Liquidity constraints, home equity and residential mortgage losses*

Hung Xuan Do[a,b†], Daniel Rösch[c], Harald Scheule[b]

[a]*School of Economics and Finance, Massey University, New Zealand*

[b]*Finance Discipline Group, UTS Business School, University of Technology, Sydney*

[c]*Department of Statistics and Risk Management, Faculty of Economics, University of Regensburg*

This version: 29 June 2018

*The authors would like to thank the participants of the seminars at the University of Regensburg, Germany, and the University of Technology, Sydney. We gladly acknowledge the valuable research assistance of Michael Tjendara. The support by the Center for International Finance and Regulation (CIFR, project number E001) is gratefully acknowledged. CIFR is funded by the Commonwealth and NSW Governments and supported by the other consortium members (see www.cifr.edu.au).

† Corresponding author.

Email addresses: h.do@massey.ac.nz (Hung Xuan Do), daniel.roesch@ur.de (Daniel Rösch), harald.scheule@uts.edu.au (Harald Scheule).

Liquidity constraints, home equity and residential mortgage losses

---

ABSTRACT

This paper analyses how borrower liquidity constraints and home equity drive the realized loss rates given mortgage default using loan-level data. We define defaulted loans with zero loss as cures and those with non-zero loss as non-cures. We find robust evidence that borrower liquidity constraints and positive equity explain cure, while negative equity explains non-zero loss. Borrower liquidity is related to the probability of cure in a v-shape fashion. The findings imply a separation of the cure and loss processes in mortgage loss models and their applications such as loan pricing and bank capital regulation.

Keywords: Cure, Loss Given Default, Liquidity Constraints, Home Equity, Mortgage, Resolution, Selection.

JEL classification: G21, G28, C19

1. Motivation

Residential mortgage loans are by far the most important asset class on bank balance sheets. Data released by the Federal Reserve Bank shows that all US commercial banks held real estate loans equivalent to $3.7 trillion in 2008 and $3.9 trillion in 2018. Mortgage credit risk is closely related to house prices and was considered to be a low risk prior to the Global Financial Crisis (GFC). Given default, mortgage loans often cure and result in zero loss rates given default (LGD). Commercial banks estimate probabilities of default (PDs) and LGDs for loan loss provisioning, regulatory capital requirements and loan pricing.

The GFC and subsequent literature have shown that borrowers may default in response to (i) liquidity constraints, (ii) negative equity, or a combination of both. For example, Elul et al. (2010) and Demyanyk et al. (2011) consider credit card utilization rates to proxy for liquidity constraints of borrowers and find that mortgage default is driven by both liquidity constraints and negative home equity. Foote et al. (2008), and Bhutta et al. (2010) document that negative equity dominates the unemployment rate in predicting the default. Contrary to this, Gerardi et al. (2018) employ data from the Panel Study of Income Dynamics to analyse the impact of liquidity shocks on mortgage default and find that loss of income due to unemployment is the most significant predictor of default during the GFC. In addition to this strand of the literature, Campbell and Cocco (2015) introduce a dynamic model that explains the mechanism through which the loan-to-value (related to home equity) and loan-to-income (related to liquidity constraints) ratio affect mortgage default.

Whilst the distinction between liquidity constraints and home equity has been analysed in the context of default prediction, the consequences on bank risk and losses from defaults, namely LGDs has not been considered. The majority of existing research focuses on unsecured credit exposures such as corporate and credit card loans without dissecting liquidity and negative equity processes. Examples are Chava et al. (2011), and Jankowitsch et al. (2014).

In terms of residential mortgage loans, the empirical LGD distribution generally has a bimodal shape with a high peak at zero. Figure 1 shows that nearly 30% of the total LGDs observations are zero-LGDs. That is, a significant fraction of defaulted loans does not generate losses. This is due to either (i) a bank's strategies in dealing with the defaulted loans such as loan modification or outright forgiveness of scheduled payments; or (ii) the property value after foreclosure exceeds the value of outstanding loan and associated resolution (i.e., workout) costs. We define defaulted loans with zero-LGD as cure loans and those with non-zero LGD as non-cure loans.[1] We hypothesize that default events which result in cures are more associated with the liquidity constraints of borrowers and those not cured to negative equity.

[Figure 1 about here]

We find that mortgage defaults that experience liquidity constraints and/or positive equity have a greater likelihood of resulting in cure events while mortgage defaults with negative equity have a greater likelihood of resulting in non-zero LGDs. The findings imply that negative equity was more causal than borrower liquidity constraints to the losses realized during the GFC, which differs from the liquidity constraints experienced by banks. Borrower liquidity is related to the probability of cure in a v-shape fashion. The findings imply a separation of the cure and loss processes in mortgage loss models and their applications such as loan pricing and bank capital regulation. The expected loss rate given default under this approach is calculated as (1-Probability of Cure)×(Non-zero LGD), which may replace the current singular treatment in bank calculations that averages over the cure and non-zero loss process. This may improve banks' efficiency in capital utilization because capital costs of the expected losses will be meaningfully lower if a portion of defaulted loans in a bank's residential

[1] In this paper, we use the two terms non-zero LGD loan (zero LGD loan) and non-cured loan (cured loan) interchangeably.

mortgage portfolio can be adequately linked with a high cure probability. A separation of the two processes allows for a better identification.

In this paper we apply a modification of the Heckman selection model (see Heckman, 1979) in order to capture the dynamics of cure and non-cure loans in different selection stages: the first stage aims to model the probability of cure and the second stage devises the magnitude of the non-zero LGD. Both stages are estimated jointly and Figure 2 illustrates this approach. This methodology has the merit of separately tracking the processes of cure and non-zero LGD. No paper to date considers the liquidity constraints and home equity, and hence the cure process, as a selecting event. Allowing a selection of cure events is important in LGD modelling as we find different dynamics of cure and non-cure loans.

[Figure 2 about here]

The remainder of this paper is organized as follows. Section 2 describes the data source and construction of variables. We provide the preliminary analyses in Section 3. Section 4 develops an econometrics framework for modelling LGDs which follows a two-step-selection mechanism for cure and non-cure loans. Section 5 discusses empirical results and Section 6 concludes the paper.

2. Data and construction of variables

2.1 Data

Our sample is provided by International Financial Research, which collects subprime loan level origination and performance data for the US non-agency residential mortgage-backed securities. The data is published to meet the information requirements of investors in these securities and provided by the servicing agents. Similar data has been used in other studies such as Rajan et al. (2015). The data has similar properties to prime mortgages in terms of risk drivers but has a higher default and loss rates. Agarwal et al. (2012) confirm that subprime loans are not exposed to adverse selection.[2]

We focus on single-family first lien loans and aggregate monthly intervals to quarterly intervals observed from 2005:Q2 to 2015:Q1. The complete dataset contains information on the loan (e.g., loan issuers, actual loan balance at origination and observation time, scheduled loan balance at the end of each period, loan age, interest rate and loan type), on the property (e.g., property location, property value at origination, whether the property is owner occupied), and on the borrower (e.g., FICO score at origination). Further, the data includes information about (i) the original loss (loss on liquidated property) and (ii) subsequent losses (loss on previously liquidated loans. Therefore, we include both measures for losses given default.

We define the default event as loan foreclosure and exclude defaulted loans that have not yet been resolved (unresolved loans, hereafter) to avoid resolution bias as unresolved loans show zero or low losses. We consider a defaulted loan as resolved if the servicing bank has excluded the loan from its portfolio and no further losses accrue to investors. Further, observations with missing values in variables are omitted and we finally record 509,408 defaulted loans that are resolved, which peaked during the crisis period (see Table 1).[3]

2.2 Cure and non-zero loss given default

LGD represents the economic loss at default and it is common to discount workout cash flows to the time of default and compute the aggregate loss rate given default. We define the LGD of a loan as the present value of losses, divided by the current outstanding loan balance

---

[2] In the study of Agarwal et al. (2012), adverse selection bias is broadly defined as the subjective decision of lenders to retain some specific types of loans (e.g., loans with high quality) in their portfolio while selling others to investors (e.g., loans with lower quality).

[3] We follow Bekaert et al. (2014) to define the start of the subprime mortgage crisis as 7 August 2007, due to the initial fall of equity markets and the first intervention of central banks to provide liquidity to financial markets. The end of the GFC is defined as 29 May 2009 with reference to the National Bureau of Economic and Research.

(current balance, hereafter) at default. We follow Qi and Yang (2009) and other contributions to discount losses to the time of default using the one-year LIBOR at default plus a spread of 3% for systematic risk. This discount rate is in line with a recent empirical study by Global Credit Data (www.globalcreditdata.org).[4] Hence, the LGD of loan i defaulted at time t_d is calculated as:

$$LGD_{i,t_d} = \frac{1}{Current\ Balance_{i,t_d}} \times \sum_{t=t_d}^{T} \frac{Losses_{i,t}}{(1+r)^{t-t_d}} \qquad (1)$$

where r is the discount rate.

Our data contains information about the losses resulting from loan defaults until the first quarter of 2015. As can be seen from Figure 3, most non-zero losses realize within 3 years of the default events. We exclude the unresolved loans from the sample. This implies that loans which have recently defaulted (e.g., 2014) have low resolution times (e.g., one year) and low LGDs. As a result, the mean LGD decreases towards the end of the observation period (see Figure 4). This resolution bias is intrinsic to all workout information and we control for it by the time from default to the last observation (TimeToEOO).

[Figure 3 and 4 about here]

Following Eq. (1), we define cure as zero LGD and non-cure as non-zero LGD, respectively. We summarize the cure rate and average LGDs per origination and observation year in Table 1. Consistent with Demyanyk and Hemert (2011), the majority of default loans are originated in years immediately prior to the GFC. Further, both average LGD (including

[4] Realised losses are generally lagging the real economy. The study shows that LGDs imply beta coefficients for North America of approximately 50%. Applying a market risk premium of 6% translates into a risk premium of 3%.

zero-LGD) and average non-zero LGD (excluding zero LGD) peak for these vintages. This behaviour is opposite to the cure rate, which experiences troughs (equivalently, peaks with non-cure rate) during these years.

[Table 1 about here]

Consistent with this, Figure 4 shows that the trend of non-zero LGDs is inverse to the cure rate (or equivalently, is in harmony with the non-cure rate). These movements suggest a negative relationship between cure rates and non-zero LGDs at *their mean level*. Comparing the average LGD and average non-zero LGD, we observe a time varying gap between them (visualized by the grey area in Figure 4), which is mainly due to the changing behaviour of the cure rate. This indicates different dynamics between cure and non-cure loans and motivates us to analyse zero and non-zero LGDs in separation.

2.3 Test variables

In this section, we describe how we construct the test variables including (i) borrower liquidity constraint and (ii) home equity. We summarize the full set of the explanatory variables in Table 2.

[Table 2 about here]

2.3.1 Borrower liquidity constraint

A borrower experiences a liquidity constraint if he or she cannot meet the loan repayment obligations. As a result, the actual current loan balance (current balance) is greater than the scheduled loan balance (scheduled balance). A larger gap between the current balance and the scheduled balance reflects a higher level of borrower liquidity constraint. We calculate the liquidity constraint measure for borrower *i* at time *t* (denoted as LC_{it}) as follows: [5]

$$LC_{it} = \frac{Current\ Balance_{it} - Scheduled\ Balance_{it}}{Scheduled\ Balance_{it}} \times 100\% \tag{2}$$

According to the construction of the LC variable, there are two economically meaningful cases: (i) the borrower meets scheduled payments ($LC_{it} = 0$) or makes curtailment payments for the loan ($LC_{it} < 0$);[6] and (ii) the borrower does not meet scheduled payments and experiences a liquidity constraint ($LC_{it} > 0$).[7] We include spline terms of LC as summarized in Table 2 to capture potential non-linear relationships between the borrower liquidity constraint and probability of cure or non-zero LGD. Further, to control the effect of outliers, we winsorize the LC values at the 5th and the 95th percentile.

2.3.2 Home equity

We employ the current loan-to-value ratio (*CLTV*) to measure home equity:

$$Home\ equity = 1 - CLTV \tag{3}$$

---

[5] We let the $LC_{it} = 0$ if the current balance and scheduled balance are both zero, indicating that the borrower has repaid the loan in full and does not experience a liquidity constraint. We consider LC_{it} as a missing value (and exclude from analyses) if the scheduled balance equals zero but the current balance is positive. This has no further implications on the results as this filter rule is applied to six defaulted loans.

