FEDERAL RESERVE BANK *of* ATLANTA | WORKING PAPER SERIES

Are Lemons Sold First? Dynamic Signaling in the Mortgage Market

Manuel Adelino, Kristopher Gerardi, and Barney Hartman-Glaser

Working Paper 2016-8a
Revised February 2017

Abstract: A central result in the theory of adverse selection in asset markets is that informed sellers can signal quality and obtain higher prices by delaying trade. This paper provides some of the first evidence of a signaling mechanism through trade delays using the residential mortgage market as a laboratory. We find a strong relation between mortgage performance and time-to-sale for privately securitized mortgages. Additionally, deals made up of more seasoned mortgages are sold at lower yields. These effects are strongest in the "Alt-A" segment of the market, where mortgages are often sold with incomplete hard information.

JEL classification: G17, G21, G23

Key words: securitization, mortgage default, adverse selection, signaling, asymmetric information

The authors thank Darren Aiello, Brendan Daley, Stuart Gabriel, Brett Green, Joseph Mason, Christopher Palmer, Anthony Pennington-Cross, Tim Riddiough, Hongfei Tang, Nancy Wallace, Paul Willen, and Basil Williams as well as seminar participants at the 2015 Southern Finance Association Conference, the 2016 American Real Estate and Urban Economics Association national conference, the 2016 Financial Intermediation Research Society conference, and the 2017 Amreican Finance Association meetings for their helpful comments and discussions. They also thank Valeria Vargas-Sejas for her outstanding research assistance. This paper was previously circulated under the title "A Test of Dynamic Signaling Models: Evidence from Mortgage Securitization." The views expressed here are the authors' and not necessarily those of the Federal Reserve Bank of Atlanta or the Federal Reserve System. Any remaining errors are the authors' responsibility.

Please address questions regarding content to Manuel Adelino, Fuqua School of Business, 100 Fuqua Drive, Durham, NC 27708, manuel.adelino@duke.edu; Kris Gerardi, Research Department, Federal Reserve Bank of Atlanta, 1000 Peachtree Street NE, Atlanta, GA 30309-4470, kristopher.gerardi@atl.frb.org; or Barney Hartman-Glaser, Anderson School of Management, 110 Westwood Plaza, Los Angeles, CA 90095, bhglaser@anderson.ucla.edu.

1 Introduction

One of the most widely studied market settings in economics is that of a seller with private information about the quality of an asset facing less-informed buyers. In the presence of such an adverse selection problem, sellers can take actions to reveal their private information, as in the classic signaling model of Spence (1973). This notion of signaling has been successfully applied in theoretical models of financial markets to explain a variety of phenomena, from the optimality of debt (DeMarzo and Duffie (1999)) to the temporary freezing of asset markets (Daley and Green (2012)). While many types of commonly observed behavior are consistent with signaling, such as the attainment of education or the propensity of underwriters to retain equity in an initial public offering, little empirical evidence indicates that agents actually engage in these activities to signal rather than for other reasons. Providing such evidence faces a fundamental challenge: a pure test of signaling theory requires the econometrician to observe agents' private information or hidden "types." We address this challenge by using unique features of the U.S. mortgage market.

We first present a simple model of mortgage sales to provide motivation for our empirical tests. In the model, sellers of high-quality mortgages face lower costs of waiting because their mortgages have lower probabilities of default. A seller privately observes mortgage quality, reflected in the probability of default. We assume that default is publicly observable and eliminates the possibility of a sale. A separating equilibrium emerges in which the time to sale of a mortgage increases in quality, a relation often referred to as the skimming property. Many recent studies, such as Fuchs and Skrzypacz (2013) and Fuchs et al. (2015), have found that the skimming property can emerge in dynamic adverse selection models of financial markets, and a number of others, such as Daley and Green (2012) and Daley and Green (2016), have found that the timing of sales in asset markets can serve as a signal of quality. More broadly, the idea that the timing of actions can reveal private information is a central prediction of many adverse selection models.[1] Thus, our model summarizes the general predictions of the literature on adverse selection and signaling.

The mortgage market is a suitable laboratory for testing the skimming property and, more generally, trade delays as signals of quality, for several reasons. First, mortgages are

[1]See also Noldeke and Van Damme (1990), Swinkels (1999), Janssen and Roy (2002), Grenadier and Wang (2005), Kremer and Skrzypacz (2007), Guerrieri et al. (2010), Grenadier and Malenko (2011), Chang (2014), and Williams (2016).

durable assets characterized by an objective measure of quality based on the probability of default. Detailed micro data are available to investors, originators, and the econometrician on the characteristics of borrowers and mortgage contracts, which together serve as a good proxy for *observable* mortgage quality at the time of the sale. While the outcomes are not known at the time of sale, they are known to the econometrician *ex post*. These *ex-post* outcomes are correlated with *unobserved* heterogeneity in asset quality that is (i) known privately by the seller, as shown in previous studies of the mortgage market,[2] (ii) unknown to potential buyers, and (iii) known to the econometrician *ex post*. The distinction between observable and unobservable asset characteristics is central to our tests and is one of the main reasons that adverse selection models are particularly difficult to test empirically.[3] Most models predict that assets that are observably better should trade faster.

Second, in the middle of the last decade, there was an active secondary market for mortgages in which investors in mortgage-backed securities (the buyers) purchased claims on large portfolios of mortgages. While there is a chain of intermediaries between the originators of mortgages and the ultimate buyers of the securities (as shown in Stanton et al. (2014) and Stanton and Wallace (2015)), we can measure the time to sale from the creation of the asset (when the mortgage is originated) to the sale of the securities that ultimately receive cash flows on those mortgages.[4]

Third, while we concentrate the majority of our empirical analysis on the relation between trade delays and mortgage quality, we are also able to (imperfectly) observe the prices at which mortgage-backed securities were sold. The combination of the availability of observed and unobserved quality measures and prices is rarely available in other contexts.

Using data on mortgages securitized in the non-agency, private-label securitization (PLS) market, we find a clear negative relation between time to sale and the component of mortgage performance that is not explained by observable mortgage characteristics. In our baseline specifications, we find that, after conditioning on all underwriting characteristics, PLS loans sold five months or more after origination are approximately 5 percentage points less likely to default relative to loans sold immediately after origination. This is an economically

[2] See, for example, Demiroglu and James (2012a), Jiang et al. (2014b), Griffin and Maturana (2016), and **?**.

[3] Fuchs et al. (2015) find evidence consistent with the skimming property in the IPO market.

[4] Note that the fact that there may be more than one transfer of a mortgage along this chain biases our tests *against* capturing the role of signaling in transmitting information.

meaningful difference, as it is approximately 30% of the average default rate in our sample (16%). Interpreting these magnitudes through the lens of our model indicates that adverse selection is severe in this market; the difference between the best and worst possible outcomes of the originator's private information is almost one third of the average outcome.

The results on *ex-post* default are in contrast to those using *ex-ante* measures of credit risk. Specifically, we construct the predicted probabilities of default using only information available to mortgage investors at the time that mortgages are sold into PLS deals. We then ask whether *ex-ante* observable credit risk is related to time to sale. We find no relation between *ex-ante* observable risk and time to sale despite the fact that this measure is highly correlated with *ex-post* performance. Put differently, while unobserved quality is related to trade delays, observable risk measures are not.

A potential alternative explanation for our findings is that a longer time to sale simply indicates that a particular mortgage did not default in the months prior to sale. In turn, the fact that a mortgage does not default in the first months after origination may indicate that it is of better quality regardless of the originator's private information. If there is some random delay in the time it takes to sell a mortgage, a longer time to sale may mechanically reflect better quality rather than an intention to signal on the part of the originator. We address this concern by restricting our analysis to mortgages that do not default in the first nine months following origination independent of when they were securitized. In this sample, observing time to sale does not contain any additional *public* information about the default history than the information that the mortgage did not default in the first nine months. Our core result is unchanged in this subsample, and it is still the case that mortgages with a longer time to sale have lower default rates *ex post*.

In contrast to the findings in the PLS segment of the market, we find no evidence of a negative relation between time to sale and mortgage default in a large sample of loans sold to the government-sponsored enterprises (GSEs) Fannie Mae and Freddie Mac. This is consistent with the institutional features of the GSE market, in which automated underwriting and the credit guarantees provided by the agencies essentially eliminate the role of asymmetric information about mortgage credit quality (although not necessarily about prepayment risk) between investors in GSE securities and originators.

Using a second loan-level dataset (CoreLogic), we show that the results are strongest in the "Alternative-A" (or "Alt-A") segment of the market, which is comprised of a majority of low-documentation loans or loans with risk characteristics that prevent them from being securitized in the conforming market. While the subprime segment of the market is riskier than the Alt-A segment, subprime mortgages are more homogeneous in their (potentially unobserved) risk characteristics. The previous literature has found private information to be especially important among low-documentation mortgages, lending further credence to an adverse selection, signaling interpretation.[5]

An additional benefit of the CoreLogic dataset is that it contains information on the identities of originators for a large subset of loans. This allows us to include originator fixed effects in our regressions, which helps address the concern that funding sources (particularly very short-term warehouse loans and repo agreements) might prevent a signaling mechanism from taking place. By estimating within-originator regressions, any variation that comes from systematic differences across originators in funding differences is absorbed by the fixed effects. To the extent that certain types of originators (particularly independent mortgage companies, as pointed out in Stanton et al. (2014) and Ganduri (2016)) relied almost exclusively on these types of funding sources, that variation is accounted for in these specifications. We find similar results to the baseline specifications that do not control for the originator.

As a final test on the default dimension, we separately estimate the correlations between time to sale and default for issuers and originators that are affiliated entities (as in Demiroglu and James (2012a) and Furfine (2014)). This helps distinguish signaling behavior from "unilateral" concerns about warehousing loans on the part of the seller. If our results simply reflected originator reluctance to hold on to bad loans without an intention to signal unobserved quality to buyers, we would expect no differences across affiliated and unaffiliated entities. Instead, we find a significantly weaker negative correlation between time to sale and default risk for the sample of mortgages in which the issuer and originator are affiliated with each other. These results indicate that a key component of informational asymmetry leading to a sales delay is between the originator of the mortgage and the issuer of the security.

We then turn to the pricing dimension to determine whether prices rise with time to sale, as predicted by the signaling model. Data on prices paid for individual mortgages are not available (to our knowledge), so we conduct an analysis of mortgage-backed security (MBS) prices. Because MBS derive their cash flows from pools of individual mortgages,

[5]See, for example, Jiang et al. (2014a), Jiang et al. (2014b), Begley and Purnanandam (2017), and Saengchote (2013)

if signaling plays an important role in the market, then we should see a positive relation between average time to sale at the pool level and MBS prices. Using data on the floating-rate, triple-A, and PLS yield spreads at origination, we find that securities made up of loans that take longer to sell (more seasoned loans) are sold at lower yields. One additional month of average loan seasoning is associated with a 1.5-2-basis-point reduction in the yield of triple-A securities (the average spread is 28 basis points).[6] Consistent with the evidence on mortgage performance, the pricing results are non-linear in seasoning and are strongest in the Alt-A segment of the market.

While our results are consistent with the idea that asymmetric information on default risk did not play an important role in the subprime segment of the PLS market (as opposed to the Alt-A segment), there is evidence from the literature that adverse selection with respect to prepayment risk may have played a role in this market (Agarwal and Yavas (2014)). It is also well known that prepayment risk was an important concern for PLS investors in the pre-crisis period.[7] Thus, as a final exercise, we regress prepayment risk on our time to sale variable, where we focus only on hybrid adjustable rate mortgages (ARMs) and define a negative prepayment event to occur when a borrower prepays more than six months before the date on which the mortgage resets. We find a negative relation between time to sale and the likelihood of an early prepayment. The majority of the hybrid ARMs that we consider are in the subprime segment of the market, which suggests that asymmetric information on prepayment risk may have been a more relevant factor in that market rather than that on credit risk.

This paper relates to the literature on adverse selection and signaling. The seminal work of Akerlof (1970) first identified that markets can break down when some participants have valuable private information. In a related work, Spence (1973) shows that informed agents can take actions to credibly reveal their private information that leads to a separating equilibrium. This insight was first applied to financial markets by Leland and Pyle (1977), who shows that the issuers of IPOs can signal information by retaining an equity stake in the

[6]We do not observe security prices at origination, so we use yield spreads as our measure of pricing (consistent with, among others, Ashcraft et al. (2011), He et al. (2012), and Begley and Purnanandam (2017)). The assumption is that the floating rate securities were almost always issued at par.

[7]For example, in a 2006 primer on mortgage-backed securities, the American Securitization Forum wrote, "Prepayment risk is the key source of cash flow uncertainty in RMBS [Residential Mortgage Backed Securities]."

IPO. DeMarzo and Duffie (1999) uses the equilibrium relation between retention and asset quality to show that debt minimizes the costs associated with the separating equilibrium and is hence an optimal security design. DeMarzo (2005) builds on this idea to show that it is optimal to first pool assets to minimize adverse selection and then to create tranches to minimize signaling costs.

This paper also contributes to the empirical literature on the effects of asymmetric information. The seminal work of Genesove (1993) finds weak evidence of adverse selection in the used car market. Another important paper is Garmaise and Moskowitz (2004), who use commercial real estate transactions to test a number of theories of asymmetric information, including the prediction that securities issuers retain a stake to signal their information. In contrast to our paper, they find no evidence that informed sellers of commercial real estate signal their information through retention. Downing et al. (2009) also consider retention and find that mortgages sold to special purpose vehicles (SPVs) tend to be of lower quality than mortgages not sold to SPVs. Agarwal et al. (2012) find no systematic difference between subprime mortgages sold in the secondary market and those retained on banks' balance sheets. Closest to our setting, Begley and Purnanandam (2017) find that higher levels of equity tranches in PLS deals (a measure of retention) are associated with lower delinquency rates and higher prices. Aiello (2016) finds evidence that borrower payment behavior during the warehouse period can be a source of private information for originators. An et al. (2011) find that information asymmetries in the secondary commercial mortgage market can lead to market break down. They argue that conduit lenders exist as a way to mitigate assymetric information.

Two studies document misrepresention in the private mortgage market. Piskorski et al. (2015) find that lenders often misrepresented loan-to-value ratios when selling mortgages and Garmaise (2015) finds that borrower's often misreport the value of their personal assets on mortgage applications. These studies indicate scope for private information in the mortgage market.

2 A Model of Signaling through Delayed Trades

To motivate our empirical tests, we present a simple model of adverse selection and delayed trade in the secondary market for mortgages. Time is infinite, continuous, and indexed by t.

The model consists of a mortgage originator and a competitive market of mortgage investors. All agents are risk neutral. At time $t = 0$, the seller originates a mortgage for potential sale to the market. This mortgage produces a cash flow of c dollars per unit of time until it defaults at some random time τ. The default time τ is an exponential random variable with parameter λ distributed on the compact interval $[\lambda_\ell, \lambda_h]$ according to the continuous density $f(\lambda)$. While $f(\lambda)$ is common knowledge, the seller privately observes λ at the origination of the mortgage. As is common in such settings, we refer to λ as the seller's type.

While both the seller and potential investors are risk-neutral, gains from trade are generated by a difference in discount rates used by the two classes of agents. Specifically, the seller discounts cash flows at rate γ, while the investors discount cash flows at rate $r < \gamma$. This difference in discount rates proxies for the difference in the investment opportunity sets of the seller and the investors. The seller has the technology to originate mortgages, while investors can only purchase mortgages in a competitive market once they have already been originated. We note that modeling these gains from trade as a difference in discount rates is convenient for the following analysis that follows, but not necessary. As long as the gains from trade between the seller and investors are monotonic in the seller's type, λ, the predictions of the model remain qualitatively unchanged.

We assume that mortgage default is publicly observable, such that if the mortgage defaults before the seller has sold it to the investors, no sale occurs. In choosing when to sell the mortgage, the seller takes market price function $P(t)$ as a given. Note that the lowest possible value of a mortgage to investors is

$$p_h = E\left[\int_t^\infty e^{-r(s-t)}\mathbb{1}(s \leq \tau)cds|\lambda_h\right] = \frac{c}{r+\lambda_h},$$

while the highest possible value is

$$p_\ell = E\left[\int_t^\infty e^{-r(s-t)}\mathbb{1}(s \leq \tau)cds|\lambda_\ell\right] = \frac{c}{r+\lambda_\ell};$$

thus, $P(t) \in [p_h, p_\ell]$.

An outcome of this game is a triple $(\lambda, t, p) \in [\lambda_\ell, \lambda_h] \times [0, \infty) \times [p_h, p_\ell]$, where λ is a realization of the seller's type and t and p respectively correspond to the time and price at which trade takes place if the mortgage has not defaulted by time t. The value for the seller of an outcome of the game is then given by

$$U(\lambda, t, p) = E\left[\int_0^t e^{-\gamma s}\mathbb{1}(s \leq \tau)cds + e^{-\gamma t}\mathbb{1}(t \leq \tau)p\big|\lambda\right]$$
$$= \frac{c}{\gamma + \lambda}\left(1 - e^{-(\gamma+\lambda)t}\right) + e^{-(\gamma+\lambda)t}p.$$

An important feature of the seller's payoff function is the so-called single-crossing property: fixing a price p, it is less costly for better (lower default risk) sellers to delay trade. Intuitively, the lower the default risk, the greater the private value of the cash flows that accrue to the seller from the mortgage before the sale, and the greater the probability that the mortgage will remain current so that it can be sold in the future. This feature of the model gives rise to the common *skimming property*, which is present in much of the literature on dynamic trading and asymmetric information,[8] and which is more broadly related to the literature on costly signaling with adverse selection.[9] In our model, the skimming property can be expressed as follows. For a given price function $P(t)$, better sellers wait (weakly) longer to trade, and thus, a trade delay can act as a signal of quality.

An equilibrium of the game is a pair of functions (T, P), where $T(\lambda)$ is the time at which a seller of type λ trades and $P(t)$ is the price of a mortgage sold at time t such that the following conditions hold:

1. Seller optimality:
$$T(\tilde{\lambda}) \in \arg\max_t U(\tilde{\lambda}, t, P(t),)$$

2. Zero profit for the investors:
$$P(T(\tilde{\lambda})) = E\left[\frac{c}{r + \tilde{\lambda}}\big|T(\tilde{\lambda})\right].$$

An equilibrium is separating if $P(T(\tilde{\lambda})) = \tilde{\lambda}$.

