Screen More, Sell Later: Screening and Dynamic Signaling in the Mortgage Market

Manuel Adelino* Bin Wei† Feng Zhao‡

March 13, 2024

Abstract

We build on previous work and provide a dynamic model of asset markets with asymmetric information where higher originator screening effort leads to more signaling through delay of sale. We test this theoretical prediction using the mortgage market as a laboratory and processing time as a measure of screening. Our findings are threefold: First, and in line with the theory, mortgage processing time and the delay of sale after origination are strongly positively related in the data. Second, processing time is longer for mortgages with higher *ex ante* credit risk, i.e., observably riskier loans are processed slower. Finally, both processing time and delay of sale are negatively related to conditional mortgage default, indicating that more screening effort leads to unobservably higher quality loans that are also sold with a longer delay.

JEL Classification: G01, G21, G23, G32, R30

Keywords: Processing time, screening, signaling, time to sale, securitization, mortgage loans, lending standards

We thank Nuno Clara, Simon Gervais, and João Thereze, as well as seminar participants at Duke for helpful comments. The views expressed in this paper are those of the authors and do not necessarily represent those of the Federal Reserve System.

*Fuqua School of Business, Duke University; NBER; CEPR. Email: manuel.adelino@duke.edu

†Research Department, Federal Reserve Bank of Atlanta. Email: bin.wei@atl.frb.org

‡Naveen Jindal School of Management, University of Texas at Dallas. Email: feng.zhao@utdallas.edu

1 Introduction

In the canonical market setting with asymmetric information of Akerlof (1970), the seller of a high-quality asset has an incentive to take costly actions to distinguish herself from those of low-quality assets (Spence, 1973). These costly actions may involve partial retention (Leland and Pyle, 1977; DeMarzo, 2005), and, in dynamic settings, the delay of sale (Janssen and Roy, 2002; Daley and Green, 2012). Partial retention or delayed sales are costly because they do not allow for the full realization of gains from trade.

This paper develops a model of origination and securitization to analyze the trade-off between the incentives to originate good assets and secondary market liquidity when securitizing and selling those assets. Our model builds on Vanasco (2017), where an informed seller can choose to screen assets at origination to improve their average quality, enhancing "productive efficiency." At the same time, asset screening worsens adverse selection and lowers market liquidity. Therefore, mortgage originators' *ex ante* incentives to screen are intricately linked to their *ex post* motivation to signal.

A robust property emerges from the combination of screening at origination and signaling before trading: The effort spent on screening for unobserved quality is positively related to the degree of signaling at the later stage. The intuition is that signaling is costly — in a dynamic setting, originators are forced to hold on to originated loans for too long, thus reducing "allocative efficiency". The anticipation of these outcomes incentivizes originators to exert more effort *ex ante*, giving rise to the positive relation between screening effort and signaling, which is the central prediction in Vanasco (2017) and this paper. The model also provides sharply distinct predictions for how private and public signals available to originators affect screening effort. While more favorable private information leads to more screening, observably better loans are associated with lower effort despite a higher likelihood of securitization.

Despite the theoretical appeal of these predictions, empirical tests that combine screening and signaling in asset markets have proven elusive. First, screening effort is rarely observable to other market participants and, by extension, to the econometrician. In addition, any test of signaling behavior requires the data available to outside observers to include asset characteristics that are correlated with seller information about quality, but are unavailable to buyers.

This paper uses the origination of private-label securitized mortgages in the U.S. between 2002 and 2006 as a unique laboratory to relate screening effort, signaling, and unobserved asset quality. We measure screening effort as the time between mortgage application and mortgage closing (which we refer to as "processing time"). Processing time is typically used by lenders to perform appraisals, obtain borrower documents, and conduct additional due diligence. This time may also be used by borrowers for purposes unrelated to lender effort, such as organizing a move, selling another home, performing home inspections, and may be related to preferences for end-of-month closing dates (Bhutta and Ringo, 2021). It is also known that processing time is associated with lender technology (Foote, Loewenstein, and Willen, 2019; Fuster, Plosser, Schnabl, and Vickery, 2019). The key assumption needed for our tests is that processing time is positively related to lender screening effort and thus leads to the origination of (unobservably) higher quality loans. We validate our measure below by showing that longer processing time is associated with better loans on unobserved dimensions in a variety of tests. Previous work by Choi and Kim (2021) comparing conforming and non-conforming loans after the collapse of the private-label securitization market in 2007 provides further evidence that processing time is a good measure of screening effort.[1]

Our model provides the foundation for an empirical test of processing time as a measure of screening that is closely related to Keys, Mukherjee, Seru, and Vig (2010). Specifically, we examine potential discontinuities in screening effort for mortgages originated around the threshold of 620 Fair Isaac Corporation (FICO) scores, a public signal that is related to the ease of securitization and observed by both originators and buyers. The key assumption in this analysis is that demand for loans is likely to be smooth around the 620 threshold, but it generates strong discontinuities in the probability of loan origination (Bubb and Kaufman, 2014) and, for a subset of those loans (low documentation loans), in the probability of loan securitization (Keys, Seru, and Vig, 2012). Keys et al. (2010) show that there is also a large jump in the probability of default, with loans below the threshold experiencing significantly lower defaults relative to those just above. We show that there is also a discontinuity in processing time in the predicted direction at exactly 620: loans with a FICO score just below have discontinuously higher processing times relative to those just above. The discontinuity in processing time is present for low documentation loans, where we also see a jump in default rates, but it is much weaker in the "full doc" subsample. Ours is the first direct evidence that the previous results in the literature on default around this threshold can be attributed to differences in lender screening effort rather than to any other unobserved source of heterogeneity.

The presence of both public and private signals of loan quality in the model at the time of loan application generates a second sharp prediction for processing time. Specifically, the model predicts that processing time should be increasing in *ex-ante* observable credit risk, but unobserved loan quality should be negatively related to processing time. We construct a measure of credit risk using backward-looking information and all observable mortgage characteristics, and show that originator processing time is strongly positively related to this measure of observable credit risk. Processing time is also higher for lower credit score borrowers, a more succinct measure of risk. However, we show that processing time is *negatively* related to conditional *ex-post* default, which is, by construction, unobservable from the perspective of market participants at the time of origination. This measure of unobservable loan quality is the same one used in Adelino, Gerardi, and Hartman-Glaser (2019) (hereafter AGH). This is an appealing measure because conditional default is related to any unobserved lender information about the mortgages, but the default itself happens well after the transactions take place, and so cannot be known to buyers or sellers at that time. By controlling for a wide array of observable characteristics available to the buyers of mortgages during this period, we can also provide a clean distinction between observable and unobservable credit risk.

[1] Bedayo, Jiménez, Peydró, and Vegas Sánchez (2020) adopt a similar measure for corporate loans and also show that it is related to ex post performance. In a similar spirit, Ben-Rephael, Carlin, Da, and Israelsen (2023) use office workday length as a measure of hard information gathering by equity analysts.

Finally, we show that the central prediction of Vanasco (2017) and our model (also in Daley, Green, and Vanasco (2020)) for the positive relation between screening effort and signaling is borne out by the data. Here, we again follow AGH and measure signaling as the time from mortgage origination (the date of closing) to the issuance of the securitization trust in which the specific mortgage is included. This measure is a good empirical analog to the notion of delay of sale used in papers like Daley and Green (2012). We show that delay of sale and processing time are positively related, i.e., loans that take longer to process also experience longer delays of sale. Both measures are also associated with lower conditional default rates when we include them together in the regressions. The "skimming property" that relates delay of sale with default and is the focus of AGH is not entirely absorbed by adding processing time to the regressions, and both variables matter for explaining *ex-post* defaults (in other words, neither variable is a sufficient statistic for the other).

Related Literature. This paper relates to the extant literature on adverse selection and signaling in the context of asset sales, started by Leland and Pyle (1977), Myers and Majluf (1984), and Gorton and Pennacchi (1995). Buyers are concerned about the presence of low quality assets ("lemons") that they cannot identify, forcing the informed seller to signal high quality by retention (Leland and Pyle, 1977; DeMarzo and Duffie, 1999; DeMarzo, 2005) or the delay of sale (Janssen and Roy, 2002; Daley and Green, 2012; Adelino et al., 2019) to obtain higher prices. However, signaling is costly in that asset cash flows are not fully allocated to the highest value party (the buyer in this case). Begley and Purnanandam (2017) study retention of equity tranches in the context of residential mortgage-backed securities and show that higher tranches are related to lower delinquency. Kremer and Skrzypacz (2007) and Daley et al. (2020) consider the introduction of ratings as a public signal that conveys information about the underlying asset and may alleviate some of these frictions, but potentially at the cost of lower underlying asset quality.

The timing of actions reveals private information in a variety of models with adverse selection. For example, Noldeke and Van Damme (1990) and Swinkels (1999) consider models of labor markets and education choice, Grenadier and Wang (2005) and Grenadier and Malenko (2011) analyze firm investment timing, Chang (2017) and Williams (2016) investigate multidimensional private information about asset quality and seller distress (or impatience), Fuchs and Skrzypacz (2019) study how the frequency of trade affects market efficiency, and Fuchs, Green, and Papanikolaou (2016) and Daley and Green (2016) study the role of adverse selection and delay of trade in generating fluctuations in liquidity in good and bad times. Fuchs and Skrzypacz (2015) consider optimal government intervention through trading restrictions in a market of distressed sellers that also has the "skimming property." Hartman-Glaser, Piskorski, and Tchistyi (2012) study a moral hazard setup where the timing of payments to an agent (in their case, the mortgage underwriter) can serve as an incentive mechanism to exert effort. Though the mechanism is fundamentally different from the one we study in this paper, this paper also generates a positive relation between the timing of sale and asset quality.

Our paper is closely related to the strand of the literature on the trade-off between incentives to originate good assets ("productive efficiency") and secondary market liquidity ("allocative efficiency"). For instance, Parlour and Plantin (2008) study the effect of loan sales on banks' origination decisions. Chemla and Hennessy (2014) analyze the effect of speculative information production or optimal regulation on the trade-off between productive and allocative efficiency (Dell'Ariccia and Marquez, 2006; Malherbe, 2012). Vanasco (2017) shows that costly retention of cash flows is essential to implement ex-ante asset screening. Daley et al. (2020) show that when informative ratings are available, there is some degree of pooling at a lower retention level—the economy endogenously shifts from a signaling equilibrium to an originate-to-distribute equilibrium. He, Jiang, and Xu (2023) develop a general equilibrium model to examine the role of information technology.

Several recent papers examine mortgage origination timeline. Foote et al. (2019) show that technology development in mortgage underwriting induces a dramatic decline in the average processing time between 1995 and 1998. Fuster et al. (2019) find that FinTech lenders shorten processing time through enhanced efficiency. The dramatic decline in processing time in our paper is more related to the rise of non-agency securitization and the associated lax lending standards.

The remainder of the paper is organized as follows. Section 2 introduces the model setup, describes the equilibrium allocations, and generates empirical predictions. In Section 4, we take the model to the data and empirically test those predictions. Section 5 concludes. All proofs are relegated to the Appendix.

2 A model of screening and signaling

In this section, we present a model of screening and signaling following Vanasco (2017) and AGH. In our model, a delayed sale is not only a signaling device for loan quality, but also allows for the implementation of screening effort as studied in Vanasco (2017). We further extend those models to include ease of securitization that is related to public information (e.g., credit scores).

The model includes a mortgage originator and a competitive market of outside investors. Both the originator and investors are risk neutral, but have different discount rates. Specifically, the originator discounts cash flows at rate γ, while the investors discount cash flows at rate r. We assume

$$\gamma > r. \tag{1}$$

The difference in discount rates generates gains from trade. That the originator has a higher discount rate may, for example, result from credit constraints that she faces, or from her preference to raise capital to originate new loans.

There are two periods. In the first period, there are two stages: the *Origination* stage and the *Securitization* stage. The originator first screens borrowers by exerting unobservable effort in the origination stage, and then sells the loan to outside investors in the securitization stage. In the second period, the state of the economy and the cash flows from the originated loan are realized.

The Origination Stage. The originator receives loan applications from borrowers of various types. Upon acceptance of an application, she lends \$1 to the borrower and originates a loan with a cash flow of c dollars per unit of time until the borrower defaults.

The originator has a technology to privately screen loan quality to obtain *soft* information about the borrower by exerting effort (e.g., information about the borrower's job security or credit constraints). We assume that the loan's default intensity, denoted $\lambda(a) \in [\lambda_g, \lambda_b]$ with $\lambda_b > \lambda_g > 0$, improves (i.e., is lower) with increased screening effort $a \in [0,1]$. That is, $\lambda'(a) < 0$. We further assume that $\lambda(0) = \lambda_b$ and $\lambda(1) = \lambda_g$.

Importantly, investors cannot observe the screening effort a exerted by the originator. Exerting effort is costly, involving nonpecuniary cost $C(a;z) : [0,1] \times [\underline{z}, \overline{z}] \to \mathbb{R}^+$, where $z \in [\underline{z}, \overline{z}]$ denotes the loan's or lender's type. We assume

(i) For a given type z,

$$C(0;z) = 0,\ C'(0;z) = 0,\ C'(a;z) > 0, C''(a;z) > 0,\ \forall a \in (0,1]; \tag{2}$$

(ii) For $a \in [0,1]$ and $z \in [\underline{z}, \overline{z}]$,

$$\frac{\partial^2 C(a;z)}{\partial z \partial a} > 0. \tag{3}$$

We make the following assumption that will hold throughout the paper. The first condition ensures that the level of effort is interior in the first-best and a benchmark case with securitization frictions only. The second condition ensures that the second-order condition holds for the market equilibrium.

Assumption 1 Functions $C(\cdot;\cdot)$ and $\lambda(\cdot)$ are such that:

(i) For any $z \in [\lambda_g, \lambda_b)$, there exists $\widetilde{a}(z) \in (0,1)$ such that $\widetilde{a}(z) = \arg\max_{a \in [0,1]} \eta(a) - C(a;z)$, where $\eta(a) \equiv \frac{c}{r+\lambda(a)}$ denotes the expected payoff from selling the loan to investors.

(ii) For any $z \in [\lambda_g, \lambda_b)$, $\left(\frac{C'(a;z)}{\rho'(a)}\right)' + \Theta(a)\lambda'(a) > 0$, $\forall a > 0$, where $\Theta(a) > 0$ is defined in (24).

The Securitization Stage. In the securitization stage, the originator arrives with private information about the loan's default intensity, $\lambda(a)$. To simplify notation, we write $\lambda(a)$ as λ and refer to it as the originator's type since $\lambda(a)$ summarizes all payoff-relevant information from a and z. Figure 1 presents a timeline summarizing the sequence of events in the model.

An outcome of the game in the securitization stage is a triple $(\lambda, t, p) \in [\lambda_g, \lambda_b] \times \mathbb{R}^+ \times [p_b, p_g]$, where t and p denote the time and price at which trade takes place. Note that the sale price

p is bounded between the lowest and highest possible values to investors, denoted p_b and p_h, respectively:

$$\begin{aligned} p_b &\equiv \mathbb{E}\left[\int_t^{\infty} ce^{-r(u-t)} 1_{\tau_d \geq u} du \middle| \lambda_b\right] = \frac{c}{r+\lambda_b}, \\ p_g &\equiv \mathbb{E}\left[\int_t^{\infty} ce^{-r(u-t)} 1_{\tau_d \geq u} du \middle| \lambda_g\right] = \frac{c}{r+\lambda_g}. \end{aligned}$$

The value for the originator is thus given by

$$\begin{aligned} U(\lambda, t, p) &\equiv \mathbb{E}\left[\int_0^{t} ce^{-\gamma u} 1_{\tau_d \geq u} du + e^{-\gamma t} 1_{\tau_d \geq t} p \middle| \lambda\right] \\ &= \frac{c}{\gamma+\lambda}\left(1 - e^{-(\gamma+\lambda)t}\right) + e^{-(\gamma+\lambda)t} p. \end{aligned} \tag{4}$$

Equation 4 says that the value to the originator is made up of the discounted cash flows while they are holding the loan, plus the discounted transaction price. The originator's value function has the so-called single-crossing property: for each $(t,p) \in \mathbb{R}^+ \times [p_b, p_g]$, $-U_t(\lambda,t,p)/U_p(\lambda,t,p)$ is strictly monotone in λ. The single-crossing property is a key necessary condition for the existence of a separating equilibrium.

We next define equilibria in the securitization market.

Definition 1. *(**Securitization-market equilibrium**) An equilibrium in the securitization market is given by a pricing function $P(t) : \mathbb{R}^+ \to \mathbb{R}^+$ and the originator's time-to-sale strategy $T(\lambda)$, satisfying the following conditions:*

(i) Originator's Optimality: $T(\lambda) \in \arg\max_t U(\lambda, t, P(t))$.

(ii) Zero Profit Condition: $P(T(\lambda)) = \mathbb{E}\left[\int_t^{\infty} ce^{-r(u-t)} 1_{\tau_d \geq u} du \middle| T(\lambda)\right]$.

(iii) An equilibrium is separating if $P(T(\lambda)) = \frac{c}{r+\lambda}$.

We focus on characterizing a separating equilibrium in which the originator fully reveals the loan's private type λ through her time-to-sale decision, implying $P(T(\lambda)) = \frac{c}{r+\lambda}$. Fixing a price p, the originator with a lower default intensity λ would like to hold the loan longer to receive the flow of payments c and to signal her type through the delayed sale. As a result, the time to sale $T(\lambda)$ would depend on her type λ. On the other hand, the investors take the originator's decision into account and price the loan accordingly: $p = P(T(\lambda))$. In equilibrium, the originator chooses the optimal time to sale $T^*(\lambda)$ to maximize her utility $U(\lambda, t, P(t))$, while the investors earn zero expected profit: $P(T(\lambda)) = \frac{c}{r+\lambda}$.

