Understanding the Subprime Mortgage Crisis

Yuliya Demyanyk, Otto Van Hemert*

This Draft: February 4, 2008
First Draft: October 9, 2007

ABSTRACT

Using loan-level data, we analyze the quality of subprime mortgage loans by adjusting their performance for differences in borrower characteristics, loan characteristics, and house price appreciation since origination. We find that the quality of loans deteriorated for six consecutive years before the crisis and that securitizers were, to some extent, aware of it. We provide evidence that the rise and fall of the subprime mortgage market follows a classic lending boom-bust scenario, in which unsustainable growth leads to the collapse of the market. Problems could have been detected long before the crisis, but they were masked by high house price appreciation between 2003 and 2005.

*Demyanyk: Banking Supervision and Regulation, Federal Reserve Bank of St. Louis, P.O. Box 442, St. Louis, MO 63166, Yuliya.Demyanyk@stls.frb.org. Van Hemert: Department of Finance, Stern School of Business, New York University, 44 W. 4th Street, New York, NY 10012, ovanheme@stern.nyu.edu. The authors would like to thank Joost Driessen, William Emmons, Scott Frame, Xavier Gabaix, Ralph Koijen, Andreas Lehnert, Chris Mayer, Andrew Meyer, Matt Richardson, Stefano Risa, Bent Sorensen, Stijn Van Nieuwerburgh, James Vickery, Jeff Wurgler, and seminar participants at the Federal Reserve Bank of St. Louis, the Florida Atlantic University, the International Monetary Fund, the second New York Fed – Princeton liquidity conference, Lehman Brothers, and the Baruch-Columbia-Stern real estate conference. The views expressed are those of the authors and do not necessarily reflect the official positions of the Federal Reserve Bank of St. Louis or the Federal Reserve System.

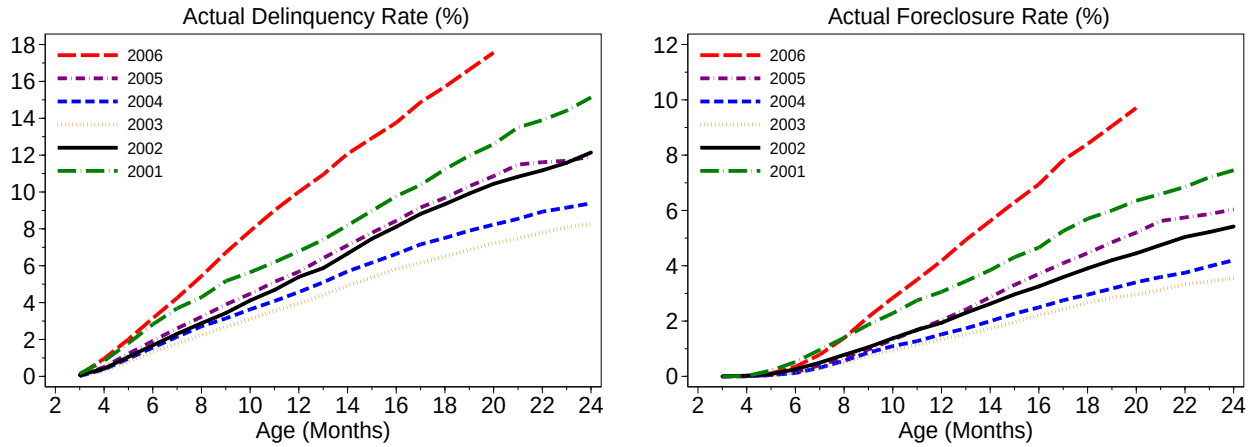
I. Introduction

The subprime mortgage crisis of 2007 was characterized by an unusually large fraction of subprime mortgages originated in 2006 being delinquent or in foreclosure only months later. The crisis spurred massive media attention; many different explanations of the crisis have been suggested. The goal of this paper is to answer the question: "What do the data tell us about the possible causes of the crisis?" To this end we use a loan-level database containing information on about half of all U.S. subprime mortgages originated between 2001 and 2006.

The relatively poor performance of vintage 2006 loans is illustrated in Figure 1. At every mortgage loan age, loans originated in 2006 show a higher delinquency rate (left panel) and a higher foreclosure rate (right panel) than loans originated in earlier years at the same ages. Note that 2001 was a fairly bad vintage year as well, ranking second, both in terms of the delinquency and the foreclosure rates.

Figure 1. Actual Delinquency and Foreclosure Rate

The figure shows the age pattern in actual delinquency and foreclosure rates for the different vintage years. Delinquency is defined as being 60 days or more late with the monthly mortgage payment, in foreclosure, or real-estate owned. Foreclosure is defined as being in foreclosure or real-estate owned. Once a foreclosure procedure for a loan is finalized and/or the loan balance becomes zero, this loan is dropped from the analysis.



We document that the poor performance of the vintage 2006 loans was not confined to a particular segment of the subprime mortgage market. For example, fixed-rate, adjustable-rate, purchase-money, cash-out refinancing, low-documentation, and full-documentation loans originated in 2006 all showed substantially higher delinquency and foreclosure rates than loans made the prior five years. This contradicts

a widely held belief that the subprime mortgage crisis was mostly confined to adjustable-rate or low-documentation mortgages.

We explore to what extent the subprime mortgage crisis can be attributed to different loan characteristics, borrower characteristics, and subsequent house price appreciation. The subsequent house price appreciation is measured as the MSA-level house price change between the period of origination and the period of loan performance evaluation. For the empirical analysis, we run logit regressions with the probability of either delinquency or foreclosure being a function of these factors.

We find that loan and borrower characteristics are very important in terms of explaining the cross-section of loan performance. However, because these characteristics were not sufficiently different in 2006 compared with the prior five years, they cannot explain the unusually weak performance of vintage 2006 loans. For example, a one-standard-deviation increase in the debt-to-income ratio raises the probability of delinquency 12 months after origination by as much as 0.73 percentage points. However, because the average debt-to-income ratio was only 0.15 standard deviations higher in 2006 than its level in previous years, it contributes very little to explain the inferior performance of vintage 2006 loans. The only variable in the considered logit regression model that contributed substantially to the crisis is the low subsequent house price appreciation for vintage 2006 loans, which can explain about a 1.1 percentage points higher-than-average delinquency rate 12 months after origination.[1] Due to geographical heterogeneity in house price changes, some areas have experienced larger-than-average house price declines and therefore have a larger explained increase in delinquency and foreclosure rates.[2]

We analyze the quality of loans based on their performance, adjusted for differences in observed loan characteristics, borrower characteristics, and subsequent house price appreciation. For the analysis, we compute the prediction error as the difference between the actual delinquency or foreclosure rate and the estimated probability of delinquency or foreclosure based on the logit regression model. In Figure 2 we plot the adjusted delinquency (left panel) and adjusted foreclosure (right panel) rates, which are obtained by adding up the prediction errors and the weighted average actual rates. This ensures having the same weighted average for the actual (Figure 1) and adjusted (Figure 2) delinquency and foreclosure rates.

As shown in Figure 2, adjusted delinquency and foreclosure rates have been steadily rising for the past six years, with no particularly large jump from 2005 to 2006. In other words, loan quality—adjusted for

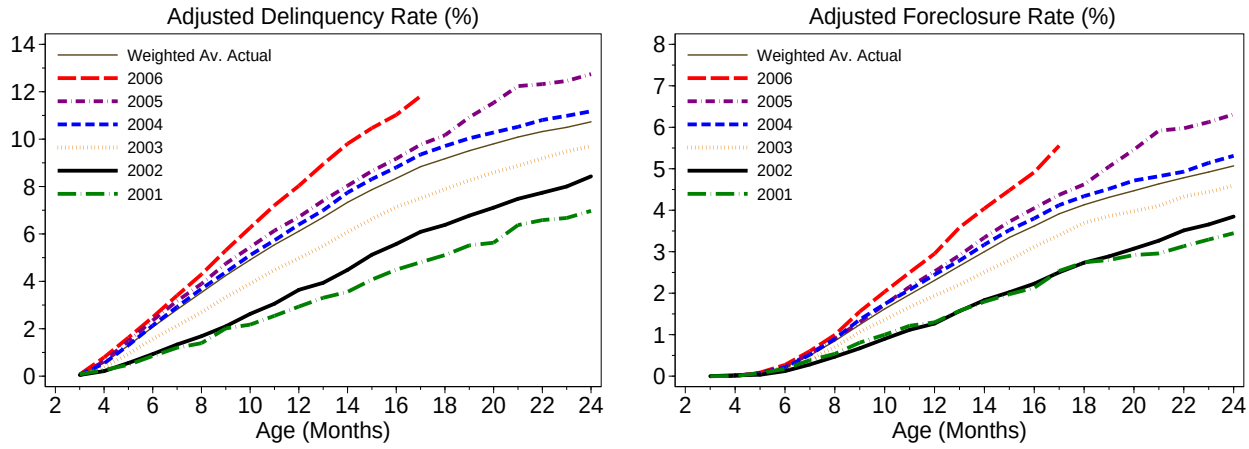
[1] Other papers that research the relationship between house prices and mortgage financing include Genesove and Mayer (1997), Genesove and Mayer (2001), and Brunnermeier and Julliard (2007).