[6] In either of these situations, the borrower does not experience a liquidity constraint. Curtailments refer to the cases where borrowers make extra principal loan repayments for their mortgages, and therefore, current balances are less than the scheduled balances (see for example, Amromin et al., 2007; Adelman et al., 2010).

[7] We note a possibility that some borrowers may choose to miss loan repayments due to strategic reasons that are particularly related to negative home equity (see for example, Mian and Sufi, 2009; Gerardi et al., 2018). Under strategic defaults, positive LC may not truly reflect the liquidity constraints of the borrowers, who have an ability to repay their mortgage but finally decide to default. We test the robustness of our results considering this possibility in our partitioned sample analyses in Table 9 and Figure 10.

Home equity has a maximum value of one and is negative if the current value of the property drops below the current loan balance (i.e., *CLTV* is greater than one) and positive if the property value is equal or greater than the outstanding loan balance (i.e., *CLTV* is lower than one).

Similar to the *LC* variable, we also include splines for *Home equity* as summarized in Table 2 to allow for non-linear relationships with regard to the probability of cure and the non-zero LGD. We control the effect of outliers by winsorizing *Home equity* values at the 5th and the 95th percentile.

As the information on the current value of the property is unavailable, we employ the House Price Index (HPI) for 401 Metropolitan Statistical Areas (MSAs) level provided by the Federal Housing Finance Agency (FHFA) to calculate *CLTV*.[8] We first map the zip-code of the property to the MSA of the HPI by using mapping data provided by the US Department of Housing and Urban Development (HUD).[9] We then calculate the *CLTV* of loan *i* at observation time *t* as:

$$CLTV_{it} = \frac{Current\ Balance_{it}}{Current\ Appraisal\ Value_{it}} \tag{4}$$

where the current appraisal value is approximated by the HPI of the MSA *m* for loan *i*:

$$Current\ Appraisal\ Value_{it} = Original\ Appraisal\ value_{it_0} \times \frac{HPI_{mt}}{HPI_{mt_0}} \tag{5}$$

[8] Glennon et al. (2018) find that the FHFA HPI index is comparable to other commercial HPI measures such as the Case-Shiller MSA index, and the Black Knight index (at the zip code-, MSA-, and state-level). We also perform a robustness check using Zillow HPI at zip-code level and obtain consistent results.

[9] We exclude loans with missing zip-codes.

2.4 Control variables

We employ several variables common in the literature to explain the dynamics of the probability of cure and non-zero LGD including loan characteristics, borrower characteristics, property characteristics and economic conditions.[10]

Although there have been few studies focusing on residential mortgage LGD, a variety of determinants of loss severities have been examined. For example, a number of loan characteristics including loan age, size, type, and purpose have been investigated (see, Clauretie and Herzog, 1990; Lekkas et al., 1993; Pennington-Cross, 2003; Qi and Yang, 2009; Zhang et al., 2010). Property characteristics including owner occupancy, property types, (e.g., single family, condominium and manufactured house) and state foreclosure laws (judicial process, statutory right of redemption, deficiency judgment) are examined in Clauretie and Herzog (1990), Pennington-Cross (2003), Qi and Yang (2009), and Zhang et al. (2010). Economic conditions including housing market conditions, unemployment rate, real economic growth and median income are considered in Clauretie and Herzog (1990), Qi and Yang (2009), and Zhang et al. (2010). However, we observe that no study to date considers the FICO score as a driver of mortgage LGDs. This might be due to a common belief that FICO score is predictive for default and not predictive for LGD. We run a preliminary test on a relationship between FICO and LGDs using OLS regressions with clustered standard errors and find an unexpected result that higher FICO scores (i.e., higher credit quality borrower) are associated

[10] We do not include the loan age in our set of control variables because of the collinearity between loan age and the *TimeToEOO* (a variable that we use to control the resolution bias), which results from standard maturities (e.g., 30 years) applied to most of the analysed mortgage contracts. For state foreclosure laws, we do not include judicial process (*JP*) in the cure equation but the non-zero LGD equation. This helps to differentiate sets of covariates used in the two equations of our two-step selection approach (see Section 4) to avoid the identification problems of the selection model.

with higher LGDs.[11] We demonstrate below that this result is reasonable in a two-step selection modelling for cure and non-cure loans.

We summarize the definition and the source of these control variables in Table 2 and explain the empirical results for all control variables in Section 5.3.

3. Descriptive analysis

In this section, we provide a descriptive analysis for the relationships between the probability of cure or non-zero LGD and the explanatory variables. Note that we only analyse default observations.

For categorical control variables, we report the relative frequency (%) of the discrete control variables for defaulted, cure and non-cure loans in Table 3. We observe a higher (lower) relative frequency of ARM (Owner Occupied) loans for non-cure loans than for cure loans. This is consistent with our expectation that ARM (Owner Occupied) loans contain higher (lower) risk than fixed rate (investment) mortgage loans. Further, it is likely that the defaulted loans that originate from states that apply statutory right of redemption or prohibit deficiency judgement are less likely to be cured. These are inferred from a lower relative frequency of cure loans originated in states that prohibit deficiency judgement or apply statutory right of redemption compared to non-cure loans.

[Table 3 about here]

We summarize the descriptive statistics of main test variables and continuous control variables as well as their pairwise correlations in Table 4 and 5.

[11] To conserve space, we do not provide estimated results of these preliminary models. Details are available upon request.

[Table 4 and 5 about here]

The pairwise correlations in Table 5 show that the relationship between the continuous control variables and non-zero LGDs are consistent with economic intuition.

Cure loans generally have higher credit quality, home equity, house price appreciation and real GDP growth than non-cure loans. These statistics are in line with the positive correlations between these control variables and the cure rate. Furthermore, cure loans have a lower current interest rate and loan size, which is consistent with negative correlations of these variables with the cure rate as observed in Table 5.

Exceptions are liquidity constraint, FICO score and the unemployment rate, which do not follow that common behaviour. We briefly discus the FICO score and unemployment results and liquidity constraint below. The FICO score represents the credit quality of a borrower, whereby a higher FICO score implies a lower credit risk. Table 4 shows that the FICO score has a lower mean for cure loans compared to non-cure loans. These statistics are also consistent with a negative correlation between the FICO and the cure rate that we observe in Table 5. In fact, this unique characteristic of the FICO score in cure loan can be explained by the common behaviour of borrowers trying to maintain their credit quality and is discussed in more detail in Section 5.3.

We observe that the unemployment rate has a higher mean in cure loans in comparison to non-cure loans. This adverse behaviour of unemployment rate towards cure might be due to an increase in the unemployment rate and may reflect a higher likelihood of experiencing a liquidity constraint, as a loss of job implies a constraint on the mortgage serviceability and hence borrower liquidity. The liquidity constraint of the borrower should be positively related to the cure loans. The pairwise correlations between cure rate, borrower liquidity constraint and unemployment rate are positive. The correlation between cure rate and borrower liquidity constraint is greater than the correlation between cure and unemployment rate.

Regarding the key test variables (i.e., borrower liquidity constraint and home equity), we find that their central measures are both significantly higher for cure loans than for non-cure loans (see Table 4). It is likely that the borrower liquidity constraint and positive home equity are more associated with cure than non-cure events. This is further confirmed by the positive correlation between these variables and the cure rate as observed in Table 5.

These relationships may not be linear and thresholds may exist. That is, the sensitivity of cure rate changes if liquidity constraints and home equity cross the zero thresholds. In terms of home equity, the zero threshold distinguishes whether the equity position is positive or negative. Regarding the borrower liquidity constraint, zero is the threshold that differentiates whether the borrower makes curtailments for the loan or experiences liquidity constraint. We visualize the relationships between borrower liquidity constraint, home equity and cure rate in Figure 5 and 6, respectively. We identify that thresholds of 0, 0.04, 0.08 and 0.52 of the borrower liquidity constraint are critical points where the relationship changes. The thresholds of -0.2, 0, 0.1 and 0.2 of home equity (equivalent with *CLTVs* of 1.2, 1, 0.9 and 0.8, respectively) were chosen in line with the literature (e.g., Qi and Yang, 2009; Gerardi et al., 2018).

[Figure 5 and 6 about here]

Figure 5 shows that the relationship between cure rate and borrower liquidity constraint follows an asymmetric v-shape with a long right-tail and a trough around 0.04%. Figure 6 displays an asymmetric tilde (~) – shape relationship between home equity and the cure rate with a very high right-tail. These behaviours are confirmed by the relative frequency of cure and non-cure rate categorized by categorical borrower liquidity constraint and home equity as shown in Table 6.

In summary, we observe from Table 6 as well as Figure 5 and 6 that the borrower liquidity constraint and positive home equity are associated with cure loans, whereas the negative home equity is more related to non-cure loans. These are also consistent with the pairwise correlations between cure rate, non-zero LGD, liquidity constraint and home equity observed in Table 5, which shows significantly positive correlations between liquidity constraint, home equity and cure rate and significantly negative correlations between home equity and non-zero LGD.

[Table 6 about here]

Borrower liquidity constraint is positively correlated with home equity, due to a time-delayed reflection between unemployment and borrower liquidity constraints. Normally, one would expect low unemployment rates and high home equity during an economic upturn as well as high unemployment rates and low home equity during an economic downturn. A negative correlation between borrower liquidity constraint and home equity is expected as the unemployment rate may lead to liquidity constraints (e.g., the loss of income may result in the inability to service a loan). Figure 7 shows that the dynamics of the unemployment rate and home equity are consistent with these expectations. However, we observe that the impact of the unemployment rate on the borrower liquidity constraint is lagged by one year (i.e., time-delayed reflection). Job losses do not translate into liquidity constraints in a contemporary fashion and the impact is lagged by approximately one year. This is reasonable as borrowers may rely on other sources of funding for some time after becoming unemployed. Besides, the lenders may also support the borrowers to rectify the loan in this period by modification of payment terms or forgiveness of scheduled payments.

[Figure 7 about here]

We check the effect of the time-delayed reflection by looking at the correlation of a lag of four with the unemployment rate and home equity. We find that the correlation between average borrower liquidity constraint and lag four of average unemployment rate (home equity) is about 70% (-56%), which is consistent with our economic intuition. This time-delayed reflection explains the positive contemporary correlation between borrower liquidity constraint and home equity that we observe in Table 5.

We inspect the sensitivity of the relationship between a borrower liquidity constraint and LGD (including cure and non-zero LGD) against the time-delayed reflection effect by looking at the scatter plots between their time averages (see Figure 8). The overall positive association between borrower liquidity constraint and the cure rate is robust regardless of the time-delayed reflection effect. However, we observe a moderate negative association between the borrower liquidity constraint and the non-zero LGD, which becomes almost insignificant when we use lag four of the non-zero LGD. This indicates that the negative contemporaneous association between borrower liquidity constraint and non-zero LGD is not economically meaningful. This confirms that the association is due to the coincidental movements between the borrower liquidity constraints and home equity caused by the time-delayed reflection effect.

[Figure 8 about here]

4. Modelling framework

To address the bimodal nature of the LGD distribution, we modify the Heckman selection model to a two-step selection mechanism of the observed loss severities that conditions on the selecting non-cure event and considers censoring of non-zero LGD. The two-step modelling includes: (i) the probability of cure for default events; and (ii) the non-zero LGD for defaults and non-cures. This selection mechanism is visually shown in Figure 2. We apply the censored regression models for the non-zero LGD in the second stage as LGD is bounded by zero and one. Hence, our model includes two equations that can be summarized as follows:

(4.1) Probability of Cure (PC)

$$C_{it}^{*} = X'_{i,t-1}\beta + u_{it} \qquad\qquad u_{it} \sim N(0,1)$$

$$C_{it} = \begin{cases} 1 & if\ C_{it}^{*} > 0 \\ 0 & if\ C_{it}^{*} \leq 0 \end{cases},$$

In our PC model (4.1), the cure variable, C_{it}, is a binary variable indicating whether the defaulted loan is cured ($C_{it} = 1$) or non-cured ($C_{it} = 0$), which is characterized by the outcome of the underlying latent process $C_{it}^{*} > 0$ or $C_{it}^{*} \leq 0$. $X_{i,t-1}$ is a vector of all cross-sectional and time-varying variables observed in the previous quarter that explains the PC, β is the vector of parameters and the error term u_{it} is assumed to follow a standard normal distribution.

