

ARE MORTGAGE BORROWERS REALLY CLUELESS? STRATEGIC BEHAVIOR, ENDOGENOUS CREDIT SCORES, AND THE NEXT FINANCIAL CRISIS

RAJEEV DAROLIA^{*}, BRENT C. SMITH[#], AND ANTHONY M. YEZER⁺

JUNE, 2015

ABSTRACT. There is a conflict between assumptions of empirical models of mortgage default and economic theories of credit market behavior. Competing hazard models of credit default and prepayment that assume loan terms, such as loan-to-value ratio (LTV), and borrower characteristics are exogenous. Decisions to price credit risk and reject applicants are assumed to have no effect on loan terms requested or credit score.

In contrast, models of rational borrowers predict that loan terms requested by borrowers are chosen strategically considering the probability of default and prepayment. Thus far, the possibility that applicants also manipulate indicators of credit risk strategically has been given little attention. This paper combines theoretical and empirical evidence that credit score is also manipulated by applicants. Empirical tests demonstrate that differences in mortgage applicant incentives to raise credit score artificially explain the subsequent decline in credit score after endorsement. Paradoxically, when lenders price increasing credit risk, borrower responses invalidate the predictions of default models used to assess risk. This provides one explanation for the high proportion of early payment defaults (EPDs often never make a single payment) in the financial crisis and for research showing that competing hazard default models substantially under-predict defaults on mortgages for 2006 through 2008 even when the actual changes in economic conditions, including falling house prices, are used in making the predictions. Given that these models are still being used, the forecast that they will fail again seems assured.

Key Words: Mortgage pricing, competing hazards, default, financial crisis

JEL Categories: D10, D11, F01, G21

^{*} Assistant Professor of Public Affairs, University of Missouri

[#] Associate Professor of Finance and Real Estate, Virginia Commonwealth University

⁺ Professor of Economics, George Washington University, <yezer@gwu.edu>

ARE MORTGAGE BORROWERS REALLY CLUELESS?: STRATEGIC BEHAVIOR, ENDOGENOUS CREDIT SCORES, AND THE NEXT FINANCIAL CRISIS

I. INTRODUCTION

Mortgage lending in the U.S. has increasingly relied on automated underwriting and price rationing. Secondary market participants promoted automated underwriting with the objective of lowering transaction costs. Price rationing took the form of subprime and Alt-A mortgage lending which allowed applicants, who would have been rejected for prime mortgages, to obtain credit at higher interest rates. Indeed, risk-based pricing even became common within the prime market. These and other financial innovations profoundly altered the mortgage credit supply curve facing households between 1990 and 2000.

Many lenders apparently assumed that the increasing use of price rationing in the supply of mortgage credit would have no effect on applicant behavior beyond well understood changes in the likelihood and timing of prepayment. Applicant characteristics and loan terms were assumed to be econometrically exogenous in joint hazard models of default and prepayment that provided assessments of credit risk for underwriting and pricing purposes. However, it is evident that borrowers reacted to the expansion of lending under price rationing. Garmaise (2013) finds the distribution of risk in the applicant pool changes significantly in ways that are not easily detected using standard underwriting when a lender adds the option ARM, a higher risk product. The result is a substantial increase in losses on non-option ARM products offered by the lender. Clearly, applicants respond to their perceptions that supply conditions have altered by changing their behavior significantly.

While strategic applicant behavior has received little attention in the empirical literature it is at the heart of many theoretical models of the loan market. The equilibrium credit rationing literature stemming from Stiglitz and Weiss (1981) is based on the assumption that applicants respond to lending criteria. In this literature, raising interest rates prompts both adverse selection and borrower incentive effects that cause credit risk to rise. Since this seminal paper, a substantial number of theory papers have been based on the presumption that applicants are informed and behave strategically. The limit of this strategic behavior is fraud in which applicants with no intention to repay enter the market.

This paper models and tests borrower efforts to lower the cost of mortgage credit by manipulating lender estimates of credit risk. Some examples of strategic applicant behavior have already been documented in the literature.¹ Borrowers clearly increase downpayments to avoid paying mortgage insurance. They secure cosigners, and provide more documentation for income or wealth to obtain favorable credit terms. Such adjustments are relatively straightforward.

The possibility that applicants go further and take actions to raise their credit scores has received little academic attention. However, a simple web search under the title “improve my credit score” yields hundreds of “proven” methods that applicants are advised to use. Moving from a market where there is non-price credit rationing to one in which there is an individual credit supply curve with rates that vary inversely with credit score, provides a strong rationale for applicants to search for strategic options to increase their scores.² Furthermore, the incentive to

1

² Obviously there is also an incentive to lower loan-to-value ratio, and payment to income ratios but the focus of this paper is on endogenous credit scores as opposed to loan terms.

engage in these efforts varies significantly when the function relating APR to credit score is non-linear. Some efforts to raise scores are fairly straightforward and impose modest cost on the applicant. Others are more elaborate. Avery, Brevoort, and Canner (2010) have an excellent analysis of the market for “piggybacking” in which authorized users are added to credit cards for a fee.³ In extreme cases, identified in Garmaise (2011) and Carrillo (2011), actual fraud was involved with the apparent cooperation of the loan officer.

In practice, the underwriting process at loan origination incorporates both hard and soft data on the real property collateral and the borrower. These are used in evaluating and pricing default potential. The credit score is one of the few variables used by the underwriter that is not directly provided by the borrower and thus, *presumed* to be outside the scope of borrower reporting bias. Credit scores have become a primary determinant of a borrower’s cost of credit. Their original intent when created in the 1950s was evaluating borrower capacity for short term finance and revolving credit. With the advent of new algorithms their use expanded to include long term indebtedness. Credit scores have become an important component of risk assessment and pricing in the residential mortgage market. Straka (2000) notes that credit scoring and automated underwriting systems combined to lower lender losses and increase profits throughout the 1990s. Fischelson-Holstein (2005) demonstrates that credit scores offer lenders more accurate predictions than those produced by manual underwriting.⁴ The seminal work by Deng, Quigley, and Van Order (2000) increased the role of formal models in mortgage pricing as they

³ The credit scoring industry has recognized the authorized user problem and taken steps to make the addition of names to accounts with excellent payment histories less influential in determining credit scores.