[2] Also, house price appreciation may differ in cities versus rural areas. See for example Glaeser and Gyourko (2005) and Gyourko and Sinai (2006).

observed characteristics and subsequent house price appreciation—deteriorated monotonically between 2001 and 2006. Interestingly, 2001 was among the worst vintage years in terms of actual delinquency and foreclosure rates, but is in fact the best vintage year in terms of the adjusted rates. High interest rates, low average FICO credit scores, and low house price appreciation created the "perfect storm" in 2001, resulting in a high actual delinquency rate; after adjusting for these unfavorable circumstances, however, the adjusted delinquency rates are low.

Figure 2. Adjusted Delinquency and Foreclosure Rate

The figure shows the age pattern in the delinquency rate (left panel) and foreclosure rate (right panel) for the different vintages, after adjusting for variation in FICO scores, loan-to-value ratios, debt-to-income ratios, missing debt-to-income ratio dummies, cash-out refinancing dummies, owner-occupation dummies, documentation levels, percentage of loans with prepayment penalties, mortgage rates, margins, house price appreciation since origination, composition of mortgage contract types, and origination amounts.



In addition to the monotonic deterioration of loan quality, we show that over time the average combined loan-to-value ratio increased, the fraction of low documentation loans increased, and the subprime-prime rate spread decreased. The rapid rise and subsequent fall of the subprime mortgage market is therefore reminiscent of a classic lending boom-bust scenario.[3] The origin of the subprime lending boom has often been attributed to the increased demand for so-called private-label mortgage-backed securities (MBSs) by both domestic and foreign investors. Our database does not allow us to directly test this hypothesis,

[3]Berger and Udell (2004) discuss the empirical stylized fact that during a monetary expansion lending volume typically increases and underwriting standards loosen. Loan performance is the worst for those loans underwritten toward the end of the cycle. Demirgüç-Kunt and Detragiache (2002) and Gourinchas, Valdes, and Landerretche (2001) find that lending booms raise the probability of a banking crisis. Dell'Ariccia and Marquez (2006) show in a theoretical model that a change in information asymmetry across banks might cause a lending boom that features lower standards and lower profits. Ruckes (2004) shows that low screening activity may lead to intense price competition and lower standards.

but an increase in demand for subprime MBSs is consistent with our finding of lower spreads and higher volume.

The logit regression specification used to compute the adjusted delinquency and foreclosure rates assumes that the regression coefficients on the different explanatory variables remain constant over time. We test the validity of this assumption for all variables and find that it was the most strongly rejected for the loan-to-value (LTV) ratio. High-LTV borrowers in 2006 were riskier than those in 2001 in terms of the probability of delinquency or foreclosure, for given values of the other explanatory variables. In fact, the increases in the adjusted delinquency and foreclosure rates are almost exclusively caused by the worsening performance of loans with a combined LTV of 80 percent or more.

Were securitizers aware of the increasing riskiness of high-LTV borrowers?[4] To answer this question, we analyze the relationship between the mortgage rate and LTV ratio (along with the other loan and borrower characteristics). We perform a cross-sectional ordinary least squares (OLS) regression, with the mortgage rate as the dependent variable, for each quarter from 2001Q1 to 2007Q2 for both fixed-rate mortgages and 2/28 hybrid mortgages. Figure 3 shows that the coefficient on the first-lien LTV variable, scaled by the standard deviation of the first-lien LTV ratio, has been increasing over time. We thus find evidence that securitizers were aware of the increasing riskiness of high-LTV borrowers, and adjusted mortgage rates accordingly.

We show that our main results are robust when allowing for interaction effects between different loan and borrower characteristics. This includes taking into account risk-layering—the origination of loans that are risky in several dimensions, such as the combination of a high LTV ratio and a low FICO score. As an extension, we estimate our regression model using data just through year-end 2005 and again obtain the continual deterioration of loan quality since 2001. This means that the seeds for the crisis were sown long before 2007, but detecting them was complicated by high house price appreciation between 2003 and 2005—appreciation that masked the true riskiness of subprime mortgages.

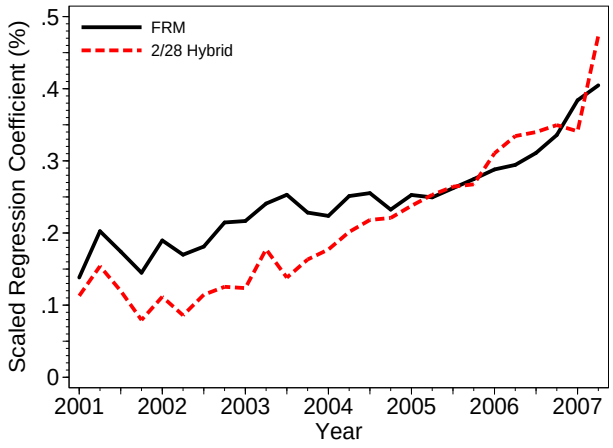
There is a large literature on the determinants of mortgage delinquencies and foreclosures, dating back to at least Von Furstenberg and Green (1974). Recent contributions include Cutts and Van Order (2005) and Pennington-Cross and Chomsisengphet (2007).[5] To the best of our knowledge, we are the first to quantify how much the different determinants have contributed to the observed high delinquency and

[4] For loans that are securitized (as are all loans in our database), the securitizer effectively dictates the mortgage rate charged by the originator.

[5] Deng, Quigley, and Van Order (2000) discuss the simultaneity of the mortgage prepayment and default option. Campbell and Cocco (2003) and Van Hemert (2007) discuss mortgage choice over the life cycle.

Figure 3. Sensitivity of Mortgage Rate to First-Lien Loan-to-Value Ratio

The figure shows the effect of the first-lien loan-to-value ratio on the mortgage rate for first-lien fixed-rate and 2/28 hybrid mortgages. The effect is measured as the regression coefficient on the first-lien loan-to-value ratio (scaled by the standard deviation) in an ordinary least squares regression with the mortgage rate as the dependent variable and the FICO score, first-lien loan-to-value ratio, second-lien loan-to-value ratio, debt-to-income ratio, missing debt-to-income ratio dummy, cash-out refinancing dummy, owner-occupation dummy, prepayment penalty dummy, origination amount, term of the mortgage, prepayment term, and margin (only applicable to 2/28 hybrid) as independent variables. Each point corresponds to a separate regression, with a minimum of 13,281 observations.



foreclosure rates for vintage 2006 loans. In addition, we uncover a trend in loan quality, determined by adjusting the delinquency and foreclosure rates for differences in loan characteristics, borrower characteristics, and subsequent house price appreciation. We provide evidence that the rise and fall of the subprime mortgage market follows a classic lending boom-bust scenario, in which unsustainable growth leads to the collapse of the market. Finally, we relate the rise in adjusted delinquency and foreclosure rates to the increasing riskiness of high-LTV borrowers.

The structure of this paper is as follows. In Section 2 we show the descriptive statistics for the subprime mortgages in our database. In Section 3 we present the econometric results and discuss explanatory factors for delinquency and foreclosure. In Section 4 we discuss the increasing riskiness of high-LTV borrowers, and the extent to which securitizers were aware of this risk. In Section 5 we analyze the subprime-prime rate spread and in Section 6 we provide robustness checks. In Section 7 we conclude.

II. Descriptive Analysis

In this paper we use and analyze a loan-level database that covers more than half of the U.S. subprime mortgage market.[6] There is no consensus on the exact definition of a subprime mortgage loan. The term subprime can be used to describe certain characteristics of the borrower (e.g., a FICO credit score less than 620),[7] lender (e.g., specialization in high-cost loans),[8] security of which the loan can become a part (e.g., high projected default rate for the pool of underlying loans), or mortgage contract type (e.g., no money down and no documentation provided). The common element across definitions of a subprime loan is a high default risk. In this paper, subprime loans are those underlying subprime securities. We do not include less-risky Alt-A mortgage loans in our analysis. We focus on first-lien loans and consider the 2001 through 2007 sample period.

We first discuss the main characteristics of the loans in our database at origination. Second, we discuss the delinquency and foreclosure rates of these loans for various segments of the subprime mortgage market.

A. Loan Characteristics at Origination

Table I provides the descriptive statistics for the subprime mortgage loans in our database that were originated between 2001 and 2006. The subprime mortgage market grew dramatically over this period. In the first block of Table I we see that over the sample period, the annual number of originated loans more than quadrupled and the average loan size almost doubled. The total dollar amount originated in 2001 was $94 billion. By 2006, originations had increased more than seven-fold, to $685 billion.

In the second block of Table I, we split the pool of mortgages into four main mortgage contract types. Most numerous are the hybrid mortgages, accounting for about half of all our subprime loans. A hybrid mortgage carries a fixed rate for an initial period (typically 2 or 3 years) and then the rate resets to a reference rate (often the 6-month LIBOR) plus a margin. The fixed-rate mortgage contract has become less popular in the subprime market over time and accounted for just 26 percent of the total number of loans in 2006. In contrast, in the prime mortgage market, most mortgage loans are of the fixed-rate type.[9]

[6] We use the First American LoanPerformance database, which contains about 85 percent of all securitized subprime mortgages. Mortgage Market Statistical Annual (2007) reports securitization shares of subprime mortgages each year from 2001 to 2006 equal to 54, 63, 61, 76, 76, and 75 percent respectively.