Value at time 0. Given the optimal time-to-sale strategy $T(\lambda)$ and pricing function $P(t)$ in the securitization stage, the value to the originator at time 0 is

$$\begin{aligned} V(a;z) &= \mathbb{E}_a\left[\int_0^{T(\lambda(a))} ce^{-\gamma u} 1_{\tau_d \geq u} du + e^{-\gamma T(\lambda(a))} 1_{\tau_d \geq t} P(T(\lambda(a)))\right] - C(a;z) \\ &= \frac{c}{\gamma + \lambda(a)}\left(1 - e^{-(\gamma+\lambda(a))T(\lambda(a))}\right) + e^{-(\gamma+\lambda(a))T(\lambda(a))}\frac{c}{r+\lambda(a)} - C(a;z) \\ &= \frac{c}{\gamma+\lambda(a)} + e^{-(\gamma+\lambda(a))T(\lambda(a))}\left(\frac{c}{r+\lambda(a)} - \frac{c}{\gamma+\lambda(a)}\right) - C(a;z) \\ &\equiv \rho(a) + (1-q(a))(\eta(a) - \rho(a)) - C(a;z), \end{aligned} \tag{5}$$

where:

(i) $\mathbb{E}_a[\cdot]$ denotes the expectations operator over the cash flows of the originator that exerted effort a;

(ii) $\rho(a) \equiv \frac{c}{\gamma+\lambda(a)}$ represents the value to the originator from holding the loan in her portfolio without selling it;

(iii) $q(a) \equiv 1 - e^{-(\gamma+\lambda(a))T(\lambda(a))}$ denotes the discount factor applied to the value from holding the loan (i.e., $\rho(a)$).

As shown in (5), the value function $V(a;z)$ can be decomposed into three components. The first component (i.e., $\rho(a)$), as we note above, is the value of holding the loan without selling it. The second component is the additional net value from selling the loan in the future, which is $\eta(a) - \rho(a)$ discounted by $(1 - q(a))$. The last component is the cost from exerting effort. If the originator sells the loan immediately (i.e., $T = 0$), then $V(a;z) = \eta(a) - C(a;z)$. If the originator does not sell the loan (i.e., $T = \infty$), then $V(a;z) = \rho(a) - C(a;z)$.

Next, we define the equilibrium of the full game.

Definition 2. *(**Full-game equilibrium**) An equilibrium is given by* $\{a^*(z), P^*(t), T^*(\lambda)\} \in [0,1] \times \mathbb{R}_+^2$ *satisfying the following conditions:*

(i) Originator's Optimality at time 0*:* $a^*(z) = \arg\max_{a\in[0,1]} V(a;z)$*, given type* z*, and* $T^*(\lambda(a))$*, and* $P^*(t)$*.*

(ii) Securitization Market Equilibrium: $\{P^*(\cdot), T^*(\cdot)\}$ *represent the securitization market equilibrium outcome as in Definition 1.*

First-best. In the following proposition we characterize the first-best in which the originator has full access to the securitization market and there are no information frictions. As shown in Proposition 1, in the first-best, the originator sells the loan immediately, independent of her type and effort choice. The first-best achieves the full allocative efficiency as investors with higher valuation hold the loan from the very beginning. The originator's effort choice equalizes the social marginal benefit of effort with its social marginal cost.

Proposition 1. *In the first-best, the originator has full access to the securitization market and sells the loan immediately, $T^{FB}(\lambda) = 0$, $\forall \lambda \in [\lambda_g, \lambda_b]$, and exerts effort $a^{FB}(z) > 0$, $\forall z \in [\underline{z}, \overline{z}]$ at time 0 given by*

$$\eta'\left(a^{FB}\right) - C'\left(a^{FB}; z\right) = 0, \tag{6}$$

where $\eta(a) \equiv \frac{c}{r+\lambda(a)}$ denotes the expected payoff from selling the loan to investors as defined in Assumption 1(i).

2.1 Market equilibria

We now characterize the market equilibrium allocations of Definitions 1 and 2. We solve the model by backward induction. First, we solve the originator's problem in the securitization stage in which she chooses the optimal time to sale, given $\lambda(a)$ determined from her effort. Then, given the optimal time-to-sale strategies in the securitization stage, we solve her screening problem in the origination stage in which she chooses the optimal level of effort to screen the loan.

Separating equilibrium. We first characterize the unique separating equilibrium in the securitization stage in the following proposition, which is the same as in AGH.

Proposition 2. *In the unique separating equilibrium, the optimal time to sale is given by*

$$T(\lambda) = \frac{1}{\gamma - r} \log \frac{r + \lambda_b}{r + \lambda}, \ \forall \lambda \in [\lambda_g, \lambda_b], \tag{7}$$

and the pricing function is given by

$$P(t) = p_b e^{(\gamma - r)t}. \tag{8}$$

The main result in AGH holds here: when the expected loan quality is lower (larger λ), $T(\lambda)$ is smaller—the originator has an incentive to sell the loan sooner. The lemons problem arises here due to information asymmetry, in that the originator can perfectly observe the loan quality that is unobservable to the investors. An originator with a better quality loan waits longer to trade and uses the delayed trade as a signal of better quality.

Optimal effort. We now solve the originator's optimal effort decision in the following proposition.

Proposition 3. *In any market equilibrium, for the z-type loan, the optimal effort must satisfy the following first-order condition:*

$$1 + (1 - q(a))\left(\frac{(\gamma - r)(\gamma + \lambda(a))T(\lambda(a))}{r + \lambda(a)} - 1\right) = \frac{C'(a; z)}{\rho'(a)}, \tag{9}$$

where $q(a) \equiv 1 - e^{-(\gamma+\lambda(a))T(\lambda(a))}$. *The second-order condition holds under Assumption 1 (ii).*

The originator chooses effort such that the marginal benefit of exerting effort equals its marginal costs. By reducing default intensity, exerting effort generates several effects on the originator's value function that are captured by the left-hand side of equation (9). First, additional effort enhances the value derived from holding the loan (i.e., $\rho' > 0$). In addition, the second term on the left-hand side includes three additional effects that simplify to the expression inside the parentheses. Two factors incentivize the originator to exert effort: (i) Additional effort improves the net present benefit from selling the loan (i.e., $(1-q)(\eta' - \rho') > 0$ because the gap in interest rates between originator and investor is positive, so the higher expected cash flows are valued more highly by the investor); (ii) More effort also decreases the likelihood of loan default before sale making the sale is more likely (i.e., $(1-q)(\eta-\rho)(-\lambda' T) > 0$). However, that the originator has to delay the sale to mitigate information frictions dampens the incentive to exert effort (i.e., $(1-q)(\eta-\rho)(-(\gamma+\lambda)T'\lambda') < 0$). We show that under certain conditions the negative impact on effort due to the delayed sale outweighs the positive effects of effort from the higher benefits from selling the loan, resulting in the second term on the left-hand side of equation (9) being negative (see Lemma 1 in Appendix A).

Screen more, sell later. A central prediction in this paper is the positive relation between screening effort and signaling. The intuition is that signaling is costly — in our model, the cost of signaling is evident in the diminished allocative efficiency, wherein originators are compelled to hold on to originated loans for an extended duration to signal the underlying loan quality and receive a higher sale price. The anticipation of these outcomes incentivizes originators to exert more effort *ex ante*, giving rise to the positive relation between screening effort and signaling. As in Vanasco (2017), a delay in selling the loan plays a dual role: first, it serves as a signal for loan quality; second, it impacts the originator's ex-ante choice of the amount of screening.

The following proposition states that higher types separate from lower types by holding the loan longer before selling it in order to deter mimicry by low types who sell sooner. With a longer delay in sale, the dampening effect on effort due to the delayed sale diminishes relatively more for higher types. Consequently, the second term on the left-hand side of equation (9) becomes less negative, incentivizing more effort for higher types, which in turn further improves loan quality.

Proposition 4. *As the type* $z \in [\underline{z}, \overline{z}]$ *improves, the optimal effort, the time to sale, and loan quality all increase. That is,*

$$\frac{da^*(z)}{dz} < 0, \tag{10}$$

$$\frac{dT^*(\lambda(a^*(z)))}{dz} < 0, \tag{11}$$

$$\frac{d\lambda(a^*(z))}{dz} > 0. \tag{12}$$

Proposition 4 highlights the tension between asset quality and market liquidity, the focus of Vanasco (2017): screening improves asset quality but exacerbates information asymmetry, causing more delay in trade of the asset cash flows. This leads to one of the main model predictions in this paper, namely that the optimal effort, the time to sale, and loan quality all increase in the type of the loan. Figure 2 plots such a positive relationship between the time to sale and processing time in a numerical example.

Prediction 1. Using processing time as a proxy for effort, the model predicts a positive relationship between processing time and the time to sale. At the same time, an increase in either of them predicts an improvement in loan quality.

2.2 Extension: Introducing Hard Information

While the baseline model focuses on the role of soft information and unobservable loan quality, we consider in this section how the presence of hard information s observed by the originator and investors that influences the probability of securitization (for example, credit scores) affects the results. Specifically, we consider the second-best when there are securitization frictions such that the originator can only securitize and sell the loan with probability $\theta(s)$, but there are no information frictions. As in the first-best, there is no need to signal due to information symmetry (buyers and sellers of the loan all observe the same information), implying immediate loan sale.

However, the originator can still exert effort to improve loan quality. Here, we obtain an intuitive result: due to securitization frictions, there is a chance that the originator has to hold the loan in her portfolio. As a result, she is less incentivized to exert effort compared to the first-best (i.e., $a^{SB}(s,z) < a^{FB}(s,z)$ when $\theta(s) < 1$), reducing "productive efficiency."

Proposition 5. *In the second-best with securitization frictions only, the loan is sold immediately,* $T^{SB}(\lambda) = 0$ *and effort* $a^{SB} > 0$ *satisfies:* $\theta(s)\eta'\left(a^{SB}\right) + (1-\theta(s))\rho'\left(a^{SB}\right) - C'\left(a^{SB}; z\right) = 0$, *which can be written as*

$$1 + \theta(s)\left(\left(\frac{\gamma + \lambda\left(a^{SB}\right)}{r + \lambda\left(a^{SB}\right)}\right)^2 - 1\right) = \frac{C'\left(a^{SB}; z\right)}{\rho'\left(a^{SB}\right)}, \tag{13}$$

where $\rho(a) \equiv \frac{c}{\gamma + \lambda(a)}$ *denotes the expected payoff from holding the loan to the originator as defined in Assumption 1(i).*

Corollary 1 below shows that in the second-best with no information frictions, the effort is increasing in the hard-information signal s. Intuitively, with a more favorable signal s, there is a higher chance that the loan will be securitized and exerting more effort further enhances gains from trade. As we will see, this result does not hold in the market equilibrium with information frictions.

Corollary 1. *The second-best effort level is increasing in the hard-information signal s, and is decreasing in the soft-information about loan type z. That is,*

$$\frac{da^{SB}(s,z)}{ds} > 0, \tag{14}$$
$$\frac{da^{SB}(s,z)}{dz} < 0. \tag{15}$$

Hard and Soft Information. One benefit of introducing the threshold in the ease of securitization is that it allows us to draw a sharp distinction between hard and soft information. In fact, our model yields distinctly different predictions regarding how private (soft and unobservable) and public (hard, observable) information influences screening. We show below that, under certain conditions, observably better loans are associated with *lower* screening effort despite a higher likelihood of securitization.

Equation (16) below is the first-order condition for optimal effort in the presence of both information frictions (the originator observes z) and the public signal s that influences the likelihood of securitization:

$$1 + \theta(s)\left(1 - q(a)\right)\left(\frac{(\gamma - r)\left(\gamma + \lambda(a)\right) T\left(\lambda(a)\right)}{r + \lambda(a)} - 1\right) = \frac{C'(a;z)}{\rho'(a)}, \tag{16}$$

Compared with equation (9), the second term on the left hand side of the first-order condition now has an additional term $\theta(s)$, which captures the impact of the ease of securitization on effort incentives. Under the conditions described in Proposition (6) below, this second term is negative, so the impact on effort due to the delayed sale outweighs the positive effect from the higher probability of sale (see Lemma 1 in Appendix A). As a result, the originator has less incentive to exert effort when she faces more ease of securitization, consistent with the findings in Keys et al. (2010).

Proposition 6. *Under the condition* $\log \frac{r+\lambda_b}{r+\lambda_g} < \frac{r+\lambda_g}{\gamma+\lambda_g}$*, all else equal, an increase in the ease of securitization as a result of a more favorable signal s, the originator exerts less effort. That is,*

$$\frac{da^*(s,z)}{ds} < 0. \tag{17}$$

Compared with Proposition 5, in the presence of information frictions, the effort level in the market equilibrium is *decreasing* in signal s, rather than *increasing* as in the second-best case. As in the second-best, the originator still has an incentive to exert more effort to increase gains from trade from selling at a higher price. However, unlike in the second-best, to mitigate information frictions, the originator has to delay the sale longer to signal better quality, dampening the incentive to exert effort ex ante. With more ease of securitization from a more favorable signal s, the dampening effect dominates, resulting in lower effort.

Securitization rule of thumb. A fundamental challenge in testing the central predictions of our model is the difficulty in measuring agents' hidden effort. As we discuss in detail in the next Section, we address this challenge by using mortgage processing time as a measure of effort. To establish evidence for processing time as a sensible measure of hidden effort, we utilize a rule of thumb in the securitization market: loans above the FICO threshold of 620 were more easily securitized during the time period of our sample (Keys et al. (2010), (Keys et al., 2012)). This rule of thumb implies a discrete positive increase in the ease of securitization across the 620 threshold.

To capture the "rule of thumb" in the model, we assume that there exists a securitization threshold, denoted by $s^* \in [\underline{s}, \overline{s}]$, such that there is a positive jump in the securitization probability at s^*, that is,

$$\theta\left(s_+^*\right) \equiv \lim_{s \downarrow s^*} \theta\left(s\right) > \theta\left(s_-^*\right) \equiv \lim_{s \uparrow s^*} \theta\left(s\right).$$

One important implication from the existence of a rule of thumb in the market and Proposition 6 is that the discontinuity in the ease of securitization around the threshold gives rise to discontinuities in effort and loan quality, or, more precisely, a negative jump in effort and a positive jump in default likelihood for loans right above the threshold than those right below.

Proposition 7. *Suppose the securitization probability jumps up at a threshold s^*, then under the condition $\log \frac{r+\lambda_b}{r+\lambda_g} < \frac{r+\lambda_g}{\gamma+\lambda_g}$, for any $z \in [\lambda_g, \lambda_b]$, in the close vicinity of the threshold s^*:*

$$\begin{aligned} a_-^* &\equiv a^*\left(s^* - \Delta s; z\right) > a_+^* \equiv a^*\left(s^* + \Delta s; z\right), & (18)\\ \lambda_-^* &\equiv \lambda\left(a_-^*; z\right) < \lambda\left(a_+^*; z\right) \equiv \lambda_+^*, & (19)\\ T_-^* &\equiv T\left(\lambda_-^*\right) > T\left(\lambda_+^*\right) \equiv T_+^*, & (20)\end{aligned}$$

where $a_-^* \equiv a^*\left(s^* - \Delta s; z\right)$ and $a_+^* \equiv a^*\left(s^* + \Delta s; z\right)$ denote the optimal effort when $s = s^* - \Delta s$ and $s^* + \Delta s$, respectively, given z.

Based on Proposition 7, there should be a discontinuity in processing time and loan quality around the securitization threshold, reminiscent of the findings in Keys et al. (2010). Figure 4 numerically illustrates this prediction.

> **Prediction 2.** If processing time indeed proxies for hidden effort, our model predicts that both processing time and loan quality drop for loans right above the 620 threshold than those right below, because of more ease of securitization for loans with FICO greater than or equal to 620.

Processing time, observable and unobservable default risk. It is worth stating that the results above generates another key prediction for default risk and processing time. Specifically, if loans with better hard information (signal s) are perceived to have lower observable default risk, then our model predicts a *negative* relationship between processing time and observable default risk. This is because the equilibrium effort is lower in response to a better hard information signal s due to the greater ease of securitization (Proposition 6).

> **Prediction 3.** Let $\widehat{\lambda}(s)$ denote the default rate predicted by observable signals s such that $\frac{d\widehat{\lambda}(s)}{ds} < 0$, then fixing loan type z, as signal s decreases, the equilibrium effort increases while the predicted default rate increases, resulting in a *positive* relationship between $\widehat{\lambda}(s)$ and $a(s; z)$. In contrast, the *unobservable* default risk $\lambda(a; z)$ is *negatively* related to processing time, because fixing signal s, $\lambda(a; z)$ is increasing in loan type z, but $a(s; z)$ is decreasing (Proposition 4).

The model implies that the originator should exert *more* effort for loans of better *unobservable* types (Prediction 1), but *less* effort for loans with *observable* better hard information signals. This highlights the tensions between observable vs. unobservable loan characteristics (or put differently, hard vs. soft information) and their different implications on effort incentives. Figure 3 illustrates the predicted relationships in a numerical example.

The tension between hard versus soft information not only exists in the relationship between default risk and processing time stated in the above prediction, but also in the relationship between time to sale and processing time.

Consider two loans: loan H has signal $s = s_H$ and while loan L has signal $s = s_L$. Suppose $s_H > s_L$ such that loan H has a higher chance to be securitized and sold than loan L. Suppose in equilibrium, these two loans have the same level of effort, denoted by a^*. Then from Propositions 4 and 6, we can prove that their unobservable loan types must satisfy $z_L > z_H$. Suppose the opposite holds, i.e., $z_L \le z_H$. Then from Proposition 4, we have $a^*(s_L, z_H) \le a^*(s_L, z_L) = a^*$. However, from Proposition 6, we have $a^* = a^*(s_H, z_H) < a^*(s_L, z_H)$. Therefore, we obtain $a^* < a^*(s_L, z_H) \le a^*$, which is a contradiction. As a result, it must be true that $z_L > z_H$. It follows that $\lambda(a^*; z_H) < \lambda(a^*; z_L)$, implying that $T(\lambda(a^*; z_H)) > T(\lambda(a^*; z_L))$. That is, the model generates the following prediction.

> **Prediction 4.** If two loans have the same equilibrium effort, then the loan with a better hard-information signal must have a better unobservable loan type (soft information), and thus will be sold with a longer delay. Therefore, controlling for processing time, time to sale has additional explanatory power regarding loan quality.

This prediction is illustrated in Figure 2. Intuitively, the originator typically has less incentive to exert effort for a loan with a better hard-information signal s, unless it also has a better soft-information signal z. So if loan H has the same effort as loan L despite its better signal $s_H > s_L$, then it must be the case it has a better soft-information loan type or $z_H < z_L$, which implies that loan H has a better overall quality and the originator would delay its sale more to signal better quality than loan L.

3 Data and summary statistics

Our data come primarily from two sources: the confidential Home Mortgage Disclosure Act (HMDA) and the CoreLogic LoanPerformance databases. We merge these two databases to examine the relationship between processing time, delay of sale, and loan default. The sample period is from 2002 to 2006.

The confidential HMDA database provides the exact application date and action date (approval or denial) for a given mortgage. We calculate processing time for a given loan—one key variable of interest in this paper—as the difference between these two dates. Note that the public version of this database cannot be used for such calculation because it only reports the year of action date.

