Collateral Pledge, Sunk-Cost Fallacy and Mortgage Default*

Sumit Agarwal
National University of Singapore, ushakri@yahoo.com

Richard K. Green
University of Southern California, richarkg@usc.edu

Eric Rosenblatt
Fannie Mae, eric_rosenblatt@fanniemae.com

Vincent Yao
Fannie Mae, Vincent_w_yao@fanniemae.com

August 2014

Abstract.

Individuals and firms pledge collateral to mitigate agency costs or contracting frictions in a world with asymmetric information. However, the option value theory suggests that once the mark-to-market asset valuation is below the current debt, the firms and individuals should default on their debt contract irrespective of the initial collateral pledged. In this paper, we estimate default models and find that after controlling for mark-to-market asset valuation, initial collateral remains an important predictor of mortgage default. Specifically, individuals that pledge higher collateral have a lower hazard to default. Our results are consistent with models of sunk cost fallacy.

Keywords: Household Finance, Mortgages, Collateral, Option Value, Financial Crisis, Sunk Cost, Credit Card Default

JEL classification: D12, D14, G11, G21

*We would like to thank Gene Amromin, Zahi Ben-David, Souphala Chomsisengphet, John Driscoll, Amit Seru and Wenlan Qian for helpful comments and suggestions. The views expressed in this research are those of the authors and do not necessarily represent the policies or positions of Fannie Mae.

Electronic copy available at: http://ssrn.com/abstract=2446748

1. Introduction

Economic theory explains the role of collateral in debt contracts as an attempt to mitigate agency costs or contracting frictions in a world with asymmetric information. Chan and Thakor (1987) argue that collateral alleviates adverse selection problem while Holmstrom and Tirole (1997) argue that collateral alleviates the moral hazard problem. Inderst and Mueller (2007) show that collateral allows local banks that use soft information to be more competitive against banks that use hard information in their lending. Benmelech, Garmaise, and Moskowitz (2005) find that liquidation values of collateralized assets are first-order determinants of loan contract terms.

Berger et. al. (2005) suggests that the use of hard information like the credit scoring technology reduces the pledged collateral in debt contracts. They argue that pledging collateral imposes a cost on the lender and the borrower. Specifically, the lenders incur the screening costs of valuing the pledged assets; the costs of monitoring the secured assets; and any enforcement/disposal expenses in the event of repossession (e.g., Leeth and Scott 1989). The use of collateral may impose opportunity costs on borrowers to the extent that it ties up assets that might otherwise be put to more productive uses. Borrowers may also suffer fluctuations in their credit availability as the values of their securable assets vary.

In this paper we use down-payment size as a measure of collateral pledged by the home buyer and test if the collateral pledged explains the mortgage default hazard, conditioned on the current loan-to-value ratios that vary across time because of changes in housing markets and paying down of the interest and principal on the house. Our thought experiment is as follows: Imagine two households looking to buy houses in two different locations. The first household puts down a 20% down payment and the second household puts down a 5% down payment. The first household by putting down more collateral is trying to signal to the lender that he is low risk. Over the next few months the first homeowner experiences a house price decline of 21% and the second homeowner experiences a house price decline of 6%. Essentially the option to default for both of them is in the money and they both should be equally likely to default. Since at that point they both have a -1% equity in the house. However, if there is some behavioral bias (like sunk cost fallacy) then it is possible that the first household is less likely to default. In other words, if households with higher initial down-payments are less prone to default, it could mean that down-payment is a proxy for commitment, or that those who put high down-payments on their houses are more likely to be motivated by nominal loss aversion.We also use measures of local house price volatility as an explanatory variable for default hazard. In the presence of foreclosure-induced transactions costs, the put option will be greater in areas with higher house price volatility, and therefore we would expect higher house price volatility to predict lower defaults.

Our results show that default risks increase with cumulative loan-to-value at origination (CLTV) conditional on the mark-to-market CLTV (MCLTV). In other words, households who have more equity in their house despite the fact that their MCLTV is negative they are less likely to default on their mortgage. Moreover, comparison of change in default probabilities along CLTV and MCLTV suggests that incremental risks owing to initial down payment and to our mark-to-market measure are similar in scale. We conduct a number of robustness tests to check if our results could be driven by other factors like the heterogeneity in house price, recourse versus non-recourse states, and macroeconomic conditions. Additionally, we replicate our analysis by matching homeowners on observable characteristics except for CLTV. Consistent with our hypothesis, we find that collateral pledged at origination does create a sunk cost fallacy in mortgage default behavior.

An alternative explanation of our results is that origination CLTV proxies for borrower characteristics that affect mortgage default directly (e.g. the borrowers use the origination CLTV to signal their type to the lender). Under this alternative interpretation, the coefficient simply captures the fact that borrowers with lower origination LTV are less likely to default because of unobserved borrower characteristics. It is difficult to completely rule out this interpretation unless we conduct a field experiment with a control and treatment groups. However, we conduct the following thought experiment to address the above concern. We look at two different borrowers with an origination CLTV of 80% and an origination CLTV of 90% on their mortgages and study their credit card usage behavior. If it is the case that they have differently default characteristics as they use the origination CLTV to signal their type, we should find the credit card default behaviour of these two borrowers to be different. However, our results show that credit card default behaviour of these two individuals is the same. This suggests that the origination CLTV may not reflect some unobserved borrower characteristics that captures their type.

Our results contribute to several strands of the literature. First, we provide new evidence consistent with prospect theory. Empirical findings consistent with prospect theory include Gensove and Mayer (2001) and List (2003, 2004) for the housing market and Benartzi and Thaler (1995), Barberis, Huang, and Santos (2001) for the stock market. Second, we also contribute to explaining consumer default behaviour on their mortgages that is above and beyond the standard explanations in the mortgage literature like Deng, Quigley and Van Order (2001), and Agarwal, Chang, and Yavas (2012). Finally, we also contribute to the literature on strategic default and show that households that have behavioral biases are less likely to default strategically. This is not necessarily inconsistent with Guiso, Sapienza and Zingales (2011), who find that strategic defaulters are rational agents.

For the remainder of this paper, section 2will review the literature, section 3 discussesthe theory of mortgage default, section 4 present our data, describes our empirical model, and present results. Finally, section 5 concludes.

2. Literature

We discuss briefly three strands of literature: (i) the literature on mortgage default, (ii) the literature on prospect theory, and how it implies that initial loan-to-value might influence mortgage default, and (iii) the recent literature on strategic default.

2.1 Default

The literature on the determinants of mortgage default (as well as corporate development of default models) has been evolving for 30 years. Foster and Van Order (1984 and 1985) were among the first to model default as a put option. When a homeowner has a mortgage, and can extinguish all of her obligation by putting a house back to the lender, (s)he has a put option as well as equity in the house. The put option's value is a function of the drift and volatility of house prices: it is more valuable when house prices are more likely to fall, and more valuable when house prices are more volatile, because under both circumstances, the chance that the house will be worth less than the mortgage, and therefore in the money, is greater. The Foster and Van Order papers used house price volatility as a covariate for predicting default, and found that it was a significant and important predictor of default.

Kau and Keenan (1993 and 1999) show that even in the absence of transactions costs, borrowers will not necessarily default immediately when the option is in the money. They derive the well-known result that under risk neutrality, the value of a mortgage M satisfies:

$$\frac{1}{2}V^2\sigma_h^2\frac{\partial^2 M}{\partial V} + \rho V\sigma_h\sigma_r\frac{\partial^2 M}{\partial h\partial r} + \frac{1}{2}\sigma_r^2\frac{\partial^2 M}{\partial r^2} + \gamma(r)\frac{\partial M}{\partial r} + (r-d)V\frac{\partial M}{\partial V} + \frac{\partial M}{\partial t} + C = rM \quad (1)$$

Where M, the mortgage, is valued based on its coupon, C, the spot interest rate r, the value of the house, V, the age of the mortgage, t, and the time to maturity, T. Let *d* be the imputed rental rate of the house. In the absence of frictions (i.e., costs of default beyond loss of the house), the optimal default strategy at time t is a function of the value of the house, V_t. Optimal default takes place when $V_t=V_t^*$. Borrowers choose V_t^* such that the value of the mortgage is minimized, subject to the condition that the value of the mortgage balance is equal to the value of the house. The value of the mortgage balance is less than the present value of the remaining payments, because the mortgage contains a prepayment option as well as a default option.