[7] The Board of Governors of the Federal Reserve System, The Office of the Controller of the Currency, the Federal Deposit Insurance Corporation, and the Office of Thrift Supervision use this definition.

[8] The U.S. Department of Housing and Urban Development uses HMDA data and interviews lenders to identify subprime lenders among them. There are, however, some subprime lenders making prime loans and some prime lenders originating subprime loans.

[9] For example Koijen, Van Hemert, and Van Nieuwerburgh (2007) show that the fraction of conventional, single-family,

Table I
Loan Characteristics at Origination for Different Vintages

Descriptive statistics for the first-lien subprime loans in the LoanPerformance database. We do not report other mortgage types, which amount to less than 0.1%.

	2001	2002	2003	2004	2005	2006
			Size			
Number of Loans (*1000)	624	974	1676	2743	3440	2646
Average Loan Size (*$1000)	151	168	180	201	234	259
			Mortgage Type			
FRM (%)	41.4	39.9	43.3	28.2	25.1	26.1
ARM (%)	0.9	1.9	1.3	4.3	10.3	12.8
Hybrid (%)	52.2	55.9	54.7	67.3	62.0	46.2
Balloon (%)	5.5	2.2	0.8	0.2	2.6	14.9
			Loan Purpose			
Purchase (%)	35.1	33.9	32.9	42.0	45.7	45.4
Refinancing (cash out) (%)	52.1	51.2	51.6	47.9	45.7	44.8
Refinancing (no cash out) (%)	12.3	14.6	15.1	10.0	8.6	9.8
			Variable Means			
FICO Score	620.1	630.5	641.4	645.9	653.7	654.7
Combined Loan-to-Value Ratio (%)	80.0	79.9	80.6	82.8	83.5	84.4
Debt-to-Income Ratio (%)	37.8	38.1	38.2	38.5	39.1	39.8
Missing Debt-to-Income Ratio Dummy (%)	41.6	44.1	38.3	35.1	39.2	31.7
Investor Dummy (%)	10.0	12.0	14.0	14.0	15.0	15.0
Documentation Dummy (%)	68.5	63.4	59.8	57.2	51.8	44.7
Prepayment Penalty Dummy (%)	66.3	63.8	61.4	60.1	60.6	61.6
Mortgage Rate (%)	9.4	8.3	7.3	6.7	6.6	7.2
Margin for ARM and Hybrid Mortgage Loans (%)	6.2	6.3	5.9	5.3	5.0	4.9

The proportion of balloon mortgage contracts jumped substantially in 2006, and accounted for 15 percent of the total number of mortgages originated that year. A balloon mortgage does not fully amortize over the term of the loan and therefore requires a large final (balloon) payment. About 13 percent of the mortgages originated in 2006 were adjustable-rate (non-hybrid) mortgages.

In the third block of Table I, we report the purpose of the mortgage loans. In about 30 to 45 percent of cases, the purpose is to finance the purchase of a house. Approximately 50 percent of our subprime mortgage loans were originated to extract cash, by refinancing an existing mortgage loan into a larger new mortgage loan. Cash-out refinancing loans were particularly prevalent between 2001 and 2003. The share of loans originated in order to refinance with no cash extraction is relatively small.

In the final block of Table I, we report the mean values for the variables that we will use in the regression analysis (see Table II for a definition of these variables). The average FICO credit score actually rose over the sample period. The combined loan-to-value (CLTV) ratio, which measures the value of all-lien loans divided by the value of the house, slightly increased over time, primarily because of the increased popularity of second-lien and third-lien loans. The (back-end) debt-to-income ratio (if provided) and the fraction of loans with a prepayment penalty were fairly constant. For about a third of the loans in our database, no debt-to-income ratio was provided; this is captured by the missing debt-to-income ratio dummy variable. The share of loans with full documentation fell sharply over the sample period, from 69 percent in 2001 to 45 percent in 2006.

The mean mortgage rate fell substantially from 2001 to 2004, consistent with sharp declines in both the 1-year and 10-year Treasury yields over the same period. Between 2004 and 2006, both the 1-year and 10-year Treasury yield increased again, but the mortgage rate on subprime mortgages, remarkably, remained fairly constant. Finally, the margin (over a reference rate) for adjustable-rate and hybrid mortgages declined from 2001 to 2004 and remained fairly constant afterward.

We do not report summary statistics on the loan source, such as whether a mortgage broker intermediated, as the broad classification used in the database rendered this variable less informative.

fully amortizing, purchase-money loans reported by the Federal Housing Financing Board in its Monthly Interest Rate Survey that are of the fixed-rate type fluctuated between 60 and 90 percent from 2001 to 2006. Vickery (2007) shows that empirical mortgage choice is affected by the eligibility of the mortgage loan to be purchased by Fannie Mae and Freddie Mac.

B. Performance of Loans by Market Segments

In Figure 1 we showed that for the subprime mortgage market as a whole, vintage 2006 loans stand out in terms of high delinquency and foreclosure rates (for variable definitions, see Table II). In Figure 4, we again plot the age pattern in the delinquency rate for vintages 2001 through 2006 and split the subprime mortgage market into various segments. As the figure shows, the poor performance of the 2006 vintage is not confined to a particular segment of the subprime market, but rather reflects a market-wide phenomenon.

In the lower four panels of Figure 4 we see that for purchase-money, cash-out refinancing, low-documentation, and full-documentation mortgage loans, the 2006 vintage shows the highest delinquency rate pattern. In general, vintage 2001 loans show the second-highest delinquency rates, and vintage 2003 loans have the lowest delinquency rates.

In the upper right panel, we see that for adjustable-rate mortgages (ARMs), the 2006 vintage clearly stands out with higher delinquency rates, while 2001 again ranks as the second worst vintage year. In the upper left panel of Figure 4 we see that, in general, the delinquency rate for fixed-rate mortgages (FRMs) is lower than that for ARMs. As with the ARMs in our sample, the delinquency rate for FRMs originated in 2006 is substantially higher than the delinquency rates of the loans originated between 2001 and 2005.

In Figure 5 we plot the delinquency and foreclosure rates of all *outstanding* mortgages. Notice that the fraction of FRMs that are delinquent or in foreclosure remained fairly constant at about five percent from 2005 on. These rates are consistent with those used in an August 2007 speech by the Chairman of the Federal Reserve System (Bernanke (2007)), who said "For subprime mortgages with fixed rather than variable rates, for example, serious delinquencies have been fairly stable at about 5-1/2 percent." It is important, though, to realize that this result is driven by an aging effect of the FRM pool, caused by a recent decrease in the popularity of FRMs (see Table I). In other words, FRMs originated in 2006 in fact performed unusually poorly (Figure 4, upper-left Panel), but if one plots the delinquency rate of outstanding FRMs over time (Figure 5, left Panel), the weaker performance of vintage 2006 loans is masked by the aging of the overall FRM pool.

Figure 4. Actual Delinquency Rate for Segments of the Subprime Mortgage Market

The figure shows the age pattern in delinquency rate for the different vintages. Each of the six panels focuses on a different segment of the subprime mortgage market.