⁴ Recently, Einav, Jenkins, and Levin (2013) have demonstrated that switching to automated credit scoring raised profits by \$1,000 per loan at a large automobile finance company. High risk borrowers were avoided and larger loans were made to low risk borrowers. Given the substantially smaller amounts borrowed on automobile loans the incentives for strategic manipulation of credit scores identified here are likely much smaller in the automobile loan market than for home mortgages.

explained the important roles of credit scores and mortgage terms in predicting the joint hazards of default and prepayment terminations.

Currently underwriting and pricing of mortgages is based on estimates of competing hazard models of default and prepayment. These models assume that loan terms and characteristics like credit score are exogenous. Put another way, the expected probability of future default or prepayment is not supposed to influence the loan terms requested. The argument made in this paper builds on an old literature, see Barth, Cordes, Yezer (1981) or Maddala and Trost (1982), which provides empirical evidence that loan terms are endogenous to the loan transaction. The extreme example of this is a fraudulent loan applicant who has no intention of repaying the loan. Such applicants only care about the loan to value ratio. Monthly payment, term to maturity, and interest rate are irrelevant because the intent is default.⁵ Clearly, in such cases loan terms are endogenous and the hazard models used to calibrate credit scoring and mortgage pricing schemes are biased and inconsistent.⁶

This paper extends the old argument about endogenous loan terms to credit scores. Applicants have an incentive to manipulate their credit scores whenever they perceive that the marginal benefits in terms of more favorable loan terms exceed the marginal costs of raising scores. The next section of this paper discusses alternative explanations for the substantial losses

⁵ Cases of early payment default (EPD) in which no payments are made by the borrower may be motivated by a larger criminal enterprise which involves sales at prices well above market value and renting the unit to unsuspecting individuals during the period before foreclosure. This is often termed “appraisal fraud.” Another form of fraud is “shotgunning” in which multiple mortgages are taken out on the same property at the same time. The FBI (2011) reports estimates of losses from mortgage fraud at \$27 billion for 2006.

⁶ Curiously, while default is treated as exogenous, the endogenous nature of the prepayment decision is well recognized in the literature where interest rates, prepayment penalties, and prepayment itself are clearly seen as jointly determined endogenous variables. When it comes to prepayment behavior, borrowers are assumed to be close to fully rational, including their expectations for future interest rates and for their credit scores to fall in a fashion that will allow refinancing out of subprime products.

experienced by mortgage lenders on the books of business endorsed after 2005 and considers the incremental contribution made by the effects of endogenous loan terms and credit scores. Then a model of applicant-lender interaction is used to develop implications of endogenous credit scores that can be the object of empirical testing. In the two empirical sections measures of the relation between credit scores and APR including the variation in this relation over time are developed, and then a test of the relation between the incentive to raise scores and the change in scores following endorsement is implemented. The potential importance of the findings is discussed in the concluding section.

II. ALTERNATIVE APPROACHES TO APPLICANT BEHAVIOR

The dominant view of credit risk in mortgage lending ignores borrower expectations, i.e. the implicit assumption is that borrowers are clueless. In these models, unexpected changes in collateral value and/or borrower cash flows cause default. These are the option-based and triggering-events theories of default. The options-based theory emphasizes the role of unanticipated changes in equity, positing that default occurs among borrowers in response to unanticipated declines in equity. Triggering-event theories, discussed in Avery, Bostic, Calem, and Canner (2000), emphasize the financial condition of the borrower and the vulnerability to significant adverse changes in income or expenditure. Ultimately default as subject to an event that negatively impacts the borrower's personal wealth and/or cash flow often related to unemployment or medical events. While these theories emphasize different determinants of mortgage default, they share the common view that the house price, wealth, or cash flow disturbances which cause default are unexpected, ex post events. This paper does not dispute the potential importance of house price, income, or expenditure shocks in explaining default. However, it does question whether these events are completely unanticipated by borrowers along

with the assumption that expectations for these events play no role in the application process. It may be that some plant closings, layoffs, medical problems, marital strife, drug dependency, etc are “negative shocks” that were not anticipated. However, the current assumption that all of life’s difficulties arrive randomly after the mortgage is endorsed and hence have no effect on pre-endorsement behavior seems incredible.

A host of papers have explained problems associated with the financial crisis in a fashion consistent with the ex-post view of mortgage default. All of these models have one element in common. Borrowers are passive, often clueless participants in the mortgage lending process. They not only fail to engage in strategic behavior, often they are completely passive, and easily persuaded to do things that are not in their best interest. Other than a few papers dealing with applicant fraud on EPDs, there has been little attention paid to the possibility that borrowers behaved strategically.⁷

One common view of the rise in default rates is the increased use of more complex mortgage instruments. In particular, Anderson and Dokko (2011) find evidence that subprime products which subject borrowers to rising monthly payments were associated with higher default rates presumably because of cash flow problems that should have been anticipated.⁸ Another view is that use of these products was associated with units where negative equity was particularly large. Willen, Pence, and Sherlund (2009) is representative of several papers concluding that falling house prices rather than exotic mortgage products were more likely to be

⁷ For empirical studies detecting fraud based on early payment default, see Carrillo ()

⁸ As noted in a recent review by Lusardi and Mitchell (2014), tests indicate that current levels of consumer financial literacy are dismal. This suggests that many borrowers signed contracts without understanding the implications for future payments. The existence of such clueless borrowers, does not imply that many borrowers fail to behave strategically as argued in this research. It may well be that some borrowers rely on the advice of others when they apply for mortgage credit.

the major cause of default and foreclosure. This reflects a victory of the option based approach over what might be termed the built in trigger event theory.

Rajan, Seru, and Vig, (RSV) (2010, 2010) have also noted the failure of models that predict default but their analysis places the blame on lenders and securitization. They argue that credit risk models include both hard (verifiable) and soft variables. Given a ready market for product from buyers of private label MBS, lenders concentrated on the quality of the hard variables that were used by investors to estimate credit and prepayment risk. The result was declining interest in the quality of soft variables. While this may appear to be similar to the arguments made in this paper it is quite different, although certainly not inconsistent. Hard variables used by investors include credit score and loan terms in the RSV analysis. These hard variables are not to blame for the high loss rates in their model. It may be that RSV are correct in claiming that deteriorating soft variables enhanced the failure of default models to predict default, but hard variable problems are at the heart of the explanation offered here.

