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 2016-8
July 2016

Abstract: A central result in the theory of adverse selection in asset markets is that informed sellers can signal quality by delaying trade. This paper uses the residential mortgage market as a laboratory to test this mechanism. Using detailed, loan-level data on privately securitized mortgages, we find a strong relation between mortgage performance and time-to-sale. Importantly, this finding is conditional on all observable information about the loans. This effect is strongest in the "Alt-A" segment of the market, where loans are often originated with incomplete documentation. The results provide some of the first evidence of a signaling mechanism through delay of trade.

JEL classification: G17, G21, G23

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

---

For helpful comments and discussions, the authors thank Brendan Daley, Joseph Mason, 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, and the 2016 Financial Intermediation Research Society conference. They also thank Valeria Vargas-Sejas for 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 that 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 fragility of over-the-counter markets (Daley and Green (2012)). However, there is remarkably little empirical evidence that agents actually engage in costly signaling to overcome informational asymmetries. This paper uses data on the U.S. mortgage market to test a central prediction of dynamic signaling models, namely that sellers signal asset quality by delaying the sale of higher quality mortgages.

In dynamic environments with adverse selection and durable assets, sellers can signal private information through the timing of their trades. Sellers of high quality assets face a lower cost of waiting because assets produce a higher interim dividend stream, or, equivalently, have a lower probability of a negative event (e.g., default). Thus, a central prediction of the costly signaling framework is that there is a positive relationship between *unobserved* asset quality and time-to-sale, often referred to as the skimming property.[1] We provide evidence of the skimming property in the originate-to-distribute mortgage market by showing that mortgages that take longer to sell default at a lower rate *after controlling for observable characteristics.*

To motivate our empirical tests, we first present a simple model of mortgage sales. This model illustrates the skimming property and shares several features with many models in the literature. In our model, a mortgage originator (the seller) faces a competitive

[1]The skimming property is one of the properties derived from the Coase (1972) analysis of pricing by a durable-goods monopolist (Coasian dnamics). Recently, many studies have found that the skimming property can emerge in dynamic adverse selection models of financial markets, see for example Daley and Green (2012),Fuchs and Skrzypacz (2013), Fuchs et al. (2015).

market of buyers for the mortgage. The seller privately observes the quality of the mortgage (measured by its probability of default) and we assume that default is publicly observable and extinguishes the possibility of sale. The fact that the gains from trade are lost if the mortgage defaults before it is sold implies a higher cost of waiting for sellers of lower quality mortgages. A separating equilibrium emerges in which time-to-sale perfectly reveals mortgage quality. This equilibrium provides the key prediction that mortgages that take longer to sell should default at a lower rate.

There are two central requirements for a test of any signaling equilibrium, including tests of the skimming property in asset markets. First, there has to be a plausible source of *unobserved* heterogeneity in asset quality that is (i) known (at least partially) by the seller, (ii) unknown to potential buyers, and (iii) known to the econometrician. That the private information of the seller is also known to the econometrician is a fundamental challenge in testing models of adverse selection. The second requirement is that at least a subset of *observable* characteristics that are available to the agents is also observable by the econometrician. The distinction between observable and unobservable asset characteristics is key to any such test, as the predictions of signaling models only apply to unobserved heterogeneity. In fact, most models predict that assets that are observably better should trade *faster*, not slower. As a result of these requirements, this core prediction of dynamic adverse selection models has yet to be empirically tested.[2]

The mortgage market serves as a unique laboratory for such a test. First, mortgages are durable assets that are characterized by an objective measure of quality based on the probability of repayment (i.e. credit risk). While outcomes were not known at the time of sale, they are known to the econometrician *ex post*. Second, during the middle of the last decade there was an active secondary market for mortgages, where issuers of mortgage-backed securities (the buyers) purchased large portfolios of mortgages from originators (the sellers). Third, there is detailed micro data available on the observable characteristics of

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

borrowers and mortgage contracts known to issuers (the buyers), the originators (the sellers), and the econometrician. This provides us a good proxy of observable mortgage quality at the time of the sale. Finally, there is evidence from previous studies that the originators of mortgages have private information that is correlated with *ex post* mortgage performance but is not available to the buyers. By conditioning on observables and tracking mortgage performance, we are able to distinguish between variation in mortgage quality that is due to observable characteristics from "excess" variation in default behavior. The central test in this paper is whether excess default is related to time-to-sale.

Using data on mortgages securitized in the non-agency, private-label securitization (PLS) market, we find a clear negative relationship between time-to-sale and the component of mortgage performance that is not explained by 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 meaningful difference, as it is approximately 30 percent of the average default rate in our sample (16 percent). Our empirical results are robust to different ways of defining default, alternative default horizons, different specifications, and, most importantly, to alternative data sources that differ in their representativeness of the PLS market. These results are consistent with previous studies that have found an important role for private information in the PLS market (Demiroglu and James (2012a), Jiang et al. (2014b), Griffin and Maturana (ming), and Piskorski et al. (2015)).

The results on *ex post* default are in contrast to those using *ex ante* measures of risk. In fact, we find no relation between summary measures of credit risk at origination and time-to-sale, even though credit risk at origination is strongly related to performance. Put differently, while unobserved quality is related to delay of trade, observable risk measures are not.

The lack of relation between observable risk and time-to-sale speaks to the interpretation of our results in the case that the buyers of mortgages (the issuers) have more information that we do as the econometrician. In fact, the validity of our tests does not rely on observing *all* information that is common to buyers and sellers in the market. Our tests require a weaker assumption, namely that credit quality as we measure it be an unbiased estimate of quality measured by issuers using their full information set. If this is the case, the results using our observable risk measure provide a good approximation of the unobserved relation between credit risk as measured by issuers and time-to-sale.

We also find no evidence of a negative relationship between time-to-sale and conditional 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, where automatic underwriting virtually eliminates the role for asymmetric information between buyers (GSEs) and sellers (originators). Stanton et al. (2014) and Stanton and Wallace (2015) show that repo agreements and warehouse lines of credit with very short maturities were a large funding source in the PLS market, which would also severely limit the ability of originators to delay the sale of mortgages. For the purposes of our tests, some originators in the PLS space must have the ability to hold on to mortgages and delay trade, even if others were limited by contractual features due to their funding sources. Even though we find that the majority of loans in the PLS market were sold 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 are sales past this time period (up to 9 months after origination).

We then turn to a secondary source of detailed loan-level data (LoanPerformance) commonly used by participants in the PLS market as both a robustness test of the main results and to implement a series of cross-sectional tests to determine whether signaling is the more likely interpretation of our findings. We confirm the negative relation between time-to-sale and ex-post default risk and, in addition, we find 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. The previous literature has found private information to be especially important among low documentation mortgages, which lends further credence to an adverse selection, signaling interpretation.[3]

The LoanPerformance data also allows us to include originator fixed effects, which helps address the concern that funding sources (in particular 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 (in particular independent mortgage companies, as pointed out in Stanton et al. (2014) and Ganduri (2015)) relied almost exclusively on these types of funding sources, that variation is accounted for in these specifications. We find similar results to the baseline regressions.

As a final test, we separately estimate the correlation 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 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.

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 related work, Spence (1973) shows informed agents can take actions to credibly reveal their private information that lead to a separating equilibrium. This insight was first applied to financial markets by Leland and Pyle (1977) who

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

showed the issuers of IPO's can signal information by retaining an equity stake in the IPO. DeMarzo and Duffie (1999) use the equilibrium relationship 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 create tranches to minimize signaling costs.

While retention is a common signaling device pointed out in the above literature on adverse selection, delay of trade serves the same function in a dynamic setting. Janssen and Roy (2002) show that, in a durable goods market in which sellers have private information, a market mechanism emerges in which prices and the quality of goods increases over time. This property of market equilibrium is the so-called *skimming property.* This property has been shown to be a general feature of equilibrium in dynamic models of adverse selection. For example, Daley and Green (2012) consider a model in which an informed party sells an asset to a market of uninformed agents. When news about asset quality arrives over time, sellers with high value assets wait to trade allowing market participants to infer that delayed trade is associated with higher value assets.

This paper also contributes to the empirical literature on the effects of asymmetric information. An important paper in this literature is Garmaise and Moskowitz (2004) who use commercial real estate transactions to test a number of theories of asymmetric information, including the core prediction of Leland and Pyle (1977) and DeMarzo and Duffie (1999) 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) 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.

The paper is organized as follows: Section 2 presents a simple model of mortgage sales with asymmetric information. Section 3 provides a brief background of the structure of the U.S. mortgage market and some important institutional features of the PLS market. Section 4 describes our test of signaling, and Section 5 provides a detailed discussion of the data sources used in the analysis. Section 6 presents the results and Section 7 concludes.

2 A Simple Model of Signaling Through Delayed Trade

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 issuers of MBS (the buyers). 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 a 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 buyers are risk neutral, there are gains from trade generated by a difference in discount rates used by the two classes of agents. Specifically, the seller discounts cash flows at a rate γ, while the buyers discount cash flows at rate $r < \gamma$. This difference in discount rates proxies for a difference in the investment opportunity set between the seller and the buyers. Indeed, the seller has the technology to originate mortgages, while buyers can only purchase mortgages in a competitive market once they have already been originated. We assume that default is publicly observable, so that if a mortgage defaults before the seller has sold it to the buyers, no sale will occur. In choosing when to sell the mortgage, the seller will take some market price function $P(t)$ as given.

Note that the lowest possible value of a mortgage to buyers is

$$p_h = E\left[\int_0^t e^{-rs}\mathbb{1}(s \le \tau)cds|\lambda_h\right] = \frac{c}{r+\lambda_h},$$

while highest possible value is

$$p_\ell = E\left[\int_0^t e^{-rs}\mathbb{1}(s \le \tau)cds|\lambda_\ell\right] = \frac{c}{r+\lambda_\ell},$$

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

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

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

An important feature of the sellers payoff function is the so-called single-crossing property; fixing a price p, delaying trade is less costly for better (lower default risk) type seller. Intuitively, the lower is the default risk, the greater is the private value of the cash flows that accrue to the seller from the mortgage prior to the sale and the greater is the probability that the mortgage will remain current so that it can be sold at some future date. This feature of the model gives rise to the common *skimming property* present in much of the literature on dynamic trading and adverse selection,[4] and is more broadly related to the literature on costly signaling in adverse selection.[5] In our model, the skimming property can be expressed

[4]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)).

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

as follows. For a given price function $P(t)$, better sellers will wait weakly longer to trade, thus a delay in trade can act as a signal of quality.

