



# Mortgage Pricing Implications of Prepayment: Separating Pecuniary and Non-pecuniary Prepayment

Lu Fang<sup>1</sup> 

Published online: 21 December 2019

© Springer Science+Business Media, LLC, part of Springer Nature 2019

## Abstract

In the study of mortgage loan pricing, prepayment and default hazards are considered. While default results in loss of initial capital, prepayment is the more frequent termination event. This study makes a distinction between pecuniary and non-pecuniary prepayments in the mortgage pricing model. Pecuniary/non-pecuniary prepayments are distinguished from each other based on whether the market interest rate at prepayment is below/above the rate at origination, and thus whether mortgage lenders/investors can reinvest the proceeds at a lower/higher interest rate using the proceeds from a prepaid loan. Using a sample of 30-year fixed-rate mortgage loans for home purchase, this study finds that pecuniary and non-pecuniary prepayments are affected differently by certain overlapping factors and are driven by some unique factors. The results also show that combining these two types of prepayments into a single prepayment measure may yield inaccurate predictions of loan termination probabilities. The results from the mortgage pricing model indicate that pecuniary and non-pecuniary prepayment risks, together with default risk, contribute separately to the pricing of a loan.

**Keywords** Mortgages · Prepayment · Option values · Interest rate

**JEL Classifications** G21 · E43 · E47 · R20

## Introduction

Prepayment hazard modelling is of great importance in forecasting the expected future cash flows from a mortgage loan, and hence is essential to the valuation of mortgage loans and mortgage-backed securities (MBS). Mortgage lenders, firms engaged in mortgage securitization, and investors may suffer losses if the actual prepayment rate

---

✉ Lu Fang  
fangl@longwood.edu

<sup>1</sup> College of Business & Economics, Longwood University, Farmville, VA 23909, USA

deviates from the expected one. While default has garnered more attention because of the loss of initial capital, mortgage prepayment occurs much more frequently.<sup>1</sup> Thus, understanding a borrower's prepayment propensity is an important issue to mortgage pricing.

The standard way to model a borrower's propensity to prepay a mortgage loan is to adopt the basic principle of option theory that a mortgage loan is embedded with both default and prepayment options. In terms of prepayment, as the interest rate drops, when the current value of a mortgage loan exceeds its par value, the call option is "in-the-money," and a rational borrower should prepay.<sup>2</sup> Accordingly, there has been extensive literature deriving the optimal time for a borrower to prepay (Dunn and McConnell, 1981a & 1981b; Shoven 1986; Hendershott and Van Order 1987; Kau et al. 1992; Follain et al. 1992; McConnell and Singh 1994; Agarwal et al. 2013; among others). However, prior research suggests that the empirically observed mortgage prepayment rates deviate from those predicted by theory (e.g., Giliberto and Thibodeau 1989; Chang and Yavas 2009; Agarwal et al. 2015).<sup>3</sup> In some cases, borrowers fail to exercise the "in-the-money" option, whereas in other cases borrowers prepay when the interest rates had risen.

Various motivations behind prepayment may help explain the differences in prepayment rates, including: 1) sale of the house caused by relocation, divorce, unexpected negative income shocks (e.g., unemployment, unexpected medical bills), and so forth; 2) refinance to obtain a lower interest rate (rate refinance); 3) refinance to extract equity from collateral for consumptions, other debt repayments, or investments (cash-out refinance); 4) refinance to reduce monthly mortgage payments due to improvements in a borrower's creditworthiness<sup>4</sup>; 5) fully repay the mortgage debt due to positive income shocks; and so on. Empirically, it is difficult to disentangle prepayments due to one motivation from those arising from others.<sup>5</sup>

<sup>1</sup> For example, Firestone et al. (2007) note that in their study the average default rate is 0.6%, while the average prepayment rate is 92%.

<sup>2</sup> The current value of a mortgage loan has incorporated the value of default and prepayment options in the future.

<sup>3</sup> For example, Giliberto and Thibodeau (1989) found that some borrowers failed to prepay although their call options were deeply "in-the-money" in the 1980s when the interest rates dropped. Chang and Yavas (2009) found that certain borrowers prepaid too early during the period from 1996 to 2003. Agarwal et al. (2015) show that approximately 57% of borrowers refinance sub-optimally – 50% choose the wrong rate, 17% wait too long, and 10% do both. Several explanations have been proposed for this discrepancy, including various prepayment motivations by borrowers, prepayment transaction costs (Stanton 1995; Follain et al. 1992; Archer et al. 1997; Dunn and Spatt 2005), institutional constraints (Archer et al. 1996; Deng and Gabriel 2006; Peristiani et al. 1997; Green and LaCour-Little 1999), borrower sophistication (Agarwal et al. 2015), rational inattention (Agarwal et al. 2015), and so forth.

<sup>4</sup> Prepayment due to improvements in a borrower's creditworthiness might be common in the subprime mortgage markets. It may occur even when the market interest rate rises.

<sup>5</sup> The difficulty in separating various types of prepayments from each other is common in most mortgage studies because the data often lack information on the reason why a borrower chooses to prepay. Although several prior studies have attempted to separate prepayments due to different motivations from each other (LaCour-Little 1999; Pavlov 2001; Clapp et al. 2000, 2001; Pennington-Cross and Chomsisengphet 2007; LaCour-Little et al. 2010), it is almost impossible to completely separate one type of prepayment from all others. For example, a borrower who would like to cash out housing equity may refinance his mortgage loan when the market interest rate drops. Thus, rate refinance cannot be completely distinguished from cash-out refinance.

Following some prior theoretical studies on mortgage terminations and pricing of mortgage loans and MBSs (Dunn and McConnell 1981a, b; Brennan and Schwartz 1985; Kau et al. 1992; Archer and Ling 1993), the current study attempts to model a borrower's prepayment behavior by making a distinction between pecuniary and non-pecuniary prepayments and to incorporate both of them in mortgage pricing. Pecuniary prepayment occurs when the market interest rate at the time of prepayment has dropped below that at the time of origination; while non-pecuniary prepayment occurs when the market interest rate at prepayment is greater than or equal to the rate at origination.<sup>6</sup> Both types of prepayment disturb the timing and amount of the expected cash flows of a mortgage loan, and prevent lenders/MBS investors from receiving the anticipated future interest rate payments. Additionally, both of them affect the effective maturity and duration of MBSs. MBS investors typically forecast the prepayment rate of a mortgage pool based on which they also anticipate effective maturity and duration of their MBSs, calculate the MBS yield, and determine the strategy to hedge against risks associated with their MBSs in their portfolio (Hancock and Passmore 2011).<sup>7</sup> As hedging MBS is costly, a failure to accurately account for pecuniary or non-pecuniary prepayment probability, which results in an inaccurate estimate of the effective maturity and duration of the MBS, would incur extra hedging costs. Besides, given the competing-risks nature of various forms of loan termination, the occurrence of non-pecuniary prepayment would limit the value of either default or prepayment option, and thus impacts mortgage/MBS price (Kau et al. 1992; Archer and Ling 1993). Overall, as stated in the prior theoretical studies (e.g. Dunn and McConnell, 1981), in an efficient market, the possible occurrences of both pecuniary and non-pecuniary prepayment will be reflected in the price of a mortgage loan/MBS, and a mortgage pricing model should incorporate not only pecuniary prepayment but also non-pecuniary prepayment.

Also indicated by those prior theoretical studies, these two types of prepayment may be determined by different factors. While pecuniary prepayment is believed to be mainly driven by a decline of market interest rate, non-pecuniary prepayment occurs due to some non-interest-rate reasons. In addition, these two types of prepayment may result in different consequences. When pecuniary prepayment occurs, lenders/investors can only reinvest the proceeds from a prepaid loan at a relatively lower interest rate. By contrast, upon the occurrence of non-pecuniary prepayment, lenders/investors can reinvest the proceeds at a relatively higher interest rate. Additionally, as pointed out by Archer and Ling (1993), in a pool of mortgage loans with heterogeneous prepayment transaction costs, due to the "burnout" phenomenon,<sup>8</sup> the occurrence of pecuniary

<sup>6</sup> Some prior studies named the pecuniary prepayment as optimal prepayment or endogenous call, and called the nonpecuniary prepayment as suboptimal (non-optimal) prepayment or exogenous prepayment (Dunn and McConnell, 1981a, b; Brennan and Schwartz 1985; Kau et al. 1992; Archer and Ling 1993).

<sup>7</sup> For example, a classic way to hedge an MBS is to purchase an interest rate swap, for which MBS investors receive the variable-rate leg of the swap and pay the fixed-portion of the swap. When interest rate rises, this interest rate swap can help MBS investors mitigate the loss in the value on the MBS. The hedging strategy normally requires an estimate of the effective maturity of the MBS to determine the term of the interest rate swap.

<sup>8</sup> Burnout refers to a phenomenon in which MBS pecuniary prepayments rate slows down over time despite lowering interest rate, because borrowers in the pool with the ability to refinance have prepaid their loans earlier in the interest rate drop cycle, leaving the remainder in the pool who are unable to do so due to some reasons including a drop in their credit scores, or lack of property equity, and so forth.

prepayment truncates the distribution of the call price after the consideration of the transaction costs, while non-pecuniary prepayment would not have an impact on it.<sup>9</sup> Thus they conclude that a prepayment model must distinguish between pecuniary and non-pecuniary prepayments.

This study utilizes a competing-risks hazard model which accounts for pecuniary and non-pecuniary prepayments as well as default hazards. The paper examines whether market conditions (interest rate change, local house price change and volatility), loan characteristics, borrower traits, underlying property characteristics, and/or neighborhood traits have different impacts on pecuniary vs. non-pecuniary prepayment hazards. In addition, a loan pricing behavior is examined when cumulative probabilities of default, pecuniary prepayment, and non-pecuniary prepayment for each loan, as well as simulated future interest rate paths based on a term structure model, are included in a contract rate determination model to test if pecuniary and non-pecuniary prepayment risks separately affect the contract rate of a loan.

The analysis is based on a sample of 30-year first-lien fixed-rate nonconforming conventional mortgage loans for home purchases in Miami-Dade County, FL originated from 1997 to 2006, with loan performance being observed between 2000 and 2010. The results clearly show that pecuniary and non-pecuniary prepayments are induced by different factors and affected differently by certain overlapping factors. Not surprisingly, the market interest rate change has a significant impact on the likelihood of pecuniary prepayment, but not on the probability of non-pecuniary prepayment. The results indicate that borrowers living in an affluent neighborhood tend to be more likely to pecuniarily prepay and less likely to non-pecuniarily prepay compared to borrowers living in a relatively poorer neighborhood. The results also show combining these two distinct types of prepayment into a single measure shifts some coefficients towards zero and hence yield inaccurate predictions of loan termination probabilities. The contract rate determination model reveals that both pecuniary and non-pecuniary prepayment risks contribute when pricing mortgages.

The empirical results of the current study hence demonstrate that separating non-pecuniary and pecuniary prepayments would help mortgage lenders, financial intermediates, and MBS investors make some improvements on mortgage termination predictions and loan pricing, especially given the recent launch of the Uniform Mortgage-Backed Security (UMBS) program. In June, 2019, under the direction of the Federal Housing Finance Agency (FHFA), Fannie Mae and Freddie Mac started issuing a new common MBS, called the UMBS, in replacement of their current offerings of To-Be-Announced (TBA)-eligible MBSs.<sup>10</sup> The major goal of this program is to improve the liquidity of the overall TBA MBSs market, and thus the liquidity of the housing finance market. A major concern on this UMBS launch is the loan quality in terms of default and prepayment risks should be very similar for Fannie and Freddie mortgage loans, otherwise investors may see a “race to the bottom” in loan quality.<sup>11</sup> The major finding of this study indicates a mortgage pricing model which distinguishes pecuniary and

<sup>9</sup> In other words, pecuniary prepayment causes the burnout phenomenon, while non-pecuniary prepayment does not. Please see Archer and Ling (1993) for details.

<sup>10</sup> For more information on the UMBS launch, please see <https://www.fhfa.gov/Media/PublicAffairs/Pages/FHFA-Announces-June-2019-Implementation-of-the-New-UMBS.aspx>

<sup>11</sup> <https://www.nationalmortgage.com/articles/uniform-mortgage-backed-security-may-prove-a-two-edged-sword-for-mortgage-investors>

non-pecuniary prepayments and incorporates both of them in loan pricing should be considered by Government-Sponsored Enterprises (GSEs) and regulators to ensure that prepayment and default risks are truly similar.

The paper is organized as follows. The second section offers an overview and a discussion of previous studies on mortgage prepayment hazard modelling. The third section explains the model in details. The fourth section describes the data and explains model specifications. The fifth section presents the results. The final section provides the conclusion.