[8] See, for example, the early literature on sequential bargaining models with one-sided incomplete information (Fudenberg and Tirole (1983), Sobel and Takahashi (1983), Cramton (1984), Fudenberg et al. (1985), Gul et al. (1986), Gul and Sonnenschein (1988), Ausubel and Deneckere (1989)), Evans (1989) and Vincent (1989). It is also present in dynamic auction models with private information (Vincent (1990)) and competitive markets models of durable goods with private information (Janssen and Roy (2002)).

[9] For example, Spence (1973) and Leland and Pyle (1977)

We focus on characterizing a separating equilibrium. Although other equilibria, such as pooling equilibria, may exist, they are eliminated by standard refinement criteria, such as the D1 refinement of Cho and Kreps (1987). The following proposition characterizes the unique separating equilibrium of the game.

Proposition 1. *The unique separating equilibrium of the game is given by*

$$T^*(\lambda) = \frac{\log(r+\lambda_h) - \log(r+\lambda)}{\gamma - r} \qquad\qquad P^*(t) = p_h e^{(\gamma-r)t}. \tag{1}$$

The method to derive the equilibrium of Proposition 1 is as follows. First, fix some candidate price function $P(t)$ and take a first-order condition for the seller's problem

$$c - (\gamma + \tilde{\lambda})P^*(t) + \frac{d}{dt}P^*(t) = 0. \tag{2}$$

Next, use the fact that for any separating equilibrium,

$$P^*(T(\tilde{\lambda})) = \frac{c}{r+\tilde{\lambda}},$$

and substitute into equation (2) to obtain the following ordinary differential equation for $P^*(t)$:

$$\frac{d}{dt}P^*(t) = (\gamma - r)P^*(t). \tag{3}$$

Finally, because the highest default risk type does not benefit from delaying trades in a separating equilibrium, we must have $T^*(\lambda_h) = 0$ and, hence, $P^*(0) = p_h$. The functions given Proposition 1 solve equations (2) and (3) with this initial condition.

To connect the equilibrium given in Proposition 1 to our empirical analysis, it is useful to consider how the type of seller changes with time to sale. We let $\lambda^*(t)$ denote the type of seller that chooses to sell at time t. Applying Proposition 1, we have

$$\lambda^*(t) = (r+\lambda_h)e^{-(\gamma-r)t} - r. \tag{4}$$

Our empirical results relate to the following key properties of the functions $\lambda^*(t)$ and $T^*(\lambda)$.

1. The default risk of the mortgage decreases with time to sale:

$$\frac{d}{dt}\lambda^*(t) < 0.$$

This means that adverse selection creates a negative relation between time to sale and default risk.

2. The price of the mortgage increases with time to sale,

$$\frac{d}{dt}P^*(t) > 0.$$

This means that adverse selection creates a positive relation between price and time to sale.

3. The maximum time to sale for a mortgage is increasing in the difference in default risk between the safest and riskiest mortgage,

$$\frac{d}{d(\lambda_h - \lambda_\ell)}T^*(\lambda_\ell) > 0.$$

This means that a more severe adverse selection problem, such as when the uncertainty about the mortgage default risk is greater, leads to longer trade delays.

Although the separating equilibrium we detail above is the unique equilibrium selected by D1, a discussion of other possible equilibria is in order. In particular, many pooling equilibria can exist in which all seller types sell at the same time. For example, if investors believe that any mortgage sold after time $t = 0$ is the riskiest type, then all seller types find it optimal to sell their mortgages at $t = 0$ because delaying the sale only leads to forgone gains from trade and does not increase the sale price. However, imposing D1 refinement eliminates this equilibrium. If investors observe an off-equilibrium-path action, such as a seller delaying a trade when investors expect an immediate sale, then D1 requires that they only place positive weight on those seller types who would gain from deviating given the largest set of prices. This set is always largest for sellers of the least risky mortgages because it is less costly for them to delay trades than for any other seller type. As such, D1 requires that investors must believe that the seller is the least risky type if she even slightly delays a trade. These beliefs thus imply that sellers of the least risky type have a profitable deviation, eliminating the simple pooling equilibrium. Thus, we focus our empirical analysis on the separating equilibrium detailed above.

Before proceeding further, a brief discussion of the relation between our model and the literature is in order. For the sake of simplicity, we have assumed that the seller can commit to a time to sale, and in that sense our game is essentially static as in the model of Spence (1973) in which students commit to a particular period of education. Swinkels (1999) shows that without commitment, the signaling equilibrium of Spence might not exist. However, a number of authors, for example Daley and Green (2012), have recently argued that dynamic concerns can restore delay in trade as a signal of quality. In a dynamic version of our model in the spirit of Fuchs and Skrzypacz (2013) or Fuchs et al. (2015), the qualitative results of our model are unchanged.

2.1 Random Delays, Defaults, and Prices

To impose further discipline on our empirical analysis, we consider a plausible variation to our model in which a correlation between delayed trades and *ex-post* performance need not be the signature of dynamic signaling or adverse selection. Intuitively, if a trade is randomly delayed, then some higher-risk mortgages may default before they can be sold. As a result, mortgages that take longer to sell are positively selected (i.e., they are of higher quality than those that could not be sold). This selection mechanism would then lead to a positive correlation between time to sale and *ex-post* performance (a negative correlation between time to sale and default rates). This implies that investors who understand this selection issue believe that mortgages that sell after a longer period of seasoning are higher quality and thus that prices increase with seasoning. Importantly, this effect does not arise from signaling, as mortgages are sold randomly into pools by assumption, but rather through a learning process. As such, a simple model comprising a randomly delayed trade and the associated selection mechanism may appear observationally equivalent to our signaling model for delayed trades. This is a key difficulty in bringing models of asymmetric information to the data: they often have similar predictions to models with symmetric information. We can overcome this difficulty in our setting by observing that the selection mechanism can be undone by conditioning the analysis on mortgages that do not subsequently (after sale) default up to a pre-specified period.

To make this intuition more precise, suppose that the mortgage seller detailed above has the same information as potential investors. Specifically, she knows that the mortgage she wants to sell has an exponential default time with an intensity $\tilde{\lambda}$ uniformly distributed on $[\lambda_l, \lambda_h]$. When she chooses to sell the mortgage, there is a delay between the point at which she lists the mortgage for sale and the moment at which the transaction is recorded, which is exponentially distributed with parameter μ. If the mortgage defaults before the transaction can be recorded, then no sale takes place. Thus, observing that the mortgage transacts at time t reveals that the mortgage did not default prior to t. Thus, the expected quality of a mortgage that transacts at time t is given by the following expression:

$$\begin{aligned} E\left[\tilde{\lambda}|\text{sold at time } t\right] &= E\left[\tilde{\lambda}|\tau > t\right] \\ &= \lambda_h + \frac{1}{t} - \frac{\lambda_h - \lambda_l}{1 - e^{-t(\lambda_h - \lambda_l)}}, \end{aligned}$$

which is increasing in the sale time t. Thus, a randomly delayed trade is associated with negative correlations between time to sale and *ex-post* default outcomes and *ex-ante* prices. These predictions are essentially the same as properties 1 and 2 of the signaling model described above, which means that to test the predictions of the signaling model in the data, we need to overcome this selection effect.

One simple way of accounting for this selection effect is to condition the analysis on loans that do not default until some exogenously specified time s, which needs to be after the period of sale, t. To see this, note that for loans that do not default before s, the event that the mortgage was sold at time $t < s$ does not contain any additional information about the default risk of the mortgage. The expected quality of a mortgage that has not defaulted by time s and is sold at time $t < s$ is given by the following expression:

$$\begin{aligned} E\left[\tilde{\lambda}|\text{sold at time } t < s \text{ and } \tau > s\right] &= E\left[\tilde{\lambda}|\tau > s\right] \\ &= \lambda_h + \frac{1}{s} - \frac{\lambda_h - \lambda_l}{1 - e^{-s(\lambda_h - \lambda_l)}}, \end{aligned}$$

which is independent of the time of sale t. Thus, in a model with a random delay and no signaling mechanism, there is no correlation between the time to sale and *ex-post* default outcomes if we condition on a sample of mortgages that do not default before s, where $s > t$.

This is in stark contrast to our model of signaling through delayed trades, in which time to sale always reveals information about the *ex-post* default risk. We explore whether such a model can explain our results in our empirical tests below.

3 Background on the U.S. Mortgage Market

Our primary focus in this paper is on loans that were sold and then securitized by private financial institutions (or issuers). This segment of the market, often referred to as the PLS (private-label securitization) market, was the source of the initial mortgage foreclosure crisis in 2007, which led to the broader financial crisis and the Great Recession. The PLS market grew rapidly during the housing boom of the mid-2000s, reaching a peak share of approximately 56% of the securitization market in 2006, before completely shutting down in the summer of 2007 when subprime mortgage defaults dramatically increased.

The PLS market is split into three broad segments according to the degree of credit risk. The three segments are referred to as "subprime," "alternative-A" (or "Alt-A"), and "prime jumbo." The collateral in prime jumbo PLS is made up of large loans to borrowers with typically very good credit scores that exceed the conforming loan limits and are thus not eligible to be securitized by the GSEs in the agency market.[10] The "Alt-A" PLS segment, also commonly referred to as "near prime," is typically characterized by loans to borrowers with slightly lower average credit scores than prime jumbo (but comparable to average credit scores in agency pools), and in which borrower income and/or assets are less than fully documented (i.e., low-documentation mortgages). These loans are also more likely to finance investor or vacation home properties. Alt-A PLS includes a mix of loans above and below the conforming loan limit. Finally, the collateral underlying subprime private-label securities is made up of loans that are usually below the conforming loan limit given to borrowers with low credit scores and includes a large fraction of cash-out refinance mortgages. The majority of subprime PLS loans did not meet the underwriting standards in the agency market and were broadly viewed as low-quality mortgages by market participants. Our primary dataset (from Lender Processing Services, described in more detail below) includes loans from all three segments of the PLS market, while our secondary source of data (CoreLogic's Loan-

[10]To be securitized by the GSEs, a mortgage must have a principal balance below the conforming loan limit, a loan-to-value ratio at or below 80%, or equivalent credit enhancements (e.g., private mortgage insurance).

Performance database, also described below) includes loans from the subprime and Alt-A segments of the market.

There is significant variation in the funding and operational models of mortgage originators in the PLS space, including independent mortgage companies, affiliated mortgage companies, and others. We refer the reader to Stanton et al. (2014) and Ganduri (2016) for detailed descriptions of the structure of the market. Stanton et al. (2014) show that repurchase agreements and warehouse lines of credit with very short maturities are a large funding source in the PLS market. This limits the originators ability to delay mortgage sales. For the purposes of our tests, we require that either originators of mortgages or issuers of PLS (or both) have the ability to hold on to mortgages and delay trades, even if some are limited by contractual features because their funding sources.[11]

We focus on loans sold into the PLS market for two reasons. First, many recent papers in the literature have documented a significant amount of private information in these markets, especially in the population of low-documentation mortgages, and have found originators to be at least partially aware of unobserved quality.[12] In contrast, private information about credit quality plays a much less important role in the agency securitization market, where the GSEs provide specific parameters regarding the underwriting criteria that they accept and agree to purchase (usually through an automated process) all loans that satisfy those criteria.

Second, our PLS data are very similar in scope to the data used by many participants in the institutional PLS market to produce valuations and to monitor performance after issuance. Some of the data we use originate from the trustees' reports provided to PLS investors in the market. Thus, our data closely match the set of underwriting characteristics that PLS issuers and investors use to make real-time purchasing decisions. This is central to the implementation of our empirical tests described below.

[11]Even though we find that the majority of loans in the PLS market were securitized within the first two months after origination, consistent with the evidence provided in Stanton et al. (2014) that warehouse loans and repurchase agreements had 30 to 45 days' maturity, the variation that is most relevant for our tests is sales past this time period (up to 9 months after origination).

[12]For example, see Demiroglu and James (2012a) and Jiang et al. (2014b).

4 Testing for Dynamic Adverse Selection Using Mortgage Data

We implement empirical tests of predictions 1 and 2 of the signaling model developed in section 2. Prediction 1 is that time to sale and mortgage quality should be positively related and that we should thus find a negative correlation between time to sale and *ex-post* default rates. Prediction 2 is that there should be a positive correlation between time to sale and mortgage prices. Given superior data and the ability to perform much richer cross-sectional tests, we focus on the analysis using loan-level default.

4.1 Time to Sale and Mortgage Default

A key issue in implementing an empirical test of the skimming property is distinguishing between observable and unobservable asset quality. Signaling in general and the skimming property in particular refer specifically to quality that the seller is informed about but is unobservable to the buyer.

We implement a strategy similar to Adelino et al. (2016) that uses conditional measures of loan performance to isolate aspects of loan quality that are unobservable to investors at the time of purchase but are correlated with the originators' (and possibly the issuers') information set (and which, by virtue of the passage of time, become observable to the econometrician). Specifically, we condition performance on a large set of loan and borrower characteristics used in mortgage underwriting models that are readily available to issuers and institutional investors in the MBS market. Our empirical specifications take the following general form:

$$Default_{ijt} = \alpha + \beta_1 * \text{Months to Sale}_{ij} + \beta_2 * X_{ijt} + \epsilon_{ijt} \tag{5}$$

where i indexes the individual mortgage, j indexes the geographic area in which each mortgage is originated, and t indexes the horizon over which we calculate the default rates. X_{ijt} is a vector of mortgage-level control variables that includes the relevant observable borrower, loan, and geographic characteristics, including detailed fixed effects. Months to Sale_{ij} is a variable that measures the time between when a mortgage is originated and when it is sold into the secondary market and securitized.

The existence of private information and signaling in the mortgage market predicts $\beta_1 <$

0. This is a joint test of two hypotheses: that (i) the seller's private information, I_{seller}, is correlated with loan quality after accounting for underwriting characteristics,

$$Corr[(E(Default_i|X_i, I_{seller}) - E(Default_i|X_i)), Default_i)] \neq 0 \tag{6}$$

and that (ii) sellers signal asset quality by delaying trades.

It is important to note that our tests do not require that we observe the full information set of the buyers. Instead, the tests require a weaker condition, namely, that our measure of *ex-ante* default risk be an unbiased estimate of the "true" credit risk. Additionally, we assume that $X_i \subseteq I_{buyer} \subset I_{seller}$, where both buyers and sellers' information sets include the mortgage characteristics we observe, and that sellers have some private information about the loans that is correlated with default. In such a setting, we can measure the relation between time to sale and observed (*ex-ante*) credit risk using our information set. To the extent that credit risk is the only variable that is systematically related with time to sale, the additional information that investors may have that we do not provides more precision for measuring credit risk but does not change the direction of the relation. Put differently, if we find no relation between observable risk and time to sale for our (very comprehensive) measure that buyers and sellers also have access to, our assumption is that this relation would not change if the public signal became more precise. This is a weaker condition than requiring that the buyers' information set be the same or a subset of ours.

4.1.1 Default Measurement and Controls

We consider two default horizons, 36 and 60 months, in our primary specifications; these are measured relative to the month of loan origination.[13] We also consider a mortgage to be in default if the borrower is either two payments behind (60+ days delinquent) or three payments behind (90+ days delinquent) at any point between origination and each default horizon. We use 60-day and 90-day delinquency cutoffs rather than the initiation of foreclosure proceedings so that our default definition reflects borrower behavior that is not confounded by the decisions of mortgage servicers.

X_{ijt} in equation 5 above accounts for a large subset of the information held by the buyers of mortgages at the time of sale. According to Stearns (2006), all issuers and most PLS

[13] We also tried a shorter horizon of 24 months; it did not make a material difference.

investors have access to detailed information at the loan level, including such data fields as original loan balance, FICO score, combined loan-to-value ratio, documentation type, occupancy type, loan purpose (refinance or purchase), property type, loan size, amortization schedule, interest rate, loan type (ARM vs. FRM), and information on the geographic location of the property.[14] Our vector of control variables includes all of these variables plus some variables that measure *ex-post* conditions in the local housing market, including the county-level unemployment rate and the level and the changes of the house price index (normalized by setting the index value for January 2000 to 100 for each county). The appendix contains a list of the exact variables that we include in our covariate set. In addition, we include a full set of state-level fixed effects and fixed effects corresponding to the year-quarter of origination and the year-quarter of loan sale.[15] Additional indicator variables are included whenever there are missing observations for any of the controls.

4.2 Time to Sale and Mortgage Spreads

We do not have access to data on individual mortgage prices.[16] Thus, we are forced to conduct our pricing analysis at the security level. While we also lack explicit data on security transaction prices at the time of issuance, we are able to construct a good proxy using yield spreads. Specifically, we focus on the average spread (quoted as a spread over the one-month LIBOR rate) of floating rate triple-A mortgage-backed securities in the PLS market. We calculate a weighted average spread at the deal-level, where we weight spreads by the face value of the securities.[17] Because we do not have information on the actual prices paid for the securities, restricting the analysis to floating rate securities virtually eliminates the

[14]This contrasts with the agency market, as the GSEs, partly because they absorb all credit risk, do not disclose as much detailed information about the mortgages that back their securities. According to Stearns (2006), "Non-agency investors have access to a wealth of data–all at the loan level–that agency investors can only dream of. "

[15]We have also experimented with a specification that includes zip code-level fixed effects to absorb any effects of unobserved geographic shocks at a very fine geographic level. We found that the results were largely unaffected. Because including such a large number of fixed effects becomes very computationally demanding, we use state fixed effects in all of the tests in the paper.

[16]To our knowledge, such data simply do not exist.

[17]Whenever a given PLS deal is made up of more than one pool of mortgages and triple-A securities have claims to cash flows from only one of the pools, the average spread and all controls are calculated at the pool level (rather than at the deal level). This follows the approach in Adelino et al. (2016), who compare outcomes across pools sold to different investors.

possibility that the securities were not issued at par. In addition, these floating rate securities have very short durations, so we can ignore the interest rate risk and the negative convexity problem that arises with fixed-rate mortgage-backed securities.

