

Article

A Multi-State Approach to Modelling Intermediate Events and Multiple Mortgage Loan Outcomes

Richard Chamboko ^{1,2,3,*}  and Jorge Miguel Bravo ^{1,4} 

¹ NOVA IMS—Information Management School, New University of Lisbon, 1070-312 Lisbon, Portugal; jbravo@novaims.unl.pt

² International Finance Corporation, World Bank Group, Washington, DC 20433, USA

³ Institute for Intelligent Systems, University of Johannesburg, Johannesburg 2006, South Africa

⁴ Department of Economics, Université Paris-Dauphine PSL, 75775 Paris, France

* Correspondence: richardchamboko@gmail.com

Received: 21 March 2020; Accepted: 2 June 2020; Published: 10 June 2020



Abstract: This paper proposes a novel system-wide multi-state framework to model state occupations and the transitions among current, delinquency, default, prepayment, repurchase, short sale and foreclosure on mortgage loans. The approach allows for the modelling of the progression of borrowers from one state to another to fully understand the risks of a cohort of borrowers over time. We use a multi-state Markov model to model the transitions to and from various states. The key factors affecting the transition into various loan outcomes are the ability to pay as measured by debt-to-income ratio, equity as marked by loan-to-value ratio, interest rates and the property type. Our findings have broader policy implications for better decision-making on granting loans and the design of debt relief and mortgage modification policies.

Keywords: credit risk; survival analysis; multi-state models; delinquency; recovery; relief programs; mortgage modification

JEL Classification: C01; C41; C53; C58; E51; G21

1. Introduction

Housing wealth is typically the largest component of wealth for many households and mortgages are their main source of credit (Campbell 2013). During the recent global financial crises, mortgage debt triggered a wave of foreclosures that impacted household consumption, balance sheets as well as the transmission of monetary policy to the real economy (Mian et al. 2015; Di Maggio et al. 2017). Consequently, households in many economies are still struggling to escape the aftermath (Mesnard et al. 2016; Ozkan and Unsal 2012; Whelan 2013). Consumer welfare has been jeopardised, and the ability of consumers to meet their loan obligations has been severely affected. As such, many economies experience rising impaired loans, which are choking the financial sector and the overall performance of their economies (Whelan 2013).

The critical role of the mortgage market in triggering the recent global financial crisis has increased policy interest, bank regulation and academic research in this area. The banking regulatory framework changes brought by the revised Basel Committee on Banking Supervision (BCBS) Accords (later adopted by national legislation in many countries and regions, for instance, the European Capital Requirement Directives and the US Regulatory Capital Rules) introduced stronger risk management requirements for banks, with capital requirements tightly coupled to estimated credit portfolio losses. The recently adopted IFRS9 and FASB's Current Expected Credit Loss (CECL) standards introduce revised expected credit loss or impairment calculation rules requiring financial institutions to calculate

expected loss for the banking book over the entire life of the exposures. Encouraged by regulators, banks devoted significant resources to develop an Internal Ratings Based approach (IRB) for the calculation of risk weighted assets for credit risk to better support decisions when granting loans, to quantify expected credit losses and to assign the mandatory economic capital.¹ Other than for banks and lenders, rigorous credit risk analysis in this area is also of importance for sound economic policy making (Kelly and O'Malley 2016) and for the design of social insurance programs (Bhutta et al. 2017) as evidenced in many countries and US states now enforcing laws to protect mortgage borrowers and mutualising some of the costs of default.

We argue that, in distressed economic environments, the focus should go beyond modelling defaults and foreclosures as the only main outcomes and should explicitly allow for other loan-level transitions both into and out of default together with other state occupation experienced by obligors. For instance, it is of paramount importance and informative for policy to understand the progression of mortgagors from normal performance to delinquency, and subsequently to default and foreclosure as well as the cure from delinquency or default to normal performance. For the US market as an example, understanding of cures is particularly important to gauge if strategies implemented by lenders and other policy interventions such as the US Treasury Department's Home Affordable Refinance Program (HARP) and the Federal Housing Administration's Home Affordable Modification Program (FHA-HAMP) were effective in cutting down losses and preventing ruthless default by underwater mortgagors (Tracy and Wright 2016; Agarwal et al. 2015, 2017; Liu and Sing 2018). Empirical evidence suggests that these debt relief programs had mixed success in preventing foreclosure. In addition, the transition into other competing and absorbing states such as repayment, repurchase or short sale is of great importance as they are significant competing risks that can affect the profitability and solvency of lending institutions. For banks and other lenders, the accurate estimation of these remain essential for pricing, interest and liquidity risk management, profit forecasting, managing delinquents and capital provisioning.

Motivated by the fixed income market transition-based credit risk assessment methods such as the McKinsey's CreditPortfolioView (McKinsey and Company 1998) and JP Morgan's Creditmetrics (Gupton et al. 1997), this paper contributes to the literature by offering a novel system-wide framework for analysing mortgage loans which exposes insights which cannot be generated by the traditional approaches. The approach allows for the modelling of the progression of borrowers from one state to another to fully understand the risks of a cohort of borrowers over time. The framework has seven allowable states (current, delinquent, default, prepayment, foreclosure, repurchase and short sale) and sixteen possible transitions to jointly model state occupations and the transitions among current, delinquency, default, prepayment, repurchase, short sale and foreclosure on mortgage loans. The framework thus combines aspects of credit risk (delinquency, default and foreclosure and self-cure/recovery) with aspects of interest rate and liquidity risk (prepayment, short sale and self-cure/recovery). A multi-state Markov model is used to model the transitions. The study thus goes beyond modelling default or foreclosure as the only risk on mortgages or the competing risks of prepayment and foreclosure to include transient states of delinquency to allow the modelling of recovery or cure from distress as well as introducing additional absorbing states of repurchase and short sale into the multi-state framework. Using this framework, we also investigate the relationship between the probability of loans transitioning to and from various loan outcomes and loan-level covariates. We empirically test the performance of the model using the US single-family mortgage loans originated during the first quarter of 2009 and were followed on their monthly repayment performance until the third quarter of 2016.

¹ The recently approved BCBS (Basel IV) reforms of the standardised (CR-SA) approach (by making it more granular and risk sensitive) and of the CR-IRB approach for the calculation of risk weighted assets for credit risk will limit the extent to which banks can reduce capital requirements through the use of internal models (e.g., by eliminating the option to use any IRB approach for equity and advanced CR-IRB for institutions and large corporations).

Our findings thus have broader policy implications for contract design, lender and borrower behaviour analysis, for the mitigation of defaults and foreclosures through the design of debt relief programs and mortgage modification policies and for the design of laws for protecting distressed borrowers. The rest of the paper is structured as follows. Section 2 reviews the literature on commonly applied methods and the factors affecting repayment of mortgage loans whilst Section 3 presents the methodological approaches used in the study. In Section 4, the statistical analysis and results are presented. Section 5 provides a summary of the study and concludes.

2. Literature Review

Many studies have looked at factors affecting mortgage repayment with most of them focusing on default or foreclosure especially in the USA real estate market and a sizable number on the UK market. Taking mortgage default as a (put) option, early literature used the [Black and Scholes \(1973\)](#) pioneered contingent claims framework. Using this approach, the key drivers of default were home values and interest rates ([Gerardi et al. 2013](#)). [Riddiough \(1991\)](#) provided early insights on the modelling of “trigger events” such as job loss, health shocks, divorce and other accidents. Similarly, [Kau et al. \(1993\)](#); [Deng et al. \(1996\)](#) and many other researchers assessed effects of these trigger events on default and foreclosure and produced mixed findings on the factors which matters the most. [Schwartz and Torous \(1993\)](#) reported loan vintage and housing index returns volatility as the key drivers of observed default behaviour. [Deng et al. \(2000\)](#) argued that negative events such as job losses and divorces were significant predictors of mortgage default. Using data on mortgages originated between 2003 and 2007, [Mayer et al. \(2009\)](#) found unemployment and house prices as the key predictors of delinquency in the USA market.

In response to the mortgage default and foreclosure crises which began in 2007, an increased number of researchers analysed and documented numerous factors as the determinants of the observed default and foreclosure behaviour. One of the key hypotheses regarding the causes of mortgage delinquency is that homeowners will not continue servicing a mortgage if they enter into negative equity, for instance, if the value of the property drops below the mortgage value ([Kau et al. 1992](#); [Kelly and O'Malley 2016](#)). In this approach, a mortgage is seen as an American option with strike price equal to the value of the mortgage and the property being the underlying asset. It is assumed that the borrower will default as soon as the option enters into the in-the-money zone, that is when the property value falls below the mortgage value. This was also referred to as ruthless or strategic default, a term used to describe a borrower in negative equity who chooses to default despite having enough financial resources to continue servicing the mortgage ([Gerardi et al. 2013](#)).

[Chan et al. \(2014\)](#) found that loan and individual characteristics such as borrowers credit history, current loan-to-value, race, ethnicity and income are key drivers of foreclosure. [Guiso et al. \(2009\)](#) found severe negative equity, gender, future employment expectations, race and morality as key determinants of ruthless or strategic default. Long term unemployment as well as falling home prices which led to negative equity were also found to be key drivers of observed mortgage defaults, foreclosures and housing vacancies ([Jones et al. 2016](#); [Tian et al. 2016](#)).

[Foote et al. \(2008\)](#) also assessed this concept of negative equity on mortgage default decisions and found that some mortgagors who were in negative equity did not default and argued that this could partly be explained by price expectations. Consistent with that, [Foote et al. \(2008\)](#) also indicated that mortgagors who were in negative equity and defaulted could have done so not only because of negative equity but due to a “double trigger” effect (negative equity combined with some adverse event such as loss of employment, health issues, death of spouse, divorce, etc.). In agreement with the double trigger hypothesis, numerous other studies also documented that mortgagors could be in negative equity and still not default ([Elul et al. 2010](#); [Bhutta et al. 2010, 2017](#)).

[Bhutta et al. \(2017\)](#) estimated the level of indebtedness and negative equity that triggers ruthless default on mortgages using U.S. data from 2007 to 2009 and found that, for most homeowners, the equity has to be deeply negative before they take the default option. [Ahlawat \(2018\)](#) also supported

the argument that many mortgage defaults are non-ruthless and that transaction costs are the key driver of the mortgage default decision. Consistently, other studies which characterise mortgage default as non-ruthless argue that transaction costs (for instance, legal fees, relocation costs and credit impairment costs) and other idiosyncratic factors are important factors driving mortgage default decisions (Vandell 1995; Ahlwat 2018).

Using UK data, Aron and Muellbauer (2010) concluded that experiencing negative equity is just one of the fundamental economic drivers of payment delinquency, along with the debt service ratio and the unemployment rates. Danis and Pennington-Cross (2005) used a two-step procedure and a seemingly unrelated bivariate probit model of mortgage outcomes to estimate probabilities of prepaid and default. The authors concluded that very delinquent loans are more likely to prepay than to default and that prepayment rates increase substantially as delinquency intensity increases. Again, using UK data, Aron and Muellbauer (2016) found that the aggregate debt-service ratio, the proportion of mortgages in negative equity and the unemployment rate have significant effects on aggregate rates of repossessions and arrears. Tian et al. (2016) documented that household and local unemployment rates were key drivers of mortgage defaults. Carranza and Estrada (2013) found house prices and debt balances as the main drivers of mortgage default in Colombia.

Another key factor explaining mortgage delinquency is the inability to reimburse. This is likely to happen when a reduction in disposable income, often triggered by unemployment spells or family events (for instance, divorce or death of a spouse) reduces the capacity to continue servicing the mortgage. Gerardi et al. (2007); Elul et al. (2010); Bajari et al. (2008); Fuster and Willen (2017) and Bhutta et al. (2010) also emphasised the role of cash flow problems or illiquidity as an important factor explaining the inability to continue paying a mortgage loan as someone who is highly illiquid may not be able to find the cash to make the loan payment and may find it costly to wait for house prices to recover. Fuster and Willen (2017) documented the effect of payment size as an important determinant of default and cure (recovery from delinquency) stating as high as 40% reduction in default rates due to a 2% reduction in interest rates. Similarly, they found a 75% increase in the cure hazard caused by a 2–2.5% reduction in interest rates. Tracy and Wright (2016) found that a reduction in monthly payment under the HARP reduced losses on credit by about 56 basis points. Bajari et al. (2013) studied the relative importance of different drivers of default and concluded that principal write downs have a huge impact on borrowers' default behaviour and welfare.

