	0.992***	0.99***
Full Documentation	1.072***	1.053***	1.068***	1.114***	1.043***	1.088***	1.071***
Fees and Points	0.996***	1.004***	1.036***	1.025***	1.039***	0.995***	0.979***
PVAR	1.006***	1.004***	0.996***	0.996***	1.06***	1.063***	1.043***
Interest Volatility	1.014***	0.99***	0.985***	1.009***	1.004***	1.019***	1.02***
House Price Volatility	1.284***	1.102***	1.155***	1.155***	1.162***	1.156***	1.209***
PosUnempG dummy	0.805***	0.849***	0.883***	0.834***	0.912***	0.902***	1.075***
Negative Equity Dummy	0.351***	0.357***	0.282***	0.371***	0.361***	0.401***	0.406***
LR test	64784	55707	94336	119860	992536	735882	109483
(p-value)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)

Table 3.3 : Estimated Cox proportional hazard rate regression (Subprime FRMs)

Table reports the hazard ratio estimates from a competing risk hazard model with loan default and prepayment as competing risks. In addition to the variables shown here, 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 and a dummy for mortgages with prepayment penalties. *Full Documentation* is a dummy variable equal to 1 if the mortgage has full documentation and zero otherwise. *Fees and Points* are the fees and discount points charged by the lender at settlement on a 30-yr FRM prime mortgage, taken from the Freddie Mac PMMS Survey. *Present Value Annualized Ratio (PVAR)* measures the ratio of the present value of the payments on mortgage principal outstanding using the existing mortgage rate to that using the current rate available on refinance. *Interest Volatility* is the standard deviation of the six-month LIBOR for the previous 24 months. *House Price Volatility* is the standard deviation of house price change in the MSA-level FHFA house price index for the previous 24 months. *PosUnempG* is a dummy variable equal to one if the MSA records and increase in the unemployment rate over the previous year and zero otherwise. *Negative Equity* is a dummy variable equal to 1 if the current LTV exceeds 100% and zero otherwise.

Panel A: Default hazard

Cohort	2000	2001	2002	2003	2004	2005	2006
FICO: 540-579	0.783***	0.778***	0.756***	0.773***	0.742***	0.788***	0.758***
FICO: 580-619	0.668***	0.63***	0.598***	0.581***	0.55***	0.6***	0.571***
FICO: 620-659	0.557***	0.526***	0.484***	0.445***	0.406***	0.464***	0.396***
FICO: 660-699	0.417***	0.398***	0.351***	0.321***	0.314***	0.363***	0.325***
FICO: 700-739	0.324***	0.278***	0.239***	0.233***	0.213***	0.288***	0.24***
FICO: >=740	0.263***	0.194***	0.203***	0.186***	0.16***	0.215***	0.199***
CLTV	1.008***	1.013***	1.014***	1.02***	1.021***	1.026***	1.037***
Full Documentation	0.847***	0.805***	0.787***	0.708***	0.745***	0.717***	0.648***
Fees and Points	1.146***	1.097***	1.044***	1.063***	1.13***	1.05***	0.958***
PVAR	1.001***	1.001***	1.018***	1.031***	1.08***	1.124***	1.056***
Interest Volatility	1.014***	1.022***	1.019***	0.997***	1.022***	1.1***	1.057***
House Price Volatility	0.669***	0.578***	0.618***	0.738***	0.759***	0.78***	0.651***
PosUnempG dummy	1.207***	1.7***	1.55***	1.313***	1.004**	1.047**	1.061**
Negative Equity Dummy	1.263**	1.269**	1.476***	1.362***	1.348**	1.173***	1.474***
LR test	181658	223959	274147	374377	51555	37695	13144
(p-value)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)

Panel B: Prepayment hazard

Cohort	2000	2001	2002	2003	2004	2005	2006
FICO: 540-579	1.288***	1.229***	1.113***	1.091***	1.078***	1.07***	1.008**
FICO: 580-619	1.49***	1.329***	1.155***	1.148***	1.117***	1.113***	1.041**
FICO: 620-659	1.621***	1.374***	1.196***	1.181***	1.156***	1.185***	1.089***
FICO: 660-699	1.707***	1.417***	1.222***	1.202***	1.186***	1.255***	1.157***
FICO: 700-739	1.766***	1.44***	1.27***	1.207***	1.236***	1.332***	1.266***
FICO: >=740	1.926***	1.386***	1.261***	1.21***	1.29***	1.39***	1.328***
CLTV	0.994***	0.995***	0.998***	0.999***	0.997***	0.993***	0.993***
Full Documentation	1.056***	1.046***	1.026***	1.016***	1.032***	1.082***	1.041***
Fees and Points	1.044***	1.028***	1.011***	1.025***	1.049***	1.002***	0.974***
PVAR	1.001***	1.002***	1.015***	1.027***	1.068***	1.077***	1.046***
Interest Volatility	1.007***	1.011***	1.005***	0.993***	1.004***	1.025***	1.027***
House Price Volatility	1.001***	1.053***	1.07***	1.113***	1.161***	1.144***	1.144***
PosUnempG dummy	0.985***	0.938***	1.073***	0.954***	0.902***	0.92***	1.14***
Negative Equity Dummy	0.538***	0.592***	0.615***	0.741***	0.729***	0.739***	0.616***
LR test	166142	230870	400187	1200083	207649	96468	16505
(p-value)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)

Table 4.1: Increase in Parametric Survival Probabilities for Improvements in FICO score (groups)—Subprime All Loans

The numbers show parametric estimates of percentage point increases in estimated survival probabilities (for the first two years after origination) of originations in the higher FICO score group relative to those in the lower FICO score group. The estimated probabilities are calculated using the corresponding competing risk proportional hazard model shown in Tables 3.1.