(4.2) Non-zero LGD

$$L_{it}^{*} = Z'_{i,t-1}\alpha + \varepsilon_{it} \qquad\qquad \varepsilon_{it} \sim N(0, \sigma_{\varepsilon}^{2})$$

$$L_{it} = \begin{cases} 1 & if \quad L_{it}^{*} \geq 1 \\ L_{it}^{*} & if \quad 0 < L_{it}^{*} < 1 \end{cases},$$

Following the selection mechanism shown in Figure 2, the magnitude of non-zero loss severity for loan *i* at time *t*, L_{it}, is only observed if defaulted loan *i* which defaults at time *t* is not cured (i.e., $C_{it} = 0$). In that case, the non-zero loss severity L_{it} is observed and follows a censored linear relationship characterized by the underlying latent process L_{it}^{*} as shown in Eq. (4.2), where, $Z_{i,t-1}$ is a vector of all time-varying determinants of the LGD observed in the previous quarter and α is the associated parameter vector. We assume a normal distribution with zero mean and σ_{ε}^{2} variance for the error term ε_{it}. We find that the empirical distribution of the error term, ε_{it}, is very close to normal, with skewness and excess kurtosis close to zero (e.g., -0.355 and 0.345 for our main model, respectively).[12]

Since these two linear equations are included in one closed system, the vector of error terms follows a multivariate normal distribution:

$$\begin{pmatrix} u_{it} \\ \varepsilon_{it} \end{pmatrix} \sim N \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho_{u\varepsilon}\sigma_{\varepsilon} \\ \rho_{u\varepsilon}\sigma_{\varepsilon} & \sigma_{\varepsilon} \end{pmatrix} \right] \tag{6}$$

where $\rho_{u\varepsilon}$ denotes the correlation between the two error terms, u_{it} and ε_{it}.

We estimate the model using Maximum Likelihood estimation and we provide a derivation of the likelihood function in the Appendix.

5. Empirical results

We examine the relationships between borrower liquidity constraint/home equity and the LGD by estimating the model with and without control variables for the full sample. As a robustness check, we separately investigate the crisis periods (i.e., from Q3:2007 to Q2:2009) as well as partitioned samples, for liquidity constraints and/or negative equity in the quarter before the default events. We analyse four sub-samples, including (i) defaulted loans with negative equity in the previous quarter (Pure NE sample), (ii) defaulted loans with both liquidity constraints and negative equity in the previous quarter (LC and NE sample), (iii) defaulted loans with either liquidity constraints or negative equity in the previous quarter (No LC and NE sample), and lastly, (iv) defaulted loans with liquidity constraints in the previous quarter (Pure LC sample).

[12] We have also tested non-linear transformation models with clustered standard errors (such as logit and probit) for non-zero LGD and find consistent results. Details are available on request.

5.1 Borrower liquidity constraints

This section analyses the relationship between borrower liquidity constraints and LGD (including probability of cure and non-zero LGD). We present the estimated results for the probability of cure and non-zero LGD equation in Table 7 and Table 8 and graph their relationships in Figure 9 for the estimated parameters to gain further insight into how borrower liquidity constraint relates to the probability of cure and non-zero LGD.

[Table 7 and 8 about here]

Probability of cure

We find a significant non-linear effect of borrower liquidity constraint on the probability of cure. The estimated parameters at every spline knot (threshold) are all statistically significant at 1% level (see Table 7), indicating that the effect of borrower liquidity constraint on the probability of cure changes significantly at these thresholds. This supports our choices of the thresholds and further confirms a non-linear relationship between borrower liquidity constraint and the probability of cure. Figure 9 consistently shows an asymmetric v-shape relationship with an increasingly long right-tail and a trough at 0.04%. This pattern is robust for different models and sample periods under consideration.

[Figure 9 about here]

We observe a decreasing left-tail of the v-shape relationship (until the borrower liquidity constraint is equal to zero), which is consistent with the economic intuition that the probability of cure is expected to be higher with greater curtailments. Furthermore, we observe that a sudden drop in the probability of cure (about 10% after controlling for other effects) associated with an increase of borrower liquidity constraint from 0 to 0.04% may be attributed to technical defaults. In the context of residential mortgage loans, technical defaults may arise from a failure to pay property taxes or homeowner's insurance premiums as described by Moulton et al. (2015). Following this research, technical defaults are not failures in loan repayment, and hence the borrower liquidity constraints associated with these defaults are by definition non-positive. Further, it is more likely that the value of property does not fall below the outstanding loan balance for technical defaults.[13] Hence, we observe a higher level of the probability of cure for borrowers with non-positive liquidity constraints relative to the positive ones as shown in Figure 9.

It is worth noting that the difference in these levels of probability of cure is also subject to the housing cycle, which has an impact on the likelihood that the technical defaults are accompanied with negative equity. For example, we may expect more technical defaults experiencing negative equity during the subprime mortgage crisis due to decreasing house prices. Our robustness check for the crisis period demonstrates this point as a smaller drop of the probability of cure compared to the full sample estimation when borrower liquidity constraints occur (see Figure 9).

The moderate increase of the v-shape relationship on the long right-tail between borrower liquidity constraint and probability of cure is consistent with our expectation. Borrower liquidity constraints are more associated with cure than non-cure loans. This result is robust regardless of the employed model specifications and sample periods. A positive association between cure loans and borrower liquidity constraints is further supported by our empirical estimation using partitioned samples as shown in Table 9. The right-tail of the v-shape relationship is captured for the LC and NE sample and the Pure LC sample. All the estimates of knots are statistically significant and show a similar pattern for the full sample and crisis sub-samples. For a better visualization, we graph the relationship between borrower liquidity constraints and probability of cure in each partitioned sample's estimation in Figure 10. We observe a larger shift in their positive association in the Pure LC sample than in the LC

[13] This is consistent with our observation that approximately 61% of defaulted loans that did not experience positive borrower liquidity constraints did not have negative equity.

and NE sample, which supports the point that defaulted loans that are cured are more associated with the borrower liquidity constraints. In addition, findings for the Pure LC sample analysis confirm the robustness of our results after controlling for strategic defaults. It is worth noting that strategic defaults are particularly related to negative equity mortgages (see Mian and Sufi, 2009; Gerardi et al., 2018).

[Table 9 and Figure 10 about here]

Non-zero LGD

We observe an increase in the level of non-zero LGD when borrower liquidity constraints occur (see Figure 9). This is consistent with technical defaults that we have discussed earlier. We expect a lower level of losses of loan foreclosures caused by technical defaults, which is equivalent with non-positive borrower liquidity constraints, compared to loan foreclosures with negative equity. Borrower liquidity constraints are negatively related to the non-zero LGD until the threshold 0.52%. That is, a higher positive borrower liquidity constraint is associated with a lower non-zero LGD. In terms of magnitude, we find that this overall trend is mostly driven by the defaulted loans that experience liquidity constraints but not negative equity in the previous quarter (see Figure 10). However, the decreasing trend is statistically insignificant after controlling for other effects (including home equity, see Table 9). This can be explained by the time-delayed reflection of borrower liquidity constraints as we discussed in Section 3. This negative association may be due to an instantaneous co-movement between borrower liquidity constraints and home equity.

5.2 Home equity

In this section, we analyse the relationship between home equity and LGD (including probability of cure and non-zero LGD). The estimated results for the probability of cure and non-zero LGD equations are presented in Table 7 and Table 8. Further, we analyze how home equity is related to the probability of cure and non-zero LGD by plotting their relationships in Figure 11 on the basis of estimated parameters shown in Table 7 and Table 8.

Probability of cure

We consistently find a non-linear relationship between home equity and the probability of cure with a continuously increasing trend. The estimated parameters at the knots -0.2, -0.1, 0 and 0.2 are statistically significant at the 1% significance level (see Table 7), indicating that the relationship changes statistically at these points. This supports our choice of thresholds and it is also consistent with the literature (e.g., Qi and Yang, 2009; Gerardi et al., 2018). Overall, we observe an asymmetric tilde (~) – shape relationship between home equity and the probability of cure with a flat relationship in the middle (see Figure 11). We find that this pattern is consistent for different model specifications and sample periods examined.

[Figure 11 about here]

Looking at the home equity on the negative and positive sides, we find that an increase in equity elevates the probability of cure while an increase in the negative equity lowers the probability of cure significantly. These results imply that the positive equity is a driver of cure loans and negative equity is a driver of non-cure loans. This is consistent with economic intuition as the positive (negative) equity indicates that the property value is above (below) the outstanding loan balance. Hence, the outstanding loan balance and resolution costs are (not) likely to be covered by the sale of the property for the case of positive (negative) equity. These results also hold for all partitioned samples as shown in Table 9.

Non-zero LGD

We find a decreasing trend of the relationship between home equity and the non-zero LGD throughout when home equity increases (see Figure 11). The result indicates a robust negative relationship between the two variables, that is, higher home equity leads to a lower non-zero LGD. This is consistent with previous literature of the residential mortgage loans (e.g., Pennington-Cross, 2003; and Qi and Yang, 2009). Greater home equity leads to a higher recovery rate (i.e., 1-LGD) and, therefore, lower non-zero LGD. This relationship between home equity and non-zero LGD is not linear as the rate of changes of non-zero LGD is not constant for changes in home equity. This is shown by statistically significant estimates for all knots of home equity at 1% significance level in Table 8 and visible in Figure 11. We find that the non-linear behaviour of the relationship between home equity and non-zero LGD is much smoother than between home equity and the probability of cure. Overall, this pattern of the home equity/non-zero LGD relationship is robust for different model specifications and sample periods.

In addition, the significantly decreasing trend of home equity/non-zero LGD relationship also implies that a higher level of negative equity leads to a higher non-zero LGD.

5.3 Control variables

Probability of cure

Estimated results for the controls of the cure equation presented in Table 7 have shown some interesting features. The foreclosed loans with higher FICO scores are less likely to be cured. FICO (2008) research shows that higher FICO scores tend to be more stable over time, which means higher FICO customers have put more effort into fulfilling their payment obligations. Borrowers may face a significant relative increase in credit costs if their credit risk lowers and this effect can last for more than ten years (see Han and Li, 2011). Generally, when high FICO customers face liquidity constraints (e.g., divorce, short-term unemployment or demotion) and miss a loan repayment, it is likely that they will try to rectify the loans at an early stage of delinquency, only giving up and allowing the loans to be foreclosed if it is no longer worthwhile maintaining their obligations and FICO profiles. This is likely to happen if they experience negative equity alongside the liquidity constraints, or in other words, their property value drops below the outstanding loan amount, which finally leads to the losses being less likely to be fully recovered.

The other control variables show significant impacts on the probability of cure, which are consistent with economic intuition. We find that defaulted loans associated with higher risk (e.g., adjustable rate mortgages) are less likely to be cured, whereas those with lower risk (e.g., owner occupied purpose loans) are more likely to be cured. Loan size has a significant concave effect on the probability of cure for the entire sample estimation, but the effect is statistically insignificant during the crisis period. This might be due to declining house prices during the GFC regardless of the house value at origination and hence, the loan size.[14] The current interest rate and state unemployment rate have a negative effect on the probability of cure while the house price appreciation and real GDP growth rate have a positive impact. A rise in the current interest rate or unemployment rate may decrease borrower ability to repay the loan. An increase in the house price appreciation or real GDP growth rate may improve this ability.

It is consistent with our expectation that the states prohibiting deficiency judgement and/or allowing for a statutory right of redemption have experienced a lower probability of cure for defaulted loans. A prohibition of deficiency judgement does not allow the lenders to collect the gap from borrowers if the repossession of a forced sale property cannot fully cover the outstanding loan balance and associated resolution costs. However, the statutory right of redemption can prolong the foreclosure and liquidation process and may result in significant extra costs.

Non-zero LGD

The estimated results for the controls of the non-zero LGD equation presented in Table 8 are intuitive and consistent with previous studies. Non-cure loans are associated with higher

---

[14] At origination, the loan size dynamics are strongly consistent with the house value since the majority of mortgage loans are originated with a loan-to-value ratio of 80%.

credit quality (i.e., higher FICO score) and have a lower non-zero LGD. This result implies that the credit score (FICO score) is not only a driver of default but also explains LGD. However, the FICO – LGD relationship may be misinterpreted if we combine both cure and non-cure loans in the estimation due to the borrower's behaviour in making decisions on whether to maintain the credit profile or to foreclose a loan. This behaviour has been discussed earlier in the analysis of the FICO effect on probability of cure. We find that defaulted loans with an adjustable mortgage rate are associated with higher loss severity in a comparison with other loan types, whereas owner-occupied mortgage loans experience a significantly lower loss severity than other loan types. These results are consistent with our expectation and supported by previous studies as ARM loans are associated with higher risk while owner-occupied loans contain less risk than other loan types (see, Qi and Yang, 2009; and Zhang et al., 2010).