While the debate over option based and trigger event models of default has raged, on other line of empirical research has produced a surprising result. Examination of the performance of default models over the period after 2006 indicates that they substantially under-predict default even when the actual change in house prices, along with other actual characteristics of the market, were used to forecast default based on models fit for data prior to that period. For example, An, Deng, Rosenblatt, and Yan, (2010) fit a competing hazard model for mortgages endorsed in 2003 and use it to forecast defaults in the 2006 vintage subprime stock. They under-predict defaults by up to 40%. Kau, Keenan, Lyubimov, and Slawson (2011) report that the parameters of default model estimates vary over time and by type of mortgage during this period. Given that these models are estimated on massive data sets using maximum

likelihood techniques, the problems in forecasting defaults in mortgages endorsed after 2005 by substituting actual house prices into the models are not likely due to failures of data quantity, or estimation method. There is clearly a problem with the forecasting model and the option and trigger event explanations that support it. If high quality academic models fail, it is reasonable to assume that the models used by lenders would be worse at predicting actual default.

Garmaise (2011) reviewed loan files carefully and documented borrower misrepresentation of assets as a problem in defaulted loans. Carrillo (2011) has demonstrated methods to identify cases of organized fraud in local mortgage markets. His approach was subsequently validated by criminal prosecution of an organized gang. These studies imply that applicants, and their collaborators, went well beyond the type of credit score manipulation needed for the results observed here. The existence of significant fraud as a response to mortgage underwriting suggests that the types of credit score manipulation required for the problems discussed here would be relatively easy to arrange. Put another way, compared to misrepresentation of income, assets, and employment, the manipulation of credit scores along the lines suggested here would be relatively easy to achieve. This research identifies strategic behavior by borrowers as an important source of credit risk.

In sum, many factors have been correctly identified as contributing to high rates of mortgage default in the housing bust following 2006.⁹ Many of the losses are accounted for by traditional credit risk models estimated assuming that loan terms and credit history are exogenous. However, it is clear that more factors were at work in producing the losses. If the

⁹ The literature review presented here has been abbreviated to show suggestive works. A full listing of all papers dealing with the role of the options based and trigger event approaches to default as well as the effects of lender forbearance would run to dozens of citations.

parameters of the hazard models predicting default and prepayment are not stable, then there is a problem with using these models going forward. The theory and empirical results in this paper suggest that the failure of credit risk models arises due to their assumption that loan terms and even borrower characteristics such as credit history are exogenous to the default option.

III. A MODEL OF THE BORROWER BEHAVIOR

The problem of endogenous credit scores and their effect on mortgage markets may be illustrated with a simple theoretical model in which rational borrowers seeking to minimize the cost of credit respond to lender attempting to price credit risk. This model identifies implications of credit score manipulation that are the object of testing in subsequent empirical sections.

A price rationing lender sets interest rates based on perceptions of credit risk of each applicant and on the opportunity cost of loan funds, which is assumed constant here. Credit risk can be expressed as the product of the probability of default and the expected loss given default:

$$r = l(p) \tag{1}$$

here r is credit risk or expected loss, l expected loss given default, and p is the probability of default from a statistical default model.

In most default models, the probability of default is found to be a function of mortgage terms, particularly the payment to income and loan to value ratios, credit score, and variables reflecting local market conditions. The general form of the default function is:

$$p = P(m, f, s, z) = P(M(i), f, s, z); P_m > 0, P_{mm} > 0, P_f > 0, P_{ff} > 0, P_s < 0, P_{ss} > 0, P_z < 0 \tag{2}$$

where m is the payment to income ratio, f is the loan to value ratio, s is a credit score, z is a vector of market conditions (particularly the expected rate of house price appreciation), i is the note rate, and $M(i)$ is a function relating the note rate and the payment to income ratio, i.e. $M_i > 0$. There are examples of empirical default models whose general form is represented by equation (2) in the academic literature discussed above and lenders have adopted these for use in underwriting models. Conditional loss functions have the following form:

$$l = L(f, z); L_f > 0, L_z < 0 \quad (3)$$

so that the expected loss given default rises with the loan to value ratio and falls with favorable economic conditions, particularly higher rates of expected housing price appreciation.

Substituting (2) and (3) into (1) yields a general expression for credit risk:

$$r = L(f, z) P(M(i), f, s, z) \quad (4)$$

Lenders who engage in risk-based pricing evaluate individual applicants and set the note rate according to $i = I(r)$, $I_r > 0$. Substituting (4) into this credit supply equation and differentiating totally gives an expression for the determinants of the supply price of credit offered to individual applicants:

$$i = I(L(f, z) P(M(i), f, s, z)) \quad (5)$$

Totally differentiating (5) and solving for the relation between interest rate and credit score yields:

$$di/ds = \lambda I_r l P_s < 0, \text{ where } \lambda = 1/(1 - I_r l P_m M_i) \quad (6)$$

where $0 < I_r l P_m M_i < 1$ is the effect of an increase in the interest rate on the interest rate. This is clearly greater than zero as all partial derivatives are positive but must be less than unity for stability of the supply relation. Accordingly $\lambda > 1$ and can become quite large as P_m rises which induces a natural limit on the payment to income ratio that lenders are willing to accept.¹⁰ Differentiating again with respect to credit score, $d^2 i / ds^2 = \lambda I_r l P_{ss} > 0$ and the credit supply relation between credit score and interest rate is convex. This relation is demonstrated later in the empirical section of this paper but it is well known in the literature on the supply of mortgage credit. For applicants who are accepted, the pricing of credit risk on the lender's rate sheet reveals that $di/ds = \lambda I_r l P_s < 0$. This provides an incentive for accepted applicants to expend effort and funds to raise s and qualify for mortgage credit at a lower interest rate.

A scan of the internet reveals that there are many steps advertised that can, without recourse to deception, raise credit scores. Other less scrupulous, but perhaps legal, opportunities exist as well. Accordingly, applicants are assumed to have a credit score opportunity function of the following form:

$$s = S(c; s^*) \quad S_c > 0, S_{cc} < 0, S_{cs^*} < 0 \quad . \quad (7)$$

where s is the final credit score achieved, c is a measure of the cost, including both explicit and implicit costs, of actions that can be taken to raise an individual's credit score and s^* is the initial credit score which would be observed by the lender if $c = 0$, i.e. $s^* = S(0; s^*)$. The general nature of the $S(c; s^*)$ function is based on the limited literature, particularly Avery, Canner, and Brevort (), which has characterized the properties of credit scoring and default equations. First, the marginal cost of raising credit scores above the "natural" or current level is always positive, $S_c >$

¹⁰ In determining the profit-maximizing interest rate, the lender will select a maximum value for the payment to income ratio that is well below the level at which $I_r l P_m M_i$ equals unity.