Quigley and Van Order (1995) illustrate the optimal default strategy when interest rates are constant. The curve OX in their Figure 1 is the highest curve that satisfies the relationship between house prices and mortgage values represented in equation 1 while assuring that the mortgage value is less than the value of a riskless mortgage. This creates a tangency point which determines V_t^*. Note that value of the house is less than (not equal to) the mortgage balance at the optimal point of default.

But the model has another important implication—that once house values pass below some threshold point, other issues, such as original loan-to-value ratio, should not affect mortgage behavior.

In a later paper, Deng, Quigley and Van Order (2001) make two further contributions: they show that default models are best modeled in a competing risk framework, where default competes with prepayment. We will discuss the econometric specification of this model below. They also hypothesize that LTV is endogenous—that those who have asymmetric information that a particular house might lose value will take on higher LTVs than those who do not.

2.1 Prospect Theory

In a recent review article, DellaVigna (2009) describes Kahneman and Tversky's prospect theory: Kahneman and Tversky (1979), in the second most cited article in economicssince 1970 (Kim, Morse, and Zingales, 2006), propose a reference-dependent model of preferences that, unlike the standard model, can fit most of the experimental evidence on lottery choice.According to prospect theory, subjects evaluate a lottery $(y, p; z, 1 - p)$ as follows: $\pi(p)\, v(y - r) + \pi(1 - p)\, v(z - r)$. Prospect theory is characterized by: (i) Reference Dependence. The value function v is defined over differences from a reference point r, instead of over the overall wealth; (ii) Loss Aversion. The value function $v(x)$ has a kink at the reference point and is steeper for losses ($x < 0$) than for gains ($x > 0$); (iii) Diminishing Sensitivity. The value function v is concave over gains and convex over losses, reflecting diminishing sensitivity to outcomes further from the reference point; (iv) Probability weighting. The decision-maker transforms the probabilities with a probability-weighting function $\pi(p)$ that overweights small probabilities and underweights large probabilities.

Empirical findings consistent with prospect theory include Gensove and Mayer (2001) and List (2003, 2004) for the housing market and Benartzi and Thaler (1995), Barberis, Huang, and Santos (2001) for the stock market. Odean (1998) found a "disposition effect," whereby sellers are more likely to sell their winners than their losers, even when tax considerations should tilt sellers to liquidating losers.

This paper continues in that tradition. While mortgage default theory does not precisely predict how far underwater borrowers need be before they default, it does suggest that past some point, down payment at origination should be irrelevant to the determination of default. A down payment may be viewed as a reference; those who put a large down payment on a house may view themselves as having “more to lose” when they default than those who do not do so. Consequently, if borrowers have nominal loss aversion, those who finance their houses with large down payment may be less likely to default, conditional on mark-to-market loan to value ratio, than those who use small down payments.

2.3 Strategic Default

Some recent papers have examined strategic default and the relationship between initial down payment and default. Kelly (2008) shows that recent work on source of down-payment and default, that households who receive grants or gifts for down-payments tend to default at higher rates than those who bring their own money to a home buying transaction. Kelly’s paper provides important insights, but relies on a data set that samples a much narrower population than ours, and that lacks the rich set of explanatory variables available to us.

Guiso, Sapienza and Zingales (2011) use survey data to study American households' propensity to default when the value of their mortgage exceeds the value of their house, even if they can afford to pay their mortgage (strategic default). They find that people who find it “immoral” to default are 77 percent less likely to do so than those who do not. While the focus of the study is different from ours, and contains the limitation of using survey data to gauge attitudes, it does examine the impact of non-financial considerations on mortgage default, and finds that such considerations matter when households make default decisions.

Oliver Wyman (2009) sorted through credit reports and argued that the United States had more than one half million strategic defaults (i.e., defaults by borrowers who were capable of repaying their loans but found themselves with negative home equity) in 2008. This paper argues that at least some borrowers are ruthless about default. Nevertheless, by mid-2009, at least 24% of borrowers "owned" houses whose value was less than their loan balances, which suggests that many, many households are not ruthless about default.

3. Theory of Mortgage Default

Recent events have allowed us to focus on how we might better understand mortgage default. In a world where transactions costs are minimal, borrowers could be "ruthless" about default: when the value of their house falls below the value of their mortgage (inclusive of the value of prepayment options), they could default on their loan and repurchase their house at the lower price. Kau and Keenan (1993 and 1999) and Foster and Van Order (1994 and 1985) were among the first to treat mortgage default as a put option with the mortgage balance being the strike price: the borrower puts the house back to the lender once the option is in the money.

But mortgage default *does* contain transaction costs. Some of them are straightforward: it costs money to move, and the point at which the house price falls below the mortgage balance is not the point at which the option is net in the money. But if moving were the only transaction cost, we would expect to see a threshold at which everyone defaulted. The fact is that we do not see such a threshold: it is a long established stylized fact that the hazard of the default is increasing in contemporaneous loan-to-value, but the relationship is smooth.

Part of the reason may be that default produces consequences for the ability to obtain credit or (possibly) to rent an apartment. Therefore households have to take into account the marginal benefit from default (i.e., the amount of debt relieved) relative to the total marginal cost. One could imagine that such a calculation is highly heterogeneous. Indeed, Deng, Quigley and van Order (2001) showed that unobserved heterogeneity is a pervasive feature of borrower behavior. It is also possible that housing market heterogeneity influences default decisions. When a household is underwater with respect to its mortgage, it holds a call option: the ability to "buy-back" the house when its value is greater than the mortgage balance. Because volatility creates option value (about which we shall say more below), the call option embedded in an underwater mortgage is higher in high volatility places than it is in low volatility places.

Beyond all of this, it is quite striking how many people decide not to default even when the cost of default is likely less than the benefits, and when the call option is worth nearly zero. Specifically, we find that loans can have loan-to-value (LTV) ratios in excess of 240 percent, and yet continue to perform. Our suspicion is that people who continue to make payments on these loans either (1) do not fully understand what is happening in their housing markets; (2) feel a strong sense of responsibility toward meeting a "moral" obligation; and/or (3) feel that they would lose the cash equity they put into their house if they chose to default. Those falling into the last category suffer from "nominal loss aversion," a characteristic discussed in Kahneman and Tversky's (1979) work on prospect theory. Genesove and Mayer (2001) showed how nominal loss aversion affects seller behavior in soft real estate markets.

4. Data, Methodology, and Results

4.1. Data

The data used in this paper comes from three sources. First, we begin with a large proprietary mortgage database, containing approximately a quarter of all national mortgage transactions over the period from 2005 through 2008. The originations before 2005 are not included because few of these homeowners have negative equity, owing to both the fact that houses purchased at that time were priced below peak and amortization.

From this mortgage data, we can identify all the original loan information and subsequent payment activity in each quarter until the loan is voluntarily prepaid, foreclosed, or still active as of 2011Q4. We randomly selected a sample of mortgages that financed owner-occupied home purchases from 2005 through 2008. The sample is further limited to single-family 1-unit houses and fixed rate mortgages with 30-year amortization terms. We also excluded loans with less than six months of payment history in total since there is not enough home price change or amortization to separate current loan to value ratio from original loan to value ratio. There are 680,223 unique loan records that contain the following variables: credit score (FICO), combined loan to value ratio (CLTV), age and monthly income of borrower, origination date, note rate, whether the property is in a recourse state and whether the loan is originated by a broker, correspondent, or retail channel. To capture potential non-linear relationship with FICO, we used four splines at 740, 700, 660 and 620. We also know the current unpaid principal balance in each quarter and the current status of the mortgage (prepaid or default) along with the exact liquidation date. The cumulativedefault (foreclosure) and prepayment rates over this set of loans are 10.4% and 65.2% respectively. With proprietary zip-code-level home price indices, we are able to construct accurate mark-to-market CLTV (MCLTV) quarter by quarter. There are 6,108,315 records in the dataset and the average duration of loans is 46.6months or about 4 years.

We also obtained MSA level unemployment rates for each quarter and each geographical area from the Bureau of Labor Statistics; we follow mortgage rates using the FreddieMac Primary Mortgage Market Survey. For each observation, we compute the difference in the market interest rate and the note rate at origination: this gives us a measure of the value of the call (refinance) option.

