Research Division
Federal Reserve Bank of St. Louis
*Working Paper Series*

# Subprime Loan Quality

**Geetesh Bhardwaj**
**and**
**Rajdeep Sengupta**

Working Paper 2008-036E
http://research.stlouisfed.org/wp/2008/2008-036.pdf

October 2008
Revised September 2011

FEDERAL RESERVE BANK OF ST. LOUIS
Research Division
P.O. Box 442
St. Louis, MO 63166

---

# Subprime Loan Quality*

Geetesh Bhardwaj† Rajdeep Sengupta‡§

August 2011

**Abstract**

This paper is an exploration of subprime mortgages over the cohorts from 2000 through 2006, especially those prior to 2004. In particular, this study contrasts subprime originations during the "boom years" of 2004-2006 with originations during an "early period" of 2000-2002. We develop a counterfactual technique to determine how originations during the early period would perform in a different environment, namely, the environment faced by originations in 2004, 2005, and 2006. We find that representative originations during the early period of 2000-2002 would not have performed significantly better than representative originations in 2004, 2005, and 2006. This result is robust to counterfactual exercises for originations with different LTVs. We conclude that mortgages of early cohorts were no less vulnerable to the environment faced by cohorts of 2004-2006.

JEL Codes: G21, D82, D86.

Keywords: mortgages, subprime, underwriting, crisis.

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

†Geetesh Bhardwaj, Director - Research, SummerHaven Investment Management. The views expressed herein are those of the individual author and and do not necessarily reflect the official positions of SummerHaven Investment Management.

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

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

# 1 Introduction

> Paulson did not see the size of the coming crisis. Nor did the others. At one point, having put on a large position, Paulson and his team thought they had made a fatal mistake because with their trades the firm had bet against [subprime mortgages that] were handed out before 2006, and were for homes that already had appreciated in value (emphasis added; Zuckerman, p. 159). They traded out of those positions and into later vintages, thinking they dodged a bullet. This was a widespread view, that subprime vintages prior to 2006 were much safer; it was supported by the data, as Paulson and Pellegrini found out. But, when the crisis came, there was no distinction between pre- and post-2006 vintages. Everything went down in value, including bonds linked to the earlier subprime vintages! [1]

Defaults on subprime mortgages in 2006 and 2007 precipitated the current housing crisis. The sheer magnitude of this problem has called into question many of the lending practices that led to this downturn. Most of the existing literature on subprime mortgages has focused on poor quality of loans originated during 2004-2007, the peak of the housing boom. The typical evidence presented relates to a widespread deterioration in subprime mortgage underwriting since 2004.[2] However, even after almost half a decade since the first problems in housing, the market for housing continues to deteriorate. The collapse and the consequent near disappearance of subprime originations have led many to reassess the depth of the subprime malaise and mortgage securitization in general (see Agarwal et al. 2011 and references therein for a recent survey). Additionally, some of the lending practices that were in vogue even before the peak of the housing bubble are now being scrutinized (Pinto, 2010).

This paper is an exploration of subprime mortgages over the cohorts from 2000 through 2006, especially those prior to 2004. In particular, our study contrasts securitized subprime originations during the "boom years" of 2004-2006 with loans originated during an "early period" of 2000-2002. In comparing originations from the same securitized subprime universe over different cohorts, our study is particularly vulnerable to the criticism in Summers (1985). Nevertheless, this study helps increase our understanding of the viability of subprime mortgage lending in general. These questions assume greater importance given that

[1] Gorton (2011, pp. 452-453)

[2] See Mayer et al. (2009) and Levitin and Wachter (2010) for a survey of this literature.

there is recent anecdotal evidence on the resurgence of subprime mortgages despite the recent subprime debacle (Androitis, 2011).

We begin by studying two important trends in observable origination characteristics that have not received much attention in the literature.[3] First, we demonstrate that there was a significant increase in the credit quality on subprime originations as measured by their origination FICO scores. We also note that this change is a subprime phenomenon and not a result of changes in creditworthiness of the U.S. population. Second, we provide evidence that the minimum criteria for obtaining a subprime loan actually increased over the cohorts. We are not the first to point this out. In fact, several observers and practitioners in the industry have commented on this trend as a move away from borrowers that were designated as subprime towards those that were more likely to be Alt-A. Some observers have described this phenomenon in terms of the creation of an Alt-B sector of the economy (see Section 2 and references therein).

This paper introduces a counterfactual technique to determine how originations during the early period would perform in a different environment, namely, the environment faced by originations in 2004, 2005 and 2006. In so doing, we observe that representative originations during the early period of 2000-2002 would not have performed significantly better than originations in 2004, 2005 and 2006. This result is robust to counterfactual exercises for originations with different values of LTV. In fact, earlier cohorts show significantly worse performance especially for high-LTV originations (LTV>90). This is largely due to fact that high-LTV originations of later cohorts had significantly higher credit scores on these originations.

Our results demonstrate that origination credit scores are an important driver of loan performance. While an environment of declining house prices can adversely affect loan performance for high credit score originations—the effect on low credit score originations can be particularly severe. Consequently, low-credit score originations of later cohorts have significantly lower survival rates. Significantly, originations of later cohorts have higher credit scores—not only in absolute terms, but also after adjusting for other attributes on the origination. In essence, this explains why the estimated survivor functions of later cohorts demonstrate higher survival rates than their corresponding counterfactual survivor functions.

[3] The significant decline in origination characteristics in terms of loans with higher loan-to-value (LTV) ratios or lacking full documentation has been well documented (see GAO (2010) and references therein). For a detailed history of such lending practices in the U.S., see Pinto (2010).

These results also suggest that the differences in real performance of originations between the two periods can be attributed to factors ex post to the origination than those ex ante. In addition, it raises the possibility that problems on subprime mortgages were not a recent phenomenon and that serious design flaws in subprime originations make them especially vulnerable to a downturn in home prices (Gorton, 2008). Our results can be interpreted as a reaffirmation of the importance of factors other than underwriting in the collapse of the subprime market (Gerardi et al., 2008; Haughwout et al., 2008).

The next section lays out the changing trends of credit quality and minimum criterion for getting a loan over the different cohorts. In section 3, we describe our methodology for counterfactual estimation, the results of which are presented in Section 4. Section 5 concludes. In an appendix to this paper, we provide further evidence on the changing trends of credit quality on subprime originations.

## 2 Data and Trends

Our principal data source is the ABS Database from Corelogic-LoanPerformance (hereafter, LP) data repository containing loan-level data on securitized mortgages.[4] We restrict our analysis to loan-level data on over nine million first-lien securitized subprime mortgages, originated between 2000 and 2006.[5] Significantly, our study comprises of originations that were securitized and within the securitized subprime universe, the coverage on these loans are fairly high especially for the later cohorts. Summary data on subprime originations over the years record a significant increase in the proportion of adjustable rate mortgages (mainly hybrid-ARMs). In addition, these years saw a sharp increase in the proportion of low-documentation loans and a significant deterioration in loan-to-value ratios (see GAO, 2010 for details). Interestingly however, and contrary to perceived wisdom, this deterioration of loan quality was hardly a secular trend. As mentioned earlier, we draw attention to two patterns that went against this trend. First, we demonstrate that there was a significant increase in the credit quality on subprime originations as measured by their origination FICO scores. We also note that this change is a subprime phenomenon and not a result

[4] This is the largest and the most comprehensive mortgage securities data repository for non-prime mortgages. Details on this database; its evolution, coverage, and comparison with other mortgage databases are available in GAO (2010).

[5] Our data include securitized subprime mortgages only. Consequently, this paper abstracts from the debate on the role of securitization on subprime loan quality (see Agarwal et al. 2011 and references therein for a survey).

of changes in creditworthiness of the U.S. population. Second, we examine the minimum criterion for obtaining a subprime loan in terms of origination characteristics over the two periods. We find that in terms of some basic observable origination characteristics, the minimum criterion to obtain a subprime loan actually strengthened over the two periods.