We also find that loans with a higher interest rate are associated with higher non-zero LGD, which supports Zhang et al. (2010). The relationship between non-zero LGD and loan size exhibits a convex parabolic shape, which is consistent to findings of Pennington-Cross (2003). Regarding state foreclosure laws on property, we find states with judicial process (e.g., Clauretie and Herzog, 1990; Pennington-Cross, 2003; and Qi and Yang, 2009), states with statutory right of redemption (e.g., Clauretie and Herzog, 1990; and Qi and Yang, 2009) and states with deficiency judgement prohibited (e.g., Clauretie and Herzog, 1990) have higher non-zero LGD than other states. The judicial process and statutory right of redemption may prolong the foreclosure and liquidation process when the defaulted loans do not cure, which leads to larger loss severities. In states that prohibit the deficiency judgement, the lenders lack an effective method to cover the difference if the foreclosure sales cannot compensate the outstanding loan balances and associated resolution costs.

Regarding the macroeconomic environment, we find that increases in the house price appreciation or real GDP growth have a negative effect on the non-zero LGD while an increase in the unemployment rate elevates the non-zero LGD. These results are consistent with previous studies, for example, Clauretie and Herzog (1990) and Zhang et al. (2010) for house price appreciation, and Clauretie and Herzog (1990) for unemployment rate.

6. Findings

We investigate the effect of borrower liquidity constraints and home equity on LGD using US residential mortgage defaulted loans observed between Q2 2005 and Q1 2015. We find robust evidence that the relationship between borrower liquidity constraints and probabilities of cure exhibits an asymmetric v-shape with an increasingly long right-tail. Further, the home equity – probability of cure relationship has an asymmetric tilde (~) – shape which is an increasing monotone on both sides and a flat in the middle. These results indicate that borrower liquidity constraints and positive equity are more associated with cure loans while the negative equity is more related to non-cure loans. Higher levels of positive equity lead to a lower non-zero LGD, while higher levels of negative equity trigger higher non-zero LGD. We find a spurious negative relationship between borrower liquidity constraints and non-zero LGD. This is due to the contemporaneous co-movement of borrower liquidity constraints and positive equity, which is caused by the time-delayed reflection effect in the borrower liquidity constraints.

The above findings together with different behaviours of some control variables (e.g., FICO scores) for cure and non-cure loans suggest a utilization of modelling structures that differentiate between the dynamics of cure and non-cure loans. One of the options can be considering cure event and non-zero LGD in two-steps. The first step models probability of cure and the second step models non-zero LGD if the defaulted loan is non-cured. This approach captures the bimodal property of the LGD distribution, in which a large fraction of defaulted loans does not generate losses. Economically, this approach is an important refinement of bank risk models and those underlying prudential regulation that considers the probability of cure as a risk mitigating factor.

References

Adelman, S. W., Cross, M. L., and Shrider, D. G. (2010). Wy do homeowners make mortgage curtailment payments? *Journal of Housing Research, 19*, 195-212.

Agarwal, S., Chang, Y., and Yavas, A. (2012). Adverse selection in mortgage securitization. *Journal of Financial Economics, 105*, 640-660.

Amromin, G., Huang, J., and Sialm, C. (2007). The tradeoff between mortgage prepayments and tax-deferred retirement savings. *Journal of Public Economics, 91*, 2014-2040.

Bekaert, G., Ehrmann, M., Fratzcher, M., and Mehl, J. A. (2014). The global crises and equity market contagion. *Journal of Finance, 69*, 2597-2649.

Bhutta, N., Dokko, J., and Shan, H. (2010). The depth of negative equity and mortgage default decisions. *Federal Reserve Board Finance and Economics Discussion Series.*

Campbell, J. Y., and Cocco, J. F. (2015). A model of mortgage default. *Journal of Finance, 70*, 1495-1554.

Chava, S., Stefanescu, C., and Turnbull, S. (2011). Modeling the loss distribution. *Management Science, 57*, 1267-1287.

Clauretie, T. M., and Herzog, T. (1990). The effect of state foreclosure laws on loan losses: evidence from the mortgage insurance industry. *Journal of Money, Credit, and Banking, 22*, 221-233.

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

Demyanyk, Y., Koijen, R. S., and Van Hemert, O. A. (2011). Determinants and consequences of mortgage default. *Federal Reserve Bank of Cleveland Working Paper.*

Elul, R., Souleles, N., Chomsisengphet, S., Glennon, D., and Hunt, R. (2010). What "triggers" mortgage default? *American Economic Review, 100*, 490-494.

FICO. (2008). Scoring your customers: How often is often enough? Available at: http://www.fico.com/en/wp-content/secure_upload/Insights_Score_Migration_2505WP.pdf.

Foote, C. L., Gerardi, K., and Willen, P. S. (2008). Negative equity and foreclosure: theory and evidence. *Journal of Urban Economics, 64*, 234-245.

Gerardi, K., Herkenhoff, K. F., Ohanian , L. E., and Willen, P. S. (2018). Can't pay or won't pay? unemployment, negative equity, and strategic default. *Review of Financial Studies, 31*, 1098-1131.

Glennon, D., Kiefer, H., and Mayock, T. (2018). Measurement error in residential property valuation: An application of forecast combination. *Journal of Housing Economics, 41*, 1-29.

Han , S., and Li, G. (2011). Household borrowing after personal bankruptcy. *Journal of Money, Credit and Banking, 43*, 491-517.

Heckman, J. (1979). Sample selection bias as a specification error. *Econometrica, 47*, 153-161.

Jankowitsch, R., Nagler, F., and Subrahmanyam, M. G. (2014). The determinants of recovery rates in the US corporate bond market. *Journal of Financial Economics, 114*, 155-177.

Lekkas, V., Quigley, J. M., and Van Order, R. (1993). Loan loss severity and optimal mortgage default. *Journal of American Real Estate and Urban Economics Association, 21*, 353-371.

Mian, A., and Sufi, A. (2009). The consequences of mortgage credit expansion: Evidence from the US mortgage default crisis. *Quarterly Journal of Economics, 124*, 1449-1496.

Moulton, S., Haurin, D. R., and Shi, W. (2015). An analysis of default risk in the home equity conversion mortgage (HECM) program. *Journal of Urban Economics*, *90*, 17-34.

Pennington-Cross, A. (2003). Subprime and prime mortgages: Loss distributions. *Working paper, Office of Federal Housing Enterprise Oversight.*

Qi, M., and Yang, X. (2009). Loss given default of high loan-to-value residential mortgages. *Journal of Banking and Finance, 33*, 788-799.

Rajan, U., Seru, A., and Vig, V. (2015). The failure of models that predict failure: Distance, incentives, and defaults. *Journal of Financial Economics*, *115*, 237-260.

Zhang, Y., Ji, L., and Liu, F. (2010). Local housing market cycle and Loss Given Default: Evidence from sub-prime residential mortgages. *Working paper International Monetary Fund.*

Appendix: Derivation of likelihood for model estimation

Following the selection mechanism shown in Figure 2, we derive the two main components for the likelihood function as follows:

1. If the defaulted loan *i* is cured ($C_{it} = 1$):

$$\Pr(C_{it} = 1) = \Pr\left(C_{it}^* > 0 \middle| X_{i,t-1}\right) = \Pr\left(X'_{i,t-1}\beta + u_{it} > 0\right) = \Pr\left(u_{it} > -X'_{i,t-1}\beta\right)$$

$$= 1 - \Phi\left(-X'_{i,t-1}\beta\right)$$

$$= \Phi\left(X'_{i,t-1}\beta\right)$$

   where $\Phi(.)$ denotes the cumulative distribution function of the standardized normal distribution.

2. If the defaulted loan *i* is non-cured ($C_{it} = 0$):

   In this case, we can observe the non-zero LGD. As the non-zero LGD is bounded by 1, we have two possibilities which build up the density of the L_{it}:

   First, $L_{it} = 1$

$$\Pr(L_{it} = 1, C_{it} = 0) = \Pr\left(L_{it}^* \geq 1, C_{it}^* \leq 0 \middle| X_{i,t-1}, Z_{i,t-1}\right)$$

$$= \Pr\left(\frac{\varepsilon_{it}}{\sigma_\varepsilon} \geq \frac{1 - Z'_{i,t-1}\alpha}{\sigma_\varepsilon}, u_{it} \leq -X'_{i,t-1}\beta\right)$$

$$= \Phi_2\left(-\frac{1 - Z'_{i,t-1}\alpha}{\sigma_\varepsilon}, -X'_{i,t-1}\beta, -\rho_{u\varepsilon}\right)$$

   where $\Phi_2(.)$ denotes the cumulative distribution function of standardized bivariate normal distribution.

   Second, $0 < L_{it} < 1$,

$$\Pr(L_{it}, C_{it} = 0) = \Pr(L_{it}, C_{it}^* \leq 0 | X_{i,t-1}, Z_{i,t-1})$$

   According to the Bayes rule, we have:

$$\Pr(L_{it}, C_{it}^* \leq 0 | X_{i,t-1}, Z_{i,t-1})$$

$$= f\left(L_{it} \middle| Z_{i,t-1}\right) \Pr\left(C_{it}^* \leq 0 \middle| L_{it}, X_{i,t-1}, Z_{i,t-1}\right) \qquad (7)$$

In Eq. (7), following density function of normal distribution it is easy to see that:

$$f\left(L_{it} \middle| Z_{i,t-1}\right) = f(\varepsilon_{it})$$

$$= \frac{1}{\sigma_\varepsilon}\phi\left(\frac{L_{it} - Z'_{i,t-1}\alpha}{\sigma_\varepsilon}\right)$$

where $\phi(.)$ denotes probability density function of standardized normal distribution.

The remaining part of Eq. (7) can be derived as follows:

$$\Pr\left(C^*_{it} \le 0 \middle| L_{it}, X_{i,t-1}, Z_{i,t-1}\right) = \Pr\left(C^*_{it} \le 0 \middle| \varepsilon_{it}, X_{i,t-1}\right) = \Pr(u_{it} \le -X'_{i,t-1}\beta \mid \varepsilon_{it}) \qquad (8)$$

The conditional distribution of u_{it} given ε_{it} can be written as:

$$u_{it}|\varepsilon_{it} \sim N\left[\frac{\rho_{u\varepsilon}}{\sigma_\varepsilon}\left(L_{it} - Z_{i,t-1}\alpha\right), 1 - \rho^2_{u\varepsilon}\right]$$

Hence, Eq. (8) is equivalent with:

$$\Pr\left(\frac{u_{it} - \frac{\rho_{u\varepsilon}}{\sigma_\varepsilon}(L_{it} - Z'_{i,t-1}\alpha)}{\sqrt{1 - \rho^2_{u\varepsilon}}} \le \frac{-X'_{i,t-1}\beta - \frac{\rho_{u\varepsilon}}{\sigma_\varepsilon}(L_{it} - Z'_{i,t-1}\alpha)}{\sqrt{1 - \rho^2_{u\varepsilon}}}\right)$$

$$= \Phi\left(-\frac{X'_{i,t-1}\beta + \frac{\rho_{u\varepsilon}}{\sigma_\varepsilon}(L_{it} - Z'_{i,t-1}\alpha)}{\sqrt{1 - \rho^2_{u\varepsilon}}}\right)$$

Combining the components of likelihood function from the two main events we have the full likelihood function of the model as:

$$L = \prod_{t=1}^{T}\prod_{i=1}^{N}\left[\Phi\left(X'_{i,t-1}\beta\right)\right]^{C_{it}} \cdot \left[\Phi_2\left(-\frac{1 - Z'_{i,t-1}\alpha}{\sigma_\varepsilon}, -X'_{i,t-1}\beta, -\rho_{u\varepsilon}\right)\right]^{(1-C_{it})\cdot I(L_{it}=1)}$$

$$\cdot \left[\frac{1}{\sigma_\varepsilon}\phi\left(\frac{L_{it} - Z'_{i,t-1}\alpha}{\sigma_\varepsilon}\right) \cdot \Phi\left(-\frac{X'_{i,t-1}\beta + \frac{\rho_{u\varepsilon}}{\sigma_\varepsilon}\left(L_{it} - Z'_{i,t-1}\alpha\right)}{\sqrt{1 - \rho^2_{u\varepsilon}}}\right)\right]^{(1-C_{it})\cdot I(0<L_{it}<1)}$$

where $I(.)$ denotes the indicator function.