The variables used in the analysis are explained in Table 1 and their descriptive statistics are reported in Table 2. All loans in the dataset are for prime conventional conforming borrowers, so their average FICO score is high at 718; their average monthly income is also high at $6,864. The average age of the borrowers in our sample is 39.5 years. In this period, the sample is dominated by wholesale channels, with 56.8% of loans originated by brokers and correspondents. Because house prices declined from the peak in 2006 and began to recover in 2009, the average MCLTV in last activity monthis86.0 percent, compared to 86.2 percent CLTV at origination due to home price change and amortization. Without controlling for economic conditions, default peaks at 6-9 quarters after origination and prepayment peaks about 2 quarters earlier. The fast prepayment speeds reflect an environment in which interest rates are falling.

Raw statistics on CLTV and MCLTV, reported in Table 3, provide a snapshot of both original loan profiles and loan activitiesby payment status and originationyears. Loans that ultimate went to foreclosureare typically low FICO, high LTV and high mark to market LTV. They have also in general experienced deterioration of unemployment rate in the local economy. Loans that were prepaid either by refinance or move are characterized of high FICO and low LTV. They have in general experienced home price appreciation and improvement of labor market. Their mark to market LTV is lower than their original LTV. Across four origination years, loans in 2006 and 2007 have the highest default and SDQ rates. They have suffered the steepest home price decline and the resulting mark to market LTVs are the highest. 2008 saw a significant improvement of underwriting standards with the highest average FICO and lowest LTV among all four years. These loans have relatively low default rate and high prepayment rate compared to other years.

Table 4 (a) reports basic borrower characteristics by CLTV categories – most interesting variable of this paper. As anticipated, borrowers putting lower down-payment at origination usually have lower FICO, make less income and are much younger. Their mortgages are more likely to be originated by broker, rather than retail banks. Debt to income ratio, however, is not directly related to down-payment. These relationships suggest that unobserved borrower characteristics could vary in a similar way, which would potentially violate the identification strategy. Both extremely low and high down-payment borrowers have higher debt ratio. Raw statistics on CLTV and MCLTV, reported in Table 2 (b), provide a snapshot of both original loan profiles and the impact of home price changes over the 4 study years. In the loan activity table, 30% of activity records have original CLTV>95%.[1] About 22% of these borrowers have MCLTVs that remained flat ((95,100]). The MCLTVs of the other 48% have well exceeded 100%; i.e., almost half of these homeowners who began with an LTV of 95 percent or higher have negative equity. Although borrowers with lower original LTV are less likely to have negative equity, none of CLTV categories are immune to negative equity. Overall, about 28% of loans in our sample have negative equity.

Figure 1 plots the default and prepayment rates by initial down-payment and current LTV. For each similar MCLTV, default rate increases in higher CLTV; for each similar CLTV, default rate increases in higher MCLTV. Because more loans in higher CLTV categories ended

[1] This overestimates the share of high CLTV loans (31%) in the loan-level data because that these loans tend to have longer durations.

up in higher MCLTV, overall performance of higher CLTV is much worse. Default rate of 95<CLTV<=100 is more than twice of those in 90<CLTV<=95, almost four times of those in 80<CLTV<=90 and 9 times of those in CLTV<=80.Prepayment relationships are less straightforward. For each similar CLTV, prepayment rate decreases in higher MCLTV. However, for each similar MCLTV, prepayment rate is indifferent to change in CLTV except for CLTV >95or MCLTV >100. This implies that the ability to refinance is largely influenced by owners' mark-to-market equity. But since 2009 US governments have instituted several prominent housing programs (e.g., Home Affordable Refinance Program or HARP) to help borrowers with MCLTV up to 125 to refinance.

4.2.Regression Models

Default and prepayment are options available to homeowners at each period. They decide whether to exercise either of the options, provided that they have perfect information on the current value of mortgages and economic conditions. Due to the nature of competing risks between prepayment and default, only the duration associated with the type that terminates first is observed. Following Van den Berg (2001), the competing risk model or multivariate mixed proportional hazard (MMPH) defines two hazard equations as

$$\begin{aligned} \theta_1(\mathrm{t} \mid \mathrm{x}, \mathrm{V}) &= \lambda_1(\mathrm{t}) \times \exp(\mathrm{x}'\beta_1) \times \mathrm{V}_1 \\ \theta_2(\mathrm{t} \mid \mathrm{x}, \mathrm{V}) &= \lambda_2(\mathrm{t}) \times \exp(\mathrm{x}'\beta_2) \times \mathrm{V}_2 \end{aligned} \quad (2)$$

where durations (T1, T2)|(x, V) are assumed independent, so that a dependence of the durations given x is modeled by way of their unobserved determinants V1 and V2 being dependent. The $\theta(t)_i$s are the hazards for prepayment and default conditional on survival up to time t and the $\lambda(t)_i$sare the baseline hazards. We include three time-varying factors in x:

MCLTV, change in mortgage rate and change in MSA-level unemployment rate since mortgage origination. Change in mortgage rate is used to measure the value of the call option for homeowners, while MCLTV is used to measure the value of the put option (Deng, Quigley and Van Order, 2000). The unemployment rate is a proxy for other "trigger events". Heckman (1989) shows that as long as the number of time-varying factors is greater than the number of equations, one can identify a competing risk model.

Box-Steffensmeyer and Jones (2004) and Box-Steffensmeyer and Zorn (2001 and 2002) present methods for estimating competing risk hazard models. The three methods are the Stratified Cox Model; the Latent Survival Time Approach; and Multiple Logit models (MNL). Much literature approaches competing risk models using the Stratified Cox approach, which assumes that we can observe multiple risks (say both a stroke and a heart attack for each individual). This is not appropriate for mortgage analysis, where the presence of one event precludes the other. Latent Survival assumes that we only observe the risk that comes first, which does seem appropriate for mortgage analysis. The corresponding approach is to estimate multiple Proportional Hazard Cox models, which is what we did. When we observe one risk, we treat the other risk as censored when we estimate hazards associated with each risk. As a check, we also run models where we exclude those that encounter one kind of hazard (say prepayment) when estimating the other type of hazard (say default). Our results are qualitatively similar using both methods, so we settle on using Cox models employing censoring techniques. The MNL approach is not as rich as the Cox approach, so we do not use it.

4.3. Results

Table 5 reports the baseline maximum likelihood estimates of the parameters of models of competing risks of mortgage default. The model tests the effect of sunk costs in mortgage risks, in addition to current equity. Consistent with the mortgage default literature, default rate decreases with credit score. Marginal effect of each FICO point is similar across different splines. Broker loans, borrowers with higher debt to income ratio as well as lower income are more likely to default. Much to our surprise, we find that older homeowners are also more likely to default. Mortgage rate change measures whether the value of mortgage is "in the money," and the magnitude of the incentive to refinance. Defaults decrease with mortgage rates – prepayment when in the money reduces default risk significantly. Higher unemployment rate increase the hazard of default. All these coefficients are consistent with the existing literature.

Estimated effects of MCLTV by CLTV from baseline regression, after controlling for other risk factors, are reported in Table 4 and also plotted in Figures 2[2]. Our baseline regression controls for state and origination year fixed effects. They confirm that sunk costs are as important as current negative equity in predicting default risk. Except for a few isolated cells, default risks increase with CLTV for a given MCLTV bucket and increases with MCLTV for a given CLTV bucket. Thus sunk costs are an important predictor of default risk. Comparison of change in default probabilities along CLTV and those along MCLTV (up to 100) suggests that incremental risks owing to initial down payment and to our mark-to-market measure are similar in scale. For instance, for starting values of MCLTV =90 and CLTV=80, default rate increases from 1.10 to 1.79 when CLTV moves up to 100 (holding MCLTV constant) and it increases from 1.10 to 1.50 when MCLTV moves up to 100 (holding CLTV constant). There is a

[2] We used state fixed effects (FE) to control for regional labor market, home price appreciation etc. We also tried using MSA fixed effects—regardless of unit of geography, our coefficients on MCLTV and CLTV remained roughly the same. Since recourse is defined by state, it is not significant when included with state FE. We also estimated hazard model based on subset of loans in recourse states vs not. Certain coefficients differ.

consistent tradeoff between MCLTV and CLTV producing the same default hazard, so that one can draw an “iso-default” boundary. Along the MCLTV spectrum, the slope of default risk is steep when MCLTV reaches 100 and then follows a straight line afterwards. Therefore, there is no apparent negative equity trigger (i.e., 110).