0. Second, the cost increases at an increasing rate because achieve small improvements in scores may be achieved by simply correcting errors in the credit report. However, large improvements in credit score required significant changes in financial condition including paying off delinquent accounts and paying firms that provide credit counseling services. Thus, $S_{cc} < 0$, the credit score improvement function is convex. It is more costly to raise a credit score that is high already because most opportunities to raise the score have been exhausted and there is a upper limit on scores so that $S_{cs^*} < 0$. The cost of increasing credit score is quite low for those with initially low scores who are seeking a small rise in score. However, the cost is much larger for a large improvement in a low score, or a modest increase in a score that is already high.

An informed, rational applicant seeking to minimize the cost of credit will respond strategically to the perceived lower cost of credit associated with raising her credit score, i.e. she will respond to $di/ds = \lambda I_r l P_s$. The applicant will expend c resources in an effort to raise her credit score until the marginal gain achieved from the lower interest rate is equal to the marginal cost or:

$$-\theta \lambda I_r l P_s = I/S_c \quad (8)$$

Here θ is a scaling parameter reflecting the relation between a change in the rate of interest and the discounted present value of the cost of the loan.¹¹ Equation (8) is easily interpreted as a necessary condition that the marginal benefit of raising credit scores, $-\theta \lambda I_r l P_s$, be equal to the marginal cost of taking the actions to raise the score, I/S_c .¹² The exact form of the functions that

¹¹ θ varies with the loan amount and the expected time until prepayment or default as well as the applicant's rate of time preference. These are all assumed constant here. A formal derivation of (8) is presented in the appendix A.

¹² Equation (8) is necessary but not sufficient because it may be that, even for small levels of c , $-\theta \lambda I_r l P_s < I/S_c$ because both P_s and S_c approach 0 as s^* rises. This case is illustrated in Figure 1 by the case of the high s^* applicant.

determine the equilibrium in (8) is unknown. However the general properties of the functions have been characterized. As noted above both the benefit and cost functions are convex as the cost of raising scores is increasing in the amount of change in score and the fall in rates with increasing scores is decreasing. This relationship suggests that efforts to raise credit scores in response to the lender's discounts for higher scores are concentrated among those with the lowest scores. Two observations support this assumption. First, the cost of incrementally raising a low score is low. Second, the return in from lower risk perceptions by lenders (interest rate savings) is highest for those with low credit scores.¹³

The applicant's decision to expend effort to change her credit score in response to lender risk pricing is demonstrated in Figure 1. The graph illustrates behavior of applicants with either high or low initial credit scores. The marginal benefit schedule is the savings in credit cost from an increase in credit score, i.e. $-\theta\lambda I_r I R_S$, and the marginal cost schedule is the cost of achieving that increase in scores, I/S_c . Marginal benefit and cost have natural limits, at zero and infinity respectively because, as s becomes sufficiently large, both R_s and S_c approach zero. For those with low initial credit scores illustrated by the solid curves in Figure 1, the marginal benefit initially exceeds marginal cost. These applicants achieve an internal maximum of net benefit from efforts to raise their scores at the point s^L . The initial excess of marginal benefit over marginal cost at low levels of s is due to the steep slope of the I_r schedule as lenders aggressively price ration at the low end of the creditworthiness spectrum. Put another way, subprime lending rates fall sharply as credit quality increases. Those with initially high credit scores, shown by the dashed lines in the figure, reach a corner solution at $c = 0$ and $s^H = s^*_{High}$. They make no attempt

¹³ Empirical estimates in a subsequent section of this paper demonstrates that, at credit scores above 720, $I_r R_S$ is essentially 0.

to raise their credit scores artificially because there is little marginal benefit from doing so and marginal cost is very high, i.e. P_s is very small for those with high credit scores, and l/S_c is large because it is difficult to raise scores that are already high (and limited from above by a ceiling).

Insert Figure 1 approximately here

Thus the first proposition arising from the theory is that optimal c is a decreasing function of initial credit score, s^* . Indeed, for sufficiently high values of s^* shown as the high score case in Figure 1, $c = 0$ because both R_s and S_c approach zero .

Now, consider the effect of changes in economic conditions, z . As z rises, lender perceptions of loss given default fall because $L_z < 0$. This has no effect on the marginal cost schedule. However, it shifts the entire marginal benefit schedule down as $-\theta \lambda l R_s$ falls. This fall in the marginal benefit schedules of applicants induces many to choose $c = 0$ and others to lower their efforts to raise credit scores, lowering the optimal s . Thus, in periods when price rationing is less aggressive, the intercept of the marginal benefit schedule will fall and the return to raising credit scores will decline. Under these circumstances, even those with low credit scores have a small incentive to expend resources to raise their scores artificially.¹⁴

The theory has implications for the use of default models estimated in one period to predict losses in another period (this is, of course necessary because losses must be forecast using past experience). The theoretical result above, that the relation between underlying borrower creditworthiness and credit score varies with the aggressiveness of the price rationing,

¹⁴ The $S(c, s^*)$ function is assumed to be relatively stable, or at least not to be a function of factors, such as the rate of house price appreciation that are important components of z . This makes variation in $-\theta \lambda l R_s$ crucial for the behavior of applicants with a given initial credit score s^* . As lender's increase $\lambda l R_s$ in response to market conditions, i.e. falling z , the marginal benefit curve shifts up and eventually rises above the marginal cost curve. At this point the option to attempt credit score improvement is in the money.

implies that the losses associated with a given credit score observed by the lender will depend on the degree of aggressiveness with which risk was being priced by lenders during the period when the loan was originated. That is, the data generating process for defaults and losses depends on the size of $-\theta \lambda I_r l P_s$ and the validity of forecasts of error losses depends on whether the future values of $-\theta \lambda I_r l P_s$ are the same as those in the past along with the possibility that S_C varies over time as applicants, third party advisors and loan officers on commission become more adept at manipulating scores.

Stated econometrically, the standard competing hazard default and prepayment models used to forecast future losses suffers from simultaneous equations bias because the credit score variable is endogenous. Furthermore, similar arguments could be made for the endogeneity of loan terms, particularly the loan-to-value and payment to income ratios over which borrowers have some control. Lenders who believe that there is no relation between their credit pricing decisions and the predictive power of their loss models along with a combination of clever applicants and fraudulent loan officers could have contributed significantly to the “extra” losses above forecasts obtained using actual house price declines. What is even more important than looking at past failures, is that this problem has been ignored and, going forward, the same difficulties will arise whenever lenders become more aggressive in pricing credit risk in mortgage markets.¹⁵

IV. EMPIRICAL EVIDENCE ON RISK-BASED MORTGAGE PRICING

This section has two purposes. The first is to estimate the partial relation between mortgage interest rates and credit scores holding other elements of credit risk constant.