Apart from the characteristics of borrowers and loans, recent research on the severity of the housing slump in the U.S. during the recession suggests that several features related to both rigidity of mortgage contracts and market frictions hampered public and private efforts to restructure or refinance households, augmenting the incidence of costly foreclosures (for instance, Piskorski et al. 2010; Di Maggio et al. 2017; Piskorski and Seru 2018; Fuster and Willen 2017). Contract rigidity, particularly the fact that most mortgage contracts were locked at high interest rates given that they were fixed rate mortgages (FRMs) prevented borrowers from receiving an automatic debt relief compared to Adjustable Rate Mortgages (ARMs) featuring no interest rate floors, increasing the likelihood of households to be delinquent on their loans. Other important documented market frictions include: (i) equity refinancing constraints (getting debt relief through refinancing of FRMs mortgages being impossible for households due insufficient equity to meet the LTV requirements mainly due to a drop in home prices); (ii) intermediary organisational constraints and lack of effective competition especially within the refinancing market resulting in limited debt relief or refinancing; (iii) agency conflicts in servicing of largely securitised mortgages; and (iv) moral hazard concerns in that by offering debt relief to distressed borrowers, many solvent borrowers could stop making payments to enjoy similar benefits (Piskorski and Seru 2018).

In terms of methods, several approaches are used for consumer credit risk assessment. The work of Altman (1968) pioneered this area with the Z score discriminant analysis model. Today, the logistic regression has become the standard for the industry (Crook et al. 2007; Noh et al. 2005; Lessmann et al. 2015). Bajari et al. (2008) developed a US sub-prime market scoring model using a

bivariate probit model allowing borrowers to default either because the mortgage to equity ratio goes above a certain value (due to, for instance, falling home prices) such that by defaulting, borrowers tend to increase their lifetime wealth or due to insufficient income and/or lack of access to other forms of credit. Many studies to date have focused on improving the default prediction accuracy by considering both traditional statistical methods and more sophisticated (e.g., advanced machine learning) modelling approaches and alternative sets of predictor features. Discriminant analysis, support vector machines, artificial neural networks, decision trees, genetic programming and standard models using external ratings provided by external credit assessment institutions have also been successfully applied (Arminger et al. 1997; Hand and Henley 1997; Kruppa et al. 2013; Lessmann et al. 2015; Butaru et al. 2016; Baesens et al. 2003; Abellán and Castellano 2017). Less attention has been given to modelling default as a dynamic process even if empirical studies have shown that models that account for the dynamism in default may be conceptually more appropriate and can result in better results (Du Jardin and Séverin 2011; Volkov et al. 2017). Grimshaw and Alexander (2011) modelled the transition matrix of movement of loans between delinquent states as a Markov chain but did not forecast transition probabilities using loan-level covariates adopting instead a term structure of credit risk spreads approach.

Above these, other methods such as survival models have been identified as superior to the former due to their ability to incorporate time varying covariates such as macroeconomic conditions which affect performance on loan payment over time (Castro 2013) and the ability to forecast event occurrence (default, recovery, prepayment, foreclosure) in the next instant of time, given that the event has not occurred until that time (Bellotti and Crook 2013; Chamboko and Bravo 2016). Commonly, survival models have been used to model the risk of defaulting (Bellotti and Crook 2013; Noh et al. 2005; Sarlija et al. 2009; Tong et al. 2012; Chamboko and Bravo 2019a). Several studies have also used the same to model foreclosure on mortgages (Gerardi et al. 2007) and also cure from delinquency to current (Ha and Krishnan 2012; Chamboko and Bravo 2016, 2019b; Ha 2010). The competing risks survival framework has also been used to model the competing risks of early payment and default on loan contracts (Deng et al. 1996; Stepanova and Thomas 2002).

The option-based model of default has also been widely used in the USA (Kau et al. 1992; Deng et al. 2000) and UK markets (Ncube and Satchell 1994) to characterise mortgage default as ruthless or mainly influence by the relation of price of houses and value of mortgages. These models define default as an American option with the strike price equal to the value of the mortgage and then assume that a borrower will default as soon as the property value falls to or below the mortgage value (Kelly and O'Malley 2016), particularly when the lender has no recourse. A major limitation of these models is that defaults are usually defined the same way as foreclosures, thus ignoring additional important options for borrowers, for instance cure or prepayment.

As observed in the literature above, most of the studies investigated mortgage delinquency, default or foreclosure as the main outcomes of interest. Limited literature is available on the modelling of the cure from delinquency to current as well as the transition to other states such as early payment, and none on repurchase and short sale as mortgage loan outcomes. Usually, these studies modelled one outcome of interest at any time, thus missing on the dynamics of the portfolio with the passage of time. By modelling the occurrence of multiple loan events at the same time and their recurrences, we offer a much richer perspective than these traditional approaches.

3. Modelling Intermediate Events and Multiple Loan Outcomes

3.1. Competing Risks

Competing risks generalise the standard survival analysis (Beyersmann et al. 2012) and refers to a situation where there is more than one cause of failure or outcomes (i.e., foreclosure, short sale or early payment). As such, competing risks models are meant to deal with situations where there is one initial state and multiple and mutually exclusive absorbing states (Deng et al. 1996;

Stepanova and Thomas 2002). Among the competing risks, only the first one to occur is observed. Diagrammatically, a simple illustrative competing risks model is presented as follows (Figure 1).

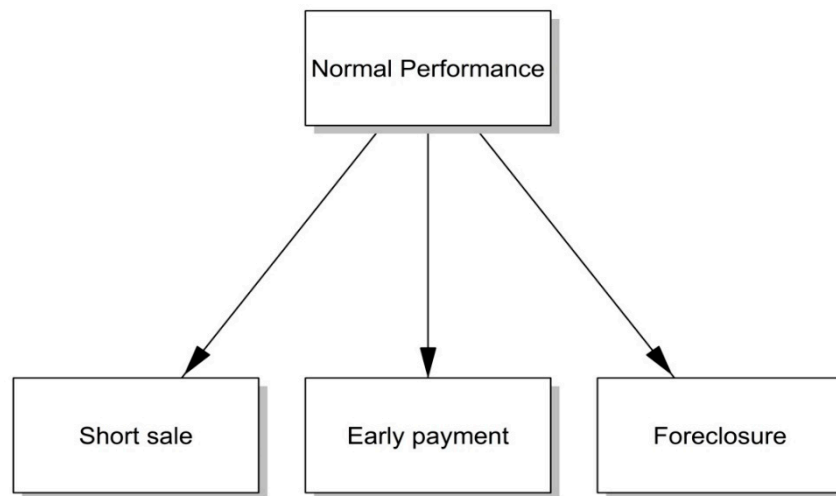


Figure 1. A competing risks scenario with three causes of failure. Notes: This is just for illustration reasons as there are many other possible causes of failure or transition from normal performance (current).

Since this approach only observes the first event as result of the competing risks, subsequent events, if available, are not considered. For instance, a typical mortgage loan gets into delinquency and default before foreclosure, thus having intermediate states which are neither initial nor absorbing. Should the intermediate and later events be of interest, the competing risks approach will fail to fully utilise the available data to understand loan repayment behaviour. In such instances, extensions of competing risk models, particularly multistate models, are applicable.

3.2. Multi-State Models

Multi-state models are extensions of the competing risk models that model events as transitions between states and include competing risks as a special case (Putter et al. 2007). They allow the modelling of events of different types as well as both intermediate and subsequent events. It is often assumed that a multi-state model is a Markov model with the Markov property stating that the transition rate is independent of both the states visited prior to the current state and the sojourn time (length of stay in current state). In other words, the future depends on history only through the present. In this paper, we focus on discrete-time Markov chains with a finite set of states since mortgage data are observed at discrete-time intervals. It is assumed that the state space of the model characterise all the possible states in which the loan can be in.

Andersen and Keiding (2002) described a multi-state process as a stochastic process $(X(t), t \in T)$ with a finite state space $S = \{1, \dots, p\}$ and with right continuous sample path: $X(t+) = X(t)$, with $T = [0, \tau]$ or $[0, \tau)$ with $\tau \leq +\infty$. The process has initial distribution $\pi_h(0) = Prob(X(0) = i), i \in S$. The multi-state process $X(\cdot)$ generated a history X_t (an σ -algebra) which contains the history of the process in the interval $[0, t]$. This history consists of both information relating to the previous states visited and the time spent in previous states. With this history, we can define transition probabilities by:

$$P_{ij}(s, t) = Prob(X(t) = j | X(s) = i, X_{S^-}) \quad (1)$$

where $P_{ij}(s, t)$ is the probability of being in state j at time t given that the subject was in state i at time s . If T is the time needed to reach state j from i , the transition intensity (hazard rate) of $i \rightarrow j$ transition is given by (Andersen and Keiding 2002):

$$\lambda_{ij}(t) = \lim_{\Delta t \rightarrow 0} \left(\frac{P_{ij}(t \leq T < t + \Delta t | T \geq t)}{\Delta t} \right) \tag{2}$$

A state $i \in S$ is absorbing if all $t \in T, i \in S, j \neq i, \lambda_{ij}(t) = 0$, otherwise i is transient. Where there is no transition, transition intensities will be zero for all t . The cumulative transition of $i \rightarrow j$ is given by Nelson–Aalen estimators as

$$\Lambda_{ij}(t) = \int_0^t \lambda_{ij}(u) du = \hat{\Lambda}_{ij}(t) = \sum_{s \leq t} \frac{N_{ij}(s)}{Y_i(s)} \tag{3}$$

where $N_{ij}(s)$ represents the number of transitions observed from state i to state j at time s and $Y_i(s)$ represents the number of uncensored individuals in state i at time s . The transition probabilities can be expressed in form of a matrix as

$$P(s, t) = \prod_{u \in (s, t)} (I + \Delta \Lambda(u)) \tag{4}$$

with $P_{ij}(s, t)$ entries, where I is the identity matrix and $\Delta \Lambda(u)$ is a matrix containing elements $\Delta \Lambda_{ij}(u)$, representing the change in the cumulative transition rate between states i and j at time u . The Aalen–Johansen estimator (Aalen and Johansen 1978) is used to estimate the transition probabilities. The Aalen–Johansen estimator $\hat{P}(s, t)$ is derived by replacing $\Lambda(u)$ with the matrix $\hat{\Lambda}(u)$ of Nelson–Aalen estimators ($\hat{\Lambda}_{ij}(t)$) as below

$$\hat{P}(s, t) = \prod_{u \in (s, t)} (I + \Delta \hat{\Lambda}(u)) \tag{5}$$

where $\Delta \hat{\Lambda}(u)$ is the observed change in matrix $\hat{\Lambda}(u)$ which represents the estimates of the cumulative hazard of transitioning from state i to state j at time u .

For mortgage repayments, all individuals start in the same state (performing/current), thus we can have M , the initial state (at time 0) for the multi-state model. We can also define the cumulative incidence function (CIF) for states Z as follows:

$$CIF_Z(t) = P(T_z \leq t) = P_{MZ}(0, t) = \left[\prod_{u \in (0, t)} (I + \Delta \hat{\Lambda}(u)) \right]_{M, Z} \tag{6}$$

where T_z is the time to transition from any other state to state Z .

The transition probabilities depend on time t , or, more generally, on a set of individual, loan-level, or macroeconomic time-dependent explanatory variables. To analyse the relationship between the characteristics of borrowers and loans and their transition rates, we model transition intensities as functions of both acquisition and performance explanatory variables. To be more specific, we use a proportional hazards Cox model and a multiplicative structure with a common baseline $i \rightarrow j$ transition intensity $P_{ij,0}(s, t)$. For an individual mortgage contract, m , with time-fixed covariates $X_m = X_{mk}$, the transition intensity is modelled as:

$$P_{ij}^{(m)}(s, t) = P_{ij}^{(0)}(s, t) \exp(\beta'_{ij} X_m) \tag{7}$$

where $P_{ij}^{(0)}(s, t)$ is the baseline hazard, which captures the shape of the hazard function and summarises how the probability of mortgage transition changes over time. The proportionality factor $\exp(\beta_{ijm})$ quantifies the effect of a given covariate X_{mk} on the transition intensity. We used R package *mstate* to fit the multi-state model and to estimate the state or transition probabilities $P(s, t)$ and the cumulative transition rates $\Lambda(u)$ whilst the *survival* package was used for fitting the transition specific prognostic survival models.