The CoreLogic LoanPerformance database provides loan performance information on whether a loan is current, delinquent, or in foreclosure for securitized residential mortgages.[2] We use loan default within fifteen months of origination as our primary loan performance measure. Following the convention in the mortgage loan industry, a loan is classified as being in default if payments on the loan are 60+ days late as defined by the Office of Thrift Supervision, or the loan is in foreclosure or real estate owned (REO) at any point within 10 to 15 months of origination.[3]

To examine the relation between processing time, delay of sale, and loan default, we merge these two databases by using the application and action dates together with the loan amount and other loan characteristics (see Appendix B1 for more details). The merged data contain detailed information about borrower and loan characteristics. Specifically, we have information on borrower credit risk characteristics at the loan origination, including the FICO score, the CLTV ratio (including first and second liens), the back-end debt-to-income (DTI) ratio, and whether the lender has complete documentation on the borrower's income and assets. As for loan specifics, the merged data also includes information on whether the loan rate is fixed or adjustable, the initial loan rate, the margin, and the first rate reset for adjustable rate loans, and whether the loan has features such as a prepayment penalty or balloon payment at maturity. We control for all of these borrower and loan characteristics in our analyses.

We supplement these two databases with additional data on macroeconomic conditions. Specifically, we collect macro variables such as local housing price appreciation, state-level unemployment

[2] As noted by Keys et al. (2010), the CoreLogic LoanPerformance database encompasses over 90% of the mortgage loans that are privately securitized by MBS issuers.

[3] Alternatively, we have also considered 90 days past due or in foreclosure for default status and obtain qualitatively similar results with both alternatives.

rate, and local median household income to control for the overall economic environment. We identify the borrower's geographic area for each loan in the sample using the five-digit ZIP code. Specifically, we compute housing price appreciation (HPA) from 36 months before loan origination using the housing price index for the borrower's county reported by the Federal Housing Finance Agency (FHFA). We use the median household income in 1999 for the borrower's ZIP code as reported by the U.S. Census Bureau in 2000. Definitions for the key variables from these databases are given in Appendix B2.

We collect affiliation information about originators from the Residential Mortgage-Backed Security (RMBS) deal offering prospectuses. The deal prospectus supplements provide information on originator and sponsor identities, and the percentage of mortgages (weighted by original loan balance) originated by each originator. We construct the deal-level originator-sponsor affiliation measures following Demiroglu and James (2012). Specifically, we define originator-sponsor affiliated deals as those in which more than 50% of the loans are originated by the sponsor-affiliated originators, and vice versa. In our sample, the affiliated (unaffiliated) deals have, on average, about 95% (5%) of the loans originated by sponsor-affiliated originators.

Our selection of the loan sample largely follows Keys et al. (2010), so the sample includes primarily "subprime" securitized loans. Specifically, we select from the CoreLogic LoanPerformance dataset loans that are for home purchase (not for refinance), owner-occupied single-family residences, townhouses, or condominiums. We exclude non-conventional properties, buy-down mortgages, Alt-A loans, and loans with missing FICO scores. We also combine limited and no-documentation borrowers and refer to them as low documentation borrowers.

3.1 Summary statistics

In Table 1, we report the summary statistics of our sample by the origination year of the mortgage loans. The sample comprises about 1.5 million loans, including 0.6 million low documentation loans and 0.9 million full documentation loans. The number of loans rose fast during our sample period and peaked in 2005. The average FICO score is 633 in our sample of subprime mortgages. The average CLTV ratio is much higher than 80% (the common 20% down payment), around 93%, and is at its highest in 2006. Comparing the low-documentation loans with those that are fully documented, we find that the FICO scores are lower, and the CLTV ratios are slightly higher for full documentation loans relative to low documentation loans.

Turning to the delay of sale, the sample average time to sale is 13.7 weeks, similar to AGH. The sample average time to sale is slightly longer for full documentation loans and, notably, full documentation loans have shorter time to sale than low documentation loans in 2006, a reversal from the previous years.

The sample average delinquency rate is 12%, rising from 8.3% in 2002 to 20.4% in 2006. The average delinquency rate is 12.6% for low documentation loans and 11.7% for full documentation loans. The delinquency rate for lower-doc loans is significantly lower than that for full documentation loans in the early part of the sample period but becomes much higher in 2006.

Of particular interest to our paper, our sample's average processing time is 3.2 weeks, similar for both low documentation and full documentation loans. The average processing time shows a downward trend even though the number of loans increases significantly over the sample period. A distinct reduction in processing time occurred starting in 2004, consistent with lower average quality of loans and the need for quicker approvals during the peak of the housing boom (Adelino, Schoar, and Severino, 2016). Figure 5 presents the distribution of processing time. We observe the long right tail in the distribution well above eight weeks.

Comparing low documentation and full documentation loans in Table 1, we observe the average processing time for low documentation loans is longer than for full documentation loans before 2004, but shorter after 2004. Delinquency rates are also significantly higher for low documentation loans in the later cohorts. These dynamics in processing time and delinquency could reflect changes in lender screening. Mortgage lenders screen applicants by collecting both "hard" information, such as credit scores that can be easily verified and credibly transmitted, and "soft" information that is not easily transmitted to investors but can help predict, for example, borrowers' future income stability. We study below the extent to which the cost borne by originators in collecting soft information can be transmitted to investors through signalling, and is related to *ex post* conditional performance.

4 Empirical tests

In this section, we present the empirical results testing the predictions from our model linking mortgage processing time to delay of sale in our sample of non-agency subprime securitized loans originated between 2002 and 2006. These tests serve the joint purpose of testing both the validity of processing time as a measure of originator effort, as well as the main predictions of the model in Section 2.

4.1 Discontinuity around the 620 FICO score

Our model predicts discontinuities in processing time as well as default intensity around the threshold that determines whether a loan can be sold or not (Proposition 7). Empirical evidence for such discontinuities in the data support the model predictions and, more importantly, validate our proposal of using processing time to proxy for lenders' screening efforts.

We exploit a key insight from Keys et al. (2010) that a FICO score of 620 can serve as a threshold for the ease of loan securitization, and that the effect of this threshold is more acute for low documentation loans (Keys et al., 2012). Fannie Mae and Freddie Mac first established a FICO score of 620 as the threshold for origination in the mid-1990s (Avery, Bostic, Calem, and Canner, 1996; Capone, 2002; Bubb and Kaufman, 2014) and required further inquiry from the lender for loans from borrowers with FICO scores below 620. As the subprime private-label securitization market grew in the early 2000s, and following the GSEs' lead, subprime mortgage-backed investors demanded securitized loans above the credit cutoff and rendered 620 as a rule of thumb in the securitized subprime lending market (Keys et al., 2012). Rating agencies and top originators widely use this cutoff in the subprime market. By comparing loans on either side of the credit score threshold with otherwise nearly identical risk profiles, we can examine whether differential access to securitization led to changes in the behavior of lenders.

We apply a regression discontinuity design (see, e.g., DiNardo and Lee, 2004; Card, Mas, and Rothstein, 2008) for mortgage processing time around the FICO cutoff score 620. When lenders screen borrowers above 620 to a lesser extent than below, we expect a negative jump in processing time for FICO scores over 620. We choose a relatively narrow range for FICO scores with 20 points on either side of the cutoff and run loan-level regressions of the following form:

$$PT_{i,t} = \alpha + \beta \times \mathbf{1}_{FICO \geq 620} + \gamma \times X_{i,t} + \delta_{i,t} + \epsilon_{i,t}, \tag{21}$$

where $PT_{i,t}$ is the processing time for loan i in period t, $\mathbf{1}_{FICO \geq 620}$ is an indicator variable that equals one if its FICO score is greater than or equal to the threshold of 620, or zero, otherwise. We include other explanatory variables in $X_{i,t}$, including controls for borrower and loan characteristics and local economic conditions. We also consider various fixed effects for loan origination year, state, and mortgage lender, labeled by $\delta_{i,t}$. The year of origination fixed effects account for the potential time trend in processing time, and the state fixed effects can account for the potentially uneven distribution of FICO scores and delinquencies across geographical locations. The lender fixed effects allow examination of within-lender variations and remove between-lender variations. We include the specifications with and without lender fixed effects. While processing time is likely to vary systematically across lenders, we show in the Internet Appendix that there is substantial variation in mortgage processing times even after accounting for lender fixed effects. The coefficient β measures the magnitude of a discontinuity if it exists in the data.

Panel A of Table 2 reports the regression discontinuity design results. We conduct our analysis for the whole sample and the low documentation and full documentation loans separately. Our loan-level findings confirm that there is a discontinuity at the 620 threshold: The loans with FICO scores 620+ are processed faster than loans with FICO scores 620-, especially for low documentation loans.

The results in Panel A should be interpreted in combination with those on defaults, i.e., we also test whether loans immediately above the threshold of 620 have a jump in default frequency. This is, in essence, simply confirming that we observe in our sample the same results as Keys et al. (2010) and Keys et al. (2012). In Panel B of Table 2, we report the results from regressions similar to equation (21) except that the dependent variable is the observed loan default. The results confirm the existence of a positive jump in default intensity by about 1.3% for the low documentation loans immediately above the threshold of 620.

The combined results in Panels A and B suggest that processing time is a valid measure of lenders' screening efforts. In fact, by directly measuring processing time and showing its association with defaults we provide direct evidence of the mechanism of lax screening by mortgage lenders proposed by Keys et al. (2010) based on the differences in defaults around the threshold.

Figure 6 presents the RDD plot of processing time, delay of sale, and loan delinquency. We use conditional versions of dependent variables in our RDD plot. Specifically, we keep the residuals from the regressions of processing time, delay of sale, or delinquency on loan and borrower characteristics, state, and year fixed effects, without the dummy variable for the FICO 620 threshold. By aggregating the regression residuals for each point in the FICO score, we generate the RDD plot for processing time and loan delinquency. The patterns in Figure 6 are consistent with our regression results in Table 2, with the addition of the jump in delay of sale, also present in Keys et al. (2012). They also represent initial evidence of the connection between screening (processing time) and signaling (delay of sale) that we propose in the model as illustrated in Figure 4.

4.2 Processing time and (un)observable default risk

We next conduct loan-level analysis on the relation between loan processing time and loan default. Loan default can depend on observable borrower and loan characteristics and unobservable information about the borrower's creditworthiness. Lenders' screening efforts can affect the unobservable component of default, including information about occupation, income volatility, unobserved neighborhood and property characteristics, among others. If processing time captures the lender's efforts, we expect processing time to negatively correlate with the *unobservable* component of default, as argued in Adelino et al. (2019) and in Proposition 4, Equation 10, and Figure 3 of this paper.

We start by regressing loan delinquency on processing time, controlling for observable loan and borrower characteristics, origination year, state fixed effects, and lender fixed effects:

$$Default_{i,t} = \alpha + \beta \times ProcessingTime + \gamma \times X_{i,t} + \delta_{i,t} + \epsilon_{i,t}, \tag{22}$$

where variables are labeled as in the previous section for equation 21.

We start by showing a version of this regression where we discretize processing time into weeks and create dummy variables from one week to eight weeks and above. In Panel A of Figure 7, we plot the coefficients of processing time dummy variables, using the loans with a processing time below one week as the base group. Given the extensive list of control variables, differences in default across loans are plausibly related to lender private information, as in Adelino et al. (2019). The fact that longer processing time is associated with lower abnormal default rates suggests that processing time is (at least in part) used for lender screening and collection of unobservable soft information. Relative to loans processed in just one week, defaults are 0.5 percentage points lower if they are processed in 4 weeks, and this effect botomms out at about 0.7 percentage points for loans processed in 8 weeks or longer. As we discuss below about Panel B of this Figure, this is in stark contrast with the relation of processing time and observable risk.

We also examine the relationship between processing time and observable risk, which we measure as the component of default that can be predicted based on loan and borrower characteristics. We generate predicted default probability for each loan based on a logistic model estimated with a two-year rolling window. Specifically, for each loan, we run logit regressions of default on all available borrower and loan variables using the previous 2 years of data (so, for 2005 loans, we use 2003 and 2004), and obtain the predicted probability of default using each loan's characteristics and the coefficients estimated in the regression.[4] In the second step, we repeat the above regression 22, but replace realized default with predicted loan default probability as the dependent variable.

In Panel B of Figure 7, we plot the coefficients for predicted defaults (observable risk) on processing time dummy variables. In stark contrast with the results on excess defaults in Panel A (which is the measure of unobservable risk), and as predicted in our model (Proposition 6), the processing time is significantly positively correlated with the predicted default probability. Predicted default is 0.07 percentage points higher for loans processed in 5 to 7 weeks relative to those processed in one week or less. Note that while this magnitude is small, and significantly smaller than the one in Panel A, the striking fact is that the figure shows the opposite general pattern to what we observe for excess defaults.

Taken together, these results suggest that when processing time is long, the observable ex-ante default risk is high, and yet the unobservable ex-post default risk is reduced. This is consistent with our model predictions and further indicates that processing time is correlated with lender screening effort.

4.3 Processing time and delay of sale

In this section, we test the model prediction of a positive correlation between mortgage processing time and delay of sale among sold loans (Proposition 4). Figure 8 presents the scatter plot between processing time and average delay of sale for each bin of processing time in the whole sample and subsamples divided by borrowers' FICO scores (split at a credit score of 640).

We observe a strong positive relation between processing time and delay of sale, and this positive relation is more pronounced at higher processing time, starting at about 4 weeks of processing time. We also observe a stronger positive relation for loans with higher FICO scores. Both empirical patterns are consistent with the model prediction illustrated in Figure 2. In particular, the model predicts a steeper relationship between lender effort and delay of sale for a high public signal of quality (Figure 2), which in the regressions is represented by an above-median FICO score.

In Table 3, we report estimates from a regression of loan-level delay of sale on processing time, controlling for loan and borrower characteristics, local macroeconomic conditions, origination year, state, and lender fixed effects. In the sample of all loans, we find a significant positive relationship between processing time and time to sale, consistent with the model prediction in Proposition 4. The coefficient estimate of processing time indicates that conditional default rates drop by 0.04 percentage points on average for each additional week of processing time. We also find that loans with higher FICO scores, higher CLTV, higher loan rates, prepayment penalties, or hybrid loan features are sold more quickly.[5]

We further test the model prediction in the subsamples of loans from originator-sponsor affiliated and unaffiliated deals, and low documentation and full documentation loans. We find a stronger

[4] The starting year with predicted default is 2003 based on loans originated in 2002, a one-year window.

[5] The estimates for control variables are reported in the Internet Appendix.

relationship between processing time and delay of sale in the subsample of loans from unaffiliated deals. As explained in Adelino et al. (2019), information asymmetry between mortgage originators and RMBS issuers is more severe among unaffiliated loans and thus likely to induce more originator signaling. Between low documentation and full documentation loans, we find the estimate of processing time is statistically and economically more significant in full documentation loans. The difference in the estimates can arise from the dependence of optimal screening efforts and delay of sales on the unobservable loan types within each subsample, as defined in Proposition 4.

To examine the convexity in the relationship between processing time and time to sale, we further split the sample by processing times into three groups: below four weeks, between four and eight weeks, and above eight weeks. We choose these cutoff values based on Figure 8. We find a flat relationship between processing time and delay of sale for the first group and a more positive relationship for the other two groups. This convexity, i.e., a higher sensitivity of delay of sale to processing time is consistent with the model predictions.

It is useful to consider to what extent the relation between processing time and delay of sale might reflect confounding factors beyond the list of control variables in our regression analysis. For instance, long loan processing time could be due to delay by borrowers to close, rather than lender screening effort. One specific such factor is that liquidity-constrained borrowers are more likely to close on a home purchase near the month-end to reduce the interest payment in the closing cost or save on the rent, as shown in Bhutta and Ringo (2021). The loans from these borrowers might have unobservably lower loan types in our model.

For robustness, we redo our analysis excluding the loans closed near the month-end. We exclude loans that close after the 25th of each month, experiment with other cutoff dates, and find similar results. An earlier cutoff date may remove many borrowers who do not delay closing, and a later cutoff date might plausibly miss some liquidity-constrained borrowers. We find that, after excluding these loans, the relation between processing time and delay of sale becomes stronger for the low documentation loans and remains the same for the full documentation loans.[6] This finding is consistent with our conjecture that borrowers with month-end closings are more likely to have unobservable lower loan types in our model.

It is important to stress that any alternative hypothesis about the correlation between unobservable loan types, processing time and delay of sale also should reconcile the evidence on observable risk (discussed in Section 4.2) and on unobservable risk, or conditional excess mortgage default, also addressed in Section 4.2 and discussed in more detail below. If we say higher quality borrowers (rather than lower quality, as we discuss above) take longer to close, it is unclear why those loans would take longer to sell. In fact, gains from trade would predict that observably higher quality loans would be sold quicker, not slower, as our model indicates. In addition, it is also unclear why the additional processing time would correlate with delay of sale and also with observable and unobservable risk in the way we observe in the data. If observables and unobservables are positively correlated, that would work against us finding a positive relationship between processing

[6] The results are reported in the Internet Appendix.

time, delay of sale, and unobservable quality.

4.4 Processing Time, Delay of Sale, and Mortgage Default

Next, we test our model prediction on the relation between loan default, processing time, and delay of sale. In Proposition 4 of our model, a better unobservable loan type is associated with longer processing time, longer delay of sale, and lower default risk, controlling for observable loan and borrower characteristics. We thus expect that processing time and delay of sale both to predict loan default, as long as one measure is not a sufficient statistic for the other. Given the noise in observing screening effort and delay of sale (for a variety of institutional constraints, e.g., time to warehouse loans, market conditions, etc.), we do not expect in our empirical exercise for one of the variables to fully absorb the other's explanatory power.

In Table 4, we examine the relation between loan delinquency after origination, loan processing time and delay of sale, controlling various loan and borrower characteristics, local housing price and macro variables, and origination year, state, and mortgage lender fixed effects.

In column (1), processing time is significantly negatively associated with loan delinquency, which is consistent with our validation of processing time as a measure of screening effort. In column (2), delay of sale is significantly negatively associated with loan delinquency, consistent with the findings in AGH. In column (3), we include both processing time and delay of sale in the regression for loan delinquency. We find that the estimates of both variables remain statistically and economically significant. Specifically, the delinquency rate increases by 0.10% (0.03%) on average when processing time (delay of sale) increases by one week. This result suggests that processing time and time to sale complement each other in predicting delinquency, as loans more carefully screened or sold with a delay have lower conditional default risks.

