Federal Reserve Bank of New York
Staff Reports

Juvenile Delinquent Mortgages: Bad Credit or Bad Economy?

Andrew Haughwout
Richard Peach
Joseph Tracy

Staff Report no. 341
August 2008

This paper presents preliminary findings and is being distributed to economists and other interested readers solely to stimulate discussion and elicit comments. The views expressed in the paper are those of the authors and are not necessarily reflective of views at the Federal Reserve Bank of New York or the Federal Reserve System. Any errors or omissions are the responsibility of the authors.

Juvenile Delinquent Mortgages: Bad Credit or Bad Economy?

Andrew Haughwout, Richard Peach, and Joseph Tracy

Federal Reserve Bank of New York Staff Reports, no. 341
August 2008
JEL classification: G21, R21

Abstract

We study early default, defined as serious delinquency or foreclosure in the first year, among nonprime mortgages from the 2001 to 2007 vintages. After documenting a dramatic rise in such defaults and discussing their correlates, we examine two primary explanations: changes in underwriting standards that took place over this period and changes in the economic environment. We find that while credit standards were important in determining the probability of an early default, changes in the economy after 2004—especially a sharp reversal in house price appreciation—were the more critical factor in the increase in default rates. A notable additional result is that despite our rich set of covariates, much of the increase remains unexplained, even in retrospect. Thus, the fact that the credit markets seemed surprised by the rate of early defaults in the 2006 and 2007 nonprime vintages becomes more understandable.

Key words: housing, mortgage default, subprime mortgages

Haughwout: Federal Reserve Bank of New York. Peach: Federal Reserve Bank of New York. Tracy: Federal Reserve Bank of New York. Address correspondence to Andrew Haughwout (e-mail: andrew.haughwout@ny.frb.org). The authors wish to thank Jonathan Stewart and especially Ebiere Okah for excellent assistance with the data. The views expressed in this paper are those of the authors and do not necessarily reflect the position of the Federal Reserve Bank of New York or the Federal Reserve System.

Gee, Officer Krupke, we're very upset;
We never had the love that ev'ry child oughta get.
We ain't no delinquents,
We're misunderstood.
Deep down inside us there is good!

"Gee, Officer Krupke" – *West Side Story*

Rapid increases in US residential mortgage defaults during 2007 and into 2008 captured the attention of researchers, the public and policy makers, and had a chilling effect on credit markets worldwide. While these increases were noted originally in the nonprime market, foreclosure increases have in more recent months begun to spill over into the prime market. This paper studies a part of this phenomenon, early defaults in the nonprime market.

Historically, four key characteristics ("risk factors" or "underwriting criteria") have been thought to determine the probability that a mortgagor will default. Those factors are the loan-to-value ratio (LTV)[1], the debt service-to-income ratio (DTI), the mortgagor's credit score, and the extent to which the mortgagor's income and assets have been verified by third party sources such as employers, tax returns, and bank account statements. To expand the potential pool of borrowers, nonprime (subprime and alt-a) mortgages by design relaxed one or more of these underwriting criteria beyond the margins required for prime mortgage loans. A direct consequence is that we would expect the default experience of these relatively new mortgage products to be worse than that of prime loans. Indeed, industry data confirm that the performance of the very first vintages of nonprime loans was significantly worse than that of prime loans.[2]

Nonetheless, as shown in Figure 1, beginning with the 2005 vintage the performance of nonprime mortgage loans became notably worse than previous vintages. The performance of the 2006 vintage deteriorated even further. By 12 months following origination, the 2005 vintage

[1] The LTV is calculated by taking the ratio of the mortgage balance to the value of the home. LTVs are typically expressed as a number ranging from 0 to 100 or higher. If the borrower has "negative equity" where the mortgage balance exceeds the value of the home, the LTV will exceed 100.

[2] The National Delinquency Survey published by the Mortgage Bankers Association of America (MBA) is one of the main sources of information on mortgage loan performance, including nonprime loans. However, it should be noted that mortgages are placed into these categories based on the servicer rather than the individual loan. Thus, if more than 50 percent of a servicer's portfolio is nonprime loans, then all of that firm's loans are lumped into the subprime category. Alt-a mortgages, according to the MBA, are divided between the prime and subprime groups. See http://www.mortgagebankers.org/files/Research/NDSFactSheet.pdf for details.

had a 90 day or more delinquency rate that was not reached by the 2003 vintage for 20 months, and the 2006 vintage at 12 months had a rate that was not reached by the 2003 vintage even by 30 months. Moreover, this sharp decline in loan performance was a surprise to investors in these loans in that to a large extent it seemed unexplained by the observed risk characteristics.

The sharp increase in defaults *very* early in the life of the loans suggests the moniker "juvenile delinquents." In the case of nonprime adjustable rate mortgages (ARMs), defaults often occurred well before the first interest rate reset while the initial "teaser" rate was still in effect. We define an "early default" as a mortgage that is 90 or more days delinquent within the first year after origination. We use this window since performance warranties by originators often covered the first year. The reasoning was that any serious underwriting problems with mortgages typically would manifest themselves within the first year. In our data, 10 percent of nonprime loans originated in 2007 experienced an early default, as compared to 2.7 percent of similar loans originated in 2003.

The purpose of this paper is to explore potential explanations for the sharp rise in early defaults of the 2005 through 2007 vintages of nonprime mortgages. We will examine how much of the deterioration in the early performance of these mortgages can be explained by changing risk characteristics of nonprime mortgages over time (i.e. "bad credit"). New and existing home sales peaked in late 2005 in many housing markets, and house prices began to soften and then to decline as these housing markets cooled. We will also explore the extent to which house price dynamics over the housing cycle as well as other local economic factors help to explain the early default behavior of the more recent vintages of nonprime mortgages (i.e. "bad economy"). Importantly, we will investigate the extent to which the effect of house price dynamics on early defaults depends on the risk profile of mortgages in a vintage – that is, are there important interaction effects that help to determine a vintage's share of juvenile delinquent mortgages in that vintage.

The next section provides a brief literature review of selective papers that are relevant to our analysis. We next describe our primary data source and discuss the evolution of the four basic risk factors for nonprime mortgages from 2001 to 2007. We provide tabulations of these risk characteristics and early default rates. We then turn to a multivariate analysis of early

defaults. The final section draws insights from our analysis for the current housing policy debate and concludes.

Review of Past Literature

Residential mortgages are complex financial instruments that confer important options on the borrower. The extensive body of previous research on residential mortgage default has adapted option theory to the study of mortgage valuation, since there exist well-developed theory and empirical methods for valuing financial derivatives and their exercise (Black and Scholes 1973).

An important feature of most residential mortgages is that they are "non-recourse" loans, either de jure or de facto. This means that in the event of a default, creditors can sell the house to cover the loan balance, but typically do not legally pursue the borrower for any deficiency.[4] This creates a "put" option for the borrower which he/she can exercise if the house value falls sufficiently relative to the loan balance. In addition to this default option, borrowers may continue to make the scheduled payments until the mortgage debt is discharged, or prepay the mortgage either by selling the house and paying off the balance on the mortgage or by refinancing into a new loan (Kau, Keenan, Mueller and Epperson, 1995). The option to prepay is often referred to as the "call" option that borrowers hold when they take out a mortgage.

Foote Gerardi and Willen (2008) succinctly summarize the prediction of option theory for default when they argue that negative equity is a necessary but not a sufficient condition for default. Borrowers with positive equity ought to rarely if ever default, since (in the event of an idiosyncratic shock such as illness, loss of job or divorce) they can sell the house or refinance the mortgage. Borrowers with negative equity, on the other hand, may default in the face of similar shocks, since the option to refinance and/or sell the house is conditional on being able to raise cash to cover the difference between the mortgage balance and the proceeds of a sale or a new mortgage (Foster and Van Order, 1984, Vandell, 1995).

Even borrowers with negative equity, however, default less frequently than simple models would predict (see Vandell 1995 for a summary of the empirical evidence and Elul 2006

[4] While legal pursuit of borrowers' assets to cover deficiencies is available in most states, it is often restricted. In California, for example, deficiency judgments are not allowed for owner-occupied one to four-family homes (Pence 2003).

for an update). For an owner occupant considering default, transactions costs include moving costs, the cost of purchasing or renting a new residence, and damage to one's credit score resulting in higher future borrowing costs. All told, some authors have argued that these costs can typically range from 15 to 30% of the value of the house, helping to explain why default appears to be underexercised relative to the simple option-theoretic prediction (Cunningham and Hendershott, 1984). Investors face fewer of these transaction costs and therefore may be more likely to default for a given LTV level.[5]

As noted by many authors, exercise of the default option will be related to the value of the prepayment option, regardless of the borrower's equity in the property (Schwartz and Torous 1993, Vandell 1993, Elul 2006). This suggests that in evaluating the prevalence of default, we must account for the value of the option to sell the house or refinance the mortgage. The typical approach to this problem is to estimate a duration model that simultaneously accounts for the competing risks of prepayment and default (Deng, Quigley and Van Order 2000). An additional advantage of this approach is that it allows insight into the value of a particular mortgage, or mortgage backed security. While our ambitions in the current study are more modest, we must remain attentive in both the specification of our models and the interpretation of our results to the fact that even early defaults may be affected by the availability of prepayment (Deng and Gabriel 2006).