A perfect Bayesian 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 for a mortgage sold at time t such that

1. Seller optimality:

$$T(\tilde{\lambda}) \in \arg\max_{t} U(\tilde{\lambda}, t, P(t),)$$

2. Zero profit for the buyers:

$$P(T(\tilde{\lambda})) = E\left[\frac{c}{r+\tilde{\lambda}} | T(\tilde{\lambda})\right].$$

We call an equilibrium separating if $P(T(\tilde{\lambda})) = \tilde{\lambda}$.

We will focus on characterizing a separating equilibrium. Although other equilibria, for example 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 sellers 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 get 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 trade 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 seller type 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 following key properties of the functions $\lambda^*(t)$ and $T^*(\lambda)$.

1. The default risk of the mortgage decreases with time-to-sale, that is

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

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

2. The maximimum 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, i.e. when the uncertainty about mortgage default risk is greater, leads to longer delays in sale.

3 Background on U.S. Mortgage Market

Broadly speaking, mortgage lending in the U.S. takes two forms: the originate-to-hold (OTH) model, in which mortgages are originated and then held on the balance sheet of a bank, and the originate-to-distribute (OTD) model, where loans are originated and then sold to an institution that typically securitizes them, and sells securities backed by the cash flows of mortgage pools to investors around the world. The OTH model has given way over time to the OTD model as the fraction of loans securitized in the U.S. rose from approximately 61 percent in 2001 to almost 75 percent in 2007 on the eve of the financial crisis, and has continued to rise in the post-crisis period, reaching almost 85 percent in 2010.[6] Before the outset of the mortgage foreclosure crisis in 2007, loans were securitized by either the housing government sponsored enterprises (GSEs), Fannie Mae and Freddie Mac, or by private financial institutions, known in the industry as private-label issuers.[7]

Our primary focus in this paper is on the latter group of loans that were sold and then securitized by private-label 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 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. There is significant variation in the funding and operational models of mortgage originators in the PLS space, including independent mortgage companies, affiliated mortgage companies

[6] 2011 Mortgage Market Statistical Annual.

[7] There is a third segment of the OTD market comprised of FHA (Federal Housing Administration) loans that are securitized by Ginnie Mae, an explicit government agency. We do not consider FHA loans in this analysis.

and others. We refer the reader to Stanton et al. (2014) and Ganduri (2015) for detailed descriptions of the structure of the market.

We focus on loans sold into the PLS market for a couple of reasons. First, there are many recent papers in the literature that have documented a significant amount of private information in these markets, and that originators were at least partially aware of unobserved quality.[8] In contrast, private information likely plays a much less important role in the agency securitization market, where the GSEs provide specific parameters regarding the underwriting criteria that they will accept, and agree to purchase (usually through an automated process) all loans that satisfy those criteria.

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

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 exceeded the conforming loan limits and were thus not eligible to be securitized by the GSEs in the agency market. [9] 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 were also more likely to finance investor or vacation home properties. Alt-A PLS included a mix of loans

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

[9]In order 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 else have equivalent credit enhancement (e.g., private mortgage insurance).

above and below the conforming loan limit. Finally, the collateral underlying subprime private label securities is made up by loans 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. 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 LoanPerformance database, also described below) includes loans from the subprime and Alt-A segments of the market.

4 Testing for Dynamic Adverse Selection Using Mortgage Data

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

We implement a strategy similar to Adelino et al. (2014) that uses conditional measures of loan performance to isolate aspects of loan quality that are unobservable to MBS issuers at the time of purchase, but are correlated with the originators' information set (and, 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 were readily available to issuers and institutional investors in the MBS market. Our empirical specifications take the following general form:

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

where i indexes the individual mortgage, j indexes the the geographic area in which each mortgage is originated, and t indexes the horizon over which we calculate default rates. X_{ijt}

is a vector of mortgage-level control variables that includes 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, namely that (i) the seller's private information, I_{seller}, is correlated with loan quality after accounting for underwriting characteristics, i.e.

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

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

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 "true credit risk. Additionally, we assume that $X_i \subseteq I_{buyer} \subset I_{seller}$, i.e. both buyers and sellers information sets include the mortgage characteristics we observe, and sellers have some private information about the loans. In such a setting, we can measure the relation between time-to-sale and credit risk using our measure of risk (which is assumed to be unbiased). To the extent that credit risk is the only variable that is systematically related with time-to-sale, the additional information in I_{buyer} is simply providing more precision for measuring credit risk, but should not change the direction of that 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 available, 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 I_{buyer} only includes the publicly available mortgage underwriting data we use in the regressions.

4.1 Measurement

We consider two different default horizons, 36, and 60 months, in our primary specifications, measured relative to the month of loan origination.[10]. 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} accounts for a large subset of the information held by the buyers of mortgages at the time of the sale. According to Stearns (2006), all issuers and most PLS investors had access to detailed information at the loan-level including data fields such as FICO score, combined LTV ratio, documentation type, occupancy type, loan purpose (refinance or purchase), property type, loan size, amortization schedule, interest rate type (ARM vs. FRM), and information on the geographic location of the property.[11] We choose our vector of control variables to include these variables, as well as some variables that measure ex-post conditions in the local housing market, which likely influence loan performance.

Specifically, our covariate set includes the combined loan-to-value (LTV) ratio, the original loan balance, the original interest rate, the borrower's credit score, the original maturity of the loan; and indicator variables for low documentation loans, interest-only loans, balloon loans, negative amortization loans, residence status (owner-occupied, investor/vacation home), loan purpose (cash-out refinance, other refinance, purchase), property type (condominium, multi-family, single-family), and the existence of a prepayment penalty.[12] We also include the county-level unemployment rate and the level of the house price index at the

[10]We have also tried a shorter horizon of 24 months, which did not make a material difference.

[11]This contrasts with the agency market, as the GSEs, in part due to the fact that 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."

[12]We estimate a fairly saturated model by including many categorical variables for the continuous variables in our covariate set like credit scores and LTV ratios. The appendix contains a list of the exact variables that we include in our covariate set.

time of origination (normalized by setting the index value for January 2000 to 100 for each county), as well as the changes in these series from the time of issuance through the end of the default horizon. In addition we include a full set of state-level fixed effects, and fixed effects corresponding to the year-quarter of origination as well as the year-quarter of loan sale.[13] Additional indicator variables are included whenever there are missing observations for any of the controls.

5 Data

In this section we describe the two loan-level datasets used in this paper. While both are similarly structured panel datasets that contain detailed information about contract characteristics and monthly loan performance, there are important differences in the scope of their coverage and in some of the underlying variables that produce advantages and disadvantages in the context of our analysis.

5.1 LPS Data

Our primary dataset comes from Lender Processing Services (LPS). The LPS dataset covers between 60 and 80 percent of the U.S. mortgage market, and contains detailed information on the characteristics and performance of both purchase-money mortgages and refinance mortgages. It includes mortgages from all segments of the U.S. mortgage market: PLS or non-agency securitized loans; loans purchased and securitized by the GSEs; and loans held in lenders' portfolios. 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 either paid off voluntarily or involuntarily via the foreclosure process. We focus on

[13]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, and found that the results were largely unaffected. Since 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.

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 makes it possible to explicitly identify if and when a loan is sold to a PLS issuer or to a GSE to be securitized. Since the purpose of this paper is to test for whether there is a positive correlation between the quality of an asset (observable only to the seller) and the time that it takes to sell the asset, we focus only on loans that are sold. Thus, we focus on loans that we identify as being transferred from a banks' portfolio to a PLS issuer or to one of the GSEs. Many loans in our LPS sample of sold mortgages begin in the portfolio of the mortgage originator and then are sold to a PLS issuer or GSE at some point after origination. In contrast, many loans in our sample are categorized by the "investor type" variable as being in a PLS or GSE security in the month of origination, in which case we assume they were immediately sold.

We adopt a few sample restrictions in our analysis in order to ensure that our results are not driven by factors that have little to do with the theory presented above. We consider only first lien mortgages originated in the 2002 – 2007 period that were sold to PLS issuers or to the GSEs.[14] We only keep loans originated in the 50 U.S. states, and we restrict the sample to loans that enter the dataset in either the same month of origination or in the month following origination.[15] In addition to these sample restrictions, 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

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

[15]That is, we throw out loans "seasoned" more than 1 month. We do this for two reasons. First, for the seasoned loans that are identified as being in PLS securities when they first enter the dataset, we are unable to determine the exact month in which they were sold. Second, since we do not observe the payment history of seasoned loans before they enter the dataset, we are unable to determine their default status in the months before they enter the dataset. The vast majority of LPS loans are seasoned 0 or 1 months, so we decide to keep loans that were seasoned 1 month. Below, we describe how we treat these in the empirical analysis.

of each respective distribution.[16]

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 look at sales to both PLS and GSE issuers. However, there are also a few important drawbacks. The biggest problem with the LPS data in our context is the lack of information on the exact identity of the financial institution that originates the mortgage. Ideally, we would want to control explicitly for the identity of the originator, as this would eliminate 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 discussed below.

5.2 CoreLogic Data

Our second source of mortgage data comes from CoreLogic's LoanPerformance (CL) PLS database, which covers virtually the entire subprime and Alt-A segments of the PLS 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. One of the main advantages, however, of using CL data is its representativeness of the PLS market.[17]

The CL database includes virtually the same mortgage and borrower characteristics (at the time of loan origination) as the LPS database, but, importantly, for a sample of CL loans (about 50% of the entire database) identity of the originating institution is provided, which allows us to examine the relationship between time-to-sale and ex-post performance using loans originated by the same lender. In addition to the identity of the originator, CL

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

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

also provides information on the identity of the mortgage servicer, as well as information on security identifiers (CUSIPs) and deal identifiers, which allows us to obtain information on the identity of the securitizer (issuer) for most loans in the sample. There is some uncertainty about whether the originator field in the LoanPerformance 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 that is in charge of 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 since it has more interaction with the mortgage borrower. To verify that the originator field in LoanPerformance indeed corresponds to the lender of record, we matched our LoanPerformance 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 LoanPerformance 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 LoanPerformance originator field. The remaining 17% are either cases in which LoanPerformance is reporting an entity other than the lender of record (most likely the aggregator) or are bad matches (there is the potential for significant matching error 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 LoanPerformance data. There is one additional difference between the CL and LPS datasets that requires some discussion. The timing for when a loan enters each dataset is different. In the LPS dataset we observe most loans from the month of origination, and can directly observe the month in which they are sold out of a bank's portfolio to PLS issuers or to the GSEs. In the CL dataset we compute time-to-sale as the difference between the date of issuance of the mortgage-backed security in which the loan is included and the reported month of origination of the mortgage.[18] In most cases, loans are transferred from the warehouse into the special purpose vehicle at the time of issuance, and so the date of issuance is a good proxy for when the mortgage credit risk is transferred from the originator to the issuers.

