Default Option Exercise over the Financial Crisis and Beyond *

Xudong An†
Federal Reserve Bank of Philadelphia

Yongheng Deng‡
University of Wisconsin - Madison

Stuart A. Gabriel§
UCLA

Abstract

We document changes in borrowers' sensitivity to negative equity and show heightened borrower default propensity as a fundamental driver of crisis period mortgage defaults. Estimates of a time-varying coefficient competing risk hazard model reveal a marked run-up in the default option beta from 0.2 during 2003-2006 to about 1.5 during 2012-2013. Simulation of 2006 vintage loan performance shows that the marked upturn in the default option beta resulted in a doubling of mortgage default incidence. Panel data analysis indicates that much of the variation in default option exercise is associated with the local business cycle and consumer distress. Results also suggest elevated default propensities in sand states and in the wake of enactment of crisis-period loan modification programs.

JEL Classification: G21; G12; C13; G18

Keywords: Mortgage default; option exercise; default option beta; time-varying coefficient hazard model

This draft: October 28, 2019

---

* We thank Sumit Agarwal, Yacine Ait-Sahalia, Gene Amromim, Linda Allen, Brent Ambrose, Bob Avery, Gadi Barlevy, Neal Bhutta, Shaun Bond, Alex Borisov, Raphael Bostic, John Campbell, Paul Calem, Alex Chinco, John Cotter, Larry Cordell, Tom Davidoff, Moussa Diop, Darrell Duffie, Ronel Elul, Jianqing Fan, Andra Ghent, Matt Kahn, Bill Lang, David Ling, Crocker Liu, Jaime Luque, Steve Malpezzi, Andy Naranjo, Raven Molloy, Kelley Pace, Erwan Quintin, Tim Riddiough, Dan Ringo, Amit Seru, Shane Sherlund, Steve Ross, Eduardo Schwartz, Joe Tracy, Alexi Tschisty, Kerry Vandell, Paul Willen, Wei Xiong, Vincent Yao, Abdullah Yavas and conference and seminar participants at AFA/AREUEA, Baruch College, the Federal Reserve Bank of Chicago, Cornell University, Federal Reserve Bank of Philadelphia, Federal Reserve Board, Georgia State University, the Homer Hoyt Institute, National University of Singapore, Tel Aviv University, Urban Economics Association, UC Irvine, UIUC, University of Cincinnati, University of Connecticut and University of Wisconsin for helpful comments. The authors acknowledge financial support from the UCLA Ziman Center for Real Estate and the NUS Institute of Real Estate Studies. The authors also gratefully acknowledge the excellent research assistance provided by Chenxi Luo and Xiangyu Guo. The views expressed here are not necessarily those of the Federal Reserve Bank of Philadelphia or the Board of Governors of the Federal Reserve System. All remaining errors are our own responsibilities.

† Department of Supervision, Regulation, and Credit, Federal Reserve Bank of Philadelphia. Email: Xudong.An@phil.frb.org;

‡ Corresponding author. Wisconsin School of Business, University of Wisconsin - Madison. Email: yongheng.deng@wisc.edu;

§ UCLA Anderson School of Management. Email: sgabriel@anderson.ucla.edu.

Default Option Exercise over the Financial Crisis and Beyond

1. Introduction

Default on residential mortgages skyrocketed during the late-2000s, giving rise to widespread financial institution failure and global financial crisis. Among factors associated with mortgage failure, analysts have pointed to the widespread incidence of negative equity, shocks to unemployment and income, lax underwriting, expansive use of risky loan products, and fraud, to name a few.[1] In this paper, we provide new evidence of heightened borrower sensitivity to default in response to negative equity and show that factor to be highly salient to crisis period defaults. The run-up in borrower propensity to default, coupled with a decline in home equity, resulted in a widespread increase in mortgage default during the 2000s crisis. Results also suggest elevated default option exercise in the wake of enactment of crisis-period loan modification programs, providing yet another example of the Lucas critique.

To empirically identify changes in borrower response to negative equity, we apply a time-varying coefficient competing risk hazard model to loan-level event-history data. We model the conditional probability of default as a function of contemporaneous borrower negative equity and a large number of other factors. We label the estimated coefficient associated with negative equity the "default option beta".[2] Contrary to existing mortgage default literature, we allow the default option beta to vary over time and place.

We estimate our models using expansive microdata on loan performance during the 2000-

[1] The long list of crisis references include but are not limited to Gerardi, et al, 2008; Mayer, Pence and Sherlund, 2009; Mian and Sufi, 2011; Keys, et al, 2010; Mian, Sufi, and Trebbi, 2010; Rajan, Seru, and Vig, 2015; Agarwal et al, 2014, 2017; An, Deng and Gabriel, 2011; Demyanyk and Van Hemert, 2011; Brueckner, Calem and Nakamura, 2012; Taylor and Sherlund, 2013; Cotter, Gabriel, and Roll, 2015, Demyanyk and Loutskina, 2016, etc.

[2] This nomenclature is consistent with prior literature on mortgage default (see, for example, Deng, Quigley, and Van Order, 2000)

2018 period. Our primary dataset includes monthly mortgage performance history on fixed-rate 30-year home mortgage loans from the Government Sponsored Enterprises (GSEs) Fannie Mae and Freddie Mac. We corroborate findings using micro-data from private-label securitizations. Results of rolling window local estimation of the hazard model show a marked run-up in the default option beta from 0.2 during 2003-2006 to about 1.5 during 2012-2013, followed by a downward retracing of roughly one-half the upward movement in the estimated beta value through 2016 (Figure 1). The upward movement in the default option beta led to substantially higher default probabilities for a given level of negative equity (Figure 2). Results of simulation show that for 2006 vintage loans, the marked upturn in the default option beta resulted in a doubling of default incidence (Figure 3). We also find substantial geographic heterogeneity in the default option beta. Figure 4 shows dramatic cyclical movements in the default option beta among all sampled states. Further, while all of the state-specific default option beta time-series rise with the onset of the crisis, they differ markedly in slope and amplitude. At peak in 2012-2013, the default option beta in hard-hit Florida, at 1.8, was one and one-half times that of Texas.

In research dating from the 1980s, mortgage default is modeled as borrower exercise of the put option (see literature reviews by Quercia and Stegman, 1992 and Kau and Keenan, 1995). Empirical findings have shown that negative equity, or the intrinsic value of the default option, is a major driver of default (see, for example, Giliberto and Ling, 1992, Quigley and Van Order, 1995, and Deng, Quigley and Van Order, 2000). Recent research, however, indicates that home equity must turn deeply negative before most borrowers exercise the default option (see, for example, Bhutta, Dokko, and Shan, 2017). Indeed, numerous authors have suggested that residential default may be less than ruthless (see, for example, Ambrose, Buttimer and Capone, 1997: Deng and Gabriel, 2006). Our results document systematic variability in default option exercise among a cross-section of states and over the economic cycle.[3]

We explore heterogeneity in borrower propensity to default via a simple theoretical framework. Our theoretical model builds on existing literature and assumes that borrowers have rational expectations and engage in default to maximize wealth (see, for example, Kau et al., 1992; Riddiough and Wyatt, 1994b; Ambrose, Buttimer and Capone, 1997; and Campbell and Cocco, 2015). Our model suggests borrower propensity to default can vary over time due to factors such as changing borrower expectations on the path of the local economy, borrowers' subjective assessment of the conditional probability of foreclosure (versus workout), changing default transaction costs (including stigma effects), and the like. For example, pessimism about the future trajectory of house prices could make the borrower more sensitive to a negative equity position. Similarly, expectations of loan modification conditional on default could also lead to elevated default option exercise.

We employ proxies for factors identified in theory to empirically assess drivers of the observed variation in the default option beta. We find that county unemployment rate shocks, reflecting cyclical fluctuations in the local economy, are highly predictive of variation in the default option beta. Conditional on controls for the local business cycle, we find that borrower default propensities are sensitive to consumer distress, where our measure of distress is orthogonalized to current economic fundamentals. Further, those factors are economically salient and together could account – via their impact on the default option beta – for over two-thirds of the increase in crisis period default risk (Figure 5). We also find evidence of a structural break in the default option beta time-series in 2009, which coincides with federal mortgage market intervention via the Home Affordable Modification Program (HAMP). Finally, while results do

[3] In related literature on corporate default, Duffie et al. (2009) find evidence of dynamic variation in the role of common latent factors in predicting firm level default.

not show significantly damped default propensities among states with recourse to borrower non-housing assets, they do indicate sizable and significantly elevated default option betas among sand states hard hit by the 2000s housing and mortgage crisis. Together, these factors explain over 70 percent of the variation in the default option beta panel.

We also seek to shed light on the structural break in default option exercise in 2009. A difference-in-differences analysis shows that those eligible for HAMP loan modification became significantly more sensitive to negative equity in the wake of program implementation, relative to the non-HAMP eligible control group. This finding is consistent with the notion that mortgage borrowers may be strategic and hence, more likely to become delinquent when they expect lenders to modify defaulted loans (see, for example, Guiso, Sapienza, and Zingales, 2013; Mayer et al., 2014).[4]

Our findings are robust to alternative model specifications and loan samples. While our primary sample is comprised of conventional conforming loans from the GSEs, we re-estimate the model using non-agency loans and confirm a similar pattern of default option beta variation. We assess the robustness of findings to book vs. market value of negative equity, controls for the non-linearity of negative equity, and size of the estimation rolling window (e.g., two vs. three years). We further evaluate robustness in the default option beta among borrowers less likely to be liquidity constrained. In addition, we test specifications of the model that account for default burnout. Finally, we estimate the model using annual cohorts to assess whether changes in the mix of borrowers may have contributed to the observed variation in the default option beta. Results throughout indicate a similar countercyclical pattern of default option beta over the crisis period

[4] Piskorski and Tchistyi (2011) also argue that bailing out the most distressed borrowers in the crisis period encourages irresponsible financial behavior during the boom. Ghent and Kudlyak (2011) find that borrowers in non-recourse states are more sensitive to negative equity.

and beyond.

Our findings contribute to the literature in several important ways. First, results provide new insights into the cyclical pattern of borrower default during the financial crisis and beyond. Among relevant crisis-related analyses (see, for example, Mian and Sufi, 2009; Keys et al., 2010; Piskorski, Seru, and Witkin, 2015; Rajan, Seru, and Vig, 2015, among many others), temporal shifts in default behavior among mortgage borrowers have received only limited attention. Here we show that changes in the propensity to exercise the mortgage default option were material to the crisis.

Second, our findings raise important issues of modeling and management of mortgage default risk in an ever-changing market environment. As evidenced in recent studies, statistical models may substantially underestimate default risk in the presence of economic fluctuations, policy intervention, and behavioral change (see, for example, An et al., 2012; Rajan, Seru, and Vig, 2015). Indeed, the assumption of a fixed and static default option beta may result in significant under-prediction of default risk (An et al., 2012). The time-varying coefficient hazard model better characterizes ongoing evolution in borrower default behavior so as to enhance risk management.

Third, our study adds to the growing literature on strategic default (see, for example, Riddiough and Wyatt, 1994a; Guiso, Sapienza and Zingales, 2013; Mayer et al., 2014). Mortgage default is more than a one-sided process and often involves strategic interaction between borrowers and lenders. Our results suggest that in anticipation of lender or servicer actions, borrowers' willingness to exercise the default option may change as well.

Finally, our study has important implications for federal policy enacted during the crisis period. While HAMP saved many defaulted borrowers from foreclosure (see, e.g., Agarwal et al., 2017), our findings suggest this program also may have had an unintended consequence of inducing some borrowers to enter into delinquency. While we are silent on the ultimate impact of HAMP on borrower well-being and social welfare, it appears that the efficacy of HAMP in mitigating home foreclosure may have been diminished by an increase in default option exercise among borrowers seeking a HAMP loan modification.

The remainder of the paper is organized as follows: in the next section, we discuss our data; in section 3, based on hazard model estimates, we document the time-series and cross-sectional variations in the default option beta; in section 4, we explore factors that drive variations in the default option beta; and section 5 provides concluding remarks.

2. Data

2.1. Data sources

Our primary dataset consists of loan-level information in the Freddie Mac's Single-Family Loan-Level Dataset and Fannie Mae's Single-Family Loan Performance Data (hereafter Freddie Mac data and Fannie Mae data, respectively, and GSE data altogether).[5] The GSE data are comprised of over 50 million loans, on which we focus on the roughly 42 million fixed-rate 30-year mortgage loans acquired by the GSEs over the 2000-2016 period. The sheer size of the dataset, as well as the richness of its content, is unparalleled. It provides detailed information on borrower and loan characteristics at origination, including the borrower's credit score, origination loan balance, note rate, loan term, loan-to-value ratio (LTV), debt-to-income ratio (DTI), loan purpose (home purchase, rate/term refinance, cash-out refinance), occupancy status, prepayment penalty indicator, and the like. The GSE data also track the performance (default, prepayment, mature, or current) of each loan every month, which is crucial to our default risk modeling.