Panel A.
The FICO score groups used below are “less than 540", “540-579", “580-619" … “700-739" and “greater than or equal to 740".

	Cohort						
Improvement in FICO score	2000	2001	2002	2003	2004	2005	2006
[< 540] to [540 – 579]	4.34	4.05	4.17	3.44	5.18	5.95	8.05
[540 – 579] to [580 – 619]	1.87	3.05	2.96	3.11	3.63	3.84	5.28
[580 – 619] to [620 – 659]	2.39	2.23	2.23	2.48	3.16	3.55	4.76
[620 – 659] to [660 – 699]	3.15	2.82	2.77	2.24	2.31	3.01	3.90
[660 – 699] to [700 – 739]	2.25	2.73	2.40	1.64	1.63	2.04	2.96
[700 – 739] to [≥740]	1.54	1.90	0.91	0.91	1.12	1.96	2.71
Average All	2.59	2.80	2.57	2.30	2.84	3.39	4.61

Panel B.
The FICO score groups used below are "less than 540", "540-559", "560-579" … "720-739" and "greater than or equal to 740".

	Cohort						
Improvement in FICO score	2000	2001	2002	2003	2004	2005	2006
[< 540] to [540 - 559]	3.33	3.22	3.17	2.37	4.09	4.29	5.29
[540 - 559] to [560 - 579]	1.16	1.64	1.92	2.01	2.09	3.02	4.98
[560 - 579] to [580 - 599]	1.35	1.81	1.55	1.33	1.75	1.50	1.84
[580 - 599] to [600 - 619]	0.80	0.89	0.98	1.61	1.72	1.98	2.45
[600 - 619] to [620 - 639]	1.51	1.22	1.11	1.15	1.71	1.83	2.76
[620 - 639] to [640 - 659]	1.10	1.39	1.54	1.42	1.48	1.82	2.14
[640 - 659] to [660 - 679]	2.41	1.58	1.45	1.00	1.03	1.42	1.77
[660 - 679] to [680 - 699]	0.29	1.18	1.15	1.11	1.20	1.52	2.61
[680 - 699] to [700 - 719]	1.49	1.51	1.42	0.73	0.83	0.78	0.97
[700 - 719] to [720 - 739]	1.46	1.24	0.73	0.60	0.23	0.91	0.99
[720 - 739] to [≥740]	0.65	1.15	0.44	0.54	0.97	1.39	2.09
Average All	1.41	1.53	1.41	1.26	1.55	1.86	2.54

Table 4.2: Increase in Parametric Survival Probabilities for Improvements in FICO score (groups)—Subprime ARM Loans

The numbers show parametric estimates of percentage point increases in estimated survival probabilities (for the first two years after origination) of originations in the higher FICO score group relative to those in the lower FICO score group. The estimated probabilities are calculated using the corresponding competing risk proportional hazard model shown in Tables 3.2.

Panel A.
The FICO score groups used below are “less than 540", “540-579", “580-619" … “700-739" and “greater than or equal to 740".

	Cohort						
Improvement in FICO score	2000	2001	2002	2003	2004	2005	2006
[< 540] to [540 – 579]	2.97	3.33	3.32	2.62	5.18	6.07	8.02
[540 – 579] to [580 – 619]	2.13	3.23	2.81	2.52	3.59	3.80	4.77
[580 – 619] to [620 – 659]	2.17	2.41	2.11	2.96	3.28	3.65	4.53
[620 – 659] to [660 – 699]	2.88	3.47	3.26	2.49	2.45	3.15	4.09
[660 – 699] to [700 – 739]	2.35	3.05	2.71	1.98	1.54	2.11	2.92
[700 – 739] to [≥740]	2.02	2.11	1.65	1.38	1.12	2.00	2.98
Average All	2.42	2.93	2.64	2.32	2.86	3.46	4.55

Panel B.
The FICO score groups used below are "less than 540", "540-559", "560-579" … "720-739" and "greater than or equal to 740".

	Cohort						
Improvement in FICO score	2000	2001	2002	2003	2004	2005	2006
[< 540] to [540 - 559]	2.89	2.40	2.49	1.70	4.10	4.35	5.47
[540 - 559] to [560 - 579]	0.18	1.84	1.63	1.75	2.11	3.15	4.60
[560 - 579] to [580 - 599]	1.76	1.82	1.57	0.64	1.65	1.31	1.64
[580 - 599] to [600 - 619]	0.58	1.04	0.96	2.04	1.82	2.16	2.22
[600 - 619] to [620 - 639]	1.73	1.49	1.14	1.29	1.77	1.86	2.56
[620 - 639] to [640 - 659]	0.35	1.02	1.29	1.83	1.52	1.83	2.30
[640 - 659] to [660 - 679]	2.99	2.14	1.95	0.98	1.12	1.51	1.72
[660 - 679] to [680 - 699]	-0.87	1.98	1.48	1.26	1.28	1.62	3.03
[680 - 699] to [700 - 719]	2.61	1.31	1.50	1.05	0.63	0.80	0.90
[700 - 719] to [720 - 739]	0.85	1.37	0.77	0.36	0.39	0.87	0.46
[720 - 739] to [≥740]	1.47	1.22	1.14	1.14	0.87	1.45	2.67
Average All	1.32	1.60	1.45	1.28	1.57	1.90	2.51