¹⁵ In the current environment, the mortgage market is characterized by massive non-price rationing and those with low credit scores are likely not even applying.

Essentially this involves measurement of $-\theta \lambda I_r / P_s$ using a massive sample of mortgages endorsed in Florida by 9 major lenders. The estimates relate APR to the credit score as well as other characteristics of the loan. Because, the data used in this estimation is from the files of actual lenders, the estimated relation should reflect the supply price of credit in the Florida market. There may be other factors that are related to risk that were ignored by these lenders, but the factors used in loan pricing were collected and retained by the lender and are available for use in the estimation.

The second objective of this section is to explore the possibility that the pattern of price rationing, i.e. the shape of the di/ds relation changed over time. In particular, for those at the lower end of the credit score spectrum, the effect of credit score on APR is expected to increase over the 2005 to 2008 period as lenders respond to perceptions that risk in mortgage lending had increased by attempting to price that risk in higher APRs. This is the basic operating rule of risk based mortgage pricing. Such an increase would generate a response by applicants who would then increase their efforts to raise credit scores.

The empirical estimates of the APR equation are achieved using ordinary least squares and a “laundry list” of variables available in the mortgage data. The effect of credit score on APR (I_r) is expected to be convex as shown above, APR falls at a decreasing rate with credit score. A stepwise linear function captures this relation using a series of score dummies at twenty point intervals. The estimating equation is given by:

$$APR_i = \alpha + \sum_j \beta_j f_{ji} + \sum_k \pi_k s_{ki} + \sum_l \rho_l z_{li} + \varepsilon_i \quad (5)$$

where f_{ji} is a vector of loan terms such as LTV, s_{ki} is a vector of credit score dummies, and z_{li} is a vector of other variables reflecting borrower characteristics. α , β , π , ρ are parameters to be

estimated and ε is an iid error term. Estimation of equations like (5) is common in the literature and the overall results are very consistent with expectations.

Table 1 contains summary results showing the relation between credit score and APR over each annual period from 2005 through 2008 as well as the entire sample period. Given the stepwise linear functional form, the effect of score on APR is simply the difference in the estimated coefficients across intervals. Table 3 confirms the expectation that $\lambda I_s R_s = di/ds$ is convex as expected with the function essentially flat as the credit score exceeds 730. Additionally, there are significant variations in the slope of the function over time. Standard errors for these parameters in the 2005 through 2007 regressions range from 2 to 5 basis points. Accordingly, it appears that, over this period, the relation between credit score and APR became more convex as the benefit from improved scores at the lower end was increasing.¹⁶ As with other research on mortgage pricing, it appears that, as the problems in credit quality approached, the lenders and secondary market participants expected higher yields to compensate for what they increasingly perceived as greater credit risk.

Insert Table 1 approximately here

The results in Table 1 can be used to compute the marginal benefit to applicants from raising their credit score based on their credit score at endorsement. This is, of course, not the same as the ideal measurement of the benefit to applicants based on their score before they contemplated application but that score is not available. This is one of several possible sources of measurement error in the marginal benefit from improved credit score. Accordingly, it is

¹⁶ Standard errors for 2008 are almost 15 basis points. This is sufficient precision to conclude that the relation between APR and credit score is convex but comparison between 2008 and other years is problematic.

likely that there is significant attenuation bias in the parameter estimate used to test for effects of marginal benefit to be discussed in the next section.

V. TESTING FOR STRATEGIC APPLICANT BEHAVIOR

The theory section developed a model which predicted that applicants with the greatest incentive to raise their credit scores based on the consequent fall in APR would be most likely to do so. The previous empirical section, provided a mechanism for estimating the expected fall in APR associated with a borrower whose credit score was in a particular range. Taken together these two sets of results imply that it is possible to identify the applicants who were most likely to have invested effort and funds to raise their credit scores before applying for a mortgage. Unfortunately, this behavior cannot be tested directly because credit scores for the period before application are not available. However, credit scores both at and after endorsement are available. To the extent that there is an incentive to raise scores artificially to lower credit cost, once this advantage is gained, households with special incentive to maintain higher scores should allow the scores fall back to their “normal” level. In terms of the theory and in contrast to expectations based on simple mean reversion, this suggests credit scores that were initially lower should fall further. Mean reversion should hold for the highest scores but be completely reversed for those homebuyers whose initial credit scores were low.¹⁷

For each mortgage, it is possible to compute the change in credit score over the first and second year after endorsement, Δs_1 and Δs_2 , and to relate this change to the initial credit score and terms of the mortgage, location of the property within Florida, and to changes in the housing

¹⁷ There is, of course, the third possibility that home purchase has an effect on credit scores. Based on government policies to encourage homeownership, this effect should be positive and hence reinforce the mean reversion argument.

market where the unit was located. In addition to all these factors which might influence changes in credit score, the theory of strategic applicant behavior predicts that the slope of the relation between APR and credit score that motivated the applicant to artificially raise her score at application, will then explain the change back to the applicant's "normal" score. A stronger financial incentive to artificially raise the score before applying, should produce a larger fall in score after endorsement. Of course, the ideal measure of the incentive to raise scores would be to observe the rate sheet of the lender serving each applicant. Instead, the marginal benefit computed in the previous section based on the estimate of $\Delta APR/\Delta s$ obtained using the estimation results from the Florida market for the year that the mortgage was endorsed.

The resulting equation for the change in credit score over time estimated by ordinary least squares for different years has the general form:

$$\Delta s_{ti} = \alpha + \sum_j \beta_j f_{ji} + \psi \Delta APR/\Delta s_i + \sum_k \rho_k M_{kti} + \varepsilon_i. \quad (6)$$

Here Δs_{ti} is the change in credit score over $t = 1, 2$ years after endorsement, f_{ji} is a vector of loan terms of associated with the mortgage at endorsement, $\Delta APR/\Delta s_i$ is the estimate of the slope of the marginal benefit curve facing the applicant at endorsement, M_{kti} is a vector of variables characterizing the market area in which the housing unit is located at time t , α , β , Ψ , and ρ are parameters or vectors of parameters to be estimated, and ε_i is an error term for applicant i .