[5] See http://www.fanniemae.com/portal/funding-the-market/data/loan-performance-data.html as well as http://www.freddiemac.com/news/finance/sf_loanlevel_dataset.html. We standardize the variables and formats between the Fannie Mae and Freddie Mac datasets.

We merge the loan-level data with proxies for the labor market, housing, and other macroeconomic conditions. For example, to obtain a measure of negative equity for each loan in each quarter, we merge the loan event history with a county-level house price index (HPI) from CoreLogic Solutions (hereafter CoreLogic). We also utilize the HPI to compute time-varying house price return volatility, which either enters the model as a standalone variable or via a volatility-adjusted default option value term. To calculate the prepayment option value for each loan in each quarter, we merge mortgage interest rates from the Freddie Mac Primary Mortgage Market Survey to our loan event history. In addition, we supplement our mortgage data with macroeconomic variables, including the county-level unemployment rate from Bureau of Labor Statistics, Treasury bond rate from the Federal Reserve Board, and CredAbility Consumer Distress Index retrieved from St. Louis Fed. For purposes of robustness, we also estimate our models using loan-level data from BlackBox Logix (BBX) for non-agency securitized mortgages.[6] Additional information on data and variable construction is found later in the paper.

2.2. Sample and descriptive statistics

In our main analysis, we focus on first-lien, full documentation, and fully-amortizing 30-year fixed-rate (FRM) mortgage loans acquired by the GSEs during 2000-2016 for Washington DC and nine representative states, including Arizona, California, Florida, Georgia, Illinois, Massachusetts, Michigan, New York, and Texas.[7] Our focus on narrowly defined loans and borrowers (only 30-year fixed-rate mortgages) allows us to draw inference on default behavior from a relatively homogeneous sample. The distribution of loans among sampled U.S. states

---

[6] The BBX data is accessed through UCLA's Ziman Center for Real Estate and the James A. Grasskamp Center for Real Estate at the Wisconsin School of Business. The data contains roughly 22 million loans. We focus on the fixed-rate fully amortizing mortgages.
[7] 15-year loans are excluded. A series of filters is also applied: we exclude loans with missing or wrong information on loan origination date, original loan balance, borrower credit score, loan-to-value ratio (LTV), or debt-to-income ratio (DTI).

allows ample cross-sectional variation in our time-series measures. We limit the analysis to larger states to ensure we have adequate loan samples for subsequent estimation of a panel data model based on state-level rolling window default option betas. Our sample contains 42,093,277 individual mortgage loans.

In Table 1, we report descriptive statistics of loan and borrower characteristics. The average loan amount at origination is $199,681, and the average note rate is 5.52 percent. The mean borrower's credit score is 737, and approximately three-quarters of the loans are for single-family properties. Cash-out refinance, and rate/term refinance mortgages comprised 26 and 30 percent of the sample, respectively. Owner-occupied loans comprise 90 percent of our sample, whereas investment property loans constitute 6 percent. About 23 percent of our sampled loans were originated pre-2003, and 28 percent of sampled loans were originated during the 2003-2007 pre-crisis boom period, whereas one-half of sampled loans were originated during the Great Recession and its aftermath through 2016. The average combined LTV is 75 percent. We also calculate a mean total debt-to-income ratio of 25 percent. All sample loans are underwritten with full documentation. Over the sample period, 68 percent of loans prepay, whereas 6 percent of loans terminate in default. At the time of data collection (March 2018), about 26 percent of our loans were still performing and hence were censored observations in our model.

3. Rise in Mortgage Default Propensities

3.1. Default hazard models

We follow the existing literature in estimating the competing risks of mortgage default and prepayment in a proportional hazard framework (see, e.g., Deng, Quigley, and Van Order, 2000; Deng and Gabriel, 2006). As noted above, the hazard model is convenient primarily because it allows us to work with the full sample of loans despite the censoring of some observations. The specification of the model is motivated by option theory, which predicts that rational mortgage borrowers will exercise the default or prepayment option to maximize their wealth. Theory suggests that mortgage borrowers will exercise the default option when the value of the mortgage exceeds the value of the collateral. Similarly, borrowers will exercise the prepayment option when the market value of the mortgage exceeds its book value. These two options compete against each other.[8] Among other drivers of default, recent research has underscored the importance of crisis-period income and liquidity shocks as default triggers (see, for example, Foote, Gerardi, and Willen, 2008; Elul et al, 2010; Gyourko and Tracy, 2014; Campell and Cocco, 2015; and Gerardi, et al, 2015 for the double trigger argument).[9] As discussed below, we also include proxies of income shocks in our model specification.

As in much of the literature, we define default as mortgage delinquency in excess of 60-days. An important consideration in this definition of default is that lenders and servicers typically intervene in the default process only after 60-day delinquency; as such, the 60-day delinquency event reflects the borrower decision-making, as is the focus of this paper. Prepayment refers to early repayment of a loan as a result of borrower relocation or refinancing for purposes of a lower interest rate, different loan terms, or cash out.

The literature typically assumes the hazard rate of mortgage loan termination at period T since origination is of the form

$$h_i^k(T, Z_{i,t}') = h_0^k(T)\exp(Z_{i,t}'\beta^k) \tag{1}$$

where k indicates default or prepayment, T is duration time, t indicates calendar time, i is the

[8] Kau et al. (1992) and Kau and Keenan (1995) have outlined the theoretical relationships among the options, and Schwartz and Torous (l993) have demonstrated their practical importance.

[9] According to the double-trigger argument, negative equity is a necessary but not sufficient condition for mortgage default. That argument further stresses the importance of income shocks to default. Low (2015) presents evidence on positive equity and default.

individual loan, and $Z'_{i,t}$ is a vector of covariates for loan i that includes all identifiable risk factors.[10] Here $h_0^k(T)$ is the baseline of the hazard function. In the proportional hazard model, changes in covariates shift the baseline hazard rate proportionally without otherwise affecting the duration pattern of default or prepayment. Covariates include proxies for default and prepayment option values, borrower credit score, payment (debt) to income ratio, loan amount, and a host of other loans, borrower, and locational characteristics.

In our analysis, we allow the coefficient of the default option in the hazard model to be time-varying so as to focus on possible intertemporal variation in the sensitivity of borrower default probability to negative equity. Therefore, our model becomes a time-varying coefficient (partially linear) model of the form

$$h_i^k(T, Z'_{i,t}) = h_0^k(T)\exp(Z'_{i,t}\beta_t^k) \tag{2}$$

To estimate a time-varying coefficient hazard model, we adopt the rolling window local estimation approach from the statistics literature (see, e.g., Cleveland, Grosse, and Shyu, 1991; Fan and Zhang, 1999). The idea is that the time-varying coefficient model can be treated as locally linear, allowing us to assume the coefficients are smooth for each short time window and to apply the usual estimation method to obtain a local estimator.[11] In that regard, we form quarterly three-year rolling windows to construct our local estimation samples. As discussed below, we also assess robustness of results to the size of the rolling window.

The hazard model is estimated using loan event-history. We construct the quarterly performance history of each loan based as reported in the GSE data, as well as a number of time-varying explanatory variables. Negative equity is defined as the percentage difference between the

[10] Notice that the loan duration time T (tau) is different from the calendar time t, which allows identification of the model.
[11] More sophisticated methods include a two-step procedure presented in Fan and Zhang (1999), and Fan and Zhang (2008).

book value of the loan and the market value of the property relative to the market value of the property. [12] The default option is defined as the cumulative distribution function value of negative equity over county house price return volatility[13]. In an alternative specification, we replace the default option variable with spline functions of negative equity and a standalone house price return volatility term. Results are robust to that transformation and are reported in Appendix figures and tables. The market value of the property is calculated based on property value at time of loan origination plus/minus any change therein as indicated by a local house price index (HPI). The prepayment option value is computed as the contemporaneous difference between the market value of the loan and its book value. The book value of the loan is the remaining mortgage balance (from the loan amortization schedule) whereas the market value of the loan is computed based on the remaining mortgage payments discounted at the current prevailing mortgage interest rate in the market (see, for example, Deng, Quigley and Van Order, 2000). We also use the change in the state unemployment rate from loan origination through termination or censor of the loan to approximate income shocks.[14] Sample statistics of the time-varying covariates are reported in Table 2.

Time-fixed covariates included in the hazard model include loan and borrower characteristics such as borrower credit score, loan-to-value ratio (LTV), debt-to-income ratio (DTI), loan amount, loan purpose, property type, occupancy type, first-time buyer status, and the like. We also include state-fixed effects and vintage-fixed effects. State-fixed effects account for the possible impact of varying state foreclosure laws on default probability, among other things,

[12] As a robustness check, we use the market value instead of the book value of the mortgage to calculate negative equity. The resulting default option beta demonstrates similar time variation.
[13] House price return volatility is a scaling factor.
[14] Butta, Dokko and Shan (2016) use local credit card delinquency rates as an alternative measure of income shocks. We test such an alternative as well as a zip-code level income change measure based on IRS data and find our results to be robust to those alternative specifications.

whereas vintage-fixed effects control for unobserved changes over time in underwriting standards. To account for potential non-linearities, we include square terms of such key variables as a change in the local unemployment rate and the default and prepayment option values. Moreover, for variables such as borrower credit score, LTV, DTI, and loan amount, we use granular buckets.

3.2. Default Option Beta Time Series

Prior to the presentation of our rolling window estimates and to assure the reasonableness of model specification, we examine a pooled-sample baseline model. Estimates of the baseline model are reported in Table 3. Standard errors clustered at the loan-level are reported in the table as well. As is evident, model coefficients conform to economic intuition and findings in the existing literature (see, e.g. Deng, Quigley, and Van Order, 2000; Deng and Gabriel, 2006). For example, the default option value is positively related to default risk. That relationship is non-linear, as reflected by the significance of the default option square term.

Similarly, as expected, the value of the prepayment option is positively related to the risk of prepayment. Consistent with the competing risk model specification, the estimated coefficients of the default option and prepayment option values also are statistically significant in the prepayment and default hazard models, respectively. As evidenced in Table 3, coefficients on LTV buckets in the default equation are significant throughout and increase monotonically over LTV levels. Further, lower levels of LTV are associated with negative default risk, whereas substantial positive coefficients are estimated for the higher LTV categories. In a similar vein, the estimated coefficients on the credit score categorical terms decline monotonically with credit score bucket. As expected, while coefficients associated with lower credit score categories are positive and significant in determination of default risk, those estimates turn negative, sizable, and significant for the higher credit score buckets. Also, monotonic and increasing coefficients are estimated for the debt-to-income categorical terms in the default hazard. Almost all other controls enter the specified competing risk equations with anticipated signs and a high level of statistical significance.

As discussed above, our focus is on the time variation in the default option beta. In that regard, we use the cumulative distribution function value of negative equity over house price return volatility to represent the default option value.[15] Given the presence of the square term in default option value, the default option beta is calculated as the coefficient of the default option term plus two times the coefficient of the default option square term times the mean value of the default option term – the first-order partial derivative of the default hazard rate with respect to value of the default option.

In Figure 1, we display rolling window estimates of the default option beta from equation (2). We plot both the estimated beta together with its confidence band. Clearly evidenced are sizable and significant intertemporal variations in the estimated beta. In that regard, the default option beta rose gradually from about 0.2 during 2003-2006 to almost 1.3 in 2009. That estimate continued to trend up in the immediate aftermath of the crisis to peak at about 1.6 during 2012-2013. Subsequent to that, a clear trending down in the default option beta was evidenced; nonetheless, as recently as 2016-Q2, the estimated beta remained elevated at close to 1.0. Overall, results indicate statistically significant countercyclical movement in the default option beta over the 2003-2016 timeframe of the analysis.[16] To assure robustness of results to loan samples, we re-estimated our models using non-agency securitized loans from BBX in place of our GSE loan sample. Results indicate the time series pattern in the default option beta is robust to loan sample.[17]

[15] See, for example, Deng, Quigley and Van Order, 2000 and Deng and Gabriel, 2006.

[16] As shown in Appendix Figure 1, a similar and marked cyclical trend is evidenced in the negative equity spline beta estimates.

[17] Some of these results appear in prior versions of our paper. Additional results are available upon request.

To provide insights as to the economic significance of changes in the mean estimated default option beta, we plot in Figure 2, the impact of borrower negative equity on default probability in the years 2006 and 2012. Interestingly, a sizable increase in negative equity had limited impact on default probability in 2006. In marked contrast, by 2012 the impact of negative equity on loan default probability was sizable. In that year, a loan with 30 percent negative equity had over a roughly 170 percent chance of entering into default as compared to a loan with 10 percent negative equity. In addition, a loan with 40 percent negative equity had over a 190 percent chance of entering into default as compared to a loan with 10 percent negative equity. Loans with negative equity in the range of 10 - 30 percent witnessed an increase in the default hazard ratio of 60 – 150 percent between 2006 and 2012.