4. Statistical Analysis and Results

4.1. Data

The study analysed 383,770 mortgage contracts from the Fannie Mae single-family mortgage loans originated during the first quarter of 2009 followed on their monthly repayment performance until the third quarter of 2016. This is an interesting period to analyse the credit performance of US mortgage loans, following the collapse of Lehman Brothers and the entrance of Fannie Mae and Freddie Mac into Conservatorship in September 2008. Besides, there was an ushering of significant monetary policy changes which lowered interest rates to historic lows levels; the introduction of two large-scale debt relief programs, namely the Home Affordable Refinancing Program (HARP) and the Home Affordable Modification Program (HAMP); and the creation of the Troubled Asset Relief Program (TARP) in October 2008. Again, this period witnessed the signing of the Dodd–Frank Wall Street Reform and Consumer Protection Act legislation in July 2010, the regulation of the financial industry, efforts to halt predatory mortgage lending and encouraging transparency for consumers to be able understand conditions relating to their mortgages before making contractual commitments. During this period, cash sales peaked, home prices hit bottom and foreclosure rates reached record high by mid 2010. The delinquency rate only dropped below 4% for first the time in March 2015 since the start of the crisis. The sample is geographically dispersed in the United States and covers loans that were originated in all states.

The primary dataset has a subset of Fannie Mae’s 30-year and less FRMs. Static acquisition and a selected number of dynamic performance variables were considered for modelling the various loan outcomes. Tables 1 and 2 present the loan acquisition and loan performance variables, respectively.

Table 1. Acquisition Variables.

Acquisition Variables	Description/Values
Loan Identifier	Mortgage loan unique identifier.
Origination Date	MM/YYYY
Loan Purpose	An indicator that denotes if a mortgage loan in a pool is either a purchase money mortgage or refinance mortgage.
Product Type	Fixed-rate mortgage or adjustable-rate mortgage.
Property Type	Single-Family (SF), Condo (CO), Co-Op (CP), Manufactured Housing (MH), PUD (PU)
Relocation Mortgage Indicator	Yes/No
Channel/ Origination Type	Retail (R), broker (B), correspondent (C)
Borrower Credit Score	300–850
Co-Borrower Credit Score	300–850

Table 1. Cont.

Acquisition Variables	Description/Values
Debt-to-Income (DTI) Ratio	Calculated at origination and is obtained by dividing the borrower's total monthly obligations by stable monthly income.
First-Time Homebuyer Indicator	Denotes if a borrower is a first-time homebuyer.
Number of Borrowers	Number of individuals who are obligated to repay the loan.
Number of Units	The number of units the mortgaged property has.
Occupancy Status	Indicates how the borrower used the mortgaged property at the time of origination.
Original Interest Rate	Mortgage original interest rate
Original Loan Term	The number of months for which a borrower's payments are due
Original Loan-to-Value (LTV)	The Original LTV reflects the loan-to-value ratio of the loan amount secured by a mortgaged property on the origination date of the underlying mortgage loan.
Original Unpaid Principal Balance (UPB)	The original amount of the mortgage loan as indicated by the mortgage documents.

Source: Fannie Mae.

Table 2. Performance Variables.

Performance Variables	Description/Values
Current Loan/Delinquency Status	The number of days (months) a borrower is delinquent, i.e., 0 = Current, or < 30 days overdue; 1 = 30–59 days; 2 = 60–89; 3 = 90–119; and so on.
Modification Flag	Y = Yes, N = No
Monthly Reporting Period	MM/DD/YYYY
Loan Age	Number of months after origination
Zero balance code	Reason mortgage loan's balance was reduced to zero, e.g., 01 = prepaid; 03 = Short Sale; and so on.

Source: Fannie Mae.

4.2. Conceptual Framework

Table 3 defines and describes the different stages or states in the progression of a mortgage loan contract included in the Fannie Mae dataset. Some states (e.g., prepayment) refer to a “terminal” status, i.e., not subject to change, while others (e.g., delinquency) indicate that loan status can change as it continues to move through its lifecycle.

As depicted in Figure 2, we propose a multi-state model framework with seven allowable states numbered 1–7 with 16 possible transitions. State 1 is the initial state and States 4–7 are absorbing (final) whilst States 1–3 are transient (intermediate). The transition from one state to the other is presented by arrows from state $i \rightarrow j$.

Table 3. States descriptions.

State	Description
Current/normal performance	This is when a borrower is up to date with payments or overdue by less than 30 days.
Delinquency	When payments are 30–59 days overdue.
Default	This is when a borrower missed payments for 60 days or more consecutively.
Prepayment	Occurs when a loan is paid in a shorter period than agreed contractually.
Mortgage foreclosure	Occurs when a borrower fails to pay in time or in full instalments and the lender repossesses the property.
Deed-in-Lieu, REO Disposition	This is when the borrower seeks release from the mortgage contract by voluntarily transferring the title of the property to the lender.
Short sale	This is when a homeowner sells a home for less than the balance remaining on a mortgage and pays off all (or a portion of) mortgage balance with the proceeds.
Recovery or cure	This is when a borrower once in delinquency or default resumes making payments.

Source: Fannie Mae.

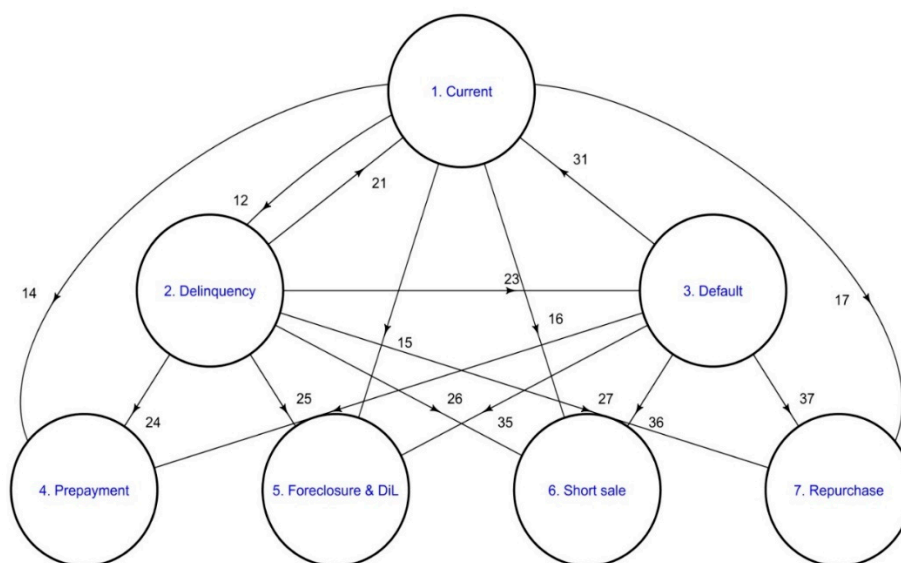


Figure 2. A multi-state model framework for analysing mortgage loans data. Source: Author’s preparation. Notes: The numbers on the arrows represent the direction of the transition. For example, 12 represents the transition from State 1 (current) to State 2 (delinquency). Similarly, number 37 means transition from State 3 (default) to State 7 (foreclosure and Dil) and so on.

4.3. Results and Discussion

4.3.1. Descriptive Statistics

Tables 4 and 5 present the characteristics of the borrowers, loans, properties purchased and some borrower behavioural characteristics. On the borrowers, only about 5% were first time homebuyers, with most of the contracts owned by two borrowers (median number of borrowers = 2). The credit scores of the borrowers ranged from 508 to 850, averaging 763, whilst that of the co-borrowers were almost the same, ranging from 505 to 850 and averaging 769.3. The debt-to-income ranged from 1% to

64%, averaging at 33%. As depicted in Figure 3, the majority of the borrower's DTI was between 20% and 50%.

Table 4. Descriptive statistics for categorical covariates.

Class of Variables	Variable	Category	Percent
Loan and borrower characteristics	Channel	Broker	13.6
		Correspondent	31.8
		Retail	54.6
	First Time Homebuyer Indicator	Yes	5.3
		No	94.7
Loan Purpose	Purchase	15.3	
	No Cash-out Refinance	53.4	
	Cash-out Refinance	31.3	
Property characteristics and purpose	Occupancy Status	Principal	93.2
		Second	4.0
		Investor	2.9
	Property Type	Condo	6.8
		Co-op	0.4
		Manufactured Housing	0.2
		Planned Urban Development	19.5
Number of Units	Single Family	73.1	
	1	98.7	
	2	1.1	
	3	0.1	
	4	0.1	
Behavioural variables	Relocation	Yes	0.2
		No	99.8
	Modification Flag	Yes	0.3
		No	99.7

Source: Author's preparation.

Table 5. Descriptive statistics for continuous covariates.

Variable	Min	1st Quantile	Median	Mean	3rd Quartile	Max	Std Dev
Borrower characteristics							
Borrowers' Credit Score	508.0	741.0	773.0	763.0	793.0	850.0	40.06
Co-borrower's Credit Score	505.0	751.0	779.0	769.3	797.0	850.0	36.99
Debt-to-Income Ratio	1	24	32	33.08	42.0	64.0	11.81
Number of Borrowers	1.0	1.0	2.0	1.61	2.0	7.0	0.498

Table 5. Cont.

Variable	Min	1st Quantile	Median	Mean	3rd Quantile	Max	Std Dev
Loan characteristics							
Loan-to-Value	3.0	60.0	73.0	68.77	80.0	97.0	15.65
Original Loan Term (months)	301	360	360	359.9	360	360	1.88
Unpaid Principal Balance	10,000	147,000	215,000	235,223.13	308,000	950,000	113,370.64
Original Interest Rate	1.88	4.75	4.875	4.98	5.125	8.625	0.361
Behavioural variables							
Loan Age at estimation (months)	0	30	41.0	45.77	58.0	92.0	24.63
Number of transitions	0	1.0	1.0	1.27	1	63.0	2.156
Sojourned time (months)	1.0	26	40.0	43.61	54	92.0	25.96

Source: Author's preparation.

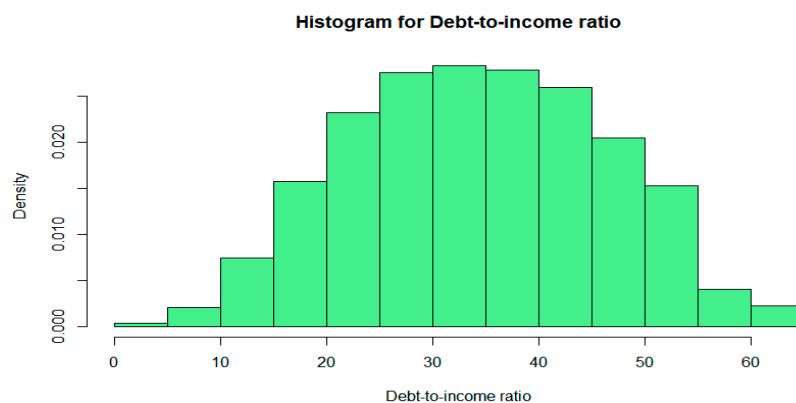


Figure 3. Distribution of debt-to-income ratio. Source: Author's preparation.

Looking at the loan and property characteristics as well as the purpose, we observed that, even though most of the borrowers were not first homebuyers, the majority (93%) used the mortgaged property as their own residence, as indicated at the origination date. Only 15% were purchase money mortgages with the other 85% as refinance mortgages. Almost three quarters of the properties were single family properties, with 20% being planned urban developments and close to 7% as condominium houses. After acquiring the mortgage, almost all borrowers did not relocate from that residency and have also not sought modification of the mortgage.

The monetary value of the properties had a median value of \$215,000. Using the loan-to-value figures, the loan amounts averaged close to 70% of the value of the property, ranging from as low as 3% to a maximum of 97% of the property value and the mean loan term was about 30 years (360 months). The origination interest rate ranged from 1.88% to 8.625% averaging at 5%.