## Literature Review

Prepayment has long been of interest to both scholars and practitioners interested in the valuation of mortgage loans and mortgage-backed securities. Earlier studies, both theoretical and empirical, emphasize the basic principle of option theory that prepayments are simply assumed to be mainly induced by interest rate fluctuations. However, the observed borrowers' prepayment behavior deviated significantly from the predictions of the option-based model (e.g., Giliberto and Thibodeau 1989).

Recent studies on prepayment have attempted to mimic the complexity of the option-based model by incorporating additional explanatory variables, especially variables describing the characteristics of the borrower and the loan (e.g., LaCour-Little 1999; Bennett et al. 2000, 2001). A number of prior studies also emphasize the role of transaction costs, which may vary across borrowers, in a borrower's decision to prepay (e.g., Follain et al. 1992; Stanton 1995; Archer et al. 1997; Dunn and Spatt 2005). Institutional constraints may also play an important role in affecting a borrower's decision to prepay a loan, as they affect a borrower's ability to qualify for a new mortgage loan. Overall, previous studies indicate that the probability of a borrower exercising the prepayment option is positively affected by their post-origination income (Archer et al. 1996; Deng and Gabriel 2006), equity in housing (Archer et al. 1996; Peristiani et al. 1997; Green and LaCour-Little 1999), and credit ratings (Peristiani et al. 1997; Bennett et al. 2001). Additionally, the role of the broker (LaCour-Little and Chun 1999), the impact of prepayment penalties (Mayer et al. 2013; Beltratti et al. 2017; etc.), the effects of points (Chang and Yavas 2009), and the consideration of a household's life cycle (Quigley 1987) may also affect a borrower's propensity to prepay.

Other recent literature attempts to distinguish among various motivations for mortgage prepayments. Using unique loan-level data, LaCour-Little (1999) focuses on the "pure" refinancing behavior after eliminating borrower mobility and liquidity demand factors, while explicitly controlling for borrower and loan characteristics.<sup>12</sup> He finds that borrowers and loan traits do affect "pure" refinancing risk, though primarily in the region where the prepayment option is "at-the-money." By matching mortgage termination data to housing transaction data, Pavlov (2001), Clapp et al. (2000, 2001) disentangle prepayments arising from mobility from those due to in-place refinancing, and investigate whether these two types of prepayments are driven by different factors within a competing-

<sup>12</sup> "Pure" refinancing by a borrower is defined as the one when a borrower prepays a mortgage loan only for rate change or term change (to alter monthly payment schedule) with the same lender.

risk model framework. Pavlov (2000) finds that in-place refinance risk is highly sensitive to interest rate change and the value of a mortgage, while mobility risk is sensitive to local economic conditions and independent of the value of the mortgage. Clapp et al. (2000) show that borrowers with high expected mobility/refinance risk are not more likely to refinance/move, indicating the necessity of differentiating movers from refiners when evaluating the prepayment risk of a loan. Clapp et al. (2001) demonstrate that refinances are driven more by financial incentives like the market value of a loan, while a borrower's characteristics including age, income, and minority status tend to have more influence on the move decision. They also show that combining these two distinct choices into a single prepayment measure may yield an inaccurate prediction of a loan's termination risk. In addition, some studies (e.g., Pennington-Cross and Chomsisengphet 2007; LaCour-Little et al. 2010; Bhutta and Keys 2016) focus on factors affecting a borrower's cash-out refinancing decision as well as refinancing amount. For example, Pennington-Cross and Chomsisengphet (2007) find that cash-out refinancing is mainly driven by house price appreciation and interest rate drops, and affected negatively by neighborhood-level median income.

Although some prior studies are able to distinguish among various types of (motivations for) prepayments due to their unique datasets, a majority of the mortgage performance studies are not able to and thus employ a single prepayment model. The division of prepayment by the expected interest rate change, as in the current study, represents a common market environment in which to view prepayment. This division is also supported by some prior theoretical studies on mortgage/MBS pricing which highlighted the importance and necessity of distinguishing pecuniary and non-pecuniary prepayments from each other. For example, Dunn and McConnell (1981a, b) and Brennan and Schwartz (1985) use a frictionless optimal call model to explain a borrower's pecuniary prepayment behavior, to which a Poisson process is added in order to model non-pecuniary prepayment. The simulation results by Dunn and McConnell (1981a, b) show that non-pecuniary prepayment has a positive effect on Ginnie Mae (GNMA) MBS price and reduces its interest rate risk and expected return. Brennan and Schwartz (1985) arrived at similar conclusions. Both by specifying a conditional probability of non-pecuniary prepayment, Kau et al. (1992) and Archer and Ling (1993) allow pecuniary and non-pecuniary prepayments to interact with each other, and to disentangle the separate effects of interest rate call and non-pecuniary prepayment on mortgage/MBS price. Kau et al. (1992) concluded that non-pecuniary prepayment in the presence of pecuniary prepayment should increase the value of a mortgage loan to a lender. The simulation results by Archer and Ling (1993) support this conclusion, however, they also found that the impact of non-pecuniary prepayment on the value of call depends on the shape of the yield curve. In addition, empirically, a distinction between pecuniary and non-pecuniary prepayments allows for the examination of important factors and differences in impacts on loan pricing of these factors. Further, this study explores the loan pricing implications of separating pecuniary and non-pecuniary prepayments. Thus, the current study is able to examine the pricing impact of default as well as pecuniary prepayment, and non-pecuniary prepayment on loan pricing.

## Model

### Mortgage Loan Performance Analysis

The borrower's loan termination behavior is modelled, on a monthly basis, using the Cox discrete-time competing-risks loan hazard model (Deng et al. 2000).<sup>13</sup> The loan performance data is restructured to have one observation for each loan and each month until the loan is terminated or until the end of the loan performance observation window. A distinction is made between pecuniary and non-pecuniary prepayments depending on whether the market interest rate at prepayment is below or above the one at loan closing. Therefore, each month, the borrower of a loan is observed to either continue, or default upon, or pecuniarily prepay, or non-pecuniarily prepay a loan. Specifically, a multinomial logit model with four outcomes is utilized to model a borrower's loan termination behavior as in Eq. (1)<sup>14</sup>:

$$\ln\left(p_{jt}/p_{0t}\right) = \delta_j\alpha_{jt} + \beta_j'\mathbf{x}_{jt} + \varepsilon_{jt} \quad j = 1, 2, 3 \quad (1)$$

where  $p_{0t}$  is the probability of a loan being current in period  $t$ , while  $p_{jt}$  is the probability of a loan being terminated either through default ( $j=1$ ), pecuniary prepayment ( $j=2$ ), or non-pecuniary prepayment ( $j=3$ ) in period  $t$  given that this loan is still current at the beginning of period  $t$ . In Eq. (1),  $t$  represents mortgage time.  $\alpha_{jt}$  refers to the baseline hazard rate of each event and is a function of mortgage time. The vector  $\mathbf{x}_{jt}$  includes a set of covariates describing the market conditions (interest rate change and house price change), the characteristics of the loan, the borrower, the underlying property, as well as the neighborhood. Some covariates are time varying while others are time invariant. This multinomial logit model thus allows one to examine whether pecuniary and non-pecuniary prepayments are driven by different factors and whether those covariates have different impacts on them.

<sup>13</sup> The discrete-time nature of this model solves the issue of left truncation and right censoring, which are common issues in mortgage literature. It allows loan default and prepayment probabilities to be estimated appropriately.

<sup>14</sup> The multinomial logit model restricts the sum of the probability of continuation, default, pecuniary prepayment, and non-pecuniary prepayment to one, thus directly controlling for the competing risks. One assumption implied by the multinomial logit model is the Independence of Irrelevant Alternatives (IIA) assumption. This assumption requires that the odds ratio for any pair of choices be independent of any third alternative (one event is not informative to the other, conditional on all of the covariates in the model). In addition, this IIA assumption requires the i.i.d. assumption for a given observation over time (choices at any point in time are independent of those at any other point in time). Thus, this assumption, empirically, requires the inclusion of a full set of covariates, especially time-varying covariates, in the hazard model. Those requirements can mostly be met in this study given the extensive data. Because of this IIA assumption, the default, pecuniary prepayment, and non-pecuniary prepayment hazard models are estimated separately treating other events as censoring, rather than estimating a single multinomial logit model using the same specifications for all of the events. This estimation approach is widely used by other studies and Allison (2010) shows that separate models perform well for most of the data. In addition, one advantage of this approach is the flexibility in specifying different models for different events.

## Mortgage Loan Pricing

On the basis of this discrete-time competing-risks loan hazard model estimation results, this study further mimics mortgage lenders' loan pricing behavior at the time of loan origination, predicting the probabilities of default, pecuniary prepayment, and non-pecuniary prepayment for each loan in the sample. Evaluating the probability of each event separately for each loan allows the investigation of whether these risks are taken into consideration as a part of mortgage pricing.

When predicting loan termination probabilities, the forecast of future market interest rate change is of particular interest to the lenders. Following prior studies, and the fact that most of the loans in the sample of this study were terminated within the first 10 years, this study adopted the 10-year treasury constant maturity yield ( $y$ ) as the benchmark market interest rate for 30-year fixed-rate fully-amortizing residential mortgage loans.<sup>15</sup> The one-factor Cox, Ingersoll and Ross (CIR) term structure model is utilized to predict the possible paths (density) of future market interest rate (10-year treasury yield) seen from loan origination  $y_t(y_0)$ .<sup>16</sup> Specifically, in this CIR term structure model, the whole yield curve is assumed to be driven by a spot interest rate ( $r(t)$ ), and this spot interest rate is assumed to follow a mean-reverting stochastic process with its volatility impacted by the level of the spot interest rate as below:

$$dr(t) = \gamma(\theta - r(t))dt + \sigma\sqrt{r(t)}dz(t) \quad (2)$$

where  $\theta$  is the long-term mean of the spot interest rate,  $\gamma$  is the reversion rate,  $\sigma$  is the volatility, and  $z(t)$  is a Wiener process. Hence,  $\gamma(\theta - r(t))$  is the deterministic part and ensures mean reversion of the spot interest rate towards its long-term mean, and  $\sigma\sqrt{r(t)}dz(t)$  describes the stochastic movements with  $\sigma\sqrt{r(t)}$  as its volatility.

Based on the estimated parameters in Equation (2),<sup>17</sup> the density of future spot interest rate for any forecast interval conditional on the spot interest rate at the time of loan origination  $dF(r(t)|r(0))$  is forecasted.<sup>18</sup> As the whole yield curve is assumed to be driven by the spot interest rate, a change in the future spot interest rate ( $r(t)$ ) would lead to a change in the future 10-year treasury yield, and would further affect the probability of loan termination. Therefore, the forecasted conditional density of future spot interest rate  $dF(r(t)|r(0))$  is used to predict the probabilities of default, pecuniary prepayment,

<sup>15</sup> All of the loans in the sample of this study are 30-year fixed-rate fully-amortizing mortgage loans for home purchase.

<sup>16</sup> The CIR term structure model is chosen as it allows one to predict the density of future interest rates, not just its mean. In addition, the CIR term structure model is widely used in mortgage literature.

<sup>17</sup> The parameters in Equation (2) were estimated with the use of 4 time series of yields with different maturities from 1987 to 2007 within the framework of the single-factor CIR term structure model. Those 4 time series are 6-month T-bill yield, 1-year Fama-Bliss bond yield, 3-year Fama-Bliss bond yield, and 5-year Fama-Bliss bond yield. Data were obtained from CRSP. The reason why this estimation period (from 1987 to 2007) is chosen is many studies have found there was a shift in Federal Reserve monetary policy in the early 1980s (Duan and Simonato (1999)) and the loan data in this study ends in 2007 based on loan origination year. The GAUSS code offered by Jin-Chuan Duan on his website is used to implement the estimation, the one he used to yield the results in Duan and Simonato (1999). This help from him is acknowledged.