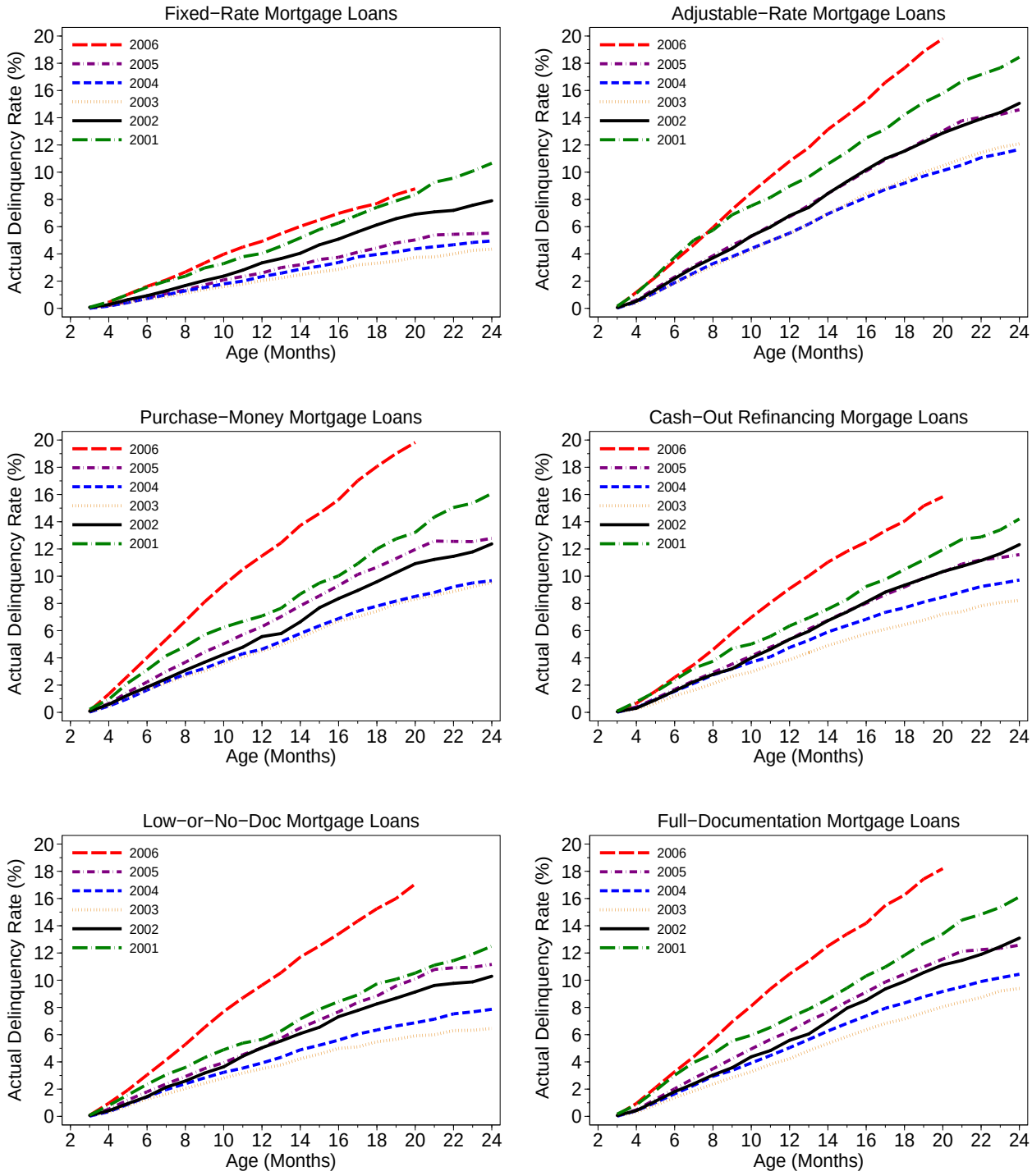
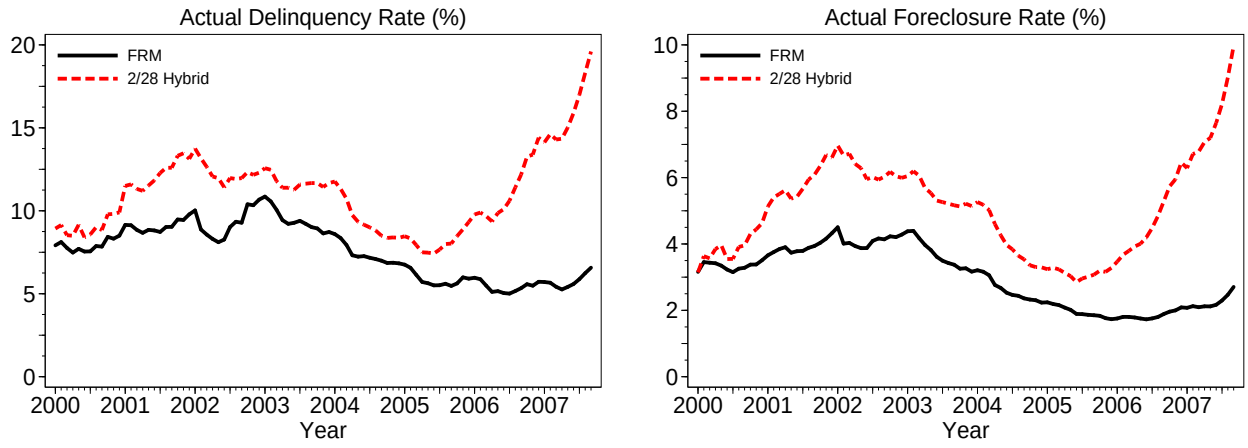


Figure 5. Actual Delinquency and Foreclosure Rates of Outstanding Mortgages

The Figure shows the actual delinquency and foreclosure rates of all outstanding FRMs and hybrids from January 2000 through September 2007.



III. Empirical Analysis of Delinquency and Foreclosure Determinants

In this section we investigate to what extent a logit regression model can explain the high levels of delinquencies and foreclosures for the vintage 2006 mortgage loans in our database. The regression coefficients are assumed to be constant over time, which allows us to interpret the time variation in the regression error term. All results in this section will be based on a random sample of one million first-lien subprime mortgage loans, originated between 2001 and 2006.

A. Empirical Model Specification

We run the following logit regression

$$\Pr(event) = \Phi(\beta' X), \tag{1}$$

where the *event* is either delinquency or foreclosure of a subprime mortgage loan after a given number of months; $\Phi(x) = 1/(1+\exp(-x))$ is the logit function; X is the vector of explanatory variables, including a constant; and β is the vector of regression coefficients. We will report the following statistics for each

explanatory variable i:

$$
\begin{aligned}
marginal_i &= \Phi(\beta'\bar{X} + \beta_i\sigma_i) - \Phi(\beta'\bar{X}) && (2)\\
deviation06_i &= (\overline{X06}_i - \bar{X}_i)/\sigma_i && (3)\\
contribution06_i &= \Phi(\beta'\bar{X} + \beta_i(\overline{X06}_i - \bar{X}_i)) - \Phi(\beta'\bar{X}) && (4)\\
&\approx marginal_i \times deviation06_i && (5)
\end{aligned}
$$

where $\bar{X}$ is the vector with mean values, σ_i is the standard deviation of the i-th variable, and $\overline{X06}_i$ is the mean value of the i-th variable for vintage 2006 loans. We define *deviation*01 and *contribution*01 for vintage 2001 loans in a similar fashion. Equation (5) emerges from a first-order Taylor approximation with the derivative of the logit function with respect to the i-th variable approximated by $marginal_i$.[10] The *marginal* statistic measures the effect of a one-standard-deviation increase in a variable (from its mean) on the probability of an event. The *deviation* statistic measures the number of standard deviations that the mean value of a variable in 2006/2001 was different from the mean value measured over the entire sample. The *contribution* statistic measures the deviation of the (average) event probability in 2006/2001 from the (average) event probability over the entire sample that can be explained by a particular variable.

For any subgroup of loans, such as a particular vintage, we can determine the predicted probability of an event by computing:

$$predicted = \sum_{j=1}^{L} \Phi(\beta' X^j)/L, \qquad (6)$$

where the superscript j refers to the loan number and L is the total number of loans in the subgroup.

B. Variable Definitions

Table II provides the definitions of the dependent and independent (explanatory) variables used in the empirical analysis. We use either the delinquency or the foreclosure dummy variable as the dependent variable. We define a loan to be delinquent if payments on the loan are 60 or more days late, the loan is in foreclosure, or the associated house is real estate owned by a lender. We define a loan to be in foreclosure if the reported status is in foreclosure or real estate owned by a lender. Once a foreclosure procedure for a loan is finalized and/or the loan balance becomes zero, we drop this loan from our analysis.

[10] Technically, we first change units by multiplying by σ_i in Equation (2) and diving by σ_i in Equation (3).

Table II: Variable Definitions

This table presents definitions of the variables used in the regression analyses. The first two variables are used as dependent variables. The other variables are used as independent variables. We report the expected sign for the independent variables in parentheses and sometimes provide a brief motivation.

Variable (Expected Sign)	Explanation
Delinquency Dummy	Payments on the loan are 60 or more days late, the loan is in foreclosure, or the loan is real estate owned.
Foreclosure Dummy	The loan is in foreclosure or is real estate owned.
FICO Score (-)	Fair, Isaac and Company (FICO) credit score at origination.
Combined Loan-to-Value Ratio (+)	Combined values of all liens divided by the value of the house at origination. A higher combined loan-to-value ratio makes default more attractive.
Debt-to-Income Ratio (+)	Back-end debt-to-income ratio, defined by the total monthly debt payments divided by the gross monthly income, at origination. A higher debt-to-income ratio makes it harder to make the monthly mortgage payment.
Missing Debt-to-Income Dummy (+)	Equals one if the back-end debt-to-income ratio is missing and zero if provided. We expect the lack of debt-to-income information to be a negative signal on borrower quality.
Cash-Out Dummy (-)	Equals one if the mortgage loan is a cash-out refinancing loan. Pennington-Cross and Chomsisengphet (2007) show that the most common reasons to initiate a cash-out refinancing are to consolidate debt and to improve property.
Investor Dummy (+)	Equals one if the borrower is an investor and does not owner-occupy the property.
Documentation Dummy (-)	Equals one if full documentation on the loan is provided and zero otherwise. We expect full documentation to be a positive signal on borrower quality.
Prepayment Penalty Dummy (+)	Equals one if there is a prepayment penalty and zero otherwise. We expect that a prepayment penalty makes refinancing less attractive.
Mortgage Rate (+)	Initial interest rate as of the first payment date. A higher interest rate makes it harder to make the monthly mortgage payment.
Margin (+)	Margin for an adjustable-rate or hybrid mortgage over an index interest rate, applicable after the first interest rate reset. A higher margin makes it harder to make the monthly mortgage payment.
House Price Appreciation (-)	MSA-level house price appreciation from the time of loan origination, reported by the Office of Federal Housing Enterprise Oversight (OFHEO). Higher housing equity leads to better opportunities to refinance the mortgage loan.
Product Type Dummies (+)	We consider four product types: FRMs, Hybrids, ARMs, and Balloons. We include a dummy variable for the latter three types, which therefore have the interpretation of the probability of delinquency or foreclosure relative to FRM. Because we expect the FRM to be chosen by more risk-averse and prudent borrowers, we expect positive signs for all three product type dummies.
Origination Amount (?)	Size of the mortgage loan. We have no clear prior on the effect of the origination amount on the probability of foreclosure and delinquency, holding constant the loan-to-value and debt-to-income ratio.