5.3 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. For example, 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 2 months after origination). While there are some sales that occur several months after origination, the number of sales drops off very quickly with time for both loan types. In implementing our tests, we would like to restrict our analysis to loans that are originated with the intent of being sold, and are concerned that the loans sold long after they were originated may not have been made with the intent of being sold (or are fundamentally different on some other dimension that is unobservable to us). Furthermore, the combination of the small number of loan sales in later months and the large number of control variables in the empirical models results in low statistical power. For these reasons, we impose one last sample restriction, which is a maximum threshold for the number of months between origination and sale. We base this threshold on the PLS sample, since that is our main focus in the analysis, and choose a threshold value of 9 months, based on the simple observation that approximately

[18]Loans typically enter the CL dataset on the issue date, so we do not see the full performance history, unless the month of origination and issuance coincide.

97% of loan sales happen within 9 months in that market.[19] 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 2 we display summary statistics for many of the control variables in the empirical models. The table displays statistics for both 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 compared to GSE loans. For example, there were 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 7 displays the distribution of months-to-sale in the CL dataset, while Table 8 provides some basic summary statistics. The first notable observation is that there are many more PLS loans in CL compared to LPS.[20] The second thing to note is that the distribution of months-to-sale in CL is similar to LPS, although there are a few subtle differences. In both datasets over 90% of loans that end up in PLS 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%). There are more dramatic differences in the summary statistics between the two datasets. 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 due to the fact that 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 average underwriting characteristics, the sample of PLS loans in CL appears to be significantly riskier than the

[19] We have experimented with higher thresholds such as 12 months, with little affect on the estimation results.

[20] The LPS sample size of 5.3 million loans listed in the tables understates the total number of PLS loans as there are some seasoned mortgages that we eliminate from the sample due to our sample restriction of only including loans for which we see a full history of performance. There are actually more than 7 million PLS loans originated between 2002 and 2007 (inclusive) in the LPS database.

LPS sample.

Unlike LPS, in CL we can distinguish between the subprime and Alt-A markets.[21] We display the distribution of months-to-sale (Table 7) and the summary statistics (Table 8) for the subprime and Alt-A loans separately. The tables show that the sample of Alt-A loans in CL looks 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 are closer to the LPS sample than the subprime loans.

6 Results

In this section we present our results on the empirical relationship between time-to-sale and conditional default rates of securitized mortgages. We present results for loans sold into both the PLS (non-agency) and GSE (agency) segments of the secondary market, as we believe these markets are likely characterized by different levels of private information and adverse selection. We also present results for different segments of the PLS market (i.e. subprime vs. alt-a), which we believe may differ in the degree to which private information plays a role.

First, in order to operationalize equation 5 it is necessary to take a stand on how the time-to-sale variable, Months-to-Sale$_{ij}$, is specified. The prediction from the model is simply that default risk decreases with time-to-sale, $\frac{d}{dt}\lambda^*(t) < 0$. Thus, we start with a simple, linear specification of the time-to-sale variable, so that Months-to-Sale$_{ij}$ takes values from 0 to 9 depending on the exact month a loan is sold relative to the month of origination.[22]

[21] There is a servicer-provided field in LPS that distinguishes Grade "A" loans and Grade "B" and "C" loans, with the grades supposedly corresponding to different levels of credit risk. We include the variable in our covariate set in the analysis. However, 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.

[22] Since we cannot distinguish between loans sold in the same month of origination and loans sold in the month following origination in the LPS dataset, we combine them and allocate a value of 0 for the linear Months-to-Sale$_{ij}$ variable. We then subtract 1 for loans with values of time-to-sale between 2 and 9 (inclusive), so that the linear Months-to-Sale$_{ij}$ variable takes values between 0 and 8 in the LPS sample.

We then change the specification to include a quadratic term, Months-to-Sale$_{ij}^2$ in order to determine if there is a non-linear relationship between mortgage credit risk and time-to-sale. Finally, we specify the Months-to-Sale$_{ij}$ variable in a non-parametric manner by including a separate indicator variable for each value of the variable, with the base case (i.e. the indicator variable left out of the regressions) being loans sold in the month of origination.[23]

The second issue that requires some discussion is the notion that the implementation of the empirical test described in section 4 is truly able to capture signaling in the mortgage market, rather than some other, unrelated artifact in the data. One potential concern is the role of early payment defaults in generating a mechanical relationship between time-to-sale and ex-post default risk due to institutional features of the PLS market. Specifically, loans in delinquency, or loans that have been delinquent at some point in the recent past, are unlikely to be sold into a securitized pool of loans. This could create a negative relationship between time-to-sale and default that is independent from a mechanism involving private information and signaling. To see this issue more clearly, assume that loans are randomly warehoused for different periods of time (and thus sold randomly at different ages), with no role for private information or adverse selection. In this world, loans that default early in their lives while still awaiting to be sold cannot be sold, and would not make it into our sample. The only defaulted loans in our sample would be mortgages that default after being sold. Thus, there could be a sample selection issue that increases in severity as time-to-sale increases. Loans that happen to be sold very quickly would be representative of the population of eligible loans in terms of default risk, whereas loans that happen to take a longer time to sell would be of higher average quality than the population of eligible loans. This would create a mechanical negative relationship between time-to-sale and default that has nothing to do with a signaling motive. In order to address this issue, we estimate our models on an alternative sample (which we refer below to as the "restricted" sample) that excludes all loans that became delinquent within 9 months of origination. For this sample,

[23]Since we cannot distinguish between loans with values of 0 and 1 for months-to-sale, our base case for the regressions estimated on LPS data includes both types of loans.

we only consider defaults that occur between the 9 month mark and the specific default horizon (36 and 60 months). This forces the sample of sold loans to be homogeneous in terms of early payment defaults across the time-to-sale distribution, and the results *cannot* be explained by the mechanical problem described above.

While this correction directly addresses the mechanical issue discussed above, there are a few significant drawbacks. First, loans that default early may still be sold, in which case the mechanical effect is not severe, and the correction would simply be throwing away information. Second, and more importantly, it may be that the private information, signaling story 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 months of origination, this "correction" to the sample effectively eliminates the variation we are most interested in. For this reason, we choose to display the correction as a robustness check below rather than adopt it as our baseline specification.

In addition to this robustness check, we also run tests that emphasize the need to distinguish between observable and unobservable risk by estimating the relationship between time-to-sale and ex-ante, observable default risk. The idea behind this exercise is to determine what, if any, relationship exists between time-to-sale and predictable default risk based on observable underwriting characteristics. As we discuss above, if the relation based on observable information is similar to what we attribute to unobservable information this would limit our ability to draw inferences from our tests. We discuss the details of this exercise below.

6.1 LPS Results for PLS Loans

Panel A of Table 13 displays results for the linear and quadratic regression specifications estimated on our sample of PLS 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) and two different default horizons (36 months and 60 months relative from the month of origination).[24] The results show a negative, statistically significant relationship 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 very consistent over the different horizons and default definitions.

The results for the quadratic specifications suggest that the relationship between time-to-sale and default rates is non-linear. The positive coefficient on the quadratic terms implies that for small values of time-to-sale the relationship is negative, but that for higher values of time-to-sale the relationship becomes significantly less negative and even turns positive.[25] We explore this non-linearity in greater detail in Table 4, where the results from the non-parametric specification are displayed. Columns 1-2 and 5-6 display the non-parametric results for the different combinations of the default definitions and horizons. The results suggest that average default rates are decreasing in time-to-sale until Months-to-Sale$_{ij} = 5$, at which point average default rates begin to moderately rise. 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 are lower by 3 - 4 percentage points on average. Again, the estimation results are quite consistent across the alternative default definitions and horizons.

Panel B of Table 13 and columns 3-4 and 7-8 in Table 4 display the same set of results for our restricted sample, where we throw out all loans that default within 9 months (inclusive) in order to address the potential sample selection bias that we discussed above. There is virtually no difference in the results for the linear specification of the Months-to-Sale$_{ij}$ variable, but there are a few subtle differences for the non-linear specifications. From the results

[24]In the Appendix we display a set of regression results that includes the coefficient estimates for most of the variables in our covariate set. Most of the estimates are consistent with the previous literature on mortgage default.

[25]The quadratic term begins to dominate the linear term when time-to-sale reaches 10 months, which is beyond the highest value for time-to-sale in our sample (9 months).

of the non-parametric specification we see that this sample restriction slightly mitigates the negative relationship between time-to-sale and default for loans sold within 4 months. However, the sample restriction appears to have the opposite effect for loans sold later as the coefficient estimates associated with loans sold between 7 and 9 months after origination become slightly more negative. This pattern is confirmed in the quadratic specifications in Table 13 as the coefficients on the linear terms become less negative while the coefficients on the quadratic terms become less positive. Overall, the sample correction appears to have a very minor effect on the results, which suggests that sample selection bias is not an important issue.

In the top left panel of Figure 1 we plot the estimated relationship between time-to-sale and ex-post PLS default risk from the non-parametric specification in column (3) of Table 4 (60+ DQ, 36-month horizon, restricted sample). The plot includes 95% confidence intervals to show the precision of the estimates. There is a clear negative trend until month 6 at which point the coefficient estimates flatten out. The estimates associated with the first 4 months are much more precise compared to the last 5 months due to the much larger sample size of loans sold early in their lives. Overall, the results in Tables 13 and 4 provide evidence of a negative relationship between time-to-sale and (conditional) ex-post default risk, which supports the existence of a signaling motive in the PLS market. Furthermore the results are robust to potential sample selection bias generated by early payment defaults.

6.2 LPS Results for GSE Loans

Tables 5 displays results for our sample of loans sold to the GSEs. The table displays results for the linear and quadratic specifications and is structured in the same manner as Table 13, which displayed the PLS results.[26] There is very little evidence of any relationship between time-to-sale and ex-post default risk in the GSE segment of the market. We plot the estimated relationship from the non-parametric specification in the top right panel in Figure

[26]For the sake of brevity we do not include a separate table containing estimation results for the non-parametric GSE specifications.

1 (the same specification as the one used to construct the PLS graph in the top left panel). The first thing to note from the plot is the stark difference in the pattern relative to 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 basically zero until the very end of the distribution when they begin to fall. In addition, the GSE estimates are much more precise, on average, compared to the PLS results due to the much larger sample size. However, the PLS estimates are fairly precise for the low values of time-to-sale where the downward trend is the most pronounced, while the GSE estimates become much more imprecise toward the end of the time-to-sale distribution when the sample size becomes significantly reduced. In general, the GSE results are consistent with our hypothesis that private information is much less of an issue in the agency market compared to the PLS market.