The loan terms in the vector f_{ji} include, initial FICO score, appraised value, LTV ratio, interest rate, det to income ratio, and dummy for LTV = 0.80 (indicates likely presence of a

second mortgage).¹⁸ The current market characteristics in the M_{kti} vector include the percent Hispanic and black population in the census tract where the property is located, whether there had been any delinquent payments on the mortgage, whether the mortgage was current, whether the mortgage had terminated due to prepayment or default, the current interest rate, and an estimate of the current LTV which serves as an indicator of the current value of the put option.¹⁹ Descriptive statistics for these variables, disaggregated by year, are given in Appendix 1 along with a full set of estimation results for equation (6).

While the theory presented here has no implications for most of the parameters in equation (6), there is a prior expectation for ψ . The model of strategic applicant behavior and endogenous credit scores predicts that Δs_{it} should be negative and significant for those applicants for which $\Delta APR/\Delta s_i$ is also negative and numerically large. Accordingly, ψ should be positive and significant. The estimation results in Table 2 indicate that ψ is never negative and generally positive and significant as expected. Given the substantial measurement error associated with matching the loan pricing schedule estimates with individual borrowers and the consequent problem of attenuation bias, this result is rather remarkable in its robustness.²⁰

¹⁸ Previous research by Ashcraft and Schuermann (2006) and Gerardi, Shapiro, and Willen (2007) identified the mass point of mortgages at 80% LTV as a likely indicator of the presence of a second mortgage. There is a strong financial incentive to choose an 80% LTV to avoid mortgage insurance on the prime mortgage.

¹⁹ The value of the current LTV is based on the computations described in detail in Smith (2011). This includes creation of a repeat sales price index for Florida counties to measure current value as described in Archer and Smith (2012) and estimated mortgage amortization. See Foote, Gerardi, and Willen (2008) for a discussion of the importance of the put option for borrower default behavior and hence for determining the current credit score.

²⁰ The magnitude of the relation between the estimated slope of the APR versus credit score relation and the subsequent fall in credit score is not large. At the mean of the sample $\Delta APR/\Delta s = 0.33$ and, given the average value of estimated $\Psi = 5.0$, the average fall in subsequent credit score due to the slope of the APR relation at the mean of the sample is only 1.5. Given that the massive potential for measurement error in $\Delta APR/\Delta s$ for each applicant, the potential for attenuation bias to reduce the size and significance of estimated Ψ is so consequential that little importance should be attached to the estimated size of the effect of the $\Delta APR/\Delta s$ on the subsequent fall in scores.

Insert Table 2 approximately here

Based on these results, it appears that borrowers in Florida during this test period were behaving strategically. Those with the most to gain from artificially raising their credit scores, generally those with the lowest initial scores, had their credit scores fall by the largest amount over the one and two-year periods after endorsement.

This result contrasts with the normal expectation based on stochastic process and public policy toward homeownership. Mean reversion implies that the lower scores should be the most likely to rise and the highest most likely to fall. Public policy promotes homeownership among those households who are least creditworthy on the presumption that it has beneficial effects on their welfare. Presumably plunging credit scores are not regarded as beneficial. Therefore, both the implications of stochastic processes and public policy would predict that homeownership has a positive effect on credit scores of marginal homebuyers. Clearly, this view is not supported in the data. However, the fall in credit scores found here does not contradict either stochastic or public policy expectations. Rather it reflects yet another process, strategic behavior by mortgage applicants, that has not previously been considered and which is sufficiently large to overwhelm the processes that would normally raise credit scores of applicants whose initial scores were below the mean.

V. SUMMARY AND CONCLUSIONS

Discussions of the high rates of mortgage defaults beginning in 2006 have emphasized the cycle in housing prices, asymmetric information and incentives in mortgage securitization, and alternative mortgage products incorporating payment shocks. Most of the literature treats borrower's as gullible, clueless, or passive. This view contrasts with much of the theoretical

literature on credit markets that preceded the mortgage market losses which treated borrowers as rational, informed agents making decisions in their self interest. This paper has followed the second approach.

Mortgage lenders engage in price rationing based on econometric model default and prepayment models that estimate the relation between loan terms and credit scores and expected future default and foreclosure loss. The major conclusion of this paper is that confidence in this approach and the results of competing hazard models of default and prepayment as a basis for assessing mortgage risk has been misplaced because, in addition to loan terms, credit score itself may be endogenous.

Furthermore there is a dynamic relation between model performance and housing market conditions that arises for two reasons. First, when house prices are rising and loss conditional on default is low, default and prepayment models perform well. Second this reinforces lenders' confidence in the models and induces them to rely on pricing rationing more and expand the range of credit scores that they will consider. This increased use of price rationing increases the incentive for applicants to artificially raise credit scores rises. Third, as lenders anticipate house prices may be at or near a peak, they raise estimates of loss given default and price this risk in interest rates. Fourth, these events which induce lenders to price credit risk more aggressively increase the marginal benefit of artificially raising credit scores and provokes borrower reactions that invalidate the assumption that loan terms and credit scores are econometrically exogenous to the probability of default. Thus, it is precisely on those rare occasions, when the applicants perceive the ex ante probability of default as being particularly high, that they take actions that invalidate the crucial assumptions of competing hazard models of default and prepayment.

The extreme case of this process is mortgage fraud in which the applicant has little interest in anything other than obtaining credit at the highest LTV possible because there is no intention to repay the loan. But, for those who do play to repay, there is still an incentive to behave strategically and artificially raise credit score if it can be done cost effectively. This provides an explanation for the documented failure of competing hazard models of prepayment and default to provide reliable predictions of default rates in the most recent financial crisis. The problem of relying on estimated default risk models that assume loan terms and credit score are exogenous to underwrite and price mortgages has been ignored in both the literature and proposals for credit market change. The empirical evidence provided here suggests that it is real and should not be neglected because the conditions for strategic behavior by mortgage applicants will surely come again.

REFERENCES

Anderson, Nathan B., and Jane K. Dokko, 2011. Liquidity problems and early payment default among subprime mortgages, Discussion Series: 2011-09, Board of Governors of the Federal Reserve System.

Ashcraft, A.B., and T. Schuermann. 2006. Understanding the Securitization of Subprime Mortgage Credit, *Foundations and Trends® in Finance*, vol 2, no 3, pp 191-309.

Avery, Robert B., Kenneth P. Brevoort, and Glenn B. Canner. (2010). "Credit where none is due? Authorized user account status and "piggybacking credit." *Finance and Economics Discussion Series*. Federal Reserve Board, Washington, D.C., Working Paper 2010-23.