Table 4.3: Increase in Parametric Survival Probabilities for Improvements in FICO score (groups)—Subprime FRM Loans

The numbers show parametric estimates of percentage point increases in estimated survival probabilities (for the first two years after origination) of originations in the higher FICO score group relative to those in the lower FICO score group. The estimated probabilities are calculated using the corresponding competing risk proportional hazard model shown in Tables 3.3.

Panel A.
The FICO score groups used below are “less than 540", “540-579", “580-619" … “700-739" and “greater than or equal to 740".

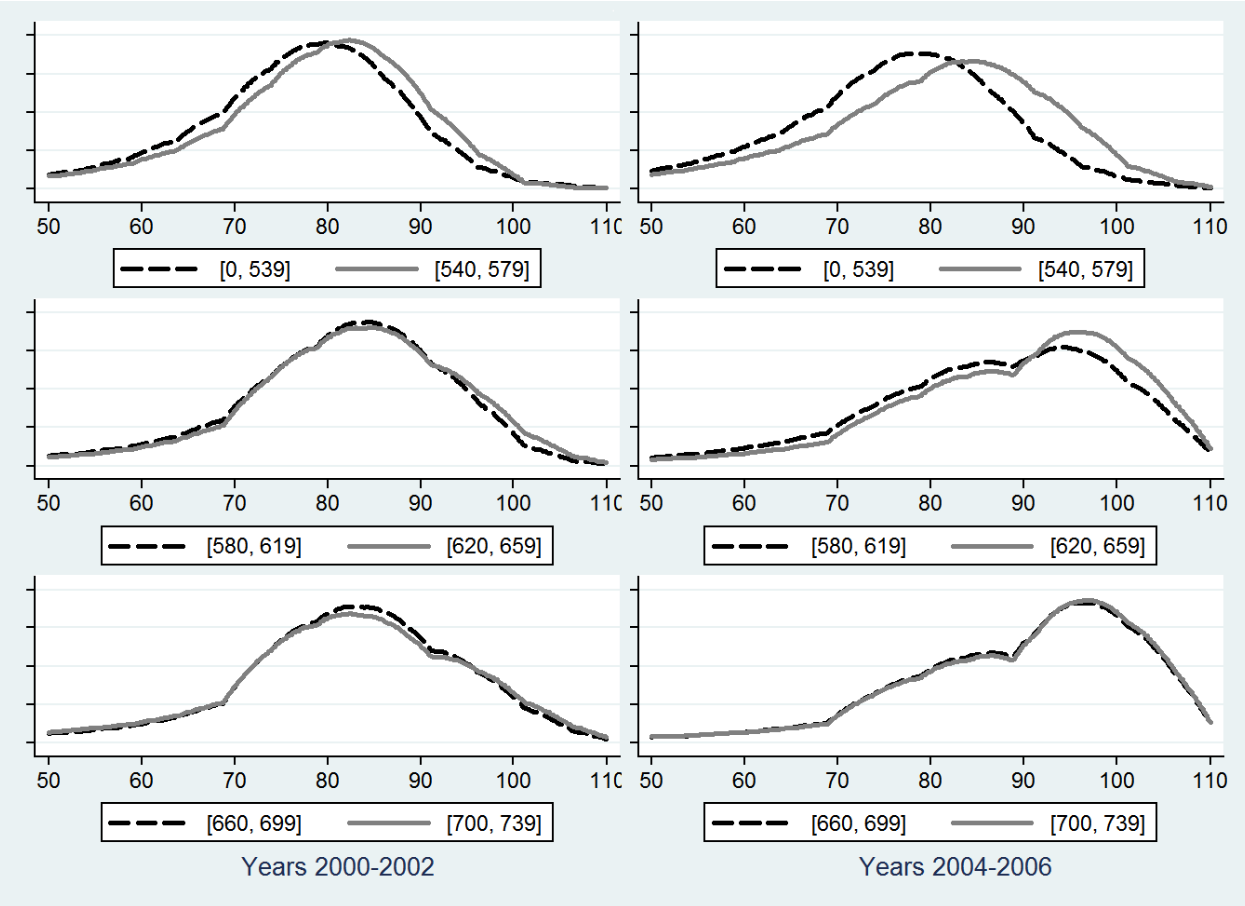
	Cohort						
Improvement in FICO score	2000	2001	2002	2003	2004	2005	2006
[< 540] to [540 – 579]	4.27	4.16	4.31	3.55	4.06	4.06	7.21
[540 – 579] to [580 – 619]	2.36	2.88	2.92	3.12	3.14	3.77	6.09
[580 – 619] to [620 – 659]	2.32	2.09	2.16	2.27	2.44	2.82	6.13
[620 – 659] to [660 – 699]	3.02	2.61	2.58	2.11	1.59	2.16	2.58
[660 – 699] to [700 – 739]	2.08	2.53	2.22	1.53	1.76	1.62	3.22
[700 – 739] to [≥740]	1.36	1.82	0.72	0.83	0.95	1.60	1.63
Average All	2.57	2.68	2.49	2.23	2.32	2.67	4.48

Panel B.
The FICO score groups used below are "less than 540", "540-559", "560-579" … "720-739" and "greater than or equal to 740".