Our empirical analysis considers the relation between average yield spreads and mortgage seasoning. The seasoning variable, which is calculated as the average months to sale in the pool, and all controls are constructed from loan-level data and aggregated to the pool level. Our specifications take the following form:

$$Spread_i = \alpha + \beta_1 * \text{Seasoning}_i + \beta_2 * X_i + \epsilon_i \tag{7}$$

where i represents a pool and X_i, which is described in detail below, includes the pool averages of all relevant loan and borrower characteristics used in the loan-level tests and the quarter of issuance fixed effects. Our model of adverse selection and signaling predicts a negative relation between average seasoning and mortgage spreads, $\beta_1 < 0$.

4.3 Data

In this section, we describe the two loan-level datasets used in this paper. While both loan-level datasets are similarly structured monthly mortgage panels, there are important differences in the scopes of their coverage and in some of the underlying variables that produce advantages and disadvantages in the context of our analysis.

The pricing data at the individual security level were obtained from Bloomberg. The data fields include security identifiers (including CUSIP and ticker), issuer name, issuance date, the identification of the loan pool that the security has claims on, the spread over one-month LIBOR at origination, and the weighted average life as advertised in the prospectus. The dataset obtained from Bloomberg covers over 90% of all subprime PLS issued in the U.S. between 2002 and 2007. We are able to combine the CoreLogic and Bloomberg datasets by merging individual security CUSIPs.

4.3.1 Lender Processing Services Data

Our primary dataset comes from Lender Processing Services (LPS). The LPS dataset covers between 60% and 80% of the U.S. mortgage market and contains detailed information on the characteristics and performance of both purchase-money mortgages and refinance mortgages. The LPS dataset is constructed using information from mortgage servicers, financial institutions that are responsible for collecting mortgage payments from borrowers. Each loan is tracked at a monthly frequency from the month of origination until it is paid off voluntarily or involuntarily via the foreclosure process. We focus on loans originated during the housing boom, from January 2002 through December 2007.

Importantly, for the purposes of this study, the dataset includes a time-varying variable, "investor type," which identifies whether a mortgage is held in a bank's portfolio, is privately securitized, or is securitized by the GSEs. This variable allows us to identify if and when a loan is securitized or sold to a GSE.

We adopt several sample restrictions in our analysis of the LPS data. We consider only first lien mortgages originated in the 2002–2007 period that were sold to PLS issuers or to the GSEs.[18] We only keep loans originated in the 50 United States and restrict the sample to loans that enter the dataset in either the same month of origination or in the month following origination. We also address outliers in the data by winsorizing the distributions of credit scores, original loan balances, LTV ratios at origination, and interest rates at origination at the 1st and 99th percentiles of each distribution.[19]

The primary advantages of using LPS data to test the skimming property are the ability to precisely identify the month of sale and the ability to consider sales to both PLS and the GSEs. However, there are several important drawbacks. The biggest problem with the LPS data in our context is a lack of information on the exact identity of the financial institution that originates the mortgage. Ideally, we want to control explicitly for the identity of the originator, as this would eliminate the potential heterogeneity in underwriting practices that is known to the PLS and GSE issuers but not to us. In addition, there is some concern that the LPS dataset may under-represent the PLS market during our sample period. For these reasons, we also use data from CoreLogic's LoanPerformance database.

[18]Thus, we eliminate loans kept in the portfolios of the mortgage originators and never sold. In addition, a small number of loans in the dataset were sold to the Federal Home Loan Banks (FHLBs), which we also eliminate from the sample.

[19]We also tried trimming instead of winsorizing the data and found that this change had little effect on the results.

4.3.2 CoreLogic Data

Our second source of mortgage data is CoreLogic's LoanPerformance (CL) database, which covers virtually the entire subprime and Alt-A segments of the private-label securitization market. Like the LPS dataset, CL contains detailed information on underwriting characteristics and monthly loan performance, but unlike LPS, CL does not have information on portfolio-held loans or loans securitized by the GSEs. however, one of the main advantages of using CL data is their representativeness of the PLS market.[20]

The CL database includes virtually the same mortgage and borrower characteristics (at the time of loan origination) as the LPS database, but, importantly, about 50% of the CL database includes the identity of the originating institution, which allows us to include originator fixed effects, such as comparing loans made by the same originator with different times to sale. In addition to the identity of the originator, CL provides information on the identity of the mortgage servicer and on security identifiers (CUSIPs) and deal identifiers, which allows us to obtain information on the identity of the securitizer (issuer) for most of the loans in the sample.

CoreLogic also allows us to distinguish between the subprime and Alt-A markets.[21] We display the distribution of months to sale (Table 2) and the summary statistics (Table 4) for the subprime and Alt-A loans separately. The tables show that the sample of Alt-A loans in CL appears more similar to the LPS sample. The Alt-A distribution of months to sale more closely resembles the LPS distribution, as a higher fraction of Alt-A loans are sold immediately compared to subprime loans. In addition, the average loan size, interest rate, and FICO score in the Alt-A sample are closer to the LPS sample compared to subprime loans.

The timing for when a loan enters each dataset is also different across the LPS and CL datasets. In LPS, we observe most loans from the month of origination, and we can directly observe the month in which they are sold out of banks' portfolios to PLS issuers or the GSEs. In CL, however, we compute time to sale as the difference between the date of issuance of the

[20] According to CoreLogic's website, the dataset contains information on mortgages that make up over 97% of outstanding non-agency PLS pool balances (http://www.corelogic.com/solutions/data-resources-for-capital-markets.aspx#rmbs).

[21] There is a servicer-provided field in LPS that distinguishes Grade "A" loans and Grade "B" and "C" loans, but loans flagged as "B" and "C" in LPS do not appear to be similar to subprime loans in CL in terms of observable underwriting characteristics.

mortgage-backed security in which the loan is included and the reported month of origination of the mortgage.[22]

4.4 Summary Statistics

Table 1 displays the distribution of the number of months between origination and sale for our sample of PLS and GSE securitized mortgages in the LPS data. It is clear from the table that the majority of both PLS and GSE securitized mortgages are sold very quickly, either immediately or only one month after origination. However, there are some important differences between the PLS and GSE distributions. Very few GSE loans (about 7%) are sold more than two months after origination, but a non-trivial fraction of PLS loans are sold later in their lives (about 20% are sold more than two months after origination). We impose a final sample restriction, which is a maximum threshold of nine months between origination and sale, to ensure that we have the power to identify the non-parametric specifications below and to ensure that the loans in the sample were originated with the intention of being sold (which might not be the case for loans sold significantly past this threshold).[23] This leaves us with a sample of over 5 million loans sold to PLS issuers and over 11 million loans sold to the GSEs.

In Table 3, we display the summary statistics for many of the control variables in the empirical models. The table displays the statistics for the sample of loans sold to PLS issuers and the sample of loans sold to the GSEs. In general, PLS loans are characterized by riskier attributes than are GSE loans. For example, there are more interest-only loans, more adjustable-rate loans, more low-documentation loans, more subprime loans, and more loans that carried prepayment penalties in the PLS sample.

We apply the same sample restrictions to the CoreLogic data that we applied to the LPS data. Table 2 displays the distribution of months to sale in the CoreLogic dataset, while Table 4 provides some basic summary statistics. The first notable observation is that there are many more PLS loans in CoreLogic than in LPS.[24] Also of note is that the distribution

[22] Loans enter the CL dataset on the issue date, so we do not see the performance history of loans before they are securitized.

[23] We have experimented with higher thresholds, such as 12 months, but these had little effect on the estimation results.

[24] The LPS sample size of 5.3 million loans listed in the tables understates the total number of PLS loans, as some seasoned mortgages are eliminated from the sample because of restriction of only including loans for

of months to sale in CL is similar to LPS, although there are several subtle differences. In both datasets, over 90% of loans that are securitized are sold within 5 months of origination, but a lower fraction of loans are sold within the first 2 months in the CL database (45%) compared to the LPS database (56%).

Table 4 shows that the CL sample is characterized by significantly lower credit scores (FICOs), higher interest rates, and lower loan amounts. There is a much higher fraction of adjustable-rate mortgages and low-documentation loans in CL. There also appears to be a large difference in the average LTV ratios, but this is likely because the LTV ratio in CL incorporates second mortgages (i.e., piggybacks), while LPS only provides the LTV ratio based on the first lien. In addition, the average (unconditional) default rates are significantly higher in the CL sample. Overall, based on the average underwriting characteristics, the sample of PLS loans in CL appears to be significantly riskier than the LPS sample.

Table 14 in the online appendix shows the summary statistics of all of the pool-level characteristics used in the pricing analysis. The average spread of triple-A securities in the data is 28 basis points, with a standard deviation of 23 basis points. This spread is computed as the pool-level average of all triple-A securities drawing cash flows from a given pool, and the sample is limited to pools with only floating rate triple-A securities. The average pool-level seasoning in the data is 3.3 months, and it is truncated at 9 months following the approach for the default analysis. About 97.5% of pools have an average seasoning below 9 months (Figure 5 shows the histogram and cumulative distribution of the pool-level seasoning variable). Pools are made up of 2,355 loans on average (the median is 1,911), with an average FICO score of 640 and a combined loan-to-value ratio of 84%.[25]

5 Results

In this section, we present the results on the empirical relation between time to sale and loan quality and the relation between time to sale and prices. Because time to sale is the key

which we have a full history of performance. More than 7 million PLS loans originated between 2002 and 2007 (inclusive) in the LPS database.

[25]We include pool-level averages of FICO and CLTV as covariates in our pricing analysis, plus categorical variables that capture the fraction of loans in each pool that fall into various FICO and CLTV categories (displayed in Table 14). In addition, we include a variable corresponding to the fraction of loans in a pool that have an LTV ratio that is exactly equal to 80% to capture the potential importance of piggyback loans, which we do not directly observe.

variable of interest, we first implement tests using simple linear specifications (consistent with the prediction in the model) so that Months to Sale_{ij} (for the loan-level default analysis) and Average Seasoning_i (the pool-level average used in the pricing regressions) take values from 0 to 9 and enter the regressions linearly. We then consider the specifications that allow for potential non-linearities. For the loan-level default regressions, we estimate a non-parametric specification in which we include separate indicator variables for each value of the months to sale variable.[26] For the pool-level pricing tests, we consider a specification that includes a quadratic term for the average seasoning of loans in a pool.

5.1 Default and Time-To-Sale

In this and the subsequent sections, we analyze mortgage quality (measured by default) as a function of time to sale. Panel A of Table 5 displays the results for the linear regression specification estimated on our sample of loans in the LPS dataset. The panel displays estimation results for our variables of interest for two different default definitions (60+ DQ and 90+ DQ, 60- and 90-day delinquency, respectively) and two different default horizons (36 months and 60 months relative to the month of origination).[27] The results show a negative, statistically significant relation between default risk and time to sale. The magnitude of the coefficient in the linear specification is approximately -0.01, which implies that a one-month increase in time to sale is associated with a 1-percentage-point decrease in the average default rate. The results appear to be consistent over the different horizons and default definitions.

In Table 6, we explore whether there is a non-linear relationship between time to sale and default. The table, which displays the results from the non-parametric specification described above, clearly shows that there is a decreasing negative relationship. Columns 1–2 ("Full Sample") suggest that the average default rates are decreasing in time to sale until Months to $\text{Sale}_{ij} = 5$, at which point the average default rates begin to rise moderately. Mortgages sold in the 5th month after origination have default rates that are approximately 6 percentage points lower than loans sold in either the month of origination or the month after origination, while mortgages sold in the 9th month after origination have default rates that

[26] Because we cannot distinguish between loans with values of 0 and 1 for months to sale, the omitted category for the regressions estimated on LPS data includes both.

[27] In Appendix Table 15, we display the coefficient estimates for all of the variables in our covariate set. Most of the estimates are consistent with the previous literature on mortgage default.

are lower by 3–4 percentage points on average. Again, the estimation results are consistent across the alternative default definitions and across different horizons (which we do not show for sake of brevity).

In Panel A of Figure 1, we plot the estimated relation between time to sale and *ex-post* PLS default risk from the non-parametric specification in column (1) of Table 6 (60+ DQ measure of default, 36-month horizon). The plot includes 95% confidence intervals. Overall, the results in Tables 5 and 6 provide evidence of a negative relation between time to sale and (conditional) *ex-post* default risk, which supports the existence of a signaling motive in the PLS market.

5.2 Accounting for "Mechanical" Effects of Random Delays

One potential concern in the default analysis above is the role of early payment defaults in generating a mechanical relation between time to sale and *ex-post* default risk due to institutional features of the PLS market. We discuss this possibility in Section 2.1 in the context of our simple model. In short, delinquent loans may be harder to sell into a securitized pool of loans, which may create a selection effect of loans sold late relative to loans sold immediately. This could create a negative relation between time to sale and default that is independent from a mechanism involving private information and signaling. Random delays would mean that loans sold quickly would be representative of the population of eligible loans, whereas loans that take longer to sell would be of higher average quality than the population of eligible loans.

To address this issue, we undo this potential selection effect by including loans in the analysis that do not default within the first 9 months of origination. Put differently, we drop all loans that are securitized between months 0 and 9 and become delinquent by month 9. We refer to this sample as the "restricted sample." This restriction forces the sample of sold loans to be homogeneous in terms of early payment defaults irrespective of whether they were sold into securitization pools, and the results *cannot* be explained by the selection effect described above.

While this correction directly addresses potential selection bias, it has several drawbacks. First, loans that default early may still be sold, in which case the selection effect is not severe, and the correction is simply wasting information. Second and more importantly, it may be that originators signaling behavior is precisely about the likelihood of early-payment default. That is, if most of the private information on loan quality concerns the likelihood of default within the first few years of origination, this "correction" would effectively eliminate the variation of greatest interest. For this reason, we choose to display the correction as a robustness check rather than to adopt it as our baseline specification.

Panel B of Table 5 and columns 3–4 in Table 6 display the same set of results for our restricted sample, where we exclude all loans that default within 9 months of origination. We find that the effects are virtually unchanged for the linear specification of the Months to Sale_{ij} variable, but there are a few subtle differences for the non-linear specifications. This sample restriction slightly mitigates the negative relation between time to sale and default for loans sold within 4 months, but it has the opposite effect for loans sold later. Overall, the sample correction appears to have a minor effect on the results, which suggests that sample selection bias is not an important issue.

5.3 Default and Time-To-Sale: Agency Loans

Table 7 displays the results for our sample of loans sold to the GSEs. We find essentially no evidence of a relation between time to sale and *ex-post* default risk in the GSE segment of the market. The estimates are all very close to zero and statistically insignificant. We plot the coefficients from a non-parametric specification in Panel B of Figure 1 (the same specification as the one used to construct the PLS graph in the top left panel). There is a stark difference in this pattern from the one displayed in the PLS graph. While there is a clear downward trend in the PLS estimates that flattens out toward the end of the time to sale distribution, the GSE coefficients are zero until the end of the distribution, when they begin to fall (although the sample size becomes significantly reduced in these later months).

In general, the GSE results are consistent with our hypothesis that private information does not play a significant role in the agency market compared to the PLS market. Specifically, the GSE market is dominated by automated underwriting systems in which the agencies pre-commit to funding loans so that originators have little incentive or need to signal loan quality through sales delays.

5.4 ***Ex-Ante*** Analysis

In this section, we estimate the empirical relation between time to sale and *ex-ante* credit risk; that is, credit risk measured at the time of issuance based on observable characteristics only. To do this, we construct *ex-ante* default probabilities for each loan using the data available in LPS and using the only performance information available at the time of origination in a manner similar to the method used in Ashcraft et al. (2010).

We focus on 36- and 60-month horizons for the default forecasts to maintain consistency with our results on *ex-post* default presented above. For each quarter in our sample, we take all loans that were originated between 48 months and 36 months prior and track those mortgages over the subsequent 36 months. We then estimate a discrete choice model (linear probability and logit, both shown in Table 8) using variables that are available in LPS to predict 36-month defaults for these loans. We use an analogous strategy for the 60-month horizon (i.e., we take all loans originated between 72 and 60 months before and track them over the subsequent 60 months). The regressions include the same set of covariates that are included in the *ex-post* default regressions above. We take the estimated coefficients from each quarterly credit risk model and apply them to the characteristics of the loans originated in the current quarter to create 36-month, loan-level, *predicted* default probabilities. This leaves us with an *ex-ante* credit risk measure that uses only information available at the time of issuance.

We take the *ex-ante* default probabilities and substitute them into equation 5 to estimate the relation between time to sale and *observable* default risk. Table 8 shows the results. We find positive coefficients for all models in Panel A (PLS loans), which is consistent with the intuition that observably better loans tend to transact *faster*. This pattern is in stark contrast to the estimated relation between *ex-post* default rates and time to sale in the PLS market in Tables 5 and 6. We observe no relation between *ex-ante* risk and time to sale for GSE loans (Panel B).

Figure 1 plots the non-parametric coefficients for using *ex-ante* risk as the outcome variable. Consistent with the regressions, the PLS figure (Panel C) shows that loans sold later are *riskier* based on the observable underwriting characteristics. Loans sold in the 2nd, 3rd, and 4th months after origination have expected default probabilities that are approximately 2–3 percentage points higher than loans sold in the month of origination or in the month immediately following origination. This difference moderates at the end of the time to sale distribution, with loans sold between 6 and 9 months having only slightly (about 1 percentage point) higher expected default probabilities, on average. The horizontal line displayed in the lower right panel in the figure implies that there is no relation between predictable default risk and time to sale in the GSE market.

5.5 Default and Time To Sale Using the CoreLogic Sample

Table 9 displays the results on the relation between *ex-post* default risk and time to sale using the sample of PLS loans in CoreLogic. One of the main reasons for using CL data is the availability of the identity of the mortgage originator and the security issuer, which allows us to account for any variation generated by heterogeneity across originators and issuers. In Table 9, we present the results corresponding to our linear specifications of equation 5 and focus on a default horizon of 36 months and a default definition based on 60+ days delinquent. Panel A focuses on the effect of controlling for originator heterogeneity, while Panel B focuses on issuer heterogeneity. Issuer information is obtained from Bloomberg; it corresponds to the private financial institution responsible for pooling and securitizing the mortgages. In each panel, we display the results for the full sample of PLS loans (columns (1)–(3)), the sample of Alt-A PLS loans (columns (4)–(6)), and the sample of subprime PLS loans (columns (7)–(9)).