Our model predicts that the loans of the lowest type are both processed and sold quickly and have the highest default risk. To test this prediction, we generate dummy variables for loans with processing time or time to sale below the median in the origination year, the interaction term of two dummy variables to capture quick processing and sale. We include the two dummy variables in column (4) and the additional interaction term in column (5). In column (4), the delinquency rates are 0.42% (0.41%) higher on average when processing time (delay of sale) is below the median. In column (5), we find that the loans have the highest default risk when both processing time and delay of sale are below the median, consistent with our model that these loans have the lowest loan type. The significance of the interaction term helps us identify the unobservable loan type in our model. We also acknowledge that both the model and the empirical setup may not fully capture other sources of signals (for example, buyer signals as in Kaya and Kim (2018) or reputation concerns as in Hartman-Glaser (2017)) that may induce a complex relation between time to sale and asset quality.

In Table 5, we examine the relation between loan default, processing time, and delay of sale for the subsamples of loans from originator-sponsor affiliated and unaffiliated deals, as well as low documentation and full documentation loans. We find that our results hold for all subsamples.

This is consistent with our finding that screening (processing time) and signaling (delay of sale) are positively correlated in all subsamples.

5 Conclusion

This paper explores the relationship between screening effort and signaling in the mortgage market, focusing on the trade-off between originating high-quality assets and maintaining secondary market liquidity that emerges from dynamic models with asymmetric information. Our model builds and extends Vanasco (2017) to show that increased screening at origination leads to more signaling, particularly through delayed sales. The model also provides a sharp distinction between the role of hard information (or observable risk) that may lead to variation in the probability of securitization and that of soft information acquired through originator effort and that is unobservable to the investors.

We use U.S. private-label securitized mortgage data from 2002 to 2006 and employ mortgage processing time as a measure of screening effort. We have three main empirical results. First, processing time is positively associated with observable credit risk but negatively correlated with conditional ex-post mortgage default, consistent with key predictions in the models. Second, we show that a discontinuity in default rates around 620 FICO scores emphasized by Keys et al. (2010) is accompanied by a discontinuity in lender effort measured by processing time. Finally, the paper establishes a positive relationship between processing time and delay of sale, suggesting that higher screening effort corresponds with more signaling in the market, and that both are related to higher unobserved quality measured by *ex post* conditional default.

Overall, we provide the first empirical test of a robust prediction in the theoretical literature linking screening effort and delay of sale. This approach also opens the door to further investigation of the role of asymmetric information and lender effort in other asset markets.

References

Adelino, M., K. Gerardi, and B. Hartman-Glaser (2019): "Are lemons sold first? Dynamic signaling in the mortgage market," *Journal of Financial Economics*, 132, 1–25.

Adelino, M., A. Schoar, and F. Severino (2016): "Loan originations and defaults in the mortgage crisis: The role of the middle class," *The Review of Financial Studies*, 29, 1635–1670.

Akerlof, G. A. (1970): "The Market for "Lemons": Quality Uncertainty and the Market Mechanism," *The Quarterly Journal of Economics*, 84, 488–500.

Avery, R. B., R. W. Bostic, P. S. Calem, and G. B. Canner (1996): "Credit Risk, Credit Scoring and the Performance of Home Mortgages," *Federal Reserve Bulletin*, 82, 621–648.

Bedayo, M., G. Jiménez, J.-L. Peydró, and R. Vegas Sánchez (2020): "Screening and loan origination time: lending standards, loan defaults and bank failures," .

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

Ben-Rephael, A., B. I. Carlin, Z. Da, and R. D. Israelsen (2023): "Security Analysis and the Collection of Hard and Soft Information," Working Paper 31936, National Bureau of Economic Research.

Bhutta, N. and D. Ringo (2021): "The effect of interest rates on home buying: Evidence from a shock to mortgage insurance premiums," *Journal of Monetary Economics*, 118, 195–211.

Bubb, R. and A. Kaufman (2014): "Securitization and moral hazard: Evidence from credit score cutoff rules," *Journal of Monetary Economics*, 63, 1–18.

Capone, C. (2002): "Research into Mortgage Default and Affordable Housing: A Primer," LISC Home Ownership Summit 2001.

Card, D., A. Mas, and J. Rothstein (2008): "Tipping and the Dynamics of Segregation," *Quarterly Journal of Economics*, 123, 117–128.

Chang, B. (2017): "Adverse selection and liquidity distortion," *The Review of Economic Studies*, 85, 275–306.

Chemla, G. and C. A. Hennessy (2014): "Skin in the Game and Moral Hazard," *Journal of Finance*, 69, 1597–1641.

Choi, D. B. and J.-E. Kim (2021): "Does securitization weaken screening incentives?" *Journal of Financial and Quantitative Analysis*, 56, 2934–2962.

Daley, B. and B. Green (2012): "Waiting for News in the Market for Lemons," *Econometrica*, 80, 1433–1504.

——— (2016): "An Information-Based Theory of Time-Varying Liquidity," *Journal of Finance*, 71, 809–870.

Daley, B., B. Green, and V. Vanasco (2020): "Securitization, Ratings, and Credit Supply," *The Journal of Finance*, 75, 1037–1082.

Dell'Ariccia, G. and R. Marquez (2006): "Lending Booms and Lending Standards," *The Journal of Finance*, 61, 2511–2546.

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

DeMarzo, P. M. and D. Duffie (1999): "A Liquidity-Based Model of Security Design," *Econometrica*, 67, 65–100.

Demiroglu, C. and C. James (2012): "How Important is Having Skin in the Game? Originator-Sponsor Affiliation and Losses on Mortgage-backed Securities," *Review of Financial Studies*, 25, 3217–3258.

DiNardo, J. and D. S. Lee (2004): "Economic Impacts of New Unionization on Private Sector Employers: 1984-2001," *Quarterly Journal of Economics*, 119, 1383–1441.

Foote, C. L., L. Loewenstein, and P. S. Willen (2019): "Technological Innovation in Mortgage Underwriting and the Growth in Credit, 1985–2015," Federal Reserve Bank of Boston Research Department Working Papers No. 19–11. https://doi.org/10.29412/res.wp.2019.11.

Fuchs, W., B. Green, and D. Papanikolaou (2016): "Adverse selection, slow-moving capital, and misallocation," *Journal of Financial Economics*, 120, 286–308.

Fuchs, W. and A. Skrzypacz (2015): "Government interventions in a dynamic market with adverse selection," *Journal of Economic Theory*, 158, 371–406.

——— (2019): "Costs and benefits of dynamic trading in a lemons market," *Review of Economic Dynamics*, 33, 105–127.

Fuster, A., M. Plosser, P. Schnabl, and J. Vickery (2019): "The role of technology in mortgage lending," *The Review of Financial Studies*, 32, 1854–1899.

Gorton, G. B. and G. G. Pennacchi (1995): "Banks and loan sales Marketing nonmarketable assets," *Journal of Monetary Economics*, 35, 389–411.

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

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

Hartman-Glaser, B. (2017): "Reputation and signaling in asset sales," *Journal of Financial Economics*, 125, 245–265.

Hartman-Glaser, B., T. Piskorski, and A. Tchistyi (2012): "Optimal securitization with moral hazard," *Journal of Financial Economics*, 104, 186–202.

He, Z., S. Jiang, and D. Xu (2023): "Technology-driven Intermediation," Working Paper.

Janssen, M. C. and S. Roy (2002): "Dynamic Trading in a Durable Good Market with Asymmetric Information," *International Economic Review*, 43, 257–282.

Kaya, A. and K. Kim (2018): "Trading Dynamics with Private Buyer Signals in the Market for Lemons," *The Review of Economic Studies*, 85, 2318–2352.

Keys, B. J., T. Mukherjee, A. Seru, and V. Vig (2010): "Did Securitization Lead to Lax Screening? Evidence from Subprime Loans," *Quarterly Journal of Economics*, 125, 307–362.

Keys, B. J., A. Seru, and V. Vig (2012): "Lender Screening and the Role of Securitization: Evidence from Prime and Subprime Mortgage Markets," *Review of Financial Studies*, 25, 2071–2108.

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

Leland, H. E. and D. H. Pyle (1977): "Informational Asymmetries, Financial Structure, and Financial Intermediation," *The Journal of Finance*, 32, 371–387.

Malherbe, F. (2012): "Market Discipline and Securitization," Working Paper, London Business School.

Myers, S. C. and N. S. Majluf (1984): "Corporate financing and investment decisions when firms have information that investors do not have," *Journal of Financial Economics*, 13, 187–221.

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

Parlour, C. A. and G. Plantin (2008): "Loan Sales and Relationship Banking," *Journal of Finance*, 63, 1291–1314.

Rosen, R. (2011): "Competition in mortgage markets: The effect of lender type on loan characteristics," Federal Reserve Bank of Chicago Economic Perspectives.

Spence, M. (1973): "Job Market Signaling," *The Quarterly Journal of Economics*, 87, 355–374.

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

Vanasco, V. (2017): "The Downside of Asset Screening for Market Liquidity," *The Journal of Finance*, 72, 1937–1982.

Williams, B. (2016): "Search, liquidity, and retention: signaling multidimensional private information," Tech. rep., Unpublished working paper, New York University.

Appendix A: Proofs

In this appendix, we first prove Lemma 1 and then provide proofs of the propositions in the main text.

Lemma 1. *Under the condition* $\log \frac{r+\lambda_b}{r+\lambda_g} < \frac{r+\lambda_g}{\gamma+\lambda_g}$, *then for any* $\lambda \in [\lambda_g, \lambda_b]$, *we have*

$$T(\lambda) < \frac{r+\lambda}{(\gamma - r)(\gamma + \lambda)}, \tag{23}$$

where $T = \frac{1}{\gamma - r} \log \frac{r+\lambda_b}{r+\lambda}$ *is defined in Proposition 2.*

Proof of Lemma 1. To simplify notation, we omit arguments of $\lambda(a(z))$ and $T(\lambda(a(z)))$ and write them as λ and T. Denote $f(\lambda) \equiv T - \frac{r+\lambda}{(\gamma-r)(\gamma+\lambda)} = \frac{1}{\gamma-r}\left(\log \frac{r+\lambda_b}{r+\lambda} - \frac{r+\lambda}{\gamma+\lambda}\right)$.

Note that $f'(\lambda) = \frac{1}{\gamma-r}\left(-\frac{1}{r+\lambda} - \frac{1}{\gamma+\lambda} + \frac{r+\lambda}{(\gamma+\lambda)^2}\right) = -\frac{1}{\gamma-r}\left(\frac{1}{r+\lambda} + \frac{\gamma-r}{(\gamma+\lambda)^2}\right) < 0$. Under the condition, $f(\lambda_g) = \frac{1}{\gamma-r}\left(\log \frac{r+\lambda_b}{r+\lambda_g} - \frac{r+\lambda_g}{\gamma+\lambda_g}\right) < 0$. Therefore, for any $\lambda \geq \lambda_g$, we have $f(\lambda) \leq f(\lambda_g) < 0$. □

Proof of Proposition 1. In the securitization stage, the originator chooses when to sell. With full information, the value for the originator in the securitization stage is

$$\max_t \mathbb{E}_a^z \left[\int_0^t ce^{-\gamma u} 1_{\tau_d \geq u} du + e^{-\gamma t} 1_{\tau_d \geq t} p(a; z)\right],$$

where $p(a;z) = \mathbb{E}_a^z\left[\int_t^\infty ce^{-r(u-t)} 1_{\tau_d \geq u} du\right] = \frac{c}{r+\lambda(a)}$ since competitive investors will price the loan at its expected value. Because $\gamma > r$, it is straightforward to show that the solution is a corner one, meaning $T^{FB}(\lambda) = 0$, $\forall \lambda \in [\lambda_g, \lambda_b]$. Since investors value more the cash flows from the loan due to their lower discount rate, selling the loan immediately at time 0 implements allocative efficiency.

Screening effort is chosen to maximize the value for the originator at time 0:

$$V(a;z) = \eta(a) - C(a;z).$$

The first-order condition is thus given by equation (6). By Assumption 1(i), the first-order condition given in (6) characterizes the solution to the problem $\max_a V(a;z)$ both in the first-best and in the full information equilibrium. □

Proof of Proposition 2. Following the proof of Proposition 1 in AGH, we can show that the originator's FOC is

$$c - (\gamma + \lambda) P + \frac{dP}{dt} = 0.$$

Because $P(T(\lambda)) = \frac{c}{r+\lambda}$, we have $\frac{dP}{dt} = (\gamma + \lambda) P - c = (\gamma - r) P$. Because $T(\lambda_b) = 0$ (and thus $P(0) = \frac{c}{r+\lambda_b} = p_b$), we have

$$P(t) = p_b e^{(\gamma - r)t}.$$

Therefore, $c - (r + \lambda) p_b e^{(\gamma - r)t} = 0$, which implies (7). □

Proof of Proposition 3. We prove this proposition for the extended model with hard information s and the probability of selling the loan $\theta(s)$. Note that in the extended model, the originator's value function at time 0 is given by:

$$V(a;s,z) = \theta(s)\left[\frac{c}{\gamma+\lambda(a)}q(a) + (1-q(a))\frac{c}{r+\lambda(a)}\right] + (1-\theta(s))\rho(a) - C(a;z).$$

The first-order condition is:

$$\begin{aligned}
& -\frac{c\lambda'}{(\gamma+\lambda)^2}(\theta(s)q + (1-\theta(s))) - \frac{c}{\gamma+\lambda}\theta(s)(1-q)\left(-\lambda' T - (\gamma+\lambda)T'\lambda'\right) \\
& +\theta(s)(1-q)\frac{c}{r+\lambda}\left(-\lambda' T - (\gamma+\lambda)T'\lambda'\right) - \theta(s)(1-q)\frac{c\lambda'}{(r+\lambda)^2} \\
= \;& C'(a;z),
\end{aligned}$$

where we simplify the notation by dropping the arguments from $\lambda(a)$ and $q(a)$ and use C' (respectively, T') to denote the derivative $\partial C(a;z)/\partial a$ (respectively, $dT(\lambda)/d\lambda$).

Note that

$$\begin{aligned}
T &= \frac{1}{\gamma-r}\log\frac{r+\lambda_b}{r+\lambda}, \\
-\lambda' T - (\gamma+\lambda)T'\lambda' &= -\lambda'\left(T - \frac{\gamma+\lambda}{(\gamma-r)(r+\lambda)}\right) = -\frac{1}{\gamma-r}\lambda'\left(\log\frac{r+\lambda_b}{r+\lambda} - \frac{\gamma+\lambda}{r+\lambda}\right).
\end{aligned}$$

We can write the condition as

$$\begin{aligned}
C'(a;z) &= -\frac{c\lambda'(\theta(s)q+(1-\theta(s)))}{(\gamma+\lambda)^2} - \left(\frac{c}{r+\lambda} - \frac{c}{\gamma+\lambda}\right)\theta(s)(1-q)\frac{1}{\gamma-r}\lambda'\left(\log\frac{r+\lambda_b}{r+\lambda} - \frac{\gamma+\lambda}{r+\lambda}\right) \\
&\quad -\theta(s)(1-q)\frac{c\lambda'}{(r+\lambda)^2} \\
&= -\frac{c\lambda'(\theta(s)q+(1-\theta(s)))}{(\gamma+\lambda)^2} - \frac{c\theta(s)(1-q)}{(r+\lambda)(\gamma+\lambda)}\lambda'\left(\log\frac{r+\lambda_b}{r+\lambda} - \frac{\gamma+\lambda}{r+\lambda}\right) - \theta(s)(1-q)\frac{c\lambda'}{(r+\lambda)^2} \\
&= -\frac{c\lambda'(\theta(s)q+(1-\theta(s)))}{(\gamma+\lambda)^2} - \frac{c\theta(s)(1-q)}{(r+\lambda)(\gamma+\lambda)}\lambda'\log\frac{r+\lambda_b}{r+\lambda} \\
&= -\frac{c\lambda'(\theta(s)q+(1-\theta(s)))}{(\gamma+\lambda)^2} - \frac{c\theta(s)(1-q)(\gamma-r)}{(r+\lambda)(\gamma+\lambda)}T\lambda' \\
&= \rho'(a)\left((\theta(s)q+(1-\theta(s))) + \theta(s)(1-q)(\gamma-r)\frac{\gamma+\lambda}{r+\lambda}T\right),
\end{aligned}$$

where in deriving the last equality we have used $\rho'(a) = -\frac{c\lambda'}{(\gamma+\lambda)^2}$.