To gain further insights regarding the economic significance of those findings, we use model estimates from 2002-2004 vintage loans to predict default in the 2006 vintage.[18] Per the above, if the default option beta was lower in 2002-2004, a model estimated with data from those years will underpredict default in the 2006 vintage. In order to isolate the impact of changes in the default option beta from that in the default option value itself, we assume perfect foresight in house price movement for the 2006 vintage. The red solid line in Figure 3 shows the actual performance of the 2006 vintage with the correct default option beta. The blue dashed line in the same figure shows the model prediction with the estimated default option beta from the 2002-2004 vintage. Over a 20-quarter horizon, the predicted default rate with the 2002-2004 beta is only about half of the actual default rate. As a comparison, the default rate of GSE 30-year FRMs during 2002-2006 was about one-third that of 2006-2010. Taken together, along with the decline in home equity, the upturn in the default option beta figured importantly in the run-up in default.

[18] See An et al (2012) for a similar simulation.

As is evident in Figure 1, the estimated movement over time in the default option beta appears to be strongly correlated with cyclical fluctuations in house prices and the broader economy. During the pre-crisis boom years and in the context of strong housing market performance, the estimated beta was small in magnitude. As boom turned to bust, the default option beta rose markedly. Finally, in the wake of the post-downturn expansion and as economic conditions improved, the household propensity to exercise the default option again declined.

For purposes of robustness, we replace the default option term with spline functions of negative equity and a standalone house price return volatility term. Results shown in Appendix Table 1 are consistent with the findings of our primary specification. For example, the spline function shows that negative equity is positively related to default probability and the relation is non-linear. Moreover, house price return volatility is negatively associated with default probability, consistent with findings in classical models of default risk that show that borrowers delay defaulting on their loan when asset volatility unexpectedly increases. Using the model shown in Appendix Table 1, we conduct rolling window estimation of the model. Results show a marked increase in the negative equity beta during 2009-2013 for borrowers with negative equity.[19] Interestingly, for borrowers with significant positive equity in their home, the relation between default probability and the depth of equity remains stable over the full sample period.

Among other robustness checks, we estimate the rolling window model using different window sizes (24 vs. 36 months). The results are robust to that transformation. Note also that the use of HPI to calculate negative equity may result in measurement error relative to the idiosyncratic house price change that a given homeowner might experience. This measurement error could bias

[19] Appropriately accounting for non-linearity helps to provide assurance that we are assessing the effects of changes in borrower tendency to default for a given level of negative equity, rather than time-variant changes in the distribution of negative equity.

the estimated default option betas. To the extent that this measurement error varies over time, the same borrower behavior could manifest with a changing beta. To assess the salience of this issue, we regressed individual house price returns (based on repeated CoreLogic real estate deeds transactions data) on the CoreLogic zip code level house price index (HPI) returns.[20] Appendix Figure 3 provides a metric of house price measurement error over our study period as represented by the mean absolute error (MAE) of that regression. As is evident, the annual variation in the house price measurement error does not coincide with the temporal variation in the default option beta as shown in Figure 1.

3.3 State Default Option Beta Panel

We further evaluate spatial heterogeneity in the default option beta time series across select states. To do so, we stratify the sample by state and estimate the rolling window model. To obtain a better picture of the spatial heterogeneity in the state-specific default option beta estimates, we plot the beta time-series for six states – including California, Florida, Georgia, Illinois, New York, and Texas – in Figure 4. As is evident, all sampled states display significant cyclical movement in the default option beta over the boom, bust, and crisis aftermath. For example, California, Florida, Georgia, and Illinois demonstrate two successive periods of upward movement in the default option beta between 2006 and 2012 prior to more recent trending down in those series. While all sampled states exhibited a peak in the default option beta series in 2012, the amplitude of those cycles varied across states. Indeed, the run-up over the crisis period and its aftermath in beta was substantially damped in Texas and New York relative to elevated levels computed for Florida.

4. What Drives Variations in Default Propensities?

[20] In this analysis, only the top 50 zip codes ranked by number of housing transactions are included.

4.1. A Theoretical Framework

As evidenced above, variations in the negative equity beta are sizable both in the time-series and in the cross-section. Below, we explore some explanations of these variations. We start with a simple theoretical framework to inform the empirical analysis.

The mortgage termination literature emanates from an option-based contingent claims framework whereby mortgage default and prepayment are options to put and call the contract, respectively (see, e.g., Kau et al., 1992; Schwartz and Torous, 1992; Ambrose, Buttimer and Capone, 1997). Recent literature has extended early papers in the context of a more general household utility/wealth maximization framework. In the broader model, mortgage borrowers exercise the default option to maximize utility/wealth, subject to liquidity constraints and other exogenous shocks (see, e.g., Campbell and Cocco, 2015; Corbae and Quintin, 2015).

As in the literature, we characterize mortgage loans as debt contracts with a compound default (put) option, such that a borrower who does not default in a given period has the right to default in the future. Consider a mortgage borrower who faces a decision at time *t* of whether to continue to make the mortgage payment or to default on the loan. Assume the property value is H_t and the remaining mortgage balance is M_t (negative equity is thus $H_t - M_t$). Default eliminates borrower's negative equity.

Building on Riddiough and Wyatt (1994b) and others, we allow for the possibility of a loan workout in the wake of default. Accordingly, if the borrower chooses to default, there are two possible outcomes, including foreclosure with probability p_t, and workout with probability $(1 - p_t)$. If foreclosed, the borrower incurs tangible transaction costs R_t, which include moving costs and credit impairment (Cunningham and Hendershott, 1984). There are also intangible foreclosure transaction costs S_t, which include stigma effects and possible psychic costs (Kau and Keenan, 1995; White, 2010). If instead, the bank agrees to work out the loan, the borrower will receive a benefit of V_t in terms of payment reduction (reduced interest rate, term extension, and the like) and/or write-off of some portion of principal balance.

Let B_t denote the benefit to the borrower of default. Then

$$B_t = p_t[-(H_t - M_t) - R_t - S_t - (1 + r_t)^{-1} E_t B_{t+1}] + (1 - p_t)V_t,$$

$$\text{where} \quad B_{t+1} = p_{t+1}[-(H_{t+1} - M_{t+1}) \cdots] \cdots \tag{3}$$

Equation (3) shows that the default benefit consists of two parts: the first part is net benefit from possible foreclosure, including the extinguishment of negative equity ($H_t - M_t$), incurrence of transaction costs ($R_t + S_t$), and loss of the option to default in the next period with a value of $E_t B_{t+1}$ discounted back to the current period with a discount rate r_t; and the second part is the net benefit of possible work out, V_t. The total benefit is just a weighted average of these two parts.

Upon loan maturity at time T, the net benefit becomes

$$B_T = p_T[-(H_T - M_T) - R_T - S_T] + (1 - p_T)V_T, \tag{4}$$

as there's no remaining next period default option.

It has long been recognized that certain exogenous shocks, such as loss of job could trigger default. Vandell and Thibodeau (1985) describe such an outcome as suboptimal default, whereas Campell and Cocco (2015) and Corbae and Quintin (2015) model default resulting from income shocks in the context of a utility/wealth maximization problem. More generally, such trigger events may be described in terms of borrower budget constraints. For the borrower to be able to continue making monthly payments, her income must be adequate to cover her mortgage payment, other debt payments, and consumption,

$$Y_t \geq P_t + D_t + C_t, \tag{5}$$

where Y_t denotes the borrower's income, P_t is the mortgage payment, D_t is other debt payment and C_t is consumption.

There is the possibility of borrower insolvency such that her income falls short of required debt payments and consumption. In such circumstances, the borrower can sell the property to pay off the loan and thus avoid default. However, there may be substantial transaction costs associated with a fire sale of the property, including real estate agent commissions and psychic distress. Alternatively, the borrower can choose to default to avoid such transaction costs. We denote such transaction costs as W_t. Further we denote the probability that the borrower becomes insolvent as q_t. The ultimate benefit of default to the borrower at decision point t is then

$$G_t = (1 - q_t)B_t + q_t(W_t + B_t) = B_t + q_t W_t. \quad (6)$$

The default condition is $G_t \geq 0$.

Model solution requires information about the full dynamics of house prices, mortgage interest rates, transaction costs, borrower income, other debt payments, consumption, the conditional probability of foreclosure given loan default, and benefits of a loan workout. While a closed-form solution is unlikely, we are able to make some inferences that inform the empirical analysis.

First, consider the probability of default. Per equation (3), a borrower benefit from default is the extinguishment of negative equity $(H_T - M_T)$. The probability of default then varies positively with that term. The probability of default also varies with the borrower's expectation of house prices and interest rates over the life of the loan, reflected in the B_{t+1} term. Finally, default probability is a function of transaction costs, borrower assessment of the likelihood of receiving a workout and magnitude of workout benefit, and borrower probability of insolvency.

Further, per above, the sensitivity of default probability to negative equity, which is the first-order partial derivative of default probability with respect to negative equity, should be a function of the borrower's expected conditional probability of foreclosure p_t. It should also be a function of borrower expectations of future house prices and mortgage interest rates.[21] This is because B_t depends on $E_t B_{t+1}$, which varies with current H_t as well as expected changes in house prices and mortgage interest rates.[22]

To summarize, the above model suggests that negative equity is a key driver of loan default. Further, as suggested above, the borrower's sensitivity to negative equity can vary with changing borrower expectations, the conditional probability of foreclosure (or workout), and other factors.

4.2. Panel data regression of state-level default option beta

In this section, informed by the above theoretical framework, we study underlying factors that drive variation in the estimated default option betas. Recall that our rolling window hazard model estimates yield a panel of default option betas by state and by quarter. As discussed above, we hypothesize that potential drivers of the default option beta include such factors as borrower changing market expectations, the future path of house prices, and the conditional probability of foreclosure.

We proxy for borrower expectations using measures of the local business cycle and consumer sentiment. Both terms are available at the state level. Following Korniotis and Kumar (2013), we use unemployment rate innovation as a measure of the local business cycle. It is computed using the BLS current quarter unemployment rate divided by its four-quarter moving average. Also, borrowers might use past evidence of house price appreciation to gauge future returns. For this reason, we consider a lagged house price return term based on the CoreLogic

[21] Here we assume negative equity is independent of borrower insolvency probability, q_t, and transaction costs (a combination of R_t, S_t and W_t).
[22] More formally if we assume house price follows a geometric Brownian motion with time varying drift, such a relation will be obvious from the first-order derivative calculation.

house price index.

We use a consumer distress index to proxy sentiment. The index comes from CredAbility Nonprofit Credit Counseling & Education, and we retrieve it from the Federal Reserve Economic Data (FRED), Federal Reserve Bank of St. Louis. It is a quarterly comprehensive measure of the average American household's financial condition. CredAbility uses more than 65 variables from government, public and private sources to convert a complex set of factors into a single index of consumer distress. Given that this distress index in part reflects economic fundamentals, which might be already reflected by unemployment rate innovation, we first regress the state-level CredAbility consumer distress index on state-level unemployment rate innovations as well as time- and state-level fixed effects to obtain a distress index orthogonalized to fundamentals. We then use the orthogonalized distress index in our analysis.

There is no consensus on how to measure borrowers' subjective assessment of the likelihood of loan modification (vs foreclosure) conditional on default. Our approach is to test for structural breaks in default option exercise coincident to enactment of major crisis-period loan modification programs, as existing literature suggests elevated borrower strategic default in the wake of such loan modification programs (see, e.g., Mayer et al., 2014).

Note that our theory suggests that while borrower income shocks are an important driver of default probability, they should not directly affect the default option beta. However, to account for the possibility that our first-stage hazard model does not fully control for this factor, we include average gross income growth from the IRS in our panel data regression as well.

We also include categorical controls for recourse states and sand states. In recourse states, lenders may pursue deficiency judgments against borrowers to the extent foreclosure proceeds fail to fully compensate for losses. It is hypothesized that even the legal threat of recourse to borrower non-housing assets may change borrower default behavior. Ghent and Kudlyak (2011) find that borrowers in non-recourse states are more sensitive to negative equity. As regards sand states, that nomenclature was adopted during the late 2000s to identify those states that experienced a marked boom-bust cycle in housing and were at the epicenter of the crisis.

We present the results of our panel data regression in Table 4. The dependent variable is the by-quarter estimate of the default option beta from the hazard model of default for each state in rolling windows (hence a panel of betas). In model 1, we include among explanatory terms the state unemployment rate innovation, the orthogonalized state consumer distress index and a time dummy. State unemployment rate innovation is positive and significant, indicating an elevated default option beta in the context of a weaker local economy. The orthogonalized MSA consumer distress index is negative and significant, suggesting elevated default option exercise in the context of higher levels of consumer distress. The time dummy is positive and significant, indicating an elevated default option beta post 2009Q3. We tested a number of other breaking points, but find post-2009Q3 provides the best fit of the data. Later in the paper we test whether this result is related to the borrower's changing view of the likelihood of receiving a loan workout in the wake of the enactment of a major mortgage modification program. Finally, in our sample, the categorical term identifying recourse states is not a significant factor in determination of the default option beta. The four variables combined explain about 68 percent of the variations in default option beta.