## 2.1 Changes in Credit Quality

Figure 1A plots (bold lines) the cumulative distribution function (cdf) of the FICO scores on subprime originations during 2000-2002 and 2004-2006 separately. It shows that the probability that a subprime borrower has a lower credit score is significantly higher on originations during 2000-2002 than on originations during 2004-2006. Stated differently, distribution of credit scores during 2004-2006 first-order stochastically dominates the distribution of credit scores during 2000-2002 (Rothschild and Stiglitz, 1970). First-order stochastic dominance is easily confirmed because the graph of the distribution for the 2004-2006 cohort lies uniformly below the graph of the distribution for the 2000-2002 cohort. More formally, we use Anderson's (1996) test of stochastic dominance to confirm this hypothesis. These results are available on request.

An important concern here is whether the observed improvement is due to a shift in the underlying distribution of FICO scores in the entire U.S. population. We confirm that changes in the credit score distribution for the credit-eligible population in the U.S. cannot explain the full improvement in the credit quality on subprime originations. We obtain credit scores for the U.S. 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 (see Lee and Van der Klaauw, 2010 for details). In Figure 1A, the cdfs of credit scores for the entire population are plotted (dotted lines) along with the cdfs of the credit scores on subprime originations.[6]

First, at higher credit scores, the credit score graphs on subprime originations are above those for the U.S. population. This is expected, since borrowers with higher credit scores are less likely to opt for a subprime mortgage. Next, for the lowest credit scores, the credit score graphs on subprime originations are below those for the population. Again, this is not surprising because there is likely to be a greater proportion of borrowers with lower credit scores in the population than among those with subprime mortgages. Finally, and

[6]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.

although this is not apparent from Figure 1A, the data reveal a marginal improvement in the credit score distribution from 2000-2002 to 2004-2006 for the population as well.

However, while we find that the improvement in credit scores on subprime originations is statistically significant across these cohorts, this is not the case for borrowers in the entire population. For example, Figure 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 cohorts of 2006 first-order stochastically dominates that on cohorts of 2000. However, with the population credit scores for the same years, we fail to establish the case for first-order stochastic dominance using the Anderson (1996) test. 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 risk for the population is at best marginal and not always significantly different across the different cohorts.

### 2.2 The Minimum Criterion for Getting a Loan

Table 1 reports the percentage of originations under certain minimum criterion for originations during 2000-2002 and then during 2004-2006. We consider originations with FICO scores not greater than three different thresholds—500, 526 (the first decile of origination FICO during 2000-2002), and 541 (the first decile of origination FICO during 2004-2006). First, the percentage of loans in each category is significantly higher during 2000-2002 than during 2004-2006. For example, 2.45 percent of originations during 2000-2002 had FICO scores no greater than 500, whereas the number drops to 0.31 percent in 2004-2006. Second, the shares of low credit quality originations with other riskier attributes (such as high LTV and low documentation) are also significantly lower for later cohorts. For example, 4.78 percent of loans originated during 2000-2002 had credit scores not greater than 541 and LTV ratios greater than 80, but the corresponding number for originations during 2004-2006 is only 2.7 percent.

These results suggest attempts to control for overall credit risk at the bottom of the risk spectrum by reducing other riskier attributes on originations of poorer credit quality. Moreover, this trend is observed across all credit score originations: that higher credit scores were used to adjust for riskier attributes on the origination (see Appendix). Interestingly, we also find that the strength of this adjustment increased signficantly over the cohorts during this period (also see Section 4.2). In summary, lenders appear to have attempted to offset riskier attributes on originations (such as lower documentation and higher LTV) by increasing the average quality of borrowers (as measured by their credit scores) to whom such loans were made. Evidence of these trends are presented in an appendix to this paper.

### 2.3 Anecdotal Evidence

With the benefit of hindsight, most observers have pointed to poor quality of originations among the worst performing cohorts—namely, originations during 2004-2006. However, it is often interesting to view some of the analysis during the years when subprime was in vogue. Anecdotal evidence appears to point to the trends described above. In their handbook chapter on non-prime mortgages, Bhattacharya et al. (2006, p. 189) observe that

> ... the demarcation between Alt-A and subprime loans has been blurred. Over time Alt-A has expanded to include loans with progressively less documentation and lower borrower credit scores. At the same time, subprime loans have, on average, experienced a slow but steady rise in average credit scores. A result of this convergence has been the creation of the so-called alt-B sector, where loans using this nomenclature were securitized in 2004.

From a different study in the same volume, Zimmerman (2006, p. 106) remarks:

> ... 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.

At the very least, this raises doubts about the popular perception of subprime lending seeking borrowers of lower and lower quality in an effort to increase market share. It would be important to recall that over 70 percent of subprime loans in each cohort were originated as refinances of existing mortgages. Presumably, some of these mortgages came from other segments of the market. Therefore, at least in theory, it is not impossible for subprime mortgages to expand market share by seeking out borrowers with higher credit scores.

# 3 Counterfactual Analysis

## 3.1 Background

In this section, we introduce a counterfactual technique aimed at comparing origination quality across cohorts while taking into account that ex post behavior of economic variables are different for each cohort. The distinguishing characteristic of loan quality as summarized by origination characteristics is that they do not change after origination and, therefore, are *time-invariant*. In contrast, factors that affect loan performance *subsequent to the origination*, such as those at the individual level (such as medical condition of the borrower, her income and consumption patterns, etc.) and the local or national level (such as unemployment conditions or house price appreciation) change with time.

A simple way of distinguishing the two sets of factors, mentioned above, would be to adopt the relative-risk hazard model (Cox, 1972):

$$h(t) = h_0(t)\varphi(\mathbf{X},\beta), \tag{1}$$

where $h_0(t)$ is the baseline hazard function. In this case, the hazard function, $h(t)$, is expressed as a product of two functions: The time-varying baseline hazard, $h_0(t)$, captures the changes in hazard function as a function of survival time. The second function, $\varphi(\mathbf{X},\beta)$, captures the relationship of hazard function as a function of time-invariant covariates. Hazard rate models using time-varying baseline hazards and deterministic covariates such as origination characteristics are useful in relative risk modeling across cohorts.

Origination characteristics $\mathbf{X}$, are predetermined and therefore exogenous to default, thereby ensuring unconfoundedness or "selection on observables" (Barnow et al., 1981; Chalak and White, 2010). Consequently, the estimated coeffiecients, $\hat{\beta}$, are consistent (Angrist and Krueger, 1999). Most important, the estimated hazard function (being the product of estimated time-varying baseline hazard $h_0(t)$ and the function $\varphi(\mathbf{X},\beta)$) has the *optimal predictive* interpretation because it provides an unbiased quasi-maximum-likelihood approximation of the hazard function (White, 2006).

The unbiasedness of the predictive estimate becomes important when one considers that our specification does not include time-varying covariates and does not model the role of prepayment in our default regression. This departure is deliberate because we are largely motivated by our desire to explain the impact of loan quality rather than explain default. Bhardwaj and Sengupta (2011) provide a competing risk hazard model using time-varying covariates to explain default for subprime mortgages.

### 3.2 Estimation Strategy

The unbiased estimate of the survivor function is used to develop a counterfactual exercise that allows us to compare the originations of one cohort with those of another. First, the relative-risk hazard model in (1) is estimated for each cohort. Next, we use the estimated relationship to evaluate the *estimated survivorship function* for a representative origination from a different cohort (Cameron and Trivedi, 2005). As mentioned above, estimating a proportional hazard model in this setting provides us with an *optimal predictive* estimate of the survivor function (White, 2006).