6.3 Ex-Ante Analysis

In this section we attempt to estimate the empirical relationship between time-to-sale and *ex-ante* credit risk in order to compare and contrast it with our results above on the relationship between time-to-sale and *ex-post* credit risk that conditioned out the set of observable underwriting characteristics. To do this, we construct ex-ante default probabilities for each loan using all of the data available in LPS in a manner that is similar in spirit to the method used in Ashcraft et al. (2010). The idea is to forecast mortgage default using only performance information available at the time of origination (i.e., from the past performance of previously originated loans).

We choose a 36-month horizon to forecast defaults in order to maintain consistency with our results above. We begin by taking each loan in our LPS sample, and determining the quarter in which it was originated. We then take all loans that were originated between 48 months and 36 months before that quarter, and track those mortgages over the subsequent 36 months, creating indicator variables that take values of one if the mortgage ever becomes

60 days delinquent at any point during the 36 month period. We then estimate a discrete choice model (linear probability and logit) using variables that are available in LPS to predict the default variable. The regressions are estimated each quarter over the period 2002–2007 and include virtually the same set of covariates that were included in the ex-post default risk regressions described above. We take the estimated coefficients from these loan-level credit risk models and apply them to the characteristics of the loans originated in the current quarter to create 36-month, loan-level, default probabilities. This leaves us with a set of ex-ante default probabilities created using only information available at the time in which the loans were originated.

We then take those ex-ante default probabilities and substitute them into equation 5 in order to estimate the relationship between time-to-sale and *observable* default risk. We display the estimation results in the lower two panels in Figure 1. The lower left panel displays the relationship between time-to-sale and ex-ante, default risk for PLS loans, while the lower right panel displays the relationship for GSE loans. The PLS results suggest that loans sold later are slightly *more* risky based on 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 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. This pattern is in stark contrast to the estimated relationship between ex-post default rates and time-to-sale in the PLS market (top left panel in Figure 1), and provides some reassurance that our ex-post conditional default measures are doing an adequate job in purging predictable default risk. The horizontal line displayed in the lower right panel in the figure implies that there is no relationship between predictable default risk and time-to-sale in the GSE market.

6.4 CoreLogic Results

Table 9 displays the core set of results on the relationship 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, which allows us to account for any variation generated by heterogeneity across originators. In Table 9 we present results corresponding to our parametric specifications of equation 5 and focus on a default horizon of 36 months and a default definition based on 60+ days delinquent. In Panel A we display results from a specification that does not control for originator heterogeneity, and thus, these are results are directly comparable to the LPS results displayed in Table 13. In Panel B, we include, for each specification, a full set of originator fixed effects. Information on the originator is available for slightly more than half of the loans in the CL dataset, so we focus our analysis on this subsample.[27]

The estimation results reported in Table 9 show a statistically significant, but slight, negative relationship between ex-post default risk and time-to-sale, which is not very sensitive to the inclusion of lender fixed effects. According to the linear specification results (column 1) an increase in time-to-sale by one month is associated with 0.28 – 0.36 percentage point increase in average default rates. While the magnitudes are significantly smaller than the LPS results discussed above, the pattern is quite similar as evidenced by the estimates from the non-parametric specification, which are displayed in the top panel of Figure 2. Average ex-post default rates decline over the first half of the time-to-sale distribution and then flatten out over the second half of the distribution in a similar manner to the LPS results plotted in the upper left panel of Figure 1.

6.4.1 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

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

to see a larger role for signaling unobservable mortgage quality in the Alt-A segment of the PLS market, since this was largely comprised of low documentation mortgages. Table 8 shows that over 70 percent of Alt-A mortgages were less than fully documented compared to 35 percent of subprime loans.

Table 9 displays the parametric specification results from separately estimating regressions for the subprime and Alt-A segments of the PLS market (columns 3-6), and the bottom panels of Figure 2 plots the results for the non-parametric specifications. The differences between the subprime and Alt-A results are fairly striking, and help to explain where the differences between the LPS and CL results are likely coming from. There is essentially no relationship between ex-post default risk and time-to-sale among subprime PLS loans (Panel C), while there is a fairly significant, negative relationship among Alt-A loans (Panel B). 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 2 and 8) it appears as though 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.

6.4.2 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 as well as for our subprime and Alt-A samples separately. The results are displayed in Table 10, with Panel A containing results for the parametric specifications and Panel B containing results for non-parametric specifications. Figure 3 plots the non-parametric results with 95 percent confidence intervals to provide a sense of the statistical significance between the low documentation and full documentation estimation results.

The results are mixed. In the sample of all PLS loans (subprime and Alt-A combined), there does appear to be a stronger negative relationship between time-to-sale and default for low documentation loans compared to full documentation loans. This negative relationship is approximately twice as large (in absolute value) in the sample of low documentation PLS loans (columns 1-2). However, Figure 3 shows that the difference in this relationship between the two types of loans is not statistically significant at conventional levels. Furthermore, based on the results displayed in Table 10 (columns 3-6) and Figure 3 (Panels B and C) there are essentially no differences between full documentation and low documentation loans within the subprime and Alt-A subsamples.

6.4.3 Affiliation Results

In this section we test whether an affiliation between the originator (seller) and issuer (buyer) plays a role in the relationship between time-to-sale and default risk. There are direct relationships between many issuers and originators in the PLS market. 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 in the case of an originator and issuer who are independent entities.[28] Thus, if signaling is driving our results, we would expect a weaker negative relationship 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

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

(PSA) associated with the PLS deals.[29]. We focus on only loans that are in deals in which either all loans were made by affiliated originators or all loans were made by unaffiliated originators.[30] Table 11 and Figure 4 displays the results. As in our analysis of documentation status above, we stratify by affiliation status in our sample of all PLS loans as well as in our Alt-A and subprime samples separately. While the results are different for the three samples, overall, the negative correlation between time-to-sale and default risk does appear to be weaker when the originator and issuer are affiliated entities. In the full sample, the correlation is more than twice as large for unaffiliated compared to affiliated issuers and originators (columns 1 - 2 in Table 11). Panel A in Figure 4 shows that this difference is statistically significant for loans sold within the 4 months of origination.

The difference in the relationship 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 6 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 B in Figure 4 shows that this difference is highly statistically significant over the entire distribution of time-to-sale. Finally, we find no differences between affiliated and unaffiliated originators in the subprime segment of the PLS market.

6.4.4 Early Prepayment Analysis

Until this point we have used default as a proxy for loan quality. We believe that this is a reasonable strategy since default is an unequivocally negative outcome from the perspective of an MBS investor. However, there are other types of negative outcomes that may be

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

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

relevant in our context, and in this section we will consider one of these alternatives, namely early prepayment risk. In addition to default, residential mortgages contain a prepayment option that allows borrowers to fully repay the outstanding principal balance of their loans before the loan reaches full maturity. Since the exercise of the prepayment option reduces the expected future cash flow of a mortgage, it also reduces the value of a mortgage security, and thus, can be considered a negative outcome from the perspective of the average MBS investor. 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 time 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, prepayments of adjustable-rate mortgages are 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, 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. The 2/28 and 3/27 hybrid ARMs were characterized by this initial period in which the interest rate was fixed, after which the interest rate would reset to a new level and begin to fluctuate, tracking a market interest rate (such as the 6-month LIBOR or the 10 year Treasury rate). Since 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, which provided further incentive for borrowers to wait until the reset date to exercise their prepayment option.[31]

[31] For an excellent reference on the PLS market in general, and especially for empirical analyses on the prepayment and default behavior of various types of PLS loans, we refer the reader to Kramer and Sinha

For these reasons, the expectations of market participants were that many 2/28 and 3/27 ARM prepayments would occur on or immediately after the reset date. Therefore, prepayments that occurred significantly before the reset date can be viewed as particularly negative outcomes. 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 prepaid several months before the interest rate reset month.[32] The 2/28 and 3/27 ARM products were by far the most popular adjustable-rate product in the PLS market, accounting for approximately 75% of all subprime and Alt-A PLS ARMs combined.[33] We consider two cutoffs of 6 months and 9 months before the reset date in defining our early prepayment indicator variables. The reason for these threshold choices is that the most common type of prepayment penalty associated with these mortgages was 6 months of interest on 80% of the principal amount prepaid. Thus, even an ARM that carried this prepayment penalty that prepaid more than 6 months before the reset date would generate lower cash flow levels compared to a loan that prepaid at the reset date, and thus can be considered as a negative outcome for a PLS investor.

Table 12 contains the results of the early prepayment analysis. Panel A displays estimation results that correspond to the parametric (quadratic) specifications while Panel B displays results for the non-parametric specifications. We show results for various corrections for the potential sample selection issue that we discussed above in the context of the LPS default analysis. Recall that our correction was to throw out all defaults that occurred within our sale period (up to 9 months after origination). We found that such a correction had a minimal impact on the results, however, the issue may be more problematic in the context of prepayment, since, by definition, a loan that is prepaid cannot possibly be sold.

(2006).

[32]We eliminate defaults from our analysis in order to isolate voluntary prepayment risk. From our analysis above we already know that there is a negative correlation between time-to-sale and (conditional) default risk. By throwing out defaults, we ensure that the results are not driven by involuntary prepayments.

[33]These products were mostly found in the subprime segment of the market, although there were a non-trivial number originated in the Alt-A segment. Many (about one-third) of Alt-A ARMs had a one month "teaser" rate that reset to a higher adjustable rate in the second month, and thus did not have prepayment profiles driven by reset concerns. See Sengupta (2010) for a detailed discussion of the composition of loans in the Alt-A and subprime PLS markets.

At the same time however, the bulk of our sample is comprised of 2/28 hybrid ARMs, which means that the early prepayment period that we are considering is often within 15 months and 18 months of origination, respectively. Therefore, throwing out all loans that prepaid in the first 9 months eliminates a significant amount of the early prepayment variation in our sample, and to the extent that investors are especially concerned with prepayments within the first year or so of origination, such a restriction could serve to attenuate the true signaling effect rather than simply correcting sample selection bias. For this reason, we display results for both a 6 and 9 month early prepayment cutoff for various sample restrictions: no restriction in columns (1) and (2), a 3-month restriction (i.e. throwing out all loans that prepay within 3 months) in columns (3) and (4), a 6-month restriction in columns (5) and (6), and finally the full 9-month restriction in columns (7) and (8). Table 12 clearly shows a negative relationship between time-to-sale and early prepayment risk. As months-to-sale increases, 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 6 months after origination are approximately 6-7% less likely to prepay early compared to loans sold immediately, while loans sold 9 months after origination are about 10-11% less likely to prepay early. The extent of the sample restriction does have a significant impact on the results. The negative relationship remains pronounced in the cases where we apply partial corrections and throw out all prepayments that occur within 3 months and 6 months of origination, respectively (columns (3) - (6)), but the most severe restriction (throwing out all prepayments that occur within 9 months of origination) significantly flattens the slope of the negative relationship between months-to-sale and early prepayment.