In model 2, we substitute a categorical control for sand states in place of that for recourse states. As expected, after controlling for state unemployment conditions, the orthogonalized state distress index, and the post-2009 policy treatment, findings indicate a significantly elevated default option beta among hard-hit sand state areas. Further, other results are robust to the inclusion of that term. In model 3, we include a full vector of state fixed effects instead of the recourse or sand state categorical terms. As is evident, findings are robust to the inclusion of state fixed effects; further, that specification explains a full three-fourths of the variation in the default option beta.

In models 4 -6, we replace the state unemployment rate innovation and orthogonalized state distress index terms with proxies for house price expectations and income shocks. We calculate house price return based on the CoreLogic house price index data and use its lagged term as a proxy for house price expectations. Borrower income shocks are approximated by the change in IRS zip code-level average adjusted gross income aggregated to the state level. Model 4 is identical to model 1 except for that substitution. Results of models 4 - 6 show that lagged HPI return is significant and negative in explanation of the default option beta.[23] To the extent lagged HPI return is a measure of borrower expectations, this result suggests that the default option betas are damped in the context of elevated expectations of house price returns. Consistent with our theory, while the income shock is a positive and significant factor in the first-stage hazard model for default probability, that same factor is insignificant in determination of the default option beta. In other words, borrower insolvency probability is a determinant of default probability but not necessarily an important factor in explaining borrower default propensities.[24] Findings associated with the estimated recourse and sand states categorical terms (see models 4 and 5, respectively) are consistent with those described above and robust to this specification.

Finally, in model 7, we include all five factors as well as the sand state categorical term and in model 6 we substitute state fixed effects for the sand state control. The results of those full specifications are consistent with those of the above models. In sum, results of panel data analysis are consistent with theory and show that those proxies for local business cycle, sentiment, and

[23] Here HPI return is calculated as $\log(HPI_t/HPI_{t-4})$, and we use one-quarter lag.
[24] Certainly, due to data limitations, our measure of income shock is not perfect. If exact information about borrower-level income and its change becomes available in the future, one can further test this hypothesis.

house price expectations capture some of the variances in the default option beta. Further, as evidenced, default propensities as embodied in the default option beta are elevated among sand states. Model 8 includes state fixed effects and provides consistent results regarding the impact of local business cycle, sentiment, and house price expectations. While the model explains over 70 percent of the variation in the default option beta, we do not rule out other possible explanations such as the increased social acceptance of default option exercise (see, e.g., Guiso, Sapienza and Zingales, 2013).

4.3 Hazard model with interaction terms

The literature on varying coefficient models suggests that if we know the determinants of time variation in the default option beta, we can simply include interaction terms between the covariate and those factors and estimate the model in linear form (see, Cai et al., 2008). In this case, the model becomes

$$h_i^k(T, Z_{i,t}^{'}) = h_0^k(T)\exp[a(t)Z_{i,t}^{'}\beta] \tag{7}$$

Here $a(t)$ is the time series factor that determines the time-varying coefficient. As the focus of this paper is the time-varying coefficient of the default option, we hold constant the coefficients of the other covariates in our interaction model.[25] As such, we have

$$a(t)Z_{i,t}^{'}\beta = \beta^1 u_t x_{i,t} + W_{i,t}^{'}\gamma, \tag{8}$$

where we decompose $Z_{i,t}$ into the default option $x_{i,t}$ and the other covariates $W_{i,t}$. Here β^1 measures how the sensitivity of borrower default to default option value varies with time series factors u_t, which include business cycle, sentiment and other indicators that we discuss in the next

[25] We conduct some tests whereby we relax this assumption and allow all the covariates to vary over time. Significant variation is evidenced only in the case of the default option.

section.

We now turn to the estimation of the competing risks proportional hazard model with interaction terms. In contrast to the 3-year moving window estimates displayed in Figure 1, here we pool all observations in estimation of the hazard model. We focus on the hypothesized drivers of the default option beta explored in section 4.2, namely unemployment rate innovations, orthogonalized MSA consumer distress index, and a time dummy for a possible structural break in 2009.

Model estimates for the default equation are reported in Table 5. While the regressions include a large number of loans, borrower, and locational controls, we focus on the table on the interaction terms. In the first column, results are based on the full sample. As is consistent with results in the panel data model, the estimated default option beta is higher for states and time-periods with higher unemployment rate innovations. In other words, borrower sensitivity to negative equity varies with the economic cycle – borrowers are more sensitive to negative equity and are more likely to pull the trigger on default in bad times.

Further, findings indicate that innovations in the unemployment rate are themselves positively associated with default probability. As is also consistent with results of panel estimation, lower levels of orthogonalized state-level consumer sentiment are associated with higher likelihoods of loan default. We similarly find evidence of a structural break in default likelihood and behavior in 2009Q3. All things equal, borrowers are more likely to default after the third quarter of 2009; further, borrowers become more sensitive to negative equity at that time.[26] As discussed below, that timing is coincident to implementation of a major loan modification program (HAMP) that likely affected borrower priors regarding receipt of a favorable loan modification

[26] We use the Wald test discussed in Andrews (1993) and test a number of alternative dates for the structural break and find 2009Q3 is the most significant structural break point.

conditional on loan default.

As discussed above, recent research has underscored the importance of crisis-period income and liquidity shocks as a default trigger. One may inquire as to whether residual income and liquidity effects not controlled for by our income change proxies bias our results. While our panel beta regression results show that is not likely the case, as the income change term is not significant in the regression, we conduct further analysis below to show that variations in the estimated default option beta are not explained by the residual income effect. First, we stratify the sample based on the borrower debt-to-income ratio and re-estimate the model using the bottom quartile of borrowers with debt-to-income ratios below 29 percent. We hypothesize that those borrowers are least likely to have liquidity issues and hence are less sensitive to income shocks. Results in the middle column of Table 5 show that even among the borrowers who are least likely to be liquidity constrained, there remain significant variations in the default option beta with unemployment rate innovations, orthogonalized state consumer distress index and the 2009Q3 time dummy.

As shown in Table 5, we also assess the robustness of results among loan samples sorted by neighborhood income growth. The positive income growth subsample includes loans in zip codes experiencing positive income growth. The sorting of loans is dynamic so that the same loan can fall into different categories based on current income growth in the zip code. We hypothesize that liquidity constraints should be least binding in neighborhoods with positive income growth. Results confirm the robustness of drivers of the variation in negative equity even among the neighborhoods with positive income growth.

We conduct a series of additional robustness checks. In so doing, we augment our model specification to assess the effects of a “woodhead”[27] measure (missed default opportunities). Results in Appendix Tables 2 show our findings regarding drivers of beta changes are highly robust to that specification. Also, we estimate the model using annual cohorts. This test addresses the concern that a changing mix of borrowers might have contributed to the observed changes in the default option beta, even after controlling for a large set of borrower characteristics. As displayed in Appendix Table 3, the estimated default option betas are robust to the cohort specification, so as to underscore the primary findings of the paper.

Finally, to assure our results are not merely driven by a specific sample of mortgage loans, we also re-run our analysis using alternative loan samples. Specifically, we re-estimate our models using a sample of private-label securitized mortgage loans from BBX. In that exercise, we run separate models by loan type for subprime, Alt-A and prime jumbo loans as well as a model with all loans pooled. We find consistent results[28].

We further conduct decomposition analysis to assess the economic significance of factors identified in the panel model. We do so by taking the aforementioned hazard model results and simulating the impact of each factor on the default option beta and default probability. The results are presented in Figure 5. The baseline results (blue line) show the impact of negative equity on the hazard rate of default in a benign economic environment. Consistent with Figure 2, the default hazard rate increases with negative equity, but the marginal impact is modest. Moving on to the red line, here we assume a recession environment indicted by sharp increase in unemployment rate. As is evident, the sensitivity of borrowers to negative equity increases significantly. Finally, we incrementally add the sentiment factor (green line) and the 2009 structural break (purple line). Overall, among primary drivers, local business cycle and consumer sentiment were each associated

[27] See Deng and Quigley (2001) for a discussion.
[28] Some of the results were shown in an earlier version of our paper, while others are available upon request.

with roughly 30 to 40 percent of the increase in default risk due to their impact on the default option beta, while the 2009 structural break contributed the remaining 20 to 30 percent, depending on the magnitude of borrower negative equity.

4.4. HAMP Program Effects

In the wake of the housing crisis, numerous government mortgage modification programs were enacted with the aim of mitigating home foreclosure. Among the most notable was the federal Home Affordable Modification Program (HAMP), which was implemented in the first quarter of 2009. The HAMP program used federal subsidies to incentivize lenders to modify loans rather than foreclose on defaulted borrowers. In the spirit of the "Lucas Critique", we suspect that enactment of a major foreclosure abeyance program may have influenced the default behavior of mortgage borrowers, e.g., borrowers may have become more likely to default to the extent a loan modification was forthcoming.

The existing literature provides ample evidence on strategic default. Riddiough and Wyatt (1994) and Guiso, Sapienza and Zingales (2013) argue that a borrower's delinquency decision may depend on the anticipated lender response (for example, the likelihood of foreclosure conditional on delinquency). Mayer et al. (2014) provide evidence of increased borrower willingness to strategically default in response to a lender loan modification program. As discussed above, in Table 5 we report on estimation of elevated default probabilities post-2009Q3. The structural break coincides with the timing of HAMP implementation.[29] Further, results show a sizable and significantly elevated default option beta for the post-2009 period. Below we report on related corroborating difference-in-differences analysis.

---

[29] While implementation of HAMP commenced in 2019Q1, substantial modification volume dates only from late 2009 and early 2010.

For a loan to qualify for modification under the HAMP program, a number of criteria must be met. First, only owner-occupied loans were eligible for modification under HAMP. Second, the loan must have been originated prior to January 2009. Third, the remaining balance on the loan must be less than $729,500. Fourth, the borrower's debt-to-income ratio at the time of modification was required to be in excess of 31 percent as the intent of the modification was to reduce borrowers' monthly housing payments to no more than 31 percent of gross monthly income. Finally, there was a HAMP implementation window, which originally was set to be from March 2009 to December 2012 but later was extended through 2016. We utilize the above eligibility rules to conduct difference-in-differences (DID) analysis of changes in borrower default option exercise in the wake of the enactment of the HAMP program. Agarwal et al (2017) use this strategy to identify the impact of HAMP on loan renegotiations.[30]

Similar to Agarwal et al (2017), our DID control group is comprised of investor property loans that did not qualify for modification under HAMP whereas our treatment group includes owner-occupied loans that may be qualified for HAMP pending other conditions. We use the 2009Q3 as the treatment date. To avoid confounding effects and consistent with HAMP program terms, we limit the sample to loans with a remaining balance below the HAMP threshold of $729,500. For similar reasons, we also exclude loans with a debt-to-income ratio below 44 percent.[31] All of our loans were originated prior to January 2009. Note that our DID test does not require a perfect identification of HAMP eligible loans or loans eventually modified via HAMP.[32] As long as one group of borrowers had a higher probability of receiving a HAMP modification

[30] In contrast to Agarwal et al (2017) our analysis focuses on borrower delinquency rather than loan modification.
[31] We do not have information on the front-end (payment-to-income) ratio in the GSE data. However, we use a 44 percent back-end ratio cutoff is to ensure that the loans included in the analysis are all HAMP eligible.
[32] Not all HAMP applications that met those five criteria were approved and some fell out of the program after the trial period.

than the other group based on *ex ante* borrower expectations, we should be able to identify HAMP effects via our DID test.

Given well-known challenges in applying DID framework in the context of non-linear models such as the Cox hazard model (Ai and Norton, 2003 and Karaca-Mandic, Norton and Dowd, 2012), we instead conduct our DID analysis using a generalized least squares estimation of a linear default model. Table 6 presents our DID regression results. The DID regression takes the form

$$Y = (\beta_1 T + \beta_2 T * After + \beta_3 After)x + Z'\gamma + \varepsilon, \qquad (9)$$

where T represents the treatment group, After represents the period after which the policy was implemented, and the Z vector represents a vector of control variables. The dependent variable Y takes value of "1" if a loan defaults in a particular quarter and "0" otherwise. Note first in Appendix Figure 2 the parallel trends exhibited in the default option beta time-series among the treated (owner-occupied) and control (investor) loans pre-treatment. However, as shown in Table 6, post-2009Q3, the treated owner-occupied loans exhibit a statistically elevated default option beta. These findings are consistent with the hypothesis that the federal program may have inadvertently resulted in elevated default propensities among borrowers in that group. The time window of our loan performance records is 2007Q3 – 2011Q3. In an alternative specification, we conduct a DID analysis where we utilize a narrower 2008-2010 version of the test window. The alternative specification yields similar results (see column 2 of Table 6).