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

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

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

For our purposes, the objective is to forecast the impact on the survivor function of cohort $v_2$ in the environment of cohort $v_1$.[7] In this specification, let $\mathbf{X}_1$ and $\mathbf{X}_2$ denote the representative originations of cohort $v_1$ and $v_2$, respectively. If unobservable characteristics and market conditions captured by the baseline survivor function are applied to the different origination characteristics, we can identify the effect of $\mathbf{X}_2$ on the survivor function in $v_1$ as follows:

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

Such a counterfactual exercise helps us test the following hypothesis:

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

First, we estimate the Cox proportional hazard model for a given cohort $v$. Next, we calculate the estimated survivor function for the representative origination of cohort $v$.

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

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

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

# 4 Results

## 4.1 Baseline Counterfactuals

With these tools in place, we can now study the performance of originations of different vintages across different environments. The null hypothesis is that mortgage applications approved during 2000-2002 are equally likely to survive an event of default as those of later cohorts—namely, 2004, 2005, and 2006—in the environment of these cohorts. Our choice of years on the counterfactual is motivated by the fact that the information set of the lender for post-2003 originations should arguably include the repayment patterns on 2000-2002 cohorts. Nevertheless, we have also conducted a reverse-counterfactual that yields qualitatively similar results. The results for the reverse-counterfactual exercise are available upon request.

Results for the relative-risk (Cox) hazard regression used in deriving the survivor function is presented in Table 2. In addition to the covariates shown in Table 2, we control for borrower attributes, loan source, lender characteristics, property type, and property location. The variable *closing rate spread* is defined here as the difference between the closing rate on the origination (the teaser rate for hybrid-ARMs) and the 30-year conventional mortgage rate. *Home Value nth Quartile* is a dummy that equals one if the value of the property lies in the $n$-th quartile of all property values in the data and zero otherwise. In addition, the covariates includes dummy variables for full documentation on the loan, owner and second-home occupancy (as opposed to investor owned properties) and cash-out and no-cash-out refinances (as opposed to purchases). As shown from the hazard ratios in Table 2, most covariates have the expected effect on default—increases in LTV, closing rate spread increases the default hazard. Moreover, full documentation, owner-occupied properties and higher credit scores reduce the default hazard.

The results of counterfactual analysis are summarized in Table 3. Table 3 has three panels corresponding to the counterfactual exercises using survivor function estimates based on 2004, 2005, and 2006 data. The numbers in parentheses are the 95% confidence intervals for the estimated and counterfactual survivor functions. The estimated survival probability after 36 months for the 2004 cohort is 0.8123, whereas the counterfactual survival probabilities of 2000, 2001 and 2002 cohorts in the 2004 environment are 0.7399, 0.7503, and 0.7787, respectively. Notably, confidence intervals of the 2004 survivor function lies above that of the 2000-2002 counterfactual survivor functions. Significantly, this feature holds true for the confidence intervals of the 2005 and 2006 survivor functions over all the loan ages considered. These results may be summarized as follows: If representative originations for each cohort during 2000-2002 were to be originated in 2004, 2005, and 2006, their performance would be no better than representative originations of 2004, 2005, and 2006 cohorts respectively.

The estimated and counterfactual survivor functions are best illustrated in terms of the survival plots in Figure 2. The left, middle and right plots in Figure 2 show the estimated survivor functions (in red) using data on cohorts of 2004, 2005, and 2006 respectively. Alongside the estimated survivor function for each cohort, we plot the counterfactual survivor functions for representative originations of 2000, 2001, and 2002 in the environment of these cohorts. In comparison to the counterfactual survivor functions—labeled here as C2000, C2001, and C2002—we find that the estimated survivor functions of the later cohorts (red lines) demonstrate higher survival rates. As discussed above, we can reject the null hypothesis in favor of the alternative that representative originations of earlier cohorts—namely, 2000-2002—would perform no better when compared with representative originations of 2004, 2005, and 2006 cohorts. In summary, the counterfactual analysis provides evidence against the hypotheses that subprime mortgages of earlier cohorts would demonstrate better performance in an environment of falling home prices. Before providing a discussion on these results, we perform a robustness check to our counterfactuals results below.

### 4.2 Robustness Check: Counterfactuals Conditioning on LTV

Some credit variables, most notably LTV, are known to exhibit nonlinear effects on default, especially in environments with adverse economic conditions such as high unemployment and declining home prices. Therefore, as a robustness check of our counterfactual estimates, we conduct the counterfactual exercise to account for some of these nonlinear effects. For this, we split our sample in terms of loan-to-value ratio into three categories. Originations with LTV less than 80 percent are grouped as category C1, those with LTV in the interval [80, 90) are grouped as category C2, and those with LTV of 90 percent or above are grouped as category C3. We conduct the same counterfactual exercises as described above for the different cohorts in each category of originations. We do this in two ways. First, we use the regressions for the full sample and use counterfactual survivor functions for each of the categories C1 through C3. Second, we re-estimate the regressions separately for loans in each of the categories and use the survivor functions in each of the categories separately. The results are qualitatively similar.

Just as shown in Table 3, Tables 4, 5, and 6 report the estimates of the survivor functions and counterfactuals by loan age for category C1, C2, and C3 respectively. The counterfactuals for each cohort are qualitatively similar to the counterfactual results for the full sample in Table 3. In addition, we present plots of the estimated and counterfactual survivor functions (similar to that in Figure 2) for the three LTV categories in each cohort. Figures 3, 4 and 5 plot the estimated survivor functions (along with their corresponding counterfactuals) for cohorts 2004, 2005, and 2006 respectively. Not surprisingly, survival probabilities at the same loan age decrease progressively as one moves from the plot for category C1 (LTV$<$80) to those for category C3 (LTV$\geq$90). Again, if one compares across cohorts, survival probabilities for the same loan age are highest for 2004 and lowest for 2006. This is easily seen by comparing the three panels in each of the Tables 3-5. Finally, Tables 3-5 confirm our earlier baseline counterfactual results for each of the categories C1 through C3: Without exception, confidence intervals of the 2004, 2005, and 2006 estimated survivor functions lie above that of their corresponding counterfactual survivor functions for 2000-2002 cohorts.

In addition, the counterfactuals reveal a striking feature of loan performance across cohorts: We observe that for the lowest LTV category C1, the difference between the estimated survivor function and its corresponding counterfactual survivor function is small. In contrast, this difference increases with LTV category and is significantly larger for the highest LTV category C3. This would appear to suggest that among high LTV loans, the performance was significantly better for originations of later cohorts. This result appears to go against perceived wisdom since later cohorts were subject to the high-default environment during a period of rapidly deteriorating house price appreciation. Why would high-LTV originations of 2000-2002 perform even worse?

The answer to this question lies in studying other attributes on the origination, namely credit quality as measured in terms of origination FICO score. Figure 6 plots the kernel density function of origination FICO for 2000-2002 cohorts and 2004-2006 cohorts separately in each of the three LTV categories. In line with our earlier assertion of overall increases in credit scores, we find that the plot for 2004-2006 cohorts (in red) almost always lies to the right of the plot for the 2000-2002 cohorts (in black). Notably, this difference in the FICO distribution between the cohorts is highest for the third LTV category—namely, loans with LTV greater than or equal to 90%.[8] For a given LTV category, this difference in FICO distribution translates in our counterfactual exercise into differences in survival rates between the estimated and counterfactual survivor functions. Therefore, the difference in the estimated and counterfactual survivor functions is the highest at LTV category C3.