In general, we believe these results on the correlation between time-to-sale and early prepayment of hybrid ARMs in the PLS market are consistent with our default analysis above, and support the existence of a motive to delay the sale of loans in order to signal their higher quality to PLS issuers and investors. 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 states of the world, and thus was an important consideration for mortgage investors.

7 Conclusion

A general feature of dynamic models of adverse selection is that the prices and (unobserved) quality of goods increases over time. This paper provides the first 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 predictable by buyers using ex-ante, observable information on underwriting characteristics. Furthermore, the positive relationship between time-to-sale and mortgage performance is not present in the agency securitization market where adverse selection between originators and issuers is not as serious of a concern. This estimated correlation appears to be strongest in the "Alt-A" segment of the PLS market, where most loans were underwritten with less than full documentation of income and/or assets, and thus, is consistent with previous studies that have found an important role of private information among low documentation mortgages.

Taken together, we believe that 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 trade.

References

Adelino, M., W. S. Frame, and K. Gerardi (2014). The effect of large investors on asset quality: Evidence from subprime mortgage securities. *Available at SSRN 2472132*.

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

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

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. K. Purnanandam (2014). Design of financial securities: Empirical evidence from private-label rmbs deals. *Available at SSRN 2022380*.

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

Coase, R. H. (1972). Durability and monopoly. *Journal of Law and Economics 15*, 143–149.

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.

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. (2015). Repo regret?

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

Griffin, J. M. and G. Maturana (Forthcoming). Who facilitated misreporting in securitized loans? *The Review of Fianncial Studies*.

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.

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.

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

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

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*.

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.

Table 1: Distribution of Months-to-Sale in LPS Sample

	PLS Loans		GSE Loans	
Months-to-Sale	# Loans	Cumulative %	# Loans	Cumulative %
0	1,607,434	29.28	1,630,348	14.14
1	1,496,668	56.55	5,369,181	60.73
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	58,619	98.55
6	56,610	95.03	29,598	98.8
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	20,283	97.18	11,916	99.33
11	18,535	97.52	10,353	99.42
12	16,356	97.82	8,881	99.5
13	13,858	98.07	7,086	99.56
14	9,098	98.24	5,823	99.61
15	5,132	98.33	3,732	99.64
16	3,961	98.4	2,898	99.67
17	2,847	98.45	2,464	99.69
18	2,366	98.5	2,506	99.71
19	1,690	98.53	2,456	99.73
20	1,468	98.55	2,028	99.75
21	1,479	98.58	1,948	99.77
22	1,883	98.62	1,577	99.78
23	1,655	98.65	1,736	99.8
24	1,463	98.67	1,549	99.81

Notes: This table displays the distribution of the # of months between the time of origination and the time of sale ("Months-to-Sale") 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).

Table 2: Summary Statistics: LPS Sample

	PLS		GSE	
	Mean	SD	Mean	SD
Loan/Borrower Characteristics				
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
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				
Unemployment rate (county-level)	4.8	1.4	4.9	1.5
36 month unemployment growth (				
Price Index (county-level)	188	53	163	46
36 month HPA (%)	43.9	26.5	31.4	23.1
Default Rates				
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 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 9 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 3: PLS Results from Parametric Specification: LPS Sample

Panel A: Full Sample

Default Horizon:	36 Months				60 Months			
Default Definition:	60+ DQ		90+ DQ		60+ DQ		90+ DQ	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Months-to-Sale	-0.0107	-0.0246	-0.0110	-0.0246	-0.0112	-0.0266	-0.0122	-0.0272
	(5.79)	(8.10)	(5.88)	(8.19)	(6.75)	(8.59)	(7.23)	(9.75)
Months-to-Sale^2		0.0027		0.0026		0.0029		0.0029
		(7.37)		(7.13)		(7.74)		(8.61)
# Loans	5,313,951	5,313,951	5,313,951	5,313,951	5,313,951	5,313,951	5,313,951	5,313,951
Adjusted R^2	0.23	0.23	0.22	0.22	0.25	0.25	0.25	0.25
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
Lender FEs?	N	N	N	N	N	N	N	N
Other Controls?	Y	Y	Y	Y	Y	Y	Y	Y

Panel B: Restricted Sample (Only Defaults Occurring After 9 Months)

Default Horizon:	36 Months				60 Months			
Default Definition:	60+ DQ		90+ DQ		60+ DQ		90+ DQ	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Months-to-Sale	-0.0105	-0.0173	-0.0101	-0.0167	-0.0112	-0.0203	-0.0115	-0.0206
	(5.90)	(6.51)	(5.99)	(6.46)	(6.39)	(6.91)	(6.83)	(7.68)
Months-to-Sale^2		0.0013		0.0013		0.0018		0.0018
		(4.57)		(4.27)		(5.39)		(5.98)
# Loans	5,143,409	5,143,409	5,143,409	5,143,409	5,143,409	5,143,409	5,143,409	5,143,409
Adjusted R^2	0.20	0.20	0.19	0.19	0.23	0.23	0.22	0.22
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
Lender FEs?	N	N	N	N	N	N	N	N
Other Controls?	Y	Y	Y	Y	Y	Y	Y	Y

This table displays 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 1-4) and over a 60-month horizon (columns 5-8). Default is defined as a loan that is 60+ days delinquent(columns 1-2 and 5-6) and 90+ days delinquent (columns 3-4 and 7-8). Months-to-Sale is defined as the number of months that elapse between origination and sale to a PLS issuer. All 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, the second row shows t-statistics.Standard errors are heteroskedasticity-robust and are clustered by year-quarter of origination.

Table 4: PLS Results from Non-Parametric Specification: LPS Sample

	Full Sample		Restricted Sample		Full Sample		Restricted Sample	
Default Horizon:	36 Months				60 Months			
Default Definition:	60+ DQ	90+ DQ	60+ DQ	90+ DQ	60+ DQ	90+ DQ	60+ DQ	90+ DQ
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Months-to-Sale = 2	-0.019	-0.019	-0.012	-0.012	-0.020	-0.020	-0.014	-0.014
	(5.12)	(5.01)	(3.82)	(3.82)	(4.60)	(4.88)	(3.07)	(3.34)
Months-to-Sale = 3	-0.038	-0.038	-0.026	-0.026	-0.040	-0.041	-0.030	-0.031
	(6.37)	(6.57)	(5.24)	(5.24)	(6.28)	(6.89)	(5.22)	(5.78)
Months-to-Sale = 4	-0.057	-0.058	-0.046	-0.046	-0.062	-0.066	-0.055	-0.055
	(7.91)	(7.93)	(6.74)	(6.74)	(8.75)	(9.52)	(7.53)	(7.88)
Months-to-Sale = 5	-0.058	-0.059	-0.052	-0.052	-0.062	-0.066	-0.059	-0.061
	(4.71)	(4.81)	(4.86)	(4.86)	(5.98)	(6.22)	(6.07)	(6.14)
Months-to-Sale = 6	-0.054	-0.054	-0.054	-0.054	-0.059	-0.064	-0.063	-0.065
	(3.56)	(3.61)	(4.49)	(4.49)	(4.32)	(4.54)	(5.31)	(5.32)
Months-to-Sale = 7	-0.044	-0.046	-0.049	-0.049	-0.047	-0.053	-0.054	-0.056
	(3.51)	(3.76)	(5.28)	(5.28)	(4.33)	(4.94)	(5.39)	(5.77)
Months-to-Sale = 8	-0.031	-0.034	-0.045	-0.045	-0.028	-0.036	-0.044	-0.047
	(2.03)	(2.38)	(3.48)	(3.48)	(2.04)	(2.49)	(3.09)	(3.29)
Months-to-Sale = 9	-0.036	-0.040	-0.049	-0.049	-0.031	-0.037	-0.045	-0.046
	(2.11)	(2.39)	(3.42)	(3.42)	(1.81)	(2.29)	(2.91)	(3.12)
# Loans	5,313,951	5,313,951	5,143,409	5,143,409	5,313,951	5,313,951	5,143,409	5,143,409
Adjusted R^2	0.23	0.22	0.19	0.19	0.25	0.25	0.23	0.23
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
Lender FEs?	N	N	N	N	N	N	N	N
Other Controls?	Y	Y	Y	Y	Y	Y	Y	Y

This table displays 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 1-4) and over a 60-month horizon (columns 5-8). Default is defined as a loan that is 60+ days delinquent(columns 1-2 and 5-6) and 90+ days delinquent (columns 3-4 and 7-8). Months-to-Sale is defined as the number of months that elapse between origination and sale to a PLS issuer. All 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, the second row shows t-statistics.Standard errors are heteroskedasticity-robust and are clustered by year-quarter of origination.

Table 5: GSE Results from Parametric Specification: LPS Sample

Panel A: Full Sample

Default Horizon:	36 Months				60 Months			
Default Definition:	60+ DQ		90+ DQ		60+ DQ		90+ DQ	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Months to Sale	-0.0001	-0.0005	-0.0010	-0.0014	-0.0002	-0.0020	-0.0012	-0.0030
	(0.11)	(0.35)	(1.60)	(1.07)	(0.19)	(1.14)	(1.72)	(1.94)
Months to Sale^2		0.0001		0.0001		0.0005		0.0005
		(0.41)		(0.39)		(1.58)		(1.65)
# Loans	11,437,522	11,437,522	11,437,522	11,437,522	11,437,522	11,437,522	11,437,522	11,437,522
Adjusted R^2	0.14	0.14	0.14	0.14	0.16	0.16	0.16	0.16
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
Lender FEs?	N	N	N	N	N	N	N	N
Other Controls?	Y	Y	Y	Y	Y	Y	Y	Y

Panel B: Restricted Sample (Only Defaults Occurring After 9 Months)

Default Horizon:	36 Months				60 Months			
Default Definition:	60+ DQ		90+ DQ		60+ DQ		90+ DQ	
Months to Sale	-0.0014	-0.0014	-0.0015	-0.0014	-0.0015	-0.0029	-0.0017	-0.0031
	(2.42)	(1.20)	(2.97)	(1.35)	(2.15)	(1.95)	(2.89)	(2.31)
Months to Sale^2		0.0000		0.0000		0.0004		0.0004
		(0.00)		(0.07)		(1.42)		(1.47)
# Loans	11,267,367	11,267,367	11,267,367	11,267,367	11,267,367	11,267,367	11,267,367	11,267,367
Adjusted R^2	0.13		0.12	0.12	0.15	0.15	0.15	0.15
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
Lender FEs?	N	N	N	N	N	N	N	N
Other Controls?	Y	Y	Y	Y	Y	Y	Y	Y

This table displays 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 1-4) and over a 60-month horizon (columns 5-8). Default is defined as a loan that is 60+ days delinquent(columns 1-2 and 5-6) and 90+ days delinquent (columns 3-4 and 7-8). Months-to-sale is defined as the number of months that elapse between origination and sale to a GSE. All 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, the second row shows t-statistics.Standard errors are heteroskedasticity-robust and are clustered by year-quarter of origination.