An, Xudong, Youngheng Deng, Eric Rosenblatt, and Vincent W. Yan, 2010. Model Stability and the Subprime Mortgage Crisis, Working Paper.

Avery, R. B., Bostic, R. W., Calem, P. S., & Canner, G. B. (2000). Credit Scoring: Statistical Issues and Evidence from Credit-Bureau Files. *Real Estate Economics*, 28(3), 523-547.

Barth Cordes Yezer 1981

Carrillo 2011

Cordes, Joseph J. James R. Barth, and Anthony M. Yezer, 1980. Financial institution regulations, redlining, and mortgage markets, in *The Regulation of Financial Institutions*, Conference Volume 21, Federal Reserve Bank of Boston, (April 1980), pp. 101-143.

Deng, Yongheng, John M. Quigley, and Robert Van Order, 2000. Mortgage terminations, heterogeneity, and the exercise of mortgage options. *Econometrica*, 68 (2), pp. 275-307.

Einav, Liran, Mark Jenkins, and Jonathan Levin, 2013. The impact of credit scoring on consumer lending. *RAND Journal of Economics*, 44(2), pp. 249-274.

Federal Bureau of Investigation, 2011. 2010 Mortgage Fraud Report, U.S. F.B.I., Washington, D.C.

Fischelson and Holstine 2005

Foote, Christopher L., Kristopher Gerardi, and Paul S. Willen. 2008. Negative equity and foreclosure: Theory and evidence. *Journal of Urban Economics* 64.2: 234-245.

Garmaise, Mark J., 2012. Borrower misrepresentation and loan performance, Working Paper UCLA.

Garmaise, Mark J., 2013. The attractions and perils of flexible mortgage lending. *Review of Financial Studies*, 26:10, pp. 2548-82.

Gerardi, K., Shapiro, A., & Willen, P. (2007). Subprime outcomes: Risky mortgages, homeownership experiences, and foreclosures, working paper.

Haughwout, Andrew, Richard Peach, and Joseph Tracy, 2008. Juvenile delinquent mortgages: Bad credit or bad economy?, *Journal of Urban Economics*, Vol. 64, pp. 246-257.

Kau, James B., Donald C. Keenan, Constantine Lyubimov, and V. Carlos Slawson, 2011. Subprime mortgage default, *Journal of Urban Economics*, Vol. 70. Pp. 75-87.

Lusardi, Annamaria, and Olivia S. Mitchell, 2014. The economic importance of financial literacy: theory and evidence. *Journal of Economic Literature*. Vol. 52 (1), pp. 5-44.

- Maddala, G.S., and Robert Trost, 1982. On measuring discrimination in loan markets, *Housing Finance Review*, 1, 245-268.
- Mayer, Christopher, Karen Pence, and Shane M. Sherlund, 2009. The rise in mortgage defaults, *Journal of Economic Perspectives*, Vol 23 (1), pp. 27-50.
- Norden, Lars, and Martin Weber. 2010. Credit line usage, checking account activity, and default risk of bank borrowers. *Review of Financial Studies*. 23: , pp. 3665-99
- Rajan, Uday, Amit Seru, and Vikrant Vig, 2010. The failure of models that predict failure” distance, incentives, and defaults,” Working Paper,
- Rajan, Uday, Amit Seru, and Vikrant Vig, 2010. Statistical default models and incentives, *American Economic Review*, Vol 100, pp. 506-510.
- Rothschild, Stiglitz 1976
- Smith, Brent C, and Wayne R. Archer. 2013. Residential Mortgage Default: The Role of House Price Volatility, Euphoria and the Borrower’s Put Option, *Journal of Real Estate Finance and Economics*.
- Smith, B. C. 2011. Stability in Consumer Credit Scores: Level and Direction of FICO Score Drift as a Precursor to Mortgage Default and Prepayment, *Journal of Housing Economics*, 20:4, 285-298 December.
- Stiglitz and Weiss 1981
- Straka, John W. (2000). A shift in the mortgage landscape: The 1990s move to automated credit evaluations. *Journal of Housing Research*, 11(2), 207-232.
- Wang and Williamson 1998
- Willen Pence Shuerlund 2009

Appendix A

The total cost of credit, Ω , is the sum of the discounted present value of payments of interest and fees and expenditures to raise the credit score:

$$\Omega = i\theta + c \quad (A1)$$

where: θ is a scaling parameter that is the ratio between the discounted present value of the expected loan payments and the interest rate, i is an interest rate reflecting the “all in cost of credit” including points and fees, and c is the cost of raising the credit score.

The consumer’s problem is to choose a level of c to maximize (A1), implying the necessary FOC condition that:

$$d\Omega/dc = (\partial i/\partial s)(\partial s/\partial c)\theta + 1 = 0 \quad (A2)$$

or:

$$-(\partial i/\partial s)\theta = 1/(\partial s/\partial c) \quad (A3)$$

The relation in (A3) has a convenient interpretation. The marginal benefit from a higher credit score is $-(\partial i/\partial s)\theta$ and the marginal cost of achieving a rise in credit score is $1/(\partial s/\partial c)$. This forms the basis for Figure 1. As s^* rises, both $(\partial i/\partial s)$ and $(\partial s/\partial c)$ approach 0 raising the possibility shown in Figure 1 as a corner solution at $c = 0$ where $-(\partial i/\partial s)\theta < 1/(\partial s/\partial c)$.