Furthermore, the second-order condition is:

$$\begin{aligned}
& \rho''(a)\left(\theta(s)q + 1 - \theta(s) + \theta(s)(1-q)(\gamma-r)\frac{\gamma+\lambda}{r+\lambda}T\right) \\
& +\rho'(a)\theta(s)(1-q)\begin{pmatrix} +(\lambda' T + (\gamma+\lambda)T'\lambda')\left(1-(\gamma-r)\frac{\gamma+\lambda}{r+\lambda}T\right) + \\ +(\gamma-r)\frac{\lambda'}{r+\lambda}T - (\gamma-r)\frac{(\gamma+\lambda)\lambda'}{(r+\lambda)^2}T + (\gamma-r)\frac{\gamma+\lambda}{r+\lambda}T' \end{pmatrix} - C''(a;z) \\
< \;& 0.
\end{aligned}$$

From the first-order condition, we have $1+\theta(s)(1-q)\left((\gamma-r)\frac{\gamma+\lambda}{r+\lambda}T-1\right)=\frac{C'(a;z)}{\rho'(a)}$, and can further simplify the second-order condition as

$$\rho''(a)\frac{C'(a;z)}{\rho'(a)}+\rho'(a)\theta(s)(1-q)\lambda'\left(\begin{array}{c}\left(T-\frac{1}{\gamma-r}\frac{\gamma+\lambda}{r+\lambda}\right)\left(1-(\gamma-r)\frac{\gamma+\lambda}{r+\lambda}T\right)\\+\frac{\gamma-r}{r+\lambda}T-\frac{(\gamma-r)(\gamma+\lambda)}{(r+\lambda)^2}T-\frac{\gamma+\lambda}{(r+\lambda)^2}\end{array}\right)-C''(a;z)<0.$$

After tedious algebra, we can show that

$$\rho''(a)\frac{C'(a;z)}{\rho'(a)}-\rho'(a)\theta(s)(1-q)\lambda'\left(((\gamma-r)T-1)^2+\frac{\gamma-r}{r+\lambda}\right)\frac{\gamma+\lambda}{(\gamma-r)(r+\lambda)}-C''(a;z)<0.$$

Define

$$\Theta(a)\equiv\theta(s)(1-q(a))\left(\left(\log\frac{r+\lambda_b}{r+\lambda(a)}-1\right)^2+\frac{\gamma-r}{r+\lambda(a)}\right)\frac{\gamma+\lambda(a)}{(\gamma-r)(r+\lambda(a))},\tag{24}$$

then the previous inequality can be written as

$$\left(\frac{C'(a;z)}{\rho'(a)}\right)'>-\theta(s)(1-q)\lambda'\left(((\gamma-r)T-1)^2+\frac{\gamma-r}{r+\lambda}\right)\frac{\gamma+\lambda}{(\gamma-r)(r+\lambda)}\equiv-\Theta\lambda',$$

which is the condition in Assumption 1 (ii). □

Proof of Proposition 4. First, note that

$$\begin{aligned}\frac{\partial\lambda(a)}{\partial a}&<0,\\ \rho'(a)&=-\frac{c\lambda'(a)}{(\gamma+\lambda(a))^2}>0,\\ \frac{\partial\left(\frac{C'(a;z)}{\rho'(a)}\right)}{\partial z}&=\frac{1}{(\rho'(a))^2}\frac{\partial C'(a;z)}{\partial z}>0.\end{aligned}$$

Therefore, from Assumption 1(ii), for the second-order condition to hold, it must be true that

$$\left(\frac{C'(a;z)}{\rho'(a)}\right)'>-\Theta\lambda'(a)>0.$$

Second, taking a total derivative of both sides of the FOC with respect to z, we have

$$\begin{aligned}&\frac{\partial\frac{C'(a;z)}{\rho'(a)}}{\partial a}\frac{da^*}{dz}+\frac{\partial\frac{C'(a;z)}{\rho'(a)}}{\partial z}\\
=&\frac{1}{\gamma-r}e^{-(\gamma+\lambda)T(\lambda)}\left(\frac{\gamma+\lambda}{r+\lambda}-\log\frac{r+\lambda_b}{r+\lambda}\right)\left(\frac{\gamma+\lambda}{r+\lambda}\log\frac{r+\lambda_b}{r+\lambda}-1\right)\frac{d\lambda}{dz}\\
&-e^{-(\gamma+\lambda)T(\lambda)}\left(\frac{\gamma-r}{(r+\lambda)^2}\log\frac{r+\lambda_b}{r+\lambda}+\frac{\gamma+\lambda}{(r+\lambda)^2}\right)\frac{d\lambda}{dz}\\
=&\left[\begin{array}{c}\frac{1}{\gamma-r}\left(\frac{\gamma+\lambda}{r+\lambda}-\log\frac{r+\lambda_b}{r+\lambda}\right)\left(\frac{\gamma+\lambda}{r+\lambda}\log\frac{r+\lambda_b}{r+\lambda}-1\right)\\-\frac{\gamma-r}{(r+\lambda)^2}\log\frac{r+\lambda_b}{r+\lambda}-\frac{\gamma+\lambda}{(r+\lambda)^2}\end{array}\right]e^{-(\gamma+\lambda)T(\lambda)}\frac{d\lambda}{dz}\\
=&-\left[\left(\log\frac{r+\lambda_b}{r+\lambda}-1\right)^2+\frac{\gamma-r}{r+\lambda}\right]\frac{1}{\gamma-r}\frac{\gamma+\lambda}{r+\lambda}e^{-(\gamma+\lambda)T(\lambda)}\frac{d\lambda}{dz}\\
=&-\Theta\frac{d\lambda}{dz},\end{aligned}$$

where Θ is defined in (24).

Third, note that $\frac{d\lambda}{dz} = \lambda' \frac{da}{dz}$, we have

$$\frac{\partial \frac{C'(a;z)}{\rho'(a)}}{\partial a} \frac{da^*}{dz} + \frac{\partial \frac{C'(a;z)}{\rho'(a)}}{\partial z} + \Theta \left(\lambda' \frac{da}{dz} \right) = 0,$$

or

$$\left(\frac{\partial \frac{C'(a;z)}{\rho'(a)}}{\partial a} + \Theta \lambda' \right) \frac{da^*}{dz} = - \left(\underbrace{\frac{\partial \frac{C'(a;z)}{\rho'(a)}}{\partial z}}_{>0} \right) < 0.$$

From the second-order condition in Assumption 1(ii), we know that the coefficient of $\frac{da^*}{dz}$ is strictly positive, implying that $\frac{da^*}{dz} < 0$ in (10).

Fourth, $\frac{d\lambda(a^*(z))}{dz} = \lambda' \frac{da^*}{dz} > 0$ in (12).

Lastly, it is straightforward to prove $\frac{dT(\lambda(a^*(z)))}{dz} < 0$ in (11), because from (7), we have

$$\frac{dT\left(\lambda\left(a^*\left(z\right)\right)\right)}{dz} = -\frac{1}{\gamma - r} \frac{1}{r + \lambda} \frac{d\lambda\left(a^*\right)}{dz}.$$

$\square$

Proof of Proposition 5. Screening effort is chosen to maximize the value for the originator at time 0:

$$V\left(a; s, z\right) = \theta(s)\eta(a) + (1 - \theta(s))\rho(a) - C\left(a\right).$$

The first-order condition is given by

$$\theta\left(s\right)\eta'\left(a^{SB}\right) + \left(1 - \theta\left(s\right)\right)\rho'\left(a^{SB}\right) - C'\left(a^{SB}; z\right) = 0, \tag{25}$$

which can be rewritten as equation (13), because $\eta'\left(a\right) = \left(\frac{\gamma + \lambda(a)}{r + \lambda(a)}\right)^2 \rho'\left(a\right)$.

Define

$$H\left(a; s, z\right) \equiv \rho'\left(a\right)\left(1 + \theta\left(s\right)\left(\left(\frac{\gamma + \lambda\left(a\right)}{r + \lambda\left(a\right)}\right)^2 - 1\right)\right) - C'\left(a; z\right).$$

Note that at the first-best allocation, $H\left(a^{SB}; s, z\right) = 0$ by Proposition 1.

$$\begin{aligned}
H_a\left(a; s, z\right) &= \rho''\left(a\right)\left(1 + \theta\left(s\right)\left(\left(\frac{\gamma + \lambda\left(a\right)}{r + \lambda\left(a\right)}\right)^2 - 1\right)\right) \\
&\quad -2\rho'\left(a\right)\theta\left(s\right)\frac{\gamma + \lambda\left(a\right)}{r + \lambda\left(a\right)}\frac{\left(\gamma - r\right)\lambda'\left(a\right)}{\left(r + \lambda\left(a\right)\right)^2} - C''\left(a; z\right) \\
&= \rho''\left(a\right)\frac{C'\left(a; z\right)}{\rho'\left(a\right)} - \rho'\left(a\right)\Theta^{SB}\left(a; s, z\right)\lambda'\left(a\right) - C''\left(a; z\right) \\
&= -\rho'\left(a\right)\left(\left(\frac{C'\left(a; z\right)}{\rho'\left(a\right)}\right)' + \Theta^{SB}\left(a; s, z\right)\lambda'\left(a\right)\right) \\
&< 0,
\end{aligned}$$

where the last inequality follows Assumption 1(ii).

On the other hand, we have

$$H_s(a;s,z) = \rho'(a)\left(\left(\frac{\gamma+\lambda(a)}{r+\lambda(a)}\right)^2 - 1\right)\theta'(s) \geq 0,$$

because $\left(\frac{\gamma+\lambda(a)}{r+\lambda(a;z)}\right)^2 > 1$ and $\theta'(s) \geq 0$.

In addition, under the conditions in (3), we have

$$\frac{\partial\left(\frac{C'(a;z)}{\rho'(a)}\right)}{\partial z} = \frac{1}{\rho'(a)}\frac{\partial C'(a;z)}{\partial z} > 0.$$

Therefore, at the second-best effort level, we have

$$H_z\left(a^{SB};s,z\right) = -\frac{\partial^2 C\left(a^{SB};z\right)}{\partial a \partial z} < 0.$$

Lastly, from $H\left(a^{SB};s,z\right) = 0$, we thus have

$$\begin{aligned}
\frac{da^{SB}(s,z)}{ds} &= -\frac{H_s\left(a^{SB};s,z\right)}{H_a\left(a^{SB};s,z\right)} \geq 0, \\
\frac{da^{SB}(s,z)}{dz} &= -\frac{H_z\left(a^{SB};s,z\right)}{H_a\left(a^{SB};s,z\right)} < 0.
\end{aligned}$$

□

Proof of Proposition 7. To prove that $a_-^* > a_+^*$, we just need to prove that $V'\left(a_-^*;s_+^*,z\right) < 0$ for a given z. First, note that a_-^* is the optimal effort when $s = s_-^*$, implying $V'\left(a_-^*;s_-^*,z\right) = 0$, or

$$\rho'\left(a_-^*\right)\left(1+\theta\left(s_-^*\right)\left(1-q\left(a_-^*\right)\right)\left(\frac{(\gamma-r)\left(\gamma+\lambda\left(a_-^*\right)\right)T\left(\lambda\left(a_-^*\right)\right)}{r+\lambda\left(a_-^*;z\right)} - 1\right)\right) = C'\left(a_-^*;z\right).$$

On the other hand, we have

$$V'\left(a_-^*;s_+^*,z\right) = \rho'\left(a_-^*\right)\left(1+\theta\left(s_+^*\right)\left(1-q\left(a_-^*\right)\right)\left(\frac{(\gamma-r)\left(\gamma+\lambda\left(a_-^*\right)\right)T\left(\lambda\left(a_-^*\right)\right)}{r+\lambda\left(a_-^*\right)} - 1\right)\right) - C'\left(a_-^*;z\right).$$

The two equations above imply

$$\begin{aligned}
V'\left(a_-^*;s_+^*,z\right) &= \rho'\left(a_-^*\right)\left(1+\theta\left(s_+^*\right)\left(1-q\left(a_-^*\right)\right)\left(\frac{(\gamma-r)\left(\gamma+\lambda\left(a_-^*\right)\right)T\left(\lambda\left(a_-^*\right)\right)}{r+\lambda\left(a_-^*\right)} - 1\right)\right) \\
&\quad -\rho'\left(a_-^*\right)\left(1+\theta\left(s_-^*\right)\left(1-q\left(a_-^*\right)\right)\left(\frac{(\gamma-r)\left(\gamma+\lambda\left(a_-^*\right)\right)T\left(\lambda\left(a_-^*\right)\right)}{r+\lambda\left(a_-^*\right)} - 1\right)\right) \\
&= \rho'\left(a_-^*\right)\left(\theta\left(s_+^*\right)-\theta\left(s_-^*\right)\right)\left(1-q\left(a_-^*\right)\right)\left(\frac{(\gamma-r)\left(\gamma+\lambda\left(a_-^*\right)\right)T\left(\lambda\left(a_-^*\right)\right)}{r+\lambda\left(a_-^*\right)} - 1\right) \\
&< 0,
\end{aligned}$$

where the last inequality follows from the condition $\log\frac{r+\lambda_b}{r+\lambda_g} < \frac{r+\lambda_g}{\gamma+\lambda_g}$ (Lemma 1).

The proof of equations (19) and (20) follows immediately from the monotonicity of $\lambda(a)$ and $T(\lambda(a))$.

□

Proof of Proposition 6. Define function $H(\cdot;s)$ by

$$H\left(a;s\right)\equiv\rho'\left(a\right)\left(1+\theta\left(s\right)\left(1-q\left(a\right)\right)\left(\frac{\left(\gamma-r\right)\left(\gamma+\lambda\left(a\right)\right)T\left(\lambda\left(a\right)\right)}{r+\lambda\left(a\right)}-1\right)\right)-C'\left(a;z\right).$$

Note that at the equilibrium allocation with positive effort a^*, $H(a^*;s)=0$ by Proposition 3. From Assumption 2(iii), we have $H_a a(a^*;s)<0$, because

$$H_a(a^*;s)=-\rho'\left(a^*\right)\left(\left(\frac{C'\left(a^*;z\right)}{\rho'\left(a^*\right)}\right)'+\lambda'\left(a^*\right)\Theta\left(a^*\right)\right)<0.$$

On the other hand, we have $H_s(a^*;s)<0$, because

$$H_s(a^*;s)=\rho'\left(a^*\right)\left(1-q\left(a^*\right)\right)\left(\frac{\left(\gamma-r\right)\left(\gamma+\lambda\left(a^*\right)\right)T\left(\lambda\left(a^*\right)\right)}{r+\lambda\left(a^*\right)}-1\right)\theta'\left(s\right)<0,$$

where the last inequality follows from the results that $\theta'\left(s\right)>0$ and $\left(\frac{(\gamma-r)(\gamma+\lambda(a^*))T(\lambda(a^*))}{r+\lambda(a^*)}-1\right)<0$ (Lemma 1). Therefore, from $H\left(a^*;s\right)=0$, we have $\frac{da^*}{ds}=-\frac{H_s}{H_a}<0$. $\square$

Appendix B: Data Appendix

Appendix B1: HMDA-LoanPerformance Merge

The merging algorithm in our paper parallels the one used in Rosen (2011) that matches the confidential HMDA database with the McDash database from Black Knight Financial Services. The most important variables used to merge these two databases include the geographic location (i.e., ZIP code) and certain loan characteristics such as the amount and closing date of the loan. Specifically, to match HMDA mortgage observations to CoreLogic LoanPerformance mortgage observations, we impose the following matching criteria. The mortgage observations in both databases are considered "matched", if (1) they have the same ZIP code;[7] (2) they have the same lien type (first or second), occupancy type (owner-occupied), purpose (home-purchase), and mortgage type (conventional); (3) their origination amounts should not differ more than \$500; (4) they have similar if not identical origination dates. Because neither database reports the closing date precisely, we use the following procedure sequentially: an exact-day match, followed by an iterative five-day difference match, and then followed by a same-month match. Our merging algorithm has a similar matching rate as in Rosen (2011) in which 50% to 80% of McDash mortgage observations can be matched with the HMDA database.

Appendix B2: Key Variables

Tables B1 and B2 report key variables from the CoreLogic LoanPerformance and the confidential HMDA databases, respectively. Table B3 reports macro variables related to macroeconomic conditions.

[7] Because the HMDA reports mortgages by census tracts, we map census tracts to ZIP codes based on the U.S. Census Bureau's approximations of ZIP codes (i.e., ZCTA5 values), available at https://mcdc2.missouri.edu/websas/geocorr2k.html.

The Home Mortgage Disclosure Act was passed into law by Congress in 1975 and expanded in 1988, to inform the public (and the regulators) about whether or not financial institutions adequately serve local credit needs. In addition, regulators use the HMDA data to help identify discriminatory lending. These data are collected by the Federal Reserve under Regulation C, and all regulated financial institutions (e.g., commercial banks, savings institutions, credit unions, and mortgage companies) with assets above $30 million must report.

The HMDA data include information on the year of the application, the identity of the lender, the dollar amount of the loan, whether or not the loan was accepted, and whether or not the lender retained the loan or sold it to a third party. In addition, the HMDA data contain information on the location of the property, as well as some information on borrower credit risk such as income and loan size. However, the HMDA data contain no information on the property value or the borrower's credit score. The detailed HMDA reporting guide is published by the Federal Financial Institutions Examination Council (FFIEC).

Table B1: Variables from the CoreLogic LoanPerformance Database

Variable List	Definition
ARM	Indicator variable for whether the loan has an adjustable rate or not
Closing Date	Loan closing date
Delinquency	Indicator variable for whether the loan is in default within fifteen months of origination: (a) payments on the loan are 60+ days late; (b) the loan is in foreclosure; or (c) the loan is real estate owned (REO)
Document Type	Loan documentation level
DTI	Back-end debt-to-income ratio
FICO	FICO score at origination
Initial Rate	Initial or original interest rate as of the loan's first payment date
Lien Type	Lien position (e.g., first lien)
Loan Amount	Loan origination amount
Loan Purpose	Purpose of the loan (e.g., purchase, refi, etc.)
Loan Type	Type of the loan (e.g., conventional)
LTV	Combined loan-to-value (CLTV) ratio (including first and second liens)
Margin	Margin (in percent) for an adjustable-rate or hybrid loan over an interest rate index, applicable after the first interest rate reset.
Balloon	Indicator variable equal to 1 for a fixed rate or adjustable rate loan where the payments are lower over the life of the loan, leaving a balloon payment at maturity.
Hybrid2	Indicator variable equal to 1 for an adjustable rate loan with the initial monthly payment fixed for the first two years. This is typically referred to as a 2/28 hybrid ARM, with the interest rate over the remaining 28 years of the loan equal to the value of an interest rate index (i.e., 6-month LIBOR) measured at the time of adjustment, plus a margin fixed for the life of the loan. The initial fixed rate is called a "teaser" interest rate because it is lower than what a borrower would pay for a 30-year fixed rate mortgage
Hybrid3	Indicator variable equal to 1 for a 3/27 hybrid ARM (i.e., the initial interest rate is fixed for 3 years)
IOflag	Indicator variable equal to 1 if the loan has an interest-only feature.
MissingDTI	Indicator variable equal to 1 if DTI is missing
Occupancy	Indicator variable for whether owner-occupied or not
Prepay Penalty	Indicator variable equal to 1 when the loan has a prepayment penalty and/or is an option ARM or negative amortization loan. These loan features make refinancing less likely in default.
Property Type	Type of the property (i.e., single-family residence)
Sale Price	Property sale price
TS	The period between loan origination and MBS closing
ZIP Code	ZIP code of the property

Table B2: Variables from the Confidential HMDA Database

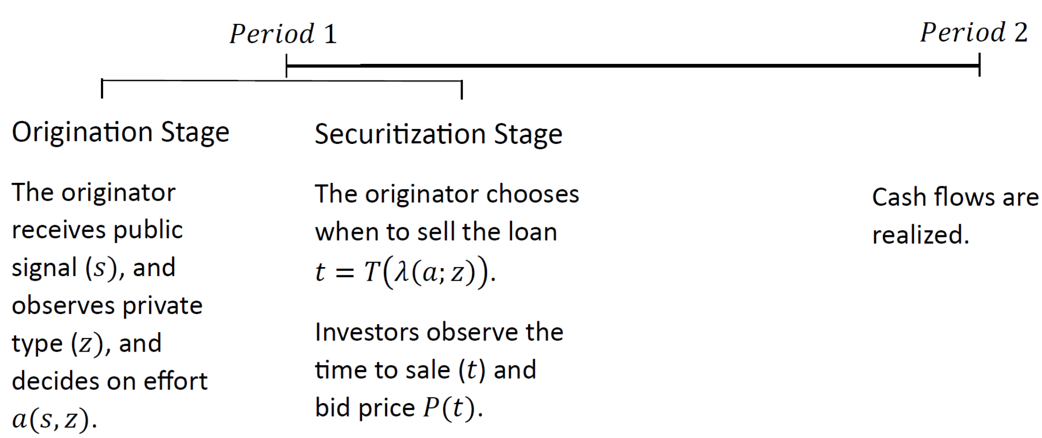
Variable List	Definition
Action Date	Date of action was taken on application
Applicant Race	Indicator variable for the race of the loan applicant (e.g., White)
Applicant Sex	Indicator variable to classify male or female
Applicant Income	Total gross annual income of applicant in thousands of dollars (nominal)
Application Date	Date of loan application
Co-applicant	Indicator variable for whether the loan includes co-applicant or not
County Code	Identify loan originated county
HMDA-ID	Unique record to identify each loan in HMDA
Jumbo loan	Indicator variable equal to one if the loan amount exceeds FHFA conforming loan limit for the month of origination
Lien Status	Indicator variable to classify loan is secured by a first lien, or a subordinate lien, or not secured by a lien
Ln(Income)	Natural log of applicant income
Ln(Loan Size)	Natural log of loan amount
Loan Amount	Loan amount granted or requested in thousands of dollars
Loan Purpose	Indicator variable for whether the loan or application was for a home purchase loan, a home improvement loan, or a refinancing loan
Loan Type	Indicator variable for whether the loan was conventional, government-guaranteed, or government-insured
Loan-to-Income	Loan amount divided by applicant income
Occupancy	Indicator variable for whether owner-occupied or not
Processing Time	Action date minus application date
Property Type	Indicator variable for whether the loan was for a manufactured home, a multifamily dwelling, or a 1- to 4-family dwelling
Purchaser Type	Indicator variable for whether the loan was subsequently sold to a secondary market entity within the same calendar year
State Code	Identify loan originated state

We supplement these databases with additional data on macroeconomic conditions. Specifically, we collect macro variables such as local housing price appreciation, state-level unemployment rate, and local median household income in order to control for the overall economic environment. For each loan in the sample, we identify the borrower's geographic area using the five-digit ZIP code.