Table 6: Ex-Ante Default Risk Results: LPS Sample

Panel A: PLS Loans

Default Horizon:	36 Months				60 Months			
Default Definition:	60+ DQ		90+ DQ		60+ DQ		90+ DQ	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Months-to-Sale	0.0058	0.0197	0.0045	0.0150	0.0057	0.0186	0.0040	0.0112
	(5.20)	(7.24)	(5.20)	(7.01)	(4.40)	(8.65)	(3.45)	(6.61)
Months-to-Sale^2		-0.0028		-0.0021		-0.0026		-0.0015
		(8.51)		(8.15)		(10.03)		(9.58)
# Loans	3,672,426	3,672,426	3,672,426	3,672,426	3,672,426	3,672,426	3,672,426	3,672,426
Adjusted R^2	0.26	0.27	0.24	0.25	0.30	0.31	0.36	0.37
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

Default Horizon:	36 Months				60 Months			
Default Definition:	60+ DQ		90+ DQ		60+ DQ		90+ DQ	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Months-to-Sale	0.0004	0.0014	0.0002	0.0004	0.0021	0.0011	0.0013	0.0008
	(1.17)	(1.78)	(0.93)	(0.96)	(3.20)	(0.68)	(3.50)	(0.78)
Months-to-Sale^2		-0.0002		-0.0001		0.0002		0.0001
		(1.57)		(0.71)		(0.64)		(0.58)
# Loans	7,378,891	7,378,891	7,378,891	7,378,891	7,378,891	7,378,891	7,378,891	7,378,891
Adjusted R^2	0.29	0.29	0.30	0.30	0.52	0.52	0.56	0.56
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 loan-level, OLS regressions where the dependent variables are the 36-month, and 60-month ex-ante default rates at the time the loan is 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 regressions include origination year-quarter fixed effects, and year-quarter of sale fixed effects. Standard errors are heteroskedasticity-robust and clustered at the quarter of issuance level. The first row for each variable shows the regression coefficient, the second row shows t-statistics.

Table 7: Distribution of Months-to-Sale in CoreLogic Sample

	All PLS		Subprime PLS		Alt-A PLS	
Months-to-Sale	# Loans	Cumulative % of Sample	# Loans	Cumulative % of Sample	# Loans	Cumulative % 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	38,190	98.6	20,454	98.6	17,736	98.6
11	30,233	98.8	16,464	98.8	13,769	98.9
12	24,564	99.0	14,094	98.9	10,470	99.1
13	19,247	99.2	11,051	99.1	8,196	99.3
14	15,630	99.3	9,301	99.2	6,329	99.4
15	14,481	99.4	9,445	99.3	5,036	99.5
16	11,835	99.5	7,744	99.4	4,091	99.6
17	13,645	99.6	9,997	99.5	3,648	99.7
18	11,432	99.6	8,364	99.6	3,068	99.7
19	10,814	99.7	7,889	99.7	2,925	99.8
20	8,602	99.8	6,245	99.7	2,357	99.9
21	7,910	99.8	6,062	99.8	1,848	99.9
22	7,574	99.9	5,790	99.9	1,784	99.9
23	7,511	100.0	5,702	100.0	1,809	100.0
24	764	100.0	3,710	100.0	1,576	100.0

Notes: This table displays the distribution of the # 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 8: Summary Statistics: CoreLogic Sample

	All PLS		Subprime PLS		Alt-A PLS	
	Mean	SD	Mean	SD	Mean	SD
Loan/Borrower Characteristics						
Term	356	37	355	34	357	42
Original Rate	7.28	1.62	7.86	1.32	6.27	1.60
Original Amount	222,701	156,989	187,331	122,637	285,146	188,403
LTV Ratio	82.8	14.7	83.8	14.0	80.9	15.7
FICO	650	72	617	60	709	50
Purchase (d)	0.416	.	0.367701	.	0.501	.
Cash Out Refinance (d)	0.472	.	0.549	.	0.335	.
Arm (d)	0.684	.	0.748031	.	0.571	.
Balloon (d)	0.055	.	0.081966	.	0.009	.
Interest Only (d)	0.213	.	0.120	.	0.376	.
Jumbo (d)	0.142	.	0.083789	.	0.246	.
Low Doc (d)	0.475	.	0.345177	.	0.704	.
Prepay Penalty (d)	0.621	.	0.740	.	0.400	.
Primary Residence (d)	0.855	.	0.917567	.	0.744	.
Single Family (d)	0.727	.	0.782	.	0.630	.
Geographic Characteristics						
Unemployment rate (county-level)	5.18	1.57	5.32	1.59	4.93	1.50
36 month unemployment growth (%)	4.7%	39.6%	9.0%	40.6%	-2.9%	36.6%
Price Index (county-level)	177	52	170	50	189	53
36 month HPA (%)	42.5%	26.5%	40.2%	26.3%	46.3%	26.4%
Default Rates						
60+ DQ, 36-month horizon	0.215	.	0.245	.	0.154	.
60+ DQ, 60-month horizon	0.269	.	0.294	.	0.227	.
90+ DQ, 36-month horizon	0.178	.	0.204	.	0.131	.
90+ DQ, 60-month horizon	0.238	.	0.255	.	0.207	.
# Loans	13,430,586		8,574,041		4,856,545	

Notes: This table displays 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 9 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 9: Baseline Parametric Results for Sample of CoreLogic PLS Loans

Panel A: No Lender Fixed Effects

Default Definition:	60+ DQ over 36 Months					
	All PLS		Alt-A		Subprime	
	(1)	(2)	(3)	(4)	(5)	(6)
Months-to-Sale	-0.0036	-0.0046	-0.0072	-0.0100	-0.0020	-0.0019
	(4.28)	(3.38)	(6.87)	(5.78)	(2.46)	(1.19)
Months-to-Sale^2		0.0002		0.0004		0.0000
		(0.93)		(1.85)		(0.05)
# Loans	7,860,499	7,860,499	1,895,618	1,895,618	5,964,881	5,964,881
Adjusted R^2	0.21	0.21	0.25	0.25	0.19	0.19
Orig Qtr FEs?	Y	Y	Y	Y	Y	Y
State FEs?	Y	Y	Y	Y	Y	Y
Sale Qtr FEs?	Y	Y	Y	Y	Y	Y
Lender FEs?	N	N	N	N	N	N
Other Controls?	Y	Y	Y	Y	Y	Y

Panel B: Lender Fixed Effects

Default Definition:	60+ DQ over 36 Months					
	All PLS		Alt-A		Subprime	
	(1)	(2)	(3)	(4)	(5)	(6)
Months-to-Sale	-0.0028	-0.0043	-0.0063	-0.01	-0.0015	-0.0005
	(3.80)	(3.96)	(5.61)	(6.73)	(2.08)	(0.48)
Months-to-Sale^2		0.0002		0.0006		-0.0002
		(1.93)		(4.09)		(1.10)
# Loans	7,860,499	7,860,499	1,895,618	1,895,618	5,964,881	5,964,881
Adjusted R^2	0.21	0.21	0.26	0.26	0.19	0.19
Orig Qtr FEs?	Y	Y	Y	Y	Y	Y
State FEs?	Y	Y	Y	Y	Y	Y
Sale Qtr FEs?	Y	Y	Y	Y	Y	Y
Lender FEs?	Y	Y	Y	Y	Y	Y
Other Controls?	Y	Y	Y	Y	Y	Y

Notes: This table displays 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 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. Specifications in Panel B include a full set of originator fixed effects. The first row for each variable shows the regression coefficient, the second row shows t-statistics. Standard errors are heteroskedasticity-robust and are clustered by year-quarter of origination.

Table 10: Documentation Results for Sample of CoreLogic PLS Loans

Panel A: 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
	(1)	(2)	(3)	(4)	(5)	(6)
Months-to-Sale	-0.0027	-0.0063	-0.0103	-0.0091	-0.0011	0.0006
	(2.43)	(4.03)	(5.08)	(5.76)	(0.99)	(0.37)
Months-to-Sale^2	0.0001	0.0003	0.0006	0.0006	0.0000	-0.0004
	(0.94)	(2.36)	(2.49)	(3.60)	(0.29)	(2.37)
# Loans	3,261,827	2,605,838	378,607	1,035,183	3,842,498	2,092,603
Adjusted R^2	0.18	0.25	0.16	0.27	0.17	0.24
Orig Qtr FEs?	Y	Y	Y	Y	Y	Y
State FEs?	Y	Y	Y	Y	Y	Y
Sale Qtr FEs?	Y	Y	Y	Y	Y	Y
Lender 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
	(1)	(2)	(3)	(4)	(5)	(6)
Months to Sale = 1	-0.0023	-0.0103	-0.0113	-0.0166	0.0001	0.0028
	(0.74)	(2.66)	(3.80)	(5.34)	(0.04)	(0.79)
Months to Sale = 2	-0.0015	-0.0138	-0.0177	-0.0219	0.0025	0.0049
	(0.54)	(4.05)	(4.39)	(7.97)	(1.12)	(1.78)
Months to Sale = 3	-0.0057	-0.0175	-0.0285	-0.0274	-0.0014	0.0005
	(1.91)	(4.05)	(5.31)	(7.49)	(0.51)	(0.13)
Months to Sale = 4	-0.0095	-0.0203	-0.0348	-0.0308	-0.005	-0.0016
	(2.75)	(3.80)	(6.88)	(6.17)	(1.46)	(0.39)
Months to Sale = 5	-0.0111	-0.0272	-0.0362	-0.0293	-0.0079	-0.0118
	(2.16)	(4.61)	(5.12)	(5.22)	(1.68)	(1.99)
Months to Sale = 6	-0.0109	-0.0292	-0.0364	-0.0347	-0.0076	-0.0116
	(1.84)	(3.90)	(5.29)	(5.54)	(1.35)	(1.28)
Months to Sale = 7	-0.0134	-0.0326	-0.048	-0.0388	-0.0087	-0.0147
	(1.81)	(3.69)	(4.41)	(4.86)	(1.27)	(1.67)
Months to Sale = 8	-0.0078	-0.0318	-0.0489	-0.0518	-0.0007	-0.0043
	(1.02)	(3.18)	(4.80)	(5.20)	(0.09)	(0.48)
Months to Sale = 9	-0.0004	-0.0339	-0.0528	-0.0583	0.0082	-0.0057
	(0.03)	(3.61)	(4.35)	(7.91)	(0.61)	(0.47)
# Loans	3,261,827	2,605,838	378,607	1,035,183	3,842,498	2,092,603
Adjusted R^2	0.18	0.25	0.16	0.27	0.17	0.24
Orig Qtr FEs?	Y	Y	Y	Y	Y	Y
State FEs?	Y	Y	Y	Y	Y	Y
Sale Qtr FEs?	Y	Y	Y	Y	Y	Y
Lender FEs?	Y	Y	Y	Y	Y	Y
Other Controls?	Y	Y	Y	Y	Y	Y

Notes: This table displays 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 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 correspond to 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, the second row shows t-statistics. Standard errors are heteroskedasticity-robust and are clustered by year-quarter of origination.