<sup>18</sup> This density is forecasted based on the transition density of the spot interest rate implied by the CIR term structure model. For details, please see Cox et al. (1985). Here, a normal distribution was used to closely approximate the true transition density.



and non-pecuniary prepayment for each loan at each mortgage time  $t$  as in Equations (3)–(5).<sup>19</sup>

$$\widehat{p}_{1t}(y_0) = \int p_{1t}[y_t(r(t))] dF(r(t)|r(0)) \quad (3)$$

$$\widehat{p}_{2t}(y_0) = \int_{y_t(r(t)) < y_0} p_{2t}[y_t(r(t))] dF(r(t)|r(0)) \quad (4)$$

$$\widehat{p}_{3t}(y_0) = \int_{y_t(r(t)) \geq y_0} p_{3t}[y_t(r(t))] dF(r(t)|r(0)) \quad (5)$$

Where  $\widehat{p}_{1t}(y_0)$  is the predicted default probability at mortgage time  $t$  seen from loan origination given that this loan has been current till the beginning of period  $t$ ,  $\widehat{p}_{2t}(y_0)$  and  $\widehat{p}_{3t}(y_0)$  are time-specific predicted conditional pecuniary and non-pecuniary prepayment probabilities seen from loan closing, respectively. Recall that, pecuniary prepayment is defined as the one when the market mortgage interest rate at prepayment (10-year treasury yield,  $y_t(r(t))$ ) drops below the rate at origination ( $y_0$ ); while non-pecuniary prepayment is defined as the one when the market interest rate at prepayment is equal to or above the rate at loan closing.

The predicted time-specific conditional probability of each event from Equation (3)–(5) is aggregated over a 5-year window to calculate a predicted cumulative probability of each event  $\widehat{P}_k$  ( $k = 1, 2, 3$ ) seen from origination as in Equation (6).<sup>20</sup>

<sup>19</sup> In this study, the probability of each event at time  $t$  ( $\widehat{p}_{kt}$ ,  $k = 1, 2, 3$ ) is forecasted based on the forecasted conditional density not the forecasted conditional mean of future spot interest rate, as only the former rather than the latter allows one to predict the probabilities of pecuniary prepayment and non-pecuniary prepayment simultaneously for each loan at each mortgage time  $t$ . Specifically, with the forecasted conditional density of future spot interest rate, in each period  $t$ , the future market mortgage interest rate (the future 10-year yield,  $y_t(r(t))$ ) can be any positive value in the spectrum, either equal to, or above, or below the rate at origination ( $y_0$ ), with the probability of each value determined by the forecasted conditional density. For each specific value of  $y_t(r(t))$ , the predicted probability of each event at time  $t$  given that specific value can be calculated as  $p_{kt}[y_t(r(t))]$ . Then considering all possible values of  $y_t(r(t))$ , the forecasted conditional density of the spot interest rate  $dF(r(t)|r(0))$  can therefore be used to calculate the predicted probability of each event at time  $t$  as in Equation (3)–(5). By contrast, with the use of the forecasted conditional mean of the spot interest rate, the expected future market mortgage interest rate (future 10-year yield) at time  $t$  would be just a single mean value ( $\bar{y}_t(r(t))$ ). This single mean value is either higher or lower than, or possibly equal to the rate at origination ( $y_0$ ). If this single mean value at time  $t$  ( $\bar{y}_t(r(t))$ ) is higher than or equal to the rate at origination, only non-pecuniary prepayment probability, not the pecuniary prepayment probability, at time  $t$  can be predicted. Similarly, if this single mean value at time  $t$  is lower than the rate at origination, only pecuniary prepayment probability in period  $t$  can be evaluated. In this study, given the need to predict the probabilities of both pecuniary and non-pecuniary prepayment at time  $t$ , the forecasted conditional density rather than the mean is adopted. Here, the integrated expectations are numerically approximated through a discretization approach in which the spot interest rate domain was divided into numerous but finite intervals.

<sup>20</sup> Here the 5-year window is chosen mainly because almost all of the non-pecuniary prepayments in the sample occurred within the first five years since loan origination, although a few defaults and pecuniary prepayments were observed in mortgage years 5–10. A capital  $P$  is used to distinguish cumulative loan termination probabilities from time-specific ones. The subscript  $k$  tells the type of the event, 1 for default, 2 for pecuniary prepayment, and 3 for non-pecuniary prepayment.

$$\widehat{P}_k = \sum_{t=1}^T \left(1 + \frac{y_0}{12}\right)^{-t} \widehat{p}_{kt} \prod_{s=1}^{t-1} \left(1 - \sum_{k=1}^3 \widehat{p}_{ks}\right) \quad k = 1, 2, 3, \quad (6)$$

where  $\widehat{P}_{kt}$  is the predicted probability of event  $k$  in period  $t$  given that the loan has been active by the beginning of period  $t$  with the corresponding continuation probability as  $\prod_{s=1}^{t-1} \left(1 - \sum_{k=1}^3 \widehat{p}_{ks}\right)$ . The time-specific unconditional predicted probability of event  $k$  in period  $t$ ,  $\widehat{p}_{kt} \prod_{s=1}^{t-1} \left(1 - \sum_{k=1}^3 \widehat{p}_{ks}\right)$ , is discounted by the 10-year yield at origination ( $y_0$ ) with the assumption that loan terminations at earlier stages of a loan would matter more to lenders when assessing loan termination risk. Thus, those time-specific discounted unconditional predicted probabilities over a 5-year window are summed up to arrive at a cumulative predicted probability of each event ( $\widehat{P}_k$ ).

In order to investigate whether lenders have incorporated each risk into mortgage pricing, the forecasted 5-year cumulative probability of each event ( $\widehat{P}_k$ ) is included in a loan contract rate determination model as in Equation (7):

$$C_0 = \alpha_0 y_0 + \beta_1 \widehat{P}_1 + \beta_2 \widehat{P}_2 + \beta_3 \widehat{P}_3 + \gamma' \mathbf{z} + \varepsilon \quad (7)$$

where  $C_0$  is the contract rate of a loan,  $y_0$  is the 10-year treasury yield (the benchmark mortgage interest rate) at the time of loan origination, and  $\mathbf{z}$  represents a full set of covariates at loan origination depicting the traits of the loan, the borrower, the collateral, and the neighborhood.<sup>21</sup>

## Data and Specifications

### Data

A sample of 30-year first-lien fixed-rate residential mortgage loans for home purchases is employed for empirical analysis. All of the mortgage loans in the sample were serviced by GMAC Residential Capital Company, LLC (GMAC ResCap) which was a finance company that specialized in servicing nonconforming conventional mortgage loans and issuing non-agency mortgage-backed securities (MBSs) before the recent financial crisis. Those loans were originated by different underwriters, sold in the secondary mortgage market, and packaged into MBSs. For each loan in the records, in addition to detailed information on the loan, the borrower, and underlying property at the time of loan origination, GMAC ResCap also reported its delinquency and prepayment status on a monthly basis, allowing the performance of loans to be tracked over time. Similar to prior studies, default is defined as the occurrence of a borrower being 90-days delinquent which eventually leads to a foreclosure or foreclosure alternative (e.g., deed in lieu of foreclosure or a short sale, etc.).

<sup>21</sup> Those covariates are incorporated mainly for the identification purpose in this system of equations.

In addition to financial incentives/constraints factors that are believed to affect the probability of a borrower prepaying a mortgage loan based on theories and prior studies (e.g., interest rate change, a borrower's FICO score and income, and so forth), some prior studies also found that a borrower's personal characteristics, especially a borrower's race and ethnicity group, have an impact on loan prepayment rate (Kelly 1995; Clapp et al. 2001; Deng and Gabriel 2006; Firestone et al. 2007; Kau et al. 2018).<sup>22</sup> To incorporate these factors, this study matches the GMAC ResCap loan data with the Home Mortgage Disclosure Act Loan/Application Register (HMDA-LAR) data to identify the race and ethnicity of the borrower of each loan in the sample.<sup>23</sup> One challenge to match these two data sets is the location of the property securing each loan is defined differently in these two data sets. Specifically, GMAC ResCap loan data records the location of each loan by its zip code, while loan applications in the HMDA-LAR data are reported based on the census tract. In addition, the definition of a census tract in HMDA-LAR data varies across time.<sup>24</sup> To solve this issue, the GMAC ResCap loan data were first matched with a data set of property transactions from Miami-Dade County, FL, which is readily available. This first-step match identifies the geographic location of the property securing each loan, and thus its 1980, 1990, and 2000 census tract identifiers. Specifically, the GMAC ResCap loan data were first restricted to 30-year fixed-rate home purchase mortgage loans from Miami-Dade, and the matching was conducted based on a series of property transaction characteristics, including property sale price (value), sale month (loan origination month), property type, and zip code.<sup>25</sup> The resulting loan-property-sale matched sample, with each loan's location by its census tract identified, was then matched to HMDA-LAR data on the basis of the loan amount (in \$1000), loan origination year, property type, property occupancy status, loan purpose, lien status (if applicable), and census tract.<sup>26</sup> Only loans with a unique HMDA-LAR match were included in the final sample for empirical analysis.<sup>27</sup>

To obtain the neighborhood characteristics of a loan, the matched loan-property-HMDA data were further augmented with the census survey data based on the 1990 census tract

<sup>22</sup> Those prior studies all find that minority borrowers are less likely to prepay than equivalent non-Hispanic white borrowers.

<sup>23</sup> The Home Mortgage Disclosure Act (HMDA) is a federal law that requires financial institutions to collect and disclose data on race, gender, and income of the applicant and the co-applicant (if applicable) for each loan application (the LAR data).

<sup>24</sup> HMDA-LAR data from 1989 to 1991 used the 1980 census tract boundaries, data from 1992 to 2002 followed the 1990 census tract definitions, and data from 2003 to 2012 adopted 2000 census tract definitions.

<sup>25</sup> Each mortgage loan was matched with property sales in the whole pool with replacements, selecting the matched pair that has the minimum difference between the appraised value of the property from the GMAC ResCap loan data and the property sale price from the property transaction data. Besides, it is also required that this minimum difference should not be larger than 5% of the average value of the appraised value and the sale price. Only a loan with a unique property sale match is included. If multiple property sales were matched to a loan with the same (minimum) difference and those properties are located at the same census tract, it is assumed that that loan has a single property sale match and is included in the final sample for empirical analysis.

<sup>26</sup> Each loan-property-sale match was only matched with approved conventional loan applications in the HMDA-LAR data. The gap in loan amount up to \$1000 was allowed, as the loan amount reported in HMDA-LAR data is in thousand dollars.

<sup>27</sup> If multiple loan applications from HMDA-LAR data were matched to one mortgage loan (meeting all of those matching criteria), and if those multiple loan applications happen to have exactly the same race and ethnicity information on the borrower and co-borrower (if any), it is assumed that that loan has a unique HMDA-LAR match and is included in the final sample of this study.

boundaries of a loan's underlying property. The 1990, 2000, and 2010 census survey data, normalized to the 1990 census tract boundaries, were employed to measure the traits of a loan's neighborhood including housing occupancy rate, poverty rate, and so forth. In order to generate time-varying neighborhood trait variables and also to track the change of a neighborhood over the life of a loan, the decennial census survey data from 1990 to 2000 and from 2000 to 2010 were linearly time-trended on a monthly basis.<sup>28</sup> Additionally, the HMDA data are aggregated annually to the census tract level to proxy a borrower's income.<sup>29</sup> Specifically, for each census tract and each year, all loan applications (approved, denied, or withdrawn) for 1-to-4 family dwelling purchases from all lenders were pooled together to calculate the median applicant income based on a three-year window (the prior year, the current year, and the next year). To generate monthly HMDA median applicant income measures, the resulting yearly HMDA measures were time-trended.

The whole pool of property transactions from Miami-Dade County, FL were utilized to generate measures that describe the characteristics and the evolution of each local housing market.<sup>30</sup> An inflation-adjusted median house price index was created for each 1990 census tract and each month of analysis. This monthly tract-level house price index was employed to track the value change of the underlying property over a loan's life.<sup>31</sup> Specifically, for each 1990 census tract and each month, all arm's length transactions of residential properties within a three-year window of that month (18 months before and 18 months after that month), were pooled together to find the median house price. Similarly, in order to measure the time-varying housing heterogeneity in a neighborhood, for each census tract and each month, the standard deviation of house sale prices based on the whole pool of transactions in that tract was calculated using a three-year window around that month.

After matching the loan-property-HMDA data with neighborhood data, there are 1404 observations of 30-year first-lien fixed-rate fully-amortizing nonconforming mortgage loans for home purchases with completed loan performance records and no missing values on the characteristics of the loan, the borrower, the underlying property, and the neighborhood. The sample is restricted to loans originated from Jan. 1997 to

<sup>28</sup> The national Gross Domestic Product (GDP) per capita is used as an inflation index to make adjustments to all monetary variables. All income/price is defined in 2009 dollars. Those inflation-adjusted values (e.g., in the year of 1990, 2000, and 2010) are then time-trended to generate monthly measures.

<sup>29</sup> As a borrower's income is not available in the GMAC ResCap loan data, the median loan applicant income in the borrower's neighborhood is used to proxy the borrower's income. Here the median loan applicant income from the aggregated HMDA data is preferred over the average household income from the decennial census survey data, because: 1) the HMDA data is updated every year while census survey data is updated every 10 years; 2) both renters and homeowners are included in the census survey, while only homebuyers (loan applicants) are included in HMDA reporting. Therefore, the aggregated HMDA data is believed to more accurately measure a borrower's income. The applicant income each year is inflation adjusted by a GDP per capita deflator. All income is defined in 2009 dollars.