Tables and figures

Table 1: Cure rates and average LGD by origination and observation year

This table reports the number of defaults (D), number of cured loans (C), average cure rates (i.e., cure per number of defaults, C/D), average LGD ($\overline{LGD}$, i.e., average of both zero and non-zero LGD) and average non-zero LGD ($\overline{LGD^*}$) by origination year and observation year. The last line reports their average values for the entire sample.

Origination Years						Observation Years					
Year	**D**	**C**	**C/D**	$\overline{LGD}$	$\overline{LGD^*}$	**Year**	**D**	**C**	**C/D**	$\overline{LGD}$	$\overline{LGD^*}$
1990	514	389	0.757	0.105	0.431						
1991	32	23	0.719	0.064	0.226						
1992	45	35	0.778	0.057	0.258						
1993	120	92	0.767	0.068	0.292						
1994	136	96	0.706	0.129	0.439						
1995	234	147	0.628	0.210	0.566						
1996	333	205	0.616	0.229	0.595						
1997	723	407	0.563	0.274	0.628						
1998	1,388	783	0.564	0.272	0.624						
1999	2,330	1,237	0.531	0.307	0.654						
2000	1,965	862	0.439	0.385	0.685						
2001	2,566	1,182	0.461	0.280	0.518						
2002	8,084	3,489	0.432	0.294	0.518						
2003	18,020	8,088	0.449	0.273	0.496						
2004	51,984	19,887	0.383	0.310	0.502						
2005	126,188	35,386	0.280	0.398	0.553	2005	15,656	6,707	0.428	0.259	0.454
2006	227,029	58,808	0.259	0.462	0.624	2006	44,839	14,817	0.330	0.327	0.488
2007	63,945	18,190	0.284	0.458	0.640	2007	102,548	19,345	0.189	0.479	0.590
2008	3,653	727	0.199	0.411	0.513	2008	136,482	25,499	0.187	0.517	0.636
2009	60	24	0.400	0.456	0.761	2009	96,033	28,746	0.299	0.424	0.606
2010	10	4	0.400	0.305	0.508	2010	43,551	15,711	0.361	0.383	0.600
2011	3	2	0.667	0.320	0.959	2011	30,183	12,292	0.407	0.339	0.571
2012	46	32	0.696	0.112	0.370	2012	18,834	9,690	0.515	0.248	0.511
						2013	13,436	9,733	0.724	0.123	0.446
						2014	7,494	7,204	0.961	0.012	0.319
						2015	352	351	0.997	0.001	0.322
Total	509,408	150,095	0.295	0.417	0.591		509,408	150,095	0.295	0.417	0.591

Table 2: Definition of the explanatory variables

Variable groups	Description
Loan characteristics	
Home equity	Home equity is calculated as, 1 – *CLTV*, where *CLTV* is the current loan-to-value ratio that is approximated using the HPI at MSA level. We construct the splines of home equity at knots (thresholds) -0.2, 0, 0.1 and 0.2 which are respectively denoted as $Home\ equity_{s-0.2}$, $Home\ equity_{s0}$, $Home\ equity_{s0.1}$, and $Home\ equity_{s0.2}$.
Loan size	Natural logarithm of loan amount at origination time, $Loan\ size_i = \ln(OB_i)$, where OB_i is the original balance of loan *i*.
Adjustable rate mortgage (*ARM*)	Indicator variable *ARM* represents adjustable rate mortgage (if *ARM*=1).
Current Interest Rate (*Current IR*)	*Current IR* represents the current interest rate associated with the loan.
Borrower characteristics	
FICO	*FICO* represents the score of a borrower at origination time, which measures a borrower's credit quality at loan approval. A higher score indicates a higher credit quality. This score is a popular indicator of the credit quality used by lenders, rating agencies and investors since the mid-1990s.
(Borrower) Liquidity constraint	The borrower liquidity constraint represent the incapacity of a borrower to fulfil the loan repayment obligations as scheduled. We construct the splines of borrower liquidity constraints at knots (thresholds) 0, 0.04, 0.08 and 0.52 denoted as $Liquidity\ constraints_{s0}$, $Liquidity\ constraints_{s0.04}$, $Liquidity\ constraints_{s0.08}$, and $Liquidity\ constraints_{s0.52}$.
Property characteristics	
State foreclosure laws on property	Variables *JP*, *SRR* and *PDJ* indicate states where judicial process is allowed (if *JP*=1), statutory right of redemption is allowed (if *SRR*=1) and deficiency judgment is prohibited (if *PDJ*=1).
Owner-occupied	An indicator variable receiving value of one if the property is owner occupied and zero otherwise.
Economic and market conditions	
House price appreciation (*HPA*)	House price appreciation is calculated as the difference of the natural logarithm of current HPI and that of HPI in previous quarter. The HPI is collected at the MSA level.
Unemployment rate (*Unemployment rate*)	Quarterly unemployment rate is collected at state level from the Bureau of Labour Statistics (BLS).
Real GDP growth rate (*Real growth rate*)	Real GDP growth rate is calculated from quarterly real GDP collected at state level from Bureau of Economic Analysis (BEA).
Time to end of observation (*TimeToEOO*)	Time gap between default events and the time of last available loss information. This variable controls for resolution bias.

Table 3: Relative frequency of control dummy variables in %

This table reports the relative frequency (%) of dummies used as control variables in both cured and non-cured loans. For a description of variables, we refer to Table 2.

	ARM	Owner Occ.	PDJ	SRR
Panel A: Defaulted loans				
0	26.8	13.1	66.4	89.0
1	73.2	86.9	33.7	11.0
Panel B: Cured loans				
0	38.8	11.1	74.3	92.1
1	61.2	88.9	25.7	7.9
Panel C: Non-cured loans				
0	21.8	13.9	63.0	87.7
1	78.2	86.1	37.0	12.3

Table 4: Descriptive statistics of continuous control variables

This table provides mean, standard deviation (Std.Dev), 5th percentile (P5) and 95th percentile (P95) of continuous variables regarding defaulted loans, cure loans and non-cure loans. For a description of the variables, we refer to Table 2.

Variable	Defaulted loans				Cure loans				Non-cure loans			
	Mean	Std.Dev	P5	P95	Mean	Std.Dev	P5	P95	Mean	Std.Dev	P5	P95
Liquidity constraints (%)	0.20	0.31	-0.36	1.03	0.28	0.36	0.00	1.03	0.17	0.28	-0.36	0.81
Equity	0.09	0.22	-0.40	0.44	0.15	0.23	-0.34	0.45	0.07	0.22	-0.40	0.37
FICO	646	67	528	756	643	71	520	757	648	66	532	756
Loan size	12.20	0.73	10.98	13.31	12.19	0.76	10.92	13.36	12.21	0.71	11.00	13.29
Current IR (%)	7.60	1.88	4.09	10.75	7.23	1.98	3.32	10.63	7.76	1.81	4.67	10.79
Unemployment rate (%)	7.12	2.52	3.80	11.70	7.30	2.50	3.90	11.60	7.05	2.53	3.80	11.70
HPA (%)	-1.97	3.53	-8.29	2.06	-0.93	3.03	-7.01	3.10	-2.40	3.63	-9.20	1.64
Real growth rate (%)	-0.12	1.37	-2.72	1.64	0.06	1.31	-2.22	1.80	-0.19	1.39	-2.93	1.62
TimeToEOO	6.19	1.85	2.50	8.75	5.56	2.36	1.00	9.00	6.46	1.52	3.50	8.75
Observations	509,408				150,095				359,313			

Table 5: Correlation matrix of continuous variables

This table provides pairwise correlations between continuous variables, cure rate and non-zero LGD. For a description of the variables, we refer to Table 2.

Variables	Cure	Non-zero LGD	Liquidity constraints	Home Equity	FICO	Loan size	Current IR	Unemployment rate	HPA	Real growth rate	TimetoEOO
Cure	1	-									
Non-zero LGD	-	1									
Liquidity constraints	0.16	-0.04	1								
Home equity	0.16	-0.22	0.18	1							
FICO	-0.03	-0.09	-0.14	-0.29	1						
Loan size	-0.02	-0.25	-0.21	-0.36	0.36	1					
Current IR	-0.13	0.2	-0.19	0.25	-0.4	-0.34	1				
Unemployment rate	0.05	0.09	0.13	-0.53	0.29	0.18	-0.38	1			
HPA	0.19	-0.19	0.25	0.39	-0.2	-0.28	-0.02	-0.09	1		
Real growth rate	0.08	-0.15	0.13	0.2	-0.1	-0.04	-0.05	-0.11	0.18	1	
TimeToEOO	-0.22	-0.03	-0.34	0.29	-0.2	-0.08	0.43	-0.58	-0.07	-0.03	1

Table 6: Relative frequencies for cure and non-cure by liquidity constraint and home equity categories

This table reports the relative frequency (%) of cure and non-cure loans for different ranges of borrower liquidity constraint and home equity. P5 denotes the 5th percentile and P95 denotes the 95th percentile.

Panel A: Borrower liquidity constraints					
	[P5, 0]	(0, 0.04]	(0.04, 0.08]	(0.08, 0.52]	[0.52, P95]
Non-cure	72.9	84.0	82.8	72.0	52.5
Cure	27.1	16.0	17.2	28.0	47.5
Panel B: Home equity					
	[P5, -0.2]	(-0.2, 0]	(0, 0.1]	(0.1, 0.2]	(0.2, P95]
Non-cure	77.8	74.6	76.0	75.0	60.9
Cure	22.2	25.4	24.0	25.0	39.1

Table 7: Estimates for probability of cure equation

This table reports the estimation outputs of the probability of cure equation with no control and with control variables for *two* samples, the Full sample (Q2/2005-Q1/2015) and the Crisis subsample (Q3/2007-Q2/2009). For a description of the variables, we refer to Table 2. FICO is divided by 1,000; Loan size, Current IR and TimeToEOO are divided by 10; Unemployment rate is a percentage; Real growth rate and HPA are in their levels. Standard errors are reported in parentheses. ***, ** and * denote that the estimates are statistically significant at 1%, 5% and 10% level. AUROC denotes the Area Under the Receiver Operating Characteristics curve, which is a popular measure for discrimination in credit risk. We do not report the estimated intercepts of these equations.

Parameter	**Full sample**		**Crisis sub-sample**	
	No control	**With controls**	**No control**	**With controls**
Borrower liquidity constraint				
Liquidity constraints	0.265***	-0.277***	-0.08***	-0.328***
	(0.025)	(0.026)	(0.031)	(0.032)
Liquidity constraint$_{s0}$	-12.556***	-7.441***	-8.93***	-5.555***
	(0.381)	(0.398)	(0.469)	(0.492)
Liquidity constraint$_{s0.04}$	17.128***	12.443***	12.268***	9.224***
	(0.763)	(0.789)	(0.939)	(0.968)
Liquidity constraint$_{s0.08}$	-3.523***	-4.592***	-2.332***	-3.245***
	(0.407)	(0.418)	(0.511)	(0.523)
Liquidity constraint$_{s0.52}$	-0.931***	-0.283***	-0.856***	-0.031
	(0.036)	(0.038)	(0.072)	(0.074)
Home equity				
Home equity	0.879***	1.325***	0.618***	0.864***
	(0.058)	(0.06)	(0.093)	(0.094)
Home equity$_{s-0.2}$	-0.98***	-1.351***	-0.875***	-0.761***
	(0.104)	(0.108)	(0.157)	(0.161)
Home equity$_{s0}$	0.194	0.565***	1.171***	1.098***
	(0.138)	(0.144)	(0.194)	(0.199)
Home equity$_{s0.1}$	-0.406**	-0.546***	-0.436*	-0.86***
	(0.16)	(0.166)	(0.227)	(0.232)
Home equity$_{s0.2}$	4.079***	4.025***	2.887***	3.179***
	(0.102)	(0.106)	(0.154)	(0.159)
Controls				
FICO		-1.147***		-0.659***
		(0.037)		(0.057)
Current IR		-0.482***		-0.446***
		(0.014)		(0.024)
Loan size		3.822***		-0.226
		(0.687)		(1.16)
Loan size square		-0.793***		0.846*
		(0.281)		(0.472)
PDJ		-0.231***		-0.236***
		(0.005)		(0.008)
SRR		-0.214***		-0.184***
		(0.007)		(0.011)
ARM		-0.297***		-0.356***
		(0.005)		(0.007)
Owner occupied		0.119***		0.099***

		(0.006)		(0.009)
TimeToEOO		-1.56***		-3.282***
		(0.015)		(0.094)
HPA		4.322***		1.28***
		(0.07)		(0.098)
Real growth rate		1.357***		1.172***
		(0.152)		(0.219)
Unemployment rate		-0.005***		-0.004
		(0.001)		(0.003)
Log Likelihood	-361,263	-288,257	-152,317	-115,612
Pseudo R^2	10.4%	21.7%	6.1%	12.7%
AUROC	0.66	0.741	0.63	0.70
Number of Observations	509,408		259,815	

Table 8: Estimates for non-zero loss equation

This table reports the estimation outputs of the non-zero loss equation with no control and with control variables for *two* samples, the Full sample (Q2/2005-Q1/2015) and the Crisis subsample (Q3/2007-Q2/2009). For a description of the variables, we refer to Table 2. FICO is divided by 1,000; Loan size, Current IR and TimeToEOO are divided by 10; Unemployment rate is in percentage; Real growth rate and HPA are in their level (i.e., in absolute terms). Standard errors are reported in parentheses. ***, ** and * denote that the estimates are statistically significant at 1%, 5% and 10% level. We do not report the estimated intercepts of these equations.