Table 11: Affiliation Results for 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
	(1)	(2)	(3)	(4)	(5)	(6)
Months-to-Sale	-0.0049	-0.0121	-0.0080	-0.0194	-0.0028	-0.0073
	(3.42)	(5.57)	(3.92)	(8.36)	(2.27)	(2.97)
Months-to-Sale^2	0.0001	0.0010	0.0004	0.0011	0.0000	0.0006
	(0.61)	(4.49)	(1.83)	(6.08)	(0.12)	(2.28)
# Loans	2,384,156	2,606,571	453,075	551,994	1,931,081	2,054,577
Adjusted R^2	0.20	0.21	0.24	0.26	0.19	0.20
Orig Qtr FEs?	Y	Y	Y	Y	Y	Y
State FEs?	Y	Y	Y	Y	Y	Y
Sale Qtr FEs?	Y	Y	Y	Y	Y	Y
Lender 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
	(1)	(2)	(3)	(4)	(5)	(6)
Months-to-Sale = 1	-0.0042	-0.0310	-0.0073	-0.0443	-0.0029	0.0039
	(1.32)	(3.80)	(3.43)	(7.54)	(1.07)	(0.82)
Months-to-Sale = 2	-0.0068	-0.0377	-0.0138	-0.0518	-0.0038	-0.0019
	(2.45)	(5.19)	(3.29)	(10.29)	(1.52)	(0.37)
Months-to-Sale = 3	-0.0140	-0.0422	-0.0253	-0.0645	-0.0086	-0.0064
	(4.40)	(5.50)	(3.29)	(10.75)	(2.92)	(1.05)
Months-to-Sale = 4	-0.0183	-0.0456	-0.0279	-0.0745	-0.0120	-0.0095
	(4.66)	(4.80)	(4.64)	(9.12)	(3.21)	(1.28)
Months-to-Sale = 5	-0.0250	-0.0490	-0.0250	-0.0770	-0.0187	-0.0132
	(4.55)	(4.76)	(4.16)	(9.70)	(3.37)	(1.68)
Months-to-Sale = 6	-0.0204	-0.0513	-0.0271	-0.0870	-0.0150	-0.0122
	(3.07)	(4.54)	(3.13)	(8.04)	(2.00)	(1.36)
Months-to-Sale = 7	-0.0297	-0.0535	-0.0308	-0.0957	-0.0294	-0.0105
	(3.66)	(4.71)	(3.46)	(7.56)	(3.00)	(1.08)
Months-to-Sale = 8	-0.0267	-0.0537	-0.0552	-0.1089	-0.0167	-0.0059
	(2.20)	(4.80)	(5.32)	(8.40)	(1.27)	(0.50)
Months-to-Sale = 9	-0.0157	-0.0482	-0.0527	-0.1185	-0.0020	0.0083
	(1.70)	(3.44)	(3.83)	(9.36)	(0.16)	(0.63)
# Loans	2,384,156	2,606,571	453,075	551,994	1,931,081	2,054,577
Adjusted R^2	0.2	0.21	0.24	0.26	0.19	0.2
Orig Qtr FEs?	Y	Y	Y	Y	Y	Y
State FEs?	Y	Y	Y	Y	Y	Y
Sale Qtr FEs?	Y	Y	Y	Y	Y	Y
Lender FEs?	Y	Y	Y	Y	Y	Y
Other Controls?	Y	Y	Y	Y	Y	Y

Notes: This table displays 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 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 correspond to those in which the originator of all 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, the second row shows t-statistics. Standard errors are heteroskedasticity-robust and are clustered by year-quarter of origination.

Table 12: Early Prepayment Results

Panel A: Parametric Specification

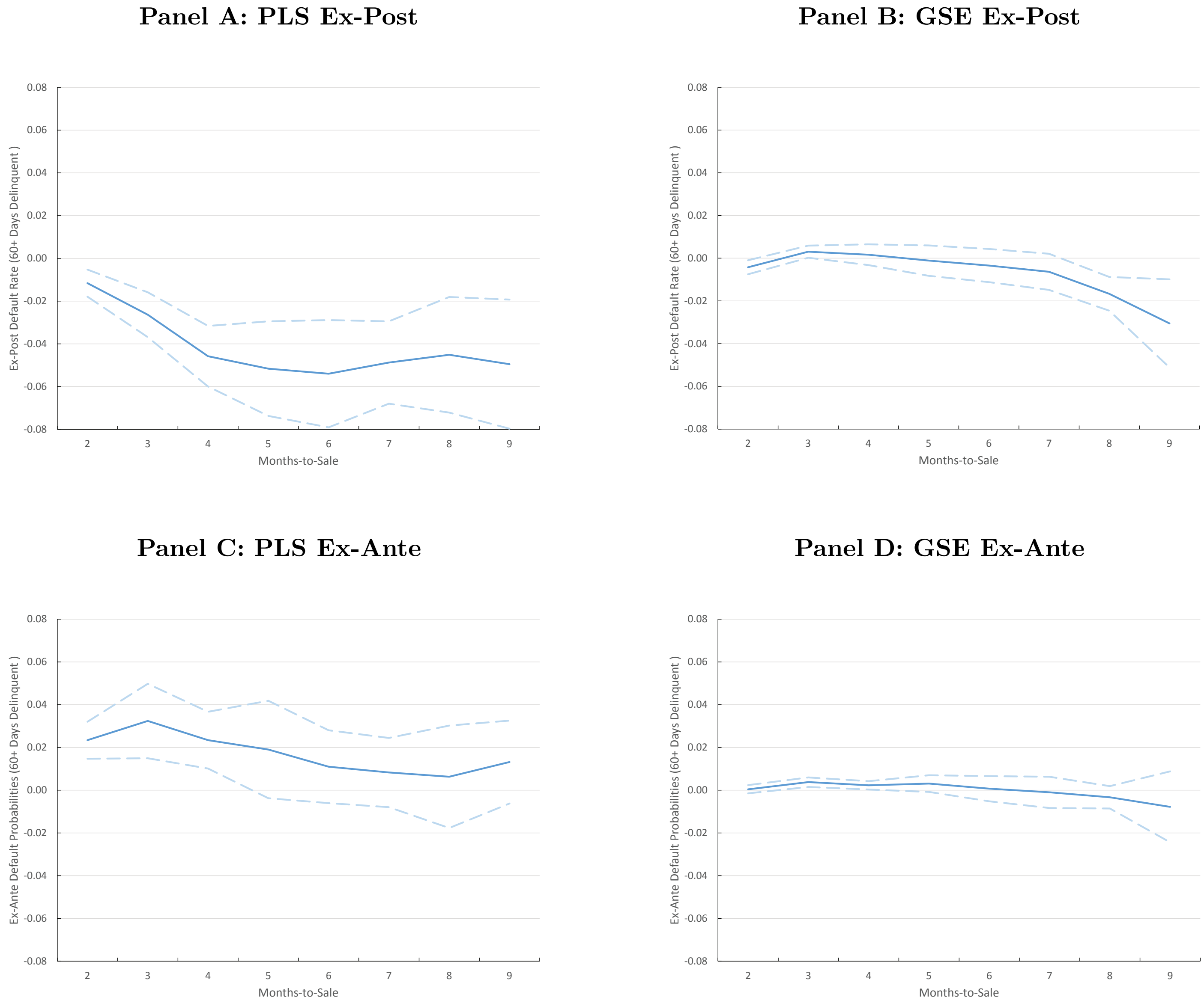
Correction:	None		$\leq$ 3 months		$\leq$ 6 months		$\leq$ 9 months	
Reset Month - Prepay Month	$\geq$ 6 Months	$\geq$ 9 Months	$\geq$ 6 Months	$\geq$ 9 Months	$\geq$ 6 Months	$\geq$ 9 Months	$\geq$ 6 Months	$\geq$ 9 Months
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Months-to-Sale	-0.0129	-0.0152	-0.0089	-0.0105	-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
Lender 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
Lender FEs?	Y	Y	Y	Y	Y	Y	Y	Y
Other Controls?	Y	Y	Y	Y	Y	Y	Y	Y

Notes: This table displays 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 3 months or 6 months before the month in which the interest rate resets from a fixed rate to an adjustable rate. All loans that prepaid within 3 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 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, the second row shows t-statistics. Standard errors are heteroskedasticity-robust and are clustered by year-quarter of origination.