Table B3: Local Macro Variables

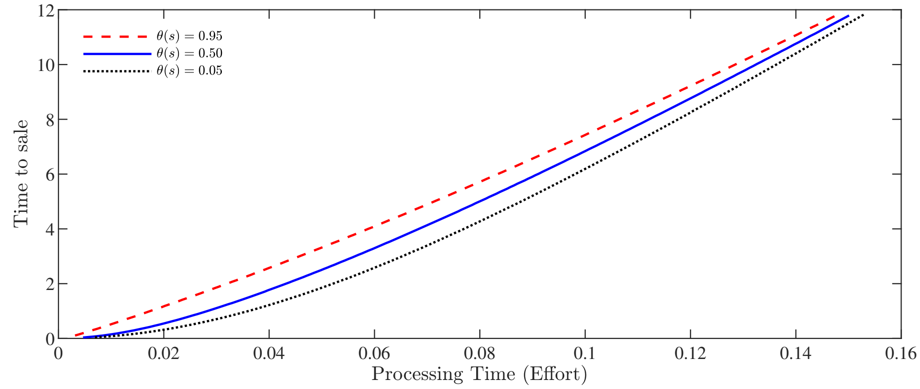
Variable List	Definition
Avg. Wage	The average wage in the borrower's county in the year of the loan origination
HPA	The 36-month change in the housing price index for the borrower's county prior to loan origination
Ln(Income)	The median household income in 1999 for the borrower's ZIP code as reported by the U.S. Census Bureau in 2000, in logarithm
Loan Number	The number of loans originated in the borrower's county in the origination year
Unemp. Rate	The unemployment rate in the borrower's state in the year of loan origination

Figure 1: Time Line of the Model



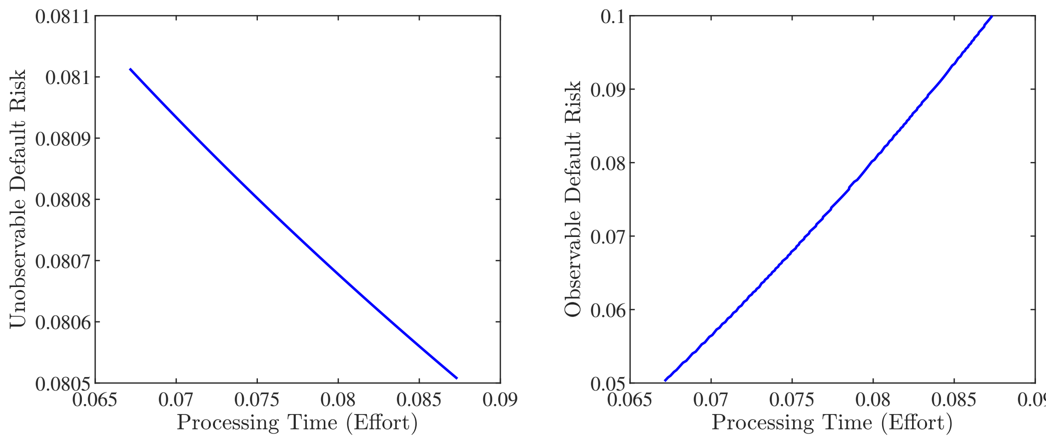
Note: This figure shows a time line summarizing the events in the model.

Figure 2: Equilibrium Relation between Time to Sale and Processing Time



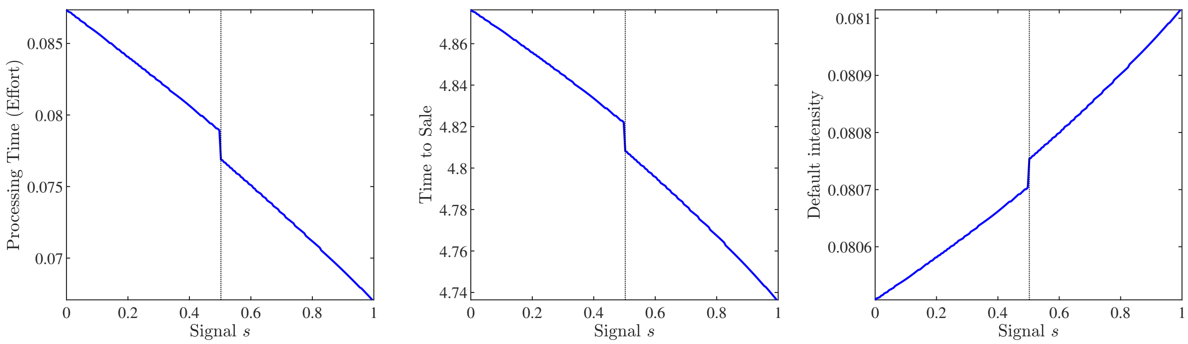
Note: This figure plots the equilibrium relationship between time to sale and processing time when $\theta(s) = 0.5$ (solid blue line), 0.95 (dashed red line), and 0.05 (dotted black line). We consider $C(a) = \frac{1}{2}ka^2$ and $\lambda(a;z) = \lambda_b + a^{\zeta}(z - \lambda_b)$ with the following parameter specification: $c = 1$, $\gamma = 0.05$, $r = 0.01$, $\lambda_b = 0.1$, $\lambda_g = 0.05$, $k = 15$, and $\zeta = 0.1$.

FIGURE 3: Equilibrium Relation between (Un)observable Default Risk and Processing Time



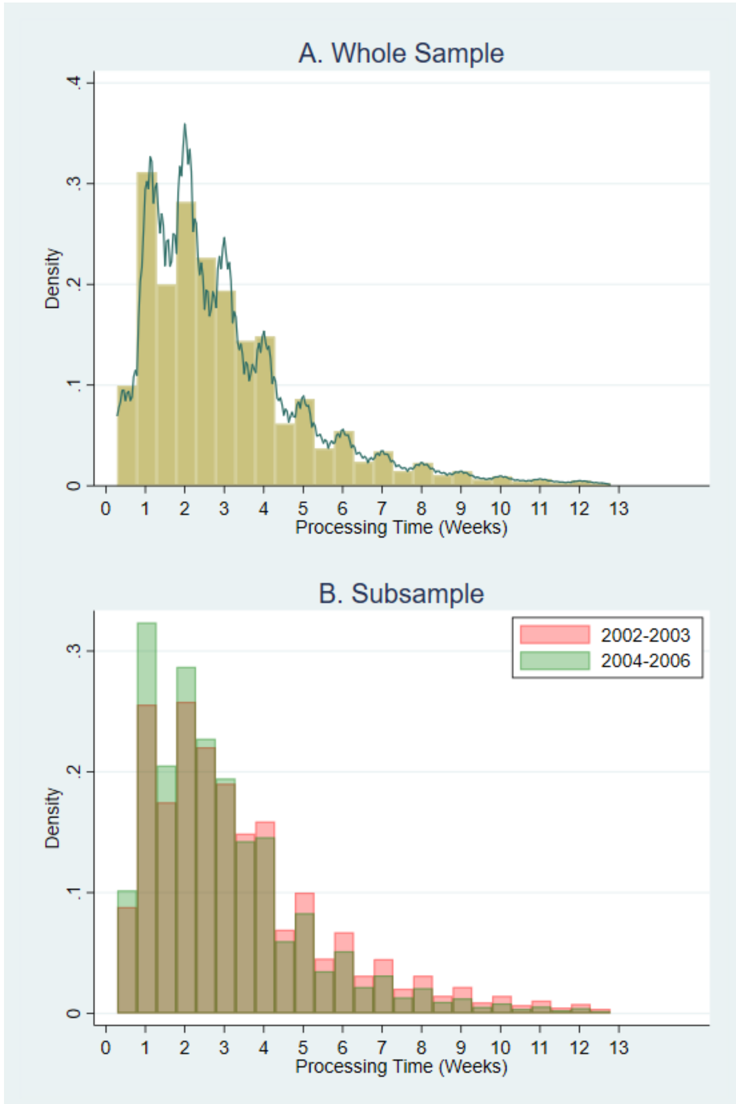
NOTE: This figure plots the equilibrium relationship between unobservable default risk and processing time in the left panel, and the relationship between observable default risk and processing time in the right panel. The unobservable default risk is defined by $\lambda(a;z) = \lambda_b + a^{\zeta}(z - \lambda_b)$, while the observable default risk is parameterized as follows: $\widehat{\lambda}(s) = \lambda_b + s(\lambda_g - \lambda_b)$. be positively related to signal s We consider $C(a) = \frac{1}{2}ka^2$ and $\lambda(a;z) = \lambda_b + a^{\zeta}(z - \lambda_b)$ with the following parameter specification: $c = 1$, $\gamma = 0.05$, $r = 0.01$, $\lambda_b = 0.1$, $\lambda_g = 0.05$, $k = 15$, and $\zeta = 0.1$.

Figure 4: Discontinuities in Processing Time, Time to Sale, and Default Intensity



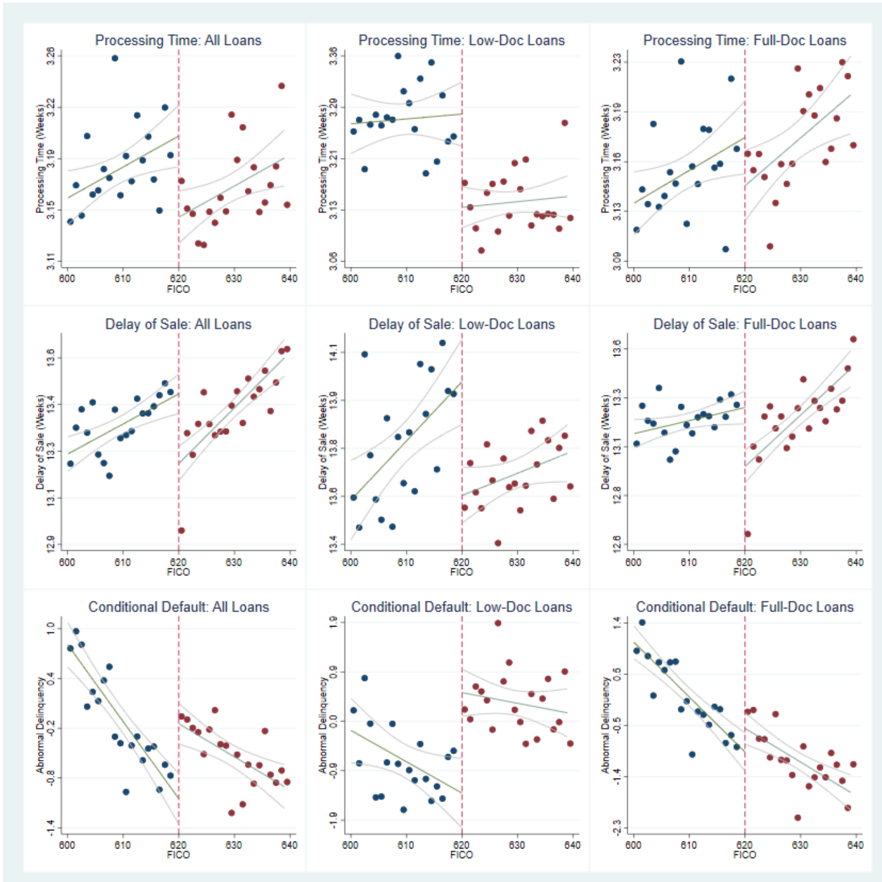
Note: This figure plots discontinuities in processing time (left panel), time to sale (middle panel), and default intensity (right panel) as a result of discontinuity in the ease of securitization. Specifically, we assume that around the threshold $s' = 0.5$, the securitization probability jumps from 0.45 to 0.55. We consider $C(a) = \frac{1}{2}ka^2$ and $\lambda(a; z) = \lambda_b + a^{\zeta}(z - \lambda_b)$ with the following parameter specification: $c = 1$, $\gamma = 0.05$, $r = 0.01$, $\lambda_b = 0.1$, $\lambda_g = 0.05$, $k = 15$, and $\zeta = 0.1$.

Figure 5: Histogram of Processing Time between 2002 and 2006



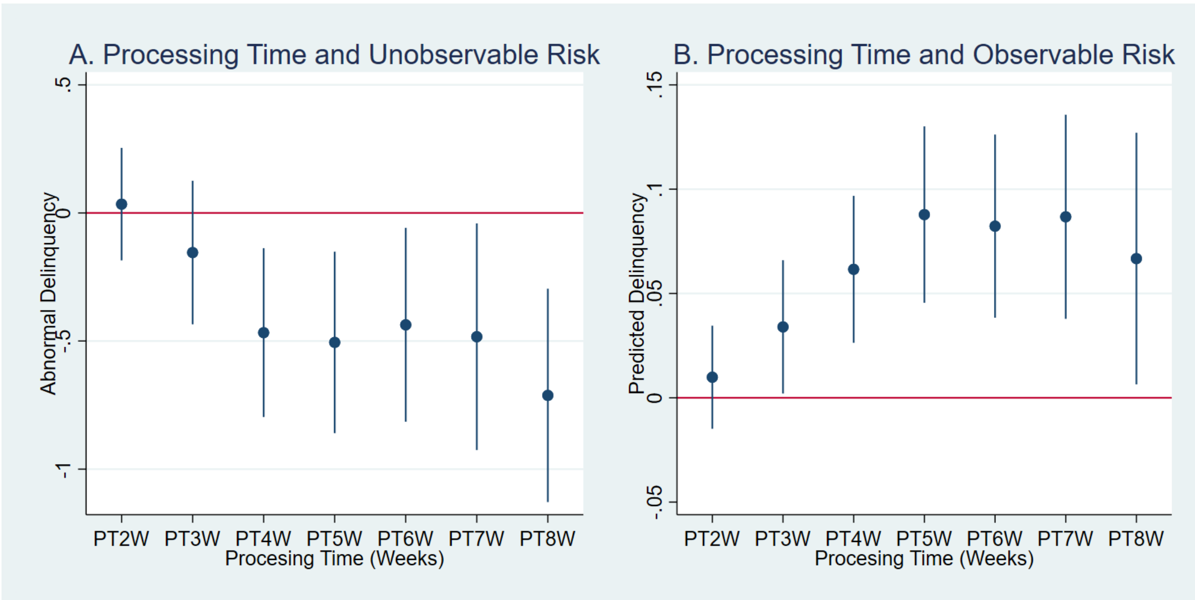
Note: This figure shows the histogram of mortgage processing time for the whole sample, 2002-2006, in Panel A and subsamples, 2002-2003 and 2004-2006, in Panel B. We add in Panel A the scaled kernel density estimate of the density estimated with the Epanechnikov kernel and asymptotically optimal bandwidth. The sample is the merged confidential HMDA and CoreLogic ABS database.

Figure 6: RDD Regression of Processing Time, Delay of Sale, and Delinquency



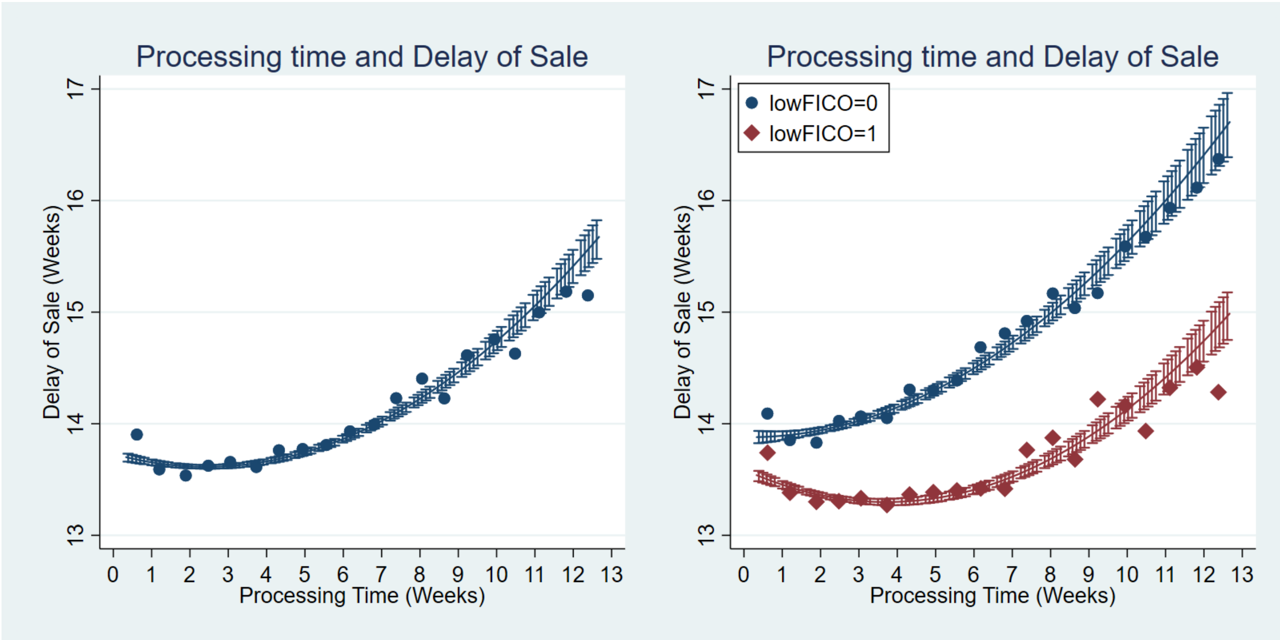
Note: This figure shows the regression discontinuity plot of processing time (the first row), delay of sale (the second row), and delinquency (the third row) for the merged confidential HMDA and CoreLogic ABS database. We first regress the processing time, delay of sale, or delinquency on borrower and loan characteristics, year, and state fixed effects. We then plot the average residuals from the regression of processing time or delinquency for one-point FICO bins between scores of 600 and 640, with a linear fit to the data on either side of the 620 cutoff and the 95% confidence interval. We add the sample average processing time to the residuals from the processing time regression. Y-axis scale in Panel C is in percentage points, so that "0.1" represents 0.1 percentage point in abnormal delinquency (or conditional default rate).