	Cohort						
Improvement in FICO score	2000	2001	2002	2003	2004	2005	2006
[< 540] to [540 - 559]	3.53	3.39	3.31	2.45	3.24	3.23	3.94
[540 - 559] to [560 - 579]	1.48	1.50	1.94	2.08	1.57	1.53	6.00
[560 - 579] to [580 - 599]	1.21	1.75	1.54	1.37	1.75	2.77	1.97
[580 - 599] to [600 - 619]	0.88	0.78	0.92	1.48	1.25	0.62	2.78
[600 - 619] to [620 - 639]	1.36	1.10	1.06	1.06	1.36	1.71	4.16
[620 - 639] to [640 - 659]	1.19	1.42	1.55	1.27	1.13	1.79	1.61
[640 - 659] to [660 - 679]	2.16	1.44	1.30	0.99	0.58	0.83	1.46
[660 - 679] to [680 - 699]	0.48	0.96	1.05	1.05	0.90	0.93	0.68
[680 - 699] to [700 - 719]	1.21	1.45	1.34	0.66	1.34	0.62	1.20
[700 - 719] to [720 - 739]	1.41	1.19	0.65	0.61	-0.19	1.19	3.89
[720 - 739] to [≥740]	0.51	1.10	0.31	0.45	1.05	0.87	-0.70
Average All	1.40	1.46	1.36	1.22	1.27	1.46	2.45

Figure 1

Kernel density plots of the LTV distribution for originations in adjacent FICO score groups. For each plot, the LTV distribution for the higher FICO score group (in gray) is shown alongside the LTV distribution for the lower FICO score group (in black, dashed). In each case, the distribution of LTV is truncated to lie in the interval [50,110] for ease of exposition.

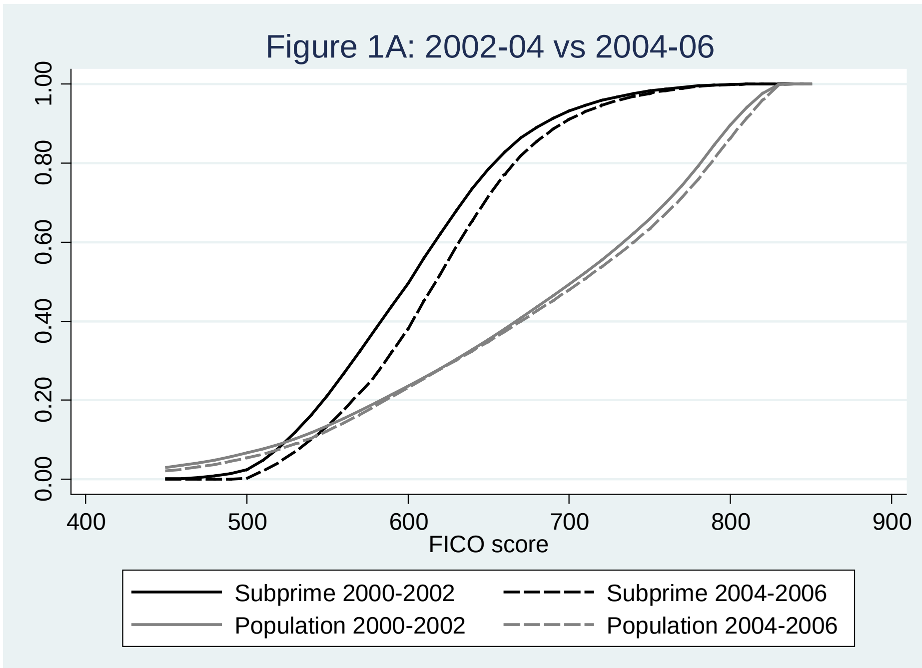


WEB APPENDIX (not meant for publication)

APPENDIX B

B.1 Population and Subprime FICO scores

Figure B. 1: Distribution of Credit Scores for Subprime and U.S. Population (2000-2002 and 2004-2006)
Plot below show the cumulative distribution functions (cdfs) of FICO scores on originations during 2000-2002 (solid lines) and then during 2004-2006 (dashed lines). The black lines show the distribution of FICO scores on subprime originations. The gray lines show the distribution of credit scores for the U.S. population with recorded credit histories.