The borrower and loan characteristics we use in the analysis are: the FICO credit score, the combined loan-to-value ratio, the value of the debt-to-income ratio (when provided), a dummy variable indicating whether the debt-to-income ratio was missing, a dummy variable indicating whether the loan was a cash-out refinancing, a dummy variable indicating whether the borrower was an investor (as opposed to an owner-occupier), a dummy variable indicating whether full documentation was provided, a dummy variable indicating whether there is a prepayment penalty on a loan, the (initial) mortgage rate, and the margin for adjustable-rate and hybrid loans.[11]

In addition, we construct a variable that measures house price appreciation from the time of origination until the time we evaluate whether the loan is delinquent or in foreclosure. To this end we use metropolitan statistical area (MSA) level house price indexes from the Office of Federal Housing Enterprise Oversight (OFHEO) and match loans with MSAs by using the zip code provided by LoanPerformance.

We also considered the change in the unemployment rate from the period of origination until the period of loan performance evaluation, which we could only measure accurately at the state-level for the entire sample. It turned out that the unemployment variable mainly picked up the time trend in the delinquency or foreclosure rate. The relationship between the (trending) unemployment rate and the (trending) loan performance, however, is spurious. When vintage dummy variables are included in the regression, the unemployment rate becomes insignificant, both statistically and economically. We therefore decided to omit the change in the unemployment rate as an explanatory variable.

In Table II we report the expected sign for the regression coefficient on each of the explanatory variables in parentheses.

C. Results

Table III shows the results of the logit regression described in Section III A with the delinquency (panel A) or foreclosure (panel B) dummy variables 12 months after origination as the dependent variables. The explanatory variables are listed in the first column. Focusing on the marginal effect of the regression coefficients, defined in Equation 2, we see that the mortgage rate is the most important factor explaining cross-sectional differences in loan performance. The positive sign means that a higher mortgage rate increases the probability of delinquency (panel A) and foreclosure (panel B). Based on the point estimate

[11] We also studied specifications that included loan purpose, reported in Table I, and housing outlook, defined as the house price accumulation in the year prior to the loan origination. These variables were not significant and did not materially change the regression coefficients on the other variables.

in panel A, a one standard deviation increase in the mortgage rate increases the probability of delinquency by 1.9 percentage points 12 months after origination.

The second most important marginal effect is associated with the FICO score. A higher FICO score decreases the probability of delinquency (panel A) and foreclosure (panel B), as one would expect. In general all the variables have the expected signs (as discussed in Table II) and all variables are statistically significant at the 1% level. We also experimented with several interaction and quadratic terms, which yielded very similar results, both qualitatively and quantitatively (see also Section VI A).

In panel C of Table III we report the "Deviation," defined as the difference between the average value of a variable in 2001/2006 and the average value over the whole sample, expressed in standard deviations (see Equation 3). In particular, the average mortgage rate was high in 2001 and the average house price appreciation was low in 2006, both compared with the average values over the whole sample period.

In panels A and B, "Contribution" measures to what extent a variable can explain why the delinquency and foreclosure rates in 2001 or in 2006 differed from the average over the entire period. We see that the high mortgage rate for vintage 2001 loans was quantitatively the most important factor explaining high observed delinquency and foreclosure rates for that year. Additionally, the low average FICO score for 2001 loans and the low house price appreciation in the months following origination help explain the poor performance of 2001. One could say that in 2001 we had "the perfect storm."

For vintage 2006 loans, low subsequent house price appreciation, in particular, contributed to their weak performance, and accounted for a 1.1 percentage point increase in delinquency rate and a 0.6 percentage point increase in foreclosure rate, 12 months after origination. The mean values in 2006 for the other variables were not sufficiently different from the sample mean to contribute much to a different delinquency or foreclosure rate in 2006. It is worth noting that the increase in the average CLTV ratio and the decrease in the fraction of loans with full documentation over time does not contribute a lot to the high delinquency and foreclosure rates in 2006.

We also computed the contributions of all explanatory factors for the other vintage years (the results are not reported). For loans originated in 2003 and 2004, the high subsequent house price appreciation between 2003 and 2005 contributed to a lower actual delinquency rate. For example, the explained change in the delinquency rate 12 months after origination was –0.3 percentage points and –0.6 percentage points for 2003 and 2004, respectively. The house price appreciation variable had the largest (absolute) contribution among all variables considered for those years. Therefore, we can say that high house price

Table III: Determinants of Delinquency and Foreclosure

The table shows the output of the the logit regression defined in Equation 1, where the event is that a loan is delinquent (panel A) or in foreclosure (panel B), 12 months after origination. The first column reports the explanatory variables (constant not reported). In panels A and B we report the marginal effect of a variable, defined in Equation 2, and the contribution of a variable to explain a different probability of delinquency or foreclosure in 2001/2006, defined in Equation 4. Panel C reports the deviation of the 2001/2006 value of a variable from the average over 2001–2006, defined in Equation 3. We have the first-order approximation *contribution* $\approx$ *marginal* $\times$ *deviation*. All coefficients are statistically significant at the 1% level.

	Panel A: Delinquency Rate			**Panel B: Foreclosure Rate**			**Panel C: Deviations**	
Explanatory Variable	Marginal Effect (%)	Contribution 2001 (%)	Contribution 2006 (%)	Marginal Effect (%)	Contribution 2001 (%)	Contribution 2006 (%)	Deviation 2001	Deviation 2006
Fico Score	−1.75	1.03	−0.26	−0.45	0.24	−0.06	−0.37	0.11
Combined Loan-to-Value Ratio	1.60	−0.26	0.16	0.58	−0.09	0.06	−0.20	0.12
Debt-to-Income Ratio	0.73	−0.07	0.10	0.19	−0.02	0.03	−0.11	0.15
Missing Debt-to-Income Ratio	0.57	0.04	−0.06	0.11	0.01	−0.01	0.08	−0.12
Cash-Out Dummy	−0.32	−0.04	0.02	−0.12	−0.01	0.01	0.11	−0.07
Investor Dummy	0.37	−0.04	0.01	0.18	−0.02	0.00	−0.12	0.03
Documentation Dummy	−0.68	−0.22	0.15	−0.25	−0.08	0.06	0.30	−0.20
Prepayment Penalty Dummy	0.10	0.01	0.00	0.05	0.01	−0.00	0.12	−0.01
Mortgage Rate	1.90	2.22	0.10	0.94	1.11	0.05	1.14	0.07
Margin	0.68	−0.07	−0.01	0.23	−0.02	−0.00	−0.11	−0.02
House Price Appreciation	−1.01	0.62	1.10	−0.44	0.31	0.56	−0.48	−0.80
Hybrid Dummy	0.07	−0.01	−0.02	0.08	−0.01	−0.02	−0.12	−0.23
ARM Dummy	0.04	−0.01	0.01	0.00	0.00	−0.00	−0.23	0.20
Balloon Dummy	0.21	0.01	0.11	0.09	0.01	0.05	0.06	0.52
Origination Amount	0.77	−0.26	0.20	0.32	−0.10	0.08	−0.39	0.27

appreciation between 2003 and 2005 masked the true riskiness of subprime mortgages.[12]

The product type has a relatively small effect on the performance of a loan, beyond what is explained by other characteristics (see Table III). In Figure 4 we showed that FRMs experience a much lower delinquency rate than ARMs, which must be driven by borrowers with better characteristics selecting into FRMs.[13]

To examine to what extent the logit regression model is capable of explaining the large observed delinquency and foreclosure rates in 2006, we compare the adjusted delinquency and foreclosure rates for different ages and different vintages in Figure 2. The adjusted rate at a given age equals the prediction error (the actual rate minus the predicted rate) plus the weighted average rate over the 2001–2006 period, with weights equal to the number of loans originated in each year. The predicted delinquency and foreclosure rates are determined using Equation 6. We add up the weighted-average actual rates to facilitate the comparison with the actual rates plotted in Figure 1. Interestingly, both the adjusted delinquency and foreclosure rates have been increasing over the past six years. In other words, loan quality deteriorated monotonically between 2001 and 2006. This picture is in sharp contrast with that obtained from actual rates, depicted in Figure 1, where 2003 was the year with the lowest delinquency and foreclosure rates, and 2001 was the year with the second-highest rates.

IV. Non-Stationarity of the Loan-to-Value Effect

The logit regression specification used in Section III assumes that the regression coefficients are constant over time. That is, the effect of a unit change in an explanatory variable on the delinquency or the foreclosure rate is the same in, for example, 2006 as it is in 2001, holding constant the values of the other explanatory variables. We test the validity of this assumption for all variables in our analysis by running cross-sectional OLS regressions for each calendar month from 2001 to 2006 and checking the stability of the regression coefficients. It turns out that the strongest rejection of a constant regression coefficient is for the CLTV ratio. In this section we first discuss this finding and then turn to the question of whether lenders were aware of the non-stationarity of the loan-to-value effect, by investigating the relationship between the loan-to-value ratio and mortgage rates over time.

[12] Shiller (2007) argues that house prices were too high compared to fundamentals in this period and refers to the house price boom as a classic speculative bubble largely driven by an extravagant expectation for future house price appreciation.

