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

Manuel Adelino* Bin Wei† Feng Zhao‡

December 21, 2025

Abstract

We develop a dynamic model of asset origination and sale with asymmetric information in which originators exert costly, unobservable screening effort and can signal loan quality through delayed sale. As in models with endogenous screening and partial retention (e.g., Vanasco, 2017), the theory predicts a positive relationship between screening effort and the strength of the signal. We test this central prediction using pre-GFC U.S. mortgage data, measuring screening effort by mortgage processing time. Consistent with the theory, we find that loans that take longer to process are sold with longer delay and are significantly less likely to default, even though observably riskier loans are processed more slowly. This contrast between observable and unobservable risk underscores the role of lender effort in generating hidden quality. Calibrating our model to pre-GFC data reveals that markets were operating in a parameter range that admits multiple equilibria and induces endogenous fragility.

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

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

We thank Dong Beom Choi, Nuno Clara, Simon Gervais, George Malikov (discussant), Yoshio Nozawa (discussant), João Thereze, and Xingtan Zhang (discussant) as well as participants at the Bank of Portugal, Duke, Federal Reserve Bank of Dallas, University of Kentucky, the University of Oklahoma, the University of Texas at Austin, the University of Texas at Dallas, 2024 EFA annual meeting, 2024 NFA annual meeting, 2024 SED Conference, and 2024 Cavalcade Asia-Pacific 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 setting with asymmetric information (Akerlof, 1970), sellers of high-quality assets have incentives to undertake costly actions to differentiate themselves from low-quality sellers, as in the classic signaling model of Spence (1973). Such costly actions may include partial retention of the asset (Leland and Pyle, 1977; DeMarzo, 2005), or, in dynamic settings, delaying the sale (Janssen and Roy, 2002; Daley and Green, 2012). While these actions help convey asset quality to uninformed buyers, they impose liquidity costs on the seller by preventing complete or immediate transactions of the asset, which diminishes potential gains from trade. This trade-off between asset quality and market liquidity is a fundamental feature of these types of transactions.

This paper studies the quality-liquidity trade-off when an informed seller can improve asset quality through *ex ante* screening. We consider an environment (motivated by mortgage lending) in which a loan originator privately exerts costly effort at origination to screen borrowers and produce a higher-quality loan. On the one hand, reduced secondary-market liquidity (e.g., the inability to quickly sell or securitize a loan) can change the incentive to screen at origination (Parlour and Plantin, 2008; Malherbe, 2012; Chemla and Hennessy, 2014). On the other hand, intensive screening may itself exacerbate adverse selection in the secondary market and reduce market liquidity (Vanasco, 2017), since better-informed sellers have an informational advantage over buyers. The interaction of these forces suggests that information frictions at origination and at sale are jointly important, though identifying them empirically is challenging. We show that the patterns observed in the U.S. private-label mortgage-backed security (PLS) market are consistent with both moral hazard (hidden effort) at loan origination and adverse selection at loan sale, and we establish a novel link between them.

We present a model of mortgage origination and securitization that embeds both stages of this process, extending Vanasco (2017). In our model, a mortgage originator can exert unobservable screening effort at origination to improve the loan's quality. After origination, the originator or lender faces a secondary market with uninformed investors and must decide when to sell the loan (immediately or after a delay). This naturally produces two information frictions: at the origination stage, hidden screening effort creates a moral hazard problem; at the securitization stage, the originator's private information about loan quality creates adverse selection among potential buyers. By construction, this setup nests the classic cases of moral hazard at origination (if secondary-market asymmetries are absent) and adverse selection at sale (if no screening effort is possible) as special cases. More importantly, it allows us to examine how these frictions interact within a single unified framework.

In equilibrium, higher-quality loans are both screened more intensively and sold later. Originators of better loans wait longer to sell in order to signal quality that is not captured by observable loan features, consistent with existing models of dynamic signaling via trading delays (e.g., Janssen and Roy, 2002; Daley and Green, 2012; Adelino, Gerardi, and Hartman-Glaser, 2019). Our theoretical contribution is to embed this dynamic signaling mechanism into a model with endogenous screening and to derive sharp, testable predictions for the joint behavior of screening effort, sale timing, and loan performance. In the model, the anticipation of the ability to signal high loan quality by delaying sale feeds back into the origination stage: when originators expect to use a delayed sale as a signal, they are more willing to exert costly screening effort up front. This creates a screening–signaling complementarity in which stronger incentives to signal in the future encourage more screening in the present, and that screening makes the signal more credible. As a result, the same loans that are ultimately sold with a delay (to signal quality) also take longer to originate (because they were more carefully screened).