Figure 1

Applicant's Choice of Credit Score Improvement

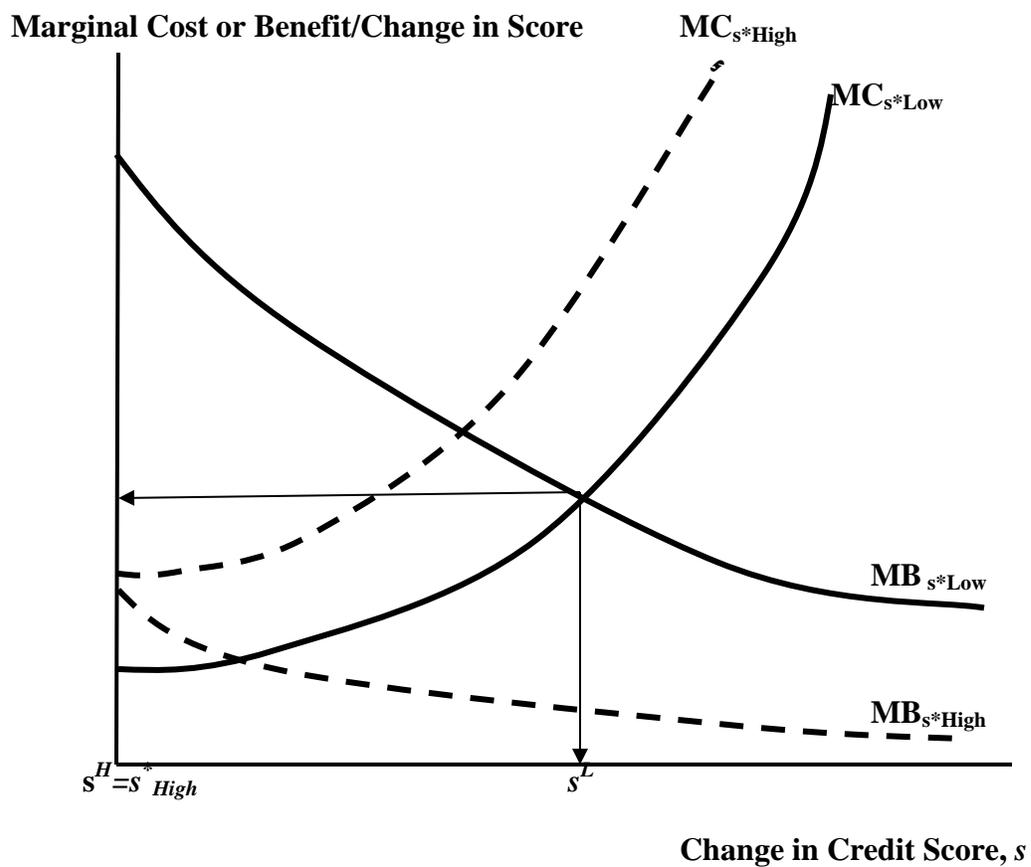


Table 1
Estimates of Fall in APR as Credit Score Rises For
Mortgages in Florida For 2005-2008

Score Interval	All Years	2005	2006	2007	2008
560-580	-0.41	-0.29	-0.38	-0.40	-1.24
600-620	-0.27	-0.14	-0.11	-0.29	-0.22
640-660	-0.14	-0.12	-0.16	-0.14	-0.09
700-720	-0.04	-0.00	-0.32	-0.35	-0.10
720-740	-0.03	-0.01	-0.03	-0.02	-0.03

Table 2

Estimated Values of the ψ Parameter

Relation Between Estimate of $\Delta\text{APR}/\Delta s$ and Subsequent $\Delta s/\Delta t$

Mortgage Type and Time	Pooled	2005	2006	2007	2008
Home Purchase, Year 1	3.55***	ns	ns	7.95**	4.28**
Home Purchase, Year 2	ns	3.97*	ns	ns	NA
Refinance, Year 1	1.16*	ns	ns	3.64**	ns
Refinance, Year 2	4.83**	ns	ns	14.38***	NA

NA – it was not possible to observe the 2 year change in credit scores for mortgages endorsed in 2008.

Table 1
Full Dataset

n=	45,370	
Variable (mean/sd)	Mean	Std Dev.
fico change	(24.357)	58.850
ΔAPR/Δs	0.332	0.565
fico_orig	708.637	60.568
appraisal	308,370.000	285,169.900
ltv ratio	81.482	11.986
hisp pop %	0.200	0.219
black pop %	0.132	0.165
terminate prepay	0.081	0.273
foreclosure	0.041	0.198
condo	0.334	0.472
seconds	0.127	0.333
unemp %	5.207	1.599
prior delinquency	0.176	0.381
current	0.888	0.315
prime fico	0.916	0.277
current rate	6.400	0.760
as of ltv	83.341	24.295

Table 2 Dataset by Year

Year	2005	2006	2007	2008
n=	11,532	12,298	15,342	6,198
Variable (mean/sd)				
fico change	(15.106)	(20.262)	(29.136)	(37.867)
	<i>46.714</i>	<i>54.148</i>	<i>64.204</i>	<i>69.576</i>
allnodes	0.199	0.321	0.395	0.443
	<i>0.374</i>	<i>0.480</i>	<i>0.543</i>	<i>0.912</i>
fico_orig	715.842	706.787	703.623	711.317
	<i>54.372</i>	<i>59.303</i>	<i>64.901</i>	<i>61.464</i>
appraisal	325,259.800	324,425.800	292,609.100	284,100.600
	<i>320,287.800</i>	<i>299,544.800</i>	<i>252,099.600</i>	<i>258,567.100</i>
ltv ratio	80.920	80.606	82.656	81.359
	<i>10.639</i>	<i>11.583</i>	<i>12.711</i>	<i>13.015</i>
hisp pop %	0.197	0.191	0.182	0.270
	<i>0.213</i>	<i>0.213</i>	<i>0.202</i>	<i>0.262</i>
black pop %	0.118	0.133	0.139	0.138
	<i>0.149</i>	<i>0.166</i>	<i>0.175</i>	<i>0.163</i>
terminate prepay	0.211	0.063	0.020	0.025
	<i>0.408</i>	<i>0.243</i>	<i>0.141</i>	<i>0.157</i>
foreclosure	0.003	0.027	0.065	0.078
	<i>0.057</i>	<i>0.161</i>	<i>0.247</i>	<i>0.268</i>
condo	0.336	0.316	0.343	0.343
	<i>0.472</i>	<i>0.465</i>	<i>0.475</i>	<i>0.475</i>
seconds	0.216	0.148	0.065	0.069
	<i>0.412</i>	<i>0.355</i>	<i>0.247</i>	<i>0.253</i>

unemp %	3.511	4.363	6.316	7.297
	<i>0.455</i>	<i>0.830</i>	<i>0.867</i>	<i>0.964</i>
prior delinquency	0.104	0.161	0.217	0.238
	<i>0.305</i>	<i>0.368</i>	<i>0.412</i>	<i>0.426</i>
current	0.972	0.911	0.840	0.806
	<i>0.166</i>	<i>0.285</i>	<i>0.367</i>	<i>0.396</i>
prime fico	0.950	0.918	0.890	0.914
	<i>0.218</i>	<i>0.274</i>	<i>0.313</i>	<i>0.280</i>
current rate	5.858	6.613	6.664	6.334
	<i>0.651</i>	<i>0.693</i>	<i>0.730</i>	<i>0.592</i>
as of ltv	62.488	85.559	93.716	92.059
	<i>15.532</i>	<i>22.152</i>	<i>22.951</i>	<i>20.809</i>