[13] Consistent with this finding, LaCour-Little (2007) shows that individual credit characteristics are important for mortgage product choice.

A. Loan-to-Value Ratio and the Delinquency Rate

Figure 6 shows the distribution of the CLTV ratio for all first-lien loans. About 30 percent have a CLTV smaller than 80 percent, about 20 percent have a CLTV of exactly 80 percent, and about 50 percent have a CLTV greater than 80 percent. The average CLTV increased slightly from 80 percent in 2001 to 84 percent in 2006 (see Table I). In addition, the distribution shifted slightly to the right over time: In 2001 the percentages of loans in these three CLTV categories was 35, 20, and 45 percent, respectively; in 2006 they were 28, 14, and 58 percent, respectively.

Figure 6. Cumulative Distribution of the Combined Loan-to-Value Ratio (CLTV)

The figure shows the cumulative distribution of the combined loan-to-value ratio for all first-lien subprime mortgage loans originated between 2001 and 2006. Loans with a combined loan-to-value ratio greater than 100 account for less than 0.5% of all loans and are omitted from the Figure.

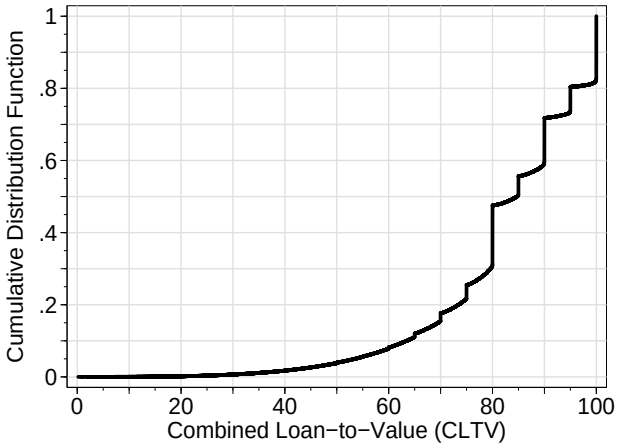


Table IV reports the actual minus the predicted delinquency rate for different CLTV value categories and different vintage years, estimated using Equation 6. In other words, the table reports the average prediction error for the three CLTV subgroups discussed above and for each origination year of loans. A positive prediction error means that the actual delinquency rate was higher than the rate predicted by the logit regression model. Consistent with Figure 2, the error increased over time. However, for the low CLTV group, the increase in the error was much smaller and, in fact, had been slightly decreasing from 2004 onward. For a CLTV ratio of 80 percent, the increase in the error was 4.3 percent, and for the CLTV ratio above 80 percent, the increase was 7.1 percent. Therefore, high CLTV ratios were increasingly associated with higher delinquency rates, beyond what is captured by the logit regression model.

Table IV
Actual Minus Predicted Delinquency Rate

This table presents the actual minus the predicted delinquency rate 12 months after origination for different vintages and combined loan-to-value (CLTV) ratios.

	2001	2002	2003	2004	2005	2006	2006–2001
$CLTV < 80\%$	−1.7%	−1.4%	−0.4%	0.6%	0.4%	0.5%	2.2%
$CLTV = 80\%$	−2.7%	−2.1%	−0.7%	0.4%	1.0%	1.6%	4.3%
$CLTV > 80\%$	−4.4%	−3.3%	−1.6%	0.2%	0.7%	2.7%	7.1%

B. Loan-to-Value Ratio and the Mortgage Rate

The combined LTV ratio rather than the first-lien LTV ratio is believed to be the main determinant of delinquency and foreclosure, because it is the burden of all the debt together that may trigger financial problems for the borrower. In contrast, the first-lien LTV is the more important determinant of the mortgage rate on a first-lien mortgage, because it captures the dollar amount at stake for the first-lien lender.[14]

In this subsection we try to determine whether lenders were aware that high LTV ratios were increasingly associated with riskier borrowers. Specifically, we test whether the sensitivity of the lender's interest rate to the first-lien LTV ratio changed over time. We perform a cross-sectional OLS regression with the mortgage rate as the dependent variable and loan characteristics, including the first-lien LTV and second-lien LTV (CLTV minus first-lien LTV), as independent variables.[15] We perform one such regression for each calendar quarter in our sample period. We can only expect to get accurate results when using relatively homogeneous groups of loans, and therefore consider fully amortizing FRM and 2/28 hybrid loans separately. Together these two contract types account for about half of all mortgage loans in our database. Each cross-sectional regression is based on a minimum of 13,281 observations.

Figure 3 shows the regression coefficient on the first-lien LTV ratio for each quarter from 2001Q1 through 2007Q2.[16] We scaled the coefficients by the standard deviation of the first-lien LTV ratio, and

[14]This is confirmed by our empirical results. To conserve space the results are not reported.

[15]Specifically, we use the FICO score, first-lien loan-to-value ratio, second-lien loan-to-value ratio, debt-to-income ratio, a dummy for a missing debt-to-income ratio, a cash-out refinancing dummy, a dummy for owner occupation, documentation dummy, prepayment penalty dummy, margin, origination amount, term of the mortgage, and prepayment term as the right-hand-side variables.

[16]Our data extends to 2007Q3, but due to a near shutdown of the securitized subprime mortgage market we lack statistical

they can therefore be interpreted as the changes in the mortgage rates when the first-lien LTV ratios are increased by one standard deviation. In the fourth quarter of 2006, a one-standard-deviation increase in the first-lien LTV ratio corresponded to about a 35-basis-point increase in the mortgage rate for both FRMs and 2/28 hybrid mortgages, keeping constant other loan characteristics. In contrast, in the first quarter of 2001, the corresponding rate increase was around 13 basis points. This provides evidence that lenders were to some extent aware of high LTV ratios being increasingly associated with risky borrowers.[17] Finally, notice that the effect of a one-standard-deviation increase in the first-lien LTV ratio on the 2/28 mortgage rate increased substantially in the wake of the subprime mortgage crisis: from 34 basis points in 2007Q1 to 47 basis points in 2007Q2.

V. Subprime-Prime Rate Spread

In general, interest rates on subprime mortgages are higher than on prime mortgages to compensate the lender for the (additional) default risk associated with subprime loans. In this section we analyze the time series of the subprime-prime rate spread, both with and without adjustment for changes in loan and borrower characteristics. We focus on fixed-rate mortgages for this exercise. For hybrid mortgages the subprime-prime comparison is more complicated because (i) both the initial (teaser) rate and the margin should be factored in, and (ii) we don't have good data on the prime initial rates and margins.

In Figure 7 we show the actual subprime-prime rate spread. The subprime rate is from the LoanPerformance database, the prime rate is the contract rate on fixed-rate mortgages reported by the Federal Housing Finance Board (FHFB) in its Monthly Interest Rate Survey.[18] The subprime-prime spread decreased substantially over time, with the largest decline between 2001 and 2004, which coincides with the most rapid growth in the number of loans originated (see Table I). In Figure 7 we also plot the yield spread between 10-year BBB and AAA corporate bonds, which we obtained from Standard and Poor's Global Fixed Income Research. Compared to the corporate BBB-AAA yield spread, the actual subprime-prime rate spread declined much more and more steadily, hence the decline cannot just be attributed to a change in the overall level of risk aversion.

We perform an OLS regression with the spread as dependent variable and the prime rate and various

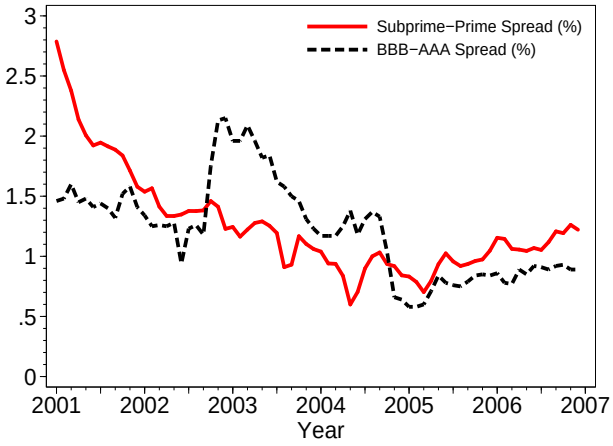
power in this quarter.

[17] The effects of other loan characteristics on mortgage rates have been much more stable over time, as unreported results suggest.

[18] Available at http://www.fhfb.gov/GetFile.aspx?FileID=6416.

Figure 7. FRM Rate Spread and Corporate Bond Yield Spread

The figure shows the FRM subprime-prime rate spread and the yield spread between 10-year BBB and AAA corporate bonds.



subprime loan and borrower characteristics as explanatory variables, using data from 2001 through 2006.[19]

$$spread = \beta_0 + \beta_1 prime + \beta_2' characteristics + error \tag{7}$$

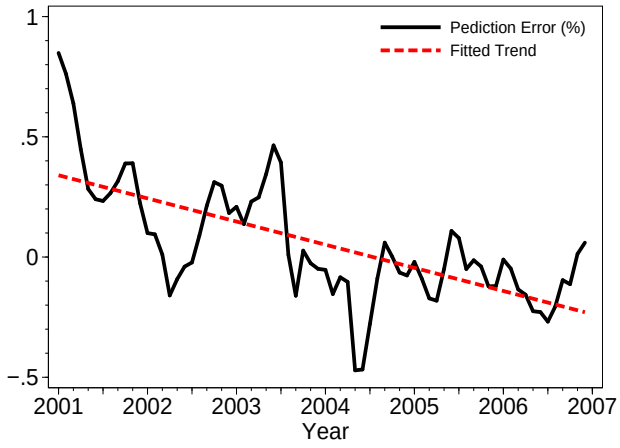
Notice that the $\beta_1 prime$ term corrects for the fact that the spread is affected by the prime rate itself, and thus changes over the business cycle, because a higher prime rate increases the default probability on subprime loans for a given spread. In Figure 8 we plot the prediction error, averaged per origination month, along with a fitted linear trend.