These results also suggest borrower characteristics (FICO, income, etc) have significant explanatory power even after controlling for mark-to-market LTV. Hence, borrower characteristics capture importantdimensions that affect mortgage default directly and variables that proxy for borrower characteristics (e.g., origination LTV) are therefore likely to predict default. We will try to isolate the effects of sunk costs from unobserved borrower characteristics using instrument variable and various robustness checks later.

Past home price appreciation could affect homeowners’ expectations on wealth and consumption. As a result, initial down payment can be endogenous to future equity. We focus on three distinct states: California and Florida that have experienced steep and persistent home price declines in the past 2 years, while Texas has experienced only a mild decline. All three states are non-recourse states for purchase money mortgages. While the peak to trough house price decline for California and Floridawas nearly -50%, Texas had nearly flat house prices for the same period. In our sample, loans originated in California have higher FICO but lower CLTV at origination, when compared to other two.

In Table 6, MCLTV in all three states had a statistically significant impact on the probability of default, and the impact generally rises in MCLTV. In Texas, however, the number of loans with very high LTVs is quite small. Generally speaking, if borrowers in Texas had an initial down payment of ten percent or more, they were not underwater. The fact that some borrowers defaulted despite having seemingly positive equity in their houseslikely reflects that fact that we can only know prices at small levels of geography, rather than at the individual household level. It is worth noting that despite the fact that the three states faced very different housing market conditions, the nature of the relationship between CLTV at origination and default probabilities, after controlling for MCLTV, is quite similar.

Because home price used in our baseline analysis is at the zip level, there could be measurement errors in Mark-to-Market CLTV. In Table 7, we used alternative MCLTV measures based on MSA, State and national HPI from Moody's economy.com. Volatility or dispersion diminishes as we increase the geographic boundary of home priced used. At the national level, very few loans are marked to negative equity while there is fairly realistic dispersion in MCLTV using MSA or State HPI. Results in Table 7 confirm that regardless of different HPI, down-payment has a direct relationship with default risk for a given MCLTV range.

Clearly, CLTV at origination is a measure of sunk cost, but it may also measure other aspects of borrowers' financial capacity and/or propensity to default. In a further attempt to show that origination CLTV captures sunk cost, we try instrumenting for CLTV, propensity score matching and look at the credit card default behavior of the same borrowers who have taken out these mortgages.

To test for the exogeneity of CLTV, we performed the Hausman-Wu (Hausman, 1978; Wu, 1973)[3]. As an instrument for individual CLTV, our final choice of instrument was Census

[3] In step 1, we estimated CLTV using the IV and obtained the predictions. We included the predictions as additional regressor in the hazard model. If CLTV is not correlated with error term, then coefficient on predictions should be statistically different from 0.

Track-level average CLTV. We also performed standard Sargan test (Sargan, 1953)[4] to check for over-identifying restrictions – whether the IV we select is correlated with residuals or not.If the IV is endogenous, then it is not valid or weak instrument[5]. Our Wu test results confirmed that CLTV is correlated with residuals, suggesting that itis correlated with unobservables. The Sargan test result can not reject the null hypothesis that tract CLTV is uncorrelated with residuals, suggesting that this it is a valid instrument. Results using the IV approach are reported in Table 8. Although the difference in the impact of CLTV after controlling for MCLTV is, as we would expect, smaller than before, it is qualitatively similar, and economically important. In the MCLTV of 110-120, the default rate among those with down payments of 5 percent or less are more than three times greater than the default rate for those with down payments of 20 percent or more. It is important to note that this Census Tract average CLTV as an instrument does not capture the within-track variations among borrowers; it captures between-track variations (e.g., preferences for schools and local amenities of the neighbourhood).If borrowers within the same neighbourhood have different values on these amenities, the Tract CLTV would not be

---

[4] The Sargan test statistic can be calculated as TR2 (the number of observations multiplied by the coefficient of determination) from the OLS regression of the residuals onto the set of exogenous variables. This statistic will be asymptotically chi-squared under the null that the error term is uncorrelated with the instruments.

[5] We have tested a number of instrument candidates, including FICO, age, income at borrower level that also measures borrower's financial capacity, as well as Census Tract classification designated by HUD for establishing Affordable Housing Goals and Census Tract median income that are exogenous of borrower's credit quality. None of these candidates have passed both Wu and Sargan tests. Our estimate of tract-level CLTV comes from the same proprietary mortgage database we have used throughout the paper. Census Tracts generally have between 2,500 and 8,000 persons and, at the time of definition, are designed to be homogeneous with respect to population characteristics, economic status, and living conditions. Census Tracts are also an important foundation for defining affordable housing goals. Tracts that have given lenders credit for meeting affordable housing goals accounted for a disproportionate chunk of defaults during the financial crisis, and foreclosures are spatially clustered (e.g., Schintler et al 2009). Consequently, tract level CLTV may be produced by an exogenous factor (i.e., qualification for affordable housing goals) and yet are also high correlated with household level CLTV. This makes it a plausible instrument.

sufficient. In our sample, Tract CLTV explains 73% variation in the borrower level LTV. Therefore, the chosen instrument is plausible.

Alternatively, we also estimate the hazard model based on matched sample using PSM. For PSM, we define 80<CLTV<=90 as the treated group while other CLTV categories as control group simply because 80-90 is more likely to have match with either high or low CLTV. We estimate the propensity of being treated using broker dummy, FICO splines, debt to income ratio. Income, borrower age, excess premium, MCLTV, state and year fixed effects. Based on the propensity score, each loan in 80<CLTV<=90 is matched with up to 4 loans from other CLTV buckets. We used only the best match along with treated group in our final hazard model and results are reported in Table 6 as well. The PSM results also confirm a strong direct relationship between CLTV and default while controlling for MCLTV. The difference across different CLTV for a given MCLTV is even greater – suggesting more importance of sunk cost in borrower's behavior.

In addition, we also run three alternative regressions for robustness checks: one using a Logit model instead of a hazard model, the second using serious delinquency (SDQ) – missing three or more payments - instead of foreclosure, the third using linear probability model instead of running hazard model. The Logit model is based on loan level instead of panel data. SDQ by definition occurs earlier than default and occurs roughly twice as often as default. Both the Logit and SDQ regressions are reported in Table 6. The Logit model produces more pronounced impacts of CLTV on default after controlling for MCLTV. This is because of the difference in the specification of dependent variables: in the Logit model, coefficients measure lifetime default rate, while in the hazard model, they measure a quarterly hazard rate. The linear probability model, on the other hand, estimates a much narrow range of coefficients although the impact of down-payment across different MCLTVs remains very significant. This is because the dependent variable is limited to 0-1.

When using SDQ as a measure of default, we find that the estimated impact of CLTV is much smaller although the direct relationship remains significant and very clear. We posit that this is because missing payments can be the result of financial stress independent of the existence of home equity, past or present; the decision to “turn in the keys,” however, is very much a product of home equity.

We also run robustness checks using alternative measures of initial home equity. Instead of CLTV at origination, we estimate two more competing risk hazard models: one using initial equity in dollars and one using initial equity in percent of annual income at origination. This is intended to address the concern over normalization that arises from specifying purchase price. In high cost areas, LTV might underestimate homeowners’ commitment because of much the high denominator used to compute CLTV.

Ourresults are broadly the same, whether we use CLTV, equity investment in dollars, or down payment as a share of income. That is, default risk decreases with initial investment whether measured in absolute dollars or in percent of annual income, controlling for MCLTV. Specifically, households with down payments of less than $5,000 default more often than others after controlling for MCLTV, and households with down payments of less than 10% of income also default more often. This last finding is particularly important, in that it shows that even for those for whom down payment was not a large burden relative to income (i.e., those for whom down payment was about a month of income) are still more likely to default controlling for

MCLTV. This strengthens our view that financial capacity is not the only unobservable driving default behavior.