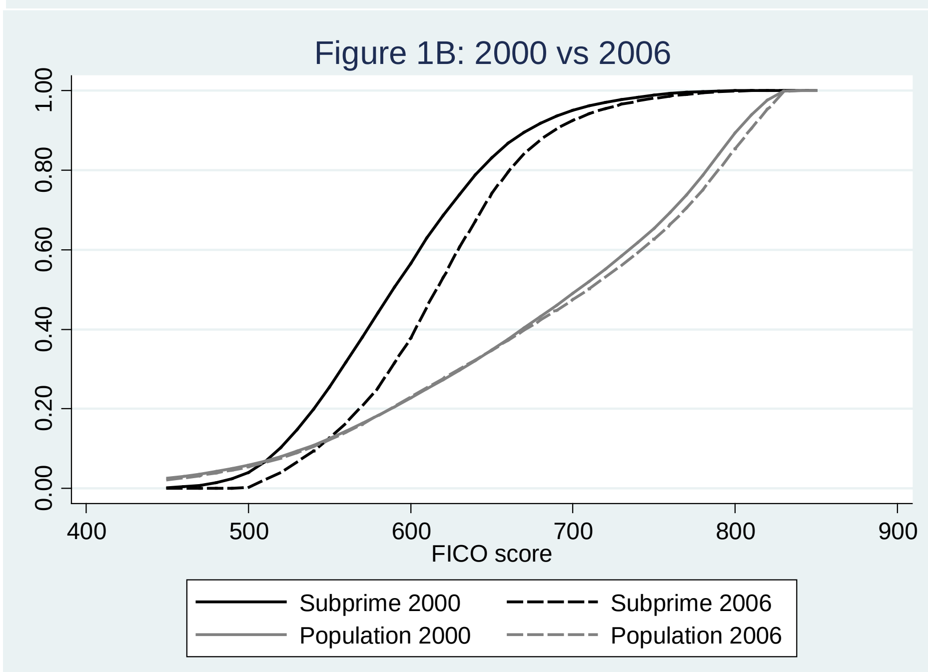


Figure B.1A shows the cumulative distribution function (cdf) of the FICO scores on subprime originations of earlier cohorts (2000-2002, in bold lines) along with those of later cohorts (2004-2006, in dotted lines). The figures show that the probability that a subprime borrower has a lower credit score is significantly higher on originations of earlier cohorts than on originations of later cohorts. Clearly, origination FICO of subprime mortgages improved significantly over time. Using more formal methods below, we verify that the distribution of credit scores on later cohorts first-order stochastically dominates the distribution of credit scores on earlier cohorts (Rothschild and Stiglitz, 1970).

An important concern here is whether the observed improvement in FICO scores on subprime mortgages was a result of the shift in the underlying distribution of FICO scores for the entire U.S. We confirm that changes in the borrower density for the credit-eligible population for the U.S. cannot explain the full improvement in the credit quality on subprime originations. To show this, we obtain credit scores for the U.S. population from the FRBNY-Equifax Consumer Credit Panel, which comprises a 5 percent random sample of U.S. individuals (aged 19 and over) with credit reports from 1999 to 2009 (Lee and Van der Klaauw, 2010).[32] In what follows, we use terms such as credit scores for the credit eligible population or consumer population or simply population credit scores to describe these data.

The results are also plotted as gray lines in Figure B.1A. At higher credit scores, the cdf of credit scores on subprime originations is above those for the consumer population. This is expected, since borrowers with higher credit scores are less likely to opt for a subprime mortgage (Pennington-Cross, 2003). For the lowest credit scores, the cdf on subprime originations in Figure B.1 is below that for the population. Again, there is likely to be a greater proportion of borrowers with lower credit scores in the population than among those with subprime mortgages. Finally, the data reveal a marginal improvement in the credit score distribution from 2000-2002 to 2004-2006 for the population as well.

A cohort by cohort comparison between the periods 2000-2002 and 2004-2006 reveals that the improvement in credit scores is statistically significant for subprime originations but not for the entire U.S. consumer population. For example, Figure B.1B compares the credit scores on subprime originations (cohorts) of 2000 and 2006 with those in the population for the same years. As before, the cdf of credit scores on subprime originations of 2006 first-order stochastically dominates that on subprime originations of 2000. However, we fail to establish the case for first-order stochastic dominance with the population credit scores for the same years.

B.2 A Formal Test of Comparison between Subprime and Population FICO

The hypotheses stated above are best verified in terms of the statistical tests for stochastic dominance developed in Anderson (1996). Following Anderson (1996), the hypothesis of *first-order stochastic dominance* of distribution $F_A(\cdot)$ over $F_B(\cdot)$ is tested by comparing the two cdfs at various points in the distribution. Let x_A and x_B be the empirical frequency vectors based on samples of size n_A and n_B, drawn respectively from the populations A and B. The Anderson test of stochastic dominance is based on a comparison of the cumulative distributions and at the deciles of the pooled sample. Anderson also shows that under the null of common population distribution (no dominance), and the assumption of independence of the two samples,

[32]Strictly speaking, the credit scores obtained from this longitudinal panel are derived from the methodology used by Equifax to mimic the proprietary algorithm used by Fair Isaac Corporation. Therefore, while they are a close match, the credit scores for each individual may not be identical under the two algorithms.

$\vartheta = \frac{x_A}{n_A} - \frac{x_B}{n_B}$ is asymptotically normally distributed with mean zero. The hypothesis of first-order dominance $F_A(\cdot)$ over $F_B(\cdot)$ requires that no element of ϑ be significantly greater than zero, while at least one element be significantly less.

Table B.1. Anderson (1996) Test for First Order Stochastic Dominance

Under the null of common population distribution (no dominance), and the assumption of independence of the two samples, $\vartheta = x_A/n_A - x_B/n_B$ is asymptotically normally distributed with mean zero. The hypothesis of first-order dominance of $F_A(\cdot)$ over $F_B(\cdot)$ requires that no element of ϑ be significantly greater than zero while at least one element is significantly less. The numbers in parentheses below report the 95% confidence interval.