These results have interesting implications in terms of the counterfactual results for the full sample. They suggest that, despite its shortcomings, origination credit scores are an important driver of loan performance. While an environment of declining house prices can adversely affect loan performance for high credit score originations—the effect on low credit score originations can be particularly severe. Consequently, low-credit score originations of later cohorts have significantly lower survival rates. Significantly, originations of later cohorts have higher credit scores—not only in absolute terms, but also after adjusting for other attributes on the origination (see Appendix for details). In essence, this explains why the estimated survivor functions of later cohorts demonstrate higher survival rates than

[8]Formally, we conduct Anderson's (1996) test of stochastic dominance for the three plots in Figure 6. We establish that the FICO distributions of later cohorts (2004-2006) first-order stochastically dominate distributions of earlier cohorts (2000-2002) in two of the three LTV categories, namely [80,90) and "greater than 90 percent."

their corresponding counterfactual survivor functions.

### 4.3 An Important Caveat

The counterfactual estimates are conducted for representative originations for the cohort and therefore independent of the number of originations in a cohort. In effect, we assume that the distribution of loan attributes (i.e., loan quality) within a cohort is independent of the number of originations in the cohort. Skeptics could argue that this assumption is unusually restrictive because loan quality is bound to deteriorate as the number of originations increase. Note that this restrictive assumption works against our results—especially, when one considers that originations during 2004-2006 were far greater than those during 2000-2002. If the assumption were violated, this implies that loan quality for 2000-2002 could deteriorate even further. This is particularly true given that the proportion of high-LTV loans originated during 2004-2006 is significantly higher than the same during 2000-2002. As described above, the two methods used in deriving our baseline counterfactuals take this into account. Significantly, both methods yield materially similar results.

## 5 Conclusion

It is important to note that our comparison is merely across cohorts but for the same category of mortgage loans—namely, securitized subprime originations. Our conclusions therefore do not extend to the mortgage market as a whole. The increasing share of subprime originations is evidence enough of declining loan quality for the overall mortgage market. Our results merely point to the fact that, in reviving subprime lending it is not advisable to go back to the "earlier period of 2000-2002" as a viable model of subprime. As the results indicate, these cohorts of mortgages were no less vulnerable to declining home prices—the environment faced by cohorts of 2004-2006.

## References

Agarwal, S., Chang Y. and A. Yavas (2011), Adverse Selection in Mortgage Securitization, Paolo Baffi Centre Research Paper No. 2010-67

Anderson, G,., (1996)."Nonparametric Tests of Stochastic Dominance in Income Distributions," Econometrica, Econometric Society, vol. 64(5), pages 1183-93, September.

Angrist, J., Krueger, A. (1999), "Empirical strategies in labor economics." In: Ashenfelter, O., Card, D. (Eds.), Handbook of Labor Economics, vol. 3A. Elsevier, Amsterdam, pp. 1277–1368.

Andriotis A., (2011) Riskier Loans Make a Comeback, as Private Firms Take the Field, Wall Street Journal, Homes, July 12.

Barnow, B., Cain, G., Goldberger, A. (1981), "Selection on observables," Evaluation Studies Review Annual 5, 43–59.

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

Bhardwaj, G., and Sengupta, R. (2011). Subprime Mortgage Design. St. Louis Fed Working Paper, 2008-039C.

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

Chalak, K. and H. White (2010), "An Extended Class of Instrumental Variables for the Estimation of Causal Effects," Canadian Journal of Economics

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

Gerardi, Kristopher, Adam Hale Shapiro, and Paul S. Willen (2007). "Subprime Outcomes: Risky Mortgages, Homeownership Experiences, and Foreclosures." Working Paper No. 07-15. Federal Reserve Bank of Boston, December, 2007.

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

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

Gorton, G. (2011), Book Review: The Big Short Shrift . Journal of Economic Literature, June 2011

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

Levitin, Adam J. and Wachter, Susan M., Explaining the Housing Bubble (August 31, 2010). University of Pennsylvania Institute for Law & Economics Research Paper No. 10-15

Haughwout Andrew , Richard Peach, Joseph Tracy (2008), Juvenile delinquent mortgages: Bad credit or bad economy?, Journal of Urban Economics, Volume 64, Issue 2, September, Pages

246-257.

Heckman, J.J. and E.J. Vyltacil (2007), "Econometric evaluation of social programs, part I: causal models, structual models and econometric policy evaluation", Handbook of Econometrics, Vol 6B, Elsevier.

Pinto, E., (2010), Government Housing Policies in the Lead-Up to the Financial Crisis: A Forensic Study. Available at SSRN: http://ssrn.com/abstract=1675959

Rothschild, M. & Stiglitz, J. E., (1970). "Increasing risk: I. A definition," Journal of Economic Theory, Elsevier, vol. 2(3), pages 225-243, September.

White, H. (2006), "Time Series Estimation of the Effects of Natural Experiments," Journal of Econometrics, 135, 527-566.

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

Summers, Lawrence H, 1985. " On Economics and Finance," Journal of Finance, American Finance Association, vol. 40(3), pages 633-35, July.

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

**Table 1: Minimum Criterion for obtaining a subprime loan**

The figures show the percentage of originations below a given credit score between 2000 and 2002. The table includes three credit score cutoffs at very low FICO scores showing the lowest quality originations. The first is chosen to be 500. The second cutoff of 526 includes the first decile of FICO on originations during the period 2000-2002. The third cutoff of 541 includes the first decile of FICO on originations during the period 2004-2006.

| Percentage of Total Originations… | 2000-2002 | 2004-2006 |
|---|---|---|
| —with FICO not greater than 500 | 2.45 | 0.31 |
| —with FICO not greater than 500 and CLTV> 80% | 0.80 | 0.06 |
| —with FICO not greater than 500 and low or no documentation | 0.34 | 0.07 |
| | | |
| —with FICO not greater than 526 (1st decile for 2000-2002) | 9.85 | 5.62 |
| —with FICO not greater than 526 and CLTV> 80% | 2.82 | 1.20 |
| —with FICO not greater than 526 and low or no documentation | 1.65 | 1.32 |
| | | |
| —with FICO not greater than 541 (1st decile for 2004-2006) | 16.40 | 10.08 |
| —with FICO not greater than 541 and CLTV> 80% | 4.78 | 2.70 |
| —with FICO not greater than 541 and low or no documentation | 2.79 | 2.41 |

**Table 2: Cox relative rate hazard rate regression: 90 day delinquency event**

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

| | **2004** | **2005** | **2006** |
|---|---|---|---|
| FICO | 0.9912*** | 0.9929*** | 0.9937*** |
| CLTV | 1.0247*** | 1.0284*** | 1.0304*** |
| Full- Documentation | 0.7672*** | 0.7273*** | 0.6831*** |
| Closing Rate Spread | 1.2130*** | 1.2234*** | 1.1695*** |
| Owner Occupied | 0.7189*** | 0.7549*** | 0.7198*** |
| Second Home | 0.6438*** | 0.7383*** | 0.7135*** |
| Refinance (Cash Out) | 0.6803*** | 0.7068*** | 0.8352*** |
| Refinance (No Cash Out) | 0.6769*** | 0.7108*** | 0.8082*** |
| Home Value First Quartile | 0.6560*** | 0.5432*** | 0.5123*** |
| Home Value Second Quartile | 0.7070*** | 0.6089*** | 0.5970*** |
| Home Value Third Quartile | 0.7833*** | 0.7903*** | 0.7836*** |
| LR test $H_0$: $\beta = 0$ (p-value) | 7552462 | 15071994 | 15316152 |

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

.

**Table 3: Counterfactual Survival Analysis**

Three panels report numbers corresponding to counterfactual exercise using survivor function estimates based on 2004, 2005, and 2006 data. The numbers in the parentheses are lower and upper confidence limits at 95 percent confidence interval for the estimated survivor function and the counterfactual survivor function.