Our last test is on the effect of initial downpayment on the performance of borrower's non-mortgage liabilities – credit card. If downpayment indeed captures borrower's broad financial capacity, it should predict borrower's credit card performance. If it measures the sunk cost in mortgage contract, we should expect no effect on credit card performance. In Table 9, there is no difference between 80-90 and 90-95 with those with CLTV below 80 in terms of predicting credit card delinquency. While coefficient on 95 above is significantly different from that on 80 below, it is negative, suggesting borrower with less downpayment perform better than those with more. The results suggest down payment may have limited correlation with broad financial capacity.

5. Conclusion

In this paper, we have examined the impact of initial loan-to-value ratio, or sunk cost, on the probability of mortgage default for *prime* mortgages. We have found that borrowers with large initial down payments are in general far less likely to default, even when they are as far below water as those with smaller initial down payments.

These results are consistent with prospect theory; they also may reflect the fact that borrowers with high down payments are in a better position to weather economic downturns, may be less subject to trigger events, and therefore less likely to default. Given that initial down payment is not a statistically significant predictor of default in Texas, however, we think it unlikely that self-selection is driving the relationship between CLTV and default probability. Specifically, if endogenous behavior were the driver behind the relationship between CLTV and default, we would expect it to manifest itself in places where house price movements did not swamp idiosyncratic borrower characteristics. Similarly, our instrumental variable results suggest that it is initial down payment per se, rather than self-selection that influences risk of default.

Finally these results are not consistent with the strategic default hypothesis put forward in Oliver Wyman (2009). If borrowers were strategic and ruthless, initial down payment should have no impact on borrower behavior. Our evidence shows that it does.

References

Agarwal, Sumit, Yan Chang, and Abdullah Yavas, 2012, Adverse selection in mortgage securitization, *Journal of Financial Economics*105 (3), 640–660.

Barberis, Nicholas, Ming Huang, and Tano Santos. 2001. Prospect Theory and Asset Prices, *Quarterly Journal of Economics*, 116(1): 1-53.

Benartzi, Shlomo and Richard H. Thaler. 1995. Myopic Loss Aversion and the Equity Premium Puzzle,*Quarterly Journal of Economics*, 100, 73-92.

Benmelech, Efraim, Mark J. Garmaise, and Tobias J. Moskowitz, 2005, Do liquidation values affect financial contracts? Evidence from commercial loan contracts and zoning regulation, *Quarterly Journal of Economics* 120(3), 1121–1154.

Box-Steffensmeier, Janet M., and Bradford S. Jones. 2004. Event History Modeling: A Guide for Social Scientists. New York: Cambridge University Press.

Box-Steffensmeier, Janet M., and Christopher J. W. Zorn. 2001. Duration Models and Proportional Hazards in Political Science, *American Journal of Political Science*, 45 (October): 972-988.

Box-Steffensmeier, Janet M., and Christopher J. W. Zorn. 2002. Duration Models for Repeated Events,*Journal of Politics*, 64 (November): 1069-1094.

Berger, Allen., Marco A. Espinosa-Vega, W. Scott Frame, and Nathan H. Miller, 2011, Why Do Borrowers Pledge Collateral? New Empirical Evidence on the Role of Asymmetric Information, *Journal of Financial Intermediation* 20, 55-70

Chan, Yuk–Shee, and Anjan V. Thakor, 1987, Collateral and competitive equilibria with moral hazard and private information, *Journal of Finance* 42, 345–363.

DellaVigna Stefano, 2009.Psychology and Economics: Evidence from The Field, *Journal of Economic Literature*, June 2009, Vol. 47, pp. 315-372

Deng, Yongheng, John M. Quigley and Robert Van Order, 2000. Mortgage Terminations, Heterogeneity and the Exercise of Mortgage Options,*Econometrica*, vol. 68(2), pages 275-308

Foster, Chester and Robert Van Order, 1985. FHA Terminations: A Prelude to Rational Mortgage Pricing,*Journal of the American Real Estate and Urban Economics Association* Vol.13, pp. 273-291

Foster, Chester and Robert Van Order, 1984.An Option-Based Model of Mortgage Default.*Housing Finance Review*, 3(4): pp. 351-372

Genesove, David and Christopher Mayer, 2001. Loss Aversion and Seller Behavior: Evidence from the Housing Market, *Quarterly Journal of Economics*, 116(4), 1233-1260.

Guiso, Luigi, Paola Sapienza and Luigi Zingales, 2011.Moral and Social Constraints to Strategic Default on Mortgages, NBER Working Paper 15145.

Hausman, J. A.,1978. Specification Tests in Econometrics,*Econometrica***46** (6): 1251–1271

Heckman, James, 1989. The Identification of the Competing Risks Model, *Biometrika*, 76(2): 325-330.

Inderst, Roman, and Holger M. Mueller, 2007, A lender-based theory of collateral, *Journal of Financial Economics* 84(3), 826–859.

Kau, J.B., D.C. Keenan and T Kim. 1993. Transactions Costs, Suboptimal Termination and Default Probabilities, *Journal of the American Real Estate and Urban Economics Association*, 24(3) 279-999.

Kau, J.B. and D.C. Keenan,1999.Patterns of Rational Default, University of Georgia Working Paper.

Kahneman, Daniel, and Amos Tversky. 1979. Prospect Theory: An Analysis of Decision under Risk, *Econometrica*, 47(2): 263-292.

Kelly, Austin, 2008. Skin in the Game: Zero down payment mortgage default,MPRA Paper 4318, University Library of Munich, Germany.

Kim, E.H., Adair Morse and Luigi Zingales, 2006. What has matted in Economics since 1970,*Journal of Economic Perspectives*, 20(4) 189-202.

List, John A, 2003.Does Market Experience Eliminate Market Anomalies?,*Quarterly Journal of Economics*, 118(1): 41-71.

List, John A, 2004.Neoclassical Theory Versus Prospect Theory: Evidence from the Marketplace,*Econometrica*, 72(2): 615-625.

Odean, Terrance, 1998.Are Investors Reluctant to Realize Their Losses?,*Journal of Finance*, 53(5): 1775-1798.

Oliver Wyman, 2009. Press release on strategic default.

Quigley, John M & Van Order, Robert, 1995.Explicit Tests of Contingent Claims Models of Mortgage Default,*The Journal of Real Estate Finance and Economics*, Springer, vol. 11(2), pages 99-117.

Sargan, John D. 1958. The Estimation of Economic Relationships Using Instrumental Variables, *Econometrica* 26, 393-415.

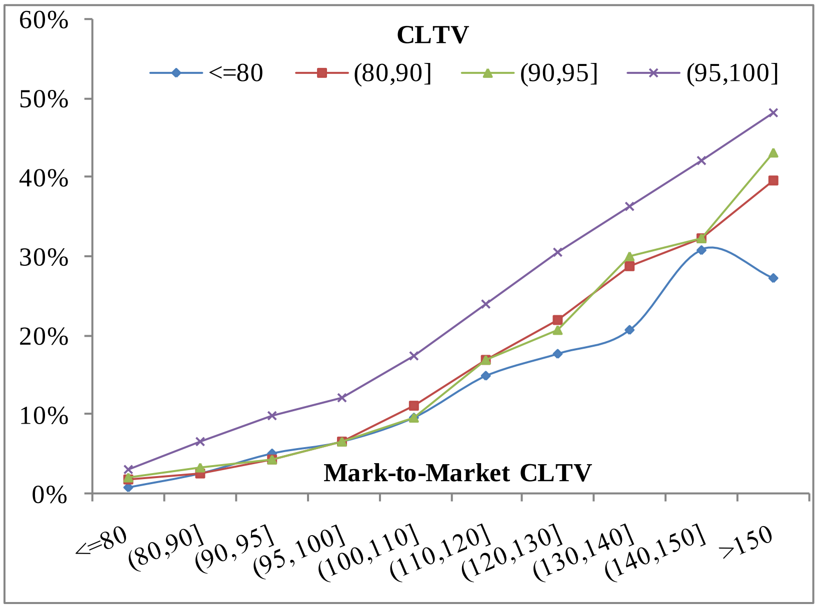
Schintler, Laurie, Emilia Istrate, Danilo Pelletiere, Giacomo Galiazzo and Rajendra Kulkarni, 2009, The Spatial Aspects of the Foreclosure Crisis: A Look at the New England Region, mimeo, George Mason University.