4.3.2. Transition Matrix

The transition matrix (Table 6) summarises the state occupation probabilities or transitions experienced by borrowers during the study period. The results show that only 11.6% of the borrowers stayed current on their payments during the study follow-up period whilst 16.9% progressed into a delinquent state. Most of the borrowers (71.4%) transitioned from current to prepayment. This high prepayment finding was echoed by Danis and Pennington-Cross (2005) who concluded that prepayment rates tend to increase as delinquency intensity rises. Again, the result is consistent with the assertion that in declining mortgage interest rates scenarios refinancing incentives are more likely worthwhile, especially given that these were fixed rate mortgages.

Table 6. Transition matrix.

		To						
		Current	Delinquency	Default	Repurchase	Prepayment	Foreclosure and Dil	Short Sale
From	Current	0.116	0.169	0.0	0.001	0.714	0.0	0.0
	Delinquency	0.706	0.0	0.271	0.0	0.023	0.0	0.0
	Default	0.749	0.0	0.0	0.011	0.054	0.127	0.06

Source: Author's estimation.

For those who entered a delinquency state, 70.6% recovered (cured) whilst 27.1% defaulted and 2.3% progressed to prepayment. Of those who reached a default state, 74.9% cured, thus transitioning back to the current state. The high recovery or cure for those both in delinquency and default states is a desirable situation which in this case signalled the ability of borrowers to escape financial distress. This could be attributed to various lenders strategies and policy interventions such as HARP which were implemented in the aftermath of the financial crises to cut down losses and preventing ruthless default by underwater mortgagors (Liu and Sing 2018). Another 12.7% of those who defaulted progressed to foreclosure, whilst another 6% and 5.4% transitioned to short sale and prepaid, respectively.

Figure 4 shows the distribution of the number of transitions experienced by borrowers. Only 11.6% of the borrowers had no transition at all, whilst about 54% transitioned (experienced event) only once, thus Figure 4 is skewed towards zero transitions. The direction or the states where the borrowers transitioned to are described in the transition matrix above (Table 6). Figure 5 shows the time spent in a state before experiencing an event or transitioning to another state. This ranged 1–92 months (Table 5). About 21% stayed in a state for three months or less before experiencing an event. On average, one stayed in a state for about 40 months (median sojourned time) before transiting into another state (Table 5). Further analysis shows that, on average, borrowers stayed in the “current” state for about 40 months before transitioning whilst they stayed about seven months in a default state before transitioning into an absorbing state (foreclosure, repurchase, short sale or prepaid).

Histogram for number of transitions

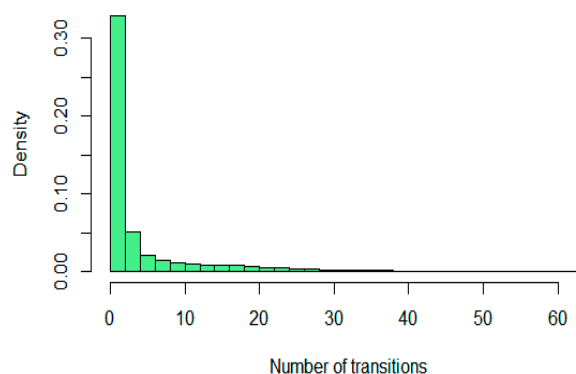


Figure 4. Number of transitions per borrower. Author's preparation.

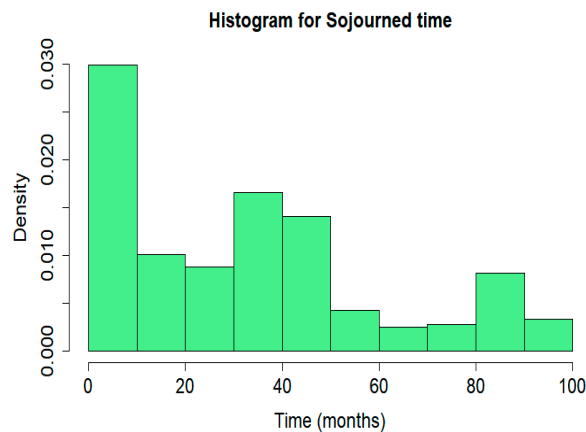


Figure 5. Sojourned time per borrower. Author’s preparation.

4.3.3. Cumulative Incidence Functions

Cumulative incidence functions giving the probability of a loan entering a given state before time t are an interesting way to consider the economic significance of transition probabilities. Since the current framework is not for a single risk event occurring per individual, but instead is a multi-state scenario involving recurrent and multiple events, and at any transient state, a borrower had competing risks likely to cause the occurrence of an event, and the cause specific failure probabilities are therefore best described by cumulative incidence curves. The cumulative incidence functions in Figures 6–9 thus quantify the cause-specific failure probabilities for selected transitions. In this case, the dependent censoring arising from the competing causes renders the Kaplan–Meier estimator (Kaplan and Meier 1958) inappropriate. As shown in Figure 6, the probability of borrowers in a current state to transition to a delinquent state remained low (below 20%) across the duration of the study. However, the probability of transitioning from current to a prepayment state started low during the first months of a mortgage and gradually rose to around 20% after 40 months and at last a sudden spike towards the end of the study period (Figure 7). This pattern could be explained by the declining interest rates during the same period, thus increasing the incentive to prepay given that these were fixed-rate mortgage loans.

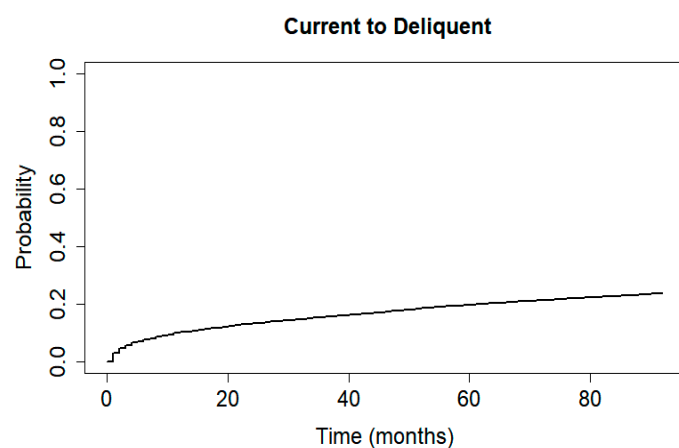


Figure 6. Cumulative incidence function (Current to delinquent). Source: Author’s estimation.

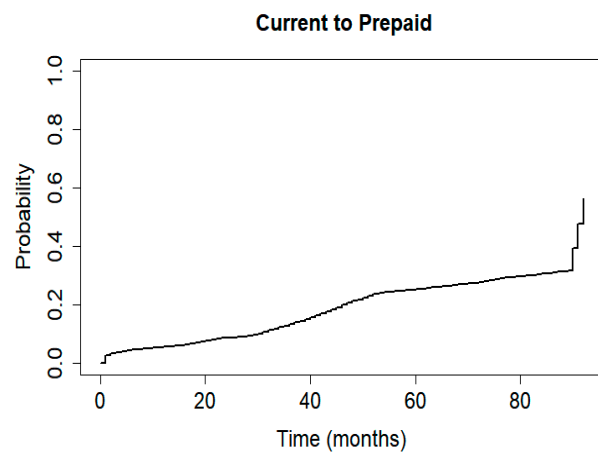


Figure 7. Cumulative incidence function (Current to prepaid). Source: Author's estimation.

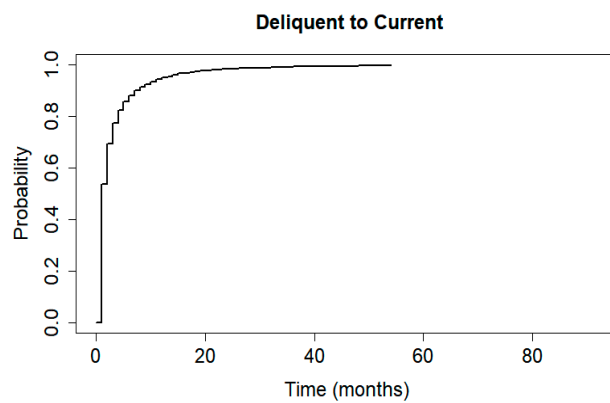


Figure 8. Cumulative incidence function (Delinquent to Current). Source: Author's estimation.

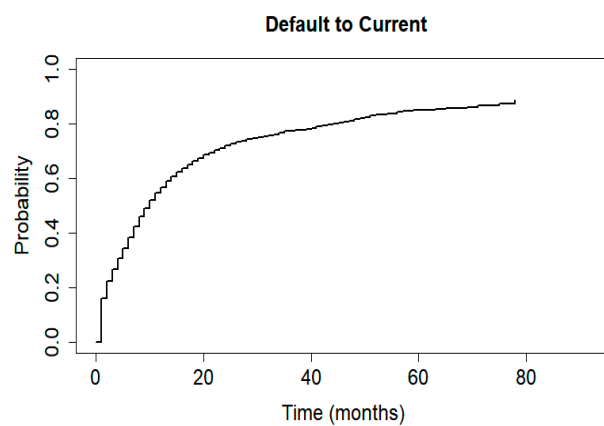


Figure 9. Cumulative incidence function (Default to current). Source: Author's estimation.

Figure 8 shows that borrowers who entered a delinquent state had a very high and encouraging probability to revert to a current state within a month or two. We also note that, for the borrowers who entered a default state, they also had encouragingly high probabilities of reverting to current (Figure 9). The chances gradually rose during the first 20 months and then stagnated.

4.3.4. Prognostic Factors for Event Specific Transitions

In this section, we present the results obtained on the estimation of prognostic factors (covariates) for event specific transitions using the proportional hazards model. We selected the most important

transitions, as shown in the transition matrix (Table 6), which are current to delinquent and/or default, current to prepaid, delinquent to current and default to current.

Current to Prepayment Transitions

Many factors potentially impact the prepayment decision including: (i) macroeconomic factors (for instance, mortgage rates, housing market inflation and consumer confidence); (ii) loan-specific factors (for instance, mortgage rate, type-FRM, ARM, Hybrid, original term, remaining term, loan size, loan-to-value ratios, credit score, insurance costs, collateral and penalties); (iii) borrower-specific factors (for instance, credit score, unemployment or loss of income, sickness or death of mortgagor or in family, divorce or other life events and borrower sentiment); and (iv) others (for instance, mortgage origination and servicing process). Additionally, one other motivation for prepaying a mortgage is housing turnover, i.e., the sale of the home triggers a prepayment. Housing turnover is explained by a multitude of factors, such as job changes and relocations, “trading up” (moving into a bigger home) of young households or downsizing of older households, unemployment spells or life events.

Typically, US mortgage borrowers have the option of making partial/full prepayments of their mortgage balances, possibly with a penalty, even when they do not have the intention to sell their houses. In these later scenarios, prepayment is mainly associated with mortgage refinancing. An important feature of the US mortgage market is that residential agency backed loans frequently allow for penalty-free prepayments, encouraging sub-optimal prepayment behaviour and the adoption of exogenous prepayment rules.²

Table 7, Column (1) presents the results on the prognostic factors for the transition from normal performance (current) to prepayment. The results suggest that higher borrower’s credit score ($HR = 0.998$, $p < 0.01$), higher co-borrower credit score ($HR = 0.998$, $p < 0.01$), higher number of borrowers ($HR = 0.905$, $p < 0.01$), a higher debt-to-income ratio ($HR = 0.999$, $p < 0.1$) and being a first-time homebuyer ($HR = 0.901$, $p < 0.01$) significantly reduced the propensity to transition from a current state to prepayment. Equally important, longer loan terms ($HR = 0.998$, $p < 0.1$), having more units ($HR = 0.753$, $p < 0.01$) and a higher loan-to-value ratio ($HR = 0.996$, $p < 0.01$) also reduced the chances of transitioning from current to prepayment whilst higher original principal balances ($HR = 1.01$, $p < 0.01$) and higher original interest rates ($HR = 1.90$, $p < 0.01$) increased the chances to transition into prepayment. As expected, the costs of mortgages and the corresponding refinancing incentives are among the most important drivers of prepayment. Given the FRMs in a declining interest rate scenario, the prepayment option is in-the-money, reflecting the positive difference between the value of the outstanding loan repayments at the current interest rate and the value of outstanding loan repayments at the original loan interest rate. The higher are the initial interest rate and principal balance, the higher is the refinancing incentive. The loan age (or seasoning) effect of prepayment is confirmed in this study, with younger loans (longer loan terms) having a lower probability of prepayment. Usually, a S-shaped relation between the rate of prepayment and the loan age is observed (see, for instance, Charlier and Bussel 2003).