*Panel 1: Counterfactual Analysis 2004*

| Age of Loan (Months) | Survivor Function 2004 | Counterfactual Survivor Function 2000 | Counterfactual Survivor Function 2001 | Counterfactual Survivor Function 2002 |
|---|---|---|---|---|
| 12 | 0.9733 | 0.9648 | 0.9660 | 0.9703 |
| | (0.973,0.9736) | (0.9643,0.9653) | (0.9655,0.9664) | (0.9699,0.9706) |
| 24 | 0.9078 | 0.8795 | 0.8834 | 0.8975 |
| | (0.9071,0.9085) | (0.8782,0.8808) | (0.8823,0.8845) | (0.8967,0.8984) |
| 36 | 0.8156 | 0.7629 | 0.7701 | 0.7962 |
| | (0.8144,0.8168) | (0.7606,0.7652) | (0.7681,0.7721) | (0.7947,0.7978) |
| 48 | 0.6982 | 0.6207 | 0.6310 | 0.6693 |
| | (0.6963,0.7) | (0.6174,0.6241) | (0.6281,0.634) | (0.667,0.6715) |
| 60 | 0.6059 | 0.5143 | 0.5262 | 0.5712 |
| | (0.6028,0.609) | (0.5097,0.5188) | (0.5221,0.5304) | (0.5676,0.5748) |

*Panel 2: Counterfactual Analysis 2005*

| Age of Loan (Months) | Survivor Function 2005 | Counterfactual Survivor Function 2000 | Counterfactual Survivor Function 2001 | Counterfactual Survivor Function 2002 |
|---|---|---|---|---|
| 12 | 0.9572 | 0.9505 | 0.9513 | 0.9523 |
| | (0.9568,0.9576) | (0.95,0.9511) | (0.9508,0.9518) | (0.952,0.9527) |
| 24 | 0.8482 | 0.8262 | 0.8287 | 0.8404 |
| | (0.8472,0.8491) | (0.8249,0.8275) | (0.8274,0.83) | (0.8396,0.8411) |
| 36 | 0.6270 | 0.5820 | 0.5870 | 0.5997 |
| | (0.6252,0.6287) | (0.5795,0.5845) | (0.5846,0.5895) | (0.5984,0.601) |
| 48 | 0.4371 | 0.3831 | 0.3890 | 0.4049 |
| | (0.434,0.4402) | (0.3794,0.3868) | (0.3854,0.3927) | (0.4022,0.4075) |

*Panel 3: Counterfactual Analysis 2006*

| Age of Loan (Months) | Survivor Function 2006 | Counterfactual Survivor Function 2000 | Counterfactual Survivor Function 2001 | Counterfactual Survivor Function 2002 |
|---|---|---|---|---|
| 12 | 0.9258 | 0.9108 | 0.9198 | 0.9217 |
| | (0.9251,0.9264) | (0.9102,0.9114) | (0.919,0.9206) | (0.9208,0.9226) |
| 24 | 0.7104 | 0.6609 | 0.6904 | 0.6967 |
| | (0.7086,0.7122) | (0.6596,0.6621) | (0.6882,0.6927) | (0.6942,0.6993) |
| 36 | 0.4423 | 0.3722 | 0.4132 | 0.4222 |
| | (0.4391,0.4454) | (0.3698,0.3747) | (0.4096,0.4169) | (0.4182,0.4264) |

**Table 4: Counterfactual Survival Analysis for category C1**

Three panels report numbers corresponding to counterfactual exercise using survivor function estimates based on 2004, 2005, and 2006 data for originations with LTV<80. The numbers in the parentheses are lower and upper confidence limits at 95 percent confidence interval for the estimated survivor function and the counterfactual survivor function.

*Panel 1: Counterfactual Analysis 2004*

| Age of Loan (Months) | Survivor Function 2004 | Counterfactual Survivor Function 2000 | Counterfactual Survivor Function 2001 | Counterfactual Survivor Function 2002 |
|---|---|---|---|---|
| 12 | 0.9807 | 0.9726 | 0.9744 | 0.9772 |
| | (0.9804,0.9809) | (0.9721,0.973) | (0.974,0.9748) | (0.9769,0.9775) |
| 24 | 0.9325 | 0.9053 | 0.9113 | 0.9207 |
| | (0.9318,0.9333) | (0.904,0.9065) | (0.9102,0.9124) | (0.9198,0.9216) |
| 36 | 0.8631 | 0.8108 | 0.8222 | 0.8403 |
| | (0.8617,0.8645) | (0.8084,0.8132) | (0.8202,0.8243) | (0.8386,0.8419) |
| 48 | 0.7715 | 0.6910 | 0.7082 | 0.7358 |
| | (0.7693,0.7736) | (0.6874,0.6946) | (0.7051,0.7114) | (0.7333,0.7384) |
| 60 | 0.6964 | 0.5972 | 0.6181 | 0.6520 |
| | (0.6932,0.6997) | (0.5924,0.602) | (0.6138,0.6224) | (0.6482,0.6557) |

*Panel 2: Counterfactual Analysis 2005*

| Age of Loan (Months) | Survivor Function 2005 | Counterfactual Survivor Function 2000 | Counterfactual Survivor Function 2001 | Counterfactual Survivor Function 2002 |
|---|---|---|---|---|
| 12 | 0.9705 | 0.9614 | 0.9630 | 0.9673 |
| | (0.9702,0.9708) | (0.9609,0.9618) | (0.9626,0.9634) | (0.9669,0.9677) |
| 24 | 0.8933 | 0.8622 | 0.8676 | 0.8824 |
| | (0.8925,0.8942) | (0.8608,0.8635) | (0.8664,0.8688) | (0.8814,0.8834) |
| 36 | 0.7263 | 0.6568 | 0.6686 | 0.7013 |
| | (0.7244,0.7282) | (0.654,0.6595) | (0.6661,0.6711) | (0.6992,0.7034) |
| 48 | 0.5673 | 0.4746 | 0.4899 | 0.5331 |
| | (0.5641,0.5706) | (0.4705,0.4788) | (0.486,0.4938) | (0.5296,0.5367) |

*Panel 3: Counterfactual Analysis 2006*

| Age of Loan (Months) | Survivor Function 2006 | Counterfactual Survivor Function 2000 | Counterfactual Survivor Function 2001 | Counterfactual Survivor Function 2002 |
|---|---|---|---|---|
| 12 | 0.9419 | 0.9372 | 0.9365 | 0.9411 |
| | (0.9413,0.9424) | (0.9363,0.938) | (0.9357,0.9372) | (0.9405,0.9418) |
| 24 | 0.7668 | 0.7500 | 0.7476 | 0.7642 |
| | (0.7651,0.7685) | (0.7473,0.7528) | (0.7452,0.75) | (0.7621,0.7663) |
| 36 | 0.5307 | 0.5035 | 0.4996 | 0.5265 |
| | (0.5275,0.534) | (0.4988,0.5082) | (0.4954,0.5038) | (0.5227,0.5303) |

**Table 5: Counterfactual Survival Analysis for category C2**
Three panels report numbers corresponding to counterfactual exercise using survivor function estimates based on 2004, 2005, and 2006 data for originations with LTV in [80, 90). The numbers in the parentheses are lower and upper confidence limits at 95 percent confidence interval for the estimated survivor function and the counterfactual survivor function.