Panel A: Stochastic Dominance of 2004-2006 over 2000-2002

FICO	Subprime (ϑ)	Subprime (t-stat)	Population (ϑ)	Population (t-stat)
1st Decile	-0.0485	-185.5142	-0.0137	-26.2394
	(-0.049; -0.048)		(-0.0148; -0.0126)	
2nd Decile	-0.0855	-236.4047	-0.0123	-20.1941
	(-0.0862; -0.0848)		(-0.0134; -0.0112)	
3rd Decile	-0.1167	-269.4439	-0.0078	-11.1864
	(-0.1175; -0.1159)		(-0.0092; -0.0064)	
4th Decile	-0.1163	-251.3688	-0.0046	-6.304
	(-0.1171; -0.1155)		(-0.006; -0.0032)	
5th Decile	-0.1055	-220.7806	-0.0022	-2.8359
	(-0.1065; -0.1045)		(-0.0037; -0.0007)	
6th Decile	-0.091	-195.1928	-0.0022	-2.748
	(-0.0919; -0.0901)		(-0.0038; -0.0006)	
7th Decile	-0.0677	-160.1472	-0.0048	-5.655
	(-0.0685; -0.0669)		(-0.0065; -0.0031)	
8th Decile	-0.0401	-116.3653	-0.0088	-10.1431
	(-0.0408; -0.0394)		(-0.0105; -0.0071)	
9th Decile	-0.0114	-57.2645	-0.0186	-21.2404
	(-0.0118; -0.011)		(-0.0203; -0.0169)	

Panel B: Stochastic Dominance of 2006 over 2000

FICO	Subprime (ϑ)	Subprime (t-stat)	Population (ϑ)	Population (t-stat)
1st Decile	-0.0818	-150.3504	-0.0038	-4.2225
	(-0.0829; -0.0807)		(-0.0055; -0.0021)	
2nd Decile	-0.1402	-187.5339	-0.0023	-2.193
	(-0.1416; -0.1388)		(-0.0044; -0.0002)	
3rd Decile	-0.189	-210.3891	-0.0003	-0.2237
	(-0.1908; -0.1872)		(-0.0027; 0.0021)	
4th Decile	-0.1891	-196.2621	0.0018	1.4277
	(-0.191; -0.1872)		(-0.0007; 0.0043)	
5th Decile	-0.1725	-174.0648	0.0031	2.2935
	(-0.1744; -0.1706)		(0.0005; 0.0057)	
6th Decile	-0.1333	-139.3142	0.0026	1.8627
	(-0.1352; -0.1314)		(-0.0002; 0.0054)	
7th Decile	-0.0906	-106.6017	-0.001	-0.6988
	(-0.0923; -0.0889)		(-0.0039; 0.0019)	
8th Decile	-0.0477	-71.2206	-0.0074	-4.9046
	(-0.049; -0.0464)		(-0.0104; -0.0044)	
9th Decile	-0.0126	-35.869	-0.0224	-14.7166
	(-0.0133; -0.0119)		(-0.0254; -0.0194)	

Panel A of Table 1 presents the results of the test for credit scores on 2004-2006 cohorts over those on 2000-2002 cohorts. The numbers in the second column show, for each decile of credit scores, the difference in the cumulative probabilities across the distributions (cohorts) for subprime originations. We also report the 95% confidence intervals for the differences in parentheses below. The third column reports t-statistics for the test of significance. The same results for the population are then presented in the fourth and fifth columns.

For both subprime originations and the general population, the distribution of credit scores on later cohorts is seen to stochastically dominate those on earlier cohorts. However, at all deciles (save the last, with highest credit scores) the difference in probabilities for the subprime originations are significantly greater than that for the population. For example, the probability of having a FICO score greater than 600 (the fourth decile) is 11.63 percent higher for later cohorts in subprime originations, but it is only 0.46 percent higher in the general population.

Panel B of Table B.1 presents the results of the Anderson (1996) test for first-order stochastic dominance of the credit score distribution of the 2006 cohort over that of the 2000 cohort. For the credit score distribution on subprime originations, the null of no dominance is rejected in favor of the alternative of first-order stochastic dominance. However, we fail to establish first-order stochastic dominance for the same years in the population credit distributions. In particular, one finds that for the fifth decile, the difference ϑ is significantly *greater* than zero.

In summary, our results indicate that the distribution of credit scores on later cohorts is seen to stochastically dominate those on earlier cohorts. For credit score distributions on subprime originations, this difference is both economically and statistically significant. However, the improvement in credit scores for the population is at best marginal and not always significantly different across the different cohorts.

B.3 FICO scores and Other Origination Attributes

The previous section showed that while origination FICO on subprime mortgages improved over the cohorts, this improvement cannot be attributed to patterns to changes in the population FICO. In this sense, the improvement in origination FICO was a pattern unique to subprime mortgages only. But what of other (non-FICO score) attributes on the origination? As has been well documented in this literature, there was an increase in the proportion of originations with higher LTV and originations lacking full documentation over the cohorts in the subprime segment (Mayer et al. 2009).

Here, we present evidence showing that this deterioration in other origination attributes was accompanied by increases in FICO scores on the origination. Stated differently, an origination with a riskier attribute (such as lower documentation and higher LTV) is more likely to have a higher FICO score in the 2004-2006 cohorts than in the 2000-2002 cohorts. Stated differently, the tradeoff or adjustment between riskier attributes on the origination and higher credit scores on the same origination grew stronger over time.