In columns (1), (4), and (7) of Panel A, we display the results from a specification that does not control for originator heterogeneity; thus, these results are directly comparable to the LPS results displayed in Table 5. In columns (2), (5), and (8), we include a full set of originator fixed effects. Information on the originators is available for slightly more than half of the loans in the CL dataset, so we focus our analysis on this subsample.[28] Finally, columns (3), (6), and (9) display the results from a specification that includes originator-by-year-quarter-of-origination fixed effects. This is a fairly demanding test, as it uses variation on months to sale and default from loans originated by the same institution in the same year-quarter to estimate the relationship between time to sale and default.

Focusing on the full sample of PLS loans, our estimate for the relation between time to sale and default is negative and statistically significant but significantly lower than the corresponding estimate obtained using LPS data. The coefficient in column (1), which cor-

[28] We do this even for the specifications that do not include originator fixed effects to isolate the effect of originator heterogeneity from the effects of changing the size and scope of the sample.

responds exactly to our LPS specification (i.e., no originator fixed effects) is -0.36 percentage points, roughly one-third of the magnitude of the estimate in Table 5 (-1.1 percentage points). We return to this comparison below when we separate the loans into the Alt-A and subprime segments of the market. In column (2), we see that the inclusion of originator fixed effects slightly decreases (in absolute magnitude) the coefficient associated with months to sale (0.36 to 0.28 percentage points), while the inclusion of originator-by-year-quarter-of-origination fixed effects (column (3)) further decreases the coefficient (0.28 to 0.15 percentage points), although the estimate remains negative and statistically significant.

In column (1) of Panel B, we display the results from a specification that includes both originator and issuer fixed effects. Compared to the specification with only originator fixed effects (column (2) in Panel A), the estimated effect increases (in absolute magnitude) from -0.28 to -0.41 percentage points. In column (2), we add issuer-by-year-quarter-of-issuance fixed effects, which approximately halves the magnitude of the coefficient. Finally, column (3) displays the results from a specification that includes both originator-by-year-quarter-of-origination and issuer-by-year-quarter-of-issuance fixed effects. The estimated effect remains negative and statistically significant, as an additional month of delay is associated with a 0.17-percentage-point decrease in the likelihood of default, all else equal.

5.5.1 Alt-A PLS vs. Subprime PLS

In addition to the information on the identities of originators, an advantage of using CL data is the ability to analyze different segments of the PLS market. *A priori*, we may expect signaling unobservable mortgage quality to have a larger role in the Alt-A segment of the PLS market because it is largely comprised of low-documentation mortgages. Table 4 shows that over 70% of Alt-A mortgages were less than fully documented, compared to 35% of subprime loans. Industry sources suggest that at least some of the loans that appear as "fully documented" may also suffer from documentation issues that prevent them from being sold in the GSE (conforming) market.

Columns (4)–(9) in Table 9 display the linear specification results from separately estimating the regressions for the subprime and Alt-A segments of the PLS market, and Figure 2 plots the results for the non-parametric specifications. The differences between the subprime and Alt-A results are fairly striking, and they help to explain where the differences between the LPS and CL results are likely originating. There is essentially no relation between *ex-post* default risk and time to sale among subprime PLS loans, while there is a significant negative relation among Alt-A loans. The estimates from the Alt-A regression are monotonically decreasing in time to sale. A loan sold to an issuer of Alt-A PLS 9 months after origination is, on average, about 6 percentage points less likely to default compared to a loan sold immediately upon origination, which is very similar to the estimated magnitudes obtained in the LPS sample. As we discussed above, when we compare the summary statistics between LPS and CL (Tables 3 and 4), it appears that the LPS sample of PLS loans is more similar to the Alt-A mortgage sample than the subprime sample in CL. This could rationalize the differences in the quantitative magnitudes of the estimates coming from each sample, as the CL Alt-A magnitudes are quite similar to those obtained from LPS. In addition, the Alt-A results are not nearly as sensitive to the inclusion of originator-by-time and issuer-by-time fixed effects. In contrast, the results for the subprime sample are highly sensitive and largely disappear with the inclusion of these fixed effects.

In Appendix Table 16 we show the results when we restrict the sample of Alt-A and Subprime loans to those that have not become delinquent 9 months after origination. This is the same correction that we implemented and discussed in Section 5.2. We find that the coefficient on Alt-A loans is reduced by approximately 0.02 in magnitude relative to the results with the whole sample (for example, without originator or issuer fixed effects, the coefficient moves from 0.072 in Table 9 to 0.054. All specifications (with and without originator and issuer fixed effects) are still highly statistically significant. The results for subprime loans are similarly affected. The results for subprime loans were already much weaker than those for Alt-A, and they are now positive and significant at between 0.002 to 0.004.

5.5.2 Documentation Results

We further explore the role of documentation standards by stratifying our PLS sample into loans with full documentation of income and assets and loans with less than full documentation ("low doc"). We stratify by documentation type for the full sample of PLS loans and for our separate subprime and Alt-A samples. The results are displayed in Table 10, with Panel A containing the results for the linear specifications and Panel B containing the results for the non-parametric specifications.

The results are mixed. In the sample of all PLS loans (subprime and Alt-A combined), there appears to be a stronger negative relation between time to sale and default for low-documentation loans compared to full-documentation loans. This negative relation is approximately 50% larger (in absolute value) in the sample of low-documentation PLS loans (columns 1–2). However, the results in columns 3–6 (breaking down loans into Alt-A and subprime) show that there are essentially no differences between full-documentation and low-documentation loans within each of the two subsamples.

5.5.3 Affiliation Results

In this section, we test whether an affiliation between the originator (seller) and the issuer (buyer) plays a role in the relation between time to sale and default risk. Many issuers and originators in the PLS market share direct relationships. In some cases, the originator and issuer are the same institution, while in others, they are part of the same vertically integrated corporation (in which case, the originator is typically a subsidiary of the issuer). *A priori*, we might expect that the scope for private information between an originator and issuer who are affiliated is less than that between an originator and issuer that are independent entities.[29] Thus, if signaling is driving our results, we expect a weaker negative relation between time to sale and default risk for the sample of loans in which the issuer and originator are affiliated with each other.

We obtained information on the identity of the issuer from Bloomberg and supplemented the Bloomberg data with hand-collected data from the pooling and service agreements (PSA) associated with the PLS deals.[30] We focus only on loans that are in deals in which either all of the loans were made by affiliated originators or all of the loans were made by unaffiliated originators.[31] Table 11 and Figure 3 display the results. As in our analysis of documentation status above, we stratify our sample of all PLS loans and our separate Alt-A and subprime samples by affiliation status.

The results are different for the three samples. The negative correlation between time to

[29]This argument is also made by Demiroglu and James (2012b) and Furfine (2014).

[30]We pulled the PSAs from the SEC's EDGAR website: http://www.sec.gov/edgar/searchedgar/companysearch.html.

[31]We decided to drop the "mixed" deals that included loans made by both affiliated and unaffiliated originators because of our lack of confidence in the identity of the originator and/or our ability to identify a relationship between the issuer and the originator (the raw data on originator identities in the CoreLogic database is somewhat messy, so we were forced to expend significant effort to clean and standardize the names to integrate the information into our empirical analysis).

sale and default risk does not appear to be sensitive to affiliation status in the full sample and the subprime sample (columns 1–2 in Table 11). However, the difference in the relation between time to sale and *ex-post* default risk for unaffiliated compared to affiliated issuers and originators is especially stark in the Alt-A segment of the market. Loans sold six months after origination by affiliated originators are approximately 3 percentage points less likely to default compared to loans sold in the month of origination (column 3 of Panel B in Table 11), while this effect increases to almost 9 percentage points for loans originated by unaffiliated originators. Panel A in Figure 3 shows that this difference is highly statistically significant over the entire distribution of time to sale.

Two issues related to the affiliation results require further discussion. First, there is some ambiguity regarding the exact place in the mortgage funding chain in which asymmetric information might play an important role. One possibility is that it occurs between the originator and the issuer, while a second possibility is that it occurs between the issuer and the ultimate PLS investors. The affiliation results shed light on this issue because the two possibilities each yield different predictions about the effect of the originator–issuer affiliation on the magnitude of the correlation between time to sale and *ex-post* default risk. As we argued above, if asymmetric information is present between the originator and issuer, we would expect to find a weaker relationship between time to sale and default for affiliated institutions. However, if asymmetric information is present between the issuer and PLS investors, then we might expect to find a stronger relationship for affiliated originators and issuers because investors may perceive that issuers are more likely to obtain private information on mortgage quality when they have an affiliation with the originators. Thus, our finding of a weaker relationship shown in Table 11 suggests that asymmetric information between originators and issuers plays a more important role.

Second, there is some uncertainty about whether the originator field in the CoreLogic database actually corresponds to the lender of record (i.e., the institution that underwrote and originated the loan) or to what is sometimes referred to as the "aggregator" or "seller," which is the institution responsible for purchasing loans from various lenders to fill the PLS mortgage pools and then selling those loans to the issuer (Stanton et al. (2014)). This is a potentially important distinction because it may be more likely that private information is obtained by the lender of record because it has more interaction with the mortgage borrower.

To verify that the originator field in CoreLogic corresponds to the lender of record, we match our CoreLogic mortgage data to a database of public mortgage filings that contains the identity of the lender of record. This database contains the universe of all residential mortgages in the state of Massachusetts during our sample period and comes from county deed registries that record information on property transactions. We compare the lender of record with the originator listed in the CoreLogic database for the sample of matched Massachusetts mortgages. In unreported tables, we find that for 83% of the matched sample, the lender of record matched the CoreLogic originator field. The remaining 17% are either cases in which CoreLogic is reporting an entity other than the lender of record (most likely the aggregator) or cases that are bad matches (there is the potential for significant matching errors because we are not able to perform a precise match using loan account numbers or social security numbers). Thus, we view the 17% figure as an upper bound on the severity of the potential issue of misidentifying the true originator in the CoreLogic data.

5.6 Security Spreads and Time To Sale

We now present evidence on the empirical relation between time to sale and security prices. The unit of observation for this analysis is a pool, a group of loans from which different triple-A securities in each PLS deal derive cash flows. Junior securities (those below triple-A) generally derive cash flows from all of the pools in a deal. If deals have only one pool of mortgages, the average spread corresponds to the weighted average spread of the triple-A securities in the deal.

Table 12 displays the results from regressing the average pool-level spreads on average pool-level seasoning. Panel A shows the results when we include only a linear term for average seasoning, while Panel B includes a quadratic term to capture potential non-linearities.[32] The results on *ex-post* default rates discussed above were significantly different in the sample of mortgages that collateralized Alt-A securities than in the sample of loans that backed subprime securities. Thus, in both panels, we show the results for the full sample of floating-rate triple-A securities (columns 1–3) and the results for Alt-A (columns 4–6) and subprime (columns 7–9) securities separately to determine whether similar patterns emerge on the pricing dimension.

In Table 12, we display the results for three different regression specifications. The first

[32] Because average seasoning at the pool level is a continuous variable, it is not possible to use the same non-parametric specification used in the loan-level default analysis above.

specification includes the quarter of issuance fixed effects but no other control variables. The second specification includes the list of pool-level controls displayed in Table 14 along with the quarter of issuance fixed effects. The third specification includes a full set of issuer fixed effects in addition to the pool-level controls and month of issuance fixed effects,.

Column (1) in Panel A shows that one additional month of average mortgage seasoning is associated with a 1.5-basis-point lower yield spread, which is about 5% of the average spread in the sample (28 basis points). When pool-level controls and both the issuer and the month of issuance fixed effects are included (column (3)), the coefficient estimate declines slightly but remains statistically significant. Similar to our findings in the default analysis above, we see in columns (4)–(9) that this effect is concentrated in the Alt-A sample. For Alt-A securities, one additional month of average mortgage seasoning is associated with a 2.4-basis-point lower yield spread.

For the non-linear specification results reported in Panel B, both the linear and the quadratic terms are significant in the full sample and the Alt-A sample. The linear terms are negative and the quadratic terms are positive, which implies a non-linear relation between time to sale and security spreads similar to the relation that we documented above between time to sale and mortgage default. Figure 5 displays the predicted security spreads as a function of the average pool-level seasoning calculated using the estimation results from the specification reported in column (6) in Panel B. The figure includes 95% confidence intervals calculated using the delta method. The plot reveals several notable takeaways. First, the minimum spread as a function of average seasoning is achieved between four and five months. Second, after five months, the spread begins to increase in seasoning; however, the confidence bands show that we begin to lose precision for seasoning greater than five months because there are so few securities in the dataset with high values of average seasoning (Figure 4).

5.7 Early Prepayment Analysis

While default is clearly undesirable from the perspective of an MBS investor, the risk of early prepayment is another potentially negative outcome for mortgage investors. Residential mortgages contain a prepayment option that allows the borrower to fully repay the outstanding principal balance of the loan before it reaches full maturity. Early prepayment risk was an important consideration for investors in the period before the housing bust and financial crisis, especially given the low levels of default rates that prevailed during that period.

It is well known in the mortgage literature that interest rate movements largely drive the prepayment behavior of borrowers with fixed-rate mortgages. In contrast, the prepayment of adjustable-rate mortgages is typically driven by life events that are unrelated to interest rate movements, such as new housing purchases driven by employment changes or changes in household size due to the birth of a child or death of a family member. In the PLS market, however, in addition to responses to life events, prepayments of adjustable-rate mortgages were often driven by specific contractual features. In particular, the prepayment behavior of 2/28 and 3/27 hybrid ARMS, the most common types of PLS ARMs (accounting for about 75% of the market), was highly correlated with the duration of the period in which the interest rate was frozen: two years for the 2/28s and 3 years for the 3/27s. After the initial period, the interest rate would reset to a new level and track a market interest rate (such as the 6-month LIBOR or the 10-year Treasury rate). Because the interest rate typically reset to a higher level, many borrowers prepaid either right at or shortly after the reset period. In addition, many ARMs in the PLS market contained prepayment penalties that expired at the same time of the interest rate reset, providing further incentive for borrowers to wait until the reset date to exercise their prepayment options.[33]

We focus on the sample of 2/28 and 3/27 ARMs that did not default and define a negative outcome to be an ARM that was prepaid several months before the interest rate reset month.[34] We consider two cutoffs, six and nine months before the reset date, in defining our early prepayment indicator variables, as the most common type of prepayment penalty associated with these mortgages was six months of interest on 80% of the principal amount prepaid. An ARM that carried this prepayment penalty and prepaid more than six months before the reset date would generate lower cash flows for investors than a loan that prepaid at the reset date, and prepayment can thus be considered as a negative outcome for a PLS investor.

Table 13 contains the results of the early prepayment analysis. Panel A displays the results for parametric (quadratic) specifications, while Panel B displays the results for the

[33]For an excellent reference on the PLS market in general and for empirical analyses on the prepayment and default behavior of various types of PLS loans in particular, see Kramer and Sinha (2006). See Sengupta (2010) for a detailed discussion of the composition of loans in the Alt-A and subprime PLS markets.

[34]We eliminate defaults from our analysis to isolate voluntary prepayment risk.

non-parametric specifications. We show the results for various corrections for the potential "mechanical" selection issue discussed in Section 5.2 above. Specifically, we exclude from the sample loans that prepay within three, six, and nine months from origination. Just as in the case of default, however, this may be an "over-correction" to the extent that investors may be especially concerned with prepayments within the first year or so after origination, and such a restriction could eliminate the true signaling effect rather than simply correct the sample selection bias.

Table 13 clearly shows a negative relation between time to sale and early prepayment risk. As months to sale increase, the likelihood of early prepayment decreases in a relatively monotonic manner. Focusing on the first two columns in the table (no correction), PLS loans sold six months after origination are approximately 6–7% less likely to prepay early compared to loans sold immediately, while loans sold nine months after origination are about 10–11% less likely to prepay early. The negative relation remains significant when we exclude prepayments that happen in the first few months after origination, but the non-parametric specification shows that the relation flattens for five months in columns (5) through (8).

In general, the results on the correlation between time to sale and early prepayment are consistent with the default analysis and support the mechanism of using sales delays to signal quality. While PLS investors were likely concerned about significant credit risk in the case of a large downturn (which, of course, occurred), prepayment risk is present in both good and bad economic conditions, and it was thus an important consideration for mortgage investors. In addition, while our results suggest that asymmetric information on default risk did not play an important role in the subprime PLS market, these results indicate that asymmetric information on prepayment risk may have been important, as the vast majority of 2/28 and 3/27 hybrid-ARMs were placed in subprime securities.[35] These findings are consistent with Agarwal and Yavas (2014), who find evidence of adverse selection with respect to prepayment risk but not default risk in the PLS market.

6 Conclusion

A general feature of models of asymmetric information and delayed trade is that the prices and (unobserved) quality of goods increase over time. This paper provides some of the first

[35] In our CL sample, approximately 96% of 2/28s and 79% of 3/27s were in subprime securities.

empirical evidence of this prediction in the context of the residential mortgage market. Using detailed loan-level data on privately securitized mortgages, we find a statistically significant and economically meaningful positive correlation between conditional *ex-post* mortgage performance and time to sale. This finding is robust to different ways of measuring performance and, importantly, is not generated by the component of mortgage performance that is predicable by buyers using *ex-ante* observable information on underwriting characteristics. Furthermore, the positive relation between time to sale and mortgage performance is not present in the agency securitization market, in which adverse selection between originators and issuers is a less serious concern. This estimated correlation appears to be strongest in the Alt-A segment of the PLS market, in which most loans were underwritten with less than full documentation of income and/or assets, which is consistent with previous studies that have found an important role of private information among low-documentation mortgages.

Taken together, the results both confirm the importance of private information in the non-agency securitization market and provide evidence consistent with a signaling mechanism by which lenders in the market are able to reveal the quality of their loans by delaying trades. Although our findings indicate that the asymmetric information problem is large in the market for secondary mortgage market, future work should attempt to quantify the welfare loss that entailed in signaling.