*Panel 1: Counterfactual Analysis 2004*

| Age of Loan (Months) | Survivor Function 2004 | Counterfactual Survivor Function 2000 | Counterfactual Survivor Function 2001 | Counterfactual Survivor Function 2002 |
|---|---|---|---|---|
| 12 | 0.9728 | 0.9609 | 0.9626 | 0.9674 |
| | (0.9725,0.9731) | (0.9603,0.9614) | (0.9622,0.9631) | (0.967,0.9678) |
| 24 | 0.9061 | 0.8668 | 0.8726 | 0.8881 |
| | (0.9054,0.9068) | (0.8655,0.8681) | (0.8714,0.8737) | (0.8872,0.889) |
| 36 | 0.8123 | 0.7399 | 0.7503 | 0.7787 |
| | (0.8111,0.8135) | (0.7375,0.7423) | (0.7482,0.7523) | (0.7772,0.7803) |
| 48 | 0.6933 | 0.5881 | 0.6027 | 0.6436 |
| | (0.6914,0.6951) | (0.5847,0.5914) | (0.5998,0.6056) | (0.6413,0.6458) |
| 60 | 0.6000 | 0.4769 | 0.4936 | 0.5408 |
| | (0.5968,0.6031) | (0.4724,0.4815) | (0.4894,0.4978) | (0.5372,0.5444) |

*Panel 2: Counterfactual Analysis 2005*

| Age of Loan (Months) | Survivor Function 2005 | Counterfactual Survivor Function 2000 | Counterfactual Survivor Function 2001 | Counterfactual Survivor Function 2002 |
|---|---|---|---|---|
| 12 | 0.9582 | 0.9447 | 0.9466 | 0.9535 |
| | (0.9578,0.9585) | (0.9441,0.9452) | (0.9461,0.9471) | (0.9531,0.9539) |
| 24 | 0.8515 | 0.8071 | 0.8133 | 0.8358 |
| | (0.8507,0.8523) | (0.8058,0.8085) | (0.8121,0.8145) | (0.8349,0.8368) |
| 36 | 0.6339 | 0.5447 | 0.5566 | 0.6014 |
| | (0.6325,0.6353) | (0.5423,0.5471) | (0.5545,0.5587) | (0.5997,0.6031) |
| 48 | 0.4457 | 0.3406 | 0.3540 | 0.4060 |
| | (0.4429,0.4486) | (0.3371,0.3442) | (0.3506,0.3573) | (0.403,0.4091) |

*Panel 3: Counterfactual Analysis 2006*

| Age of Loan (Months) | Survivor Function 2006 | Counterfactual Survivor Function 2000 | Counterfactual Survivor Function 2001 | Counterfactual Survivor Function 2002 |
|---|---|---|---|---|
| 12 | 0.9199 | 0.9127 | 0.9125 | 0.9144 |
| | (0.9192,0.9206) | (0.9118,0.9137) | (0.9116,0.9133) | (0.9138,0.915) |
| 24 | 0.6908 | 0.6671 | 0.6663 | 0.6726 |
| | (0.6889,0.6926) | (0.6646,0.6698) | (0.664,0.6686) | (0.6713,0.6739) |
| 36 | 0.4137 | 0.3807 | 0.3796 | 0.3882 |
| | (0.4105,0.4169) | (0.3768,0.3847) | (0.376,0.3831) | (0.3857,0.3907) |

**Table 6: Counterfactual Survival Analysis for category C3**

Three panels report numbers corresponding to counterfactual exercise using survivor function estimates based on 2004, 2005, and 2006 data for originations with LTV≥ 90. The numbers in the parentheses are lower and upper confidence limits at 95 percent confidence interval for the estimated survivor function and the counterfactual survivor function.

*Panel 1: Counterfactual Analysis 2004*

| Age of Loan (Months) | Survivor Function 2004 | Counterfactual Survivor Function 2000 | Counterfactual Survivor Function 2001 | Counterfactual Survivor Function 2002 |
|---|---|---|---|---|
| 12 | 0.9673 | 0.9540 | 0.9566 | 0.9644 |
| | (0.9669,0.9677) | (0.9534,0.9546) | (0.956,0.9571) | (0.9639,0.9648) |
| 24 | 0.8879 | 0.8449 | 0.8530 | 0.8782 |
| | (0.887,0.8888) | (0.8433,0.8464) | (0.8515,0.8544) | (0.8771,0.8792) |
| 36 | 0.7783 | 0.7010 | 0.7152 | 0.7605 |
| | (0.7768,0.7799) | (0.6983,0.7037) | (0.7127,0.7178) | (0.7586,0.7623) |
| 48 | 0.6430 | 0.5347 | 0.5540 | 0.6172 |
| | (0.6407,0.6453) | (0.5311,0.5383) | (0.5505,0.5575) | (0.6145,0.6199) |
| 60 | 0.5402 | 0.4176 | 0.4388 | 0.5102 |
| | (0.5365,0.5438) | (0.4129,0.4225) | (0.4341,0.4436) | (0.5062,0.5142) |

*Panel 2: Counterfactual Analysis 2005*

| Age of Loan (Months) | Survivor Function 2005 | Counterfactual Survivor Function 2000 | Counterfactual Survivor Function 2001 | Counterfactual Survivor Function 2002 |
|---|---|---|---|---|
| 12 | 0.9465 | 0.9325 | 0.9364 | 0.9436 |
| | (0.946,0.947) | (0.9316,0.9333) | (0.9354,0.9375) | (0.9432,0.9441) |
| 24 | 0.8130 | 0.7686 | 0.7810 | 0.8038 |
| | (0.8119,0.8142) | (0.7662,0.7709) | (0.778,0.784) | (0.8028,0.8048) |
| 36 | 0.5561 | 0.4741 | 0.4961 | 0.5384 |
| | (0.5541,0.5581) | (0.4702,0.4781) | (0.4909,0.5014) | (0.5368,0.54) |
| 48 | 0.3534 | 0.2664 | 0.2887 | 0.3337 |
| | (0.3501,0.3566) | (0.2619,0.271) | (0.2829,0.2946) | (0.3308,0.3367) |

*Panel 3: Counterfactual Analysis 2006*

| Age of Loan (Months) | Survivor Function 2006 | Counterfactual Survivor Function 2000 | Counterfactual Survivor Function 2001 | Counterfactual Survivor Function 2002 |
|---|---|---|---|---|
| 12 | 0.9086 | 0.8895 | 0.8937 | 0.8981 |
| | (0.9079,0.9093) | (0.8887,0.8902) | (0.8927,0.8947) | (0.8971,0.8991) |
| 24 | 0.6539 | 0.5950 | 0.6077 | 0.6211 |
| | (0.6522,0.6556) | (0.5936,0.5965) | (0.6052,0.6102) | (0.6186,0.6236) |
| 36 | 0.3629 | 0.2898 | 0.3047 | 0.3210 |
| | (0.3601,0.3659) | (0.2873,0.2923) | (0.3013,0.3082) | (0.3175,0.3246) |

**Figure 1: Distribution of Credit Scores for Subprime and U.S. Population: 2000-2002 and 2004-2006 cohorts**

Plot below show the cumulative distribution function (cdf) of FICO scores on originations during 2000-2002 (in red) and then during 2004-2006 (in black). The bold lines show the distribution of FICO scores on subprime originations. The dotted lines show the distribution of credit scores for the U.S. population with recorded credit histories.