<sup>30</sup> The whole pool includes residential property sales in Miami-Dade County, FL, from 1990 to 2013. Each house sale price is inflation adjusted by a GDP per capita deflator first and then pooled together to generate the median house price index. All prices are defined in 2009 dollars.

<sup>31</sup> In this study, the generated monthly census-tract-level median house price index is preferred over those commonly used MSA-level house price index or the yearly FHFA zip-code-level house price index, as the former is updated more frequently (monthly) and varies across neighborhoods of a smaller scale (census tract is the smallest unit among MSA, zip code, and census tract in Miami-Dade County). Therefore, this generated monthly census-tract-level median house price index enables one to more closely track the value change of each individual house.

Dec. 2006.<sup>32</sup> Their loan performance is observed from Jan. 2000 to Oct. 2010, a period covering the recent financial crisis.

## Specifications

As default and different forms of prepayment may be induced by different financial incentives and encumbered by various financial constraints, the multinomial logit model as represented by Equation (1) is estimated through three separate logit models, one for each event, with different baseline hazard rates ( $\alpha_{jt}$ ) and different specifications.<sup>33</sup> In the default hazard model, a scaled Standard Default Assumption schedule (SDA) is used as the baseline hazard rates  $\alpha_{jt}$ , while in prepayments hazard models, mortgage year (age) fixed effects are adopted for  $\alpha_{jt}$  which allow for more flexibility in baseline hazard rates.

Indicated by the option-based model of “financial” mortgage terminations (Kau et al. 1992), the two most prominent factors determining whether the default or prepayment option is “in-the-money,” and thus a borrower’s tendency to terminate a loan, are the change in the value of the collateral and market interest rate change. As discussed above, the 10-year treasury constant maturity yield ( $y$ ) is used as the benchmark market interest rate for 30-year fixed-rate residential mortgage loans. Therefore, the gap between the 10-year yield at the time of loan origination and the 10-year yield at time  $t$  lagged by 2 periods ( $y_0 - y_{t-2}$ ) is utilized to measure the market interest rate change at time  $t$ .<sup>34</sup> The value change of the collateral is measured by the change of the generated monthly census-tract-level median house price index. Specifically, the ratio of the median house price index at time  $t$  to the index at the time of loan origination (time 0) is calculated and employed to measure how the value of the collateral at time  $t$  changes since loan origination. This ratio is named as the relative house price at time  $t$  ( $RHP_t$ ). In order to account for any correlation between interest rate change and house price change at time  $t$ , an interaction term of market interest rate change ( $y_0 - y_{t-2}$ ) and relative house price ( $RHP_t$ ) is also included. All of the three market characteristics variables (interest rate change, relative house price, and the interaction term) are included in the default and prepayment hazard models.

A set of covariates depicting the characteristics of the loan, the borrower, the underlying property, and the neighborhood are also included in this competing-risks hazard model. Variables describing the characteristics of the loan include the contract rate spread at origination ( $C_0 - y_0$ ),<sup>35</sup> the original loan-to-value (LTV) ratio, the original loan amount,

<sup>32</sup> The sample cannot be extended, as there were only a few loans originated before 1997 and also after 2006 due to the impact of the occurrence of the recent financial crisis.

<sup>33</sup> If the specifications are the same for the three hazards (default, pecuniary prepayment, and non-pecuniary prepayment), the three separate logit models would yield the same estimation results as a single multinomial logit model. One advantage of estimating the multinomial logit model through three separate logit models is the flexibility in specifying different models for different events. Here, to investigate whether the covariates have different impacts on pecuniary vs. non-pecuniary prepayments, this study adopts the same specifications for both of them, but different specifications for default.

<sup>34</sup> The 10-year yield at time  $t$  lagged by 2 periods instead of the 10-year yield at time  $t$  is used because in practice when a borrower decides to refinance his mortgage loan while observing an interest rate drop, it usually takes time to apply for and obtain a new mortgage loan. For the first two mortgage months, the 10-year yield at time  $t$  lagged by 2 periods is simply the yield at origination (time 0).

<sup>35</sup> The contract rate spread is defined as the difference between the fixed contract rate ( $C_0$ ) and the 10-year treasury constant maturity yield at origination ( $y_0$ ). This contract rate spread measures both the risk premiums of a loan as well as the points paid at loan closing which are used to buy down the contract rate.

a borrower's credit score (FICO), a dichotomous variable indicating whether a borrower provided full income documentation at loan closing, another dichotomous variable indicating whether a loan is encumbered by a prepayment penalty at time  $t$ , and loan origination season fixed effects. Property characteristics variables include property type fixed effects and property owner occupancy status fixed effects. A set of time-varying neighborhood traits variables are also incorporated in the hazard model including tract-level house price heterogeneity, housing occupancy rate, poverty rate, and relative median applicant income at time  $t$ .<sup>36</sup> In addition, a borrower's race and ethnicity are also included given the findings by prior studies that borrowers of different racial or/and ethnical groups may have different loan default/prepayment patterns.<sup>37</sup>

Following prior studies, considering possible nonlinear relationships between a covariate and loan termination hazard, a few covariates take a nonlinear form. The original LTV ratio (in percentage) is transformed into categorical variables with the following cutoffs: 80%, 90%, and 100%. A quadratic function form of the original loan amount is employed. With the assumption that once a borrower's FICO score is above a threshold (e.g., 700), an increase in FICO score might have a(n) less/insignificant marginal impact on loan termination rate, a continuous linear spline function form of the FICO score with a knot point at 700 is adopted.<sup>38</sup> In addition, please note the dichotomous variable indicating whether a loan is encumbered by a prepayment penalty at time  $t$  is excluded in the default hazard model, as this variable is not believed to directly affect loan default probability.<sup>39</sup>

Turning our attention to the contract rate determination model represented by Equation (7), in addition to the three predicted cumulative loan termination probabilities, the current study has included the 10-year treasury yield at origination ( $y_0$ ) as well as a full set of covariates indicated by theories to affect loan pricing. They are the original LTV ratio, a borrower's FICO score, the original loan amount,<sup>40</sup> a borrower's income documentation status fixed effects, prepayment penalty fixed effects,<sup>41</sup> loan origination season fixed effects, property type fixed effects, property owner occupancy status fixed effects, a borrower's race and ethnicity group fixed effects, and several variables depicting neighborhood traits at the time of loan

<sup>36</sup> The variable of the tract-level relative median applicant income at time  $t$  is defined as the ratio of the median applicant income in a census tract at time  $t$  to the median applicant income in Miami-Dade County, FL at time  $t$ .

<sup>37</sup> Several prior studies have found that minority borrowers tend to be less likely to prepay their mortgage loans (Kelly 1995; Clapp et al. 2001; Deng and Gabriel 2006; Firestone et al. 2007; Kau et al. 2018).

<sup>38</sup> Specifically, the FICO linear spline function was specified as follows:  $FICO_{(FICO \leq 700)} = \text{minimum}(FICO, 700)$ ; and  $FICO_{(FICO > 700)} = \text{maximum}(FICO, 700) - 700$ . Therefore, coefficient on  $FICO_{(FICO \leq 700)}$  measures the effects of FICO score on dependent variable when  $FICO \leq 700$ ; while coefficient on  $FICO_{(FICO > 700)}$  measures the marginal effects of FICO score when  $FICO > 700$ . The results are robust with various FICO score knot points, e.g., 720, 750, and so forth.

<sup>39</sup> This study has tested whether this prepayment penalty variable impacts default rate, and the results show that the impact is not significant.

<sup>40</sup> Similar to the specifications in the loan hazard model, the original LTV ratio variable is in its categorical form, the FICO score variable is in its continuous linear spline function form with a FICO score of 700 as the knot point, and the original loan amount is in its quadratic form.

<sup>41</sup> Loans in the sample are classified into three categories based on whether and how long a loan is encumbered by a prepayment penalty: 1) no prepayment penalty; 2) prepayment penalty for the first 1–3 years; 3) prepayment penalty for the first 5 years.

origination, including tract-level past housing price appreciation rate,<sup>42</sup> heterogeneity in housing price, housing occupancy rate, poverty rate, and relative median applicant income. Additionally, a calendar time trend variable is also included.<sup>43</sup>

## Endogeneity Issue

The issue of endogeneity arises in both the discrete-time competing-risks loan hazard model (Equation (1)) and the contract rate determination model (Equation (7)). In the competing-risks loan hazard model, one covariate – the contract rate spread ( $C_0 - y_0$ ) might be endogenous, as this variable might be correlated with the error term in Equation (1). The reason for this concern is in empirical mortgage loan performance analysis, it is normally very challenging to measure and thus incorporate every loan termination risk affecting factor. If a loan termination risk affecting factor is not observed in the data, but is observed by lenders/underwriters, it would be captured in the error term in Equation (1), while lenders/underwriters would take it into consideration when setting up the contract rate. In this scenario, the contract rate spread would be correlated with the error term. Given the non-linear nature of this competing-risks loan hazard model, following Petrin and Train (2003), a standard control function (CF) method is employed to address this potential endogeneity issue. Specifically, a linear reduced-form contract rate model is estimated first in which the contract rate of a loan is regressed against all exogenous variables in the system. The residual from this reduced-form contract rate model is obtained and then included in the competing-risks loan hazard model together with other covariates.

The final contract rate determination model (Equation (7)) also suffers from endogeneity issue. Recall that the contract rate spread is included in the competing-risks loan hazard model (Equation (1)), and the three predicted cumulative loan termination probabilities ( $\hat{P}_k$ ) are generated based on the estimation results of this competing-risks loan hazard model. Hence, each of the three predicted cumulative termination probabilities is a function of the contract rate (spread), while the dependent variable in Equation (7) is the contract rate of a loan. Given that this contract rate determination model is a linear one, a set of generated IVs ( $\tilde{P}_k$ ) are calculated and included in the estimation of the final contract rate model in place of the three original predicted probabilities (generated variables,  $\hat{P}_k$ ). To generate each IV ( $\tilde{P}_k$ ), the same procedures to calculate its corresponding generated variable ( $\hat{P}_k$ ) is utilized, but replacing the actual contract rate spread ( $C_0 - y_0$ ) with a predicted rate spread ( $\hat{C}_0 - y_0$ ). This predicted contract rate spread is simply the difference between a predicted contract rate and the 10-year yield at the time of loan origination. This predicted contract rate is obtained from the estimation of the linear reduced-form contract rate equation used in the CF method. As this predicted contract rate ( $\hat{C}_0$ ) is a linear function of all exogenous variables in the system, the three generated IVs ( $\tilde{P}_k$ ) would also be exogenous, with each as a valid IV for its corresponding generated variable ( $\hat{C}_k$ ). Using those generated IVs, 2SLS is employed

<sup>42</sup> The tract-level past housing price appreciation rate is defined as the ratio of the median housing sale price in a census tract in a three-year period prior to the month of loan origination to the median housing sale price in the same census tract in another three-year period prior to the three-year pre-origination period, then minus 1.

<sup>43</sup> This time trend variable equals 1 if a loan was originated in year 1997, and equals 2 for a loan originated in year 1998, and so forth.

to estimate the final contract rate model represented by Equation (7). Additionally, given that the three predicted probabilities ( $\hat{P}_k$ ) are generated variables, in order to correct the standard errors of the covariates estimates in this contract rate model, the method outlined in Appendix 6A of Wooldridge (2010) is employed.

## Results

### Statistical Descriptions

A brief description of the traits of the loans in the pooled sample at the time of loan origination is presented in Table 1, including the characteristics of the loan, the borrower, the underlying property, and the neighborhood.<sup>44</sup> Roughly 9% of loans in the sample were defaulted upon, 56% of the loans were prepaid when market interest rates dropped (pecuniarily prepaid), and 24% of the loans were prepaid when the rates had risen (non-pecuniarily prepaid). As all of loans in this sample are nonconforming loans, their high-risk nature is reflected by the statistics of the following variables. The average contract rate is high (around 8% annually), approximately 300 basis points higher than the prevailing 10-year treasury constant maturity yield. The original LTV ratio is 85% on average, and the mean FICO score of the borrowers in the sample is about 700. Less than half of the borrowers (48%) were able to provide full income documentation while applying for a loan, and a great portion of the borrowers (33%) were encumbered by a prepayment penalty for the first few years. For most of the variables describing the neighborhood characteristics, both the standard deviation and the range are large relative to the mean value, indicating a substantial level of spatial variation in the performance of the local housing markets and in the socioeconomic status of the local residents/homeowners. In this pooled sample, approximately 7% of loans were extended to African American borrowers, and 22% of them were originated to non-Hispanic white borrowers. The majority of the borrowers (71%) were Hispanics.<sup>45</sup>

<sup>44</sup> In this study a restricted sample is also generated only including loans within those census tracts where both the occurrence of pecuniary prepayments and that of non-pecuniary prepayments were observed. The 1990 census tract definitions are adopted here. There are 1119 loans in this restricted sample. All of the results based on this restricted sample are very close to those based on the pooled sample (full sample). Results on this restricted sample are omitted here but are available upon request.