² The early repayment of a mortgage significantly impacts both the bank’s profitability, via reduced interest margins, the loss of foregone interest payments (less the risk of a default on the outstanding debt) and the bank’s interest rate and liquidity positions. Prepayment creates a reinvestment risk problem for banks that must be addressed by appropriate Asset–Liability Management (e.g., immunization) techniques (see, e.g., Bravo and Silva 2006).

Table 7. Event specific transaction factors.

	Current to Prepaid Transitions (1)	Current to Delinquent/Default (2)	Delinquent to Current (3)	Default to Current (4)
Covariate	HR [Estimate] (SE)	HR [Estimate] (SE)	HR [Estimate] (SE)	HR [Estimate] (SE)
Modification Y	2.37 [0.863 ***] (0.0145)	5.57 [1.72 ***] (0.0299)	Omitted	Omitted
Purpose P	0.979 [−0.0208 ***] 0.00778	0.819 [−0.199 ***] (0.0136)	1.16 [0.152 ***] (0.0147)	1.36 [0.305 ***] (0.0367)
Purpose R	1.08 [0.08 ***] (0.00435)	0.852 [−0.16 ***] (0.0136)	1.06 [0.0564 ***] (0.00983)	1.03 [0.0281] (0.0245)
Property Type CP	0.78 [−0.249 ***] (0.0517)	1.13 [0.123] (0.147)	1.23 [0.207 ***] (0.0862)	0.989 [−0.0112] (0.228)
Property Type MH	0.654 [−0.424 ***] (0.0408)	0.938 [−0.0636] (0.110)	1.04 [0.0346] (0.0861)	1.02 [0.0199] (0.223)
Property Type PU	1.10 [0.0988 ***] (0.00111)	0.993 [−0.00682] (0.0372)	0.95 [0.0515 **] (0.0212)	1.27 [0.236 ***] (0.0535)
Property Type SF	1.12 [0.112 ***] (0.0104)	1.21 [0.190 ***] (0.0343)	0.927 [−0.0762 ***] (0.0188)	1.27 [0.242 ***] (0.0417)
Relocation Y	1.29 [0.253 ***] (0.0442)	0.952 [−0.0495] (0.135)	1.08 [0.0773] (0.106)	1.83 [0.604 ***] (0.216)
Occupancy P	1.49 [0.399 ***] (0.0128)	1.10 [0.094 ***] (0.0335)	0.818 [−0.201] (0.0216)	1.04 [0.0404] (0.0573)
Occupancy S	1.32 [0.274 ***] (0.0155)	1.02 [0.0218] (0.0441)	0.925 [−0.0783 **] (0.0491)	1.01 [0.00872] (0.0832)
First time homebuyer U	0.575 [−0.553 *] (0.2.36)	0.92 [−0.11] (0.0113)	Omitted	Omitted
First time homebuyer Y	0.901 [−0.104 ***] (0.0124)	0.791 [−0.234 ***] (0.0355)	0.985 [−0.0172] (0.0202)	0.927 [−0.0753] (0.0475)
Units	0.753 [−0.283 ***] (0.0143)	0.854 [−0.1.58 ***] (0.0361)	0.988 [−0.0116] (0.0216)	0.828 [−0.189 ***] (0.0542)
Channel C	1.03 [0.0267 ***] (0.00635)	0.897 [−0.109 ***] (0.0192)	0.978 [−0.0219 *] (0.0129)	1.03 [0.0266] (0.0315)
Channel R	0.911 [−0.0933 ***] (0.00587)	0.785 [−0.242 ***] (0.0179)	0.941 [−0.0861 ***] (0.0123)	0.972 [−0.0284] (0.0303)
Co-borrower credit score	0.998 [−0.00163 ***] (0.0000712)	0.993 [−0.00714 ***] (0.000197)	Omitted	Omitted
Borrower credit score	0.998 [−0.00174 ***] (0.0000706)	0.989 [−0.0112 ***] (0.000199)	1.0 1 [0.00202 ***] (0.0000895)	1.03 [0.00174 ***] (0.000219)
DTI	0.999 [−0.000439 *] (0.000175)	1.02 [0.0214 ***] (0.000558)	0.992 [−0.0081 ***] (0.00038)	0.996 [−0.00415 ***] (0.000983)
Number of borrowers	0.905 [−0.0999 ***] (0.0181)	0.982 [−0.0186] (0.0518)	1.04 [0.0378 ***] (0.00865)	1.09 [0.0846 ***] (0.0219)
Loan-to-Value (LTV)	0.996 [−0.00358 ***] (0.000128)	1.01 [0.012 ***] (0.000462)	0.994 [−0.0059 ***] (0.000317)	0.991 [−0.0094 ***] (0.00147)
Loan term	0.998 [−0.00234 *] (0.000977)	1.03 [0.0294 ***] (0.00525)	1.0 [0.000631] (0.00407)	0.992 [−0.00842] (0.0091)
Original principal balance	1.01 [0.0000019 ***] (0.000000177)	1.01 [6.62e−07 ***] (0.0000000589)	1.0 [−0.000000001] (0.000000067)	1.01 [−0.000001 ***] (0.000000018)
Original interest rate	1.90 [0.644 ***] (0.00551)	1.48 [0.39 ***] (0.048)	0.868 [−0.141 ***] (0.00998)	0.804 [−0.294 ***] (0.0401)

*** significant at 1%, ** significant at 5%, * significant at 10%. Notes: HR represent hazard ratios and SE stands for standard errors. A variable is omitted in the model when it does not carry enough data to run that model.

With respect to the property type securing the mortgage loan, the results show that, compared to condominiums, single-family home ($HR = 1.1, p < 0.01$) and planned urban development ($HR = 1.1, p < 0.01$) were more likely to transition from current to prepayment and the opposite was true for cooperative share ($HR = 0.78, p < 0.01$) and manufactured home ($HR = 0.65, p < 0.01$). Property type has been shown to be significant in explaining mortgage prepayment and default hazard rates in prior studies (see Vandell et al. 1993). The results also show that purchase money mortgages were less likely to transition from current to prepayment compared to refinance mortgages ($HR = 0.979, p < 0.01$). Similarly, borrowers who had their mortgage loans modified were twice as likely to transition from a current state to prepayment ($HR = 2.37, p < 0.01$). Lenders can modify the loan terms in many possible ways (for instance, reducing the interest rate, reducing the principal balance, extending the maturity date of the loan) to offer financial relief for borrowers, enabling them to resume their regularly scheduled payments. In any case, modifying the loan could mean having issues with the loan or inability to repay, thus incentivising the need to get rid of the mortgage. Several government programs were launched in the U.S. through 2008 and 2009 to encourage mortgage modification as foreclosures spiked. Mortgage loans made to borrowers whose employers relocate their employees were more likely to transition to prepayment ($HR = 1.29, p < 0.01$), confirming the hypothesis that relocating for higher paying job opportunities is an important factor triggering house sales and mortgage prepayment.

On loan origination channel used, loans originated by retailers were less likely to transition to prepayment compared to those originated by brokers ($HR = 0.911, p < 0.01$) and the opposite was true for those originated by correspondents ($HR = 1.03, p < 0.001$). Considering how mortgaged property was used at the time of origination, the results suggest that those who used property as the principal residence ($HR = 1.49, p < 0.01$) or second home ($HR = 1.32, p < 0.01$) were more likely to transition to prepayment compared to those who regarded these as investment property. This result can be explained both by the fact that investors in income property do not relocate and by the fact that investment property loans often carry a prepayment penalty consisting of a percentage of the remaining principal outstanding.

Current to Delinquency and Default Transitions

Table 7, Column (2) presents the results on the factors affecting the transition from current to a delinquency state suggesting difficulties in repaying mortgage loans. Considering the definition applied in this study, delinquency is the state when a loan obligor had missed payments for 30–59 days whilst consecutively missing payments for 60 days or more was defined as having defaulted. It is therefore reasonable to say that the factors which affect the transition from current to delinquency are the same for the delinquency to default transition since it is the same path with one-month difference. As expected, higher credit scores for borrowers ($HR = 0.989, p < 0.01$) or co-borrowers ($HR = 0.993, p < 0.01$) reduced the propensity to transition to a delinquency state.

Similarly, being a first-time homebuyer ($HR = 0.791, p < 0.01$) significantly reduced the propensity to get into a delinquency state. Typically, first-time homebuyers correspond to younger households with lower income, lower home equity and lower credit scores when compared to repeat homebuyers but this does not mean they necessarily default at a higher rate. Our results show that first-time homebuyers prepay at a lower rate whilst at the same time do not default at a higher rate. For lenders, our results show that, taking into account the borrower and loan characteristics at origination and pricing appropriately the borrower's risk and ability to reimburse, there is no evidence that first-time homebuyer mortgages are intrinsically riskier than average repeat homebuyer mortgages. This conclusion has potentially important policy implications since many (national, regional and local) government policies provide numerous first-time homebuyers programs and opportunities to make increase the affordability of homes (for instance, Good Neighbor Next Door, Fannie Mae or Freddie Mac).

As also expected, a higher debt-to-income ratio ($HR = 1.02$, $p < 0.01$), longer loan term ($HR = 1.03$, $p < 0.01$), higher loan-to-value ratio ($HR = 1.01$, $p < 0.01$), higher original loan amount ($HR = 1.01$, $p < 0.01$) and higher original interest rate ($HR = 1.48$, $p < 0.01$) increased the chances of borrowers to transition into a delinquency state. These results confirm previous evidence that the level of household mortgage debt and debt-servicing ratio has a significant impact on household default (see, for instance, [Di Maggio et al. 2017](#)). Moreover, our results suggest that the design of mortgages is important to reduce default, increase welfare and to minimise the impact of foreclosures in the economy. The typical FRM design in the US prevented many distressed borrowers from refinancing in a declining interest (and inflation) rate scenario, triggering delinquency and default, impacting household outcomes. Given the cyclicity of the economy, indexing mortgage payments to current interest rates through, for instance, ARMs generates uncertainty in the stream of payments but has the potential to alleviate distressed borrowers in crisis periods, improving household outcomes and welfare. Standard ARMs reduce the debt burden of borrowers in crisis periods by automatically passing interest rate reductions to households, by delivering larger payment reductions due to front-loading and by mitigating the price declines and the price-default spiral ([Guren et al. 2018](#)).

Mortgage loans originated by retailers ($HR = 0.785$, $p < 0.01$) or correspondents ($HR = 0.897$, $p < 0.01$) were less likely to get into a delinquency state compared to those originated by brokers. The purpose of the mortgaged property at the origination date was also an important factor explaining transiting into delinquency. Borrowers with properties which were their principal residence ($HR = 1.10$, $p < 0.01$) had higher chances to fall into delinquency compared to those who regarded these as investment property. Home homeownership is a form of wealth which owners can liquidated should the need arises. However, homeowners have however to spend money on maintenance costs to maintain the value of their investment over time. This means non-monetary costs (for instance, utility loss of the foreclosed home, necessity to move, lower credit rating and moral or ethical breach) create a disincentive to default on primary residences but that may not be sufficient to prevent default if borrower's are unable to make payments

On the property type securing the mortgage loan, borrowers who had planned urban developments ($HR = 0.993$, $p = 0.855$), cooperative share ($HR = 1.13$, $p = 0.402$) and manufactured homes ($HR = 0.938$, $p = 0.565$) were not significantly different from those with condominiums whilst single-family homebuyers ($HR = 1.1$, $p < 0.01$) were more likely to transition from current to a delinquency state. Compared to cash-out refinance, purchase money ($HR = 0.819$, $p < 0.01$) and no cash-out refinance ($HR = 0.852$, $p < 0.01$) mortgages were less likely to transition into a delinquency state. Borrowers who had their mortgage loans modified were five times more likely to transition from current to a delinquency state ($HR = 5.57$, $p < 0.01$). Being a borrower, whose employers relocate their employees ($HR = 0.952$, $p = 0.713$) and the number of borrowers per contract ($HR = 0.982$, $p = 0.713$) were not important factors explaining the transition to delinquency.