In addition, Kau, Keenan and Kim (1993) argue that the apparent underexercise of the default option is partly explained by the fact that continuing to make payments preserves the borrower's ability to default or sell the house in the future. That is, exercise of the default option depends on the future value of both living in the house and selling the house, either in the marketplace or through default (effectively selling the house to the lender).

Much of the empirical research conducted on mortgage default over the last two decades has focused on the large market for prime/conforming or FHA mortgages. Most relevant for our study is recent work on the FHA market segment, which serves borrowers who are similar to the nonprime sample we study.

[5] Besides not incurring any moving costs, investors may also not face the same increase in future borrowing costs in the event of a default. If the housing investment is held in a limited liability corporation, then a default would not affect the owner's personal credit rating. In this case, we are more confident that the borrower would be identified as an "investor" in the LoanPerformance data.

Deng and Gabriel (2006) and An, Bostic, Deng and Gabriel (2007) present competing risk estimates for a sample of FHA purchase loans originated between 1992 and 1996. In these studies, lower FICO scores are associated with higher default rates and somewhat lower prepayment rates. Higher local unemployment rates have little effect on default rates. However, higher local unemployment is associated with lower prepayments among FHA borrowers. Finally, higher LTVs – or measures of the probability that the put option is "in the money" – raise both the default risk, particularly for borrowers with low FICO scores.

Very recently, studies of subprime mortgages have become more common, as their market share has expanded. Pennington-Cross and Ho (2006) provide a detailed analysis of the performance of subprime mortgages over the period leading up to the crash in the housing market. Their analysis provides insights into the behavior of these mortgages in a general environment of rising house prices. They use LoanPerformance data on subprime mortgages that were originated between 1998 and 2005. Loans were followed for up to five years or the end of 2005. Fixed rate mortgages are contrasted to the hybrid 2/28 adjustable mortgage. Like the two FHA studies, Pennington-Cross and Ho find that borrowers with lower credit scores are more likely to default, and that local unemployment rates seem to have little effect on default. In an interesting contrast with the FHA results, Pennington-Cross and Ho find that among subprime borrowers, higher LTVs raise the default risk but lower the prepayment risk.

Gerardi, Shapiro and Willen (2007) take a non-traditional approach to analyzing mortgage default. Using Massachusetts deed records from January 1987 to August 2007, they compile a panel data set tracking purchases, refinances, sales and foreclosures on all residential properties in the state. The strength of this data is the ability to follow the same property though different owners, as well as across different mortgages for a given owner. This is in contrast to the typical loan level data which tracks a given mortgage over its life, but does not permit linking mortgages over time for the same borrower or the same property.

A weakness of the deed based data is that it lacks information on the characteristics of the borrower and some characteristics of the mortgage. The deed records indicate the mortgage originator. The authors identify subprime mortgages by matching the originator to a Department of Housing and Urban Development list of subprime lenders, and estimate a competing risk model of the outcomes of distinct "ownership experiences." Starting with a home purchase, they

[9] As noted earlier, they control for average household income for the census tract.

follow the household until the home is sold or goes into foreclosure. A key finding by Gerardi *et al* (2007) is the important role of house price appreciation on the likelihood of a foreclosure. Cumulative price appreciation since the date of the house purchase exerts a sizeable downward effect on the probability that the ownership experience ends in a foreclosure. The authors impose symmetry in the effect of price appreciation and price depreciation. It would be useful to know if nominal losses are relatively more important at generating defaults than nominal gains are at preventing defaults. The authors do not include a negative equity indicator for the current mortgage due to worries over the endogeneity of this indicator variable.

Gerardi *et al* (2007) report that ownership experiences that begin with a subprime mortgage are much more likely to end in a default than observably similar ownership experiences that begin with a prime mortgage. Here it is important to keep in mind that the authors are not able to control for some key characteristics of the borrower such as income and credit score.[9] It is not clear, then, how much of the difference in default rates across ownership experiences could be explained by differences in these borrower specific risk factors. Foote, Gerardi and Willen (2008) use the same data set to examine the specific role of negative equity in default behavior. The authors confirm many of the results in Gerardi *et al* (2007), including the higher likelihood of default among subprime borrowers, and find additional evidence that borrowers in a negative equity position – whether the mortgage is prime or not – are more likely to default. Our focus on the nonprime market segment will allow us to determine the effect of negative equity on borrowers in particularly high cost mortgages.

Demyanyk and van Hemert (2008) adopt an approach similar to ours in their analysis of subprime mortgages. Using LoanPerformance data, which we discuss in more detail below, they examine the likelihood that a mortgage is either 60 or more days delinquent or in foreclosure within the first twelve months following origination. Their principal aim is to see to what extent changes in the distribution of risk factors can explain the deterioration in the early performance of subprime mortgages in 2005 and 2006. We will contrast our findings to theirs in greater detail later in the paper. The authors conclude that declining underwriting standards, particularly reflected in increasing LTVs at origination, are the dominant force explaining the rapid rise in early delinquency and defaults among subprime borrowers.

One important difference between our approach and that of Demyanyk and van Hemert involves the treatment of house price appreciation. While we follow previous literature by

controlling for an updated LTV using current house price information, Demyanyk and van Hemert control for the initial LTV and treat house price appreciation as an independent effect on the default probability. This approach misses what we believe is an important interaction between house price dynamics and origination LTV, and complicates the evaluation of the borrower's put option. We discuss the separate effects of updated LTV and house price appreciation below. In addition, Demyanyk and van Hemert impose that house price increases and decreases have a symmetric impact on defaults. We test for asymmetric effects.

A second difference in approaches is that Demyanyk and van Hemert include the mortgage rate as a control variable in their empirical specifications. The mortgage rate has the largest standardized marginal effect in their specifications. The coefficient on the mortgage rate reflects the impact on early performance from variation in mortgage rates that is orthogonal to the other risk factors in their specification. One possibility is that this residual rate variation reflects risk factors observed in the underwriting but not fully reflected in the data by the recorded risk measures. Alternatively, this variation could reflect the degree of competition in the local lending markets, differences in bargaining skill across different borrowers, or lagged performance of mortgages in the local lending market. It is not clear, then, what specifically the mortgage rate is capturing in their analysis. We do not include the loan-specific mortgage rate in our empirical specifications, since our aim is to evaluate the extent to which standard observable risk factors can account for the rise in early defaults.

Nonprime Mortgage Data and Tabulations on Early Defaults

Loan Performance Data

Our mortgage data come from FirstAmerican CoreLogic's Loan Performance Data, a proprietary data base which, as of June 2008, provides loan-level information at a monthly frequency on approximately seven million active, securitized subprime and alt-a loans, carrying balances of over $1.6 trillion.[10] Subprime mortgages are small loans (compared to alt-a loans) and are often made to borrowers with some blemish on their credit history, or who are willing to commit large shares of their incomes to debt service. Alt-a mortgages are typically larger value

[10] See http://www.loanperformance.com/data-power/default.aspx

loans made to more credit-worthy borrowers who, for a variety of reasons, may choose not to provide the income or asset verification required to obtain a prime mortgage. Both types of nonprime mortgages are typically higher-cost than prime conforming loans.

The database consists of information on securitized subprime and alt-A mortgage loans. A large share of outstanding subprime and alt-A mortgages are securitized, with the balance remaining in lender portfolios, and LoanPerformance captures upwards of 90% of the securitized loans (Mayer and Pence 2008). Pennington-Cross (2002) argues that securitized subprime mortgages differ systematically from those retained in portfolio. Since our data are limited to securitized loans, any inferences should be limited to this set of loans.

The LoanPerformance dataset is a rich source of information on the characteristics of these securitized loans. The dataset includes information on the date of origination, the zip code in which the collateral property is located, details of the mortgage contract (including term, initial interest rate, and rate adjustment schedule), and underwriting information (including borrower credit score, debt to income ratio, the loan to value ratio for senior and junior liens, and the extent of income and asset verification provided by the borrower). Also included are monthly updates of "dynamic" information including the current interest rate, mortgage balance and the borrower's payment record.

We analyze a one percent random sample of the first-lien subprime and Alt-A loans reported in the data as of our most recent monthly update, for April 1, 2008.[11] The universe from which we sample includes all loans, whether they are currently active or have been paid off. From this set of loans, we select those that originated after January 1, 2001 and for which we observe at least twelve months of performance; thus, our "youngest" loans originated during April 2007. For each origination, we use the payment history for the first twelve months in order to determine whether it defaulted during that time period. The result is a dataset consisting of about 117,000 loans. We classify these loans as subprime or Alt-A based on the designation of the security in which they were packaged. Approximately 2,000 of our loans are missing this designation in the data, bring our total for analysis to around 115,000 loans.We combine this loan-level information with economic data measured at the metropolitan area level. These data

[11] Since observations in the LoanPerformance dataset are loans coded to the zip code, we choose our panel based on first-lien loans only. This avoids the possibility of double counting subordinate lien loans on the same property.

include measures of house price appreciation (the widely-used OFHEO repeat-sale index[12]) and labor market conditions from the Bureau of Labor Statistics. The principal characteristics of our data are described in Table A-1.