Van den Berg, G.J. 2001. Duration models: Specification, identification, and multiple durations. in J.J. Heckman and E. Leamer, editors, Handbook of Econometrics, Volume V North Holland, Amsterdam.

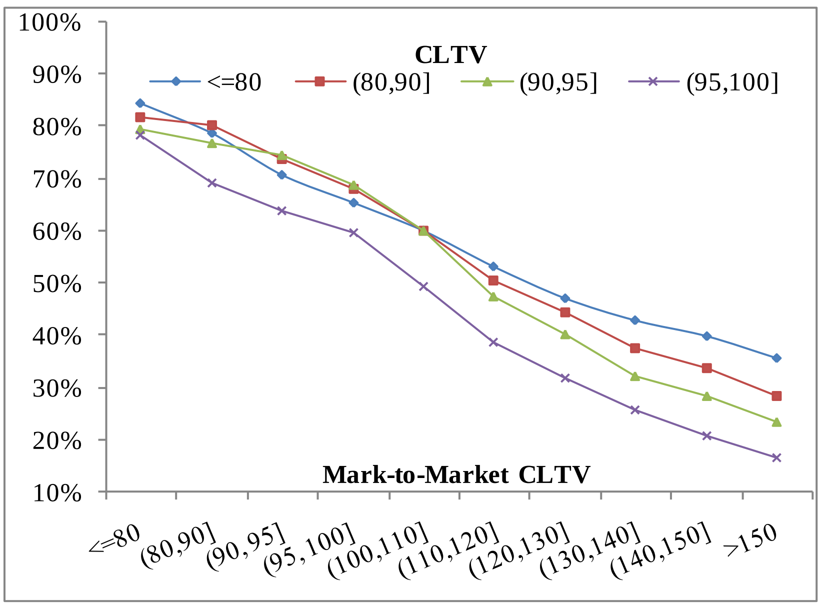
Wu, D. Min. 1973. Alternative Tests of Independence Between Stochastic Regressors and Disturbances. *Econometrica*, 41, 733-750.

Figure 1 Raw Statistics

(a) Raw Default Rate



(b) Raw Prepayment Rate



Note: both charts are based on status at the termination date.

Figure 2 Estimated Coefficients from Proportional Hazard Model

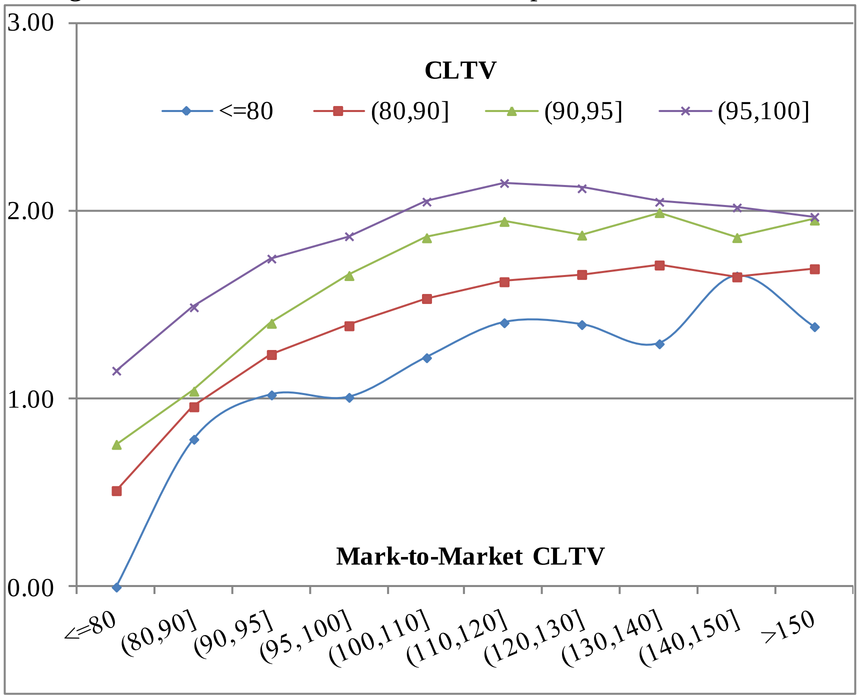


Table 1 Data Description

Variable	Description
Loan-level data	
Note Rate	Gross note rate of the mortgage that borrower is obligated to pay.
FICO	Credit score at origination.
Spline FICO>740	Equals FICO if FICO > 740; 0 otherwise.
Spline 700<=FICO<740	Equals FICO if 700<=FICO<740; 0 otherwise.
Spline 660<=FICO<700	Equals FICO if 660<=FICO<700; 0 otherwise.
Spline 620<=FICO<660	Equals FICO if 700<=FICO<740; 0 otherwise.
Spline FICO<620	Equals FICO if 700<=FICO<740; 0 otherwise.
Combined LTV (CLTV)	Combined loan to value ratio.
Broker Indicator	1 if loan is originated by broker or correspondent lender; 0 otherwise. Broker loans have less skin in the game because they sell loans to big banks.
Debt to Income Ratio	Backend ratio – defined as housing expense plus other debt payment divided by monthly income.
Monthly Income	Borrower's monthly income.
Borrower Age	Borrower's age.
Excess Premium	Difference between mortgage note rate and expected rate. The latter is based on a regression of the former on CLTV, FICO, mortgage product and year-quarter dummies
Mark-to-Market CLTV (MCLTV)	Current balance divided by current property value, adjusted by ratio of combined LTV to LTV.
Loan Age	Duration, in months, from origination to last payment date. For defaults, it's from origination to liquidation date.
Default	1 if property has been foreclosed by the bank; 0 otherwise. Most but not all the SDQ loans lead to foreclosure. Some of them will cure prior to the foreclosure proceeding.
Prepayment	1 if the loan is voluntarily prepaid by borrower (refinance or move); 0 otherwise.
SDQ	Seriously delinquent, defined as borrower missed 3 or more payments. It is considered by alternative definition of default due to lengthy process of foreclosure.
Loan by quarterly activity data	
Default	1 if loan is foreclosed in that quarter; 0 otherwise.
Prepayment	1 if loan is prepaid in that quarter; 0 otherwise.
MCLTV	Mark to market CLTV in that quarter.
Change in Unemployment	Change in unemployment rate of that MSA from origination to current quarter.
Change in Interest Rate	Change in mortgage rate from origination to current quarter.

Table 2 Summary Statistics

Variable	N	mean	sd	min	p25	p50	p75	max
Loan-level data								
Note Rate	680,223	0.064	0.006	0.027	0.060	0.064	0.066	0.130
FICO	680,223	718.311	61.755	330	677	726	770	898
Spline FICO>740	680,223	331.180	383.730	0	0	0	770	898
Spline 700<=FICO<740	680,223	152.874	294.333	0	0	0	0	739
Spline 660<=FICO<700	680,223	120.576	259.994	0	0	0	0	699
Spline 620<=FICO<660	680,223	72.267	202.704	0	0	0	0	659
Spline FICO<620	680,223	41.414	150.638	0	0	0	0	619
Combined LTV (CLTV)	680,223	86.201	15.613	1	80	90	100	100
Indicator CLTV<80	680,223	0.217	0.412	0	0	0	0	1
Indicator 80<=CLTV<90	680,223	0.146	0.353	0	0	0	0	1
Indicator 90<=CLTV<95	680,223	0.156	0.362	0	0	0	0	1
Indicator CLTV>95	680,223	0.316	0.465	0	0	0	1	1
Broker Indicator	680,223	0.568	0.495	0	0	1	1	1
Debt to Income Ratio	680,223	0.418	6.719	0.01	0.31	0.40	0.48	4727
Monthly Income	680,223	6864	4692	1	3963	5833	8417	174870
Borrower Age	680,223	39.452	12.382	18	30	37	47	99
Excess Premium	680,223	0.000	0.004	-0.048	-0.002	0.000	0.002	0.055
Mark-to-Market CLTV (MCLTV)	680,223	86.042	15.823	1.03	79.9	90	98.77	402.85
Loan Age	680,223	46.604	21.890	6	29	48	63	103
Default	680,223	0.104	0.305	0	0	0	0	1
Prepayment	680,223	0.652	0.476	0	0	1	1	1
SDQ	680,223	0.202	0.402	0	0	0	0	1
Credit Card Delinquency	680,223	0.863	0.343	0	1	1	1	1
Loan by quarterly activity data								
Default	6,108,315	0.011	0.107	0	0	0	0	1
Prepayment	6,108,315	0.072	0.259	0	0	0	0	1
MCLTV	6,108,304	86.219	20.274	0	76.46	88.83	97.84	403.75
Change in Unemployment	6,108,315	0.958	1.731	-14.55	-0.07	0.25	1.68	15.37
Change in Interest Rate	6,108,315	-0.197	0.600	-1.74	-0.43	-0.07	0.20	0.97