<sup>45</sup> One issue arises with the use of the race and ethnicity information from the HMDA-LAR data, as the reporting rules on the classification of race and ethnicity changed significantly in 2004. Before 2004, "Hispanic" was a category of race, while since 2004 "Hispanic" is no longer a race category but an ethnicity category. In order to align the pre- and post-2004 race and ethnicity data, a hierarchy is used among all of the race and ethnicity classifications considering the demographic characteristics of the population in Miami-Dade County, FL. If a post-2004 applicant self-reported as an African American/American Indian/Asian/Native Hawaiian, this applicant is treated as an African American/American Indian/Asian/Native Hawaiian, no matter whether (s)he self-identified as a Hispanic or not. If an applicant self-reported White for race and Hispanic for ethnicity, (s)he is treated as a Hispanic. Therefore, white borrowers in this sample are all non-Hispanic whites. In addition, if there is a co-applicant in the HMDA-LAR record, a similar hierarchy ordering is employed to determine the race and ethnicity variable for each loan in the sample. Specifically, if either of the applicant and co-applicant self-identified as an African American, that loan is considered to be given to an African American borrower. If one applicant is a Hispanic and the other one is a non-Hispanic white, that loan is treated as one taken by a Hispanic. If both of the two applicants self-reported as non-Hispanic whites, that mortgage loan is considered to be extended to a non-Hispanic white.



**Table 1** Descriptive Statistics of Mortgage Loans at Loan Origination

Variable Name	Mean	Std. Dev.	Min.	Max.
Pecuniary Prepay (0,1)	0.5605	0.4965	0.0000	1.0000
Non-pecuniary Prepay (0,1)	0.2436	0.4294	0.0000	1.0000
Default (0,1)	0.0869	0.2818	0.0000	1.0000
Loan Characteristics				
Contract rate at origination ( $C_0$ )	7.9622	1.1282	5.2500	12.5000
10-year treasury yield at origination ( $y_0$ )	4.9201	0.8163	3.3300	6.8900
Contract rate spread at origination ( $C_0 - y_0$ )	3.0422	0.9476	1.0100	6.4650
Original LTV	84.8048	11.8555	36.0000	107.0000
FICO at origination	700.0370	53.2957	483.0000	822.0000
Original loan amount (in \$10,000)	18.1243	12.1792	2.0000	80.0000
Full income documentation (0,1)	0.4765	0.4996	0.0000	1.0000
Without prepayment penalty (0,1)	0.6702	0.4703	0.0000	1.0000
Prepayment penalty for 1 to 3 years (0,1)	0.2236	0.4168	0.0000	1.0000
Prepayment penalty for 5 years (0,1)	0.1061	0.3081	0.0000	1.0000
Property Characteristics				
Property owner occupied (0,1)	0.7778	0.4159	0.0000	1.0000
Property condo (0,1)	0.3376	0.4731	0.0000	1.0000
Neighborhood-Level Characteristics at Origination <sup>a</sup>				
Recent housing price appreciation rate <sup>b</sup>	0.1949	0.2227	-0.4519	2.8150
Heterogeneity in housing price (in \$10,000) <sup>c</sup>	11.2656	7.8752	1.1012	43.3243
Housing occupancy rate (from Census Survey) <sup>d</sup>	0.9094	0.0858	0.5373	0.9883
Poverty rate (from Census Survey) <sup>d</sup>	0.1441	0.0826	0.0216	0.5988
Relative median applicant income (from HMDA) <sup>e</sup>	1.2188	0.5786	0.5483	4.8597
Borrower Race				
Non-Hispanic White	0.2201	0.4145	0.0000	1.0000
Hispanic	0.7087	0.4545	0.0000	1.0000
African American	0.0712	0.2573	0.0000	1.0000
Sample Size	1404			

<sup>a</sup> The neighborhood of a loan's underlying property is defined based on the property's 1990 census tract boundaries

<sup>b</sup> Recent housing price appreciation rate is defined as the ratio of the median housing sale price in a census tract in a three-year period prior to the month of loan origination to the median housing sale price in the same census tract in another three-year period prior to the three-year pre-origination period, then minus 1

<sup>c</sup> Heterogeneity in housing price is defined as the standard deviation of the housing sale price in a census tract over a three-year window (in \$10,000)

<sup>d</sup> Housing occupancy rate and poverty rate were generated from the decennial census survey data in 1990, 2000, and 2010

<sup>e</sup> Relative median applicant income was generated from the HMDA data aggregated on the census tract level on a yearly basis. It is defined as the ratio of the median applicant income in a census tract at origination to the median applicant income in Miami-Dade County, FL at origination

Table 2 provides descriptive statistics of the pooled sample by the eventual termination event by Oct., 2010 (censored, defaulted, pecuniarily prepaid, non-pecuniarily prepaid). Among the four event groups, the average contract rate at origination ( $C_0$ ) is the highest for loans that were pecuniarily prepaid. This finding is somewhat puzzling, however comparison on the average benchmark mortgage interest rate (10-year yield at origination,  $y_0$ ) across these four event groups reveals that it is because those loans that were eventually pecuniarily prepaid were originated in a period with relatively higher interest rates. The descriptive statistics seem to imply that loans that were pecuniarily prepaid differ from those that were non-pecuniarily prepaid, indicating the importance of making a distinction between these two types of prepayment. Loans in these two types of prepayment groups appear to be different from each other indicated by the following variables, including the average original LTV ratio (83% vs. 89%), the average FICO score at origination (700 vs. 691), the proportion of the borrowers with full income documentation (54% vs. 44%), the proportion of the borrowers without prepayment penalties (75% vs. 52%), the mean tract-level poverty rate (13% vs. 16%), and the average tract-level relative median applicant income (1.32 vs. 1.05).

### What Impacts Pecuniary Vs. Non-pecuniary Prepayments?

The prepayment hazard model estimates are presented in Table 3. Model (1) is for a prepayment hazard model in which pecuniary and non-pecuniary prepayments are combined; while model (2) and (3) are for pecuniary and non-pecuniary prepayment hazards, respectively. Comparisons on the estimation results between this combined prepayment model and the separated ones underscore the importance of separating these two types of prepayment from each other. In the combined prepayment model, the estimated coefficient of the market interest rate change variable ( $y_0 - y_{t-2}$ ) is not significantly different from zero. However, in the separated ones, this coefficient is positive and significant in the pecuniary prepayment model, while it is negative but insignificant in the non-pecuniary prepayment model. This finding is consistent with the anticipation that observing a declining market interest rate, a borrower would have more financial incentives to refinance the mortgage loan to lower the overall borrowing costs. Although the value change of the collateral (relative house price at time  $t$ ,  $RHP_t$ ) is found to have a significant and positive impact in each of the three models, the estimated coefficient in the pecuniary/non-pecuniary prepayment hazard model is slightly smaller than that in the combined model. Overall, this finding indicates that regardless of the interest rate environment (rising or declining), as the value of the collateral increases, a borrower tends to be more likely to prepay a loan.

Some additional differences between the models appear when comparing the estimation results of covariates describing the traits of the loan, the property, and the neighborhood. Although the estimated coefficient of the contract rate spread variable is positive and significant in the combined and non-pecuniary models, the coefficient in the combined model is only approximately one-third of that in the non-pecuniary prepayment hazard model. The contract rate spread variable captures the impacts of points and hence a borrower's expectation on how long the property would be held on non-pecuniary prepayment probability. Borrowers that plan to move soon in the near future would rationally self-select less points and thus a higher contract rate at origination to lower the effective borrowing cost; while borrowers expecting to stay

**Table 2** Descriptive Statistics of Mortgage Loans at Loan Origination by Event

Variable Name	Event			
	Continued	Pecuniary Prepayment <sup>a</sup>	Non-pecuniary Prepayment <sup>a</sup>	Default
	Mean	Mean	Mean	Mean
<b>Loan Characteristics</b>				
Contract rate at origination ( $C_0$ )	6.9958	8.3335	7.6083	7.7716
10-year treasury yield at origination ( $y_0$ )	4.4237	5.3750	4.1999	4.6267
Contract rate spread at origination ( $C_0 - y_0$ )	2.5721	2.9585	3.4084	3.1448
Original LTV	85.2614	82.7370	89.0175	85.7623
FICO at origination	719.0327	700.5947	691.4327	696.7377
Original loan amount (in \$10,000)	19.6167	17.4869	17.0751	23.3055
Full income documentation (0,1)	0.3725	0.5375	0.4386	0.3197
Without prepayment penalty (0,1)	0.6471	0.7522	0.5205	0.5902
Prepayment penalty for 1 to 3 years (0,1)	0.2810	0.1601	0.3158	0.3033
Prepayment penalty for 5 years (0,1)	0.0719	0.0877	0.1637	0.1066
<b>Property Characteristics</b>				
Property owner occupied (0,1)	0.7255	0.7891	0.7807	0.7623
Property condo (0,1)	0.4837	0.3202	0.2953	0.3852
<b>Neighborhood-Level Characteristics at Origination <sup>b</sup></b>				
Recent housing price appreciation rate <sup>c</sup>	0.3162	0.1340	0.2087	0.3969
Heterogeneity in housing price (in \$10,000) <sup>d</sup>	11.2457	11.9827	9.4626	11.7188
Housing occupancy rate (from Census Survey) <sup>e</sup>	0.9013	0.9103	0.9191	0.8867
Poverty rate (from Census Survey) <sup>e</sup>	0.1476	0.1324	0.1628	0.1625
Relative median applicant income (from HMDA) <sup>f</sup>	1.1523	1.3198	1.0520	1.1186
<b>Borrower Race</b>				
Non-Hispanic White	0.1111	0.2681	0.1784	0.1639
Hispanic	0.8235	0.6696	0.7398	0.7295
African American	0.0654	0.0623	0.0819	0.1066
Number of Loans	153	787	342	122

<sup>a</sup> Pecuniary prepayment is defined as the occurrence of a borrower prepaying a loan when the market interest rate at prepayment drops below the rate at the time of loan origination; while non-pecuniary prepayment is defined as the one when the rate at prepayment is equal to or above the rate at origination

<sup>b</sup> The neighborhood of a loan's underlying property is defined based on the property's 1990 census tract boundaries

<sup>c</sup> Recent housing price appreciation rate is defined as the ratio of the median housing sale price in a census tract in a three-year period prior to the month of loan origination to the median housing sale price in the same census tract in another three-year period prior to the three-year pre-origination period, then minus 1

<sup>d</sup> Heterogeneity in housing price is defined as the standard deviation of the housing sale price in a census tract over a three-year window (in \$10,000)

<sup>e</sup> Housing occupancy rate and poverty rate were generated from the decennial census survey data in 1990, 2000, and 2010

<sup>f</sup> Median applicant income was generated from the HMDA data aggregated on the census tract level on a yearly basis. It is defined as the ratio of the median applicant income in a census tract at origination to the median applicant income in Miami-Dade County, FL at origination