Delinquent to Current Transition (Cure or Recovery)

Despite its intuitive appeal, empirical data show that lenders generally renegotiate a relatively small proportion of their delinquent mortgages. Information gaps between borrowers and lenders, institutional aspects and several market frictions (for instance, contract rigidity, equity refinancing constraints, intermediary organisational constraints, state specific laws and regulations, mortgage servicing incentives, loss of information in mortgage securitisation and bank-held or securitised mortgage) have emerged as possible explanations for the scarcity of restructuring efforts. Lenders typically ponder the costs and benefits of avoiding a foreclosure, including the possibility of borrowers self-curing from delinquency without a renegotiation and redefault risk (see, for instance, [Piskorski et al. 2010](#); [Fuster and Willen 2017](#)). Repossessing collateral to resolve delinquent loans is often a long and costly process for lenders. In economic terms, renegotiating a delinquent loan makes sense when both parties consider a renegotiated loan as a better outcome than repossession of the property or postponing a solution, in which case the arrears continue to rise. On the one hand, the moral-hazard

cost of mortgage renegotiations may incentivise non-distressed borrowers to miss payments so that they get a better deal (Mayer et al. 2014). Additionally, in most US states, laws governing mortgage defaults and foreclosure allows lenders to use deficiency judgments processes to claim the debtor's other assets or earnings to settle any differences in case of foreclosure sales. This should in principle discourage default and motivate more borrowers to resume to normal performance since default puts the debtor's other assets at risk.

In this study, we use loan, borrower and property characteristics to understand the probability of cure including. Table 7, Column (3) presents the factors affecting the transition from a delinquency state to a current state. This transition represents the cure or recovery of distressed borrowers back normal performance (current). Mortgage loans with higher credit scores are more likely to be cured from a delinquency state ($HR = 1.01$, $p < 0.01$). This means higher credit score borrowers strive to keep their payments current even when they face financial constraints (for instance, short-term unemployment and divorce) that may force them to miss a loan/interest repayment. The borrower credit score can be considered to some extent a proxy for an individual's ability to pay and willingness to pay since it reflects past credit payment history, the extent of indebtedness, the length of credit history, recent credit taken and the borrower's resilience to income or liquidity shocks or life events. Similarly, a higher number of borrowers per loan also increased the chances of recovering ($HR = 1.04$, $p < 0.01$). This can be interpreted as the positive role of social capital on mortgage repayment.

As also expected, higher debt-to-income ratio ($HR = 0.992$, $p < 0.01$), higher loan-to-value ratio ($HR = 0.994$, $p < 0.01$) and higher original interest rate ($HR = 0.868$, $p < 0.01$) reduced the chances of recovering for the distressed borrowers. From this, it is clear that a prudent policy for increasing the likelihood of cures in the event of delinquency and default should involve considering lower loan-to-value ratio at loan origination and offering some debt relief in distressed periods, for instance, through ARMs instead of FRMs. Similar conclusions were obtained in related studies (see, for instance, Ahlwat 2018). Other factors which were significant include the type of the property, the purpose of the loan, the type of occupancy and the origination channel. Purchase loans have higher probabilities of cure than do refinance loans. Mortgage loans originated by retailers ($HR = 0.941$, $p < 0.01$) or correspondents ($HR = 0.978$, $p < 0.01$) were less likely to recover compared to those originated by brokers. Borrowers with properties which were their principal residence ($HR = 0.818$, $p < 0.01$) or second home ($HR = 0.91$, $p = 0.0193$) had lower chances of recovering from delinquency compared to those who regarded these as investment property.

On the property type securing the mortgage loan, borrowers who had planned urban developments ($HR = 0.95$, $p < 0.01$) and single-family homebuyers ($HR = 0.927$, $p = 0.015$) were less likely to recover compared to those with condominiums. Those with cooperative share ($HR = 1.23$, $p = 0.024$) were more likely to recover whilst those with manufactured homes ($HR = 1.04$, $p = 0.6876$) were not significantly different from those with condominiums. Compared to cash-out refinance, purchase money ($HR = 1.16$, $p < 0.01$) and no cash-out refinance ($HR = 1.06$, $p < 0.001$) mortgages were less likely to recover from a delinquency state. In this transition, it was shown that the loan term ($HR = 1.0$, $p = 0.8768$), the original principal balance ($HR = 1.0$, $p = 0.9732$), being a borrower whose employer relocates their employees ($HR = 1.08$, $p = 0.4639$) and the number of units ($HR = 0.988$, $p = 0.592$) were not significant.

Default to Current Transition (Cure or Recovery)

There are several ways by which lenders or loan servicers can resolve a defaulted mortgage. Usually, the recovery of distressed borrowers from a default state back to current or normal performance allowing borrowers to keep their homes is a desirable policy objective. Designing successful mortgage modification and cure policies requires knowing why homeowners default and which factors increase the likelihood to cure. Table 7, Column (4) presents the factors explaining recovery of distressed borrowers from a default state back to a current state. This transition also represents the cure or recovery of distressed borrowers from an advanced stage of mortgage delinquency back normal performance

(current). As expected, higher borrowers' credit scores ($HR = 1.03$, $p < 0.01$) and higher number of borrowers per loan ($HR = 1.09$, $p < 0.01$) significantly increased the chances of borrowers to recover, since they usually reflect higher ability and willingness to service the mortgage. As also expected, higher debt-to-income ratio ($HR = 0.996$, $p < 0.01$), higher loan-to-value ratio ($HR = 0.991$, $p < 0.01$), higher principal balance ($HR = 1.01$, $p < 0.01$), higher number of units ($HR = 0.828$, $p < 0.01$) and higher original interest rate ($HR = 0.804$, $p < 0.01$) reduced the chances of recovering for defaulted borrowers. This result suggests that the extent of indebtedness and mortgage costs plays an important role in explaining the likelihood of mortgage cure.

Compared to cash-out refinance, purchase money ($HR = 1.36$, $p < 0.01$) were more likely to recover whilst no cash-out refinance ($HR = 1.03$, $p < 0.01$) mortgages were not significantly different. It was also shown that the loan term, the origination channel, being a first-time homebuyer and the type of occupancy were not significant factors.

4.4. Model Validation

While Table 7 presents the factors that affect transition into various loan outcomes, the evaluation of the performance of these models is essential. The traditional classification problem is based on the cross-sectional classification of subjects into a simple binary outcome, typically the presence or absence of default. In classifying individuals as defaulted or not, a marker is prone to two types of error and research is typically conducted to minimise these two errors by using covariates that maximise both high sensitivity and specificity. The model's classification accuracy is commonly quantified using a single-number summary measure such as the area under the ROC curve (AUROC) (see, e.g., Kelly and O'Malley 2016; Kruppa et al. 2013; Bellotti and Crook 2013; Chamboko and Bravo 2016). This approach plots the models' performance based on the true and false positives. The mostly accepted minimum AUROC curve for a model to be deemed good is 0.7 whilst an AUROC curve of 0.5 would mean that the model does not have discriminant ability and is not different from randomness (Hosmer et al. 2013; Chamboko and Bravo 2019b). Implicit in the use of traditional diagnostic measures are current-status definitions of default. In a more general case of multi-state models, the states in the progression of a mortgage loan change with time and adjustments are necessary to include state transition timing in definitions of prognostic errors rates, i.e., time-dependent ROC curve methods extending the traditional concepts of sensitivity and specificity are needed to characterise prognostic accuracy. Two extensions for classification measures are commonly proposed in the literature. The first corresponds to the consideration of cumulative (prevalent) cases recruited over a fixed period baseline or a landmark starting time point and a future follow-up time point to define cases (see, e.g., Heagerty et al. 2000). In a multi-state approach, the binary classification error concepts have to be extended to risk sets leading to the adoption of an alternative incident cases definition where loans who experience an event at time t are the time-specific cases of interest (see, e.g., mboxcitealpB48-risks-765210). This approach is better suited to the proportional hazards Cox regression model which is based on the fundamental concept of a time-varying risk set of individuals, and associated time-specific "cases" or subjects who experience the event (e.g., delinquency) at a given time. The latter approach was used in this study. For the four transition models analysed, the set of loans at risk of an event were partitioned into a training set of imminent cases (loans which experience the transition event) and a test set of "controls" (loans which have not yet experienced the event). Table 8 shows that the four transition models had good and acceptable predictive performance almost or above the 0.7 threshold. We note that the discriminative power of the models for the initial transitions from current to delinquent and from current to prepaid is higher than that for the recovery process transitions from delinquent or default to current.

Table 8. AUROC curves.

Model	AUROC Curve	Standard Error	95% Confidence Interval
Current to Prepaid	0.7830	0.0021	0.77876–0.78718
Current to Delinquent	0.7968	0.0015	0.79380–0.79980
Delinquent to Current	0.6913	0.0083	0.68359–0.71604
Default to Current	0.7110	0.0024	0.70639–0.71569

The AUROC curves can also be graphically illustrated, as shown on Figures 10–13. The further the ROC curve is to the right of 45° diagonal of the ROC area, the less predictive is the model. On the other hand, the more the curve is to the left, the more predictive it is.

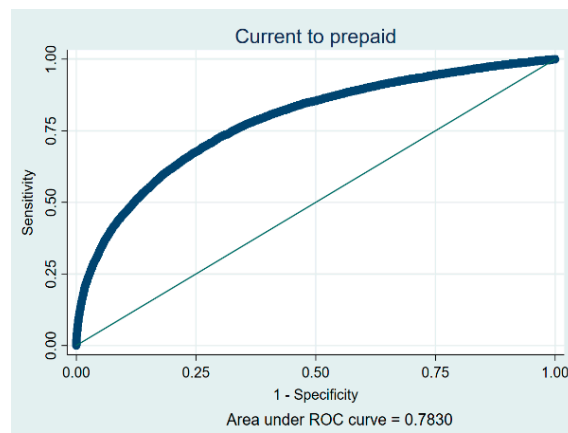


Figure 10. Current to prepaid.

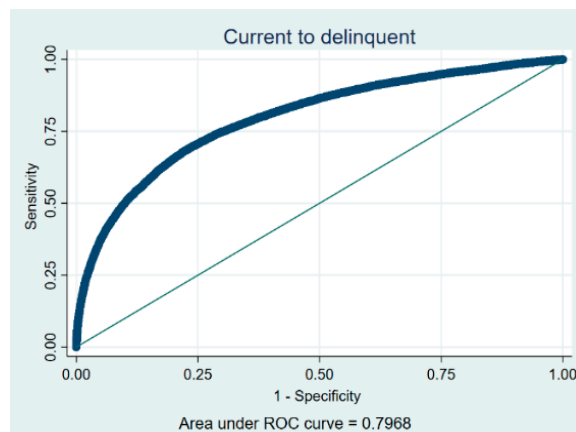


Figure 11. Current to delinquent.

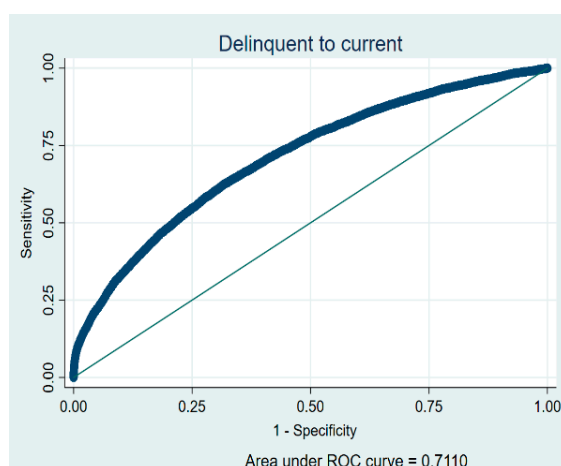


Figure 12. Delinquent to current.

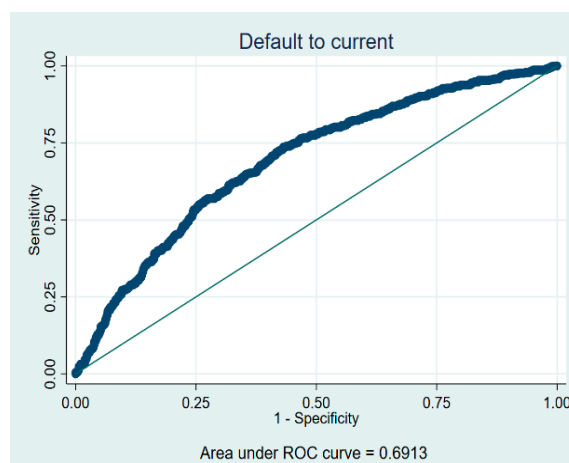


Figure 13. Default to current.