We further conduct a number of placebo tests of our difference-in-differences test. As shown in Table 7, we first run the linear default model with a random breakpoint (2008Q3) where there is no policy change so as to evaluate whether the DID regression results might simply reflect uncontrolled differences between our control and treatment groups. In the second placebo test, both the "treatment" and control groups are loans with DTI below 29 percent and thus are both

HAMP ineligible. Results in Table 7 indicate a lack of significance associated with the treatment group beta in either placebo test.

We acknowledge that it is challenging to ascertain the exact impact of HAMP as we do not have a perfect counter-factual. However, our aforementioned results suggest that such a nation-wide program coupled with intense media coverage of default and default assistance programs could have affected borrower behavior, as we argue in this paper.

5. Conclusion

In the wake of the late-2000s implosion in house values, mortgage default skyrocketed. While crisis period default commonly has been ascribed to the sizable run-up in borrower negative equity, we show those loan terminations also were importantly precipitated by elevated default option exercise. Results of time-varying coefficient hazard model estimation indicate that for a given value of the mortgage default option, borrower propensity to default rose markedly during the period of the financial crisis, especially in hard-hit states. Panel data analysis indicates that much of the variation in default option exercise can be explained by the local business cycle, consumer distress, and federal policy intervention.

Our findings have implications for mortgage underwriting and pricing. From the perspective of credit risk management, results underscore the importance of model instability and the appropriateness of time-varying coefficient models. Our study also provides guidance on factors governing cross-section and time-series variation in estimated default option betas. Mortgage originators, investors, and regulators need to account for such shifts in predicted default behavior in their business planning and practice.

Our findings also have implications for macroprudential policy. In that regard, there has been substantial debate on whether government should bail out borrowers via mortgage modification. Arguments against such programs point to borrower moral hazard, whereby anticipated bailout of distressed borrowers may encourage irresponsible financial behavior. Our findings suggest that federal foreclosure prevention and loan work-out programs may have inadvertently incented higher levels of default propensity, in turn suggesting adverse, unintended consequences of policies designed to mitigate mortgage failure.

References

Agarwal, Sumit, Gene Amromin, Itzhak Ben-David, Souphala Chomsisengphet, and Douglas D. Evanoff. 2014. Predatory Lending and the Subprime Crisis. *Journal of Financial Economics*, 113(1), 29-52.

Agarwal, Sumit, Gene Amromin, Itzhak Ben-David, Souphala Chomsisengphet, Tomasz Piskorski, and Amit Seru. 2017. Policy Intervention in Debt Renegotiation: Evidence from the Home Affordable Modification Program. *Journal of Political Economy* 125(3), 654-712.

Ai C. and E. C. Norton. 2003. Interaction Terms in Logit and Probit Models. *Economics Letters*, 80(1), 123–9.

Ambrose, Brent W., Richard J. Buttimer, Jr., and Charles A. Capone. 1997. Pricing Mortgage Default and Foreclosure Delay. *Journal of Money, Credit and Banking*, 29(3), 314-325.

An, Xudong, Yongheng Deng, and Stuart A. Gabriel. 2011. Asymmetric Information, Adverse Selection and the Pricing of CMBS. *Journal of Financial Economics*, 100(2), 304-325.

An, Xudong, Yongheng Deng, Eric Rosenblatt, and Vincent W. Yao. 2012. Model Stability and the Subprime Mortgage Crisis. *Journal of Real Estate Finance and Economics*, 45(3), 545-568.

Andrews, D. 1993. Tests for Parameter Instability and Structural Change with Unknown Change Point. *Econometrica* 61 (4): 821–856.

Brueckner, Jan K, Paul S. Calem, and Leonard I. Nakamura. 2012. Subprime Mortgages and the Housing Bubble. *Journal of Urban Economics*, 71(2), 230-243.

Bhutta, Neil, Jane Dokko, and Hui Shan. 2017. Consumer Ruthlessness and Mortgage Default during the 2007 to 2009 Housing Bust. *Journal of Finance*, forthcoming.

Cai, J., J. Fan, J. Jiang, and H. Zhou. 2008. Partially Linear Hazard Regression with Varying-Coefficients for Multivariate Survival Data. *Journal of Royal Statistical Society B*, 70, 141-158.

Campbell, J.Y. and J.F. Cocco. 2015. A Model of Mortgage Default. *Journal of Finance*, 70(4): 1495-1554.

Cleveland, W. S., Grosse, E., and Shyu, W. M. 1991. Local regression models. In *Statistical Models in S* (J. M. Chambers, and T. J. Hastie, eds.) 309–376. Wadsworth / Brooks-Cole, Pacific Grove, CA.

Corbae, Dean, and Erwan Quintin. 2015. Leverage and the Foreclosure Crisis. *Journal of Political Economy*, 123, 1-65.

Cotter, John, Stuart Gabriel, and Richard Roll. 2015. Can Metropolitan Risk be Diversified? A Cautionary Tale of the Housing Boom and Bust. *Review of Financial Studies*, 28(3), 913-936.

Demyanyk, Y., and O. Van Hemert. 2011. Understanding the subprime mortgage crisis. *Review of Financial Studies*, 24(6), 1848-1880.

Demyanyk, Y., and E. Loutskina. 2016. Mortgage Companies and Regulatory Arbitrage. *Journal of Financial Economics* 122(2): 328–351.

Deng, Yongheng. 1997. Mortgage Termination: An Empirical Hazard Model with Stochastic Term Structure. *Journal of Real Estate Finance and Economics*, 14 (3), 309-331.

Deng, Yongheng, and Stuart A. Gabriel. 2006. Risk-based Pricing and the Enhancement of Mortgage Credit Availability among Underserved and Higher Credit-Risk Populations. *Journal of Money, Credit and Banking*, 1431-1460.

Deng, Yongheng, and John M. Quigley. 2001. Woodhead Behavior and the Pricing of Residential Mortgages. *SSRN Working paper 288694*.

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

Duffie, Darrell, Andreas Eckner, Guillaume Horel, and Leandro Saita. 2009. Frailty Correlated Default. *Journal of Finance*, 64, 2089-2124.

Elul, Ronel, Nicholas S. Souleles, Souphala Chomsisengphet, Dennis Glennon, and Robert Hunt. 2010. What "Triggers" Mortgage Default? *American Economic Review*, 100(2), 490-94.

Fan, Jianqing, and Wenyang Zhang. 1999. Statistical Estimation in Varying Coefficient Models. *Annals of Statistics*, 27(5), 1491-1518.

Fan, Jianqing, and Wenyang Zhang. 2008. Statistical Methods with Varying Coefficient Models. *Statistics and Its Inference*, 1, 179-195.

Foote, Chris, Kristopher S. Gerardi, and Paul S. Willen. 2008. Negative Equity and Foreclosure: Theory and Evidence. *Journal of Urban Economics*, 64(2), 234–245.

Gerardi, Kristopher, Kyle Herkenhoff, Lee O'Hanian, and Paul Willen. 2015. Can't Pay or Won't Pay? Unemployment, Negative Equity, and Strategic Default. FRB Atlanta Working Paper.

Ghent, Andra C., and Kudlyak, Marianna. 2011. Recourse and Residential Mortgage Default: Evidence from U.S. States. *Review of Financial Studies*, 24(9), 3139-3186.

Guiso, L., Sapienza, P., and Zingales, L. 2013. The determinants of attitudes toward strategic default on mortgages. *Journal of Finance*, 68(4), 1473-1515.

Gyourko, J., and J. Tracy. 2014. Reconciling theory and empirics on the role of unemployment in mortgage default. *Journal of Urban Economics* 80, 87–96.

Karaca-Mandic, P., E. C. Norton, and B. Dowd. 2012. Interaction Terms in Non-Linear Models. *Health Service Research*, 47(1), 255-274.

Kau, J. B., D. C. Keenan, W. Mueller III, and J. F. Epperson. 1992. A Generalized Valuation Model for Fixed-Rate Residential Mortgages. *Journal of Money, Credit, and Banking*, 24(3), 279-99.

Keys, Benjamin, Tanmoy Mukherjee, Amit Seru, and Vikrant Vig. 2010. Did Securitization Lead to Lax Screening? Evidence from Subprime Loans. *Quarterly Journal of Economics*, 125(1), 307-362.

Korniotis, G. M., and A. Kumar. 2013. State-Level Business Cycles and Local Return Predictability. *Journal of Finance* 68(3), 1037-1096.

Low, D. 2015. Mortgage Default with Positive Equity. NYU Working Paper.

Mayer, Christopher, Karen Pence, and Shane M. Sherlund. 2009. The Rise in Mortgage Defaults. *Journal of Economic Perspectives*, 23(1), 27-50.

Mayer, C., E. Morrison, T. Piskorski, and A. Gupta. 2014. Mortgage Modification and Strategic Behavior: Evidence from a Legal Settlement with Countrywide. *American Economic Review*, 104(9), 2830-57.

Mian, A., and A., Sufi. 2009. The Consequences of Mortgage Credit Expansion: Evidence from the U.S. Mortgage Default Crisis. *Quarterly Journal of Economics*, 124 (4), 1449-1496.

Mian, Atif, Amir Sufi, and Francesco Trebbi. 2010. The Political Economy of the U.S. Mortgage Default Crisis. *American Economic Review*, 100, 67-98.

Nadauld, Taylor D., and Shane M. Sherlund (2013). "The Impact of Securitization on the Expansion of Subprime Credit," *Journal of Financial Economics*, 107(2), 454-476.

Piskorski, Tomasz, Amit Seru, and James Witkin. 2015. Asset Quality Misrepresentation by Financial Intermediaries: Evidence from RMBS Market. *Journal of Finance*, 70(6), 2635-2678

Piskorski, T., and A. Tchistyi. 2011. Stochastic House Appreciation and Optimal Mortgage Lending. *Review of Financial Studies*, 24, 1407-1446.

Quercia, R. G., and M. A. Stegman. 1992. Residential Mortgage Default: A Review of the Literature. *Journal of Housing Research*, 3, 341-379.

Quigley, J. M., & Van Order, R. 1995. Explicit tests of contingent claims models of mortgage default. *Journal of Real Estate Finance and Economics*, 11(2), 99-117.

Rajan, U., A. Seru, and V. Vig. 2015. The Failure of Models that Predict Failure: Distance, Incentives and Defaults. *Journal of Financial Economics*, 115(2), 237-260.

Riddiough, T. J., and S.B. Wyatt. 1994a. Wimp or Tough Guy: Sequential Default Risk and Signaling with Mortgages. *Journal of Real Estate Finance and Economics*, 9, 299-321.

Riddiough, T. J., and S.B. Wyatt. 1994b. Strategic Default, Workout, and Commercial Mortgage Valuation. *Journal of Real Estate Finance and Economics*, 9(1), 5-22.

Schwartz, Eduardo S., and Walter N. Torous. 1992. Prepayment, Default, and the Valuation of Mortgage Pass-through Securities. *Journal of Business*, 65(2), 221-239.

Vandell, K. D. 1993. Handing over the keys: a perspective on mortgage default research. *Real Estate Economics*, 21(3), 211-246.

Vandell, K. D., and T. Thibodeau. 1985. Estimation of Mortgage Defaults Using Disaggregate Loan History Data. *Journal of the American Real Estate and Urban Economics Association*, 13(3), 292-316.

White, B. T. 2010. Underwater and not walking away: shame, fear, and the social management of the housing crisis. Wake Forest L. Rev., 45, 971.

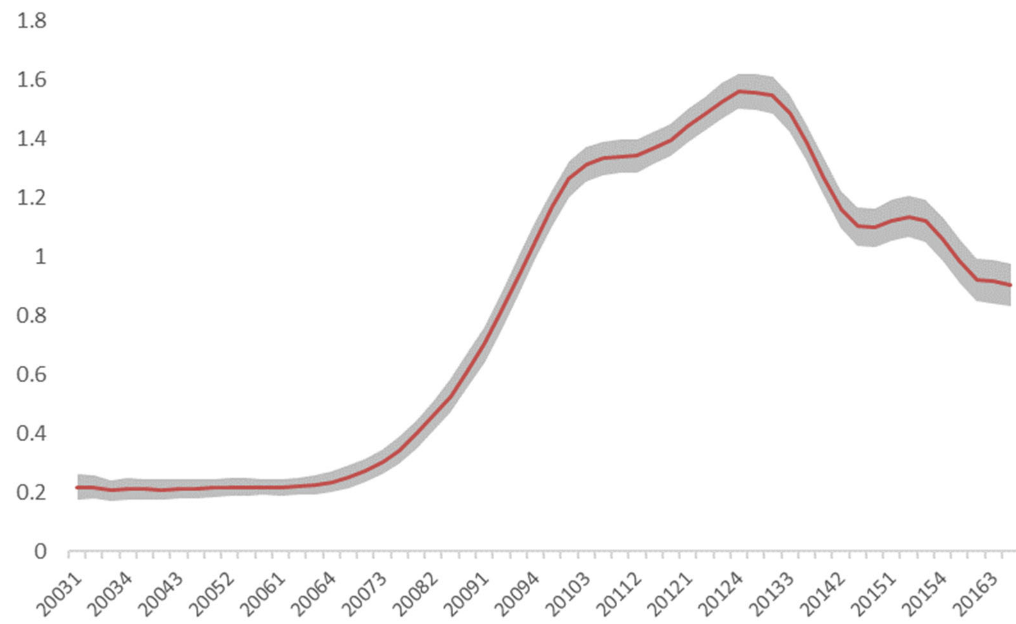


Figure 1 Rolling Window Estimates of the Default Option Beta

Notes: Based on the Freddie Mac Single-Family Loan-Level Dataset and Fannie Mae Single-Family Loan Performance Data (hereafter GSE data). This figure shows the estimates of default option beta in a hazard model. The estimation is based on three-year rolling window samples of first-lien, full-documentation, and fully amortizing 30-year fixed-rate mortgages acquired by the GSEs during 2000 – 2017. A random sample of the GSE loans is used in the estimations. The dark line shows the point estimates and the shaded area shows the confidence interval.