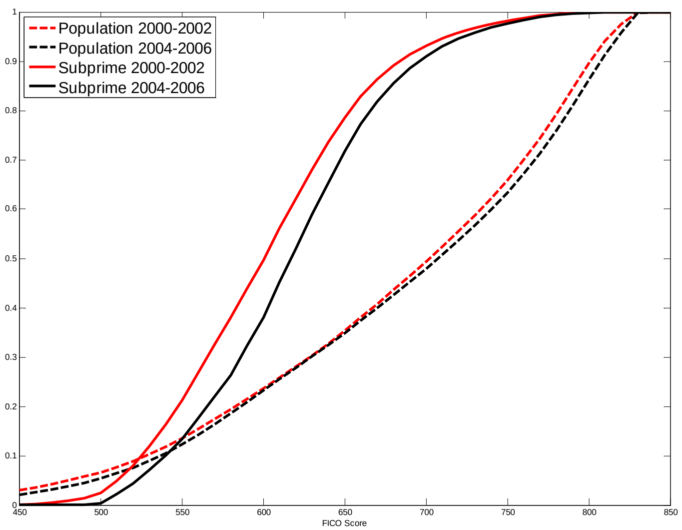


**Fig. 1A**

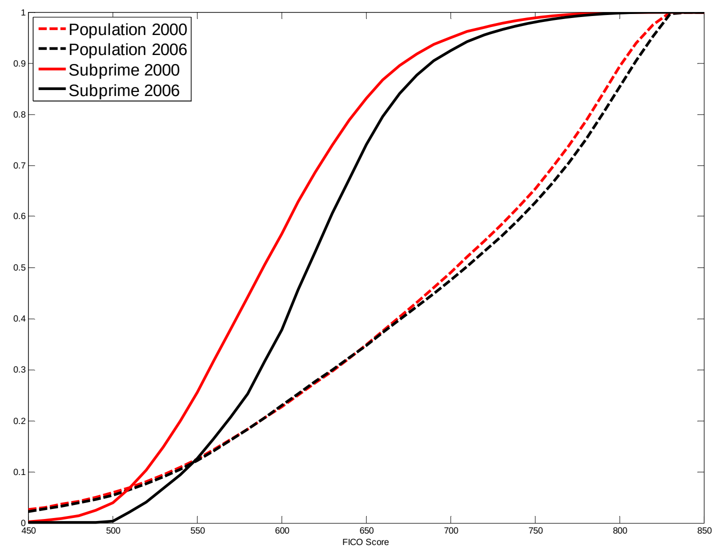


**Fig. 1B**

**Figure 2: Counterfactual analysis for 2004, 2005 and 2006 cohorts**

The figures show the estimated proportional hazard survivorship function for representative originations from different cohorts. The three plots correspond to the counterfactual exercises using survivor function estimates based on origination data from 2004, 2005 and 2006 respectively. In each of the three plots, C2000, C2001 and C2002 denote the counterfactual estimates for the cohorts 2000, 2001 and 2002 respectively.

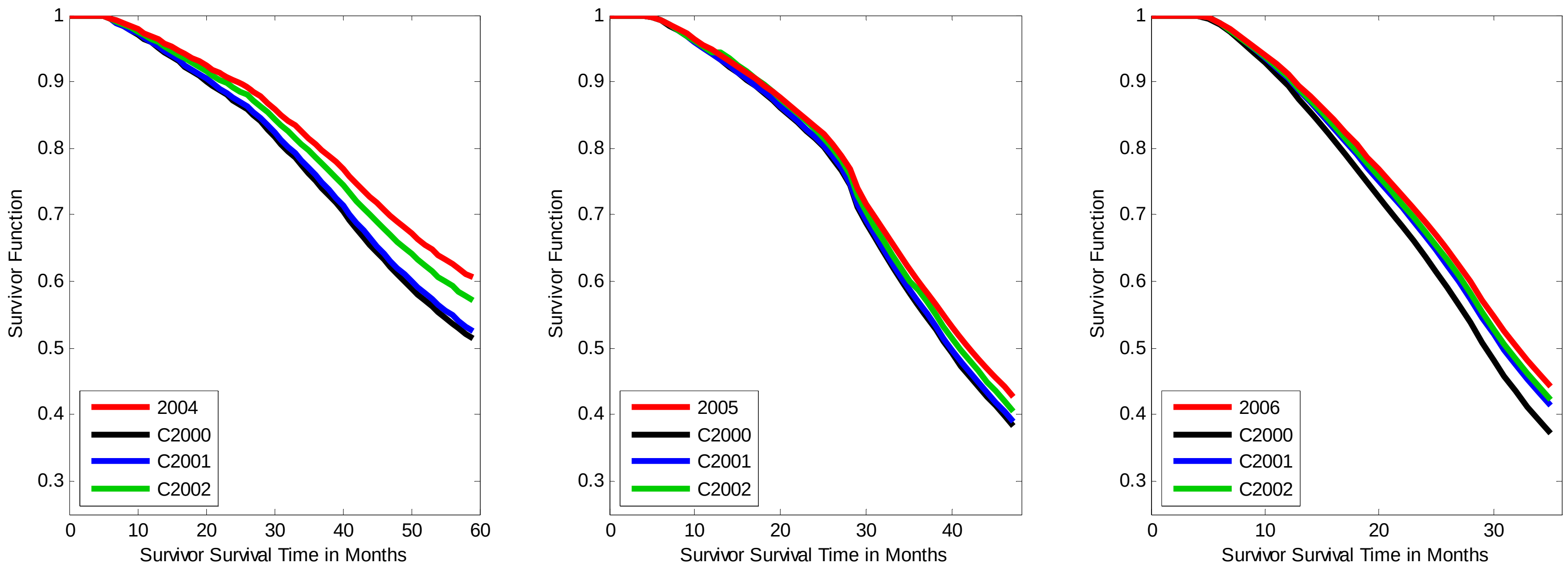

**Figure 3: Counterfactual analysis for 2004, different LTV ratios**

The plots below show the estimated proportional hazard survivorship function for representative originations from different cohorts. The three plots correspond to the counterfactual exercises using survivor function estimates based on origination data from 2004 for LTV<80 or C1, for LTV in [80, 90) or C2, for LTV≥90 or C3 respectively. In each of the three plots, C2000, C2001 and C2002 denotes the counterfactual estimates for the cohorts 2000, 2001 and 2002 respectively.

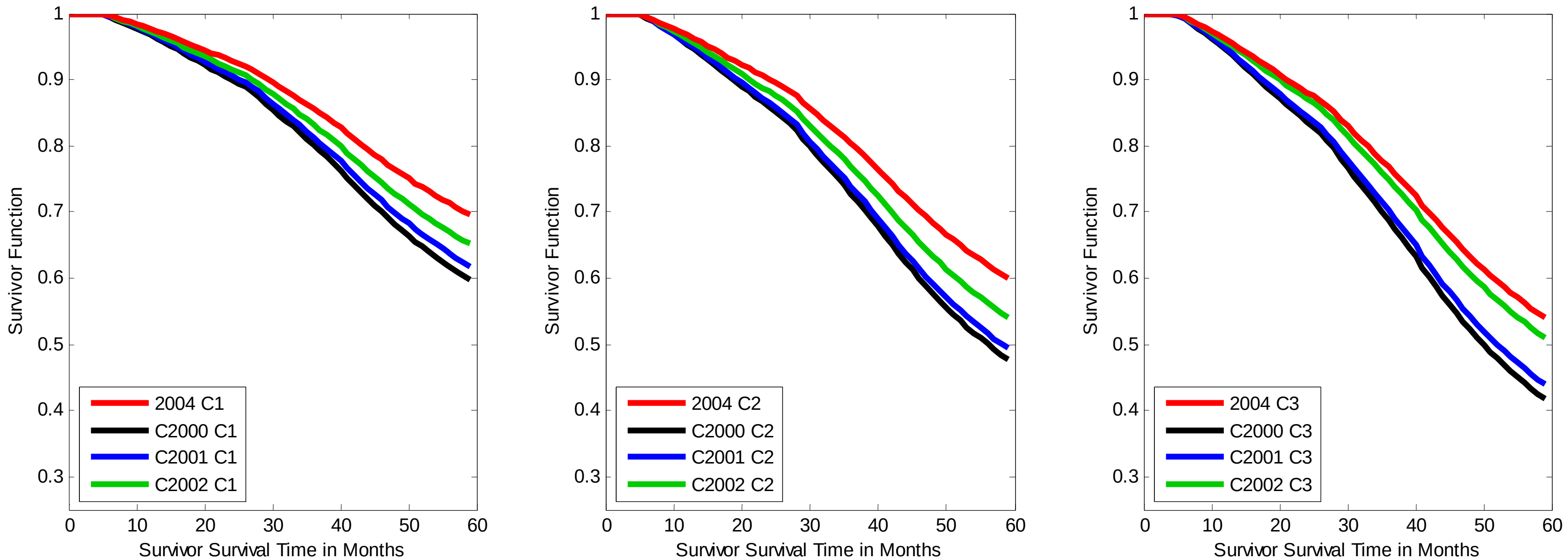

**Figure 4: Counterfactual analysis for 2005, different LTV ratios**

The plots below show the estimated proportional hazard survivorship function for representative originations from different cohorts. The three plots correspond to the counterfactual exercises using survivor function estimates based on origination data from 2005 for LTV<80 or C1, for LTV in [80, 90) or C2, for LTV≥90 or C3 respectively. In each of the three plots, C2000, C2001 and C2002 denotes the counterfactual estimates for the cohorts 2000, 2001 and 2002 respectively.