FIGURE 7: Processing time and (un)observable mortgage default risk



NOTE: This figure shows the coefficient estimates of processing time, grouped into dummy variables from one week to 8+ weeks, in the regression of loan delinquency (Panel A) or predicted default (Panel B), borrower and loan characteristics, year, state, and lender fixed effects. The predicted default probability is estimated with a logistic model using borrower and loan characteristics with a 2-year rolling window. The point estimates and the 95% confidence intervals are plotted. Y-axis scale is in percentage points, so that "0.1" represents 0.1 percentage point difference in default rate relative to loans processed in one week or less. The sample is the merged confidential HMDA and CoreLogic ABS database.

FIGURE 8: Processing Time and Delay of Sale



NOTE: This figure shows the scatter plot of processing time and delay of sale for the whole sample (the left panel) and subsamples grouped by FICO scores (the right panel). We average the delay of sale for each processing time bin. We plot a quadratic fit along with the 95% confidence interval. The LowFICO variable is an indicator variable for FICO scores below 640. The sample is the merged confidential HMDA and CoreLogic ABS database.

Table 1: Summary Statistics

Year		PT (Weeks)	TS (Weeks)	FICO	CLTV	Delinq. (%)	N
2002	All Loans	3.6	17.9	622	86.3	8.3	95639
	Low-Doc	3.7	18.2	648	85.3	6.9	32046
	Full-Doc	3.6	17.8	610	86.8	9.0	63593
2003	All Loans	3.7	14.4	636	90.4	6.8	179133
	Low-Doc	3.7	14.3	656	89.1	5.9	66921
	Full-Doc	3.6	14.5	623	91.2	7.3	112212
2004	All Loans	3.3	13.0	636	92.6	7.8	358025
	Low-Doc	3.3	12.6	656	91.6	7.1	139460
	Full-Doc	3.3	13.2	623	93.3	8.3	218565
2005	All Loans	3.1	14.2	636	93.6	11.0	478141
	Low-Doc	3.1	13.9	658	93.1	10.8	194257
	Full-Doc	3.2	14.4	621	94.0	11.1	283884
2006	All Loans	3.0	13.2	630	95.0	20.4	394034
	Low-Doc	2.9	13.7	653	94.3	23.5	159346
	Full-Doc	3.1	12.8	615	95.4	18.4	234688
All years	All Loans	3.2	13.7	633	92.9	12.0	1504972
	Low-Doc	3.2	13.7	655	92.2	12.6	592030
	Full-Doc	3.2	13.8	619	93.3	11.7	912942

Note: This table reports summary statistics based on the merged confidential HMDA and CoreLogic ABS database.

Table 2: Loan-level Regression of Discontinuity

	Whole sample		Low Doc		Full Doc	
Panel A: Processing Time						
	(1)	(2)	(3)	(4)	(5)	(6)
$\mathbf{1}[FICO \geq 620]$	-0.43^{***}	-0.16	-0.93^{***}	-0.59^{***}	-0.20	-0.03
	(-3.23)	(-1.48)	(-3.64)	(-3.14)	(-1.40)	(-0.28)
Adjusted R^2	0.051	0.184	0.064	0.184	0.048	0.188
N	423129	422378	128733	128211	294396	293733
Panel B: Delinquency						
	(7)	(8)	(9)	(10)	(11)	(12)
$\mathbf{1}[FICO \geq 620]$	0.70^{***}	0.63^{***}	1.20^{**}	1.27^{***}	0.40	0.31
	(2.70)	(2.60)	(2.47)	(2.79)	(1.41)	(1.22)
Adjusted R^2	0.052	0.057	0.077	0.081	0.038	0.043
N	423129	422378	128733	128211	294396	293733
State, Year FEs	Yes	Yes	Yes	Yes	Yes	Yes
Lender FE	No	Yes	No	Yes	No	Yes
Loan&Borr. Char	Yes	Yes	Yes	Yes	Yes	Yes

Note: This table reports the results of the loan-level regression of discontinuity based on the merged ABS and HMDA data for low documentation loans with FICO between 600 and 640. $\mathbf{1}[FICO \geq 620]$ is an indicator that takes a value of 1 at $FICO \geq 620$ and a value of zero if $FICO < 620$. Standard errors are clustered by lender and year, and t-statistics are reported in parentheses. Significance level: *($p < .10$); **($p < .05$); and ***($p < .01$).

Table 3: Processing Time and Time to Sale

	Whole Sample	Subsample						
		Originator-Sponsor		Documentation		PT (weeks)		
		Affiliated	Unaffiliated	Low-Doc	Full-Doc	$(0,4]$	$[4,8)$	$[8,\infty)$
PT	0.04***	0.01	0.04***	0.02*	0.05***	0.00	0.04**	0.04*
	(3.42)	(1.22)	(2.79)	(1.89)	(3.26)	(0.16)	(2.05)	(1.90)
Adjusted R^2	0.213	0.474	0.240	0.243	0.205	0.211	0.218	0.256
N	1335209	420588	914017	512792	821785	1019613	250689	63752
State, Year FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Lender FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Loan&Borr. Char	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Note: This table reports the results of a loan-level regression of time to sale on loan processing time based on the merged ABS and HMDA data. The dependent variable is standardized time to sale. The control variables are defined in Appendix B. Standard errors are clustered by lender and year, and t-statistics are reported in parentheses. Significance level: *($p < .10$); **($p < .05$); and ***($p < .01$).

Table 4: Processing Time, Time to Sale, and Loan Default

	(1)	(2)	(3)	(4)	(5)
PT	-0.09^{***}		-0.10^{***}		
	(-4.30)		(-4.34)		
TS		-0.04^{***}	-0.03^{***}		
		(-3.16)	(-3.13)		
LowPT				0.42^{***}	0.29^{***}
				(4.22)	(2.70)
LowTS				0.41^{*}	0.27
				(1.74)	(1.15)
LowPT*LowTS					0.26^{**}
					(2.33)
Adjusted R^2	0.072	0.071	0.071	0.072	0.072
Observations	1437070	1335209	1335209	1437070	1437070
State, Year FEs	Yes	Yes	Yes	Yes	Yes
Lender FE	Yes	Yes	Yes	Yes	Yes
Loan&Borr. Char	Yes	Yes	Yes	Yes	Yes

Note: This table reports the results of a loan-level regression of loan default on time to sale and processing time based on the merged ABS and HMDA data. The dependent variable is loan delinquency within 10 to 15 months of loan origination. The control variables are defined in Appendix B. Standard errors are clustered by lender and year, and t-statistics are reported in parentheses. Significance level: *($p < .10$); **($p < .05$); and ***($p < .01$).

Table 5: Processing Time, Time to Sale, and Loan Default: Subsample Analysis

Panel A: Originator-Sponsor Affiliation						
	Affiliated			Unaffiliated		
	(1)	(2)	(3)	(4)	(5)	(6)
PT	-0.08^{***}		-0.11^{***}	-0.10^{***}		-0.10^{***}
	(-3.77)		(-4.17)	(-3.49)		(-3.40)
TS		-0.07^{***}	-0.06^{***}		-0.03^{**}	-0.03^{**}
		(-6.28)	(-6.26)		(-2.25)	(-2.22)
Adjusted R^2	0.074	0.075	0.075	0.071	0.071	0.071
Observations	521651	420588	420588	914752	914017	914017
Panel B: Loan Documentation Level						
	Low Doc Loans			Full Doc Loans		
	(7)	(8)	(9)	(10)	(11)	(12)
PT	-0.11^{***}		-0.12^{***}	-0.07^{***}		-0.08^{***}
	(-3.88)		(-3.90)	(-3.37)		(-3.37)
TS		-0.04^{**}	-0.04^{**}		-0.04^{***}	-0.04^{***}
		(-2.13)	(-2.12)		(-3.85)	(-3.82)
Adjusted R^2	0.090	0.090	0.090	0.064	0.064	0.064
Observations	550672	512792	512792	885722	821785	821785
State, Year FEs	Yes	Yes	Yes	Yes	Yes	Yes
Lender FE	Yes	Yes	Yes	Yes	Yes	Yes
Loan&Borr. Char	Yes	Yes	Yes	Yes	Yes	Yes

Note: This table reports the results of a loan-level regression of loan default on time to sale and processing time based on the merged ABS and HMDA data for subsample of loans by originator-sponsor affiliation and loan documentation. The dependent variable is loan delinquency within 10 to 15 months of loan origination. The control variables are defined in Appendix B. Standard errors are clustered by lender and year, and t-statistics are reported in parentheses. Significance level: *($p < .10$); **($p < .05$); and ***($p < .01$).

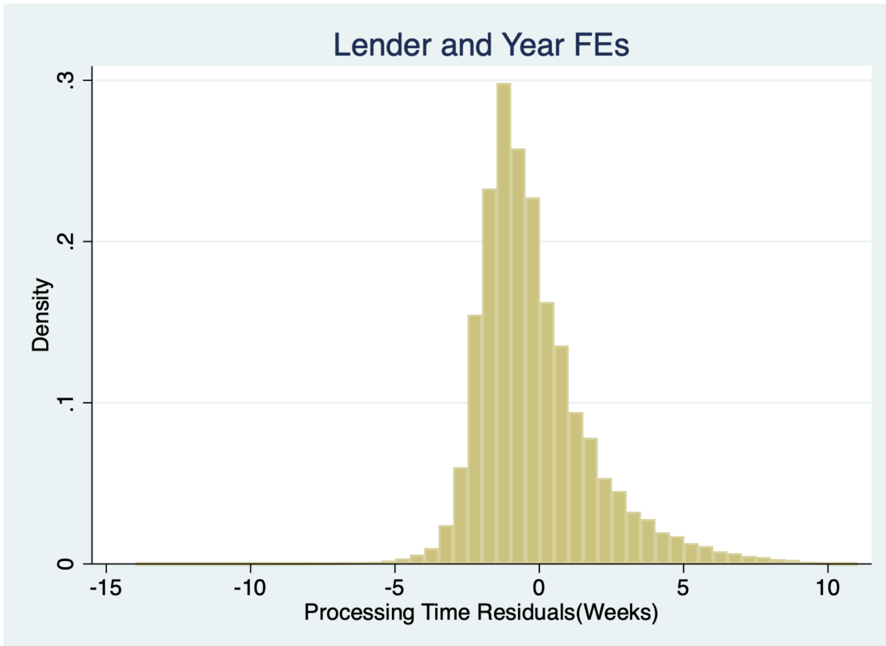
INTERNET APPENDIX
Screen More, Sell Later: Screening and Dynamic Signaling in the Mortgage Market

Manuel Adelino, Bin Wei, Feng Zhao

A Additional Figures

This section provides the additional figures referenced in the main text. Figure A.1 shows the histogram of the residuals from the regression of mortgage processing time on lender and origination year fixed effects for the whole sample, 2002-2006.

Figure A.1: Histogram of Processing Time Residuals Between 2002 and 2006



Note: This figure shows the histogram of the residuals from the regression of mortgage processing time on lender and origination year fixed effects for the whole sample, 2002-2006. The sample is the merged confidential HMDA and CoreLogic ABS database.

B Summary Statistics and Coefficient Estimates of Control Variables

In this section, we provide summary statistics and the coefficient estimates for the control variables for baseline regressions in the main manuscript. The summary statistics for the control variables are reported in Table B.1. The coefficient estimates for Table 2 in the main draft are reported in Table B.2 for the processing time regression and in Table B.3 for the delinquency regression. The coefficient estimates for Table 3 in the main draft are reported in Table B.4. The coefficient estimates for Table 4 in the main draft are reported in Table B.5.

Table B.1: Summary Statistics

This table reports the summary statistics of all variables based on the merged ABS and HMDA data. The variables are defined in Appendix C.

Variable	Mean	SD	Median	p10	p90
PT	3.23	2.61	2.43	1.00	6.14
TS	13.74	9.35	12.29	4.71	23.43
FICO	633.48	55.35	630.00	564.00	708.00
CLTV	92.89	9.34	99.58	80.00	100.00
Delinq(%)	12.04	32.54	0.00	0.00	100.00
Low Doc	0.39	0.49	0.00	0.00	1.00
Affiliated	0.32	0.47	0.00	0.00	1.00
DTI	29.13	19.88	37.70	0.00	49.30
missing DTI	0.29	0.45	0.00	0.00	1.00
Initial Rate	7.80	1.27	7.70	6.25	9.50
Margin	5.19	2.33	5.88	0.00	7.29
Prepayment Penal.	0.73	0.44	1.00	0.00	1.00
ARM	0.65	0.48	1.00	0.00	1.00
Hybrid2	0.70	0.46	1.00	0.00	1.00
Hybrid3	0.14	0.35	0.00	0.00	1.00
IO	0.16	0.36	0.00	0.00	1.00
Balloon	0.07	0.26	0.00	0.00	0.00
Wage	40.73	9.31	39.41	30.33	53.94
Unemployment	0.05	0.01	0.05	0.04	0.07
Median Income	4.12	0.54	4.09	3.43	4.80
Housing Price Runup	0.15	0.10	0.12	0.04	0.31
Loan Amount	5.00	0.59	4.96	4.22	5.82

Table B.2: Loan-level Processing Time Regression of Discontinuity

This table reports the results of the loan-level regression of discontinuity based on the merged ABS and HMDA data for low-doc loans with FICO between 600 and 640. $\mathbf{1}[FICO \geq 620]$ is an indicator that takes a value of 1 at $FICO \geq 620$ and a value of zero if $FICO < 620$. Standard errors are clustered by lender and year, and t-statistics are reported in parentheses. Significance level: *($p < .10$); **($p < .05$); and ***($p < .01$).

	Whole sample		Low Doc		Full Doc	
	(1)	(2)	(3)	(4)	(5)	(6)
FICO>=620	-0.43***	-0.16	-0.93***	-0.59***	-0.20	-0.03
	(-3.23)	(-1.48)	(-3.64)	(-3.14)	(-1.40)	(-0.28)
Housing price runup	-0.37*	-0.07	0.12	0.36***	-0.53**	-0.21*
	(-1.95)	(-0.63)	(0.69)	(2.77)	(-2.51)	(-1.74)
FICO	0.05	-0.08	-0.66	-0.63	0.39	0.17
	(0.16)	(-0.30)	(-1.26)	(-1.43)	(1.06)	(0.56)
Low-Doc	-0.98***	-0.85***				
	(-3.26)	(-6.84)				
CLTV	-0.87***	-0.56***	-0.65***	-0.26***	-0.92***	-0.65***
	(-5.42)	(-8.38)	(-4.22)	(-3.79)	(-5.36)	(-9.15)
Loan amount	-0.42***	-0.58***	-0.14	-0.27**	-0.47***	-0.73***
	(-3.19)	(-5.83)	(-1.02)	(-2.32)	(-2.97)	(-6.29)
DTI	-0.13	-0.11	-0.46**	-0.19	0.10	-0.03
	(-0.79)	(-1.12)	(-2.35)	(-1.06)	(0.51)	(-0.25)
Missing DTI	0.91	0.00	-0.43	-0.54	1.59	0.36
	(1.15)	(0.00)	(-0.76)	(-1.33)	(1.60)	(0.95)
Initial interest rate	0.55***	0.27***	-0.09	0.16	0.91***	0.27**
	(2.84)	(2.76)	(-0.61)	(1.56)	(3.75)	(2.30)
Margin	-1.34*	-0.34	-0.27	-0.25	-1.84**	-0.31
	(-1.74)	(-1.45)	(-0.58)	(-1.09)	(-1.98)	(-1.16)
Prepayment Penalty	-0.38	0.01	-1.08***	-0.20	-0.21	0.01
	(-0.97)	(0.09)	(-3.04)	(-1.04)	(-0.47)	(0.07)
ARM	1.94	0.22	-1.06	-1.07*	3.28**	0.69*
	(1.44)	(0.58)	(-0.95)	(-1.88)	(2.08)	(1.69)
Hybrid2	-0.51	-0.70*	-1.68**	0.03	-0.01	-1.19**
	(-0.47)	(-1.66)	(-2.31)	(0.06)	(-0.01)	(-2.50)
Hybrid3	0.01	-0.43	-1.30*	0.19	0.64	-0.84*
	(0.01)	(-1.01)	(-1.72)	(0.41)	(0.56)	(-1.72)
Interest Only	2.13*	-0.18	-0.15	-1.65***	3.21**	0.39
	(1.89)	(-0.45)	(-0.15)	(-2.69)	(2.38)	(0.90)
Balloon	0.22	-1.13***	-1.65*	-2.19***	0.80	-0.80*
	(0.20)	(-2.79)	(-1.82)	(-3.68)	(0.62)	(-1.80)
Wage	-0.43***	-0.47***	-0.40***	-0.41***	-0.38***	-0.46***
	(-5.50)	(-9.24)	(-4.44)	(-5.44)	(-4.22)	(-7.96)
Unemployment	-0.12	0.06	-0.10	-0.00	-0.11	0.08
	(-1.50)	(1.05)	(-0.85)	(-0.01)	(-1.32)	(1.37)
Median income	0.23**	0.29***	-0.39***	-0.19*	0.46***	0.50***
	(2.32)	(3.35)	(-3.41)	(-1.94)	(4.24)	(5.63)
Constant	21.50***	22.91***	26.08***	23.75***	19.30***	22.47***
	(11.60)	(39.36)	(24.04)	(40.14)	(8.29)	(34.64)
County, Year FEs	Yes	Yes	Yes	Yes	Yes	Yes
Lender FE	No	Yes	No	Yes	No	Yes
Adjusted R^2	0.051	0.184	0.064	0.184	0.048	0.188
Observations	423129	422378	128733	128211	294396	293733

Table B.3: Loan-level Delinquency Regression of Discontinuity

This table reports the results of the loan-level regression of discontinuity based on the merged ABS and HMDA data for low-doc loans with FICO between 600 and 640. $\mathbf{1}[FICO \geq 620]$ is an indicator that takes a value of 1 at $FICO \geq 620$ and a value of zero if $FICO < 620$. Standard errors are clustered by lender and year, and t-statistics are reported in parentheses. Significance level: *($p < .10$); **($p < .05$); and ***($p < .01$).