References

Adelino, M., W. S. Frame, and K. Gerardi (2016). The effect of large investors on asset quality: Evidence from subprime mortgage securities. Working papaer, Federal Reserve Bank of Atlanta.

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

Agarwal, Sumit, Y. C. and A. Yavas (2014). Adverse selection in mortgage securitization. *Journal of Financial Economics 105*(3), 640–660.

Aiello, D. (2016). Information exploitation? a pre-crisis rmbs issuer's private information. Working paper, University of California Los Angeles.

Akerlof, G. A. (1970). The market for lemons: Quality uncertainty and the market mechanism. *The Quarterly Journal of Economics*, 488–500.

An, X., Y. Deng, and S. A. Gabriel (2011). Asymmetric information, adverse selection, and the pricing of cmbs. *Journal of Financial Economics 100*(2), 304–325.

Ashcraft, A., P. Goldsmith-Pinkham, P. Hull, and J. Vickery (2011). Credit ratings and security prices in the subprime mbs market. *The American Economic Review 101*(3), 115–119.

Ashcraft, A. B., P. Goldsmith-Pinkham, and J. I. Vickery (2010). Mbs ratings and the mortgage credit boom. *FRB of New York Staff Report* (449).

Ausubel, L. M. and R. J. Deneckere (1989). Reputation in bargaining and durable goods monopoly. *Econometrica*, 511–531.

Begley, T. A. and A. Purnanandam (2017). Design of financial securities: Empirical evidence from private-label rmbs deals. *Review of Financial Studies 30*(1), 120–161.

Chang, B. (2014). Adverse selection and liquidity distortion. Working paper, University of Wisconsin.

Cho, I. and D. Kreps (1987). Signaling games and stable equilibria. *The Quarterly Journal of Economics*, 179–221.

Cramton, P. C. (1984). Bargaining with incomplete information: An infinite-horizon model with two-sided uncertainty. *The Review of Economic Studies 51*(4), 579–593.

Daley, B. and B. Green (2012). Waiting for news in the market for lemons. *Econometrica 80*(4), 1433–1504.

Daley, B. and B. Green (2016). An information-based theory of time-varying liquidity. *Journal of Finance 71*(2), 809–870.

DeMarzo, P. M. (2005). The pooling and tranching of securities: A model of informed intermediation. *Review of Financial Studies 18*(1), 1–35.

DeMarzo, P. M. and D. Duffie (1999). A liquidity-based model of security design. *Econometrica 67*(1), 65–100.

Demiroglu, C. and C. James (2012a). How important is having skin in the game? originator-sponsor affiliation and losses on mortgage-backed securities. *Review of Financial Studies 25*(11), 3217–3258.

Demiroglu, C. and C. James (2012b). How important is having skin in the game? originator-sponsor affiliation and losses on mortgage-backed securities. *Review of Financial Studies 25*, 3217–3258.

Downing, C., D. Jaffee, and N. Wallace (2009). Is the market for mortgage-backed securities a market for lemons? *Review of Financial Studies 22*(7), 2457–2494.

Evans, R. (1989). Sequential bargaining with correlated values. *The Review of Economic Studies 56*, 499–510.

Fuchs, W., B. S. Green, and D. Papanikolaou (2015). Adverse selection, slow moving capital and misallocation. *Journal of Financial Economics*.

Fuchs, W. and A. Skrzypacz (2013). Costs and benefits of dynamic trading in a lemons market. *Working Paper*.

Fudenberg, D. and J. Tirole (1983). Sequential bargaining with incomplete information. *The Review of Economic Studies*, 221–247.

Furfine, C. H. (2014). Complexity and loan performance: Evidence from the securitization of commercial mortgages. *Review of Corporate Finance Studies 2*, 154–187.

Ganduri, R. (2016). Repo regret? Working paper, Emory University.

Garmaise, M. J. (2015). Borrower misreporting and loan performance. *The Journal of Finance 70*(1), 449–484.

Garmaise, M. J. and T. J. Moskowitz (2004). Confronting information asymmetries: Evidence from real estate markets. *Review of Financial Studies 17*(2), 405–437.

Genesove, D. (1993). Adverse selection in the wholesale used car market. *Journal of Political Economy*, 644–665.

Grenadier, S. R. and A. Malenko (2011). Real options signaling games with applications to corporate finance. *Review of Financial Studies 24*(12), 3993–4036.

Grenadier, S. R. and N. Wang (2005). Investment timing, agency, and information. *Journal of Financial Economics 75*(3), 493–533.

Griffin, J. M. and G. Maturana (2016). Who facilitated misreporting in securitized loans? *The Review of Fianncial Studies 29*(2), 384–419.

Guerrieri, V., R. Shimer, and R. Wright (2010). Adverse selection in competitive search equilibrium. *Econometrica 78*(6), 1823–1862.

Gul, F. and H. Sonnenschein (1988). On delay in bargaining with one-sided uncertainty. *Econometrica*, 601–611.

Gul, F., H. Sonnenschein, and R. Wilson (1986). Foundations of dynamic monopoly and the coase conjecture. *Journal of Economic Theory 39*(1), 155–190.

He, J. J., J. Q. Qian, and P. E. Strahan (2012). Are all ratings created equal? the impact of issuer size on the pricing of mortgage-backed securities. *The Journal of Finance 67*(6), 2097–2137.

Janssen, M. C. and S. Roy (2002). Dynamic trading in a durable good market with asymmetric information. *International Economic Review 43*, 257–282.

Jiang, W., A. A. Nelson, and E. Vytlacil (2014a). Liar's loan? effects of origination channel and information falsification on mortgage delinquency. *Review of Economics and Statistics 96*(1), 1–18.

Jiang, W., A. A. Nelson, and E. Vytlacil (2014b). Securitization and loan performance: A contrast of ex ante and ex post relations in the mortgage market. *Review of Financial Studies 27*(2), 454–483.

Kramer, B. and G. Sinha (2006). Bear stearns quick guide to non-agency mortgage-backed securities. Technical report, Bear Stearns.

Kremer, I. and A. Skrzypacz (2007). Dynamic signaling and market breakdown. *Journal of Economic Theory 133*(1), 58–82.

Leland, H. E. and D. H. Pyle (1977). Informational asymmetries, financial structure, and financial intermediation. *The Journal of Finance 32*(2).

Noldeke, G. and E. Van Damme (1990). Signalling in a dynamic labour market. *The Review of Economic Studies 57*(1), 1–23.

Piskorski, T., A. Seru, and J. Witkin (2015). Asset quality misrepresentation by financial intermediaries: evidence from the rmbs market. *The Journal of Finance 70*(6), 2635–2678.

Saengchote, K. (2013). Soft information in the subprime mortgage market. In *26th Australasian Finance and Banking Conference*.

Sengupta, R. (2010). Alt-a: The forgotten segment of the mortgage market. *Federal Reserve Bank of St. Louis Review 92*(1), 55–71.

Sobel, J. and I. Takahashi (1983). A multistage model of bargaining. *The Review of Economic Studies 50*(3), 411–426.

Spence, M. (1973). Job market signaling. *The Quarterly Journal of Economics*, 355–374.

Stanton, R., J. Walden, and N. Wallace (2014). The industrial organization of the us residential mortgage market. *Annu. Rev. Financ. Econ. 6*(1), 259–288.

Stanton, Richard, W. J. and N. Wallace (2015). Mortgage loan-flow networks and endogenous financial norms.

Stearns, B. (2006). Bear stearns quick guide to non-agency mortgage-backed securities. *Residential Mortgage Criteria Report*.

Swinkels, J. M. (1999). Education signalling with preemptive offers. *The Review of Economic Studies 66*(4), 949–970.

Vincent, D. R. (1989). Bargaining with common values. *Journal of Economic Theory 48*, 47–62.

Vincent, D. R. (1990). Dynamic auctions. *The Review of Economic Studies 57*, 49–61.

Williams, B. (2016). Search, liquidity, and retention: Signaling multidimensional private information. Technical report, Working Paper, NYU.

Table 1: Distribution of Time to Sale in the LPS Sample

	PLS Loans		GSE Loans	
Months to Sale	# Loans	Cumulative %	# Loans	Cumulative %
0/1	3,104,102	56.55	6,999,528	60.72
2	1,261,872	79.54	3,700,677	92.83
3	518,156	88.98	471,520	96.92
4	191,413	92.47	128,404	98.04
5	84,131	94.00	58,619	98.55
6	56,610	95.03	29,598	98.80
7	41,849	95.79	18,733	98.96
8	30,881	96.36	16,243	99.11
9	24,969	96.81	14,203	99.23
≥ 10	174,972	100.00	88,931	100.00

Notes: This table displays the distribution of the number of months between the time of origination and the time of sale (months to sale) for privately securitized mortgages in the LPS dataset. The LPS sample includes only first-lien mortgages originated between January 2002 and December 2007. The sample is further restricted to only mortgages seasoned less than two months (i.e., loans that entered the dataset in either the month of origination or the month following origination).

Table 2: Distribution of Time to Sale in the CoreLogic PLS Sample

	All PLS		Subprime		Alt-A	
Months to Sale	# Loans	Cum. % of Sample	# Loans	Cum. % of Sample	# Loans	Cum. % of Sample
0	2,446,106	17.9	1,079,646	12.4	1,366,460	27.7
1	3,675,646	44.8	2,296,307	38.7	1,379,339	55.6
2	2,952,576	66.4	2,026,277	62.0	926,299	74.3
3	2,064,585	81.6	1,521,350	79.4	543,235	85.3
4	1,149,410	90.0	861,916	89.3	287,494	91.1
5	571,103	94.2	415,989	94.1	155,114	94.3
6	286,959	96.3	201,827	96.4	85,132	96.0
7	140,231	97.3	86,683	97.4	53,548	97.1
8	87,131	97.9	51,849	98.0	35,282	97.8
9	56,839	98.3	32,197	98.4	24,642	98.3
≥ 10	228,536	100.0	85,146	100.0	143,390	100.0

Notes: This table displays the distribution of the number of months between the time of origination and the time of sale (months to sale) for privately securitized mortgages in the CoreLogic dataset. The CoreLogic sample includes only first-lien mortgages backing subprime and Alt-A PLS that were originated between January 2002 and December 2007. The time of sale corresponds to the month in which the PLS security was issued.

Table 3: Summary Statistics: LPS Sample

	PLS Loans		**GSE Loans**	
	Mean	SD	Mean	SD
Loan/Borrower Characteristics (continuous variables)				
Term	354	49	333	63
Original Rate	5.96	1.97	6.17	0.77
Original Amount	299,218	204,952	176,680	90,235
LTV Ratio	73.1	15.0	74.0	18.3
FICO	700	68	713	63
Loan/Borrower Characteristics (dummy variables)				
	Mean		Mean	
Purchase (d)	0.440		0.432	
Cash Out Refinance (d)	0.208		0.140	
Arm (d)	0.519		0.127	
Balloon (d)	0.008		0.003	
Interest Only (d)	0.234		0.064	
"B" or "C" Grade (d)	0.178		0.012	
Jumbo (d)	0.296		0.005	
Low Doc. (d)	0.146		0.131	
Prepay Penalty (d)	0.279		0.012	
Primary Residence (d)	0.868		0.876	
Single Family (d)	0.822		0.847	
Geographic Characteristics				
	Mean	SD	Mean	SD
Unemployment rate (county-level)	4.8	1.4	4.9	1.5
36-month unemployment growth (%)	(4.4)	36.8	4.5	42.1
Price Index (county-level)	188	53	163	46
36-month HPA (%)	43.9	26.5	31.4	23.1
Default Rates				
	Mean		Mean	
60+ DQ, 36-month horizon	0.160		0.090	
60+ DQ, 60-month horizon	0.225		0.133	
90+ DQ, 36-month horizon	0.136		0.071	
90+ DQ, 60-month horizon	0.204		0.111	
# Loans	5,313,983		11,437,525	

Notes: This table displays the summary statistics for both privately securitized mortgages (PLS) and mortgages acquired by the housing GSEs (Fannie Mae and Freddie Mac) in the LPS dataset. The LPS sample includes only first-lien mortgages originated between January 2002 and December 2007. The sample is further restricted to only mortgages seasoned less than two months (i.e., loans that entered the dataset in either the month of origination or the month following origination). In addition, the sample only includes loans that were sold to either PLS issuers or the GSEs within nine months of origination (inclusive). All of the variables in the table are included in the set of model covariates. For a full list of covariates, see the Online Appendix.

Table 4: Summary Statistics: CoreLogic Sample

	All PLS		**Subprime**		**Alt-A**	
	Mean	SD	Mean	SD	Mean	SD
Loan/Borrower Characteristics (continuous variables)						
Term	356	37	355	33	358	47
Original Rate	7.48	1.57	7.87	1.33	6.26	1.62
Original Amount	214,855	150,813	190,628	125,503	291,003	192,566
LTV Ratio	83.0	14.3	83.7	14.0	80.8	15.1
FICO	639	70	617	61	710	48
Loan/Borrower Characteristics (dummy variables)						
	Mean		Mean		Mean	
Purchase (d)	0.395		0.363		0.495	
Cash Out Refinance (d)	0.500		0.552		0.339	
Arm (d)	0.741		0.763		0.669	
Balloon (d)	0.070		0.090		0.009	
Interest Only (d)	0.184		0.117		0.391	
Jumbo (d)	0.129		0.089		0.257	
Low Doc. (d)	0.442		0.351		0.728	
Prepay Penalty (d)	0.661		0.745		0.394	
Primary Residence (d)	0.870		0.919		0.716	
Single Family (d)	0.743		0.782		0.622	
Geographic Characteristics						
	Mean	SD	Mean	SD	Mean	SD
Unemployment rate (county-level)	5.23	1.58	5.34	1.59	4.88	1.49
36-month unemployment growth (\%)	5.4%	39.3%	8.8%	40.3%	-5.2%	33.7%
36-month HPA (\%)	42.3%	26.3%	40.7%	26.1%	47.1%	26.3%
Default Rates						
	Mean		Mean		Mean	
60+ Days Delinquent, 36-month horizon	0.304		0.333		0.215	
60+ Days Delinquent, 60-month horizon	0.372		0.390		0.318	
90+ Days Delinquent, 36-month horizon	0.251		0.272		0.186	
90+ Days Delinquent, 60-month horizon	0.327		0.339		0.291	
\# Loans	7,868,492		5,969,285		1,899,207	

Notes: This table displays the summary statistics for loans backing subprime and Alt-A PLS in the CoreLogic dataset. The CoreLogic sample includes only first-lien mortgages originated between January 2002 and December 2007. In addition, the sample only includes loans that were sold to PLS issuers within nine months of origination (inclusive). All of the variables in the table are included in the set of model covariates.

Table 5: Baseline Parametric Results for the Sample of PLS Loans in LPS

Panel A: Full Sample

Default Horizon:	36 Months		60 Months	
Default Definition:	60+ DQ	90+ DQ	60+ DQ	90+ DQ
Months to Sale	-0.0107	-0.0110	-0.0112	-0.0122
	(5.79)	(5.88)	(6.75)	(7.23)
# Loans	5,313,951	5,313,951	5,313,951	5,313,951
Adjusted R^2	0.23	0.22	0.25	0.25
Orig Qtr FEs?	Y	Y	Y	Y
State FEs?	Y	Y	Y	Y
Sale Qtr FEs?	Y	Y	Y	Y
Other Controls?	Y	Y	Y	Y

Panel B: Restricted Sample

Default Horizon:	36 Months		60 Months	
Default Definition:	60+ DQ	90+ DQ	60+ DQ	90+ DQ
Months to Sale	-0.0105	-0.0101	-0.0112	-0.0115
	(5.90)	(5.99)	(6.39)	(6.83)
# Loans	5,143,409	5,143,409	5,143,409	5,143,409
Adjusted R^2	0.20	0.19	0.23	0.22
Orig Qtr FEs?	Y	Y	Y	Y
State FEs?	Y	Y	Y	Y
Sale Qtr FEs?	Y	Y	Y	Y
Other Controls?	Y	Y	Y	Y

This table displays the results from the estimation of equation 5 on PLS loans in the LPS dataset originated in the 2002–2007 period. The dependent variable is an indicator variable for loans that default over a 36-month horizon (columns 1 and 2) and over a 60-month horizon (columns 3 and 4). Default is defined as a loan that is 60+ days delinquent (columns 1 and 3) and 90+ days delinquent (columns 2 and 4). Months to sale is defined as the number of months that elapse between origination and sale to a PLS issuer. All of the regressions include origination year-quarter fixed effects, state fixed effects, year-quarter of sale fixed effects, and the detailed list of covariates described in the text. The first row for each variable shows the regression coefficient and the second row shows the t-statistic. The standard errors are heteroskedasticity-robust and are clustered by year-quarter of origination. The restricted sample only includes loans that do not default within 10 months of origination.

Table 6: Baseline Non-Parametric Results for the Sample of PLS Loans in LPS

Default Horizon:	36 Months			
	Full Sample		**Restricted Sample**	
Default Definition:	60+ DQ	90+ DQ	60+ DQ	90+ DQ
Months to Sale = 2	-0.019	-0.019	-0.012	-0.012
	(5.12)	(5.01)	(3.82)	(3.82)
Months to Sale = 3	-0.038	-0.038	-0.026	-0.026
	(6.37)	(6.57)	(5.24)	(5.24)
Months to Sale = 4	-0.057	-0.058	-0.046	-0.046
	(7.91)	(7.93)	(6.74)	(6.74)
Months to Sale = 5	-0.058	-0.059	-0.052	-0.052
	(4.71)	(4.81)	(4.86)	(4.86)
Months to Sale = 6	-0.054	-0.054	-0.054	-0.054
	(3.56)	(3.61)	(4.49)	(4.49)
Months to Sale = 7	-0.044	-0.046	-0.049	-0.049
	(3.51)	(3.76)	(5.28)	(5.28)
Months to Sale = 8	-0.031	-0.034	-0.045	-0.045
	(2.03)	(2.38)	(3.48)	(3.48)
Months to Sale = 9	-0.036	-0.040	-0.049	-0.049
	(2.11)	(2.39)	(3.42)	(3.42)
# Loans	5,313,951	5,313,951	5,143,409	5,143,409
Adjusted R^2	0.23	0.22	0.19	0.19
Orig Qtr FEs?	Y	Y	Y	Y
State FEs?	Y	Y	Y	Y
Sale Qtr FEs?	Y	Y	Y	Y
Other Controls?	Y	Y	Y	Y

This table displays the results from the estimation of equation 5 on PLS loans in the LPS dataset originated in the 2002–2007 period. The dependent variable is an indicator variable for loans that default over a 36-month horizon. Default is defined as a loan that is 60+ days delinquent (columns 1 and 3) and 90+ days delinquent (columns 2 and 4). Months to sale is defined as the number of months that elapse between origination and sale to a PLS issuer. All of the regressions include origination year-quarter fixed effects, state fixed effects, year-quarter of sale fixed effects, and the detailed list of covariates described in the text. The first row for each variable shows the regression coefficient and the second row shows the t-statistic. The standard errors are heteroskedasticity-robust and are clustered by year-quarter of origination. The restricted sample only includes loans that do not default within 10 months of origination.