Tabulations on Early Default

The value of the default put option discussed earlier will likely depend on the initial LTV on a mortgage and the pattern of house price appreciation since the origination of the loan. Table 1 provides information on the initial LTV for subprime and alt-a mortgages originated from 2001 to 2007. For each type of mortgage, we provide descriptive statistics on the distribution of the cumulative LTV at origination of the first-lien mortgage. The cumulative LTV reflects the total mortgage balance of the first-lien mortgage and any subordinate lien loans if they exist at origination. We also report the fraction of transactions that involved one or more 2nd-lien loans. What is striking for both classes of mortgages is the significant rise in the incidence of 2nd-liens from 2001 to 2006. For subprime mortgages over this period the incidence of 2nd-liens rises from 3.2 percent to 29.4 percent, and for alt-a mortgages the incidence rises from 2.2 percent to 43.9 percent. A consequence is that 30.1 percent of subprime mortgages originated in 2006 had an initial cumulative LTV of at least 100 – that is, the borrower at origination did not have *any* equity in the house. Excluding the 2nd-lien loans, for subprime mortgages in 2006 the median LTV was 80, while the the 90th percentile LTV was 95.[13] While from 2004 to 2007 the incidence of 2nd-liens was higher for alt-a mortgages, the distribution of initial LTV was not as significantly affected. It is also clear that the number of nonprime mortgages in our data rises sharply over most of the time period, reflecting both increases in originations and more complete coverage by LoanPerformance (Mayer and Pence 2008).

Over the period from 2003 to 2007, the incidence of early defaults more than quadrupled for both subprime and alt-a mortgages. Comparing subprime to alt-a mortgages in Table 2, for any given initial range of LTV the early default rate for subprime mortgages tended to be higher than for alt-a mortgages. As the average early default rates were rising, the data show for both classes of mortgages increases in the early default rates within each LTV range. For alt-a

[12] See http://www.ofheo.gov/hpi.aspx for details.
[13] For alt-a mortgages in 2006 based just on the first lien mortgage the median through 90th percentile LTV was 80.

mortgages originated from 2003 to 2007, the data also show a rise in the average early default rate as one moves from lower to higher initial LTV intervals. This pattern is less consistent for subprime mortgages over this same time period.

The housing boom in the first half of this decade was concentrated in a small number of states. Figure 2 shows a scatter plot of the average annual house price appreciation rate from the third quarter of 2001 to third quarter of 2006 on the vertical axis and appreciation rate from the first quarter of 2007 to the first quarter of 2008 on the horizontal axis. Nine states experienced double digit house price appreciation sustained over a five year period that were followed by house price declines in the year following. In four of these states (AZ, CA, FL, and NV) the reversal has been especially sharp. [14] Three states (IN, MI and OH) presents a different picture where instead of experiencing a housing boom and bust, these states have been suffering from relative economic weakness and soft housing markets during this entire decade.

The rapid house price increases in the boom/bust states prior to the downturn would act to keep the put option for default out-of-the-money. Even where the lender finances most or all of the borrower's down payment with a 2nd lien loan, twelve months of double-digit house price appreciation will generate more than sufficient equity to cover the transactions costs of selling the house. Similarly, in cases where a borrower in a boom/bust state suffers a job loss, divorce or significant health problem during the boom period, we would not expect to see this result in a default. The borrower would have a financial incentive to sell the house and prepay the mortgage rather than default. Finally, as discussed earlier, owners may be less likely to exercise the default put option than investors other things equal.

We take a preliminary look in Table 3 at the likely interplay between house price dynamics, local economic conditions, investor status and initial LTV. For owners in economically depressed states, the incidence of early defaults tends to be higher than the overall average for both classes of mortgages. In contrast, for subprime owners in those states that experienced a house price boom, the incidence of early defaults is low relative to owners in the "other" states except for the highest LTV interval. This pattern does not emerge for alt-a owners where the data indicate that the default rate in the boom/bust states was generally higher than for

[14] Both the OFHEO and the other widely-used house price index, S&P/Case-Shiller (CS), are derived using the repeat sales methodology. However, there are important differenced in the construction of these indices which result in quite different assessments of the behavior of home prices, particularly at the national level. See Leventis (2008) for details.

alt-a owners in the "other" states. Consistent with investors having a lower threshold for exercising the default put option, the overall early default rate is higher for investors than for owners for initial LTVs greater than 90. Comparing investors across the boom/bust states and the economically weak states, nonprime investors have higher early default rates in the economically weak states, with the exception of alt-a investors who start out with no equity in the deal.

The second risk factor is the DTI ratio. This ratio is meant to capture a borrower's capacity to pay even in the face of transitory shocks to his/her personal finances.[15] Table 4 gives the evolution of the DTI ratios for nonprime mortgages over our sample period. The first thing to note is that subprime borrowers relative to alt-a borrowers have a distribution of DTI that is less concentrated in the DTI range below 30 and more concentrated in the DTI range above 40. That said, the distributions of DTIs were reasonably constant over the time period with the exception of 2006/2007 where there was a noticeable increase in high DTI subprime and alt-a mortgages. The early default performance by different initial DTI intervals and year is presented in Table 5. Subprime mortgages display a stronger relationship than alt-a mortgages between DTI and early defaults across the years in our sample. Similar to our finding for initial LTV, from 2003 onwards the incidence of early defaults rises over time within each DTI range for both types of mortgages, but in this case the change is relatively common across the DTI intervals.

The "willingness" to pay is captured by the borrower's FICO score. Table 6 provides information on the distribution of FICO scores over time for the nonprime mortgages in our sample. Similar to the DTI risk measure, subprime borrowers are more concentrated than alt-a borrowers in the below 600 interval of FICO scores and much less concentrated in the above 660 inverval. The likelihood that a subprime borrower had a FICO score below 600 was decreasing from 2001 to 2007. While few alt-a borrowers had FICO scores below 619, the incidence was declining as well over this time period. In contrast to the initial LTV and DTI risk factors, the distributions of nonprime FICO scores were not deteriorating over the period leading up to the sharp rise in early defaults.

Table 7 shows early default rates across time for the different FICO ranges. In each year early defaults typically are a declining function of FICO scores. Generally, borrowers with a FICO score of less than 600 are at least three times more likely to experience an early default as

[15] A possibility is that DTI is less predictive for early defaults than it is for overall defaults, as the likelihood of a significant financial shock is less over the first year of a mortgage as compared to the expected life of the mortgage.

borrowers with a FICO score of over 660. The exception is for subprime borrowers in 2006 and 2007 where the early default rates tended to converge across the range of FICO scores. The data do not display this same convergence in early defaults for alt-a mortgages. For subprime borrowers, the relative increase in early defaults since 2003 within a range of FICO scores was larger for the two higher FICO score ranges than for the two lower ranges. The relative increase in early defaults within a FICO interval for alt-a borrowers is more varied across the different ranges of credit scores.

The final standard risk factor is the level of documentation used in the underwriting of the mortgage. The data classify underwriting into one of three categories: full documentation, low documentation ("limited-doc") and no documentation ("no-doc"). Table 8 provides the distribution of the documentation level over time for subprime and alt-a mortgages. The use of lower documentation was much more prevalent for underwriting alt-a mortgages than for subprime mortgages. Despite the focus in the press made on no-doc mortgages, in each year the incidence of no-doc mortgages was in single digits, and was declining over the sample period. What is more notable is the shift in composition from fully documented to limited documented underwriting. From 2001 to 2006, the share of fully documented subprime mortgages fell from 77.8 percent to 61.7 percent, while the share of fully documented alt-a mortgages fell from 36.8 percent to 18.9 percent.

Table 9 gives the yearly average early default rates for subprime and alt-a mortgages broken down by the level of documentation. In each year for subprime mortgages, early defaults are more prevalent for limited as compared to fully-documented mortgages. For alt-a mortgages, the incidence of early defaults in each year generally increases as one moves from fully-documented to limited doc mortgages, and from limited doc to no-doc mortgages. From 2005 to 2007 as the overall incidence of early defaults among alt-a mortgages was rising, the incremental effect on early defaults associated with moving from fully documented mortgages to limited documented mortgages was higher than the incremental effect associated with moving from limited documented mortgages to mortgages with no documentation.

Determinants of Early Default

Econometric specification

For each mortgage in our data, the outcome of interest is an indicator for whether the mortgage experienced an early default. We adopt a linear index model in order to examine the determinants of early default. For each mortgage *i* originated in year and month t_m in metropolitan area *j*, we assume that there is a latent index I^*_{ijt} that captures the net benefit to the borrower from an early default. We specify this latent index as a linear index of the observable risk factors (*X*) and local economic conditions (*Z*).

$$I^*_{ijt} = X_{it}\beta + Z_{jt}\delta + \alpha_{jt} + \varepsilon_{ijt}$$

Given the paucity of data that we have to control for local economic conditions, we assume that the performance of mortgages originated in a given metropolitan area and year/quarter may be affected by a common random error component α_{jt} that captures any unobserved economic shocks that impacted this local housing market. We account for the possible presence of the error component α_{jt} in calculating the standard errors of our estimates.

Let I_{ijt} denote our observed indicator variable for whether a mortgage experiences an early default.[16] We assume that this occurs whenever the unobserved latent index takes on a positive value.

$$I_{ijt} = \begin{matrix} 1 & \text{if } I^*_{ijt} \geq 0 \\ 0 & \text{otherwise} \end{matrix}$$

The probability of an early default is given by the following.

$$\Pr(I_{ijt} = 1 \mid X_{it}, Z_{jt}) = \Pr(I^*_{ijt} > 0 \mid X_{it}, Z_{jt})$$

[16] It is important to distinguish this early default indicator from a default hazard in a competing risk specification. If a mortgage prepays in the first year, this would censor the default hazard at the prepayment date. In contrast, in our setup an early prepayment is treated in the same manner as a mortgage that is still ongoing without an early default after 12 months. Both outcomes would generate a value of zero for the early default indicator.