**Table 3** Prepayment Hazard Model Estimates

Variables	Model (1)		Model (2)		Model (3)	
	Prepayment Combined		Pecuniary Prepayment <sup>a</sup>		Non-pecuniary Prepayment <sup>a</sup>	
	Coef.	<i>P</i> value	Coef.	<i>P</i> value	Coef.	<i>P</i> value
Intercept	-8.5281	<.0001	-8.8589	0.0004	-11.0318	0.0014
Market Characteristics						
Market interest rate change at time $t$ ( $y_{0-t}$ ) <sup>b</sup>	0.3040	0.1244	0.6540	0.0181	-0.4906	0.2610
Relative house price at time $t$ (RHP) <sup>b</sup>	1.3164	<.0001	0.9976	0.0003	1.2264	<.0001
$(y_{0-t}) \times$ (RHP) <sup>b</sup>	0.1076	0.4801	0.3702	0.0887	-0.4340	0.2085
Loan Characteristics						
Contract rate spread at origination ( $C_0-y_0$ )	0.3293	0.0407	0.1272	0.5265	0.9029	0.0012
Original LTV categories (Base group: Original LTV <=80)						
80 < Original LTV <=90	0.1411	0.1506	0.0430	0.7259	0.1860	0.2790
90 < Original LTV <=100	0.4368	0.0013	0.4117	0.0166	0.5229	0.0209
100 < Original LTV	0.4151	0.0587	0.2552	0.3820	0.5573	0.1180
Original loan amount (in \$10,000)	0.0205	0.0960	0.0035	0.8113	0.0749	0.0019
Square term of original loan amount (in \$10,000)	-0.0003	0.0971	-0.0002	0.5148	-0.0010	0.0232
FICO at origination continuous linear splines						
Minimum (FICO, 700)	0.0013	0.5384	0.0028	0.3064	0.0007	0.8476
Maximum (FICO, 700)-700	-0.0024	0.0828	-0.0020	0.2358	-0.0032	0.2000
Full income documentation (0,1)	-0.0981	0.2158	0.1091	0.2641	-0.3626	0.0115
Within prepayment penalty period at time $t$ (0,1) <sup>b</sup>	-0.5568	<.0001	-0.5589	0.0001	-0.7965	<.0001
Property Characteristics						
Property owner occupied (0,1)	-0.0071	0.9398	0.1901	0.1075	-0.3291	0.0434
Property condo (0,1)	-0.2276	0.0153	-0.2468	0.0363	0.0112	0.9435
Neighborhood-Level Characteristics						
Housing price heterogeneity at time $t$ (in \$10,000) <sup>b</sup>	0.0115	0.1380	-0.0056	0.5707	0.0423	0.0007
Housing occupancy rate at time $t$ <sup>b</sup>	0.5498	0.2841	0.1660	0.7856	0.8782	0.3786
Poverty rate at time $t$ <sup>b</sup>	-0.1074	0.8362	-1.0661	0.1099	0.9734	0.2670
Relative median applicant income at time $t$ (HMDA) <sup>b</sup>	0.0528	0.5882	0.2884	0.0098	-0.7194	0.0013
Borrower Race (Base group: Non-Hispanic white)						
African Americans	-0.2669	0.0825	-0.3245	0.0981	-0.2796	0.2755
Hispanics	-0.2079	0.0131	-0.2245	0.0242	-0.2019	0.2012
Control Function Variable						
Residual <sup>c</sup>	0.1491	0.3814	0.3358	0.1152	-0.4563	0.1245
Baseline Hazard Rate						
Mortgage year fixed effects <sup>d</sup>	YES		YES		YES	
Origination seasons fixed effects <sup>d</sup>	YES		YES		YES	
Likelihood Ratio (Chi-Square)	581.1177		814.2394		504.3571	

<sup>a</sup> Pecuniary prepayment is defined as the occurrence of a borrower prepaying a loan when the market interest rate at prepayment drops below the one at the time of loan origination; while non-pecuniary prepayment is defined as the one when the rate at prepayment is equal to or above the rate at origination

<sup>b</sup> Denotes time-varying variables

<sup>c</sup> The residual comes from the contract-rate reduced-form estimation (the control function method)

<sup>d</sup> Mortgage year fixed effects estimates and the loan origination season fixed effects estimates are omitted here, but are available upon request

longer would choose the opposite combination. Thus, a positive association between contact rate spread and non-pecuniary prepayment probability would be expected, with which the empirical results here are consistent.

The original loan amount is normally used to capture the scale effects of loan size (Clapp et al. 2001). Given the assumption that loan prepayment transaction costs tend to be fixed, a larger mortgage loan would provide a larger dollar incentive to prepay, keeping other variables the same. The results indicate that the original loan size appears to have a positive impact on non-pecuniary prepayment probability but at a decreasing rate, while its impact on pecuniary prepayment probability, and in the combined model, is not significant.

A borrower providing full income documentation upon loan application tends to be less likely to prepay when interest rates rise than an equivalent borrower with limited/no income documentation. The income documentation status variable may capture the impact of job stability on a person's residential mobility and hence non-pecuniary prepayment rate. The results also indicate that prepayment penalties deter prepayment, a result that holds across all three models.

Results on the two property traits variables indicate that a borrower purchasing a house as the primary residence appears to be less likely to prepay in an environment of rising interest rates than a similar borrower buying the house as an investment. The results also indicate that a borrower living in a condo tends to be less likely to pecuniarily prepay than a borrower living at a single-family house. One possible explanation for this finding is the financial benefits from pecuniary prepayments (lower monthly payments) might be less for a loan used to purchase a condo given its smaller original loan amount and a shorter expected holding period on average.<sup>46</sup> Please note here the property owner occupancy status variable does not have a significant impact on pecuniary prepayment probability, and the property type variable does not significantly affect non-pecuniary prepayment rate.

The estimation results on neighborhood characteristics variables indicate that pecuniary and non-pecuniary prepayments tend to occur in neighborhoods with different traits. These traits are shown to lack significance in the combined model. The probability of non-pecuniary prepayment tends to be higher in a census tract with a higher level of house price heterogeneity and a lower level of median applicant income. By contrast, borrowers living in a more affluent neighborhood appear to be more likely to prepay when interest rates drop, *ceteris paribus*. In addition, results of a borrower's race and ethnicity group indicate that minority borrowers are less likely to take advantage of a lower interest rate through prepayment than equivalent non-Hispanic white borrowers. However, there is no evidence of significant differences in non-pecuniary prepayment rates across racial and ethnic groups.<sup>47</sup>

<sup>46</sup> In the sample of this study, the average original loan amount for loans used to purchase a single-family house is approximately \$210,000, while that for loans used to buy a condo is roughly \$130,000. The dollar amount of benefits from refinance to reduce monthly payments tends to increase as the loan balance increases. Besides, prior studies (e.g., Follain et al. 1992; Archer et al. 1997) find that borrowers with relatively shorter expected holding periods of the property may not prepay their mortgage loans even when the prepayment option is "in-the-money" (the interest rate drops), since the paycheck period would be too short to realize savings from refinancing and to recoup refinance costs.

<sup>47</sup> Some prior studies (Kelly 1995; Clapp et al. 2001; Deng and Gabriel 2006; Firestone et al. 2007; Kau et al. 2018) find that minority borrowers tend to be less likely to prepay. By separating pecuniary prepayments from non-pecuniary prepayments, the results show minority borrowers are more likely not to prepay when it is financially beneficial.

The comparisons indicate that separating pecuniary from non-pecuniary prepayments allows one to identify variations in the impacts of the covariates on each termination decision. The separation by prepayment types also helps to avoid the pitfalls of the combined prepayment model which fails to capture the significance of some coefficients.

Given the substitute nature of default and prepayment, the default hazard model is also estimated within the framework of a competing-risks loan hazard model. Table 4 reports the default hazard estimation results. Overall, the results are very consistent with both theories and prior studies. Default is mainly induced by a drop in the value of the collateral, as the variable of relative house price at time  $t$  ( $RHP_t$ ) is shown to have a significantly negative impact on default probability. The original loan amount variable impacts the default rate in a positive way, but at a decreasing rate. Borrowers with a relatively higher FICO score, with full income documentation upon loan application, with the underlying property as the primary residence, or borrowers living at a neighborhood of higher income tend to be less likely to default. However, the marginal impact of a higher FICO score on default rate becomes insignificant when the FICO score is above 700. In addition, a borrower's race and ethnicity is not shown to significantly impact default probability.

### Loan Prepayment Probabilities Simulation Results

To further illustrate how identified factors impact the overall pecuniary vs. non-pecuniary prepayment probabilities, this study simulated 5-year cumulative probability of each event (default, pecuniary prepayment, and non-pecuniary prepayment) by the factors that are shown to impact pecuniary and non-pecuniary prepayment rates differently.<sup>48</sup> The simulations are conducted in two interest rate change environments: declining interest rates and rising interest rates.<sup>49</sup> Specifically, the 5-year predicted cumulative probability of each event for a sample base loan is calculated first. Here, a base loan is assumed to be taken by a non-Hispanic white borrower with limited income documentation to purchase a single-family house for investment purposes. The values of the time-invariant covariates are set at the full sample median, while the values of the time-varying covariates are set at the sample median of each period. Each factor is then varied to see how the 5-year predicted cumulative probabilities vary.<sup>50</sup>

<sup>48</sup> Only covariates that are shown to have different impacts on pecuniary vs. non-pecuniary prepayments (significant vs. insignificant, positive vs. negative) are discussed here.

<sup>49</sup> As discussed above, the CIR term structure model is employed to predict the possible paths (density) of future interest rates seen from loan origination. Based on the prediction results, a mortgage loan originated in January 2000 would be expected to have a greater chance of experiencing future interest rate drops subsequent to loan origination, while a loan originated in January 2004 would be expected to have a greater chance of facing interest rate rises since loan origination. Therefore, in the declining interest rates environment, a loan is assumed to be originated in January 2000; whereas in the rising interest rate environment, a loan is assumed to be originated in January 2004. Please notice here, a declining interest rate environment does not mean that the future interest rate will definitely (100%) drop. In that environment, the interest rate might increase in the future, however, the possibility is relatively low. That is the same with a rising interest rate environment.

<sup>50</sup> The varying covariates of interest include the contract rate spread at origination, original loan amount, borrower income documentation status (0,1), underlying property owner occupancy status (0,1), property type (0,1), tract-level house price heterogeneity, median applicant income, and a borrower's race and ethnicity group.

**Table 4** Default Hazard Model Estimates

Variables	Coef.	<i>P</i> value
Intercept	5.5525	0.2641
Market Characteristics		
Market interest rate change at time $t$ ( $y_0 - y_{t-2}$ ) <sup>a</sup>	-0.4646	0.4195
Relative house price at time $t$ (RHP) <sub><math>t</math></sub> <sup>a</sup>	-3.3867	<.0001
( $y_0 - y_{t-2}$ ) $\times$ (RHP) <sub><math>t</math></sub> <sup>a</sup>	0.8426	0.1214
Loan Characteristics		
Contract rate spread at origination ( $C_0 - y_0$ )	0.1838	0.6087
Original LTV categories (Base group: Original LTV $\leq 80$ )		
80 < Original LTV $\leq 90$	-0.0842	0.7832
90 < Original LTV $\leq 100$	0.3797	0.3497
100 < Original LTV	1.1537	0.0664
Original loan amount (in \$10,000)	0.1241	0.0003
Square term of original loan amount (in \$10,000)	-0.0010	0.0282
FICO at origination continuous linear splines		
Minimum (FICO, 700)	-0.0112	0.0340
Maximum (FICO, 700)-700	-0.0013	0.7467
Full income documentation (0,1)	-0.6474	0.0103
Property Characteristics		
Property owner occupied (0,1)	-0.5237	0.0479
Property condo (0,1)	0.2550	0.3509
Neighborhood-Level Characteristics		
Housing price heterogeneity at time $t$ (in \$10,000) <sup>a</sup>	0.0454	0.0503
Housing occupancy rate at time $t$ <sup>a</sup>	-2.0557	0.1701
Poverty rate at time $t$ <sup>a</sup>	2.2451	0.1244
Relative median applicant income at time $t$ (HMDA) <sup>a</sup>	-1.7626	0.0001
Borrower Race (Base group: Non-Hispanic white)		
African Americans	0.3622	0.3650
Hispanics	-0.0467	0.8627
Control Function Variable		
Residual <sup>b</sup>	0.5072	0.1784
Baseline Hazard rate		
SDA	0.8096	0.0296
Origination seasons fixed effects <sup>c</sup>	YES	
Likelihood Ratio	222.6313	

<sup>a</sup> Demotes time-varying variables

<sup>b</sup> The residual comes from the contract-rate reduced-form estimation (the control function method)

<sup>c</sup> Loan origination season fixed effects estimates are omitted here, but are available upon request

Table 5 demonstrates the predicted 5-year cumulative probabilities of pecuniary and non-pecuniary prepayments by those covariates. In a declining interest rate environment, the probability of a baseline borrower pecuniarily prepaying a loan is 54%, and the predicted cumulative non-pecuniary prepayment probability by year 5 is around 17%. By contrast, in a rising interest rate environment, for the base loan, the predicted cumulative probabilities of pecuniary and non-pecuniary prepayments by year 5 are approximately 5% and 56% respectively.