Table 3 Statistics by Status and Origination Year

(a) Stats by Status

	Defaulted		Prepaid		Other	
Variables	mean	std	mean	std	mean	std
Loan-level data						
Note Rate	0.068	0.006	0.063	0.005	0.064	0.006
FICO	676.011	62.447	732.751	55.485	697.763	63.240
Combined LTV (CLTV)	95.240	7.841	83.497	16.333	89.574	13.938
Broker Indicator	0.688	0.463	0.550	0.497	0.565	0.496
Debt to Income Ratio	0.473	7.141	0.409	7.801	0.418	0.876
Monthly Income	5,776	3,631	7,442	4,989	5,787	3,944
Borrower Age	37.672	12.211	39.711	12.443	39.515	12.228
Excess Premium	0.001	0.005	0.000	0.004	0.000	0.005
Mark-to-Market CLTV (MCLTV)	95.369	8.376	83.281	16.459	89.443	14.257
Loan Age	47.589	19.365	41.746	23.015	59.147	12.959
SDQ	0.984	0.127	0.013	0.113	0.374	0.484
Credit Card Delinquency	0.673	0.469	0.913	0.281	0.811	0.391
% of Obs	10%		65%		24%	
Loan by quarterly activity data						
Default	0.105	0.307	0.116	0.320	6.271	4.019
Prepayment	0.000	0.000	5.837	3.889	89.538	14.177
MCLTV	101.482	20.717	81.923	18.866	90.034	19.480
Change in Unemployment	1.349	1.980	0.839	1.633	1.076	1.810
Change in Interest Rate	-0.244	0.584	-0.185	0.596	-0.206	0.614

(b) Stats by Origination Year

	2005		2006		2007		2008	
Variables	mean	std	mean	std	mean	std	mean	std
Loan-level data								
Note Rate	0.060	0.005	0.066	0.005	0.065	0.005	0.061	0.005
FICO	717.018	61.019	713.554	62.795	711.180	64.197	739.019	51.387
Combined LTV (CLTV)	84.638	16.700	86.789	16.123	88.595	14.667	82.937	14.422
Broker Indicator	0.479	0.500	0.581	0.493	0.650	0.477	0.509	0.500
Debt to Income Ratio	0.420	6.615	0.428	11.302	0.414	0.126	0.407	3.822
Monthly Income	6,247	3,935	6,630	4,450	6,853	4,537	7,918	5,790
Borrower Age	39.833	12.710	39.175	12.512	38.952	12.108	40.278	12.246
Excess Premium	0.000	0.004	0.000	0.004	0.000	0.004	0.000	0.004
Mark-to-Market CLTV (MCLTV)	83.409	16.821	86.436	16.247	88.833	14.900	83.571	14.713
Loan Age	60.116	24.842	49.034	21.249	42.979	17.803	34.122	15.755
Liquidation Default	0.077	0.267	0.121	0.326	0.141	0.349	0.044	0.206
Liquidation Prepayment	0.673	0.469	0.652	0.476	0.593	0.491	0.731	0.443
SDQ	0.154	0.361	0.220	0.414	0.269	0.443	0.115	0.318
Credit Card Delinquency	0.865	0.341	0.856	0.351	0.849	0.358	0.897	0.304
% of Obs	22%		26%		34%		19%	
Loan by quarterly activity data								
Default	0.006	0.075	0.011	0.106	0.019	0.136	0.011	0.102
Prepayment	0.049	0.216	0.062	0.241	0.079	0.270	0.175	0.380
MCLTV	78.969	18.769	87.394	20.276	93.276	20.159	86.995	16.849
Change in Unemployment	0.346	1.611	0.929	1.682	1.512	1.738	1.594	1.481
Change in Interest Rate	0.169	0.519	-0.332	0.501	-0.400	0.566	-0.468	0.656

Table 4Statistics by CLTV

(a) Borrower Characteristics by CLTV Category

	<=80	(80,90]	(90,95]	(95,100]	All
FICO	745.67	734.92	727.22	695.86	718.31
Broker Indicator	0.52	0.55	0.55	0.60	0.57
Debt to Income Ratio	0.42	0.41	0.39	0.43	0.42
Monthly Income	7827.24	7554.11	7785.23	5857.65	6864.35
Borrower Age	45.82	41.01	38.91	36.08	39.45

(b) Distribution by CLTV and Mark-to-Market CLTV (column percent and subtotals)

MCLTV	CLTV				
	<=80	(80,90]	(90,95]	(95,100]	Total
<=80	81.6	36.1	8.1	3.1	26.8
(80,90]	10.2	39.1	28.0	11.1	19.4
(90, 95]	2.2	8.3	24.3	15.6	12.3
(95, 100]	1.4	4.5	14.0	22.2	13.0
(100,110]	2.0	4.8	13.0	28.6	15.9
(110,120]	1.2	2.8	5.0	9.1	5.6
(120,130]	0.7	2.0	3.2	4.2	2.9
(130,140]	0.4	1.2	2.1	2.7	1.8
(140,150]	0.2	0.7	1.3	2.0	1.3
>150	0.1	0.5	0.9	1.5	0.9
Total	19.8	23.0	12.3	44.9	100.0

Table 5 Proportional Hazard Model of Default

	Baseline
	coefficients
Spline: FICO>=740	-0.006***
	(-36.52)
Spline: 700<=FICO<740	-0.006***
	(-33.08)
Spline: 660<=FICO<700	-0.006***
	(-31.24)
Spline: 620<=FICO<760	-0.006***
	(-29.30)
Spline: FICO<620	-0.006***
	(-26.15)
Broker dummy	0.266***
	(30.16)
Debt to Income Ratio	0.001*
	(2.02)
Monthly income	-0.000***
	(-21.67)
Borrower age	0.004***
	(11.53)
Excess premium	48.550***
	(65.66)
Unemployment change	0.188***
	(55.67)
Interest savings	-1.690***
	(-115.41)
State FE	Yes
Origination Year FE	Yes
CLTV x MCLTV Dummies	Yes, next table
Observations	6081056
Pseudo-R2	0.120
Loglikelihood	-696155.593
Chi-square	190217.841

Table 6 Coefficients on CLTV x MCLTV from Proportional Hazard Model by Region

MCLTV	Baseline: Entire Sample				CA			
	<=80	(80,90]	(90,95]	(95,100]	<=80	(80,90]	(90,95]	(95,100]
<=80	**0.00**	**0.70**	**0.90**	**1.28**	**0.00**	**0.56**	**3.08**	**2.68**
(80,90]	**1.10**	**1.27**	**1.35**	**1.79**	**1.28**	**1.83**	**2.09**	**4.10**
(90, 95]	**1.45**	**1.65**	**1.78**	**2.12**	**1.82**	**2.12**	**2.98**	**3.59**
(95, 100]	**1.50**	**1.87**	**2.09**	**2.30**	**1.92**	**2.12**	**2.82**	**3.34**
(100,110]	**1.82**	**2.12**	**2.38**	**2.56**	**2.31**	**2.58**	**2.97**	**3.29**
(110,120]	**2.14**	**2.35**	**2.60**	**2.79**	**2.60**	**2.92**	**3.32**	**3.49**
(120,130]	**2.27**	**2.53**	**2.65**	**2.89**	**2.81**	**3.02**	**3.29**	**3.67**
(130,140]	**2.31**	**2.71**	**2.91**	**2.95**	**2.95**	**3.26**	**3.53**	**3.76**
(140,150]	**2.80**	**2.79**	**2.91**	**3.05**	**3.34**	**3.37**	**3.43**	**3.81**
>150	**2.67**	**2.97**	**3.12**	**3.12**	**3.32**	**3.59**	**3.58**	**3.87**

MCLTV	FL				TX			
	<=80	(80,90]	(90,95]	(95,100]	<=80	(80,90]	(90,95]	(95,100]
<=80	**0.00**	**0.84**	**1.61**	**2.53**	**0.00**	**0.53**	**0.42**	**1.05**
(80,90]	**0.91**	**1.33**	**2.01**	**2.91**	**1.01**	**0.88**	**1.20**	**1.51**
(90, 95]	**1.23**	**1.47**	**1.86**	**2.48**		**1.07**	**1.58**	**1.97**
(95, 100]	**1.41**	**1.64**	**1.74**	**2.30**		**2.27**	**1.65**	**2.16**
(100,110]	**1.55**	**1.95**	**2.11**	**2.46**			**1.09**	**2.26**
(110,120]	**2.02**	**2.08**	**2.25**	**2.44**				**1.94**
(120,130]	**2.00**	**2.20**	**2.30**	**2.57**				**2.90**
(130,140]	**2.22**	**2.51**	**2.58**	**2.66**				
(140,150]	**2.63**	**2.56**	**2.69**	**2.69**				
>150	**1.88**	**2.56**	**2.90**	**2.87**				

Numbers in bold are statistically significant (p-value <= 10%).