We report linear probability estimates which facilitate calculating the decomposition of the change in aggregate defaults into credit and economy effects.

The first risk factor is the LTV. We start with the initial combined LTV at origination. If there is a 2nd-lien loan, we include it in the calculation of this initial LTV. To update the LTV we take into account any paydown of the principal from the 1st lien mortgage over the next year. We then use the metro area OFEHO repeat-sale house price index to update the house value and LTV for each of the next four quarters following the origination. We take a simple average of these updated LTVs. To allow for potential nonlinear effects of LTV on early default, we enter the updated LTV as a series of indicators for different intervals. The left-out interval covers all LTVs below 80. We include indicators for the following LTV intervals: 80 to 84, 85 to 89, 90 to 94, 95 to 99, and 100 or higher (100+). To test whether investors react differently to the put option for default, we interact the investor indicator with the two highest LTV indicators, 95 to 99 and 100+.

The next risk factor is the DTI. LoanPerformance includes the "back-end ratio" which is calculated as the ratio of the sum of the annual mortgage principle and interest, property taxes and insurance, and other debt (such as car loans and student loans) to the income of the borrower. This variable is missing for 37% of the data. Appendix table A2 provides the results from estimating a probit on an indicator for whether DTI is missing on the other explanatory variables we use in this study. No-doc loans are much more likely to have a missing DTI, which suggests that lenders were less likely to code the variable if it is a "stated" item. There is a pattern to the missing DTI suggesting that as other risk factors of a loan deteriorate (ie higher LTV and/or lower FICO), the DTI is more likely to be reported in the data. We treat the missing DTI in two different manners to check for robustness. First, we include an indicator for whether DTI is missing and the DTI for those mortgages where it is reported. Second, we regress the reported DTI values on the other explanatory variables and replace the missing DTI values with the predicted DTI value from this regression. We also allow for nonlinearity in the effect of DTI on early defaults by entering DTI as a series of indicators for the following intervals: 40 – 44, 45 – 49, 50+. The left-out interval is for DTI less than 40.[17]

[17] We tested for whether creating subintervals of DTI below 40 would improve the fit. The data supported collapsing these into a single interval.

Next we turn to the credit or FICO score for the borrower. Freddie Mac uses three broad intervals for the FICO score for classifying mortgages: under 620, 620 to 680, and over 680. Their analysis suggests that only movements between these intervals are materially important in predicting performance of mortgages. We use a finer set of intervals that allow us to test whether variation in FICO scores within those broad intervals are informative for predicting early defaults for nonprime mortgages. We create a series of indicator variables for the following FICO intervals: less 560, 560 to 589, 590 to 619, 620 to 649, 650 to 679, and 680 to 719. The left-out interval is for FICO scores of 720 or higher.

The last of the four primary risk factors discussed in the introduction is the level of documentation used in underwriting the mortgage. We code an indicator for a limited-documentation mortgage and an indicator for a mortgage with no documentation. It is important to keep in mind when interpreting the coefficient estimates on these indicators for documentation level that the estimation procedure treats the stated risk factors for these lower documentation mortgages as being equivalent to the verified risk factors for fully documented mortgages.

We include a few additional variables related to the mortgage in our credit (X) vector. While the interest rates for all mortgages are fixed during the first twelve months that we follow their performance, we include an indicator for a fixed-rate mortgage. We include two indicators for mortgages that result from a refinance as opposed to a purchase – one for when no cash is taken out, and a second for when cash is taken out. As noted earlier, we include an indicator for an investor property as well as an indicator for a 2nd home.[18] All of our specifications include six indicators for different property types for the underlying collateral: condo, two-four unit, townhouse, planned-urban-development, manufactured housing, and other.[19]

Turn now to the variables we use to capture changes in the local economic conditions (Z) that may affect early defaults. For each mortgage, we take the average of the metro area unemployment rate over the twelve months following the origination. This variable should pick up income stresses that arise through job loss. We noted earlier that we control for the updated LTV based on changes in metropolitan house prices over the year following origination. To see

[18] It is likely that the investor variable is underreported in the data. It is reasonable to assume that misreporting is one-sided. That is, some investors hide this fact and are coded as owner-occupied; but owner-occupied borrowers are unlikely to misreport themselves as an investor. Appendix table A2 reports the results from running a probit of the investor indicator on our other control variables.
[19] Single family houses are the left-out property type.

if there are additional effects of changes in house prices beyond how they affect the LTV, we include the change in metropolitan house prices over the year following the origination of the mortgage, and we allow house price increases and decreases to have asymmetric effects on early defaults. In addition, we interact this house price appreciation with the investor indicator to test if investors react to house price movements in a different manner from owner-occupied borrowers. We also include year effects in all of our specifications.

Our specification is a reduced-form representation of the competing risks to which loans are exposed in their first year of life. To control for the risk of prepayment, we experimented with variables measuring changes in the interest rate environment borrowers faced when considering refinancing their loans. Specifically, we calculate for each month subsequent to mortgage origination the ratio of prime 30-year (and nonprime) mortgage rates to those prevailing at origination. We take the minimum of this ratio as a measure of the opportunity to refinance into a lower cost mortgage. In addition, we control for the presence of a prepayment penalty on the mortgage, which affects the probability of default by reducing the attractiveness of prepayment. Only the prepayment penalty was significantly related to early default behavior. In interpreting our results, it is important to bear in mind that the effects of house price declines on early default subsume both direct effects – determining the value of the put option - and indirect effects - reducing the availability of the prepayment option and thus extending borrower exposure to the risk of default (Caplin, Freeman and Tracy 1997). We return to these issues below.

Empirical Results

The linear probability model results are given in Table 10. The first specification is estimated with our sample of subprime mortgages and the second specification is estimated with our sample of alt-a mortgages. Descriptive statistics for each of these samples are provided in appendix table A1. The standard errors for both specifications use clustering on mortgages originated in the same metropolitan area, year and quarter.

The current LTV exerts a strong influence on early defaults for both classes of mortgages. The LTV marginal effects are relative to mortgages with LTVs below 80. The data

indicates that as the LTV increases, the likelihood of an early default rises by a similar amount for subprime and alt-a mortgages, with the incremental effect magnified for those mortgages that have negative equity. Compared to subprime borrowers with a current LTV below 80, borrowers with negative equity have an 6.8 percentage point higher early default rate (relative to an average early default rate of 6.6 percent). For alt-a borrowers, those with negative equity have a 6.9 percentage point higher early default rate compared to borrowers with a current LTV below 80 (relative to an average early default rate of 2.1 percent). While the absolute size of the negative equity effect is similar across subprime and alt-a loans, the relative impact is considerably larger for alt-a loans.

As discussed in the literature review, there may be differences between how "ruthless" investors are versus owners in exercising the default option on the mortgage. To test this hypothesis, we interact the indicator variables for the two highest LTV intervals with the investor indicator variable. For borrowers with negative equity, the data indicate that investors appear to be much more likely than owners to default. The point estimate for the incremental effect on the default rate is over 24.6 percentage points for subprime investors and 20.3 percentage points for alt-a investors.[20]

The second risk factor is the DTI which measures the ability of the borrower to make the mortgage payments. For each sample of mortgages, the coefficient on the linear DTI effect was similar across the two methods for handling the missing value for DTI. We report the specification in Table 10 that includes an indicator variable for a missing DTI. The data indicated that there is little effect of changes in DTI below 40 on early defaults. Increases in DTI above 40 have a small effect on increasing the likelihood of early default for subprime loans. Subprime borrowers who appear to be financially stretched as indicated by a DTI above 50, have a rate of early defaults that is 1.3 percentage points higher than borrowers who start out with a DTI below 40.[21] For alt-a loans, variations in DTI have a small and inconsistent pattern of effects on early default rates.

[20] If the investor variable is mismeasured along the lines discussed earlier, then these estimated differential effects for negative equity on early defaults are likely to be conservative.

[21] We also check to see if there is any interaction between high DTI and high LTV. We interact the indicator for DTI exceeding 50 with the indicators for LTV of 95-99 and LTV of 100+. Neither interaction was significant.