As would be expected, the 5-year cumulative probability of non-pecuniary prepayment rises substantially with the contract rate spread at origination in both an environment of declining interest rates and that of rising interest rates, and the increase is more substantial in a rising interest rate environment. Specifically, in a rising interest rate environment, the simulated non-pecuniary prepayment probability with a 75% percentile contract rate spread is about 32 percentage points higher than that with a 25% percentile contract rate spread; while in a declining interest rate environment, this gap is only around 16 percentage points. The original loan amount has a similar impact on the 5-year cumulative non-pecuniary prepayment probability. In a rising interest rate environment, the simulated non-pecuniary prepayment probability rises from 48% to 64% as the original loan amount increases from the 25% percentile to the 75% percentile. In a declining interest rate environment, that increase is only 8 percentage points (14% to 22%). By contrast, the impacts of the original contract rate spread and loan amount on the 5-year cumulative pecuniary prepayment probability are relatively much smaller, especially in a rising interest rate environment.<sup>51</sup>

A borrower's income documentation status (0,1), the underlying property owner occupancy status (0,1), property type (0,1), and a borrower's race and ethnicity group are all shown to affect the predicted cumulative probabilities of pecuniary or/and non-pecuniary prepayments. Recall that the predicted 5-year cumulative probabilities of pecuniary and non-pecuniary prepayments of a base loan are 54% and 17% respectively in a declining interest rate environment, and are 5% and 56% in a rising interest rate environment. The base loan is assumed to be taken by a non-Hispanic white borrower with limited income documentation to purchase a single-family house for investment purposes. Keeping others the same, the 5-year cumulative probability of non-pecuniary prepayment among borrowers with full income documentation drops from 17% to 12% in a declining interest rate environment, and from 56% to 45% in a rising interest rate environment relative to borrowers with limited income documentation. The underlying property owner occupancy status is shown to have a very similar impact on the simulated 5-year non-pecuniary prepayment probability as the factor of the income documentation status. Turning our attention to the impact of the underlying property type on the predicted cumulative pecuniary prepayment probability, this simulated probability among loans with the underlying property as a condo drops from 54% to 47% and from 5% to 4% in a declining/rising interest rate environment respectively, in contrast to that among loans with the underlying property of a single-family house. Additionally, the results in Table 5 clearly demonstrate the lower pecuniary prepayment probability by either Hispanics or African Americans relative to their non-Hispanic white

<sup>51</sup> The major reason why the predicted 5-year cumulative pecuniary prepayment probability also varies with the original loan contract rate spread and loan amount is because all hazards compete with each other (as the likelihood of one hazard increases, the likelihood of others declines), although results in Table 4 show that neither of them has a significant impact on pecuniary prepayment rate.



**Table 5** Predicted 5-year Cumulative Prepayment Probabilities <sup>a</sup>

	Declining Interest Rate Environment <sup>b</sup>		Rising Interest Rate Environment <sup>b</sup>	
	Pecuniary Prepayment Probability	Non-pecuniary Prepayment Probability	Pecuniary Prepayment Probability	Non-pecuniary Prepayment Probability
A Base loan <sup>c</sup>	53.63%	17.07%	4.95%	55.94%
Contract Rate Spread				
25% Percentile	56.16%	11.40%	5.30%	41.47%
50% Percentile	53.63%	17.07%	4.95%	55.94%
75% Percentile	47.63%	27.69%	4.17%	73.51%
Original Loan Amount				
25% Percentile	57.25%	13.56%	5.44%	47.88%
50% Percentile	53.63%	17.07%	4.95%	55.94%
75% Percentile	46.83%	21.83%	4.16%	63.92%
Income Documentation				
Limited income doc.	53.63%	17.07%	4.95%	55.94%
Full income doc	60.39%	12.27%	6.15%	45.46%
Property Owner Occupancy Status				
Not owner occupied	53.63%	17.07%	4.95%	55.94%
Owner occupied	61.67%	12.26%	6.57%	46.18%
Underlying Property Type				
Single family house	53.63%	17.07%	4.95%	55.94%
Condo	47.21%	18.29%	3.87%	56.30%
Housing Price Heterogeneity <sup>d</sup>				
25% Percentile	56.20%	14.65%	5.36%	50.72%
50% Percentile	53.63%	17.07%	4.95%	55.94%
75% Percentile	48.39%	22.16%	4.20%	64.78%
Relative Median Applicant Inc. <sup>d</sup>				
25% Percentile	50.07%	19.19%	4.44%	59.52%
50% Percentile	53.63%	17.07%	4.95%	55.94%
75% Percentile	59.60%	13.36%	5.99%	48.34%
Borrower Race				
Non-Hispanic White	53.63%	17.07%	4.95%	55.94%
Hispanic	50.31%	15.38%	4.25%	50.31%
African American	47.29%	14.50%	3.89%	47.42%

<sup>a</sup> Only covariates that are shown to have different impacts on pecuniary vs. non-pecuniary prepayment hazards are presented. Predicted cumulative probabilities of default are omitted here but available upon request

<sup>b</sup> Based on the CIR term structure model prediction, a mortgage loan originated in January, 2000 is predicted to have a greater chance of experiencing future interest rate drops subsequent to loan origination, while a loan originated in January, 2004 is predicted to have a greater chance of facing future interest rate rises. Therefore, the predicted possible paths (density) of future interest rates since January, 2000/January, 2004 over a 5-year window are used to simulate a declining/rising interest rate environment, respectively. Please notice here, a declining interest rate environment does not mean that the future interest rate will definitely (100%) drop. In

that environment, the interest rate might increase in the future, however, the possibility is relatively low. That is the same with a rising interest rate environment

<sup>c</sup> For a base loan, probabilities are evaluated at the median value of each covariate (time-varying medians are calculated for time-varying covariates) based on the full sample except for those covariates: 1) the baseline borrower is assumed to be a non-Hispanic white with limited income documentation; 2) the underlying property for a base loan is assumed to be a single-family house not owner occupied. Predicted cumulative probabilities in italics are those for the base loan

<sup>d</sup> As house price heterogeneity and relative median applicant income are time-varying variables, median values (50% percentiles), 75% percentiles, and 25% percentiles are based on the full sample over a five-year window since loan origination

counterparts. For example, in a declining interest rate environment, the cumulative pecuniary prepayment probability is 54% for non-Hispanic whites, 50% for Hispanics, and 47% for African Americans.<sup>52</sup>

A higher level of housing price heterogeneity is shown to elevate non-pecuniary prepayment likelihood. Five years after loan origination, the non-pecuniary prepayment likelihood with a 75% percentile housing price heterogeneity is about 7.5(14) percentage points higher than that with a 25% percentile housing price heterogeneity in a declining(rising) interest rate environment. The tract-level relative median applicant income serves to elevate pecuniary prepayment likelihood and to damp non-pecuniary prepayment propensity. For example, in a declining interest rate environment, as the relative median applicant income increases from the 25% percentile to the 75% percentile, the simulated 5-year pecuniary prepayment probability increases from 50% to 60%, while the predicted 5-year non-pecuniary prepayment likelihood decreases from 19% to 13%. Similarly, in a rising interest rate environment, this gap is +2% for pecuniary prepayment probability and -11% for non-pecuniary prepayment probability. Overall the results are consistent with the findings in Table 3, which underscores the importance of distinguishing pecuniary prepayments from non-pecuniary ones for mortgage loan performance analysis and loan pricing.

### Implications for Loan Pricing

In order to further investigate the impacts of the three risks on the pricing of mortgage loans, the 5-year cumulative probability of each event (default, pecuniary prepayment, and non-pecuniary prepayment) is predicted for each loan in the sample and included in the loan contract rate determination model as in Equation (7).<sup>53</sup> The results in Table 6 show that all of the three risks have positive and significant impacts on the contract rate. Specifically, a 10-percentage point increase in default probability would lead to an

<sup>52</sup> As discussed above, because of the competing nature of the three hazards (default, pecuniary prepayment, and non-pecuniary prepayment), the simulated 5-year pecuniary (non-pecuniary) prepayment probability seems to vary by the borrower income documentation status and the underlying property owner occupancy status (the underlying property type and the race and ethnicity of a borrower), although the hazard model results in Table 4 reveal that each of those covariates does not have a significant impact on the corresponding hazard rate.

<sup>53</sup> Please note that while calculating the 5-year predicted cumulative probability of each event for each loan in the sample, a borrower's race/ethnicity is not allowed to impact those predicted probabilities (zero out the coefficients of the race and ethnicity variables), strictly following the current fair-lending laws and regulations (e.g., the Equal Credit Opportunity Act), although the results in Table 4 show that minority borrowers tend to be less likely to prepay than equivalent non-Hispanic white borrowers when market interest rates drop. Please see Ross and Yinger (2002) and Kau et al. (2018) for details.

**Table 6** Loan Contract Rate Estimates - 2SLS <sup>a</sup>

Variable	Parameter	<i>P Value</i> <sup>b</sup>
Intercept	11.1394	<.0001
Predicted Termination Probability		
IV <sup>c</sup> - Predicted default probability	1.0035	0.0014
IV <sup>c</sup> - Predicted pecuniary prepayment probability	1.2719	<.0001
IV <sup>c</sup> - Predicted non-pecuniary prepayment probability	0.8181	0.0001
Loan Characteristics		
10-year treasury yield at origination ( $y_0$ )	0.7455	<.0001
Original LTV categories (Base group: Original LTV $\leq$ 80)		
80 < Original LTV $\leq$ 90	0.1878	<.0001
90 < Original LTV $\leq$ 100	0.3839	<.0001
100 < Original LTV	0.6285	<.0001
FICO at origination continuous linear splines		
Minimum (FICO, 700)	-0.0100	<.0001
Maximum (FICO, 700)-700	-0.0007	0.3221
Original loan amount (in \$10,000)	-0.0319	<.0001
Square term of original loan amount (in \$10,000)	0.0004	0.0003
Full income documentation (0,1)	-0.1190	0.0159
Prepayment penalty categories (Base group: No prepayment penalty)		
Prepayment penalty for 1–3 years	0.6357	<.0001
Prepayment penalty for 5 years	0.3833	<.0001
Property Characteristics		
Property owner occupied (0,1)	-0.1000	0.0466
Property condo (0,1)	0.0588	0.2522
Neighborhood-Level Characteristics at Origination		
Recent housing price appreciation rate at origination	-0.3681	0.0009
Heterogeneity in housing price at origination (in \$10,000)	-0.0075	0.1053
Housing occupancy rate at origination	-0.3941	0.1199
Poverty rate at origination	0.1529	0.5858
Relative median applicant income at origination (HMDA)	0.0218	0.7282
Borrower Race (Base group: Non-Hispanic Whites)		
African Americans	0.1807	0.0360
Hispanics	-0.0219	0.5886
Market/Time Characteristics		
Time Trend	-0.0161	0.2379
Origination season fixed effects <sup>d</sup>	YES	
# of Loans	1404	
Adjusted R <sup>2</sup>	0.6865	

<sup>a</sup> Following the Equal Credit Opportunity Act (ECOA), a borrower's race or ethnicity is not allowed to be considered by loan underwriters. Therefore, while calculating those predicted loan termination probabilities, the impact of race and ethnicity on loan performance is ignored

<sup>b</sup> The method outlined in Appendix 6A of Wooldridge (2010) was employed to correct the standard errors and to calculate the *p*-values to account for the presence of generated regressors

<sup>c</sup> The predicted loan termination probabilities ( $\hat{P}_k$ ) are generated variables and are endogenous in this contract rate equation being a function of the contract rate (spread). To address this issue, a set of generated IVs ( $\hat{P}_k$ ) are used, each serves as a valid IV for its corresponding generated variable ( $\hat{P}_k$ ). Each generated IV ( $\hat{P}_k$ ) is calculated as its corresponding generated variable ( $\hat{P}_k$ ), but with the actual contract rate spread ( $C_0 - y_0$ ) being replaced with a predicted contract rate spread ( $\hat{C}_0 - y_0$ ). The predicted contract rate ( $\hat{C}_0$ ) is obtained from the estimation of a reduced-form contract rate equation

<sup>d</sup> The loan origination season fixed effects estimates are omitted here, but are available upon request

increase in the original contract rate by 10 basis points. The increase in the contract rate due to a 10-percentage-point rise in pecuniary prepayment probability is 13 basis points, while that from a 10-percentage-point increase in non-pecuniary prepayment probability is 8 basis points (5 basis points smaller than that of the pecuniary one). This finding indicates that pecuniary and non-pecuniary prepayment risks do not appear to have the *same* impact on the contract rate, and lenders appear to charge a marginally higher risk premium for pecuniary prepayment risk than for a non-pecuniary one. Although it would generally be hypothesized that non-pecuniary prepayment would lower mortgage contract rates,<sup>54</sup> the results here indicate positive risk premiums for both prepayment types. The positive risk premium of non-pecuniary prepayment needs to be explained through the lens of the subprime market and non-agency MBS. Mayer et al. (2013) find that in a rising interest rate environment, one of the major reasons for a subprime borrower to non-pecuniarily prepay a mortgage loan is an improvement in the borrower's creditworthiness or a positive wealth shock, and the propensity of non-pecuniary prepayment would be higher for riskier borrowers with initial lower credit scores.<sup>55</sup> Thus, when a borrower non-pecuniarily prepays his mortgage loan with the same lender due to an increase in his creditworthiness (or in property equity), the contract rate on the new loan would probably be lower than that on the existing one. This occurrence would bring a negative impact on a lender's net cash inflows. Accordingly, a rational lender who anticipated this type of risk upon loan origination would charge a risk premium for it.<sup>56</sup> In addition, as previously mentioned, the hedging strategy adopted by MBS holders depends on their estimates of the prepayment rates and thus the effective maturity and duration of the MBSs. As non-pecuniary prepayment is not driven by the call option but rather some "exogenous" reasons (e.g., relocation, improvement in borrowers' creditworthiness, etc.) and seems more random, it is relatively more difficult to accurately predict non-pecuniary prepayment rate. Therefore, the risk of non-pecuniary prepayment may make it more challenging to effectively hedge the portfolio or may incur extra hedging costs to MBS investors, which may have been incorporated into MBS/mortgage pricing, especially given the fact that approximately 25% of the loans in the sample of this study were non-pecuniarily prepaid. Based on those considerations, non-pecuniary prepayment risk may raise the contract rate of a loan. Overall, the empirical results indicate that the potential positive impact of non-pecuniary prepayment risk on loan pricing outweighs its negative effect based on this sample of the subprime loans.