We note that, since the loans status changes over time, as does its risk characteristics and other exogenous conditions, the model's ability to discriminate between cases and controls also changes and cannot be directly compared with that of traditional binary outcome models.

5. Conclusions

In this paper, we adopt a multi-state approach to modelling the progression of borrowers from one state to another to fully understand the risks of a cohort of borrowers over time. We introduce a multi-state framework with seven allowable states: current (normal performance), delinquent, default, repurchase, foreclosure, short sale and prepaid. Additionally, we investigate the relationship between the probability of loans transitioning to and from various loan outcomes using both acquisition and performance explanatory variables. We tested this framework using the Fannie Mae data with a cohort of borrowers whose loans were initiated during the first quarter of 2009 and were followed for 92 months until the third quarter of 2016. Our findings have broader policy implications for contract design, lender and borrower behaviour analysis; for the mitigation of defaults and foreclosures through the design of debt relief programs and mortgage modification policies; for pricing of mortgage-backed securities; and for the design of laws protecting distressed borrowers.

The transition matrix shows that about 11.6% of the borrowers did not transition to any other state but remained current on their payments during the study follow-up period. Conversely, this means the other 88.4% of the borrowers transitioned into some risky state according to the contractual agreement.

As high as 71.4% of the borrowers transitioned from the current state into prepaid, thus having paid the mortgage loan more quickly than the contractually agreed time. This shows a massive risk to lenders as potential income from interest is lost because of prepayment. Moreover, since high-quality borrowers are more likely to prepay (or renegotiate), there is a real risk that the prepayment option is exercised by high-quality borrowers, thus lowering the average quality of the loan pool. However, the probability of prepayment rose towards the end of the study follow up period, suggesting that prepayment penalties could be in effect helping to preserve the contract provisions for the first few years. The results also show that 16.9% of the borrowers missed at least one payment, thus transitioned from current into a delinquent state. Of those who entered a delinquent state, most of them (70.6%) recovered (cured), while 27.1% defaulted. Importantly, about three quarters of those who entered a default state also recovered to a current state. Overall, there were reasonable recovery rates for those who entered delinquency and default state, signalling the ability of borrowers to recover from financial distress. Even though recovery of the defaulters was reasonably high, another quarter transitioned to an absorbing state with most of them (12.7%) progressing into foreclosure, whilst 6% and 5.4%, respectively, transitioned to short sale and prepaid. In terms of the factors affecting the transition into various loan outcomes, we see the cross-cutting importance of ability to pay as measured by debt-to-income ratio, equity as marked by loan-to-value ratio, interest rates, and the property type.

We conclude that jointly modelling the described state occupations and transitions allows a system-wide helicopter view, which provides a holistic understanding of the dynamics of a mortgage loan portfolio more than just modelling prepayment, foreclosure, delinquency, recovery, repurchase and short sale separately. We therefore recommend that, during times of economic distress, the focus should go beyond modelling defaults and foreclosures as the only main outcomes, especially on mortgage loans as other transitions and state occupation experienced by borrowers becomes critically essential.

One powerful feature of the multi-state class of models is the potential for the inclusion of individual, loan-level or macroeconomic time-dependent covariates which explain the transition intensity between states. The loan-specific probabilities of transitioning between states change in response to variations in loan and borrower characteristics, property characteristics or behavioural variables. They also change in response to changes in macroeconomic conditions. For instance, in our model, the household's income enters indirectly via the debt-to-income (DTI) ratio, house price information enters via the Loan to Value ratio (LTV) and the yield curve dynamics enters through loan-specific interest rate information. In this paper, we model intensities as functions of both acquisition and performance loan level explanatory variables considering only idiosyncratic factors as well as dynamic variables that depends on past transition history.

For retail portfolios, a natural extension of this paper being worked out is to explicitly include common macroeconomic time variables (e.g., employment rates, gross domestic product change, consumption, stock market performance, consumer price index and share of non-performing loans) as covariates to explain the transition intensity between states. The recently introduced IFRS9 standards require financial institutions to calculate expected loss for the banking book over the entire life of the exposures, conditional on macroeconomic factors, on a point-in-time basis (and not through-the-cycle, i.e., neutralising economic fluctuations as required under the Basel framework), including forward-looking information. The inclusion of common macroeconomic variables as drivers of loan or portfolio default and of credit risk migration can be done in several ways, e.g., by combining a point-in-time credit scoring model with only idiosyncratic factors (e.g., age, income, past credit behaviour, delinquency history and loan amount) with an ex-post inclusion of the common macroeconomic time variables. One example is [Stein et al. \(2010\)](#) which use a Cox proportional hazard model with time-varying covariates to represent the systematic risk factors for mortgage portfolios and model survival times. Another option is to consider general multinomial scoring transition models for each row of the transition matrix (see, e.g., [Nyström and Skoglund 2006](#)). Another approach entails the use of dynamic Markov credit scoring models which explicitly incorporate idiosyncratic and common macroeconomic credit factors as well as past transition behaviour (see, e.g., [Skoglund and Chen 2016](#)).

These models typically entail more complex path-dependent structures in which delinquency history is important to forecast credit losses, repayment behaviour, revenue generation, liquidity, capital requirements and other measures in a total balance sheet approach. For simpler applications, further research may be conducted empirically to compare the predictive accuracy of the multi-state model with that of traditional two-state credit scoring models tracking default.

Author Contributions: Conceptualization, R.C.; methodology, R.C. and J.M.B.; software, R.C.; validation, R.C.; software and formal analysis, R.C.; investigation, R.C.; resources, R.C. and J.M.B.; data curation, R.C.; writing—original draft preparation, R.C.; writing—review and editing, R.C. and J.M.B.; visualization, R.C. and J.M.B.; supervision, J.M.B. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by national funds through FCT (Fundação para a Ciência e a Tecnologia) under the project UIDB/04152/2020 - Centro de Investigação em Gestão de Informação (MagIC).

Conflicts of Interest: The authors declare no conflict of interest.

References

- Altman, Edward I. 1968. Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *Journal of Finance* 23: 589–609. [\[CrossRef\]](#)
- Aalen, Odd O., and Søren Johansen. 1978. An empirical transition matrix for nonhomogeneous Markov chains based on censored observations. *Scandinavian Journal of Statistics* 5: 141–50.
- Abellán, Joaquín, and Javier G. Castellano. 2017. A comparative study on base classifiers in ensemble methods for credit scoring. *Expert Systems with Applications* 73: 1–10. [\[CrossRef\]](#)
- Agarwal, Sumit, Gene Amromin, Itzhak Ben-David, Souphala Chomsisengphet, Tomasz Piskorski, and Amit Seru. 2017. Policy Intervention in Debt Renegotiation: Evidence from Home Affordable Modification Program. *Journal of Political Economy* 125: 654–712. [\[CrossRef\]](#)
- Agarwal, Sumit, Gene Amromin, Souphala Chomsisengphet, Tim Landvoigt, Tomasz Piskorski, Amit Seru, and Vincent Yao. 2015. *Mortgage Refinancing, Consumer Spending, and Competition: Evidence from the Home Affordable Refinancing Program*. NBER Working Paper No. 21512. Cambridge: National Bureau of Economic Research.
- Ahluwat, Samit. 2018. Evaluation of Mortgage Default Characteristics Using Fannie Mae’s Loan Performance Data. *Journal of Real Estate Finance & Economics* 97: 58–69.
- Andersen, Per Kragh, and Niels Keiding. 2002. Multi-state models for event history analysis. *Statistical Methods in Medical Research* 11: 91–115. [\[CrossRef\]](#)
- Arminger, Gerhard, Daniel Enache, and Thorsten Bonne. 1997. Analyzing credit risk data: A comparison of logistic discrimination, classification tree analysis, and feedforward networks. *Computational Statistics* 12: 293–310.
- Aron, Janine, and John Muellbauer. 2010. *Modelling and Forecasting UK Mortgage Arrears and Possessions*. London: Department for Communities and Local Government.
- Aron, Janine, and John Muellbauer. 2016. Modelling and forecasting mortgage delinquency and foreclosure in the UK. *Journal of Urban Economics* 94: 32–53. [\[CrossRef\]](#)
- Baesens, Bart, Tony Van Gestel, Stijn Viaene, Maria Stepanova, Johan Suykens, and Jan Vanthienen. 2003. Benchmarking State-of-the-Art Classification Algorithms for Credit Scoring. *The Journal of the Operational Research Society* 54: 627–35. [\[CrossRef\]](#)
- Bajari, Patrick, Chenghuan Sean Chu, and Minjung Park. 2008. *An Empirical Model of Subprime Mortgage Default from 2000 to 2007*. NBER Working Paper 14625. Cambridge: National Bureau of Economic Research.
- Bajari, Patrick, Chenghuan Sean Chu, Denis Nekipelov, and Minjung Park. 2013. *A Dynamic Model of Subprime Mortgage Default: Estimation and Policy Implications*. Working Paper 18850. Cambridge: National Bureau of Economic Research.
- Bellotti, Tony, and Jonathan Crook. 2013. Forecasting and stress testing credit card default using dynamic models. *International Journal of Forecasting* 29: 563–74. [\[CrossRef\]](#)
- Beyersmann, Jan, Arthur Allignol, and Martin Schumacher. 2012. *Competing Risks and Multi-State Models with R*. New York: Dordrecht: Heidelberg: London: Springer.
- Black, Fischer, and Myron Scholes. 1973. The pricing of options and corporate liabilities. *Journal of Political Economy* 81: 637–54. [\[CrossRef\]](#)