Parameter	**Full sample**		**Crisis sub-sample**	
	No control	**With controls**	**No control**	**With controls**
Borrower liquidity constraint				
Liquidity constraints	-0.042***	-0.01**	0.018***	0.01**
	(0.006)	(0.005)	(0.006)	(0.005)
Liquidity constraint$_{s0}$	2.969***	1.112***	2.572***	0.782***
	(0.076)	(0.067)	(0.081)	(0.072)
Liquidity constraint$_{s0.04}$	-2.835***	-1.658***	-2.232***	-1.238***
	(0.15)	(0.131)	(0.16)	(0.14)
Liquidity constraint$_{s0.08}$	-0.326***	0.443***	-0.535***	0.375***
	(0.081)	(0.071)	(0.088)	(0.077)
Liquidity constraint$_{s0.52}$	0.193***	0.162***	0.277***	0.118***
	(0.01)	(0.008)	(0.016)	(0.014)
Home equity				
Home equity	-0.201***	-0.332***	-0.169***	-0.217***
	(0.013)	(0.011)	(0.017)	(0.015)
Home equity$_{s-0.2}$	0.12***	0.24***	0.001	0.052**
	(0.023)	(0.021)	(0.029)	(0.025)
Home equity$_{s0}$	-0.121***	-0.227***	-0.042	-0.149***
	(0.031)	(0.027)	(0.036)	(0.031)
Home equity$_{s0.1}$	-0.5***	-0.222***	-0.399***	-0.329***
	(0.037)	(0.032)	(0.043)	(0.037)
Home equity$_{s0.2}$	0.268***	-0.302***	0.124***	-0.27***
	(0.026)	(0.023)	(0.033)	(0.029)
Controls				
FICO		-0.228***		-0.223***
		(0.008)		(0.01)
Current IR		0.187***		0.146***
		(0.003)		(0.004)
Loan size		-15.296***		-14.653***
		(0.162)		(0.218)
Loan size square		5.677***		5.36***
		(0.066)		(0.089)
JP		0.119***		0.128***
		(0.001)		(0.001)
PDJ		0.042***		0.027***
		(0.001)		(0.002)
SRR		0.037***		0.023***
		(0.001)		(0.002)
ARM		0.018***		0.011***
		(0.001)		(0.001)
Owner occupied		-0.076***		-0.07***

		(0.001)		(0.002)
TimeToEOO		-0.016***		1.017***
		(0.005)		(0.017)
HPA		-1.374***		-1.044***
		(0.014)		(0.016)
Real growth rate		-1.127***		-1.25***
		(0.03)		(0.038)
Unemployment rate		0.006***		0.024***
		(0.000)		(0.000)
σ_{ε}	0.271***	0.235***	0.251***	0.217***
	(0.000)	(0.000)	(0.000)	(0.000)
$\rho_{u,\varepsilon}$	-0.02	0.035***	-0.008	-0.056***
	(0.012)	(0.013)	(0.016)	(0.015)
Log Likelihood	-361,263	-288,257	-152,317	-115,612
Adjusted R^2	8.1%	29.6%	8.3%	30.3%
Number of Observations	509,408		259,815	

Table 9: Estimated results for the samples partitioned for liquidity constraint and negative equity

This table reports the estimation outputs of the probability of cure equation (PC) and non-zero loss equation ($\overline{LGD^*}$) with control variables for samples partitioned by liquidity constraint (i.e., *Liquidity constraint* > 0) and negative equity (i.e., *CLTV* >1). *Pure NE sample* includes the defaulted loans that only experienced negative equity but no liquidity constraints in previous quarter. *LC and NE sample* includes the defaulted loans that experienced both liquidity constraints and negative equity in the previous quarter. *No LC and NE sample* includes the defaulted loans that did not experience any of liquidity constraints and negative equity in the previous quarter. *Pure LC sample* includes the defaulted loans that only experienced liquidity constraints but not negative equity in the previous quarter. For a description of the variables, we refer to more details in Table 2. FICO is divided by 1,000; Loan size, Current IR and TimeToEOO are divided by 10; Unemployment rate is in percentage; Real growth rate and HPA are in their level (i.e., in absolute terms). Standard errors are reported in parentheses. ***, ** and * denote the estimates are statistically significant at 1%, 5% and 10% level. For brevity, we do not report the estimations of intercepts and control variables. The set of control variables is the identical and estimates consistent with those presented in Table 7 and 8.

Parameter	Pure NE sample		LC and NE sample		No LC and NE sample		Pure LC sample	
	PC	$\overline{LGD^*}$	PC	$\overline{LGD^*}$	PC	$\overline{LGD^*}$	PC	$\overline{LGD^*}$
Borrower liquidity constraint								
Liquidity constraint$_{(\leq 0)}$	-0.067*	-0.04***			-0.502***	-0.057***		
	(0.038)	(0.005)			(0.041)	(0.008)		
Liquidity constraint$_{s0}$			-6.15***	-0.399*			-2.141*	-0.303
			(1.473)	(0.215)			(1.159)	(0.236)
Liquidity constraint$_{s0.04}$			7.911***	0.437			6.565***	-0.079
			(2.074)	(0.292)			(1.49)	(0.302)
Liquidity constraint$_{s0.08}$			-1.531*	-0.074			-4.36***	0.267***
			(0.823)	(0.111)			(0.5)	(0.1)
Liquidity constraint$_{s0.52}$			-0.608***	0.105***			-0.176***	0.151***
			(0.076)	(0.013)			(0.044)	(0.012)
Home equity								
Home equity$_{(\leq 0)}$	1.342***	-0.344***	1.139***	-0.237***				
	(0.085)	(0.013)	(0.088)	(0.014)				
Home equity$_{s-0.2}$	-1.083***	0.116***	-0.778***	-0.02				
	(0.169)	(0.023)	(0.167)	(0.027)				
Home equity$_{s0}$					0.939***	-0.144***	0.445**	-0.299***

					(0.22)	(0.042)	(0.178)	(0.038)
Home equity$_{s0.1}$					-1.543***	-0.185***	-0.047	-0.202***
					(0.331)	(0.063)	(0.257)	(0.056)
Home equity$_{s0.2}$					4.152***	-0.415***	3.56***	-0.383***
					(0.196)	(0.043)	(0.133)	(0.034)
Intercept	Yes		Yes		Yes		Yes	
Controls	Yes		Yes		Yes		Yes	
σ_{ε}	0.162***		0.204***		0.233***		0.27***	
	(0.001)		(0.001)		(0.001)		(0.001)	
$\rho_{u,\varepsilon}$	0.03		0.828***		-0.011		-0.008	
	(0.067)		(0.006)		(0.028)		(0.02)	
Log Likelihood	-16,167		-21,174		-63,951		-165,566	
Pseudo/Adjusted R^2	12.1%	29.9%	23%	34.2%	17.7%	24%	25.6%	29.2%
Number of Observations	73,802		72,786		115,601		247,219	

Figure 1: Distribution of Loss Given Default
This figure shows the bimodal LGD distribution with a large fraction of defaulted loans that do not generate losses (zero LGDs or cures). About 30% of defaulted loans are cured.

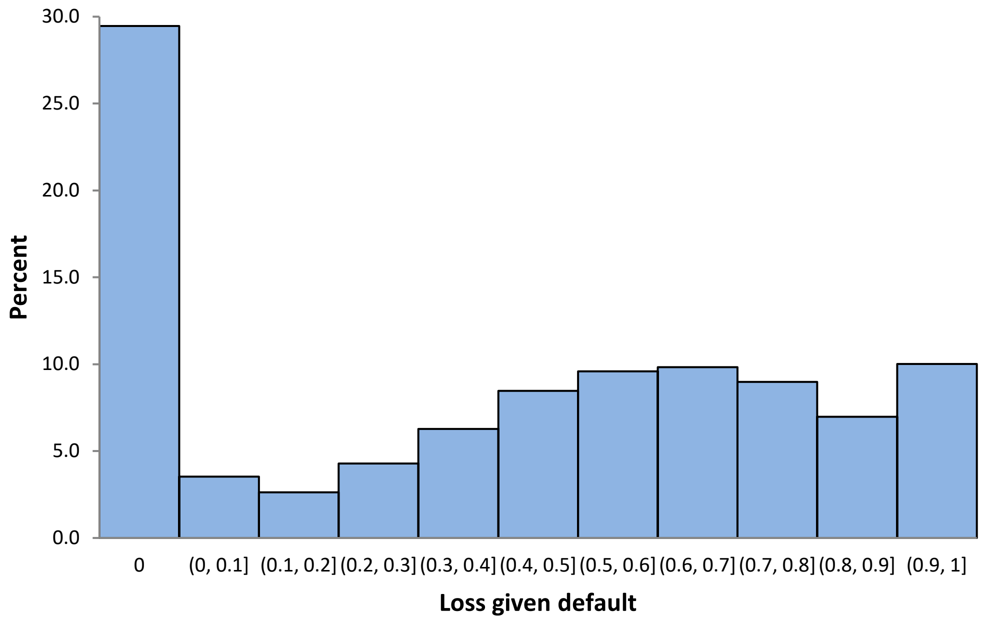


Figure 2: Selection mechanism for cure and non-zero LGD
This figure shows the selection mechanism for cure and non-zero LGD that our model is based on. The mechanism indicates that the cure events are observed if loan *i* defaults, while the non-zero LGD can only be observed if the defaulted loan *i* is non-cured. This mechanism is implemented by a joint probability framework between cure and non-zero losses for modelling purposes in this paper.

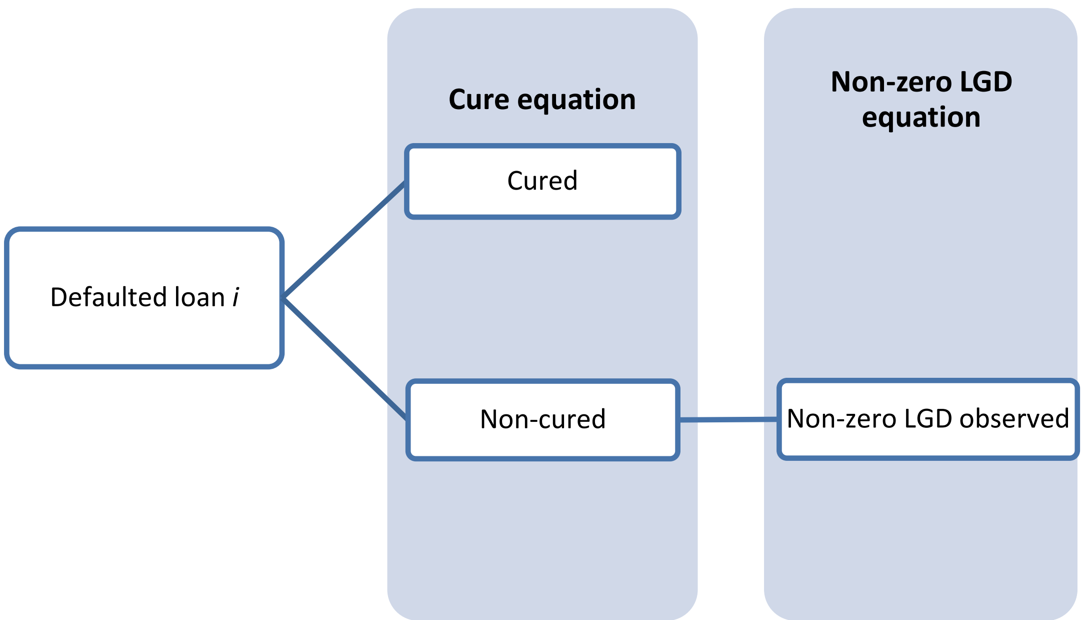


Figure 3: Distribution of time to resolution

This figure displays distribution of time to resolution. It shows that most non-zero LGDs are resolved within 3 years after default. The highest frequency of non-zero LGDs is observed one year after the default.