Table 7 Coefficients on CLTV x MCLTV from Proportional Hazard Model by Different HP

MCLTV	Baseline				MSA HPI			
	<=80	(80,90]	(90,95]	(95,100]	<=80	(80,90]	(90,95]	(95,100]
<=80	**0.00**	**0.70**	**0.90**	**1.28**	**0.00**	**0.90**	**0.95**	**1.22**
(80,90]	**1.10**	**1.27**	**1.35**	**1.79**	**0.87**	**1.14**	**1.56**	**1.93**
(90, 95]	**1.45**	**1.65**	**1.78**	**2.12**	**1.15**	**1.42**	**1.73**	**2.26**
(95, 100]	**1.50**	**1.87**	**2.09**	**2.30**	**1.23**	**1.68**	**1.83**	**2.44**
(100,110]	**1.82**	**2.12**	**2.38**	**2.56**	**1.40**	**1.92**	**2.15**	**2.46**
(110,120]	**2.14**	**2.35**	**2.60**	**2.79**	**1.68**	**2.03**	**2.48**	**2.62**
(120,130]	**2.27**	**2.53**	**2.65**	**2.89**	**1.97**	**2.16**	**2.48**	**2.81**
(130,140]	**2.31**	**2.71**	**2.91**	**2.95**	**2.11**	**2.46**	**2.62**	**2.90**
(140,150]	**2.80**	**2.79**	**2.91**	**3.05**	**2.38**	**2.46**	**2.77**	**2.94**
>150	**2.67**	**2.97**	**3.12**	**3.12**	**2.63**	**2.60**	**2.82**	**2.83**
MCLTV	State HPI				National HPI			
	<=80	(80,90]	(90,95]	(95,100]	<=80	(80,90]	(90,95]	(95,100]
<=80	**0.00**	**0.90**	-0.45	**-0.41**	**0.00**	**0.91**	**-1.06**	0.04
(80,90]	**0.85**	**1.20**	**1.56**	**1.01**	**0.95**	**1.20**	**1.61**	-0.01
(90, 95]	**0.95**	**1.35**	**1.71**	**2.00**	**1.08**	**1.29**	**1.75**	**1.91**
(95, 100]	**1.05**	**1.57**	**1.93**	**2.31**	**1.67**	**1.47**	**1.77**	**2.22**
(100,110]	**1.07**	**1.63**	**1.97**	**2.33**		**1.44**	**1.70**	**2.23**
(110,120]	**1.10**	**1.77**	**2.05**	**2.35**		**2.32**	**1.79**	**2.09**
(120,130]	**1.55**	**1.82**	**2.15**	**2.44**			**2.55**	**2.39**
(130,140]	**1.60**	**1.82**	**2.09**	**2.43**				**2.64**
(140,150]	**1.68**	**1.84**	**2.27**	**2.38**				**2.20**
>150	**1.97**	**2.11**	**2.28**	**2.32**				1.06

Numbers in bold are statistically significant (p-value <= 10%).

Table 8 Coefficients on CLTV x MCLTV from Alternative Models

MCLTV	Instrument Variable				Propensity Score Matching (PSM)			
	<=80	(80,90]	(90,95]	(95,100]	<=80	(80,90]	(90,95]	(95,100]
<=80	**0.00**	**0.57**	**0.75**	**0.96**	**0.00**	**0.75**	**0.98**	**1.66**
(80,90]	**1.21**	**1.35**	**1.40**	**1.77**	**1.19**	**1.28**	**1.37**	**2.15**
(90, 95]	**1.75**	**1.81**	**1.89**	**2.20**	**1.47**	**1.74**	**1.83**	**2.48**
(95, 100]	**2.06**	**2.15**	**2.26**	**2.40**	**1.58**	**2.01**	**2.14**	**2.67**
(100,110]	**2.30**	**2.41**	**2.57**	**2.68**	**1.86**	**2.30**	**2.51**	**2.94**
(110,120]	**2.53**	**2.67**	**2.82**	**2.91**	**2.35**	**2.60**	**2.84**	**3.21**
(120,130]	**2.83**	**2.93**	**2.97**	**3.05**	**2.54**	**2.67**	**2.85**	**3.33**
(130,140]	**2.83**	**3.10**	**3.18**	**3.19**	**2.51**	**2.95**	**3.15**	**3.37**
(140,150]	**3.06**	**3.21**	**3.26**	**3.29**	**2.86**	**3.10**	**3.12**	**3.49**
>150	**3.31**	**3.41**	**3.49**	**3.43**	**3.29**	**3.29**	**3.40**	**3.59**

MCLTV	SDQ				Logit			
	<=80	(80,90]	(90,95]	(95,100]	<=80	(80,90]	(90,95]	(95,100]
<=80	**0.00**	**0.34**	**0.51**	**0.77**	**0.00**	**0.76**	**0.90**	**1.03**
(80,90]	**0.86**	**0.79**	**0.84**	**1.15**	**1.19**	**1.33**	**1.47**	**1.83**
(90, 95]	**1.11**	**1.22**	**1.17**	**1.39**	**1.63**	**1.78**	**1.90**	**2.28**
(95, 100]	**1.09**	**1.38**	**1.52**	**1.55**	**1.73**	**2.07**	**2.28**	**2.52**
(100,110]	**1.39**	**1.54**	**1.80**	**1.79**	**2.09**	**2.43**	**2.63**	**2.84**
(110,120]	**1.58**	**1.72**	**2.00**	**2.04**	**2.47**	**2.76**	**2.97**	**3.22**
(120,130]	**1.78**	**1.85**	**2.03**	**2.15**	**2.63**	**2.96**	**3.09**	**3.43**
(130,140]	**1.81**	**1.99**	**2.15**	**2.18**	**2.73**	**3.29**	**3.49**	**3.60**
(140,150]	**1.97**	**2.06**	**2.27**	**2.25**	**3.28**	**3.37**	**3.53**	**3.78**
>150	**1.98**	**2.18**	**2.32**	**2.29**	**3.00**	**3.64**	**3.91**	**3.93**

MCLTV	Linear Probability			
	<=80	(80,90]	(90,95]	(95,100]
<=80	**0.00**	**0.00**	**0.00**	**-0.01**
(80,90]	**0.01**	**0.02**	**0.02**	**0.02**
(90, 95]	**0.02**	**0.03**	**0.03**	**0.05**
(95, 100]	**0.02**	**0.04**	**0.05**	**0.07**
(100,110]	**0.05**	**0.07**	**0.08**	**0.12**
(110,120]	**0.09**	**0.12**	**0.13**	**0.18**
(120,130]	**0.12**	**0.16**	**0.16**	**0.24**
(130,140]	**0.14**	**0.22**	**0.24**	**0.29**
(140,150]	**0.24**	**0.25**	**0.26**	**0.34**

Numbers in bold are statistically significant (p-value <= 10%).

Table 9 Impact of Initial Down Payment on Credit Card Delinquency

CLTV	
(80, 90]	-0.001
	(-0.41)
(90, 95]	-0.003
	(-1.94)
> 95	-0.020***
	(-16.21)
Other Controls	Yes
State FE	Yes
Origination Year FE	Yes
Observations	678,256
Adj R2	0.051