Panel A of Table B.2 reports the percentages of loans with and without full documentation under various FICO score groups. Panel B shows the percentage of loans under various FICO score groups for different intervals of LTV. Among the FICO score groups in Table B.2 (both panels), there is a decline in the percentage of borrowers in the lowest FICO group (<620) over the cohorts with an increase in the percentage of borrowers in the next two categories, namely, 620-659 and 660-719. The percentage of borrowers in the highest category (720) remains roughly the same for all cohorts. Moreover, the overall trends for FICO scores in relation to documentation (Panel A) are similar to those in relation to LTV (Panel B). These results also suggest an overall increase in credit score on both high-risk and low-risk originations.

Table B.2: Distribution of FICO vis-à-vis Other Origination Attributes

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 all other originations are classified as *Low doc*. CLTV denotes the combined loan-to-value ratio on the origination.

Panel A: FICO distribution conditional on Documentation level on loan by cohort

	Full doc loans				Low-doc			
Cohort	< 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

Panel B: Distribution of FICO scores conditional on CLTV by cohort

	CLTV ≤ 80				80 < CLTV ≤ 90				90 < CLTV ≤ 100			
Cohort	< 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

APPENDIX C

Why do we explain FICO performance in terms of a relative default measure?

First, both academic and practitioner literature have emphasized that credit scores are not absolute but relative measures of credit risk. For example, Keys et al. (2010, p. 316, emphasis added) note that, "FICO scores provide *a ranking of potential borrowers* by the probability of having some negative credit event in the next two years." In the realm of practitioners, this feature of credit scoring is common knowledge: In their public release statements, Transunion, the third largest credit bureau in the U.S., claim that:

> *Credit scores are not an absolute statement of risk for an individual consumer, rather they state a consumers' risk in relation to other consumers.*[33]

Second, the difficulty with analyzing absolute measures of default is largely due to the fact that we are analyzing default rates in two different regimes under which the different loans perform. As noted in the paper, the early period in the subprime market is characterized by a low default regime (shown here as the bold line in Fig C.1) whereas the later period is characterized by high default regime (broken line). A key feature of the data is that the upward shift of the line is greater for low FICO borrowers than high FICO borrowers leading to a rotation of the default pattern from the early cohorts of subprime originations (regime 1) to the later cohorts (regime 2).

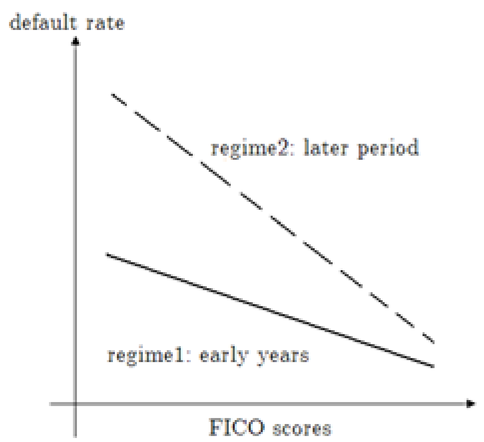


Figure C.1

This change in the default pattern across regimes can be explained in several ways.[34] Our explanation characterizes the regimes as periods with different house price growth (HPG). We borrow results from Table 11 in Bhardwaj and Sengupta (2012). The median house price growth for mortgages in our sample is roughly 8%. We split the HPG variable into three dummies: one each for HPG that is greater than the median of 8%, HPG greater than 1 but less than or equal to 8% and HPG less than or equal to 1%. The results show the default hazard ratios from a competing risk hazard framework that includes FICO scores by quartiles and HPG as covariates.

[33] http://www.transunion.com/docs/rev/business/financialservices/VantageScore_CreditScoreBasics-Part1.pdf

[34] For example, Keys et al. (2010) explains this phenomenon as being caused by reduced incentives to generate soft information under securitization. In their version, Regime 2 is characterized as the securitization regime which caused a reduction in borrower quality across the spectrum, largely generated by a low screening intensity. Since this phenomenon was more severe for borrowers for whom soft information was more valuable, namely low FICO score borrowers, we observe a rotational shift of the lines in the two regimes.

That HPG has been the prime mover of defaults on subprime mortgages across cohorts has been well documented in the subprime literature. However, what these results show is that the interaction between HPG and FICO scores demonstrates remarkable non-linearity. Within each HPG group (movement along the curve), a lower FICO quartile origination shows a higher default hazard. However, across HPG groups (shit of the curve) the impact of house price declines on default is significantly severe on originations with low FICO scores.