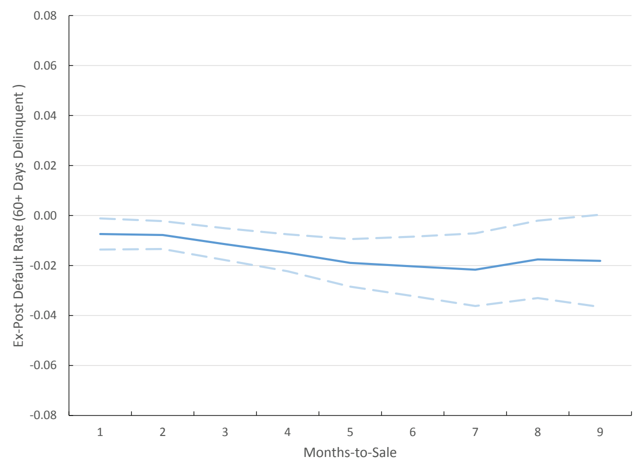
Figure 1: Ex-Ante vs. Ex-Post LPS Results



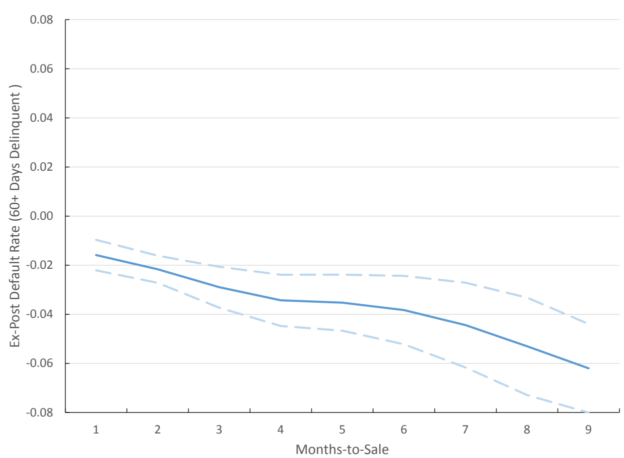
Notes: This figure displays results from the estimation of the non-parametric version of equation 5 for both PLS and GSE loans in the LPS dataset originated in the 2002 - 2007 period. Panels A and B correspond to ex-post default rates, while panels C and D correspond to ex-ante predicted default rates. 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 PLS issuer. Dotted lines correspond to 90 percent confidence intervals.

Figure 2: CoreLogic PLS Results

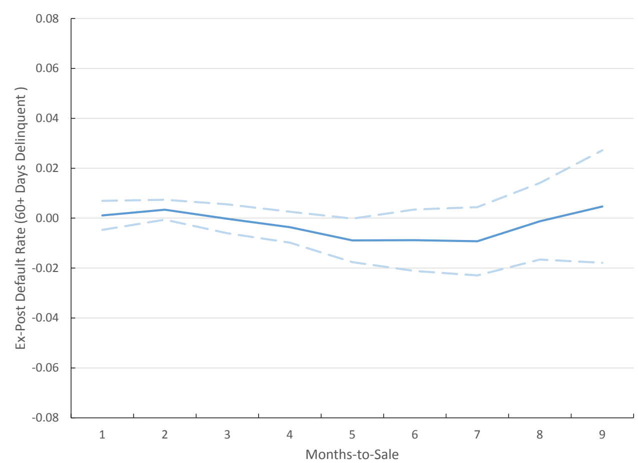
Panel A: All PLS



Panel B: Alt-A PLS



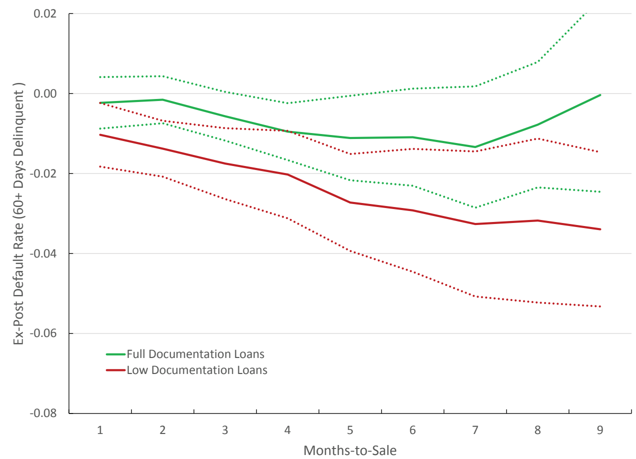
Panel C: Subprime PLS



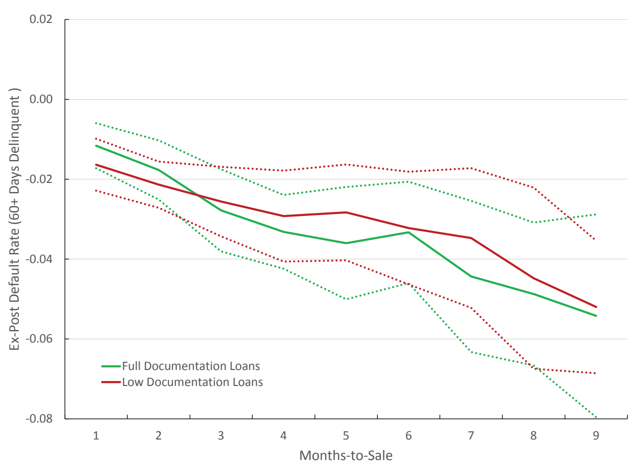
Notes: This figure displays results from the estimation of the non-parametric version of equation 5 for PLS loans in the CoreLogic dataset originated in the 2002 - 2007 period. Panel A corresponds to all PLS loans, while panels B and C correspond to Alt-A and Subprime loans, respectively. Default is 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. Dotted lines correspond to 90 percent confidence intervals.

Figure 3: CoreLogic PLS Documentation Results

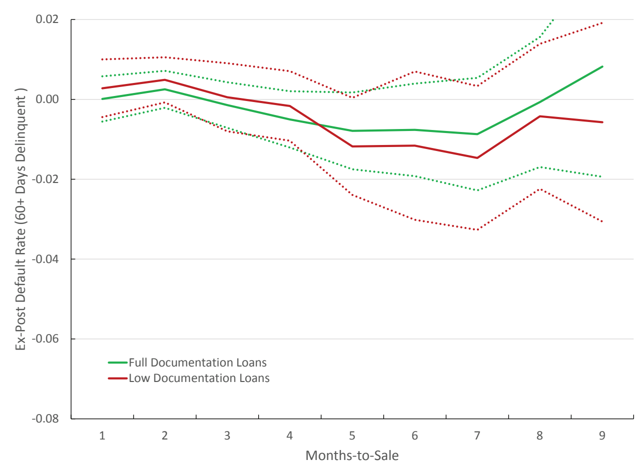
Panel A: **All PLS**



Panel B: Alt-A PLS



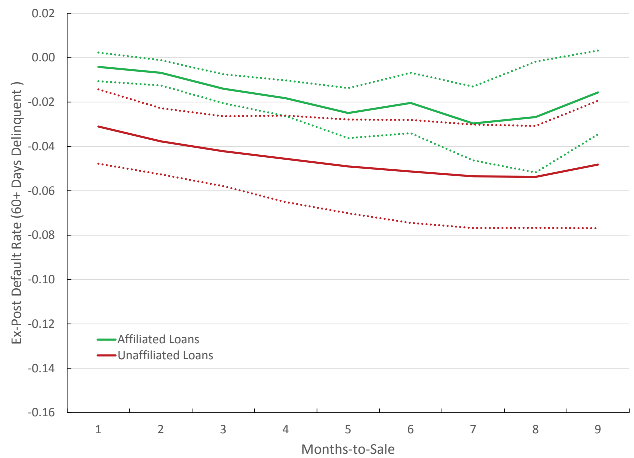
Panel C: Subprime PLS



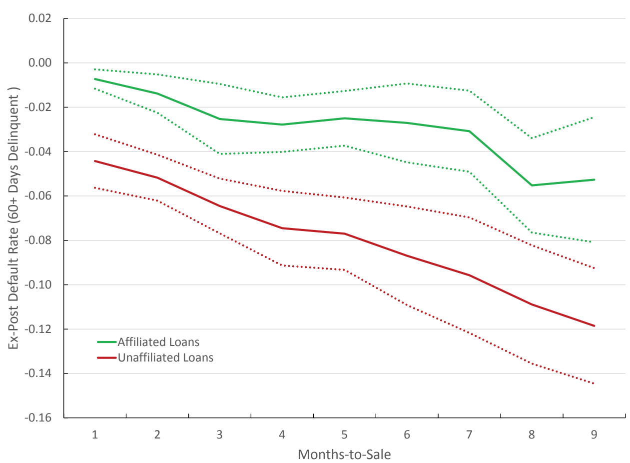
Notes: This figure displays results from the estimation of the non-parametric version of equation 5 for PLS loans in the CoreLogic dataset originated in the 2002 - 2007 period. Panel A corresponds to all PLS loans, while panels B and C correspond to Alt-A and Subprime loans, respectively. Default is 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. Dotted lines correspond to 90 percent confidence intervals.

Figure 4: CoreLogic PLS Affiliation Results

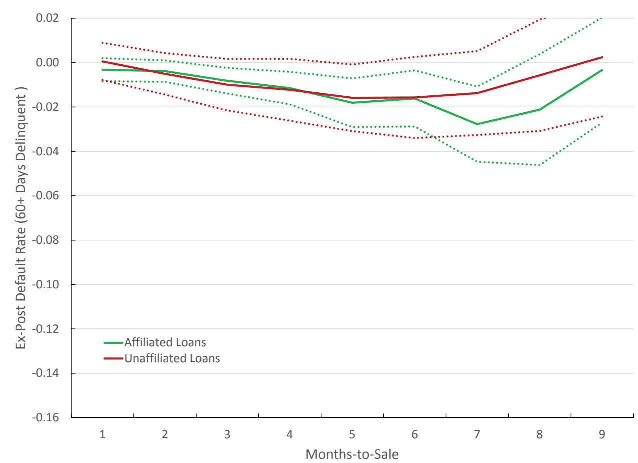
Panel A: All PLS



Panel B: Alt-A PLS



Panel C: Subprime PLS



Notes: This figure displays 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 all PLS loans, while panels B and C correspond to Alt-A and Subprime loans, respectively. 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. Dotted lines correspond to 90 percent confidence intervals.

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 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 models include both the continuous FICO variable as well as a set of indicator variables corresponding to 5 FICO intervals: FICO < 580, $580 \leq$FICO<620, $620 \leq$FICO<660, $660 \leq$FICO<700, FICO≥ 700.

House Prices: County-level house price indices from CoreLogic. We include both the level of prices in the county in the month of origination as well as 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 lien and the second lien. All models include both the continuous LTV variable as well as a set of indicator variables corresponding to 5 LTV intervals: LTV < 70, $70 \leq$LTV<80, $80<$LTV<90, $90 \leq$LTV<100, LTV≥ 100. An indicator variable for 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 the 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 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 level of rates in the county in the month of origination as well as the cumulative growth in the unemployment rate from the month of mortgage origination, calculated over the default horizon.

Table 13: 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 $\geq$ 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
Lender 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 regressions include origination year-quarter fixed effects, state fixed effects, year-quarter of sale fixed effects. The first row for each variable shows the regression coefficient, the second row shows t-statistics. Standard errors are heteroskedasticity-robust and are clustered by year-quarter of origination.