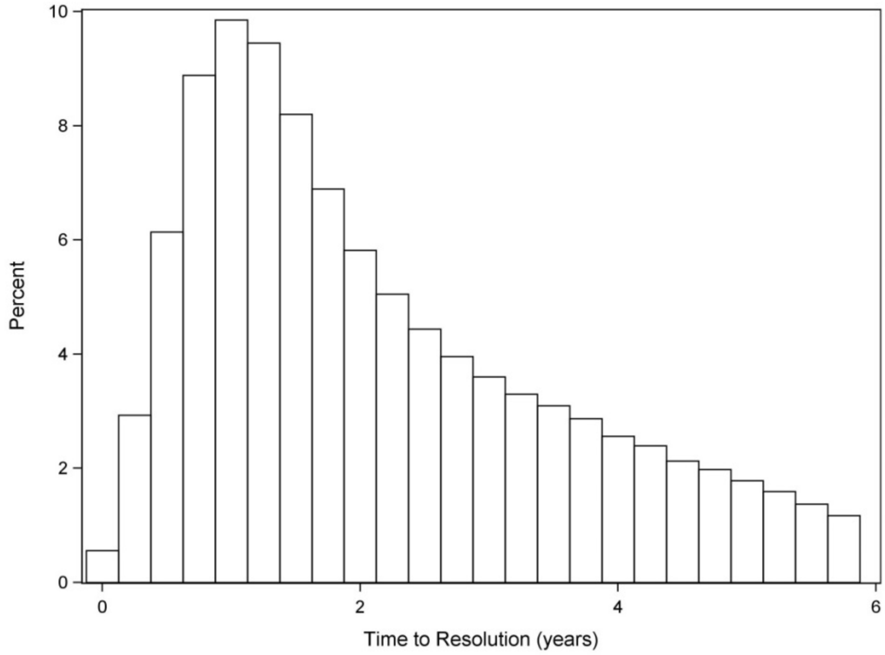


Figure 4: Non-cure rate and Loss Given Default by time

This figure shows average non-cure rate, average LGD and average non-zero LGD per quarter. The grey area displays the gap between average non-zero LGD and average LGD. The three series tend to co-move over time with different degrees. The gap shows a time varying behaviour which is primarily due to changes of the cure rate.

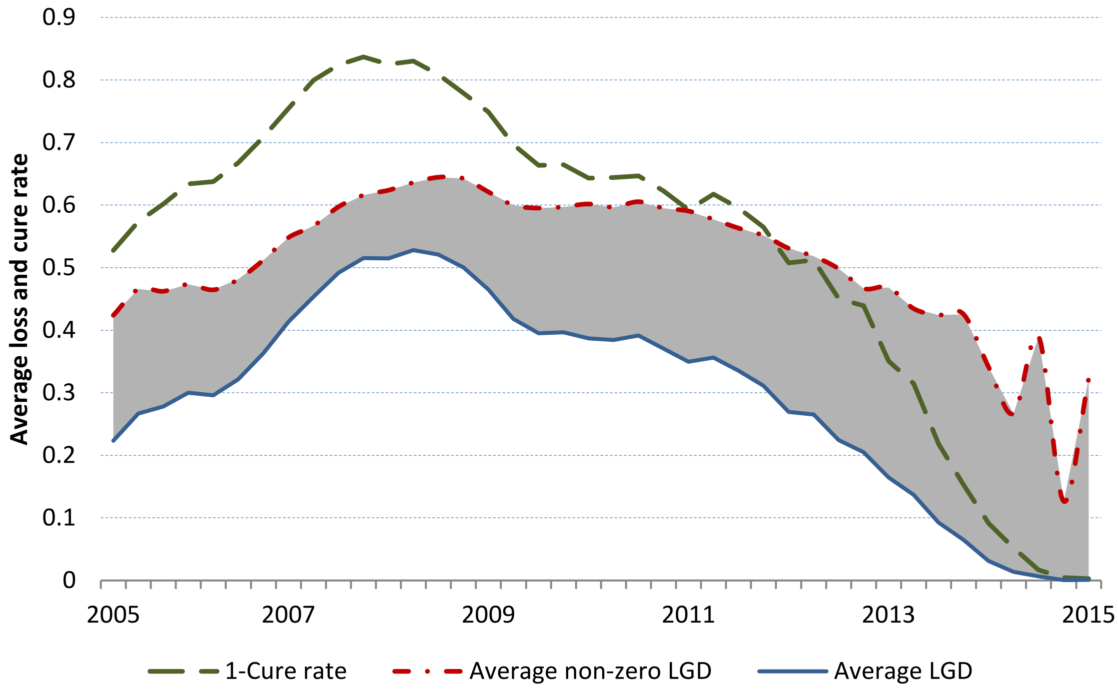


Figure 5: Average cure rate by ranges of liquidity constraint

This figure shows average cure rates by ranges of borrower liquidity constraints, which is marked by their mid-points on the left vertical axis and the number of observations in each range on the right vertical axis. The dashed bounds on each bar represent the 95% confidence interval for the cure rate.

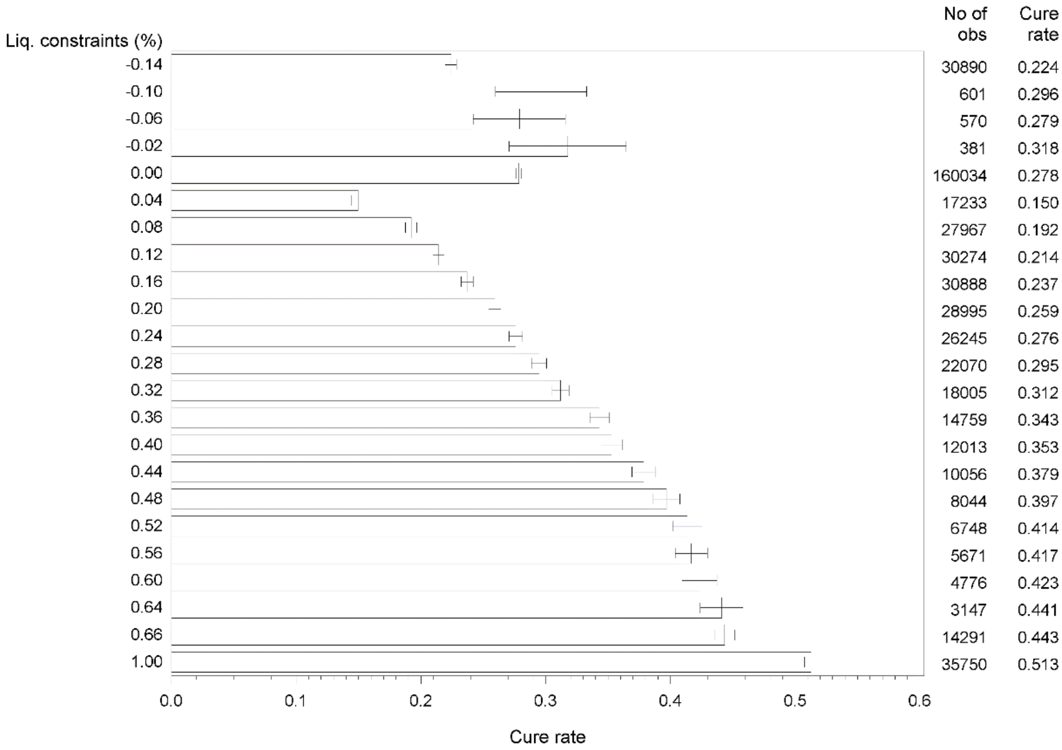


Figure 6: Average cure rate by ranges of home equity

This figure shows the average cure rate by ranges of home equity, which is marked by their mid-points on the left vertical axis and the number of observations in each range on the right vertical axis. The dashed bounds on each bar represent the 95% confidence interval for the cure rate.

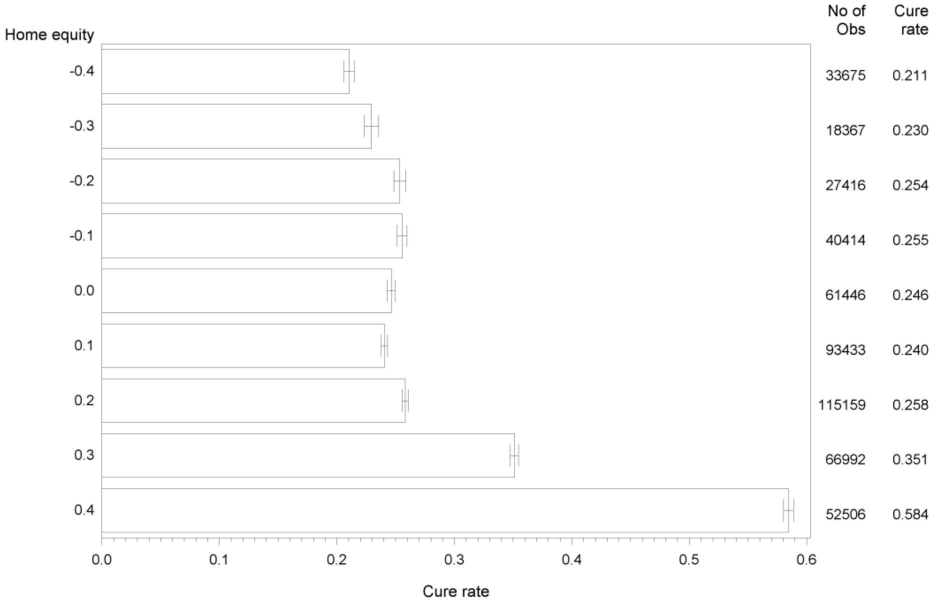


Figure 7: Average unemployment rate, borrower liquidity constraint and home equity by time

This figure shows the average unemployment rate, borrower liquidity constraint and home equity by time. The time varying behaviour of unemployment rate is negatively related with home equity, while the borrower liquidity constraint co-moves with home equity in some periods. There is a time-delayed reflection of the impact of the unemployment rate on the borrower liquidity constraint by one year.

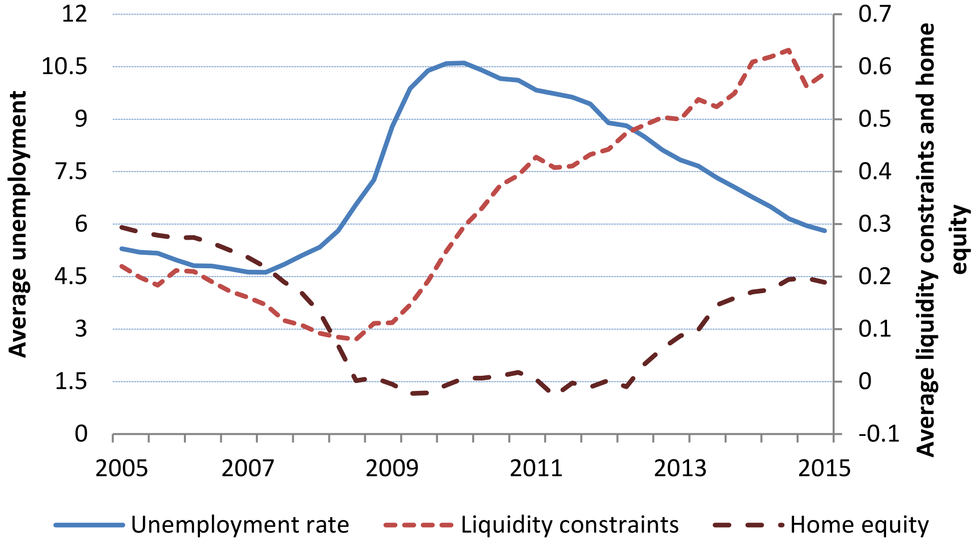


Figure 8: Scatter plots of average liquidity constraint, cure rate and average non-zero LGD by time

This figure shows scatter plots of average liquidity constraint, cure rates by time and average non-zero LGD by time. The first two plots relate to the cure rate and the second two plots to the non-zero LGD. The cure rate and non-zero LGD is contemporary for the first and third plot and lagged by four periods for the second and fourth plot. The ellipses show 95% confidence interval and the solid lines are the regression lines.