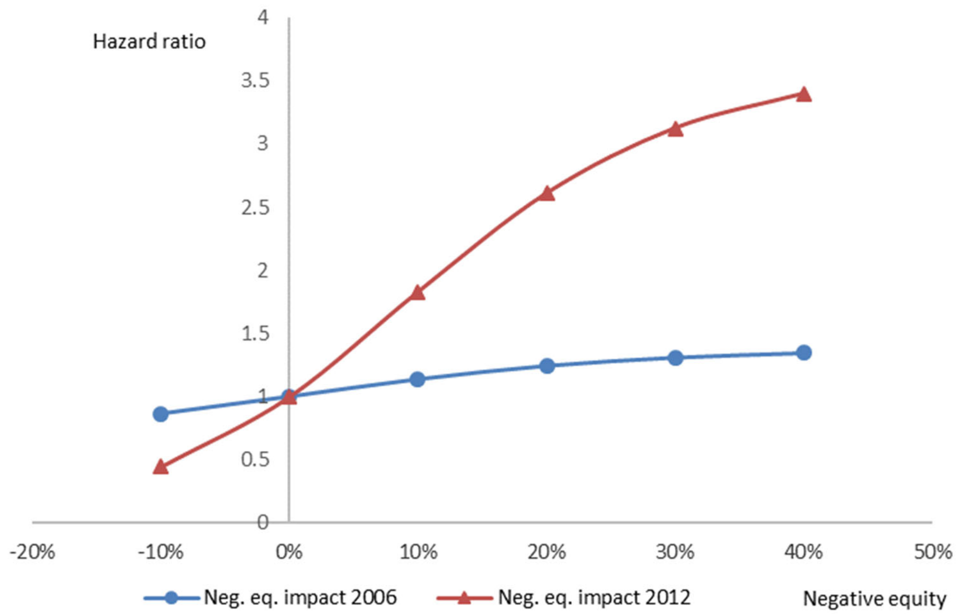


Figure 2 The Impact of Negative Equity on Mortgage Default Probability

Notes: This figure shows the simulated impact of negative equity on default probability in different years. Simulations are based on the default option beta estimates shown in Figure 1.

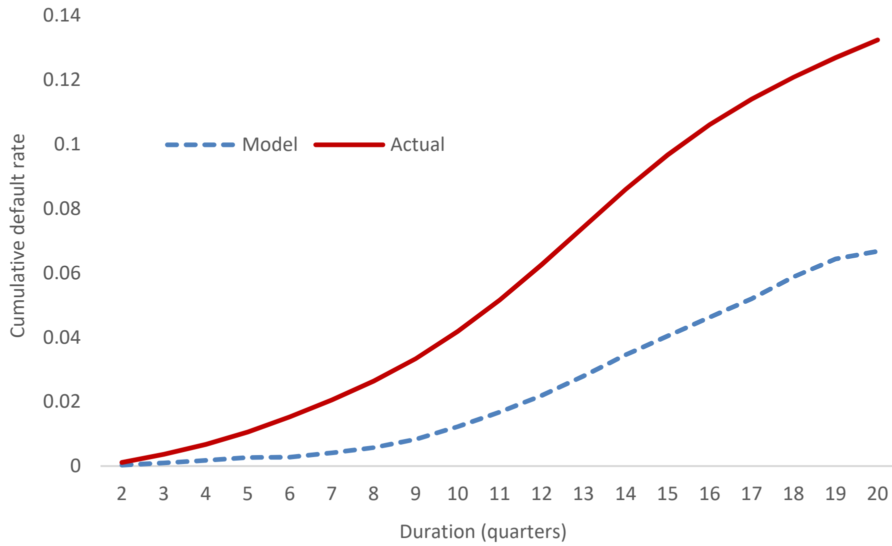


Figure 3 The Impact of Default Option Beta on Default Predictions

Notes: Based on the GSE data. Here we use a model estimated using the 2002-2006 performance record of the 2002-2004 vintages to predict default of the 2006 vintage assuming a perfect foresight of house price movement. The red solid line shows the actual performance of the 2006 vintage (with the actual beta), while the blue dashed line shows the model prediction with the estimated beta using the 2002-2006 performance of the 2002-2004 vintages. Over the 20-quarter horizon, the predicted default rate with the estimated beta is only about half of the actual default rate. As a comparison, default rate of GSE 30-year FRMs during 2006-2010 was about three times as high as that during 2002-2006.

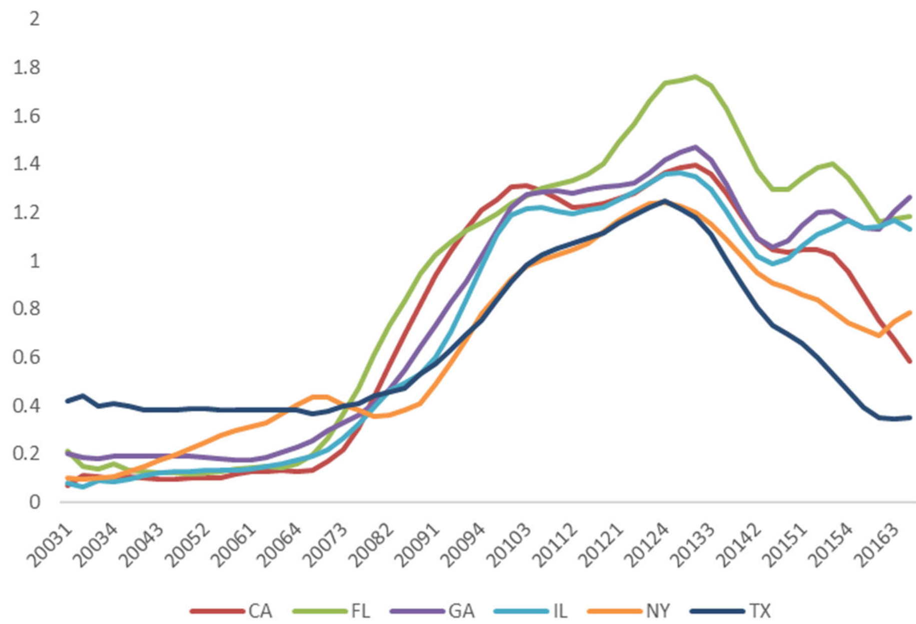


Figure 4 Default Option Beta Time Series for the Selected States

Notes: Based on the GSE data. This figure shows the by-state point estimates of the default option beta based on three-year rolling window samples of loans in the selected states. The estimations are based on the full sample of GSE loans and the betas in different states are estimated separately.

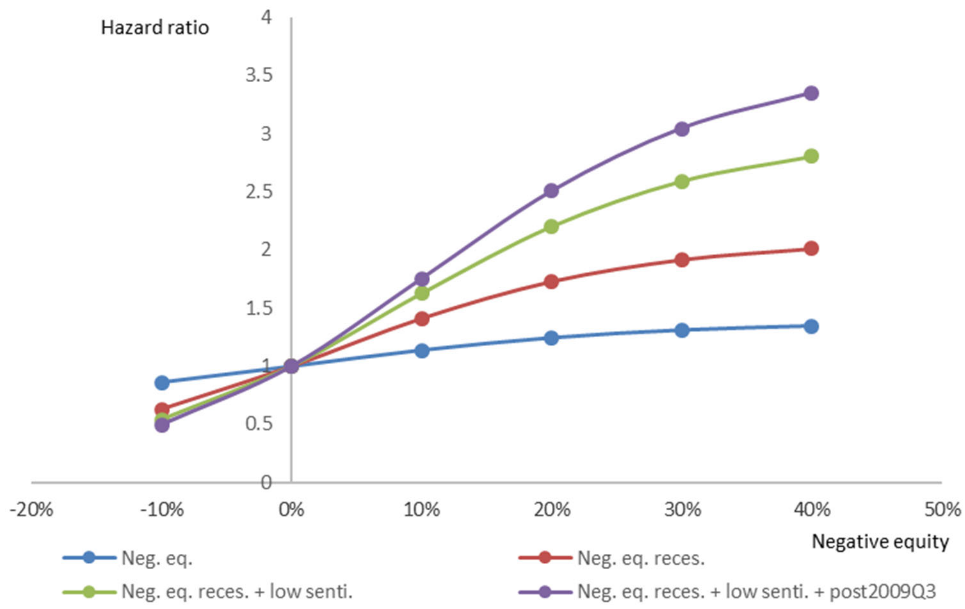


Figure 5 Decomposition of the Various Drivers of Default Option Beta

Notes: This figure shows the simulated impact of negative equity on default probability to illustrate the impact of the various drivers of the default option beta. Simulations are based on the estimates shown in Table 5.

Table 1 Summary Statistics of the Loan Sample

Variable	Mean	STD	Min.	Median	Max.
Original loan amount	199,681	109,602	5,000	176,000	1,470,000
Note rate (%)	5.52	1.26	1.88	5.63	13.50
Loan-to-value ratio (LTV, %)	74	15	25	79	97
Combined LTV (%)	75	15	25	79	200
Borrower credit score	737	54	300	748	850
Debt-to-income ratio (DTI, %)	34	11	9	34	61
First time home buyer	13.2	–	0	–	1
Single family	73.3	–	0	–	1
Condominium	8.6	–	0	–	1
Planned-unit Development	18.1	–	0	–	1
Owner-occupied	89.5	–	0	–	1
Second home	4.1	–	0	–	1
Investment property	6.4	–	0	–	1
Home purchase	44.9	–	0	–	1
Cash out refinance	25.7	–	0	–	1
Rate/term refinance	29.5	–	0	–	1
Originated prior to 2003	23.3	–	0	–	1
Originated 2003-2007	28.2	–	0	–	1
Originated 2008-2012	24.9	–	0	–	1
Originated after 2012	23.6	–	0	–	1
Defaulted	5.95	–	0	–	1
Prepaid	68.02	–	0	–	1
Current	26.03	–	0	–	1
Total number of loans			42,093,277		

Notes: Based on the Freddie Mac Single-Family Loan-Level Dataset and Fannie Mae Single-Family Loan Performance Data (hereafter GSE data). GSE loans included here are first-lien, full-documentation, and fully amortizing 30-year fixed-rate mortgages acquired by the GSEs during 2000 – 2016. We exclude loans with missing or obvious wrong information on loan origination date, original loan balance, borrower credit score, loan-to-value ratio (LTV), or debt-to-income ratio (DTI). The data cutoff date is March 2018. Default is defined as 60- day delinquency. Prepayment refers to early repayment of a loan as a result of borrower move or refinancing for lower interest rates, different loan terms or cash out. Current (censor) means that the loan is performing at date of data cutoff date. For definitions of variables, see http://www.fanniemae.com/portal/funding-the-market/data/loan-performance-data.html and http://www.freddiemac.com/news/finance/sf_loanlevel_dataset.html. The fields and codes have been normalized across the two datasets when possible.