Table 7: Baseline Results for the Sample of GSE Loans in LPS

Panel A: Linear Specification

Default Horizon:	36 Months		60 Months	
Default Definition:	60+ DQ	90+ DQ	60+ DQ	90+ DQ
Months to Sale	-0.0001	-0.0010	-0.0002	-0.0012
	(0.11)	(1.60)	(0.19)	(1.72)
Loans	11,437,522	11,437,522	11,437,522	11,437,522
Adjusted R^2	0.14	0.14	0.16	0.16
Orig Qtr FEs?	Y	Y	Y	Y
State FEs?	Y	Y	Y	Y
Sale Qtr FEs?	Y	Y	Y	Y
Other Controls?	Y	Y	Y	Y

Panel B: Non-Parametric Specification

Default Horizon:	36 Months		60 Months	
Default Definition:	60+ DQ	90+ DQ	60+ DQ	90+ DQ
Months to Sale = 2	-0.006	-0.005	-0.007	-0.007
	(3.17)	(3.08)	(3.54)	(3.65)
Months to Sale = 3	0.007	0.005	0.006	0.004
	(3.70)	(3.23)	(2.87)	(2.38)
Months to Sale = 4	0.023	0.007	0.024	0.008
	(6.57)	(2.91)	(5.81)	(2.51)
Months to Sale = 5	0.012	0.006	0.012	0.007
	(2.73)	(1.78)	(2.55)	(1.68)
Months to Sale = 6	0.005	0.000	0.009	0.003
	(1.03)	(0.04)	(1.57)	(0.68)
Months to Sale = 7	0.000	-0.003	0.003	0.000
	(0.06)	(0.76)	(0.58)	(0.04)
Months to Sale = 8	-0.016	-0.015	-0.007	-0.008
	(3.39)	(3.37)	(1.40)	(1.71)
Months to Sale = 9	-0.030	-0.030	-0.022	-0.024
	(2.85)	(3.03)	(2.11)	(2.32)
Loans	11,437,522	11,437,522	11,437,522	11,437,522
Adjusted R^2	0.14	0.14	0.16	0.16
Orig Qtr FEs?	Y	Y	Y	Y
State FEs?	Y	Y	Y	Y
Sale Qtr FEs?	Y	Y	Y	Y
Other Controls?	Y	Y	Y	Y

This table displays the results from the estimation of equation 5 on GSE loans in the LPS dataset originated in the 2002–2007 period. The dependent variable is an indicator variable for loans that default over a 36-month horizon (columns 1-2) and over a 60-month horizon (columns 3-4). Default is defined as a loan that is 60+ days delinquent (columns 1 and 3) and 90+ days delinquent (columns 2 and 4). Months to sale is defined as the number of months that elapse between origination and sale to a GSE. All of the regressions include origination year-quarter fixed effects, state fixed effects, year-quarter of sale fixed effects, and a detailed list of the covariates described in the text. The first row for each variable shows the regression coefficient and the second row shows the t-statistic. The standard errors are heteroskedasticity-robust and are clustered by year-quarter of origination.

Table 8: Ex-Ante Default Risk Results

Panel A: PLS Loans

Model:	Linear Probability				Logit			
Default Horizon:	36 Months		60 Months		36 Months		60 Months	
Default Definition:	60+ DQ	90+ DQ	60+ DQ	90+ DQ	60+ DQ	90+ DQ	60+ DQ	90+ DQ
Months to Sale	0.0058	0.0045	0.0057	0.0040	0.0047	0.0015	0.0031	0.0010
	(5.20)	(5.20)	(4.40)	(3.45)	(2.63)	(0.99)	(4.35)	(1.34)
# Loans	3,672,426	3,672,426	3,672,426	3,672,426	3,660,474	3,660,474	3,613,121	3,613,121
Adjusted $R\hat{2}$	0.26	0.24	0.30	0.36	0.41	0.42	0.59	0.67
Orig Qtr FEs?	Y	Y	Y	Y	Y	Y	Y	Y
Sale Qtr FEs?	Y	Y	Y	Y	Y	Y	Y	Y

Panel B: GSE Loans

Model:	Linear Probability				Logit			
Default Horizon:	36 Months		60 Months		36 Months		60 Months	
Default Definition:	60+ DQ	90+ DQ	60+ DQ	90+ DQ	60+ DQ	90+ DQ	60+ DQ	90+ DQ
Months to Sale	0.0004	0.0002	0.0021	0.0013	0.0001	-0.0004	0.0007	0.0002
	(1.17)	(0.93)	(3.20)	(3.50)	(0.06)	(0.36)	(0.98)	(0.27)
# Loans	7,378,891	7,378,891	7,378,891	7,378,891	7,121,472	7,121,458	7,378,462	7,377,410
Adjusted $R\hat{2}$	0.29	0.30	0.52	0.56	0.26	0.25	0.51	0.46
Orig Qtr FEs?	Y	Y	Y	Y	Y	Y	Y	Y
Sale Qtr FEs?	Y	Y	Y	Y	Y	Y	Y	Y

This table shows the loan-level, OLS, and logit regressions, where the dependent variables are the 36-month and 60-month *ex-ante* default rates at the time the loan was originated, where the *ex-ante* default rates are calculated using the extensive information in the data on loan and borrower characteristics at the time of origination for the previous three years for the 36-month *ex-ante* rates and five years for the 60-month *ex-ante* rates. Default is defined as a loan being 60 days and 90 days delinquent or more at any point since origination. The independent variable of interest is months to sale, which is defined as the number of months that elapse between origination and sale to a PLS issuer or GSE. All of the regressions include origination year-quarter fixed effects and year-quarter of sale fixed effects. The standard errors are heteroskedasticity-robust and are clustered at the quarter of issuance level. The first row for each variable shows the regression coefficient and the second row shows the *t*-statistic.

Table 9: Baseline Parametric Results for the Sample of CoreLogic PLS Loans

Panel A: Including Originator Fixed Effects

	Default Definition: 60+ DQ over 36 Months								
		All PLS			**Alt-A**			**Subprime**	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Months to Sale	-0.0036	-0.0028	-0.0015	-0.0072	-0.0063	-0.0057	-0.0020	-0.0015	0.0005
	(8.57)	(6.41)	(3.66)	(10.55)	(10.54)	(11.46)	(5.15)	(3.65)	(1.11)
# Loans	7,860,499	7,858,236	7,855,810	1,895,618	1,894,861	1,893,617	5,964,881	5,963,091	5,961,433
Adjusted R^2	0.21	0.21	0.22	0.25	0.26	0.27	0.19	0.19	0.20
Orig YQ FEs?	Y	Y	Y	Y	Y	Y	Y	Y	Y
State FEs?	Y	Y	Y	Y	Y	Y	Y	Y	Y
Issue YQ FEs?	Y	Y	Y	Y	Y	Y	Y	Y	Y
Originator FEs?	N	Y	Y	N	Y	Y	N	Y	Y
Originator x Orig-YQ Fes?	N	N	Y	N	N	Y	N	N	Y
Other Controls?	Y	Y	Y	Y	Y	Y	Y	Y	Y

Panel B: Including Issuer Fixed Effects

	Default Definition: 60+ DQ over 36 Months								
		All PLS			**Alt-A**			**Subprime**	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Months to Sale	-0.0041	-0.0019	-0.0017	-0.0063	-0.0047	-0.0051	-0.0027	-0.0003	0.0005
	(8.29)	(4.15)	(3.79)	(11.83)	(9.28)	(9.94)	(6.16)	(0.70)	(1.09)
# Loans	7,725,370	7,725,369	7,722,948	1,849,905	1,849,904	1,848,674	5,875,179	5,875,179	5,873,522
Adjusted R^2	0.21	0.22	0.22	0.26	0.27	0.27	0.19	0.20	0.20
Orig YQ FEs?	Y	Y	Y	Y	Y	Y	Y	Y	Y
State FEs?	Y	Y	Y	Y	Y	Y	Y	Y	Y
Issue YQ FEs?	Y	Y	Y	Y	Y	Y	Y	Y	Y
Originator FEs?	Y	Y	Y	Y	Y	Y	Y	Y	Y
Originator x Orig-YQ Fes?	N	N	Y	N	N	Y	N	N	Y
Issuer FEs?	Y	Y	Y	Y	Y	Y	Y	Y	Y
Issuer x Issue -YQ FEs?	N	Y	Y	N	Y	Y	N	Y	Y
Other Controls?	Y	Y	Y	Y	Y	Y	Y	Y	Y

Notes: This table displays the results from the estimation of equation 5 on PLS loans in the CoreLogic dataset originated in the 2002–2007 period. The dependent variable is an indicator variable for loans that default over a 36-month horizon. Default is defined as a loan that is 60+ days delinquent. Months to sale is defined as the number of months that elapse between origination and sale to a PLS issuer. All of the regressions include origination year-quarter fixed effects, state fixed effects, year-quarter of sale fixed effects, and the

Table 10: Documentation Results for the Sample of CoreLogic PLS Loans

Panel A: Linear Results

Default Definition:	60+ DQ over 36 Months					
	All PLS		**Alt-A**		**Subprime**	
	Full Doc	Low Doc	Full Doc	Low Doc	Full Doc	Low Doc
Months to Sale	-0.0033	-0.0051	-0.0061	-0.0051	-0.0025	-0.0034
	(7.78)	(7.35)	(9.56)	(8.71)	(5.94)	(4.81)
# Loans	4,275,516	3,408,451	493,756	1,344,859	3,781,606	2,063,379
Adjusted R^2	0.18	0.25	0.16	0.28	0.17	0.24
Orig Qtr FEs?	Y	Y	Y	Y	Y	Y
Originator FEs?	Y	Y	Y	Y	Y	Y
Issue Qtr FEs?	Y	Y	Y	Y	Y	Y
Issuer Fes?	Y	Y	Y	Y	Y	Y
State FEs?	Y	Y	Y	Y	Y	Y
Other Controls?	Y	Y	Y	Y	Y	Y

Panel B: Non-Parametric Results

Default Definition:	60+ DQ over 36 Months					
	All PLS		**Alt-A**		**Subprime**	
	Full Doc	Low Doc	Full Doc	Low Doc	Full Doc	Low Doc
Months to Sale = 1	-0.0032	-0.0112	-0.0125	-0.0181	-0.0008	0.0025
	(2.55)	(5.35)	(6.80)	(10.59)	(0.64)	(1.35)
Months to Sale = 2	-0.0068	-0.0173	-0.0194	-0.0227	-0.0025	-0.0005
	(4.96)	(7.74)	(8.46)	(11.39)	(1.78)	(0.24)
Months to Sale = 3	-0.0145	-0.0244	-0.0291	-0.0272	-0.009	-0.0083
	(8.23)	(8.72)	(9.76)	(11.19)	(5.14)	(3.10)
Months to Sale = 4	-0.0173	-0.0273	-0.0343	-0.0319	-0.0119	-0.0093
	(8.40)	(7.78)	(9.81)	(10.42)	(5.80)	(2.75)
Months to Sale = 5	-0.0199	-0.0347	-0.0359	-0.0313	-0.0148	-0.0196
	(8.18	(9.15)	(8.43)	(9.74)	(6.14)	(4.94)
Months to Sale = 6	-0.0194	-0.0358	-0.0332	-0.0365	-0.0141	-0.0188
	(6.42)	(7.18)	(7.01)	(8.05)	(4.73)	(3.51)
Months to Sale = 7	-0.0208	-0.0339	-0.0428	-0.0381	-0.0151	-0.0194
	(5.29)	(5.61)	(7.04)	(6.53)	(3.79)	(3.11)
Months to Sale = 8	-0.0118	-0.0275	-0.0464	-0.0455	-0.0049	-0.0033
	(2.82)	(4.08)	(7.63)	(6.49)	(1.08)	(0.51)
Months to Sale = 9	-0.0033	-0.032	-0.0535	-0.0526	0.0097	-0.0004
	(0.64)	(5.08)	(7.24)	(9.48)	(1.70)	(0.06)
# Loans	4,275,516	3,408,451	493,756	1,344,859	3,781,606	2,063,379
Adjusted R^2	0.18	0.25	0.16	0.28	0.17	0.24
Orig Qtr FEs?	Y	Y	Y	Y	Y	Y
Originator FEs?	Y	Y	Y	Y	Y	Y
Issue Qtr FEs?	Y	Y	Y	Y	Y	Y
Issuer Fes?	Y	Y	Y	Y	Y	Y
State FEs?	Y	Y	Y	Y	Y	Y
Other Controls?	Y	Y	Y	Y	Y	Y

Notes: This table displays the results from the estimation of equation 5 on PLS loans in the CoreLogic dataset originated in the 2002–2007 period. The dependent variable is an indicator variable for loans that default over a 36-month horizon. Default is defined as a loan that is 60+ days delinquent. Months to sale is defined as the number of months that elapse between origination and sale to a PLS issuer. All of the regressions include origination year-quarter fixed effects, state fixed effects, year-quarter of sale fixed effects, originator fixed effects, and the detailed list of covariates described in the text. "Full Doc" loans correspond to those in which the borrower's income and assets were not fully documented at the time of origination, while "Low Doc" loans are those in which either the borrower's income or assets (or both) were not fully documented. The first row for each variable shows the regression coefficient and the second row

Table 11: Affiliation Results for the Sample of CoreLogic PLS Loans

Panel A: Parametric Results

Default Definition:	60+ DQ over 36 Months					
	All PLS		**Alt-A**		**Subprime**	
	Affiliation	No Affiliation	Affiliation	No Affiliation	Affiliation	No Affiliation
Months to Sale	-0.0046	-0.0049	-0.0049	-0.0100	-0.0029	-0.0031
	(6.93)	(8.02)	(7.12)	(10.84)	(4.96)	(5.21)
# Loans	3,176,715	3,473,338	603,234	735,374	2,573,481	2,737,861
Adjusted R^2	0.20	0.21	0.24	0.26	0.19	0.20
Orig Qtr FEs?	Y	Y	Y	Y	Y	Y
Originator FEs?	Y	Y	Y	Y	Y	Y
Issue Qtr FEs?	Y	Y	Y	Y	Y	Y
Issuer Fes?	Y	Y	Y	Y	Y	Y
State FEs?	Y	Y	Y	Y	Y	Y
Other Controls?	Y	Y	Y	Y	Y	Y

Panel B: Non-Parametric Results

Default Definition:	60+ DQ over 36 Months					
	All PLS		**Alt-A**		**Subprime**	
	Affiliation	No Affiliation	Affiliation	No Affiliation	Affiliation	No Affiliation
Months to Sale = 1	-0.0039	-0.0313	-0.007	-0.0442	-0.0014	0.0021
	(2.12)	(8.66)	(3.40)	(13.42)	(1.11)	(0.68)
Months to Sale = 2	-0.0071	-0.0396	-0.0114	-0.0513	-0.0027	-0.0065
	(3.34)	(11.78)	(4.91)	(14.97)	(1.46)	(2.16)
Months to Sale = 3	-0.0153	-0.0468	-0.0206	-0.0615	-0.0082	-0.0136
	(5.77)	(12.26)	(4.34)	(15.34)	(3.64)	(4.15)
Months to Sale = 4	-0.0197	-0.0482	-0.0243	-0.0712	-0.0112	-0.0142
	(6.56)	(11.14)	(6.76)	(13.69)	(4.22)	(3.77)
Months to Sale = 5	-0.0271	-0.0523	-0.02	-0.075	-0.018	-0.0185
	(6.58)	(11.04)	(4.24)	(14.32)	(5.37)	(4.51)
Months to Sale = 6	-0.024	-0.054	-0.0221	-0.0817	-0.0147	-0.0192
	(5.12)	(9.74)	(4.11)	(11.91)	(3.37)	(3.83)
Months to Sale = 7	-0.0326	-0.0555	-0.0263	-0.0923	-0.0274	-0.0178
	(5.03)	(9.52)	(3.97)	(12.01)	(4.20)	(3.31)
Months to Sale = 8	-0.0316	-0.0507	-0.0491	-0.0971	-0.0212	-0.009
	(4.49)	(8.68)	(6.31)	(10.61)	(2.83)	(1.49)
Months to Sale = 9	-0.0152	-0.052	-0.0428	-0.1128	-0.0035	-0.0021
	(1.81)	(8.11)	(4.24)	(12.40)	(0.32)	(0.31)
# Loans	3,176,715	3,473,338	603,234	735,374	2,573,481	2,737,861
Adjusted R^2	0.20	0.21	0.24	0.26	0.19	0.20
Orig Qtr FEs?	Y	Y	Y	Y	Y	Y
Originator FEs?	Y	Y	Y	Y	Y	Y
Issue Qtr FEs?	Y	Y	Y	Y	Y	Y
Issuer Fes?	Y	Y	Y	Y	Y	Y
State FEs?	Y	Y	Y	Y	Y	Y
Other Controls?	Y	Y	Y	Y	Y	Y