The third risk factor is the FICO credit score which is meant to capture a borrower's "willingness to pay." We tested whether the three broad intervals used by Freddie are appropriate for nonprime mortgages in predicting early defaults. The data strongly rejected the hypothesis that variation of the FICO score within the Freddie intervals was not predictive for early defaults. FICO scores exert a strong most influence on the early default behavior for owners. The reported marginal effects for each interval are relative to borrowers who have a FICO score of 720 or higher. The data indicates that as the FICO score declines below 680 for subprime loans the likelihood of early default rises. Subprime borrowers with a FICO score of 560 to 589 have a early default rate that is 6.9 percentage points higher than borrowers with a FICO above 720. This differential early default rate increases to 10.3 percentage points as we move to FICO scores below 560.[22] The data indicate a similar pattern of marginal effects of FICO scores on early default rates for alt-a owner-occupied loans. However, for alt-a investor loans there is no significant impact of variation of FICO scores on likelihood of a loan experiencing an early default.[23]

The final basic risk factor is the degree of documentation carried out in underwriting the mortgage. Controlling for the "stated" risk factors in these mortgages, the data indicate that low-doc underwriting is associated with a higher early default rate of around 3 percentage points for subprime loans and 1.3 percentage points for alt-a loans. Both of these effects exceed the simple weighted average of the early default differences listed in table 9. This is consistent with a degree of systematic bias in the statement of the risk profile of these mortgages relative to the risk profiles for fully documented mortgages.[24]

The next set of variables reported in table 10 refers to characteristics of the mortgage and whether the mortgage was for a purchase or a refinance. Controlling for the observed risk factors, borrowers who select a fixed rate mortgage have from a 0.6 to a 1.2 percentage point

[22] We explore whether the poor performance for low FICO borrowers is affected by the prevailing house price appreciation in the local housing market. We interact the house price appreciation with the three indicators for FICO scores below 620. The interaction is only statistically significant for the lowest FICO score interval, but the magnitude was small relative to the direct effect of this FICO score interval.
[23] We also tested for differences between owner-occupied and investor subprime loans on the impact of FICO scores and did not find them to be statistically significant.
[24] An obvious risk factor that may be biased downward is the DTI. Note, though, in appendix table A2 that DTI is often missing for low and no documentation mortgages.

lower incidence of early default.[25] We control for whether the mortgage has a prepayment penalty. The prepayment penalty applies if the mortgage is paid off during the first year, but is immaterial in the case of an early default. The data indicate that prepayment penalties are associated with a 0.7 percentage point higher early default rate for subprime loans, but have no significant impact on the early default rate for alt-a loans.[26] Over 60 percent of the subprime mortgages were initiated as a refinance rather than for a purchase. The data indicate that, holding constant the observed risk factors, early defaults were less likely for subprime mortgages that involved a refinance regardless of whether the borrower extracted cash or not.[27] For alt-a mortgages, only the cash out refinances were associated with a significantly lower early default rate, though the magnitude of the effect is smaller than for subprime cash out refinances.

Many of the discussions for providing assistance to preventing foreclosures exclude investors and owners of 2nd homes from the proposals. Justifications offered include that the costs of foreclosure differ for a household that is living in the home versus an investor or a vacation home. In addition, concerns over moral hazard from any government intervention tend to be magnified for investors and 2nd homes. We include indicators for investors and for 2nd homes to see if their average early default rates differ from owner-occupied primary residences. The data indicate that controlling for observed risk factors investors are more likely to default in the first year. On average, subprime investors have an early default rate that is 2.7 percentage points higher than for owner-occupied primary residences, while alt-a investors have a 1.3 percentage point higher early default rate. The data does not find any significant difference in early default rates for 2nd homes.

The final variables reported in Table 10 attempt to capture differences in the local economies and housing markets. We control for the change in house prices over the year since origination using the OFEHO repeat-sale house price for the MSA. We also control for the average unemployment rate in the MSA over the course of the year. It is important to keep in

[25] Future research should explore the extent to which this may reflect a selection effect on the borrowers and/or on the lenders. We are currently working with Professor Chris Mayer of Columbia University to merge in borrower characteristics by matching the LoanPerformance data with the HMDA data.
[26] One interpretation of this finding is that in the absence of the prepayment penalties some of the early subprime defaults might have refinanced instead during the first year.
[27] The data does not support the idea of "strategic early defaults" where borrowers sensing the turn in the housing markets extract all of the equity possible through cash-out refinances and then quickly default.

mind that we adjusted the initial LTV for the house price appreciation.[28] Even after factoring in the house price appreciation into the borrower's LTV, the data suggests a strong independent and asymmetric effect of price appreciation on early defaults. If house prices over the year rise by 10 percentage points in a local market, early defaults are reduced by 1.4 percentage points for subprime mortgages held by owners, and are reduced by 2.7 percentage points for subprime mortgages held by investors .[29] In contrast, if house prices decline by 10 percentage points in a local market, early defaults rise by 4.8 percentage points for subprime owners, and rise by 10.3 percentage points for subprime investors.[30] For alt-a owners, only house price declines exert an independent impact beyond the current LTV on early defaults. Finally, while we control for the initial DTI, we do not observe shocks to the borrower's income over time. To proxy for this, we include the average local unemployment rate. The data indicate that a 1 percentage point increase in the local unemployment rate is associated with only a quarter of a percentage point rise in early defaults for subprime mortgages, and no significant difference in early defaults for alt-a mortgages. This is consistent with the muted response of early defaults to the initial DTI.

Bad credit or bad economy?

Table 11 examines the question of the relative importance of credit effects versus economy effects in explaining the sharp rise in early defaults. The first column of the table reports differences from 2003 in the average early default rate by year starting with the rise in early defaults in 2005 (note that 2003 has the lowest average rate of early defaults for our sample period, 3.4% of subprime loans and 0.7 % of alt-a loans). This is the overall change in early defaults that the empirical model is trying to explain. The second column reports the difference in average early defaults predicted by our linear probability model, while the third column

[28] In contrast, Demyanky and Van Hemert (2008) control for the *initial* LTV and subsequent house price appreciation. Given data limitations, we have to assume that each borrower experienced the average house price appreciation for the msa based on the price index. If we had access to the estimated variances for these MSA price indices, we could generate a distribution of updated LTVs and calculate the probability that the borrower's LTV is in each of our intervals.

[29] Demyanky and Van Hemert (2008) report a similar marginal effect of house price appreciation. This is surprising since in their empirical specification the marginal effect of house price appreciation includes the indirect effect through the loan-to-value. Generally, we find the larger marginal effects for credit factors than they report.

[30] The incremental effect of house price decline for subprime investors relative to subprime owners is large but not precisely estimated in the data.

reports the fraction of the total change in average early default rates that is predicted by our model.

The final two columns disaggregate the explained rise in early defaults into components that are attributable to differences across years in underwriting standards and to economic conditions. The LTV indicators reflect aspects of both credit – the initial LTV – and economy – the effect of house price appreciation on the current LTV. To separate out these two influences, we create a set of counter factual LTV averages where for each year we take the initial LTV for a mortgage and adjust it using the house price appreciation from 2003 for the same metropolitan area and quarter of origination. Differences in these counter factual averages for the LTV indicators from their 2003 average reflect only the initial distribution of LTV in each year – a credit effect. The difference between the actual LTV averages for a given year and the counter factual averages for that year reflect only the differences in house price appreciation rates for that year and 2003 – an economy effect.

The major difference between 2003 and 2005-2007 was a dramatic change in house price appreciation. After rising nearly 14% in 2003, the OFHEO index accelerated to 16% in 2004 before slowing and eventually reversing. For 2005-2007, OFHEO grew 10%, 1% and –4% respectively.[31] The decomposition indicates that changes in economic variables, particularly this reversal in house price appreciation, from 2003-2007 account for the bulk of our explanation for observed increases in early defaults. In 2006, we estimate that changes in the economy added 2.4 percentage points to the average early default rate for subprime loans, while in 2007 that figure rises to 4.1 percentage points.

"Bad Credit," on the other hand, contributes less to our explained rise in average early defaults. Had the economy continued to produce unemployment and house price appreciation rates in 2005 through 2007 like those in 2003, our model predicts that changes in the credit profiles of new nonprime mortgages in each year would result in an increases in average early default rates for subprime loans of less than a percentage point in each year. For example, the model predicts that the average early default rate in 2006 for subprime loans would have increased by 94 basis points due to credit related factors. Of this, 44 basis points reflects changes in the initial LTV distribution; 8 basis points reflects worsening DTI by borrowers; changes in

[31] These figures refer to national average growth rates, but we use the MSA-level growth rates to conduct the experiment.

the distribution of FICO scores would have reduced early defaults by 17 basis points; shifts away from fully documented underwriting would have contributed 16 basis points; and the balance of 43 basis points reflects changes in the other credit factors.[32]

Since our benchmark year of 2003 produced remarkably strong house price appreciation that was subsequently sharply reversed, it may not seem surprising that we find a large role for the economy and a relatively minor role for changes in credit standards in our explanation of the rise in early defaults. After all, very high house price appreciation is sufficient to offset even a substantial upward shift in the initial LTV distribution. For example, the 2003 growth rate was sufficient to bring a property with an initial LTV between 100 to 109 to a current LTV below 95 where the incremental effect of LTV changes on early default is minimal. This raises the question of whether our result that "bad credit" played a relatively minor role in explaining early defaults would continue to hold if we used a more "normal" house price appreciation experience as our benchmark.

We tested this proposition by replicating the experiment described above, but subjecting each property value in both the base year (2003) and the comparison years to the 1985-2000 annual average OFHEO growth rate of 4.2 percent. This change only marginally affects the role of "bad credit" as an explanation for the rise in early defaults. In 2006, for example, the overall rise in defaults predicted by changes in credit standards rises from 94 basis points (shown in Table 11) to 100 basis points. This points out that even relatively modest increase in house prices would have been sufficient to have offset much of the upward drift in the distribution of initial LTV in 2006 and 2007.

Conclusion

We use loan-level data on securitized nonprime mortgages to examine what we refer to as "juvenile delinquency": default or serious delinquency in the first year following a mortgage's origination. Early default became much more common for loans originated in 2005-2007. Two complementary explanations have been offered for this phenomenon. The industry-standard

[32] Similar to Demyanyk and van Hemert (2008), we find that changes in the distribution of LTV accounts for a larger share of the credit related rise in early defaults than FICO, DTI and the level of documentation. However, in our results the contribution of LTV is smaller than that of changes in house prices.

explanation of default behavior focuses attention on a relaxation of lending standards after 2003. We see evidence of this in our data, as some underwriting criteria, particularly loan-to-value ratios at origination, deteriorated. At the same time, however, the housing market experienced a sharp and pervasive downturn, a factor which has received attention in recent research. Our results suggest that while both of these factors – bad credit and bad economy - played a role in increasing early defaults starting in 2005, changes to the economy appear to have played the larger role.