The estimation results on all other covariates in this contract rate model are consistent with expectations. A borrower would face a higher contract rate with a

<sup>54</sup> As previously stated, when non-pecuniary prepayment occurs, lenders/investors can reinvest the proceeds from prepayment at a relatively higher interest rate. Besides, the occurrence of non-pecuniary prepayment would limit the value of default and prepayment options given the competing-risks nature of those terminations. Therefore, it would generally be anticipated that non-pecuniary prepayment would reduce contract rates.

<sup>55</sup> Recall that, the empirical results on the non-pecuniary prepayment model in Table 3 indicate the same conclusion that riskier borrowers (e.g., without full income documentation, with less income) tend to be more likely to prepay when interest rates rise.

<sup>56</sup> The prevalence of the prepayment penalties in the subprime mortgage market versus the rareness of the penalties in the prime mortgage market also supports this consideration by the lenders. Recall that, in the sample of the current study, more than half of the borrowers were encumbered by a prepayment penalty for the first few years. As found by Mayer et al. (2013), lenders originating subprime loans used a prepayment penalty as a commitment device to have the borrowers to credibly remain with them for a longer period of time.

higher original LTV ratio, a lower FICO score, without full income documentation, with a prepayment penalty, and with the underlying property as an investment. As the original loan amount rises, the contract rate decreases at a decreasing rate. Also, the results show that lenders rely on information on the local recent housing price appreciation rate to set up the contract rate. In addition, the results show that African American borrowers in this study pay a significantly higher contract rate than their non-Hispanic white counterparts.

## Conclusion

This study explores mortgage pricing behavior and prepayment when a distinction is made between pecuniary and non-pecuniary prepayments. These two types of prepayments, which occur in different interest rate environments (falling & rising), are separated from each other, as they may result in different effects on the future reinvestment rate of return of a mortgage lender/investor. Within the analysis framework of a competing-risks hazard model, the factors driving pecuniary and non-pecuniary prepayment rates are investigated.

Based on a sample of 30-year first-lien fixed-rate nonconforming mortgage loans for home purchase from Miami-Dade County, FL, the results show that different factors drive these two prepayment decisions, and certain variables describing the characteristics of the loan, the borrower, the collateral, and the neighborhood have different impacts on pecuniary vs. non-pecuniary prepayment likelihoods. Overall, both pecuniary and non-pecuniary prepayments are induced by house price appreciation and deterred by prepayment penalties. However, pecuniary prepayment is mainly driven by the change (decline) of market interest rates, and affected by a borrower's income and race group, as well as the type of the underlying property. By contrast, non-pecuniary prepayment propensity is more affected by the traits of the loan (original contract rate spread and original loan amount), the characteristics of the borrower (income documentation status and income), the features of the underlying property (owner occupancy status), as well as the local housing market conditions (house price heterogeneity). A comparison of the estimation results between the hazard models where pecuniary and non-pecuniary prepayments are separated, and the combined model, demonstrates that the separated prepayment hazard models would be preferred, as the combined model does not reflect both interest rate environments. This finding is also supported by the simulation results.

Those findings have important implications for mortgage pricing as the market value of a mortgage loan is affected by various risks. By including the three predicted cumulative termination probabilities (default, pecuniary prepayment, and non-pecuniary prepayment) into a contract rate determination model, this study finds that all of the three termination risks impact the contract rate of a loan positively. This finding underscores the importance of taking into account the three termination risks separately, but simultaneously in mortgage pricing.

**Acknowledgements** I would like to thank Henry J. Munneke and James B. Kau for their helpful comments and suggestions. I acknowledge and appreciate the constructive comments of an anonymous referee.

## References

- Allison, P. D. (2010). *Survival analysis using SAS: A practical guide* (2nd ed.). SAS Institute.
- Agarwal, S., Driscoll, J. C., & Laibson, D. I. (2013). Optimal mortgage refinancing: A closed-form solution. *Journal of Money, Credit and Banking*, 45(4), 591–622.
- Agarwal, S., Rosen, R. J., & Yao, V. (2015). Why do borrowers make mortgage refinancing mistakes? *Management Science*, 62(12), 3494–3509.
- Archer, W. R., & Ling, D. C. (1993). Pricing mortgage-backed securities: Integrating optimal call and empirical models of prepayment. *Real Estate Economics*, 21(4), 373–404.
- Archer, W. R., Ling, D. C., & McGill, G. A. (1996). The effect of income and collateral constraints on residential mortgage terminations. *Regional Science and Urban Economics*, 26(3–4), 235–261.
- Archer, W. R., Ling, D. C., & McGill, G. A. (1997). Demographic versus option-driven mortgage terminations. *Journal of Housing Economics*, 6(2), 137–163.
- Beltratti, A., Benetton, M., & Gavazza, A. (2017). The role of prepayment penalties in mortgage loans. *Journal of Banking & Finance*, 82, 165–179.
- Bennett, P., Peach, R., & Peristiani, S. (2000). Implied mortgage refinancing thresholds. *Real Estate Economics*, 28(3), 405–434.
- Bennett, P., Peach, R., & Peristiani, S. (2001). Structural change in the mortgage market and the propensity to refinance. *Journal of Money, Credit and Banking*, 33(4), 955–975.
- Bhutta, N., & Keys, B. J. (2016). Interest rates and equity extraction during the housing boom. *American Economic Review*, 106(7), 1742–1774.
- Brennan, M. J., & Schwartz, E. S. (1985). Determinants of GNMA mortgage prices. *Real Estate Economics*, 13(3), 209–228.
- Chang, Y., & Yavas, A. (2009). Do borrowers make rational choices on points and refinancing? *Real Estate Economics*, 37(4), 635–658.
- Clapp, J. M., Harding, J. P., & LaCour-Little, M. (2000). Expected mobility: Part of the prepayment puzzle. *The Journal of Fixed Income*, 10(1), 68–78.
- Clapp, J. M., Goldberg, G. M., Harding, J. P., & LaCour-Little, M. (2001). Movers and Shuckers: Interdependent prepayment decisions. *Real Estate Economics*, 29(3), 411–450.
- Cox, J. C., Ingersoll Jr., J. E., & Ross, S. A. (1985). A theory of the term structure of interest rates. *Econometrica*, 53(2), 385–407.
- Deng, Y., & Gabriel, S. (2006). Risk-based pricing and the enhancement of mortgage credit availability among underserved and higher credit-risk populations. *Journal of Money, Credit and Banking*, 38(6), 1431–1460.
- Deng, Y., Quigley, J. M., & Van Order, R. (2000). Mortgage terminations, heterogeneity and the exercise of mortgage options. *Econometrica*, 68(2), 275–307.
- Duan, J. C., & Simonato, J. G. (1999). Estimating and testing exponential-affine term structure models by Kalman filter. *Review of Quantitative Finance and Accounting*, 13(2), 111–135.
- Dunn, K. B., & McConnell, J. J. (1981a). A comparison of alternative models for pricing GNMA mortgage-backed securities. *The Journal of Finance*, 36(2), 471–484.
- Dunn, K. B., & McConnell, J. J. (1981b). Valuation of GNMA mortgage-backed securities. *The Journal of Finance*, 36(3), 599–616.
- Dunn, K. B., & Spatt, C. S. (2005). The effect of refinancing costs and market imperfections on the optimal call strategy and the pricing of debt contracts. *Real Estate Economics*, 33(4), 595–617.
- Firestone, S., Van Order, R., & Zorn, P. (2007). The performance of low-income and minority mortgages. *Real Estate Economics*, 35(4), 479–504.
- Follain, J. R., Scott, L. O., & Yang, T. T. (1992). Microfoundations of a mortgage prepayment function. *The Journal of Real Estate Finance and Economics*, 5(2), 197–217.
- Giliberto, S. M., & Thibodeau, T. G. (1989). Modeling conventional residential mortgage refinancing. *The Journal of Real Estate Finance and Economics*, 2(4), 285–299.
- Green, R. K., & LaCour-Little, M. (1999). Some truths about ostriches: Who Doesn't prepay their mortgages and why they Don't. *Journal of Housing Economics*, 8(3), 233–248.
- Hancock, D., & Passmore, W. (2011). Did the Federal Reserve's MBS purchase program lower mortgage rates? *Journal of Monetary Economics*, 58(5), 498–514.
- Hendershott, P. H., & Van Order, R. (1987). Pricing mortgages: An interpretation of the models and results. *Journal of Financial Services Research*, 1(1), 19–55.

- Kau, J. B., Fang, L., & Munneke, H. J. (2018). An unintended consequence of mortgage financing regulation—a racial disparity. *The Journal of Real Estate Finance and Economics*, 59, 549–588. <https://doi.org/10.1007/s11146-018-9683-y>.
- Kau, J. B., Keenan, D. C., Muller, W. J., & Epperson, J. F. (1992). A generalized valuation model for fixed-rate residential mortgages. *Journal of Money, Credit and Banking*, 24(3), 279–299.
- Kelly, A. (1995). Racial and ethnic disparities in mortgage prepayment. *Journal of Housing Economics*, 4(4), 350–372.
- LaCour-Little, M. (1999). Another look at the role of borrower characteristics in predicting mortgage prepayments. *Journal of Housing Research*, 10(1), 45–60.
- LaCour-Little, M., & Chun, G. (1999). Third party originators and mortgage prepayment risk: An agency problem? *Journal of Real Estate Research*, 17(1), 55–70.
- LaCour-Little, M., Rosenblatt, E., & Yao, V. (2010). Home equity extraction by homeowners: 2000–2006. *Journal of Real Estate Research*, 32(1), 23–46.
- Mayer, C., Piskorski, T., & Tchisty, A. (2013). The inefficiency of refinancing: Why prepayment penalties are good for risky borrowers. *Journal of Financial Economics*, 107(3), 694–714.
- McConnell, J. J., & Singh, M. (1994). Rational prepayments and the valuation of collateralized mortgage obligations. *The Journal of Finance*, 49(3), 891–921.
- Pavlov, A. D. (2001). Competing risks of mortgage termination: Who refinances, who moves, and who defaults? *The Journal of Real Estate Finance and Economics*, 23(2), 185–211.
- Pennington-Cross, A., & Chomsisengphet, S. (2007). Subprime refinancing: Equity extraction and mortgage termination. *Real Estate Economics*, 35(2), 233–263.
- Petrin, A., & Train, K. (2003). Omitted product attributes in discrete choice models. (no. w9452). *National Bureau of Economic Research*.
- Peristiani, S., Bennett, P., Monsen, G., Peach, R., & Raiff, J. (1997). Credit, equity, and mortgage Refinancings. *Economic Policy Review*, 3(2), 83–99.
- Quigley, J. M. (1987). Interest rate variations, mortgage prepayments and household mobility. *The Review of Economics and Statistics*, 69(4), 636–643.
- Ross, S. L., & Yinger, J. (2002). *The color of credit: Mortgage discrimination, research methodology, and fair-lending enforcement*. MIT Press.
- Shoven, J. B. (1986). The effects of interest rates on mortgage prepayments. *Journal of Money, Credit, and Banking*, 18(1), 41–59.
- Stanton, R. (1995). Rational prepayment and the valuation of mortgage-backed securities. *The Review of Financial Studies*, 8(3), 677–708.
- Wooldridge, J. M. (2010). *Econometric analysis of Cross section and panel data*. Cambridge, MA: MIT Press.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.