Notes: This table displays the results from the estimation of equation 5 on PLS loans in the CoreLogic dataset originated in the 2002–2007 period. The dependent variable is an indicator variable for loans that default over a 36-month horizon. Default is defined as a loan that is 60+ days delinquent. Months to sale is defined as the number of months that elapse between origination and sale to a PLS issuer. All of the regressions include origination year-quarter fixed effects, state fixed effects, year-quarter of sale fixed effects, originator fixed effects, and the detailed list of covariates described in the text. "Affiliated" PLS deals are those in which the originator of all of the mortgages in the deal is affiliated with the issuer (either the same company or part of the same vertical corporation). The first row for each variable shows the regression coefficient and the second row shows the *t*-statistic. The standard errors are heteroskedasticity-robust and

Table 12: Pricing Analysis Results

Panel A: Linear Specification

Dependent Variable: Pool-level Average Yield Spread (Triple-A Securities Only)

	All Securities			Alt-A Securities			Subprime Securities		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Seasoning	-0.015***	-0.003	-0.010***	-0.022*	-0.024*	-0.024*	0.002	0.002	0.002
	(0.004)	(0.003)	(0.003)	(0.013)	(0.014)	(0.014)	(0.002)	(0.002)	(0.002)
Pool Covariates?	N	Y	Y	N	Y	Y	N	Y	Y
Issue Qtr FE?	Y	Y	Y	Y	Y	Y	Y	Y	Y
Issuer FE?	N	N	Y	N	N	Y	N	N	Y
Observations	3,532	3,532	3,513	909	909	909	2,623	2,615	2,615
Adjusted R^2	0.17	0.33	0.45	0.09	0.16	0.16	0.67	0.71	0.71

Panel B: Non-Linear Specification

Dependent Variable: Pool-level Average Yield Spread (Triple-A Securities Only)

	All Securities			Alt-A Securities			Subprime Securities		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Seasoning	-0.095***	-0.051***	-0.035**	-0.177***	-0.169***	-0.169***	-0.003	-0.008	-0.009
	(0.019)	(0.018)	(0.014)	(0.047)	(0.052)	(0.052)	(0.006)	(0.007)	(0.008)
Seasoning^2	0.010***	0.006***	0.003**	0.020***	0.019***	0.019***	0.001	0.001	0.001
	(0.002)	(0.002)	(0.002)	(0.007)	(0.007)	(0.007)	(0.001)	(0.001)	(0.001)
Pool Covariates?	N	Y	Y	N	Y	Y	N	Y	Y
Issue Qtr FE?	Y	Y	Y	Y	Y	Y	Y	Y	Y
Issuer FE?	N	N	Y	N	N	Y	N	N	Y
Observations	3,532	3,532	3,513	909	909	909	2,623	2,615	2,615
Adjusted R^2	0.17	0.33	0.45	0.09	0.16	0.16	0.67	0.71	0.71

This table displays the results from the estimation of equation 7. The sample includes triple-A, floating rate subprime, and Alt-A securities issued between January 2002 and December 2007. The dependent variable is the weighted average spread over the 1-month LIBOR of all triple-A securities with claims on cash flows for a given mortgage pool. Seasoning is the average age (# months) of all mortgages in a pool at the time of issuance. All of the regressions include month-of-issue fixed effects. The set of pool-level covariates corresponds to the variables included in Table 14, which are all pool-level averages. The first row for each variable shows the regression coefficient and the second row shows the t-statistic. The standard errors are heteroskedasticity-robust and are clustered at the deal-level. Statistical significance is denoted by stars, with the following mapping: *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 13: Early Prepayment Results

Panel A: Parametric Specification

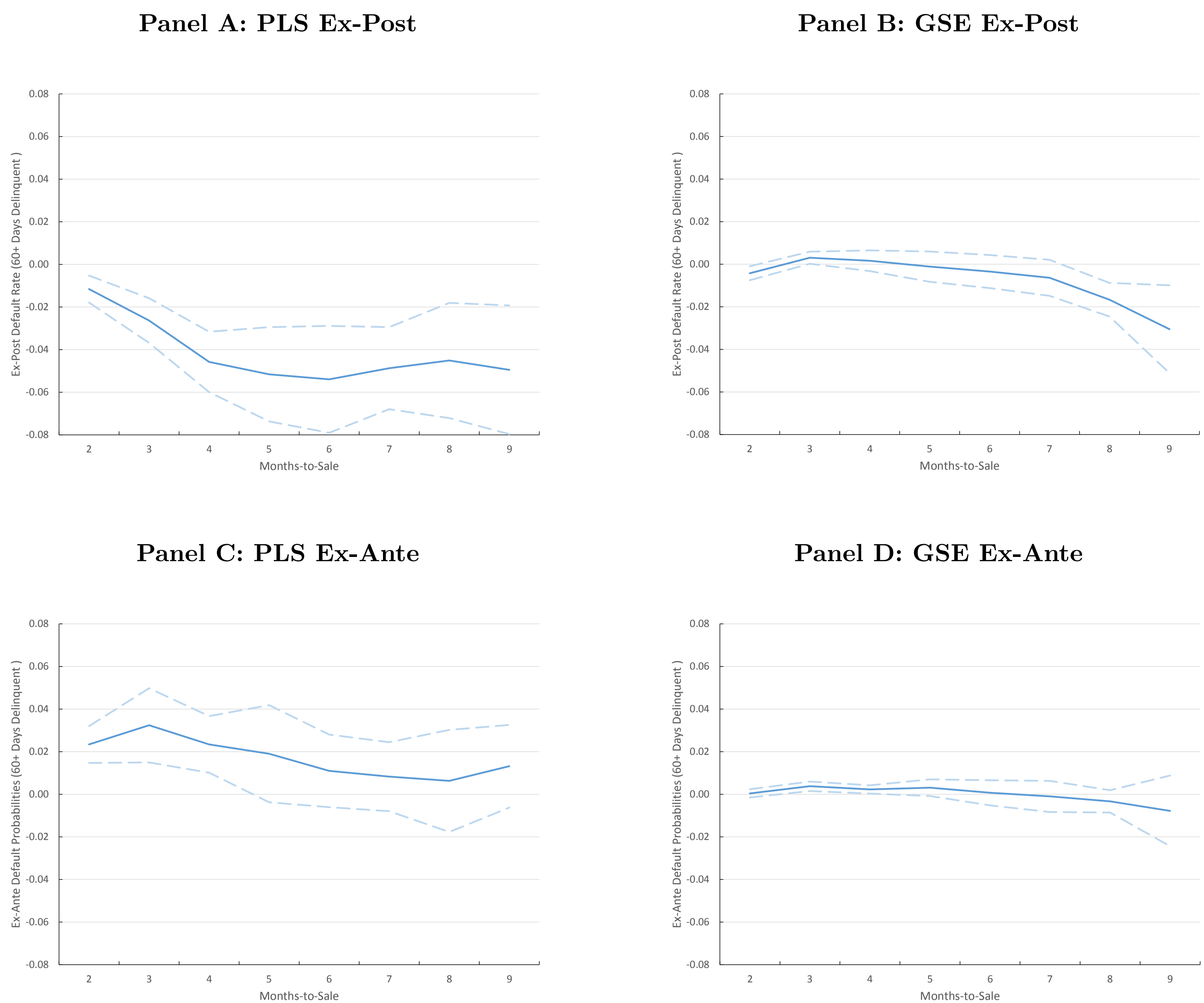
Correction:	None		$\leq$ 3 months		$\leq$ 6 months		$\leq$ 9 months	
Reset Month - Prepay Month	$\geq$ 6 Months	$\geq$ 9 Months	$\geq$ 6 Months	$\geq$ 9 Months	$\geq$ 6 Months	$\geq$ 9 Months	$\geq$ 6 Months	$\geq$ 9 Months
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Months to Sale	-0.0129	-0.0152	-0.0089	-0.0105	-0.0111	-0.0131	-0.0144	-0.0169
	(6.20)	(6.28)	(4.11)	(4.15)	(4.76)	(4.75)	(5.66)	(5.57)
Months to Sale ^{2}	0.0007	0.0009	0.0004	0.0005	0.0012	0.0015	0.0019	0.0023
	(2.56)	(2.83)	(1.36)	(1.58)	(3.75)	(4.03)	(5.07)	(5.36)
# Loans	4,024,361	4,024,361	3,968,227	3,968,227	3,701,607	3,701,607	3,302,260	3,302,260
Adjusted R^2	0.13	0.13	0.13	0.13	0.12	0.11	0.10	0.08
Orig Qtr FEs?	Y	Y	Y	Y	Y	Y	Y	Y
State FEs?	Y	Y	Y	Y	Y	Y	Y	Y
Sale Qtr FEs?	Y	Y	Y	Y	Y	Y	Y	Y
Originator FEs?	Y	Y	Y	Y	Y	Y	Y	Y
Other Controls?	Y	Y	Y	Y	Y	Y	Y	Y

Panel B: Non-parametric Specification

Correction:	None		$\leq$ 3 months		$\leq$ 6 months		$\leq$ 9 months	
Reset Month - Prepay Month	$\geq$ 6 Months	$\geq$ 9 Months	$\geq$ 6 Months	$\geq$ 9 Months	$\geq$ 6 Months	$\geq$ 9 Months	$\geq$ 6 Months	$\geq$ 9 Months
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Months to Sale = 1	-0.024	-0.027	-0.023	-0.025	-0.024	-0.027	-0.025	-0.028
	(4.90)	(4.87)	(4.41)	(4.34)	(4.51)	(4.39)	(4.07)	(3.87)
Months to Sale = 2	-0.033	-0.038	-0.028	-0.032	-0.030	-0.034	-0.031	-0.035
	(6.90)	(6.88)	(5.70)	(5.63)	(5.81)	(5.69)	(5.17)	(4.98)
Months to Sale = 3	-0.039	-0.045	-0.030	-0.035	-0.032	-0.037	-0.034	-0.039
	(7.09)	(7.07)	(5.19)	(5.13)	(5.36)	(5.25)	(5.13)	(4.89)
Months to Sale = 4	-0.043	-0.049	-0.034	-0.038	-0.029	-0.033	-0.030	-0.033
	(7.24)	(7.48)	(5.36)	(5.47)	(4.51)	(4.53)	(4.50)	(4.38)
Months to Sale = 5	-0.049	-0.056	-0.040	-0.045	-0.026	-0.028	-0.028	-0.030
	(9.32)	(9.35)	(7.06)	(7.02)	(4.43)	(4.21)	(4.69)	(4.26)
Months to Sale = 6	-0.059	-0.066	-0.049	-0.055	-0.024	-0.024	-0.027	-0.027
	(8.59)	(8.93)	(6.93)	(7.15)	(3.03)	(2.88)	(3.24)	(3.02)
Months to Sale = 7	-0.064	-0.072	-0.054	-0.060	-0.027	-0.028	-0.014	-0.012
	(7.97)	(7.83)	(6.65)	(6.54)	(3.22)	(3.01)	(1.50)	(1.14)
Months to Sale = 8	-0.082	-0.090	-0.073	-0.078	-0.046	-0.047	-0.017	-0.011
	(10.65)	(11.38)	(8.99)	(9.56)	(5.57)	(5.63)	(1.91)	(1.22)
Months to Sale = 9	-0.096	-0.108	-0.085	-0.097	-0.059	-0.065	-0.011	-0.008
	(9.67)	(9.07)	(8.58)	(8.00)	(5.84)	(5.44)	(1.01)	(0.58)
# Loans	4,024,361	4,024,361	3,968,227	3,968,227	3,701,607	3,701,607	3,302,260	3,302,260
Adjusted R^2	0.13	0.13	0.13	0.13	0.12	0.11	0.10	0.08
Orig Qtr FEs?	Y	Y	Y	Y	Y	Y	Y	Y
State FEs?	Y	Y	Y	Y	Y	Y	Y	Y
Sale Qtr FEs?	Y	Y	Y	Y	Y	Y	Y	Y
Originator FEs?	Y	Y	Y	Y	Y	Y	Y	Y
Other Controls?	Y	Y	Y	Y	Y	Y	Y	Y

Notes: This table displays the results from the estimation of equation 5 on adjustable-rate PLS loans in the CoreLogic dataset originated in the 2002–2007 period. The dependent variable is an indicator variable for loans that prepay more than three months or six months before the month in which the interest rate resets from a fixed rate to an adjustable rate. All of the loans that are prepaid within three months of origination are eliminated from the sample. Months to sale is defined as the number of months that elapse between origination and sale to a PLS issuer. All of the regressions include origination year-quarter fixed effects, state fixed effects, year-quarter of sale fixed effects, originator fixed effects, and the detailed list of covariates described in the text. The first row for each variable shows the regression coefficient and the second row shows the t-statistic. The standard errors are heteroskedasticity-robust and are clustered by year-quarter of origination.

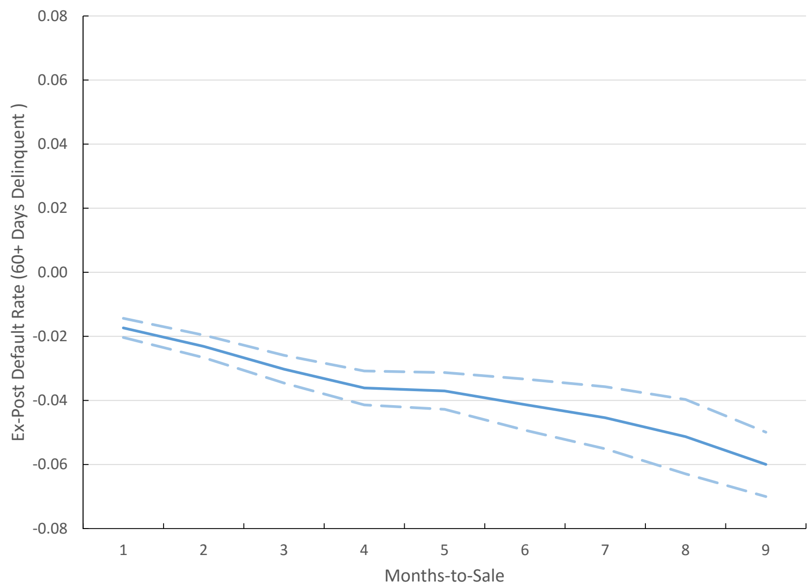
Figure 1: Ex-Ante vs. Ex-Post LPS Results



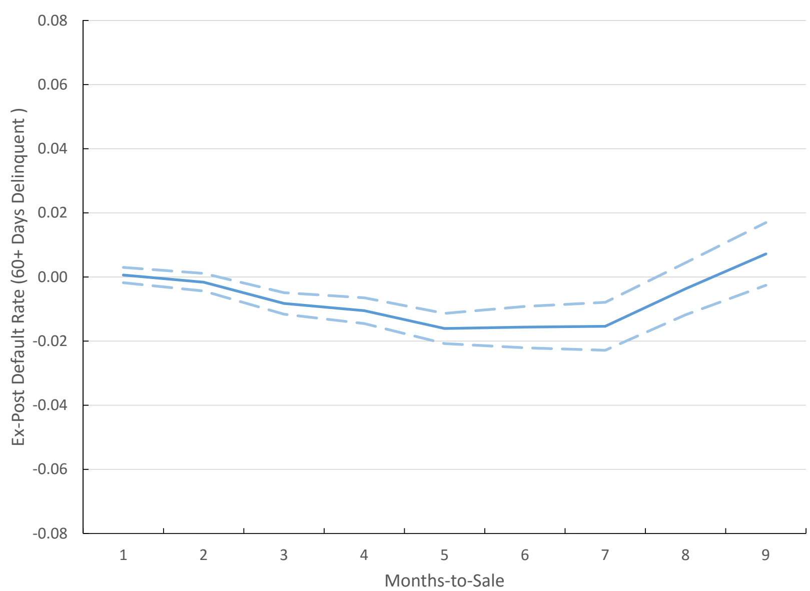
Notes: This figure displays the results from the estimation of the non-parametric version of equation 5 for both the PLS and GSE loans in the LPS dataset originated in the 2002–2007 period. Panels A and B correspond to the *ex-post* default rates, while panels C and D correspond to *ex-ante* predicted default rates. Default is defined as a loan that becomes 60 days delinquent over a 36-month horizon measured from origination. Months to sale is defined as the number of months that elapse between origination and sale to a

Figure 2: CoreLogic PLS Results

Panel A: Alt-A PLS



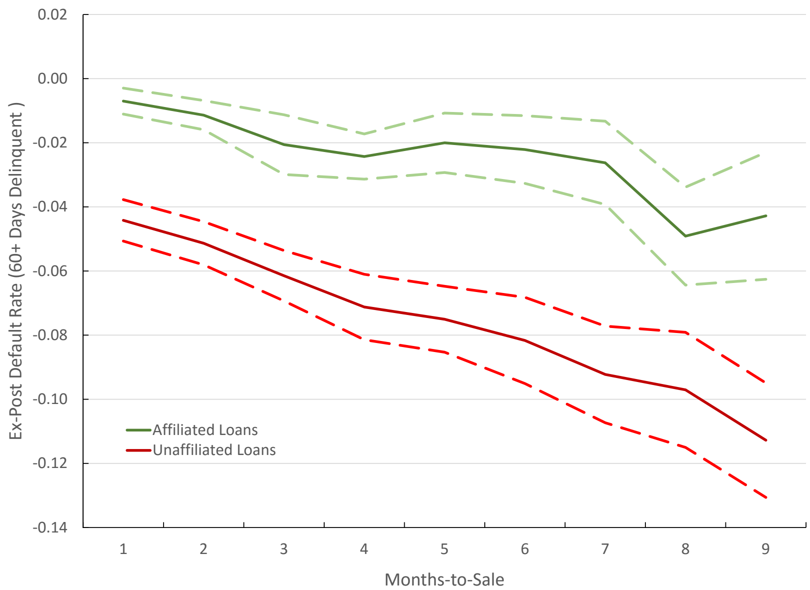
Panel B: Subprime PLS



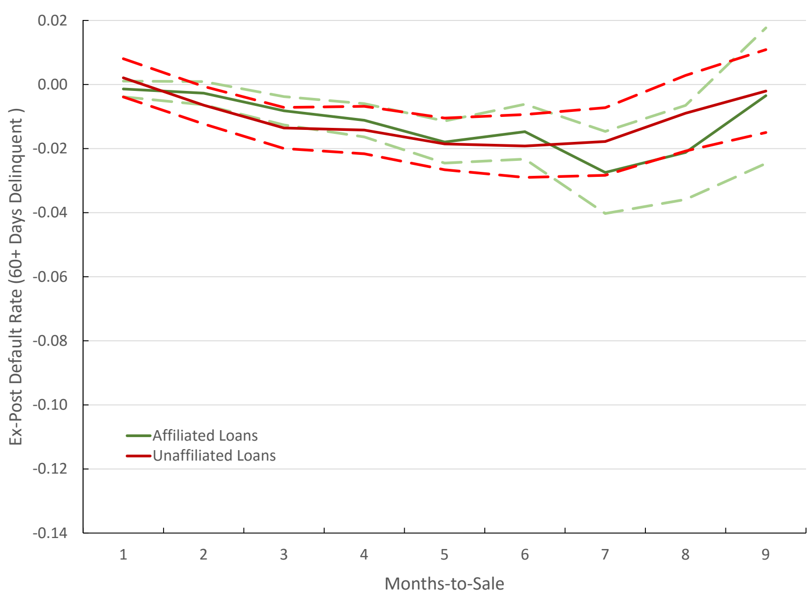
Notes: This figure displays the results from the estimation of the non-parametric version of equation 5 for PLS loans in the CoreLogic dataset originated in the 2002–2007 period. Panel A corresponds to Alt-A PLS loans and panel B corresponds to subprime loans. Default is defined as a loan that becomes 60 days delinquent over a 36-month horizon measured from origination. Months to sale is defined as the number of months that elapse between origination and sale to a PLS issuer. The dotted lines correspond to 90% confidence intervals.