The downward trend in Figure 8 indicates that the subprime-prime spread, after adjusting for differences in observed loan and borrower characteristics, declined. In Figure 2 we showed that loan quality, obtained by adjusting loan performance for differences in loan and borrower characteristics and subsequent house price appreciation, deteriorated over the period, and thus the (adjusted) riskiness of loans rose. Therefore, on a per-unit-of-risk basis, the subprime-prime mortgage spread decreased even more than the level of the spread.

[19]The explanatory factors in the regression are the FICO credit score, a dummy variable that equals one if full documentation was provided, a dummy variable that equals one if prepayment penalty is present, origination amount, value of debt-to-income ratio, a dummy variable that equals one if debt-to-income was not provided, a dummy variable that equals one if loan is a refinancing, a dummy variable that equals one if a borrower is an investor, loan-to-value ratio based on a first-lien, and loan-to-value ratios based on a second, third, etc. liens if applicable.

Figure 8. Prediction Error in the Subprime-Prime Rate Spread

The figure shows the prediction error in the subprime-prime rate spread, determined in a regression of the spread on the prime rate and the following loan and borrower characteristics: FICO credit score, a dummy variable that equals one if full documentation was provided, a dummy variable that equals one if a prepayment penalty is present, origination amount, value of debt-to-income ratio, a dummy variable that equals one if debt-to-income was not provided, a dummy variable that equals one if the loan is a refinancing, a dummy variable that equals one if a borrower is an investor, loan-to-value ratio based on a first lien, and loan-to-value ratio based on a second, third, etc. liens if applicable.



VI. Robustness and Extensions

In this section we discuss the robustness of reported results and extensions of the empirical model used in Sections III and IV.

A. Allowing for Interaction Terms

In this subsection we consider a logit regression with the delinquency rate as the dependent variable. As independent variables we use those from the baseline case presented in Table III, plus the 10 interaction and quadratic terms that can be constructed from the four most important independent variables: the FICO score, the CLTV ratio, the mortgage rate, and subsequent house price appreciation.

Allowing for the above interaction terms, we take into account the effect of risk-layering—such as, for example, the effect of a combination of a borrower's low FICO score and a high CLTV ratio—on the probability of delinquency. In the above example, it is a priori not clear what the sign on the FICO-CLTV interaction variable is. A negative sign means that a low FICO and a high CLTV reinforce each other and give rise to a predicted delinquency probability that is higher than when the interaction is ignored.

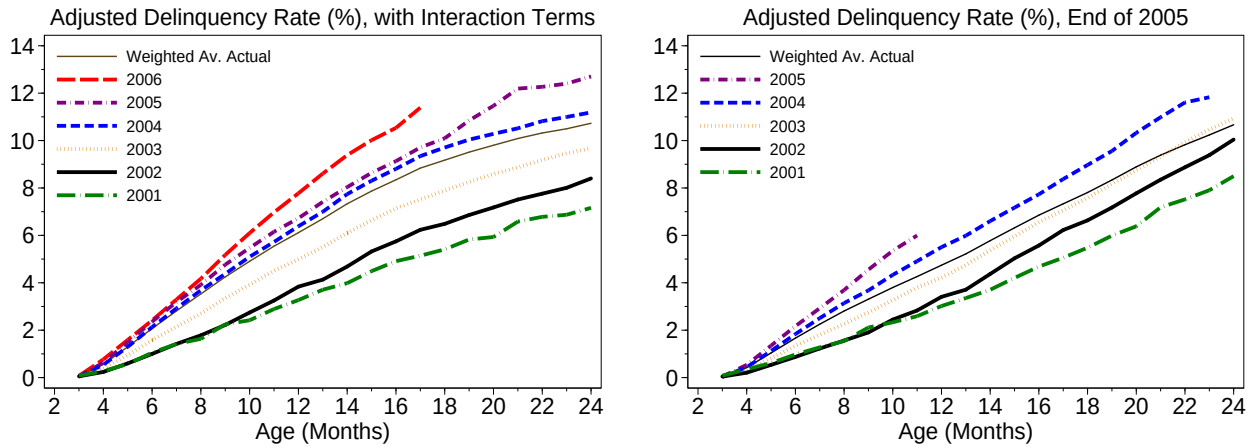
A positive sign could be explained by lenders who originate a low FICO and high CLTV loan only if they have positive private information on the loan or borrower quality. It turns out that the coefficient on the FICO-CLTV interaction term is positive and significant at the 1% level for delinquency 12 months after origination.

More certain is the sign we expect on the HPA-CLTV variable. Low house price appreciation is expected to especially give rise to a higher delinquency probability for a high CLTV ratio, because the borrower is closer to a situation with negative equity in the house (combined value of the mortgage loan larger than the market value of the house). Consistent with this intuition, we find a negative and significant (at the 1 percent level) coefficient on this interaction term for delinquency 12 months after origination.

Using this alternative regression specification, we plot in Figure 9 (left panel) the adjusted delinquency rates, as we did in Figure 2 (left panel) for the baseline case specification. Comparing Figure 2 (left panel) and Figure 9 (left panel) we see that adding the interaction and quadratic terms hardly changes the adjusted delinquency rates. Therefore, our result of a continual increase in the adjusted delinquency rate is robust to explicitly taking into account the effect of risk-layering and other interaction effects.

Figure 9. Adjusted Delinquency Rate, Alternative Specifications.

The figure shows the adjusted delinquency rate under two alternative specifications; it should be compared to the adjusted delinquency rates presented in Figure 2, the baseline case. In the left panel, the adjusted delinquency rates are obtained using the variables of the baseline case plus 10 additional quadratic and interaction terms that can be constructed from the four most important independent variables—the FICO score, the CLTV, the mortgage rate, and house price appreciation. In the right panel, we estimated the baseline case regression using the data from 2001 to 2005 only.



B. Adjusted Delinquency Rate as Observed Year-End 2005

In this subsection we study the following question: Based on information available at the end of 2005, was the dramatic deterioration of loan quality since 2001 already apparent? Notice that we cannot answer this question by simply inspecting vintages 2001 through 2005 in Figure 2 (left panel), because the computation of the adjusted delinquency rate for, say, vintage 2001 loans, makes use of a regression model estimated using data from 2001 through 2006.

Hence, we re-estimate the logit regression model underlying Figure 2 (left panel) making use of only 2001–2005 data. The resulting age pattern in delinquency rates is plotted in Figure 9 (right panel). We again obtain the result that the adjusted delinquency rate rose monotonically from 2001. We therefore conclude that the dramatic deterioration of loan quality should have been apparent by the end of 2005.[20]

C. Non-Linearity in the Sensitivity of the Mortgage Rate to the LTV

In Figure 3 we plotted the sensitivity of the fixed-rate and 2/28 hybrid mortgage rates to the first-lien LTV ratio. The sensitivity is defined as the regression coefficient on the first-lien LTV (scaled by the standard deviation) in a regression with the mortgage rate as dependent variable and the first-lien LTV, the second-lien LTV, and the other loan and borrower characteristics listed in Subsection VI B, as independent variables.

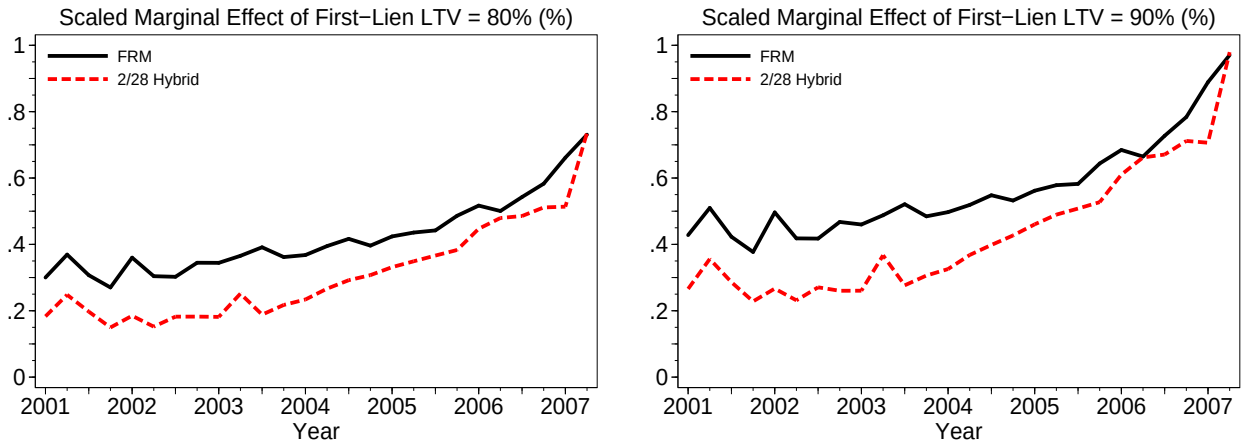
In this subsection we study the robustness of this result to adding the square of the first-lien LTV and the square of the second-lien LTV as independent variables, therefore allowing for a non-linear functional form. In Figure 10 we report the resulting scaled marginal effect of the first-lien LTV for fixed-rate and 2/28 hybrid mortgages evaluated at a first-lien LTV of 80 percent (left panel) and 90 percent (right panel). Without non-linear terms the marginal effect is simply given by the regression coefficient. This is what we plotted in Figure 3. With the quadratic terms, the marginal effect is given by $\beta_{LTV} + 2\beta_{LTV^2}X$, where the βs are the regression coefficients and X is the first-lien LTV ratio at which the marginal effect is evaluated.