Perhaps as important a finding is that, in spite of the set of covariates we control for, our model predicts at most 43 percent of the annual increase in subprime early defaults during the 2005-2007 period. Observable changes in standard underwriting standards and key economic measures appear to be unable to explain the majority of the run-up in early defaults.[33] The fact, noted in our introduction, that many participants in the industry appeared to have been surprised by the degree of the increase in early defaults is in some sense verified here: observable characteristics of the loans, borrowers and economy seem to leave much unexplained, even with the benefit of hindsight. The difference between what we predict, conditional on observables, and what we actually observe is the difference between a bad few years for lenders/investors and a full-blown credit crunch.

The data does indicate a significant difference in behavior between owners and investors, especially in terms of how they respond to downward movements in house prices and negative equity situations. This has implications for underwriting. First, there may be payoffs to increased efforts at determining the true occupancy status of the borrower as part of the underwriting process. Second, originators may want to require additional equity up front from investors to reduce the likelihood that future house price declines could push the investor into negative equity.

An aim of our future research will be to improve the ability of the model to track the changes in average early default rates. A first step is to add more extensive demographic controls for the borrower. Second, given the estimated nonlinear response of the current LTV on early defaults, being able to estimate a distribution of current LTVs for each loan may be quite

[33] The model does a better job of explaining the rise in early defaults for alt-a mortgages in 2005 and 2006 with the model capturing two-thirds of the increase. However, in 2007 only 42 percent of the change is explained by the model.

important during periods of house price declines.[34] A third area of investigation is the possibility that a significant number of the investors misrepresented their status as an "owner".[35] The data indicate that investors appear to have a much stronger reaction in their early default decisions to negative equity and to declines in house prices. Identifying likely cases where the investor status is misrepresented could lead to significant improvements in the ability of the estimated model to track the rise in early defaults. Finally, it is possible that part of the rise in early defaults reflects changes in the composition of unobserved risk factors of borrowers. Further progress on this possibility would be facilitated by panel data that follows a borrower across multiple mortgages.

How much of the rise in juvenile delinquent mortgages reflected bad credit or a bad economic environment? Based on our evidence to date, a definitive answer is still elusive since too much of the rise in early defaults is remains unexplained. More work is needed to narrow the gap between actual and explained performance for nonprime mortgages. Of what is explained, Officer Krupke might have to concede that many of these juvenile delinquent mortgages "never had the house price appreciation that ev'ry mortgage oughta get."

[34] We are in discussions with OFEHO regarding the use of the metro-area house price variances.
[35] Fitch reports that in a small sample of subprime loans that defaulted 66% of the borrowers misrepresented their occupancy status.

References

An, Xudong, Raphael W. Bostic, Deng Yongheng, and Stuart A. Gabriel. "GSE Loan Purchases, the FHA, and Housing Outcomes in Targeted, Low-Income Neighborhoods." In *Brookings-Wharton Papers on Urban Affairs*, edited by Gary Burtles and Janet Rothenberg Pack, 205-240. Washington DC, Brookings Institution Press, 2007.

Black, Fischer, and Myron Scholes. "The Pricing of Options and Corporate Liabilities." *Journal of Political Economy* 81 (May 1973): 637-654.

Caplin, Andrew, Charles Freeman, and Joseph Tracy. "Collateral Damage: Refinancing Constraints and Regional Recessions." *Journal of Money, Credit, and Banking* 29 (November 1997): 496-516.

Cunningham, Donald F., and Patric H. Hendershott. "Pricing FHA Mortgage Default Insurance." *Housing Finance Review* 3 (October 1984): 373-392.

Demyanyk, Yuliya, and Otto Van Hemert. "Understanding the Subprime Mortgage Crisis." Working Paper. New York University, Stern School of Business, February, 2008.

Deng, Yongheng, and Stuart Gabriel. "Risk-Based Pricing and the Enhancement of Mortgage Credit Availability Among Underserved and Higher Credit-Risk Populations." *Journal of Money, Credit and Banking* 38 (September 2006): 1431-1460.

------, John M. Quigley, and Robert van Order. "Mortgage Terminations, Heterogeneity and the Exercise of Mortgage Options." *Econometrica* 68 (March 2000): 275-307.

Elul, Ronel. "Residential Mortgage Default." *Federal Reserve Bank of Philadelphia Business Review*, Third Quarter 2006, 21-30.

Fitch Ratings. *The Impact of Poor Underwriting Practices and Fraud in Subprime RMBS Performance*. US Residential Mortgage Special Report, 2007.

Foote, Christopher L., Kristopher Gerardi, and Paul S. Willen. "Negative Equity and Foreclosure: Theory and Evidence." Public Policy Discussion Papers 08-3. Federal Reserve Bank of Boston, 2008.

Foster, Chester, and Robert van Order. "An Option-Based Model of Mortgage Default." *Housing Finance Review* 3 (October 1984): 273-291.

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

Kau, James B., Donald C. Keenan, and Taewon Kim. "Transaction Cost, Supoptimal Termination, and Default Probabilities for Mortgages." *Journal of the American Real Estate and Urban Economics Association* 21 (Fall 1993): 247-263.

------, Donald C. Keenan, Walter J. Muller, and James F. Epperson. "The Valuation at Origination of Fixed Rate Mortgages with Default and Prepayment." *Journal of Real Estate Finance and Economics* 11 (July 1995): 5-36.

Leventis, Andrew. "Revisiting the Differences Between the OFHEO and S&P/Case-Shiller House Price Indexes: New Explanations." Research Paper. Office of Federal Housing Enterprise Oversight, January, 2008.

Mayer, Christopher, and Karen Pence. "Subprime Mortgages: What, Where, and to Whom?" FEDS Working Paper 2008-29. Board of Governors of the Federal Reserve System, June, 2008.

Pence, Karen M. "Foreclosing on Opportunity: State Laws and Mortgage Credit." FEDS Working Paper 2003-16. Board of Governors of the Federal Reserve System, May, 2003.

Pennington-Cross, Anthony. "Subprime Lending in the Primary and Secondary Markets." *Journal of Housing Research* 13, no. 1 (2002): 31-50.

------, and Giang Ho. "The Termination of Subprime Hybrid and Fixed Rate Mortgages." Working Paper 2006-042A. Federal Reserve Bank of St Louis, Research Division, July, 2006.

Schwartz, Eduardo S., and Walter N. Torous. "Mortgage Payment and Default Decisions: A Poisson Regression Approach." *Journal of the American Real Estate and Urban Economics Association* 21 (Winter 1993): 431-449.

Vandell, Kerry D. "How Ruthless Is Mortgage Default? A Review and Synthesis of the Evidence." *Journal of Housing Research* 6, no. 2 (1995): 245-264.

Table 1. Combined initial LTV

	Subprime						
Year	10th	25th	50th	75th	90th	% w. 2nd	N
2001	63	75	80	89	90	3.16	3,984
2002	65	75	80	90	97	3.56	6,636
2003	64	75	84	90	100	7.40	11,210
2004	65	78	85	95	100	16.07	17,093
2005	66	80	87	100	100	24.55	19,816
2006	66	80	90	100	100	29.36	14,756
2007	65	78	87	95	100	17.93	1,556

	Alt-A						
Year	10th	25th	50th	75th	90th	% w. 2nd	N
2001	64	75	80	90	100	2.17	1,107
2002	60	73	80	90	100	3.28	2,013
2003	51	67	80	90	100	12.36	3,746
2004	60	74	80	95	100	29.95	7,613
2005	60	74	80	95	100	36.39	12,164
2006	62	75	85	95	100	43.88	11,556
2007	60	75	80	95	100	34.82	2,079

Notes: Loan Performance data, authors calculations

Table 2. Early Defaults by Initial LTV and Year

		Subprime				
Year	Overall	< 80	80 – 89	90 – 94	95 – 99	100+
2001	6.70	6.97	6.98	5.50	5.37	8.99
2002	4.73	4.76	5.51	3.76	1.85	5.30
2003	3.40	3.23	3.45	3.42	2.95	4.03
2004	4.70	3.76	4.77	5.11	5.44	5.30
2005	6.60	5.35	6.15	7.15	7.83	7.60
2006	11.26	6.95	11.03	12.91	11.23	14.19
2007	14.52	9.20	14.18	16.94	14.58	19.94

		Alt-A				
Year	Overall	< 80	80 – 89	90 – 94	95 – 99	100+
2001	2.98	0.99	3.80	8.05	2.80	1.54
2002	1.39	0.74	1.75	2.68	1.18	1.72
2003	0.67	0.33	0.38	1.05	1.57	1.72
2004	0.85	0.28	0.69	0.80	1.64	2.13
2005	1.34	0.58	1.09	1.41	2.10	2.99
2006	3.21	1.21	2.32	3.64	4.51	6.47
2007	6.93	2.47	4.79	8.56	11.11	15.61