Figure 3: CoreLogic PLS Affiliation Results

Panel A: Alt-A PLS

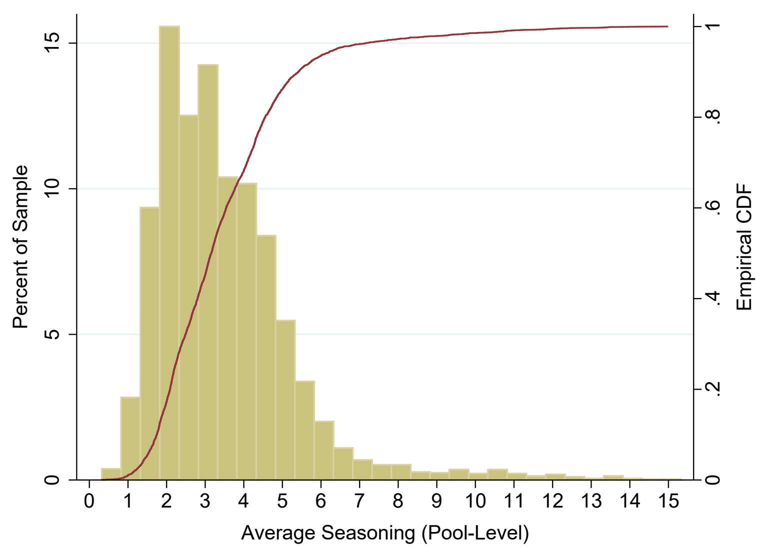


Panel B: Subprime PLS



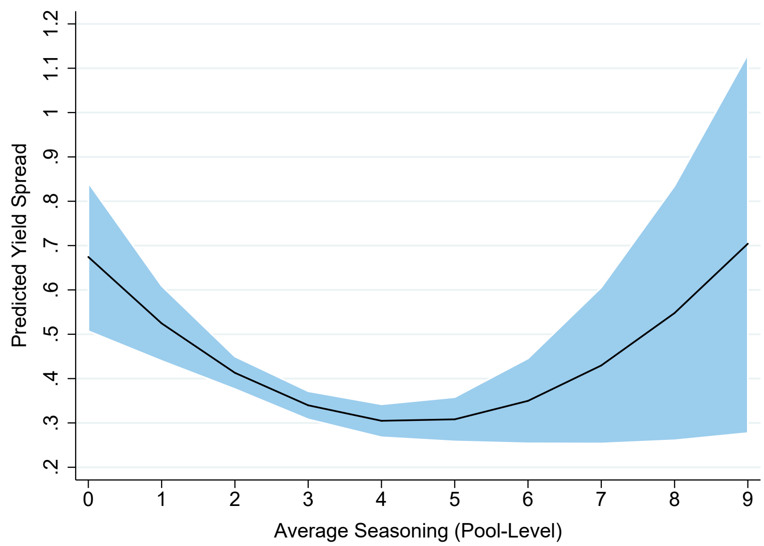
Notes: This figure displays the results from the estimation of the non-parametric version of equation 5 for PLS loans in the CoreLogic dataset originated in the 2002–2007 period. Panel A corresponds to Alt-A PLS loans and panel B corresponds to subprime loans. Default is defined as a loan that becomes 60 days delinquent over a 36-month horizon measured from origination. Months to sale is defined as the number of months that elapse between origination and sale to a PLS issuer. The dotted lines correspond to 90% confidence intervals.

Figure 4: Distribution of Pool-Level Seasoning



Notes: This figure displays the density and cumulative distribution of the average months of seasoning in the sample of floating-rate, triple-A, Subprime and Alt-A securities issued between January 2002 and December 2007 used in the pricing analysis in section 5.6.

Figure 5: Predicted Yield Spread as Function of Seasoning



Notes: This figure displays the predicted security spreads (over the 1-month LIBOR) as a function of the average pool-level seasoning calculated using the estimation results from the specification reported in column (6) in panel B of Table 12. The shaded area corresponds to 95% confidence intervals calculated using the delta method.

Appendix

Variable Definitions

ARM: An indicator variable that takes a value of 1 if the mortgage has an adjustable rate and 0 if it has a fixed rate.

Balance: The natural logarithm of the principal balance of the loan at origination.

Balloon: An indicator variable that takes a value of 1 if the mortgage is characterized by a balloon payment at the end of its term and 0 if it is a fully amortizing mortgage.

Condo: An indicator variable that takes a value of 1 if the property is a condominium or a townhouse and 0 otherwise.

FICO: The credit score of the borrower at origination. All of the models include both the continuous FICO variable and a set of indicator variables corresponding to 5 FICO intervals: FICO < 580, $580 \leq$FICO < 620, $620 \leq$FICO < 660, $660 \leq$FICO < 700, and FICO ≥ 700.

House Prices: County-level house price indices from CoreLogic. We include both price level in the county in the month of origination and the cumulative growth in prices from the month of mortgage origination calculated over the default horizon.

Interest-Only: An indicator variable that takes a value of 1 if the loan requires payments of only interest for a specified period of time and 0 otherwise.

Jumbo: An indicator variable that takes a value of 1 if the loan amount at origination exceeds the conforming loan limit set by statute that limits the size of mortgages eligible to be insured by the GSEs (during the vast majority of our sample period, the limit was $417,000 for mortgages on single-family properties) and 0 otherwise.

Loan-to-Value (cumulative): The loan-to-value ratio at origination computed using information on the first and second liens. All of the models include both the continuous

LTV variable and a set of indicator variables corresponding to 5 LTV intervals: LTV < 70, $70 \leq$ LTV < 80, $80 <$ LTV < 90, $90 \leq$LTV< 100 and LTV ≥ 100. An indicator variable for the LTV ratios exactly equal to 80 is also included as a proxy for unreported second liens.

Low Documentation: An indicator variable that takes a value of 1 if the borrower's income and assets are not fully documented in the underwriting process and 0 if they are fully documented.

Month to Sale: The number of months after the date of origination in which a loan is sold to a PLS issuer or acquired by one of the GSEs. In the LPS dataset, the variable is based on a field that is updated monthly and shows the current holder of the loan. In the CoreLogic LoanPerformance database, the variable is based on the length of time between the month of origination and the month in which the corresponding PLS security is issued.

Multi-family: An indicator variable that takes a value of 1 if the property is a 2–4-family house and 0 otherwise.

Negative Amortization: An indicator variable that takes a value of 1 if the loan requires payments of less than interest and principal for a specified period of time and 0 otherwise.

Prepayment Penalty: An indicator variable that takes a value of 1 if the mortgage contains a prepayment penalty and 0 otherwise.

Primary Residence: An indicator variable that takes a value of 1 if the property is the primary residence of the borrower and a value of 0 if the property is either an investment or a second home.

Purchase Loan: An indicator variable that takes a value of 1 if the loan is used to purchase property and 0 otherwise.

Refinance (traditional): An indicator variable that takes a value of 1 if the loan is used to refinance previous mortgage debt without converting any equity into cash and 0 otherwise.

Refinance (cashout): An indicator variable that takes a value of 1 if the loan is used to refinance previous mortgage debt with a portion of the equity converted to cash and 0 otherwise.

Single Family: An indicator variable that takes a value of 1 if the property is a detached single-family home and 0 otherwise.

Term: The maturity length of the mortgage in months.

Unemployment: County-level unemployment rates from the Bureau of Labor Services (BLS). We include both the levels of the rates in the county in the month of origination and the cumulative growth in the unemployment rate from the month of mortgage origination calculated over the default horizon.

Table 14: Pricing Analysis Summary Statistics

	Mean	Standard Dev.	Minimum	25th Perc.	Median	75th Perc.	Maximum
Yield Spread	0.28	0.23	0.04	0.16	0.23	0.32	2.09
Months to Sale	3.3	1.4	0.3	2.2	3.1	4.2	9.0
# Loans	2,355	1,833	55	1,108	1,911	3,078	18,190
Log Loan Balance	12.2	0.4	11.0	11.9	12.1	12.4	14.9
FICO	640	43	413	609	624	682	764
FICO $<$ 580	0.20	0.15	0.00	0.01	0.22	0.31	0.87
580 $\leq$ FICO $<$ 620	0.19	0.12	0	0.05	0.22	0.27	0.67
620 $\leq$ FICO $<$ 660	0.23	0.08	0	0.19	0.24	0.28	0.68
660 $\leq$ FICO $<$ 700	0.18	0.09	0.01	0.11	0.15	0.25	0.72
FICO $\geq$ 700	0.20	0.21	0	0.06	0.10	0.35	0.92
CLTV	84	6	39	80	84	88	102
CLTV $<$ 70							
70 $\leq$ CLTV $<$ 80	0.15	0.07	0	0.10	0.14	0.19	0.49
80 $\leq$ CLTV $<$ 90	0.28	0.13	0	0.20	0.27	0.36	0.92
90 $\leq$ CLTV $<$ 100	0.24	0.10	0	0.18	0.23	0.29	0.97
CLTV $\geq$ 100	0.20	0.20	0	0.02	0.16	0.32	0.96
LTV = 80	0.16	0.12	0	0.08	0.12	0.20	0.91
Term	359	15	120	356	359	360	480
Purchase Loan	0.42	0.20	0	0.27	0.40	0.57	1
Cashout Refinance	0.48	0.19	0	0.33	0.50	0.62	1
Primary Residence	0.87	0.13	0	0.85	0.91	0.95	1
Single-Family Property	0.73	0.11	0	0.68	0.75	0.80	0.99
Condominium	0.08	0.04	0	0.05	0.07	0.09	0.36
ARM	0.83	0.18	0	0.76	0.85	1	1
Interest-Only	0.21	0.28	0	0	0.10	0.26	1
Negative Amortization	0.10	0.30	0	0	0	0	1
Low Documentation	0.47	0.23	0	0.31	0.41	0.61	1
Balloon	0.08	0.15	0	0	0	0.05	1
Jumbo	0.19	0.24	0	0	0.10	0.27	1
Prepayment Penalty	0.69	0.21	0	0.65	0.74	0.81	1
Fraction in CA	0.26	0.17	0	0.13	0.23	0.34	1
Unemployment Rate	5.14	0.61	1.73	4.66	5.06	5.63	6.83
Predicted WAL	2.59	0.61	0	2.23	2.52	2.90	6.61
Subordination	1.00	3.10	0	0.81	0.85	0.91	103.35
# Securities				3,532			

Notes: This table displays the summary statistics for the variables included in the pricing analysis presented in section 5.6. All of the mortgage characteristics correspond to averages that are calculated at the pool-level in the sample of CoreLogic loans, which includes mortgages backing subprime and Alt-A triple-A floating rate securities issued between January 2002 and December 2007. Yield Spread is the weighted average spread over the 1-month LIBOR of all triple-A securities with claims on cash flows for a given mortgage pool. Seasoning is the average age (# months) of all mortgages in a pool at the time of issuance. Predicted WAL is a model-based calculation of the expected weighted average life. Subordination is calculated as the ratio of the total face value of all triple-A securities associated with a pool to the sum of the remaining principal balances of all of the loans in the pool in the month of issuance.

Table 15: Model Coefficient Estimates

Dependent Variable: Indicator for 60+ DQ within 36 months of origination	
Months to Sale	-0.0107
	(5.79)
Primary Residence (d)	-0.0012
	(0.49)
Prepayment Penalty (d)	0.0687
	(7.70)
ARM (d)	0.0281
	(2.24)
Balloon Payment (d)	0.0890
	(4.74)
Low Documentation (d)	0.0515
	(9.74)
Missing Documentation (d)	0.0119
	(1.80)
B or C Grade Mortgage (d)	0.1091
	(9.38)
Single Family Property (d)	-0.0010
	(0.69)
Missing Property Type (d)	0.0302
	(7.12)
Interest-Only (d)	0.0130
	(1.44)
Purchase Loan (d)	0.0015
	(0.22)
Refinance (cash-out) (d)	0.0141
	(3.04)
Missing Loan Type (d)	0.0141
	(3.04)
Term	0.0001
	(2.81)

LTV	0.0010
	(3.96)
Missing LTV (d)	0.1632
	(4.23)
$70 \leq$ LTV < 80 (d)	0.0352
	(4.19)
LTV $= 80$ (d)	0.0257
	(7.33)
$80 <$ LTV < 90 (d)	0.0443
	(4.75)
$900 \leq$ LTV < 100 (d)	0.0608
	(5.72)
LTV ≥ 100 (d)	0.0459
	(4.04)
FICO	-0.0011
	(8.59)
Missing FICO (d)	-0.8955
	(8.54)
FICO < 580 (d)	-0.0614
	(3.22)
$580 \leq$ FICO < 620 (d)	-0.0482
	(4.53)
$620 \leq$ FICO < 660 (d)	-0.0149
	(5.86)
$660 \leq$ FICO < 700 (d)	-0.0128
	(2.72)
Interest Rate (at origination)	0.0110
	(6.53)
Jumbo (d)	0.0217
	(2.55)
Unemployment Rate (at origination)	0.0041
	(7.63)
Cumulative Change in Unemployment Rate (36 months)	0.0244
	(5.75)

House Price Level (at origination)	0.0016
	(12.36)
Cumulative Change in House Prices (36 months)	-0.1583
	(7.65)
# Loans	5,313,951
Adjusted R^2	0.23
Orig Qtr FEs?	Y
State FEs?	Y
Sale Qtr FEs?	Y
Originator FEs?	N

Notes: This table displays the full set of results for the specification in Table 3, column (1). The dependent variable is an indicator variable for loans that default over a 36-month horizon. Default is defined as a loan that is 60+ days delinquent. Months to sale is defined as the number of months that elapse between origination and sale to a PLS issuer. All of the regressions include origination year-quarter fixed effects, state fixed effects, and year-quarter of sale fixed effects. The first row for each variable shows the regression coefficient and the second row shows the t-statistic. The standard errors are heteroskedasticity-robust and are clustered by year-quarter of origination.

Table 16: Correcting for Potential Selection Bias: Alt-A and Subprime PLS Loans

Panel A: Effect of Lender Fixed Effects

	Default Definition: 60+ DQ over 36 Months					
	Alt-A			**Subprime**		
Months to Sale	-0.0054	-0.0042	-0.0035	0.0020	0.0026	0.0039
	(7.62)	(6.86)	(6.92)	(4.83)	(5.72)	(8.60)
# Loans	1,848,602	1,847,871	1,846,633	5,426,811	5,425,136	5,423,582
Adjusted R^2	0.24	0.25	0.26	0.17	0.18	0.18
Orig YQ FEs?	Y	Y	Y	Y	Y	Y
State FEs?	Y	Y	Y	Y	Y	Y
Issue YQ FEs?	Y	Y	Y	Y	Y	Y
Lender FEs?	N	Y	Y	N	Y	Y
Lender x Orig-YQ Fes?	N	N	Y	N	N	Y
Other Controls?	Y	Y	Y	Y	Y	Y

Panel B: Effect of Issuer Fixed Effects

	Default Definition: 60+ DQ over 36 Months					
	Alt-A			**Subprime**		
Months to Sale	-0.0040	-0.0023	-0.0026	0.0020	0.0040	0.0043
	(7.37)	(4.41)	(5.09)	(4.00)	(7.82)	(8.03)
# Loans	1,803,941	1,803,940	1,802,714	5,344,226	5,344,226	5,342,673
Adjusted R^2	0.25	0.26	0.26	0.18	0.18	0.18
Orig YQ FEs?	Y	Y	Y	Y	Y	Y
State FEs?	Y	Y	Y	Y	Y	Y
Issue YQ FEs?	Y	Y	Y	Y	Y	Y
Lender FEs?	Y	Y	Y	Y	Y	Y
Lender x Orig-YQ Fes?	N	N	Y	N	N	Y
Issuer FEs?	Y	Y	Y	Y	Y	Y
Issuer x Issue -YQ FEs?	N	Y	Y	N	Y	Y
Other Controls?	Y	Y	Y	Y	Y	Y

Notes: This table displays results from the estimation of equation 5 on Alt-A and Subprime PLS loans in the CoreLogic dataset that do not default within 10 months of origination. The specifications are identical to those in Table 9 in the text. The dependent variable is an indicator variable for loans that default over a 36-month horizon. Default is defined as a loan that is 60+ days delinquent. Months to sale is defined as the number of months that elapse between origination and sale to a PLS issuer. All of the regressions include origination year-quarter fixed effects, state fixed effects, year-quarter of sale fixed effects, and the detailed list of covariates described in the text. The first row for each variable shows the regression coefficient and the second row shows the t-statistic. The standard errors are heteroskedasticity-robust and are clustered by year-quarter of origination.