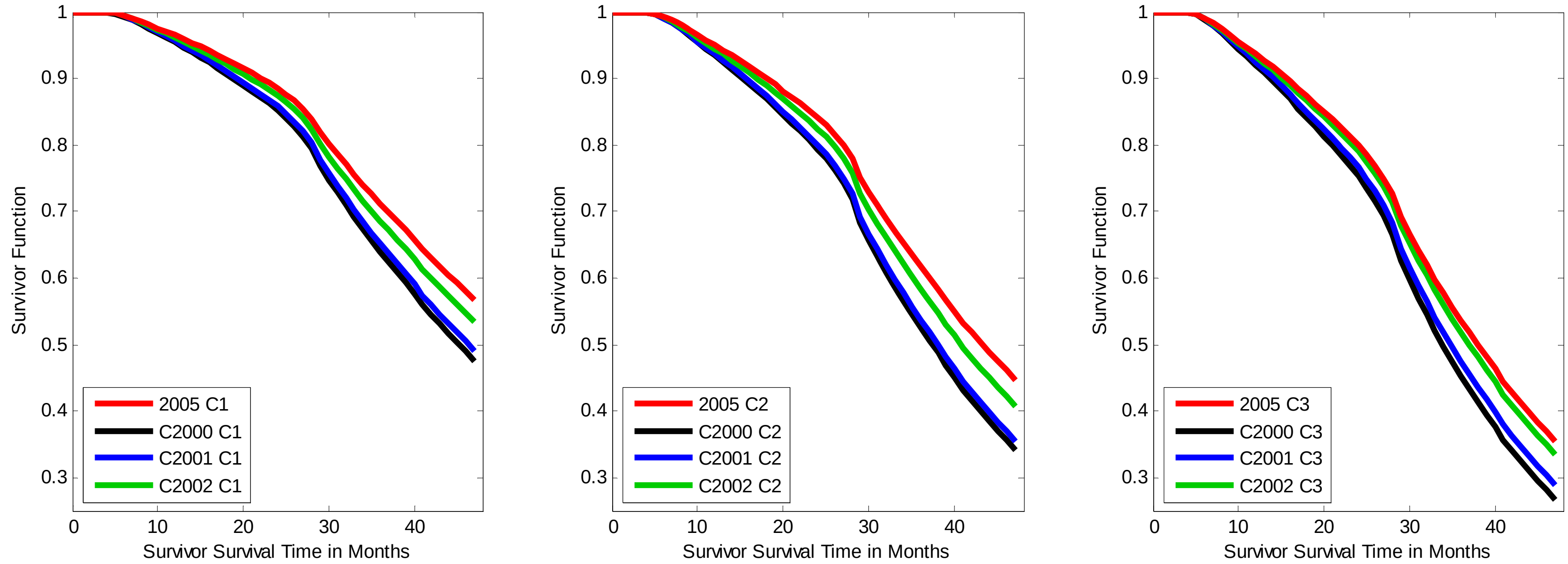

**Figure 5: Counterfactual analysis for 2006, different LTV ratios**

The plots below show the estimated proportional hazard survivorship function for representative originations from different cohorts. The three plots correspond to the counterfactual exercises using survivor function estimates based on origination data from 2006 for LTV<80 or C1, for LTV in [80, 90) or C2, for LTV≥90 or C3 respectively. In each of the three plots, C2000, C2001 and C2002 denotes the counterfactual estimates for the cohorts 2000, 2001 and 2002 respectively.

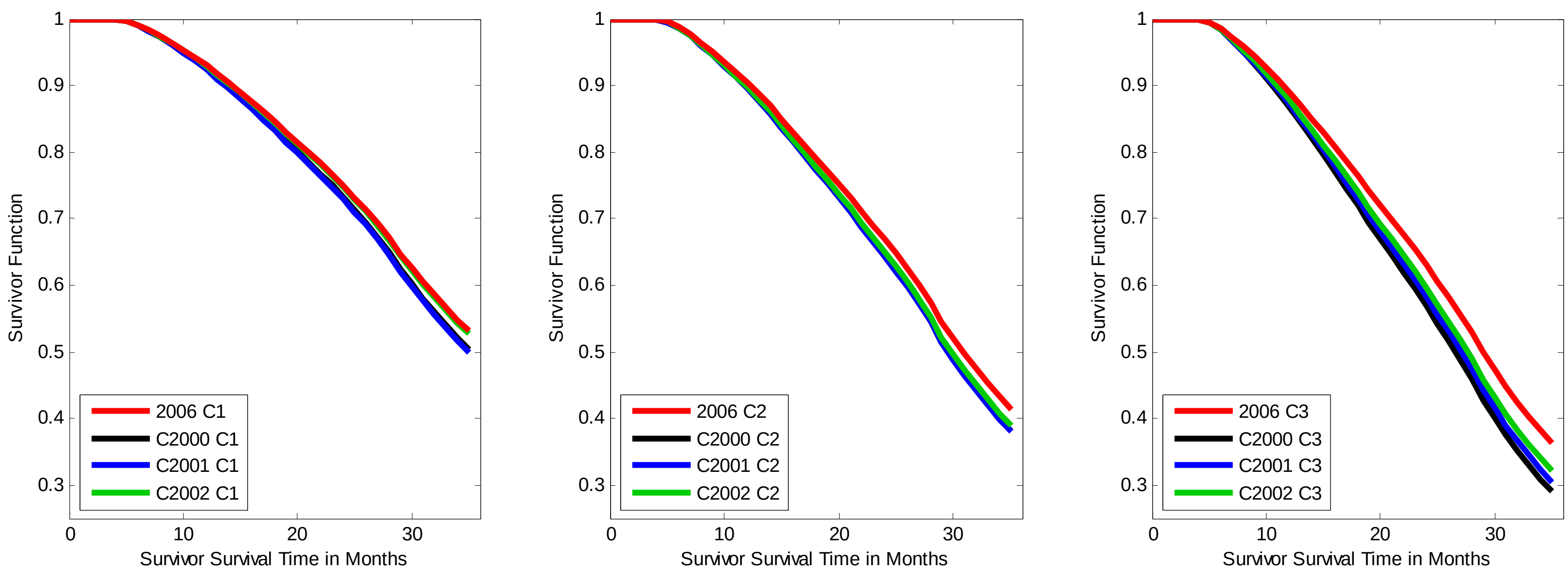

**Figure 6: Kernel Density of Origination FICO scores by LTV across different cohorts**

The figures below shows the kernel density plots of origination FICO scores for different LTV, less than 80 percent, between 80 and 90 percent, and 90 percent or above, respectively. Each plot shows the kernel densities for 2000-2002 cohorts (in black) and for 2004-2006 cohorts (in red) separately.