Table 2 Summary Statistics of the Event History Sample

	Defaulted					Prepaid					Current				
Variable	Mean	STD	Min.	Median	Max.	Mean	STD	Min.	Median	Max.	Mean	STD	Min.	Median	Max.
Default option (book)	0.21	1.11	0.00	0.00	1.00	0.05	2.27	0.00	0.00	1.00	0.04	4.64	0.00	0.00	1.00
Default option (market)	0.33	1.31	0.00	0.03	1.00	0.11	3.75	0.00	0.00	1.00	0.07	6.51	0.00	0.00	1.00
Prepayment option	0.11	0.30	-0.12	0.11	0.45	0.10	1.27	-0.12	0.10	0.45	0.05	2.45	-0.13	0.04	0.44
Negative equity (book)	-0.17	0.84	-0.93	-0.20	1.57	-0.33	2.69	-0.91	-0.32	1.23	-0.36	5.95	-0.93	-0.36	1.35
Negative equity (market)	-0.08	1.02	-0.92	-0.13	1.91	-0.26	3.28	-0.90	-0.27	1.53	-0.33	6.84	-0.92	-0.34	1.67
HPA volatility	0.08	0.11	0.00	0.08	0.25	0.07	0.48	0.01	0.07	0.25	0.07	0.99	0.00	0.07	0.25
Change in unemp. rate	2.22	9.87	-17.10	1.90	20.20	0.52	36.91	-14.30	0.40	19.40	0.13	79.41	-17.50	-0.20	20.20
Number of loan quarters	612,490,108														

Notes: Based on the performance history of loans described in Table 1. Negative equity (book) is the difference between the book value of the loan and the market value of the property over the market value of the property. Negative equity (market) is the difference between market value of the loan and market value of the property over the market value of the property. Default option is the cumulative distribution function value of negative equity over house price return volatility. Prepayment option is the difference between market value of the loan and book value of the loan. The market value of the property is calculated based on property value at origination plus change therein, as indicated by a local house price index (HPI). For more details, see Deng, Quigley and Van Order (2000). The book value of the loan is the remaining balance, and market value is calculated as the present value of the remaining mortgage payments using the current prevailing mortgage interest rate as the discount rate. For loans with junior lien(s), book and market values of the loan account for the junior lien(s). Change in unemployment rate is from loan origination to the performance date. HPI is from CoreLogic. The mortgage interest rate is from the St. Louis Fed. The unemployment rate is from Bureau of Labor Statistics (BLS).

Table 3 MLE Estimates of the Baseline Competing Risks Hazard Model

Dependent variable: default/prepay hazard	Default			Prepayment		
	Estimate		S.E.	Estimate		S.E.
Default option (book)	0.747	***	0.015	0.058	***	0.014
Default option squared	-0.091	***	0.005	-0.078	***	0.006
Prepayment option	0.273	***	0.007	0.987	***	0.008
Prepayment option squared	0.083	***	0.003	0.006		0.004
Change in unemp. rate	0.174	***	0.006	-0.185	***	0.006
Change in unemp. rate squared	0.027	***	0.003	0.003		0.003
LTV < 60	-0.385	***	0.014	0.059	***	0.011
LTV 60-70	-0.130	***	0.014	0.017		0.012
LTV 80-90	0.090	***	0.012	-0.039	***	0.010
LTV > 90	0.355	***	0.014	-0.082	***	0.013
Credit score < 580	0.744	***	0.037	-0.035		0.049
Credit score 580-620	0.376	***	0.020	-0.037		0.026
Credit score 660-700	-0.486	***	0.013	0.051	***	0.015
Credit score 700-740	-0.975	***	0.013	0.104	***	0.015
Credit score 740-780	-1.531	***	0.013	0.159	***	0.014
Credit score > 780	-2.062	***	0.015	0.154	***	0.015
DTI < 20	-0.344	***	0.017	-0.009		0.013
DTI 20-30	-0.217	***	0.011	0.008		0.009
DTI 40-50	0.213	***	0.011	-0.027	**	0.010
DTI> 50	0.335	***	0.014	-0.048	***	0.013
Loan amount < 10k	0.199	***	0.015	-0.577	***	0.013
Loan amount 10-15k	-0.007		0.015	-0.282	***	0.012
Loan amount 15-20k	-0.048	**	0.015	-0.117	***	0.012
Loan amount 25-30k	0.030		0.019	0.095	***	0.015
Loan amount 30-35k	0.096	***	0.023	0.104	***	0.018
Loan amount > 35k	0.181	***	0.023	0.314	***	0.016
First time home buyer	-0.063	***	0.015	-0.109	***	0.013
Condominium	-0.109	***	0.016	-0.015		0.013
Planned-unit development	-0.104	***	0.014	0.068	***	0.011
Investment property	0.105	***	0.017	-0.198	***	0.015
Second home	-0.202	***	0.025	-0.108	***	0.018
Cash out refinance	0.429	***	0.012	-0.130	***	0.010
Rate/term refinance	0.256	***	0.012	-0.089	***	0.010
State FE		Yes			Yes	
Vintage FE		Yes			Yes	
Flexible baseline function		Yes			Yes	
N			1,034,009			

Notes: Based on the GSE data. These are MLE estimates of the competing risks hazard model for default and prepayment based on a random sample of the event history data described in Table 2. The hazard model is in the form of $\mathrm{h}_i^k(T, Z'_{i,t}) = \mathrm{h}_0^k(T)\exp(Z'_{i,t}\beta^k)$, where k indicates default or prepayment, T indicates duration time, t indicates calendar time, i indicates individual loan, and $Z'_{i,t}$ are the risk factors reported in this table. The baseline $\mathrm{h}_0^k(T)$ is estimated non-parametrically and not reported here. State- and vintage-fixed effects are not reported here, either, but they are available upon request. Variable definitions are discussed in Tables 1 and 2. Standard errors are clustered at the loan-level. ***, ** and * indicate 0.1%, 1% and 5% significance, respectively.

Table 4 OLS Estimates of the Panel Data Model of the Mortgage Default Option Beta

Dependent variable: default option beta	Model 1		Model 2		Model 3		Model 4		Model 5		Model 6		Model 7	
	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.
Intercept	0.232***	0.013	0.205***	0.009			0.245***	0.008	0.217***	0.009			0.212***	0.009
Unemp. rate innovation	0.062***	0.006	0.061***	0.006	0.062***	0.006							0.032***	0.009
Orth cons distress index	-0.023***	0.006	-0.022***	0.006	-0.013*	0.007							-0.013*	0.006
Post 2009Q3	0.386***	0.014	0.385***	0.013	0.387***	0.012	0.343***	0.014	0.342***	0.013	0.346***	0.012	0.360***	0.014
Lagged HPA							-0.071***	0.007	-0.072***	0.007	-0.064***	0.006	-0.045**	0.010
Income growth							0.010	0.007	0.011	0.007	0.009	0.006	0.011	0.007
Recourse state	-0.019	0.014					-0.018	0.014						
Sand state			0.048***	0.014					0.058***	0.014			0.052***	0.013
State FE	No		No		Yes		No		No		Yes		No	
N	410		410		410		410		410		410		410	
Adjusted R-Squared	0.683		0.691		0.755		0.685		0.696		0.749		0.707	

Notes: OLS estimates of the panel data model of the default beta similar to what is shown in Figure 4. The dependent variable is the default option beta estimate from the hazard model for default (the first stage analysis) for each state in each rolling window (thus a panel of betas). Loans included in the first stage hazard model estimation are GSE loans described in Table 1. Due to the availability of the consumer distress index, our sample cutoff date here is March 2013, and we focus on nine representative states and the District of Columbia: Arizona, California, DC, Florida, Georgia, Illinois, Massachusetts, Michigan, New York, and Texas. Recourse states in our sample include DC, Florida, Georgia, Illinois, Massachusetts, Michigan, New York, and Sand states refer to Arizona, California, and Florida. Unemployment rate innovation is the current quarter unemployment rate divided by its four-quarter moving average and is based on Bureau of Labor Statistics (BLS) data. Consumer distress index is a quarterly comprehensive measure of the average American household's financial condition compiled by CredAbility and made available by St. Louis Fed. Orthogonalized consumer distress index is the residual from a regression where state-level consumer distress index is regressed on the state-level unemployment rate innovation, state-fixed effect and year-fixed effect. In addition, the HPI return is calculated based on the CoreLogic home price index; change in average AGI is based on IRS data. For the structural break, we test a number of breaking points, but find 2009Q3 is the best breaking point based on model fit. Other variable definitions are discussed under Tables 1 and 2. ***, ** and * indicate 0.1%, 1% and 5% significance, respectively.

Table 5 Default Option Exercise and Business Cycle, Sentiment and Structural Break

Dependent variable: default hazard	All loans		Low DTI		Positive Inc. Growth	
	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.
Default option * unemployment rate innovation	0.038***	0.009	0.035**	0.012	0.159***	0.044
Default option * orthogonalized consumer distress	-0.039***	0.009	-0.030*	0.012	-0.008*	0.004
Default option * post 2009Q3	0.106***	0.022	0.136***	0.030	0.133***	0.029
Default option	0.743***	0.026	0.863***	0.038	0.793***	0.076
Default option squared	-0.114***	0.012	-0.123***	0.016	-0.311***	0.042
Unemployment rate innovation	0.300***	0.010	0.261***	0.012	0.405***	0.018
Orthogonalized consumer distress index	-0.180***	0.007	-0.167***	0.008	-0.024***	0.001
Post 2009Q3	0.211***	0.029	0.092*	0.037	0.062***	0.014
Control variables	Prepayment option, prepayment option squared, change in unemployment rate, change in unemployment rate squared, LTV buckets, FICO buckets, DTI buckets, loan amount buckets, fist time home buyer indicator, property type, occupancy type, loan purpose, state FE, vintage FE					
N	397,750		103,992		252,605	

Notes: Based on the GSE data. These are MLE estimates of the competing risks hazard model for default and prepayment based on a random sample of the event history data described in Table 2. Unemployment rate innovation is the current quarter unemployment rate divided by its four-quarter moving average and is based on Bureau of Labor Statistics (BLS) data. Consumer distress index is a quarterly comprehensive measure of the average American household's financial condition compiled by CredAbility and made available by the St. Louis Fed. Orthogonalized consumer distress index is the residual from a regression where state-level consumer distress index is regressed on the state-level unemployment rate innovation, state-fixed effect and year-fixed effect. Due to the availability of the consumer distress index, our sample cutoff date here is March 2013 and we focus on nine representative states and the District of Columbia: Arizona, California, DC, Florida, Georgia, Illinois, Massachusetts, Michigan, New York, and Texas. For the structural break, we test a number of breaking points, but find 2009Q3 is the best breaking point based on model fit. The low DTI subsample is loans with DTI less than 29% (the lower quartile). The positive income growth subsample is loans in zip code that experience positive income growth. The sorting of loans is dynamic, so the same loan can fall into different categories based on the current income growth in the zip code. Income growth is calculated based on IRS AGI data. Other variable definitions are discussed in Tables 1 and 2. Standard errors are clustered at the loan-level. ***, ** and * indicate 0.1%, 1% and 5% significance, respectively.

Table 6 Diff-in-Diff Tests of the HAMP Eligibility Effect

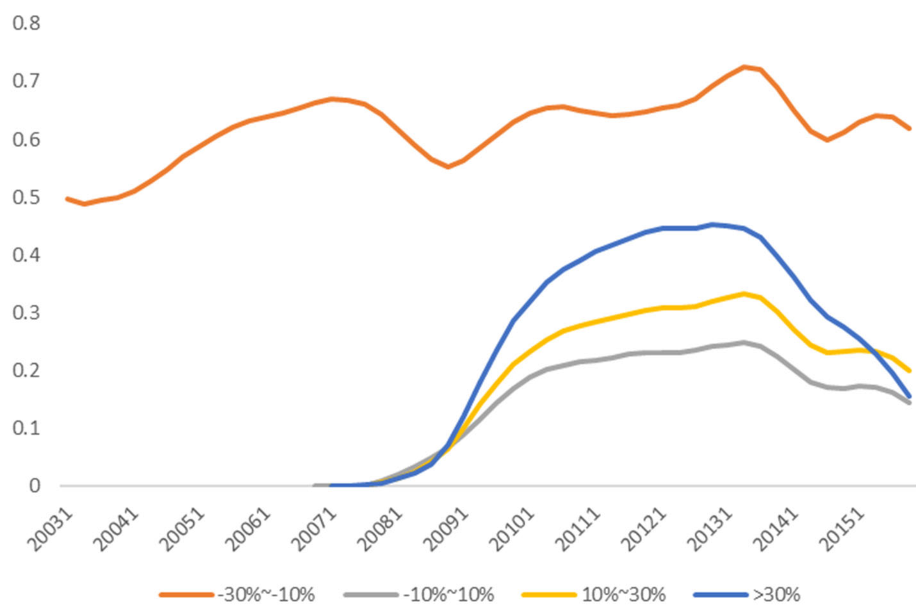
Dependent variable: 0/1 default indicator	2007-2011 data		2008-2010 data	
	Estimate	S.E.	Estimate	S.E.
Treatment group beta	-0.007*	0.003	-0.008	0.005
Treatment group * Post event beta	0.002**	0.001	0.004**	0.002
Post event beta	0.011***	0.003	0.008**	0.003
Control variables	Default option, Default option squared, Default option * unemployment rate innovation, Default option * orthogonalized consumer distress index, unemployment rate innovation, orthogonalized consumer distress index, post 2009Q3, prepayment option, prepayment option squared, change in unemployment rate, change in unemployment rate squared, LTV buckets, FICO buckets, DTI buckets, loan amount buckets, fist time home buyer indicator, property type, occupancy type, loan purpose, state FE, vintage FE			
N	14,226		9,258	

Notes: Based on the GSE data. These are generalized least square (GLS) estimates of a linear default model that uses the difference-in-differences (DID) approach to test the HAMP eligibility effect on borrower default option exercise. The DID test is in the form of $Y \ = (\beta_1 \, T + \beta_2 \, T * After + \beta_3 \, After)x + Z'\gamma$, where T represents the treatment group, $After$ represents the period after which the policy was implemented, and the Z vector represents a vector of control variables described in the table. Loans included in the test are limited to those originated before January 2009 with debt-to-income ratio (DTI) above 44 percent and a remaining balance of no more than $729,500. The HAMP payment-to-income ratio cutoff is 31 percent, but we only observe DTI in our data, so we choose DTI cutoff of 44 percent to ensure the debt service ratio of the selected loans is high enough to meet HAMP requirement. The treatment group is owner-occupied property loans, which satisfy the HAMP occupancy requirement. The control group is investor property loans that are not HAMP eligible. The event date is 2009Q3. In the first test, the time window of our loan performance records is from 2007Q3 to 2011Q3, while in the second test, the time window is from 2008Q3 to 2010Q3. Standard errors are clustered at the loan-level. ***, ** and * indicate 0.1%, 1% and 5% significance, respectively.