	Whole sample		Low Doc		Full Doc	
	(1)	(2)	(3)	(4)	(5)	(6)
FICO>=620	0.70***	0.63***	1.20**	1.27***	0.40	0.31
	(2.70)	(2.60)	(2.47)	(2.79)	(1.41)	(1.22)
Housing price runup	3.46***	3.05***	3.24***	2.99***	3.27***	2.93***
	(9.52)	(9.31)	(7.04)	(6.40)	(9.89)	(9.69)
FICO	-7.44***	-7.09***	-4.55***	-4.40***	-8.82***	-8.34***
	(-12.60)	(-12.20)	(-4.27)	(-4.05)	(-12.21)	(-12.17)
Low-Doc	2.59***	1.60***				
	(7.51)	(6.18)				
CLTV	1.55***	1.58***	2.47***	2.17***	0.99***	1.15***
	(7.58)	(7.99)	(8.99)	(8.08)	(5.56)	(6.50)
Loan amount	3.09***	3.17***	3.47***	3.74***	2.97***	2.99***
	(13.07)	(13.74)	(11.39)	(12.60)	(11.49)	(12.42)
DTI	1.17***	1.40***	1.29***	1.00**	0.84***	1.23***
	(5.02)	(7.19)	(2.64)	(2.37)	(3.40)	(5.54)
Missing DTI	2.61***	3.16***	2.26**	2.20**	2.15***	2.93***
	(3.73)	(6.77)	(2.28)	(2.55)	(2.64)	(5.42)
Initial interest rate	3.05***	3.34***	4.18***	4.36***	2.46***	2.82***
	(13.92)	(20.95)	(17.44)	(15.87)	(10.89)	(18.44)
Margin	1.11***	-0.30	1.34**	0.76	0.99***	-0.62***
	(2.72)	(-0.98)	(2.15)	(1.34)	(2.59)	(-2.81)
Prepayment Penalty	0.25	0.57**	0.98**	1.07***	-0.11	0.25
	(1.10)	(2.54)	(2.53)	(2.71)	(-0.42)	(1.15)
ARM	-2.05*	-0.11	-3.14*	-2.55*	-1.47*	0.84*
	(-1.93)	(-0.15)	(-1.71)	(-1.73)	(-1.82)	(1.84)
Hybrid2	2.48***	4.17***	2.76***	3.28***	2.06***	4.10***
	(5.01)	(8.73)	(3.04)	(3.50)	(3.65)	(7.83)
Hybrid3	1.90***	2.77***	2.22**	2.24**	1.60**	2.67***
	(3.11)	(4.77)	(2.31)	(2.23)	(2.40)	(4.24)
Interest Only	-2.65***	-0.65	-4.19**	-3.35**	-2.03***	0.33
	(-2.67)	(-0.81)	(-2.18)	(-2.06)	(-3.03)	(0.66)
Balloon	1.22	2.92***	1.19	1.12	0.80	3.01***
	(1.43)	(4.68)	(0.90)	(0.88)	(0.96)	(5.54)
Wage	0.42***	0.32***	0.53***	0.42***	0.33***	0.25**
	(4.42)	(3.51)	(3.55)	(3.01)	(2.95)	(2.29)
Unemployment	1.23***	1.19***	1.67***	1.62***	1.05***	1.01***
	(10.19)	(10.33)	(8.94)	(8.80)	(8.26)	(8.54)
Median income	-1.02***	-1.05***	0.14	0.04	-1.58***	-1.56***
	(-5.73)	(-5.91)	(0.41)	(0.11)	(-9.98)	(-9.85)
Constant	8.63***	5.65***	10.31***	9.45***	8.92***	4.89***
	(8.59)	(7.46)	(7.14)	(8.19)	(8.73)	(7.13)
County, Year FEs	Yes	Yes	Yes	Yes	Yes	Yes
Lender FE	No	Yes	No	Yes	No	Yes
Adjusted R^2	0.052	0.057	0.077	0.081	0.038	0.043
Observations	423129	422378	128733	128211	294396	293733

Table B.4: Processing Time and Time to Sale

This table reports the results of a loan-level regression of time to sale on loan processing time based on the merged ABS and HMDA data. The dependent variable is standardized time to sale. The control variables are defined in Appendix C. Standard errors are clustered by lender and year, and t-statistics are reported in parentheses. Significance level: *($p < .10$); **($p < .05$); and ***($p < .01$).

	Whole Sample	Subsample						
		Originator-Sponsor		Documentation		PT (weeks)		
		Affil.	Unaffil.	Low-Doc	Full-Doc	$(0,4]$	$[4,8)$	$[8,\infty)$
PT	0.04***	0.01	0.04***	0.02*	0.05***	0.00	0.04**	0.04*
	(3.42)	(1.22)	(2.79)	(1.89)	(3.26)	(0.16)	(2.05)	(1.90)
Housing price runup	0.22**	0.10	0.11	0.16**	0.22**	0.20**	0.29**	0.29
	(2.36)	(0.61)	(1.60)	(2.13)	(1.98)	(2.38)	(2.07)	(1.55)
FICO	-0.22*	-0.16	-0.18**	-0.23***	-0.23	-0.23*	-0.20*	-0.27*
	(-1.94)	(-1.10)	(-2.15)	(-2.71)	(-1.44)	(-1.88)	(-1.91)	(-1.81)
Low-Doc	0.18	-0.21	0.27***	0.00	0.00	0.17	0.18	0.09
	(1.48)	(-1.37)	(2.70)	(.)	(.)	(1.47)	(1.25)	(0.46)
CLTV	-0.24**	0.17	-0.36***	-0.19*	-0.26*	-0.22*	-0.27**	-0.40***
	(-2.19)	(0.98)	(-3.43)	(-1.66)	(-1.90)	(-1.90)	(-2.29)	(-3.58)
Loan amount	-0.19	-0.09	0.08	-0.10	-0.22	-0.17	-0.25**	-0.34***
	(-1.41)	(-0.56)	(1.52)	(-1.30)	(-1.26)	(-1.15)	(-2.40)	(-2.85)
DTI	-0.18	-0.42	-0.11	0.01	-0.30*	-0.20*	-0.07	-0.09
	(-1.58)	(-1.30)	(-1.38)	(0.11)	(-1.88)	(-1.86)	(-0.49)	(-0.60)
Missing DTI	-3.75***	-2.68***	-3.02***	-3.09***	-4.18***	-3.89***	-3.34***	-3.44***
	(-7.09)	(-2.75)	(-5.48)	(-6.54)	(-6.38)	(-7.35)	(-5.67)	(-5.41)
Initial interest rate	-0.66**	-0.58	-0.39***	-0.45***	-0.75**	-0.56**	-0.73***	-1.51***
	(-2.55)	(-1.21)	(-3.51)	(-3.28)	(-2.24)	(-2.04)	(-3.24)	(-5.18)
Margin	-0.35*	-0.52**	-0.06	-0.28	-0.35	-0.36*	-0.44**	0.05
	(-1.66)	(-2.03)	(-0.33)	(-1.47)	(-1.46)	(-1.72)	(-1.98)	(0.16)
Prepayment Penalty	-0.93***	-0.33**	-1.09***	-0.32***	-1.35***	-0.59***	-1.34***	-3.17***
	(-4.34)	(-2.21)	(-3.90)	(-2.85)	(-4.50)	(-3.69)	(-4.67)	(-5.40)
ARM	1.62***	1.71*	0.81**	1.13**	1.83***	1.46***	1.87***	2.62***
	(3.36)	(1.95)	(1.98)	(2.14)	(3.23)	(2.98)	(3.62)	(3.49)
Hybrid2	-2.12***	-1.45**	-1.79***	-1.68***	-2.35***	-1.70***	-2.39***	-5.56***
	(-3.27)	(-2.13)	(-3.59)	(-3.93)	(-2.66)	(-2.84)	(-3.07)	(-4.42)
Hybrid3	-1.67***	-1.36**	-1.25**	-1.46***	-1.74**	-1.15**	-2.09***	-5.98***
	(-2.64)	(-1.98)	(-2.31)	(-3.17)	(-2.07)	(-1.97)	(-2.84)	(-4.44)
Interest Only	1.86***	1.26	1.12***	1.58***	1.97***	1.69***	2.19***	2.71***
	(4.06)	(1.43)	(2.66)	(3.07)	(3.78)	(3.68)	(4.35)	(3.63)
Balloon	1.55***	1.13	1.33***	1.11**	1.71***	1.43***	1.49***	2.88***
	(3.23)	(1.10)	(2.72)	(2.27)	(2.97)	(2.97)	(2.83)	(3.71)
Wage	-0.02	0.00	-0.06***	-0.04**	-0.02	-0.01	-0.01	-0.10
	(-0.65)	(0.11)	(-2.81)	(-2.31)	(-0.61)	(-0.49)	(-0.29)	(-1.43)
Unemployment	0.02	-0.04	0.06*	-0.05*	0.04	0.00	0.08*	0.06
	(0.48)	(-0.65)	(1.76)	(-1.91)	(0.61)	(0.03)	(1.73)	(0.67)
Median income	0.09	0.06	-0.11***	-0.04	0.14	0.09	0.10	0.16
	(0.86)	(0.66)	(-2.76)	(-0.75)	(1.08)	(0.78)	(1.24)	(1.35)
Constant	15.4***	12.5***	16.6***	14.9***	16.0***	15.0***	15.6***	19.7***
	(23.81)	(20.42)	(28.00)	(28.05)	(19.92)	(25.15)	(20.37)	(15.27)
Adjusted R^2	0.213	0.474	0.240	0.243	0.205	0.211	0.218	0.256
N	1335209	420588	914017	512792	821785	1019613	250689	63752
State, Year FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Lender FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Loan&Borr. Char	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

TABLE B.5: Processing Time, Time to Sale, and Loan Default

This table reports the estimates of control variables of a loan-level regression of loan default on time to sale and processing time based on the merged ABS and HMDA data. The dependent variable is loan delinquency within 10 to 15 months of loan origination. The control variables are defined in Appendix C. Standard errors are clustered by lender and year, and t-statistics are reported in parentheses. Significance level: *($p < .10$); **($p < .05$); and ***($p < .01$).

	(1)	(2)	(3)	(4)	(5)
PT	-0.09***		-0.10***		
	(-4.30)		(-4.34)		
TS		-0.04***	-0.03***		
		(-3.16)	(-3.13)		
LowPT				0.42***	0.29***
				(4.22)	(2.70)
LowTS				0.41*	0.27
				(1.74)	(1.15)
LowPT*LowTS					0.26**
					(2.33)
Housing price runup	2.16***	2.39***	2.39***	2.17***	2.17***
	(7.80)	(8.20)	(8.20)	(7.83)	(7.84)
FICO	-3.34***	-3.51***	-3.52***	-3.34***	-3.34***
	(-20.92)	(-21.66)	(-21.60)	(-20.88)	(-20.88)
Low-Doc	1.74***	1.86***	1.85***	1.74***	1.74***
	(8.16)	(8.23)	(8.23)	(8.15)	(8.15)
CLTV	1.47***	1.57***	1.56***	1.46***	1.46***
	(8.72)	(8.90)	(8.89)	(8.62)	(8.62)
Loan amount	3.52***	3.64***	3.64***	3.51***	3.51***
	(15.24)	(14.65)	(14.70)	(15.22)	(15.22)
DTI	0.88***	0.82***	0.82***	0.88***	0.88***
	(4.98)	(4.46)	(4.45)	(4.99)	(4.99)
Missing DTI	2.16***	1.93***	1.93***	2.08***	2.08***
	(5.52)	(4.60)	(4.60)	(5.24)	(5.24)
Initial interest rate	3.58***	3.52***	3.53***	3.58***	3.58***
	(20.32)	(21.27)	(21.26)	(20.42)	(20.43)
Margin	-0.36	-0.29	-0.30	-0.37	-0.37
	(-1.54)	(-1.22)	(-1.24)	(-1.59)	(-1.59)
Prepayment Penalty	0.66***	0.50**	0.50**	0.66***	0.66***
	(3.48)	(2.35)	(2.36)	(3.47)	(3.48)
ARM	0.42	0.50	0.51	0.45	0.45
	(0.72)	(0.83)	(0.84)	(0.77)	(0.77)
Hybrid2	3.16***	2.77***	2.76***	3.16***	3.16***
	(6.82)	(5.48)	(5.45)	(6.79)	(6.79)
Hybrid3	1.93***	1.49***	1.48***	1.94***	1.94***
	(3.68)	(2.60)	(2.58)	(3.67)	(3.67)
Interest Only	-0.09	-0.07	-0.08	-0.05	-0.05
	(-0.14)	(-0.11)	(-0.11)	(-0.08)	(-0.08)
Balloon	3.43***	3.37***	3.35***	3.45***	3.45***
	(6.68)	(6.10)	(6.09)	(6.73)	(6.73)
Wage	0.17***	0.20***	0.20***	0.17***	0.17***
	(2.75)	(3.34)	(3.25)	(2.74)	(2.73)
Unemployment	1.24***	1.25***	1.25***	1.24***	1.24***
	(12.10)	(11.52)	(11.51)	(12.10)	(12.09)
Median income	-1.39***	-1.44***	-1.43***	-1.39***	-1.39***
	(-9.95)	(-9.90)	(-9.90)	(-9.93)	(-9.93)
Constant	7.64***	8.57***	8.89***	6.93***	7.00***
	(13.56)	(14.08)	(14.95)	(11.51)	(11.72)
Adjusted R^2	0.072	0.071	0.071	0.072	0.072
Observations	1437070	1335209	1335209	1437070	1437070
State, Year, Lender FEs	Yes	Yes	Yes	Yes	Yes
Loan&Borr. Char	Yes	Yes	Yes	Yes	Yes

C Excluding Loans with Month-End Closing Dates

This section provides a robustness analysis of our main findings when excluding the loans with month-end closing dates. The borrowers may choose month-end closing because of liquidity constraints, and thus lengthen the processing time. We exclude the loans that are closed after the 25th day of the month. The results corresponding to Table 2 in the main draft are reported in Table C.1. The results corresponding to Table 3 in the main draft are reported in Table C.2. The results corresponding to Table 4 in the main draft are reported in Table C.3.

Table C.1: Loan-level Regression of Discontinuity

This table reports the results of the loan-level regression of discontinuity based on the merged ABS and HMDA data for low-doc loans with FICO between 600 and 640. $\mathbf{1}[FICO \geq 620]$ is an indicator that takes a value of 1 at $FICO \geq 620$ and a value of zero if $FICO < 620$. Standard errors are clustered by lender and year, and t-statistics are reported in parentheses. Significance level: *($p < .10$); **($p < .05$); and ***($p < .01$).

	Whole sample		Low Doc		Full Doc	
Panel A: Processing Time						
	(1)	(2)	(3)	(4)	(5)	(6)
$\mathbf{1}[FICO \geq 620]$	-0.47***	-0.22*	-1.07***	-0.70***	-0.19	-0.07
	(-3.15)	(-1.72)	(-3.90)	(-3.18)	(-1.15)	(-0.50)
Adjusted R^2	0.053	0.185	0.067	0.184	0.051	0.188
N	291406	290713	89985	89552	201421	200822
Panel B: Delinquency						
	(7)	(8)	(9)	(10)	(11)	(12)
$\mathbf{1}[FICO \geq 620]$	0.74**	0.68**	0.78	0.82	0.64*	0.56*
	(2.34)	(2.24)	(1.22)	(1.35)	(1.89)	(1.80)
Adjusted R^2	0.053	0.057	0.077	0.08	0.039	0.044
N	291406	290713	89985	89552	201421	200822
State, Year FEs	Yes	Yes	Yes	Yes	Yes	Yes
Lender FE	No	Yes	No	Yes	No	Yes
Loan&Borr. Char	Yes	Yes	Yes	Yes	Yes	Yes

Table C.2: Processing Time and Time to Sale

This table reports the results of a loan-level regression of time to sale on loan processing time based on the merged ABS and HMDA data. The dependent variable is standardized time to sale. The control variables are defined in Appendix C. Standard errors are clustered by lender and year, and t-statistics are reported in parentheses. Significance level: *($p < .10$); **($p < .05$); and ***($p < .01$).

	All Loans	Low-Doc	Full-Doc	Subsample by PT (weeks)		
				$(0, 4]$	$[4, 8)$	$[8, \infty)$
PT	0.05***	0.03**	0.05***	0.02	0.04*	0.04
	(3.65)	(2.42)	(3.36)	(0.77)	(1.73)	(1.55)
Adjusted R^2	0.218	0.247	0.21	0.216	0.22	0.259
N	926032	359592	565899	705043	175946	44074
State, Year FEs	Yes	Yes	Yes	Yes	Yes	Yes
Lender FE	Yes	Yes	Yes	Yes	Yes	Yes
Loan&Borr. Char	Yes	Yes	Yes	Yes	Yes	Yes

Table C.3: Processing Time, Time to Sale, and Loan Default

This table reports the results of a loan-level regression of loan default on time to sale and processing time based on the merged ABS and HMDA data. The dependent variable is loan delinquency within 10 to 15 months of loan origination. The control variables are defined in Appendix C. Standard errors are clustered by lender and year, and t-statistics are reported in parentheses. Significance level: *($p < .10$); **($p < .05$); and ***($p < .01$).

	(1)	(2)	(3)	(4)	(5)
PT	-0.08***		-0.09***		
	(-3.27)		(-3.43)		
TS		-0.03***	-0.03***		
		(-2.87)	(-2.84)		
LowPT				0.38***	0.22*
				(3.25)	(1.71)
LowTS				0.38	0.20
				(1.59)	(0.85)
LowPT*LowTS					0.34**
					(2.43)
Adjusted R^2	0.072	0.072	0.072	0.072	0.072
Observations	994401	926032	926032	994401	994401
State, Year FEs	Yes	Yes	Yes	Yes	Yes
Lender FE	Yes	Yes	Yes	Yes	Yes
Loan&Borr. Char	Yes	Yes	Yes	Yes	Yes