Notes: LoanPerformance data, authors calculations

Table 3. Early Defaults by Initial LTV, Investor status and Geography

	Subprime				
	< 80	80 – 89	90 – 94	95 – 99	100+
Owner	5.07	6.26	6.58	6.71	8.99
Boom/bust (AZ, CA, FL, NV)	4.22	5.47	5.40	6.19	9.39
Weak economy (IN, MI, OH)	7.82	9.21	10.24	9.68	9.74
Other	5.47	6.33	6.57	6.43	8.63
Investor	4.64	6.18	9.17	9.30	11.79
Boom/bust (AZ, CA, FL, NV)	2.90	3.51	6.18	8.51	3.45
Weak economy (IN, MI, OH)	11.56	12.65	18.33	13.43	22.81
Other	4.31	5.92	7.02	8.44	9.09

	Alt-A				
	< 80	80 – 89	90 – 94	95 – 99	100+
Owner	0.70	1.65	2.36	2.95	4.34
Boom/bust (AZ, CA, FL, NV)	0.66	1.56	2.46	3.83	5.28
Weak economy (IN, MI, OH)	0.00	3.68	2.68	3.97	4.15
Other	0.81	1.58	2.22	2.26	3.63
Investor	0.97	1.40	3.15	3.45	8.62
Boom/bust (AZ, CA, FL, NV)	0.67	1.30	2.58	5.42	17.58
Weak economy (IN, MI, OH)	2.74	3.62	4.76	7.69	5.71
Other	1.17	1.22	3.34	1.99	6.51

Notes: LoanPerformance data, authors calculations

Table 4. Distribution of Debt-to-Income Ratios – by Year

	Subprime			
Year	< 30	30 – 34	35 – 39	40 +
2001	49.05	9.06	9.76	32.13
2002	46.88	8.88	10.76	33.48
2003	42.05	9.49	11.01	37.45
2004	37.85	9.12	12.34	40.69
2005	40.54	7.82	11.00	40.64
2006	29.54	8.28	11.70	50.47
2007	37.53	7.78	10.54	44.15

	Alt-A			
Year	< 30	30 – 34	35 – 39	40 +
2001	78.05	5.33	6.87	9.76
2002	76.45	6.36	8.49	8.69
2003	75.04	6.54	7.82	10.60
2004	65.87	8.05	10.84	15.24
2005	65.42	7.88	11.16	15.54
2006	57.61	8.19	14.17	20.04
2007	58.01	6.97	13.56	21.45

Notes: LoanPerformance data, authors calculations

Table 5. Early Defaults by Debt-to-Income and Year

	Subprime			
Year	< 30	30 – 34	35 – 39	40 +
2001	6.09	7.76	6.94	7.27
2002	4.56	3.90	5.04	5.09
2003	3.20	3.10	3.57	3.64
2004	4.73	3.59	4.74	4.90
2005	6.48	5.87	5.69	7.09
2006	10.03	10.15	10.60	12.31
2007	13.87	14.88	15.24	14.85

	Alt-A			
Year	< 30	30 – 34	35 – 39	40 +
2001	3.59	0	1.32	0.93
2002	1.43	1.56	1.17	1.14
2003	0.71	0.41	0.34	0.76
2004	0.70	1.14	1.09	1.21
2005	1.12	1.15	1.84	2.01
2006	3.05	3.59	3.24	3.50
2007	5.72	4.83	7.09	10.76

Notes: LoanPerformance data, authors calculations

Table 6. Distribution of FICO Scores – by Year

	Subprime			
Year	< 600	600 – 619	620 – 659	660+
2001	51.53	13.50	19.78	15.19
2002	45.16	13.32	21.47	20.04
2003	39.68	12.58	23.68	24.06
2004	38.06	13.53	24.50	23.91
2005	35.59	14.34	25.76	24.30
2006	35.75	16.22	27.00	21.03
2007	39.91	17.16	24.04	18.89

	Alt-A			
Year	< 600	600 – 619	620 – 659	660+
2001	1.81	2.89	17.43	77.87
2002	2.24	2.78	14.41	80.58
2003	0.83	1.63	13.93	83.61
2004	0.71	1.29	14.93	83.07
2005	0.41	1.18	14.03	84.37
2006	0.14	0.84	16.79	82.23
2007	0.00	0.43	15.92	83.65

Notes: LoanPerformance data, authors calculations

Table 7. Early Defaults by FICO Scores and Year

	Subprime			
Year	< 600	600 – 619	620 – 659	660+
2001	9.16	5.02	4.82	2.31
2002	7.21	2.60	3.65	1.73
2003	5.51	3.55	1.85	1.37
2004	7.24	3.81	3.84	2.03
2005	9.92	5.91	5.35	3.45
2006	13.31	10.82	11.14	8.25
2007	15.14	15.36	15.24	11.56

	Alt-A			
Year	< 600	600 – 619	620 – 659	660+
2001	15.00	3.13	6.74	1.86
2002	2.22	3.57	2.76	1.05
2003	12.90	3.28	0.96	0.45
2004	1.85	4.08	1.58	0.66
2005	6.00	5.56	2.34	1.09
2006	12.50	10.31	4.95	2.77
2007	0.00	22.22	13.29	5.64

Notes: LoanPerformance data, authors calculations

Table 8. Distribution of Documentation Level – by Year

	Subprime		
Year	Full	Low	None
2001	77.84	21.76	0.40
2002	71.13	28.30	0.57
2003	67.02	32.52	0.46
2004	65.37	34.34	0.29
2005	62.28	37.47	0.24
2006	61.71	38.00	0.29
2007	64.20	35.48	0.32

	Alt-A		
Year	Full	Low	None
2001	36.77	55.56	7.68
2002	40.64	51.96	7.40
2003	35.50	57.26	7.23
2004	37.75	55.72	6.53
2005	31.11	64.44	4.46
2006	18.92	76.56	4.53
2007	16.84	77.49	5.68

Notes: LoanPerformance data, authors calculations

Table 9. Early Defaults by Documentation Level and Year

	Subprime		
Year	Full	Low	None
2001	6.58	7.15	6.25
2002	4.49	5.43	0.00
2003	3.31	3.54	5.77
2004	4.62	4.87	2.04
2005	6.05	7.54	0.00
2006	9.26	14.55	4.65
2007	11.91	19.20	20.00

	Alt-A		
Year	Full	Low	None
2001	0.98	3.41	9.41
2002	0.98	1.34	4.03
2003	0.45	0.75	1.11
2004	0.77	0.94	0.60
2005	0.77	1.57	2.03
2006	2.06	3.47	3.63
2007	3.14	7.57	9.32

Notes: LoanPerformance data, authors calculations

Table 10. Probability of an Early Default

Variable	Subprime (1)		Alt-A (2)	
LTV: 80 – 84	0.95**	(0.30)	0.26	(0.24)
85 – 89	1.40**	(0.28)	1.24**	(0.26)
90 – 94	2.02**	(0.36)	1.38**	(0.29)
95 – 99	3.07**	(0.42)	2.91**	(0.38)
100+	6.84**	(1.26)	6.92**	(0.89)
Investor • 95 – 99	1.78	(2.45)	3.77*	(1.61)
• 100+	24.59*	(12.35)	20.29**	(5.96)
DTI missing	0.41	(0.22)	0.37*	(0.16)
DTI: 40 – 44	0.71*	(0.28)	0.59*	(0.31)
45 – 49	0.77**	(0.28)	0.47	(0.44)
50+	1.26**	(0.39)	–0.90*	(0.42)
FICO: <560	10.35**	(0.42)	11.81**	(4.57)
560 – 589	6.92**	(0.40)	6.23**	(2.29)
590 – 619	4.47**	(0.36)	4.77**	(1.14)
620 – 649	3.32**	(0.35)	2.98**	(0.33)
650 – 679	1.72**	(0.32)	1.75**	(0.23)
680 – 719	0.09	(0.34)	0.68**	(0.14)
Investor • <560			–11.84**	(4.59)
• 560 – 589			–6.97**	(2.37)
• 590 – 619			–4.68**	(1.18)
Limited documentation	2.96**	(0.22)	1.27**	(0.15)
No documentation	0.73	(1.08)	2.14**	(0.38)
Fixed rate mortgage	–1.17**	(0.20)	–0.57**	(0.17)
Prepayment penalty	0.70**	(0.21)	0.28	(0.18)
Refinance – no cash	–2.65**	(0.35)	–0.11	(0.20)
Refinance – cash out	–3.11**	(0.25)	–0.61**	(0.16)
Investor	2.73**	(0.65)	1.29**	(0.35)
2nd – home	1.37	(0.96)	0.51	(0.44)
House price appreciation − positive (10%)	–1.42**	(0.14)	–0.02	(0.11)
House price appreciation − negative (10%)	4.80**	(1.03)	1.66*	(0.69)
Investor • house price appr − positive (10%)	–1.32**	(0.39)	–0.44*	(0.18)
Investor • house price appr − negative (10%)	5.57	(4.57)	–0.74	(1.25)
Local unemployment rate	0.25**	(0.08)	–0.04	(0.06)
Root mean square error	0.243		0.139	
Mean early default rate	6.61		2.06	
Observations	75,051		40,278	

Notes: Linear probability estimates with standard errors given in parentheses. Standard errors use clustering at the msa•year•quarter level. LoanPerformance data, 1 percent random sample. Year effects and six property type fixed effects are included.