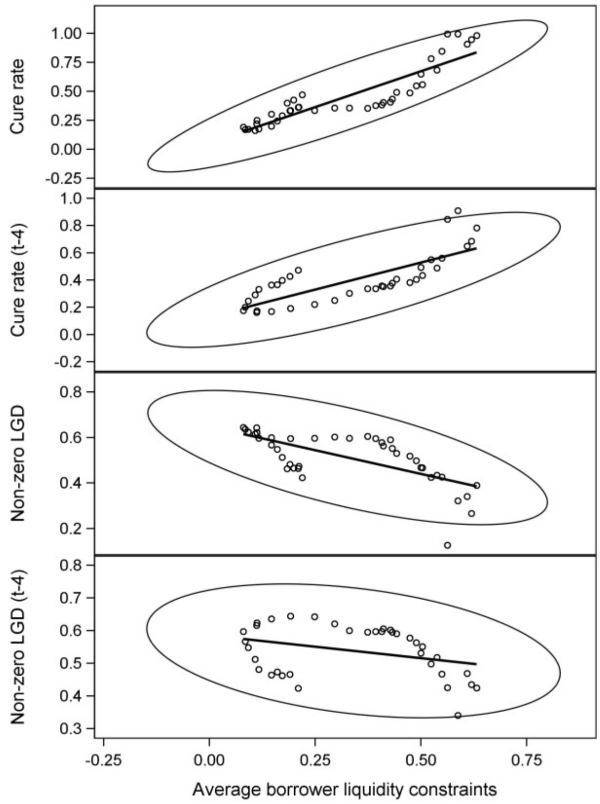


Figure 9: Relationship between liquidity constraint, probability of cure and non-zero LGD

This figure depicts the relationships between borrower liquidity constraint, probability of cure and non-zero LGD for the estimated parameters from Table 7 and 8.

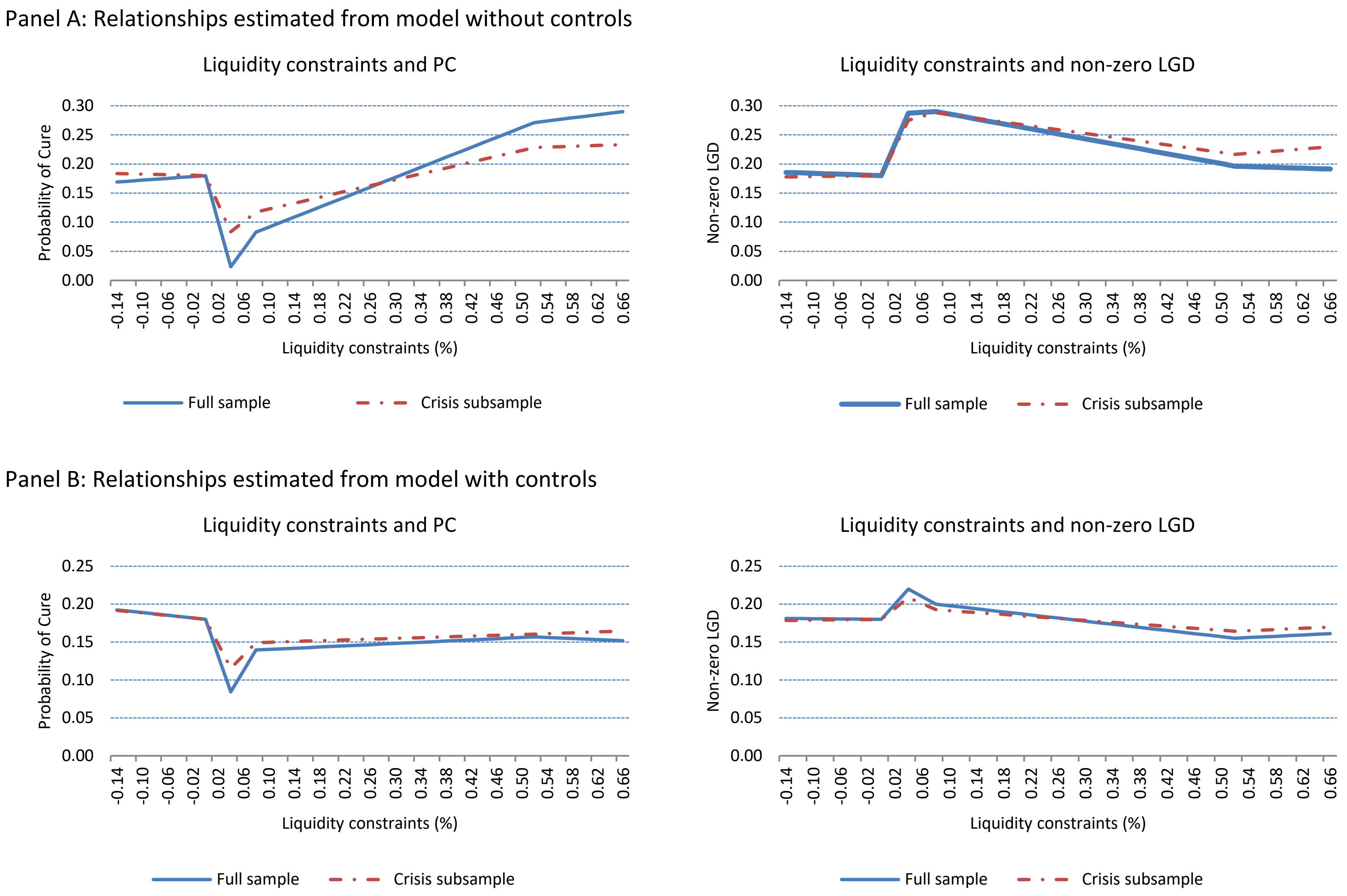


Figure 10: Relationship between liquidity constraint, probability of cure and non-zero LGD for partitioned samples

This figure shows the relationships between borrower liquidity constraints, probability of cure and non-zero LGD for four partitioned samples that distinguish positive and negative liquidity constraints on the horizontal direction, positive and negative equity on the vertical direction. The relationships are visualized for the estimated parameters from Table 9.

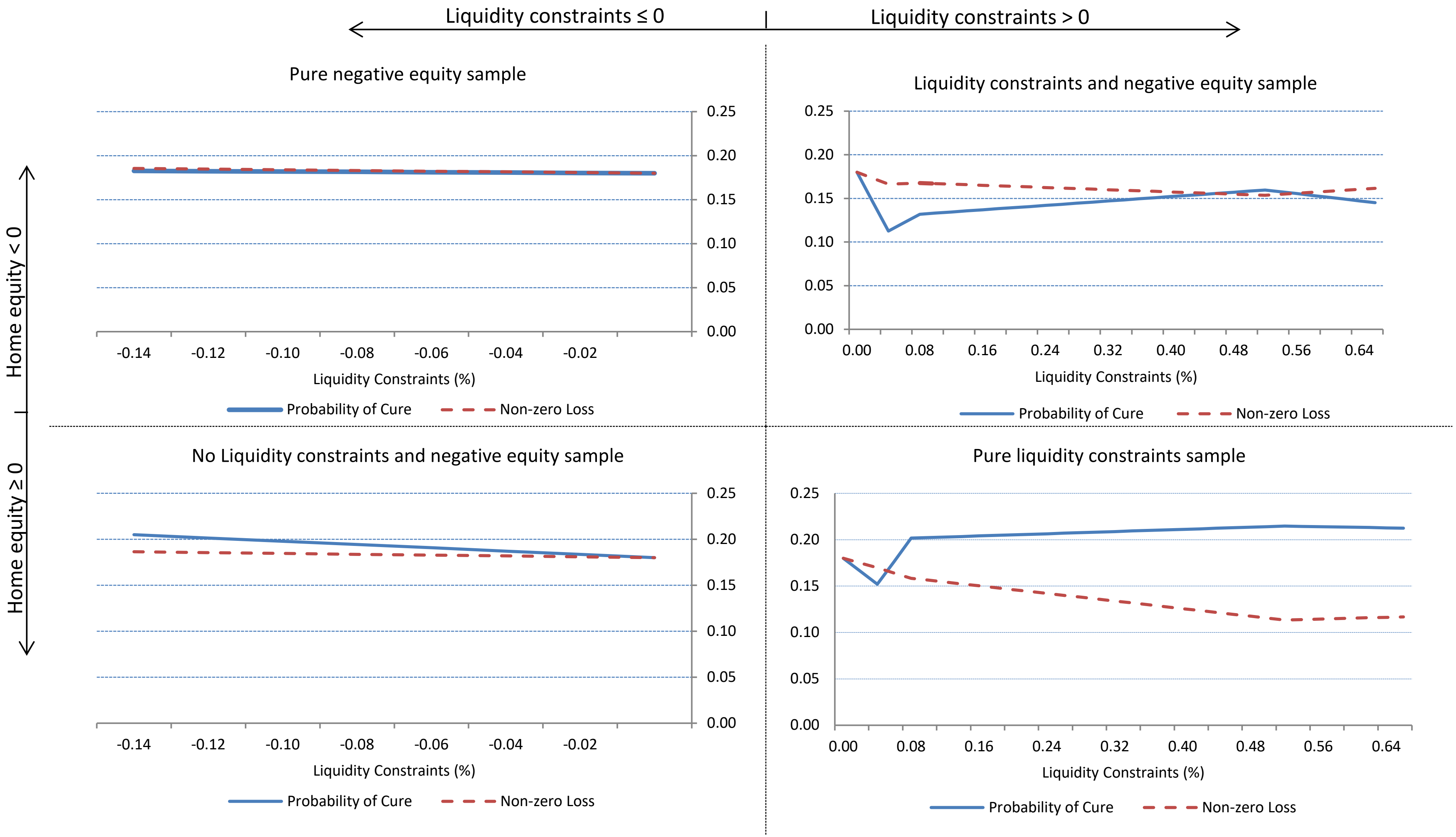
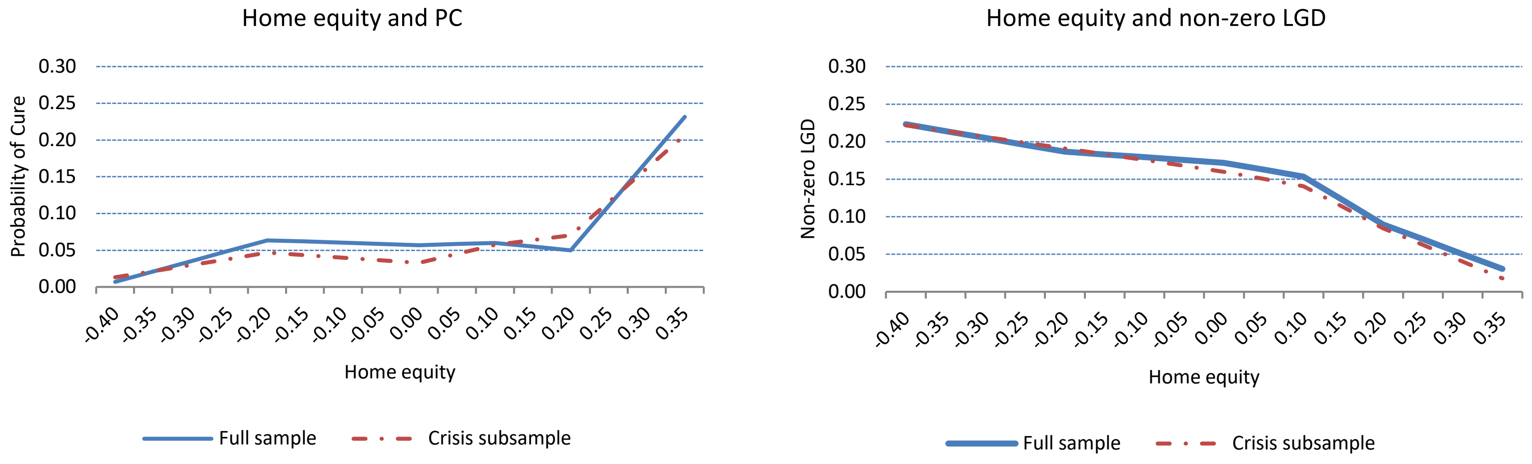


Figure 11: Relationship between home equity, probability of cure and non-zero LGD

This figure depicts the relationships between home equity, probability of cure and non-zero LGD for the estimated parameters from Table 7 and 8.

Panel A: Relationships estimated from model without controls



Panel B: Relationships estimated from model with controls