Plotting the default rates on ARM2 loans in Table 11 below, we can create a plot (Figure C.2) similar to the one above (Figure C.1), where each cohort is characterized by a different house price regime. The pattern is similar for all subprime mortgage products: fixed rate, ARM with teaser rates for two years (ARM2) and three years (ARM3) respectively. Clearly, the absolute changes in default rates in our story can be explained in largely terms of differences in house price regimes—establishing the point that resorting to absolute measures of default to judge FICO performance suffers from severe identification problems.[35]

Table 11: Default Hazard Ratios for interactions of House Price Growth with credit variables by product type

	ARM2		ARM3		Fixed	
Interactions of House Price Growth (HPG)	Default	Prepay.	Default	Prepay.	Default	Prepay.
… with Credit Scores						
FICO in 4th Quartile and HPG > 8 (Baseline)						
FICO in 4th Quartile and $1 < HPG \leq 8$	8.761***	0.509***	9.138***	0.519***	8.73***	0.734***
FICO in 4th Quartile and $HPG \leq 1$	18.785***	0.273***	23.028***	0.264***	25.18***	0.384***
FICO in 3th Quartile and HPG > 8	1.806***	1.003	1.729***	1.021***	1.862***	1.036***
FICO in 3th Quartile and $1 < HPG \leq 8$	13.181***	0.454***	13.994***	0.466***	14.635***	0.611***
FICO in 3th Quartile and $HPG \leq 1$	28.262***	0.222***	33.361***	0.241***	42.556***	0.28***
FICO in 2nd Quartile and HPG > 8	2.757***	0.986***	2.561***	1.019***	2.832***	1.042***
FICO in 2nd Quartile and $1 < HPG \leq 8$	17.581***	0.428***	17.667***	0.438***	18.501***	0.565***
FICO in 2nd Quartile and $HPG \leq 1$	37.175***	0.217***	45.696***	0.201***	64.548***	0.218***
FICO in Bottom Quartile and HPG > 8	4.918***	0.97***	4.269***	1.010*	4.873***	1.032***
FICO in Bottom Quartile and $1 < HPG \leq 8$	25.925***	0.39***	24.771***	0.394***	26.819***	0.468***
FICO in Bottom Quartile and $HPG \leq 1$	56.303***	0.185***	71.657***	0.155***	110.661***	0.123***
… with Loan-to-Value Ratios						
LTV < 80 and HPG > 8 (Baseline)						
LTV < 80 and $1 < HPG \leq 8$	5.281***	0.513***	6.035***	0.506***	5.979***	0.68***
LTV < 80 and $HPG \leq 1$	7.297***	0.346***	12.633***	0.302***	18.458***	0.317***
$80 \leq LTV < 90$ and HPG > 8	1.074***	0.972***	1.121***	0.98***	1.144***	1.019***
$80 \leq LTV < 90$ and $1 < HPG \leq 8$	7.061***	0.429***	7.959***	0.434***	8.267***	0.552***
$80 \leq LTV < 90$ and $HPG \leq 1$	18.021***	0.189***	24.431***	0.168***	33.999***	0.145***
$90 \leq LTV < 100$ and HPG > 8	1.171***	0.954***	1.261***	0.993	1.358***	0.998
$90 \leq LTV < 100$ and $1 < HPG \leq 8$	8.655***	0.398***	9.756***	0.399***	10.484***	0.488***
$90 \leq LTV < 100$ and $HPG \leq 1$	21.531***	0.171***	27.294***	0.165***	37.775***	0.142***
$LTV \geq 100$ and HPG >8	1.848***	0.89***	1.774***	0.937***	1.965***	0.961***
$LTV \geq 100$ and $1 < HPG \leq 8$	13.044***	0.325***	13.976***	0.357***	14.586***	0.437***
$LTV \geq 100$ and $HPG \leq 1$	27.324***	0.149***	35.122***	0.146***	41.115***	0.178***

[35] It also follows that the defaults in the subprime mortgage market can have complementary explanations. While Keys et al. (2010) emphasize the important role of lax screening as discussed in footnote 2, these results show that the role of house prices and hard information characteristics can hardly be neglected.

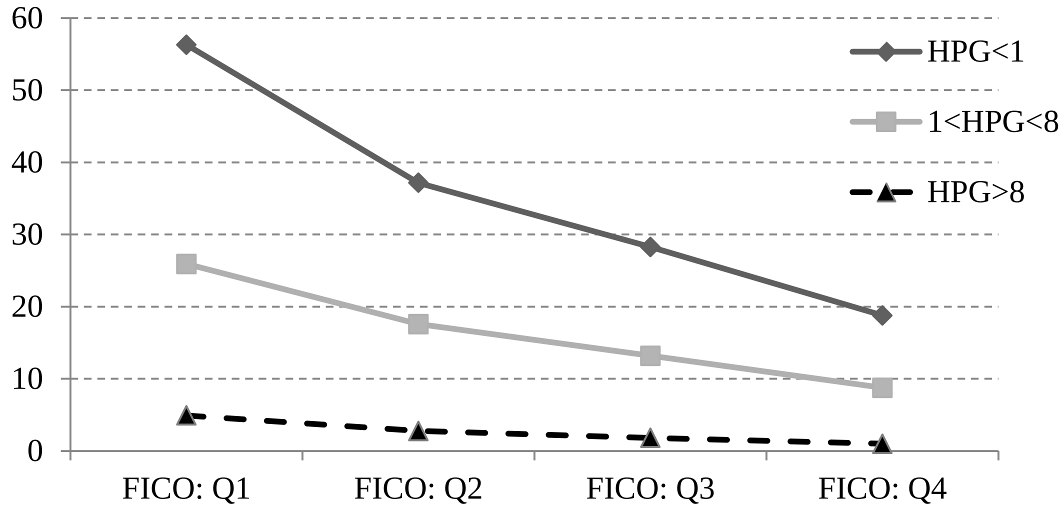


Figure C.2: Default Hazard on Subprime ARM2 loans by House Price Growth (HPG) and FICO quartiles