** significant at the 1 percent level

* significant at the 5 percent level

Table 11. Decomposition of Rise in Early Default Rates

	Subprime				
	Early Default		Percent	Due to:	
Year	Difference	Explained	Explained	Credit	Economy
2005	3.20	0.90	28.1	0.58	0.32
2006	7.86	3.39	43.1	0.94	2.45
2007	11.13	4.80	43.2	0.72	4.08

	Alt-A				
	Early Default		Percent	Due to:	
Year	Difference	Explained	Explained	Credit	Economy
2005	0.67	0.44	65.8	0.20	0.24
2006	2.54	1.70	66.8	0.55	1.14
2007	6.26	2.61	41.7	0.49	2.12

Notes: All differences expressed relative to 2003. LoanPerformance data, authors calculations using the estimates reported in table 10 and sample means by year for the variables.

Table A1: Descriptive Statistics

	Subprime		Alt-A	
	Mean	Standard Deviation	Mean	Standard Deviation
Early default	6.61	24.84	2.06	14.20
Loan size ($000)	189.9	123.4	291.8	225.3
Initial interest rate	7.86	1.43	5.60	2.18
LTV	74.37	19.69	73.21	19.73
80 – 84	0.105	0.306	0.079	0.269
85 – 89	0.122	0.327	0.095	0.293
90 – 94	0.098	0.297	0.099	0.299
95 – 99	0.114	0.318	0.103	0.304
100+	0.018	0.133	0.028	0.165
DTI	28.92	19.68	16.41	19.10
40 – 44	0.150	0.357	0.098	0.298
45 – 49	0.188	0.390	0.051	0.221
50+	0.083	0.276	0.018	0.132
FICO	617.49	60.54	708.37	48.10
< 560	0.184	0.388	0.001	0.034
560 – 589	0.147	0.354	0.003	0.052
590 – 619	0.188	0.391	0.010	0.101
620 – 649	0.193	0.395	0.104	0.305
650 – 679	0.138	0.345	0.177	0.381
680 – 719	0.091	0.288	0.298	0.457
Limited documentation	0.344	0.475	0.654	0.476
No documentation	0.003	0.058	0.054	0.226
Fixed rate mortgage	0.232	0.422	0.436	0.496
Prepayment penalty	0.722	0.448	0.368	0.482
Refinance – no cash	0.083	0.276	0.162	0.369
Refinance – cash out	0.547	0.498	0.327	0.469
Investor	0.073	0.259	0.213	0.409
2^{nd} – home	0.010	0.100	0.034	0.182
House price appreciation	9.559	9.138	8.402	10.108
Local unemployment rate	5.126	1.309	4.851	1.260
Origination year: 2001	0.053	0.224	0.027	0.163
2002	0.088	0.284	0.050	0.218
2003	0.149	0.356	0.093	0.290
2004	0.228	0.419	0.189	0.391
2005	0.264	0.441	0.302	0.459
2006	0.197	0.397	0.287	0.452
2007	0.021	0.142	0.052	0.221

Note: LoanPerformance data, 1 percent random sample. Sample sizes: 75,051 subprime loans and 40,278 alt-a loans.

Table A2. Probability of an Investor and Missing DTI

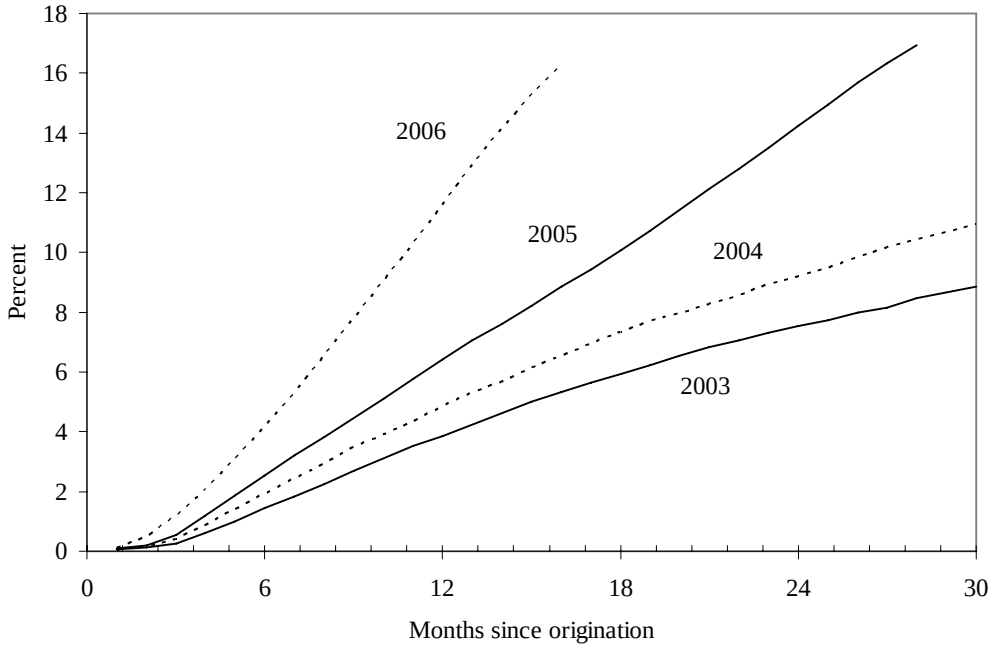
Variable	Investor (1)		Missing DTI (2)	
LTV: 80 – 84	–0.35	(0.30)	–2.50**	(0.51)
85 – 89	–1.41**	(0.23)	–2.67**	(0.48)
90 – 94	–6.53**	(0.16)	–2.29**	(0.52)
95 – 99	–7.03**	(0.13)	–4.31**	(0.54)
100+	–8.48**	(0.12)	–8.29**	(1.02)
Investor • 95 – 99			–4.15	(2.13)
Investor • 100+			0.28	(6.08)
DTI missing	–1.40**	(0.19)		
DTI: 40 – 44	–3.34**	(0.22)		
45 – 49	–3.01**	(0.23)		
50+	–2.25**	(0.33)		
FICO: <560	–9.97**	(0.13)	–14.88**	(0.52)
560 – 589	–8.74**	(0.13)	–10.44**	(0.57)
590 – 619	–8.23**	(0.15)	–9.93**	(0.52)
620 – 649	–6.94**	(0.16)	–10.52**	(0.48)
650 – 679	–4.99**	(0.18)	–7.84**	(0.48)
680 – 719	–2.21**	(0.20)	–4.29**	(0.48)
Limited documentation	–0.55**	(0.17)	1.89**	(0.32)
No documentation	–4.98**	(0.29)	56.12**	(0.63)
Fixed rate mortgage	2.17**	(0.20)	16.51**	(0.34)
Prepayment penalty	–1.24**	(0.18)	–12.34**	(0.31)
Refinance – no cash	–3.38**	(0.20)	–2.91**	(0.50)
Refinance – cash out	–5.03**	(0.19)	–3.58**	(0.36)
Investor			0.62	(0.66)
2nd – home			1.89	(1.09)
House price appreciation	–0.16**	(0.01)	–0.24**	(0.02)
Investor • house price appr			0.10*	(0.05)
Local unemployment rate	0.05	(0.07)	–0.69**	(0.13)
Mean of dependent variable	12.21		37.57	

Notes: Probit marginal effects with standard errors given in parentheses. Left-out year is 2003. LoanPerformance data, 1 percent random sample. Year effects and six property effects are included. Each marginal effect reflects a percentage point change in the dependent variable.

** significant at the 1 percent level

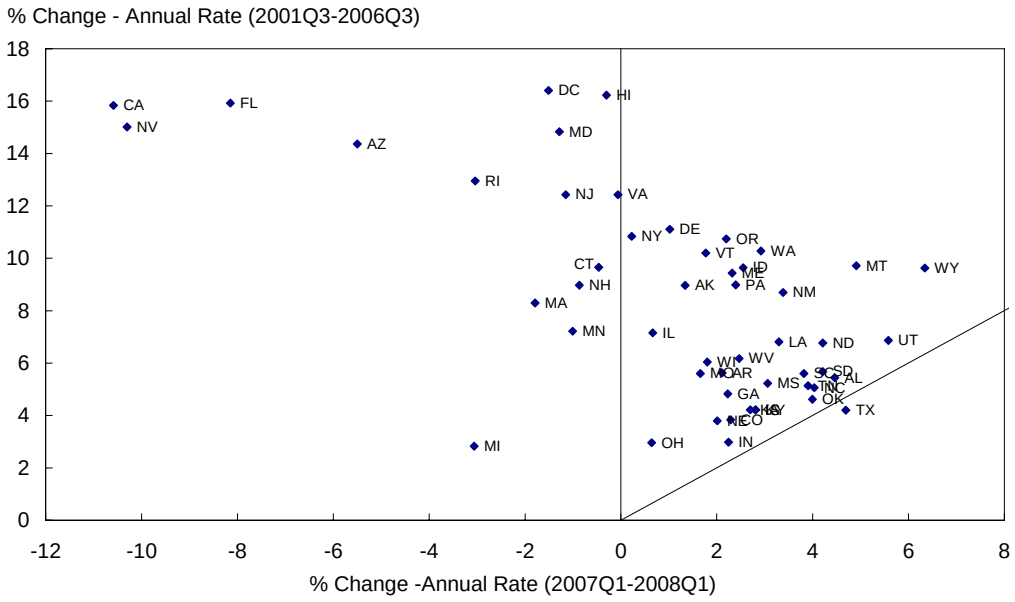
* significant at the 5 percent level

Figure 1. Nonprime 90+ Days Delinquencies – by vintage



Notes: FirstAmerican CoreLogic LoanPerformance

Figure 2. House Price Appreciation Over Time – by State



Source: Office of Federal Housing Enterprise Oversight