- Bhutta, Neil, Jane Dokko, and Hui Shan. 2010. *The Depth of Negative Equity and Mortgage Default Decisions*. FEDS Series. Washington, DC: Board of Governors of the Federal Reserve System, pp. 2010–35.
- Bhutta, Neil, Jane Dokko, and Hui Shan. 2017. Consumer Ruthlessness and Mortgage Default during the 2007 to 2009 Housing Bust. *The Journal of Finance* 72: 2433–66. [\[CrossRef\]](#)
- Bravo, Jorge Miguel Ventura, and Carlos Manuel Pereira da Silva. 2006. Immunization Using a Stochastic Process Independent Multifactor Model: The Portuguese Experience. *Journal of Banking and Finance* 30: 133–56. [\[CrossRef\]](#)
- Butaru, Florentin, Qingqing Chen, Brian Clark, Sanmay Das, Andrew W. Lo, and Akhtar Siddique. 2016. Risk and risk management in the credit card industry. *Journal of Banking and Finance* 72: 218–39. [\[CrossRef\]](#)
- Campbell, John Y. 2013. Mortgage Market Design. *Review of Finance* 17: 1–33. [\[CrossRef\]](#)
- Carranza, Juan Esteban, and Dairo Estrada. 2013. Identifying the determinants of mortgage default in Colombia between 1997 and 2004. *Annals of Finance* 9: 501–18. [\[CrossRef\]](#)
- Castro, Vítor. 2013. Macroeconomic determinants of the credit risk in the banking system: The case of the GIPSI. *Economic Modelling* 31: 672–83. [\[CrossRef\]](#)
- Charlier, Erwin, and Arjan Van Bussel. 2003. Prepayment behavior of Dutch mortgages: An empirical analysis. *Real Estate Economics* 31: 165–204. [\[CrossRef\]](#)
- Chamboko, Richard, and Jorge M. Bravo. 2016. On the modelling of prognosis from delinquency to normal performance on retail consumer loans. *Risk Management* 18: 264–87. [\[CrossRef\]](#)
- Chamboko, Richard, and Jorge M. Bravo. 2019a. Frailty correlated default on retail consumer loans in Zimbabwe. *International Journal of Applied Decision Sciences* 12: 271–87. [\[CrossRef\]](#)
- Chamboko, Richard, and Jorge M. Bravo. 2019b. Modelling and forecasting recurrent recovery events on consumer loans. *International Journal of Applied Decision Sciences* 12: 257–70. [\[CrossRef\]](#)
- Chan, Sewin, Claudia Sharygin, Vicki Been, and Andrew Haughwout. 2014. Pathways After Default: What Happens to Distressed Mortgage Borrowers and Their Homes? *Journal of Real Estate Finance and Economics* 48: 342–79. [\[CrossRef\]](#)
- Crook, Jonathan N., David B. Edelman, and Lyn C. Thomas. 2007. Recent developments in consumer credit risk assessment. *European Journal of Operational Research* 183: 1447–65. [\[CrossRef\]](#)
- Danis, Michelle A., and Anthony Pennington-Cross. 2005. A Dynamic Look at Subprime Loan Performance. *Journal of Fixed Income* 15: 28–39. [\[CrossRef\]](#)
- Deng, Yongheng, John M. Quigley, and Robert Van Order. 2000. Mortgage Terminations, Heterogeneity and the Exercise of Mortgage Options. *Econometrica* 68: 275–307. [\[CrossRef\]](#)
- Deng, Yongheng, John M. Quigley, Robert Van Order, and Freddie Mac. 1996. Mortgage default and low downpayment loans: The costs of public subsidy. *Regional Science and Urban Economics* 26: 263–85. [\[CrossRef\]](#)
- Di Maggio, Marco, Amir Kermani, Benjamin J. Keys, Tomasz Piskorski, Rodney Ramcharan, Amit Seru, and Vincent Yao. 2017. Interest Rate Pass-Through: Mortgage Rates, Household Consumption, and Voluntary Deleveraging. *American Economic Review* 107: 3550–88. [\[CrossRef\]](#)
- Du Jardin, Philippe, and Eric Séverin. 2011. Predicting corporate bankruptcy using a self-organizing map: An empirical study to improve the forecasting horizon of a financial failure model. *Decision Support Systems* 51: 701–11. [\[CrossRef\]](#)
- Elul, Ronel, Nicholas S. Souleles, Souphala Chomsisengphet, Dennis Glennon, and Robert Hunt. 2010. What “Triggers” Mortgage Default? *The American Economic Review* 100: 490–94. [\[CrossRef\]](#)
- Foote, Christopher L., Kristopher Gerardi, and Paul S. 2008. Negative equity and foreclosure: Theory and evidence. *Journal of Urban Economics* 64: 234–45. [\[CrossRef\]](#)
- Fuster, Andreas, and Paul S. Willen. 2017. Payment Size, Negative Equity, and Mortgage Default. *American Economic Journal: Economic Policy* 9: 167–91. [\[CrossRef\]](#)
- Gerardi, Kristopher, Adam Shapiro, and Paul Willen. 2007. *Subprime Outcomes: Risky Mortgages, Homeownership and Foreclosure*. Technical Report Working Paper 07–15. Atlanta: Federal Reserve Bank of Atlanta.
- Gerardi, Kristopher, Kyle Herkenhoff, Lee E. Ohanian, and Paul Willen. 2013. *Unemployment, Negative Equity and Strategic Default*. Working Papers Series. Atlanta: Federal Reserve Bank of Atlanta.
- Grimshaw, Scott D., and William P. Alexander. 2011. Markov chain models for delinquency: Transition matrix estimation and forecasting. *Applied Stochastic Models in Business Industry* 27: 267–79. [\[CrossRef\]](#)
- Guiso, Luigi, Paola Sapienza, and Luigi Zingales. 2009. *Moral and Social Constraints to Strategic Default on Mortgages*. NBER Working Paper No. 15145. Cambridge: National Bureau of Economic Research.

- Gupton, Gred M., Christopher Clemens Finger, and Mickey Bhatia. 1997. *CreditMetricsTM—Technical Document*. New York: JP Morgan, pp. 1–212.
- Guren, Adam M., Arvind Krishnamurthy, and Timothy J. McQuade. 2018. *Mortgage Design in an Equilibrium Model of the Housing Market*. NBER Working Paper 24446. Cambridge: National Bureau of Economic Research.
- Ha, Sung Ho. 2010. Behavioral assessment of recoverable credit of retailer's customers. *Information Sciences* 180: 3703–17. [[CrossRef](#)]
- Ha, Sung Ho, and Ramayya Krishnan. 2012. Predicting repayment of the credit card debt. *Computers and Operations Research* 39: 765–73. [[CrossRef](#)]
- Hand, David J., and William E. Henley. 1997. Statistical Classification Methods in Consumer Credit Scoring: A Review. *Journal of the Royal Statistical Society Series A (Statistics in Society)* 160: 523–41. [[CrossRef](#)]
- Heagerty, Patrick J., Thomas Lumley, and Margaret S. Pepe. 2000. Time-dependent ROC curves for censored survival data and a diagnostic marker. *Biometrics* 56: 337–44. [[CrossRef](#)]
- Heagerty, Patrick J., and Yingye Zheng. 2005. Survival model predictive accuracy and ROC curves. *Biometrics* 61: 92–105. [[CrossRef](#)]
- Hosmer, W. David, Stanley Lemeshow, and X. Rodney Sturdivant. 2013. *Applied Logistic Regression*, 3rd ed. Wiley Series in Probability and Statistics. Hoboken: Wiley.
- Jones, Timothy, Dean Gatzlaff, and G. Stacy Sirmans. 2016. Housing Market Dynamics: Disequilibrium, Mortgage Default, and Reverse Mortgages. *Journal of Real Estate Finance and Economics* 53: 269–81. [[CrossRef](#)]
- Kaplan, Edward L., and Paul Meier. 1958. Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association* 53: 457–81. [[CrossRef](#)]
- Kau, James B., Donald C. Keenan, Walter J. Muller, and James F. Epperson. 1992. A Generalized Valuation Model for Fixed-Rate Residential Mortgages. *Journal of Money, Credit and Banking* 24: 279–99. [[CrossRef](#)]
- Kau, James B., Donald C. Keenan, and Taewon Kim. 1993. Transaction costs, suboptimal termination and default probabilities. *Real Estate Economics* 21: 247–63. [[CrossRef](#)]
- Kelly, Robert, and Terence O'Malley. 2016. The good, the bad and the impaired: A credit risk model of the Irish mortgage market. *Journal of Financial Stability* 22: 1–9. [[CrossRef](#)]
- Kruppa, Jochen, Alexandra Schwarz, Gerhard Armingier, and Andreas Ziegler. 2013. Consumer credit risk: Individual probability estimates using machine learning. *Expert Systems with Applications* 40: 5125–31. [[CrossRef](#)]
- Lessmann, Stefan, Bart Baesens, Hsin-Vonn Seow, and Lyn C. Thomas. 2015. Benchmarking state-of-the-art classification algorithms for credit scoring: An update of research. *European Journal of Operational Research* 247: 124–36. [[CrossRef](#)]
- Liu, Bo, and Tien Foo Sing. 2018. "Cure" Effects and Mortgage Default: A Split Population Survival Time Model. *Journal of Real Estate Finance and Economics* 56: 217–51. [[CrossRef](#)]
- Mayer, Christopher, Karen Pence, and Shane M. Sherlund. 2009. The Rise in Mortgage Defaults. *Journal of Economic Perspectives* 23: 27–50. [[CrossRef](#)]
- Mayer, Christopher, Edward Morrison, Tomasz Piskorski, and Arpit Gupta. 2014. Mortgage modification and strategic behavior: Evidence from a legal settlement with Countrywide. *American Economic Review* 104: 2830–57. [[CrossRef](#)]
- McKinsey and Company. 1998. *CreditPortfolioViewTM Approach Documentation and User's Documentation*. Zurich: McKinsey and Company.
- Mesnard, Benoit, A. Margerit, Cairen Power, and Marcel Magnus. 2016. Non-Performing Loans in the Banking Union: Stocktaking and Challenges. European Parliament, IPOL-EGOV. Available online: [http://www.europarl.europa.eu/RegData/etudes/BRIE/2016/574400/IPOL_BRI\(2016\)574400_EN.pdf](http://www.europarl.europa.eu/RegData/etudes/BRIE/2016/574400/IPOL_BRI(2016)574400_EN.pdf) (accessed on 20 June 2019).
- Mian, Atif, Amir Sufi, and Francesco Trebbi. 2015. Foreclosures, House Prices, and the Real Economy. *Journal of Finance* 70: 2587–634. [[CrossRef](#)]
- Nyström, Kaj, and Jimmy Skoglund. 2006. A Credit Risk Model for Large Dimensional Portfolios with Application to Economic Capital. *Journal of Banking & Finance* 30: 2163–97.
- Noh, Hyun Ju, Tae Hyup Roh, and Ingoo Han. 2005. Prognostic personal credit risk model considering censored information. *Expert Systems with Applications* 28: 753–62. [[CrossRef](#)]

- Ncube, Mthuli, and Stephen E. Satchell. 1994. *Modelling UK Mortgage Defaults Using a Hazard Approach Based on American Options*. Working Paper 8. Cambridge: Department of Applied Economics, University of Cambridge.
- Ozkan, Mr F. Gulcin, and D. Filiz Unsal. 2012. *Global Financial Crisis, Financial Contagion, and Emerging Markets*. IMF Working Papers. Washington, DC: International Monetary Fund, vol. 12, p. 1.
- Piskorski, Tomasz, Amit Seru, and Vikrant Vig. 2010. Securitization and Distressed Loan Renegotiation: Evidence from the Subprime Mortgage Crisis. *Journal of Financial Economics* 97: 369–97. [[CrossRef](#)]
- Piskorski, Tomasz, and Amit Seru. 2018. *Mortgage Market Design: Lessons from the Great Recession*. Brookings Papers on Economic Activity, Economic Studies Program. Washington, DC: The Brookings Institution, vol. 49, pp. 429–513.
- Putter, Hein, Marta Fiocco, and Ronald B. Geskus. 2007. Tutorial in biostatistics: Competing risks and multi-state models. *Statistics in Medicine* 26: 2389–430. [[CrossRef](#)] [[PubMed](#)]
- Riddiough, Timothy J. 1991. Equilibrium Mortgage Default Pricing with Non-Optimal Borrower Behavior. Ph.D. dissertation, University of Wisconsin, Madison, WI, USA.
- Sarlija, Natasa, Mirta Bencic, and Marijana Zekic-Susac. 2009. Comparison procedure of predicting the time to default in behavioural scoring. *Expert Systems with Applications* 36: 8778–88. [[CrossRef](#)]
- Schwartz, Eduardo S., and Walter N. Torous. 1993. Mortgage Prepayment and Default Decisions: A Poisson Regression Approach. *Real Estate Economics* 21: 431–49. [[CrossRef](#)]
- Skoglund, Jimmy, and Wei Chen. 2016. The Application of Credit Risk Models to Macroeconomic Scenario Analysis and Stress Testing. *Journal of Credit Risk* 12: 1–45. [[CrossRef](#)]
- Stein, R. M., A. Das, Y. Ding, and S. Chinchalkar. 2010. *Moody's Mortgage Metrics Prime: A Quasi-Structural Model of Prime Mortgage Portfolio Losses*. New York: Moody's Research Labs.
- Stepanova, Maria, and Lyn Thomas. 2002. Survival analysis methods for personal loan data. *Operations Research* 50: 277–89. [[CrossRef](#)]
- Tian, Chao Yue, Roberto G. Quercia, and Sarah Riley. 2016. Unemployment as an Adverse Trigger Event for Mortgage Default. *Journal of Real Estate Finance and Economics* 52: 28–49. [[CrossRef](#)]
- Tong, Edward NC, Christophe Mues, and Lyn C. Thomas. 2012. Mixture cure models in credit scoring: If and when borrowers default. *European Journal of Operational Research* 218: 132–39. [[CrossRef](#)]
- Tracy, Joseph S., and Joshua Wright. 2016. Payment changes and default risk: The impact of refinancing on expected credit losses. *Journal of Urban Economics* 93: 60–70. [[CrossRef](#)]
- Vandell, Kerry D., Walter Barnes, David Hartzell, Dennis Kraft, and William Wendt. 1993. Commercial Mortgage Defaults: Proportional Hazards Estimation Using Individual Loan Histories. *Real Estate Economics* 21: 451–80. [[CrossRef](#)]
- Vandell, Kerry D. 1995. How ruthless is mortgage default? A review and synthesis of the evidence. *Journal of Housing Research* 6: 245–64.
- Volkov, Andrey, Dries F. Benoit, and Dirk Van den Poel. 2017. Incorporating sequential information in bankruptcy prediction with predictors based on Markov for discrimination. *Decision Support Systems* 98: 59–68. [[CrossRef](#)]
- Whelan, Karl. 2013. *Ireland's Economic Crisis: The Good, the Bad and the Ugly*. UCD Working Paper. Dublin: University College Dublin.