As shown in Figure 10, the marginal effect is rising over time, consistent with the baseline case results presented in Figure 3. Moreover, we find that there is a statistically and economically significant non-linear effect of the first-lien LTV on the mortgage rate. Comparing the left and right panels in Figure

[20] One reason why investors did not massively start to avoid or short subprime-related securities is that the timing of the subprime market downturn may have been hard to predict. Moreover, a short position is associated with a high cost of carry (Feldstein (2007)).

Figure 10. Sensitivity of Mortgage Rate to First-Lien LTV Ratio Allowing for Non-Linearity

The figure shows the scaled marginal effect of the first-lien loan-to-value (LTV) ratio on the mortgage rate for first-lien fixed-rate and 2/28 hybrid mortgages, evaluated at a first-lien LTV of 80% (left panel) and 90% (right panel). The effect is determined using an OLS regression with the interest rate as dependent variable and the FICO score, first-lien LTV (and the square), second-lien LTV (and the square), debt-to-income ratio, missing debt-to-income ratio dummy, cash-out refinancing dummy, owner-occupation dummy, prepayment penalty dummy, origination amount, term of the mortgage, prepayment term, and margin as independent variables.



10, the higher the first-lien LTV ratio, the more sensitive is the mortgage rate to changes in the first-lien LTV. The largest difference between the results based on specifications with and without non-linearity is observed for 2/28 hybrid mortgages in 2007 at a first-lien LTV of 90 percent (right panel). The scaled marginal effect increases by 27 basis points over the course of 3 months in 2007 when a model allows for non-linearity. In contrast, ignoring the non-linearity, as in Figure 3, the increase in the scaled marginal effect is only 13 basis points.

VII. Concluding Remarks

The subprime mortgage market experienced explosive growth between 2001 and 2006. Angell and Rowley (2006) and Kiff and Mills (2007), among others, argue that this was facilitated by the development of private-label mortgage backed securities, which do not carry any kind of credit risk protection by the Government Sponsored Enterprises. Investors in search of higher yields kept increasing their demand for private-label mortgage-backed securities, which also led to sharp increases in the subprime share of the mortgage market (from around 8 percent in 2001 to 20 percent in 2006) and in the securitized share of the subprime mortgage market (from 54 percent in 2001 to 75 percent in 2006).

In this paper we show that during the dramatic growth of the subprime (securitized) mortgage market, the quality of the market deteriorated dramatically. We analyze loan quality as the performance of loans, adjusted for differences in borrower characteristics (such as credit score, level of indebtedness, ability to provide documentation), loan characteristics (such as product type, amortization term, loan amount, interest rate), and subsequent house price appreciation.

The decline in loan quality has been monotonic, but not equally spread among different types of borrowers. Over time, high-LTV borrowers became increasingly risky (their adjusted performance worsened more) compared to low-LTV borrowers. Securitizers seem to have been aware of this particular pattern in the relative riskiness of borrowers: We show that over time mortgage rates became more sensitive to the LTV ratio of borrowers. In 2001, for example, the premium paid by a high LTV borrower was close to zero. In contrast, in 2006 a borrower with a one standard deviation above-average LTV ratio paid a 30 basis point premium compared to an average LTV borrower.

In principal, the subprime-prime mortgage rate spread (subprime mark-up) should account for the default risk of subprime loans. For the rapid growth of the subprime mortgage market to have been sustainable, the increase in the overall riskiness of subprime loans should have been accompanied by an increase in the subprime mark-up. In this paper we show that this was not the case: Subprime mark-up—adjusted and not adjusted for changes in differences in borrower and loan characteristics—declined over time. With the benefit of hindsight we now know that indeed this situation was not sustainable, and the subprime mortgage market experienced a severe crisis in 2007. In many respects, the subprime market experienced a classic lending boom-bust scenario with rapid market growth, loosening underwriting standards, deteriorating loan performance, and decreasing risk premiums.[21] Argentina in 1980, Chile in

[21] A more detailed discussion, theory, and empirical evidence on such episodes is available in Dell'Ariccia and Marquez

1982, Sweden, Norway, and Finland in 1992, Mexico in 1994, Thailand, Indonesia, and Korea in 1997 all experienced the culmination of a boom-bust scenario, albeit in different economic settings.

Were problems in the subprime mortgage market apparent before the actual crisis showed signs in 2007? Our answer is yes, at least by the end of 2005. Using the data available only at the end of 2005, we show that the monotonic degradation of the subprime market was already apparent. Loan quality had been worsening for five consecutive years at that point. Rapid appreciation in housing prices masked the deterioration in the subprime mortgage market and thus the true riskiness of subprime mortgage loans. When housing prices stopped climbing, the risk in the market became apparent.

(2006), Demirgüç-Kunt and Detragiache (2002), Gourinchas, Valdes, and Landerretche (2001), and Kamisky and Reinhart (1999), among many others.

References

Angell, Cynthia, and Clare D. Rowley, 2006, Breaking new ground in U.S. mortgage lending, FDIC: Outlook Summer 2006 Federal Deposit Insurance Corporation.

Berger, Allen, and Gregory Udell, 2004, The institutional memory hypothesis and the procyclicality of bank lending behavior, *Journal of Financial Intermediation* 12, 458–495.

Bernanke, Ben, 2007, Housing, housing finance, and monetary policy, Speech Federal Reserve Bank of Kansas City's Economic Symposium.

Brunnermeier, Markus, and Christian Julliard, 2007, Money illusion and housing frenzies, *Review of Financial Studies* forthcoming.

Campbell, John Y., and João F. Cocco, 2003, Household risk management and optimal mortgage choice, *Quarterly Journal of Economics* 118, 1449–1494.

Cutts, Amy Crews, and Robert A. Van Order, 2005, On the economics of subprime lending, *The Journal of Real Estate Finance and Economics* 30, 167–196.

Dell'Ariccia, Giovanni, and Robert Marquez, 2006, Lending booms and lending standards, *Journal of Finance* 61, 2511–2546.

Demirgüç-Kunt, Asli, and Enrica Detragiache, 2002, Does deposit insurance increase banking system stability? An empirical investigation, *Journal of Monetary Economics* 49, 1373–1406.

Deng, Yongheng, John M. Quigley, and Robert A. Van Order, 2000, Mortgage terminations, heterogeneity and the exercise of mortgage options, *Econometrica* 68, 275–308.

Feldstein, Martin S., 2007, Housing, credit markets and the business cycle, NBER working paper 13471.

Genesove, David, and Christopher Mayer, 1997, Equity and time to sale in the real estate market, *American Economic Review* 87, 255–269.

Genesove, David, and Christopher Mayer, 2001, Loss aversion and seller behavior: Evidence from the housing market, *Quarterly Journal of Economics* 116, 1233–1260.

Glaeser, Edward, and Joseph Gyourko, 2005, Urban decline and durable housing, *Journal of Political Economy* 113, 345–375.

Gourinchas, Pierre-Oliver, Rodrigo Valdes, and Oscar Landerretche, 2001, Lending booms: Latin America and the world, *Economia* 1, 47–99.

Gyourko, Joseph, Mayer Christopher, and Todd Sinai, 2006, Superstar cities, NBER working paper 12355.

Kamisky, Graciela L., and Carmen M. Reinhart, 1999, The twin crises: The causes of banking and balance-of-payment problems, *American Economic Review* 89, 473–500.

Kiff, John, and Paul Mills, 2007, Money for nothing and checks for free: Recent developments in U.S. subprime mortgage markets, IMF working paper.

Koijen, Ralph, Otto Van Hemert, and Stijn Van Nieuwerburgh, 2007, Mortgage timing, NBER working paper 13361.

LaCour-Little, Michael, 2007, The home purchase mortgage preferences of low-and-moderate income households, *Real Estate Economics* forthcoming.

Pennington-Cross, Anthony, and Souphala Chomsisengphet, 2007, Subprime refinancing: Equity extraction and mortgage termination, *Real Estate Economics* 35, 233–263.

Ruckes, Martin, 2004, Bank competition and credit standards, review of financial studies, *Review of Financial Studies* 17, 1073–1102.

Shiller, Robert J., 2007, Understanding recent trends in house prices and home ownership, NBER working paper 13553.

Van Hemert, Otto, 2007, Household interest rate risk management, Working paper.

Vickery, James, 2007, Interest rates and consumer choice in the residential mortgage market, Working paper.

Von Furstenberg, George M., and R. Jeffery Green, 1974, Home mortgage delinquencies: A cohort analysis, *Journal of Finance* 29, 1545–1548.