Table 7 Placebo Test of the Diff-in-Diff Test of the HAMP Eligibility Effect

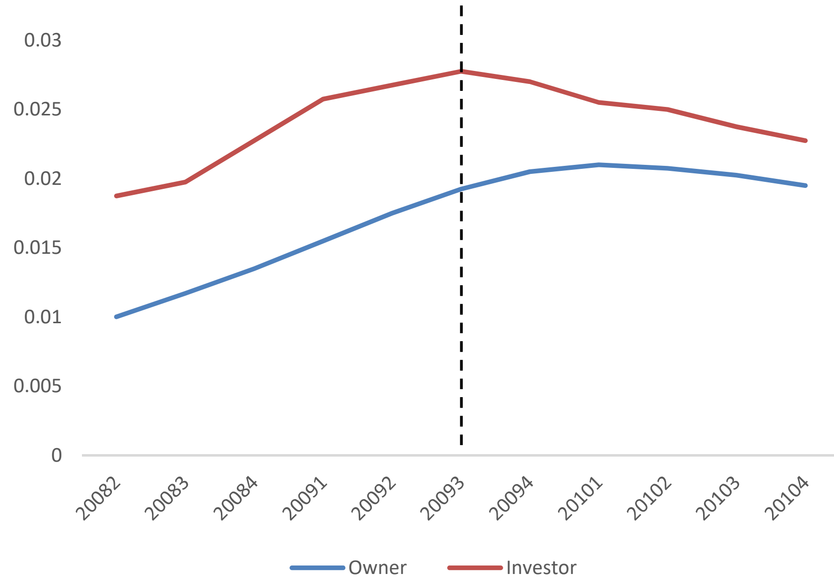
Dependent variable: 0/1 default indicator	Irrelevant event date		Low DTI sample	
	Estimate	S.E.	Estimate	S.E.
Treatment group beta	-0.005	0.003	-0.004*	0.002
Treatment group * Post event beta	0.001	0.002	0.002	0.002
Post event beta	0.003	0.002	0.003*	0.002
Control variables	Default option, Default option squared, Default option * unemployment rate innovation, Default option * orthogonalized consumer distress index, unemployment rate innovation, orthogonalized consumer distress index, post 2009Q3, Prepayment option, Prepayment option squared, change in unemployment rate, change in unemployment rate squared, LTV buckets, FICO buckets, DTI buckets, loan amount buckets, first time home buyer indicator, property type, occupancy type, loan purpose, state FE, vintage FE			
N	16,492		9,985	

Notes: Based on the GSE data. These are GLS estimates of a linear default model. The tests here are in the same form as those in Table 6 except that in the first test we pick a random breakpoint (2008Q3) where there is no policy change and in the second test both the "treatment" group and the control group are loans with DTI below 29 percent and thus are both HAMP ineligible. Loans included in the test are also limited to those originated before January 2009 with a remaining balance of no more than \$729,500. Standard errors are clustered at the loan-level. ***, ** and * indicate 0.1%, 1% and 5% significance, respectively.



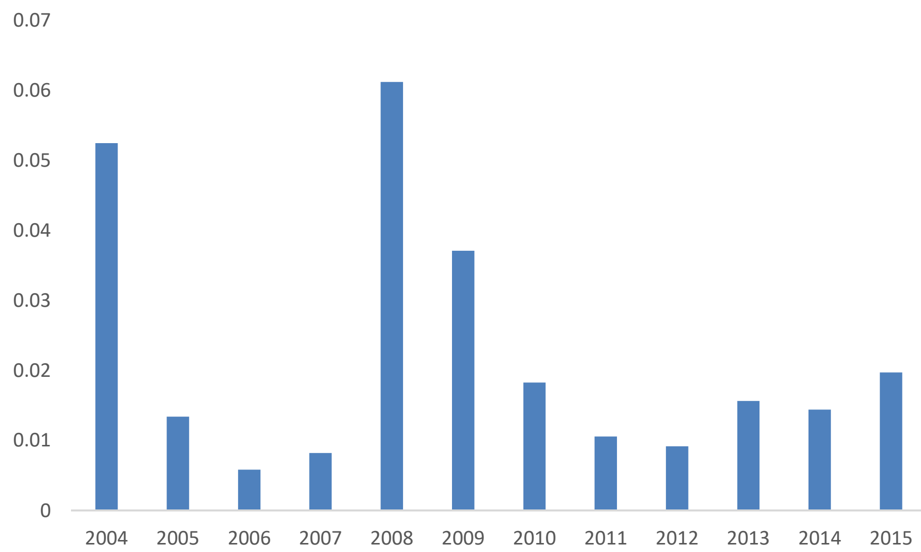
Appendix Figure 1 Negative Equity Spline Beta Estimates

Notes: Based on the GSE data. Instead of using the default option calculated based on the equity position and house price return volatility in the model we use negative equity and house price return volatility as standalone variables. A spline function is used for negative equity with cutoffs at -30%, -10%, 10% and 30%. During 2003 and 2007, the betas for certain segments of the spline function are not estimated due to too few loans with negative equity in those segments.



Appendix Figure 2 Parallel Trend Test for the Difference-in-Differences Test

Notes: Based on the GSE data. Parallel trend test results for the DID test in Table 6. It shows the default option betas of investor loans and owner loans, respectively, prior to and after HAMP. These are estimates of a linear default model similar to those in Table 6. The difference between the two groups of loans was stable prior to HAMP, as shown here.



Appendix Figure 3 House Price Measurement Error over Our Study Period

Notes: Using the CoreLogic Solutions real estate deeds data and HPI, we regress individual house price returns (based on repeated transactions) on house price index (HPI) return. This figure plots the mean absolute error (MAE) of such a regression. Only the top 50 zip codes ranked by the number of housing transactions are included. The chart shows that the magnitude of house price measurement error does not coincide with the variation in default option beta as shown in Figure 1.

Appendix Table 1 Negative Equity and House Price Return Volatility in the Competing Risks Hazard Model

Dependent variable: default/prepay hazard	Default			Prepayment		
	Estimate		S.E.	Estimate		S.E.
Negative equity <-30%	5.088	***	0.079	2.062	***	0.065
Negative equity -30~-10%	5.085	***	0.121	1.482	***	0.105
Negative equity -10~10%	4.407	***	0.395	0.588		0.397
Negative equity 10~30%	2.936	***	0.259	-2.070	***	0.289
Negative equity >30%	2.107	***	0.143	-1.511	***	0.168
House price return volatility	-0.005	*	0.002	-0.079	***	0.007
Call option	0.180	***	0.013	0.908	***	0.011
Call option squared	0.072	***	0.007	0.038	***	0.007
Change in unemp. rate	-0.026	*	0.012	-0.293	***	0.011
Change in unemp. rate squared	0.038	***	0.007	-0.009		0.005
LTV < 60	0.487	***	0.034	0.504	***	0.023
LTV 60-70	0.146	***	0.031	0.175	***	0.021
LTV 80-90	-0.089	***	0.023	-0.162	***	0.018
LTV > 90	-0.072	*	0.028	-0.423	***	0.024
Credit score < 580	0.805	***	0.084	-0.064		0.085
Credit score 580-620	0.391	***	0.038	-0.004		0.043
Credit score 660-700	-0.465	***	0.026	0.068	*	0.026
Credit score 700-740	-0.961	***	0.025	0.099	***	0.025
Credit score 740-780	-1.490	***	0.025	0.163	***	0.025
Credit score > 780	-1.986	***	0.027	0.189	***	0.026
DTI < 20	-0.367	***	0.039	0.042		0.022
DTI 20-30	-0.200	***	0.021	-0.001		0.016
DTI 40-50	0.215	***	0.020	-0.016		0.016
DTI> 50	0.349	***	0.027	-0.035		0.024
Loan amount < 10k	0.287	***	0.030	-0.529	***	0.023
Loan amount 10-15k	0.023		0.028	-0.270	***	0.021
Loan amount 15-20k	-0.032		0.029	-0.104	***	0.021
Loan amount 25-30k	-0.021		0.036	0.093	***	0.026
Loan amount 30-35k	0.054		0.041	0.074	*	0.030
Loan amount > 35k	0.058		0.044	0.263	***	0.029
First time home buyer	-0.104	***	0.028	-0.145	***	0.022
Condominium	-0.133	***	0.029	-0.051	*	0.023
Planned-unit development	-0.086	***	0.026	0.070	***	0.019
Investment property	0.194	***	0.032	-0.118	***	0.025
Second home	-0.143	**	0.045	-0.075	*	0.032
Cash out refinance	0.440	***	0.023	-0.100	***	0.018
Rate/term refinance	0.277	***	0.022	-0.091	***	0.017
State FE		Yes			Yes	
Vintage FE		Yes			Yes	
Flexible baseline function		Yes			Yes	
N			1,034,009			

Notes: Based on the GSE data. Instead of using the default option calculated based on the equity position and house price return volatility in the model, we use negative equity and house price return volatility as standalone variables. A spline function is used for negative equity with cutoffs at -30%, -10%, 10% and 30%. Standard errors are clustered at the loan-level. ***, ** and * indicate 0.1%, 1% and 5% significance, respectively.

Appendix Table 2 Default Burnout and Option Exercise

Dependent variable: default hazard	Estimate	S.E.
Put option * woodhead	-0.086**	0.020
Put option * unemployment rate innovation	0.042***	0.005
Put option * orthogonalized consumer distress index	-0.030**	0.008
Put option * post 2009Q3	0.028***	0.005
Put option	0.932***	0.037
Put option squared	-0.166***	0.022
Woodhead	0.303***	0.082
Unemployment rate innovation	0.310***	0.013
Orthogonalized consumer distress index	-0.201***	0.010
Post 2009Q3	0.206***	0.043
Control variables	Call option, call option squared, change in unemployment rate, change in unemployment rate squared, LTV buckets, FICO buckets, DTI buckets, loan amount buckets, fist time home buyer indicator, property type, occupancy type, loan purpose, state FE, vintage FE	
N	348,659	

Notes: Based on the GSE data. These are MLE estimates of the competing risks hazard model for default and prepayment based on a random sample of the event history data described in Table 2. The specification is the same as those in Table 5 except that we include an additional variable, "Woodhead", which is measured as the number of missed default opportunities since loan origination by comparing negative equity and the payment status in each period. Standard errors are clustered at the loan-level. ***, ** and * indicate 0.1%, 1% and 5% significance, respectively.

Appendix Table 3 Vintage Hazard Model Results

Dependent variable: default hazard	2003		2005		2007	
	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.
Put option * unemployment rate innovation	0.027**	0.008	0.062***	0.016	0.018*	0.007
Put option * orthogonalized consumer distress index	-0.062***	0.017	-0.042*	0.018	-0.032	0.022
Put option * post 2009Q3	0.076**	0.021	0.054*	0.020	0.007	0.038
Put option	0.846***	0.056	1.106***	0.033	1.036***	0.032
Put option squared	-0.077***	0.012	-0.165***	0.027	-0.002	0.026
Unemployment rate innovation	0.034	0.021	0.152***	0.029	0.171***	0.032
Orthogonalized consumer distress index	-0.243***	0.010	-0.188***	0.013	-0.188***	0.019
Post 2009Q3	-0.008	0.069	0.120*	0.046	0.219***	0.062
Control variables	Call option, call option squared, change in unemployment rate, change in unemployment rate squared, LTV buckets, FICO buckets, DTI buckets, loan amount buckets, fist time home buyer indicator, property type, occupancy type, loan purpose, state FE, vintage FE					
N	324,370		261,918		248,852	

Notes: Based on the GSE data. These are MLE estimates of the competing risks hazard model for default and prepayment based on a random sample of the event history data described in Table 2. The model specification is exactly the same as in Table 5, but here we run the model for each of the selected vintages. Standard errors are clustered at the loan-level. ***, ** and * indicate 0.1%, 1% and 5% significance, respectively.