Relative to Vanasco (2017), who studies how costly retention can sustain screening in a static setting, we consider the *timing* of sale as the primary signaling instrument in a continuous-time environment tailored to the mortgage PLS market, following (Adelino et al., 2019, AGH).[1] A second theoretical innovation is to explicitly model how the ease of securitization depends on a public signal of borrower quality—such as FICO scores—through a "rule-of-thumb" securitization cutoff. This feature allows us to generate predictions about discontinuities in screening effort, sale timing, and loan quality around the cutoff, which we can take directly to the data to validate processing time as an empirical proxy for effort. Finally, we calibrate the model in a quantitative exercise to illustrate endogenous fragility in the PLS market due to the existence of multiple equilibria.

We test this prediction using the U.S. private-label mortgage-backed securities (PLS) market during 2002-2007 as a laboratory. We measure screening effort as the time between mortgage application and mortgage closing (which we refer to as "processing time"). Lenders typically use this interval 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, or performing home inspections, and may be related to preferences for end-of-month closing dates (Bhutta and Ringo, 2021). It is also well 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

[1] Fuchs, Gottardi, and Moreira (2025) discuss in detail why the timing of sale can be a more robust signaling mechanism.

screening effort and thus leads to the origination of (unobservably) higher quality loans.[2]

We start by validating processing time as a proxy for screening effort. We exploit a sharp change in screening incentives around a Fair Isaac Corporation (FICO) credit score of 620, a widely used credit score cutoff. The key assumption in this analysis is that borrowers' demand for loans is smooth around the 620 threshold, but that the threshold 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, Mukherjee, Seru, and Vig (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, and suggest that this is evidence of a sharp change in screening effort.

Consistent with our model and the results in the literature regarding performance, we find that loans with FICO scores just below 620 take significantly longer to process than those just above 620. This discontinuity in processing time is most pronounced for low documentation loans with limited borrower documentation, but is more muted in segments with full documentation. The fact that otherwise similar loans below the cutoff undergo additional processing provides direct evidence that longer processing times reflect greater lender effort, rather than just unobserved borrower traits or bureaucratic delays.

Turning to the full sample, we show that this result is not limited to the region around the 620 FICO cutoff: loans with higher processing time are generally associated with lower *ex post* default once we control for all available loan and borrower characteristics. Securitized mortgages are an appealing asset to consider the role of unobserved asset quality because of the availability of a large set of observable characteristics that are known to be related to default. After accounting for all these variables, we show that processing time is still related to performance, consistent with the hypothesis that originators gather information about loan quality during processing over and above the characteristics that are recorded in the datasets.

Importantly, processing time for *observably* worse loans is longer, not shorter. We build a proxy for observable credit risk by combining all available loan characteristics into a credit model and constructing a measure of predicted default. We find that higher predicted default is associated with more days of processing by the lender, the opposite of what we find for the unobservable component. This suggests that the effects we find operate through quality

[2] Choi and Kim (2021) compare conforming and non-conforming loans after the subprime crisis and provide evidence that processing time is a good measure of screening effort. 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.

differences that are unobservable to the buyers in this market, precisely the ones we would expect to lead to signaling at the time of sale.

Next, we show that the central prediction for the positive relationship between screening effort and signaling is borne out by the data. Following AGH, we 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, for example, 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. This positive relation holds in the Alt-A and subprime subsamples, and holds after accounting for very fine origination, issuance and state fixed effects.

Moreover, both processing time and time-to-sale are predictive of better loan performance after accounting for a large set of observable characteristics: loans with longer processing time and/or held longer before sale display lower default rates, all else equal. Notably, each of these variables contains independent information about loan quality. In regressions predicting default, neither processing time nor time-to-sale fully subsumes the effect of the other. This implies that early screening and post-origination signaling are each related to reduced unobservable default risk. In particular, the "skimming" property—i.e., loans sold later perform better—still holds in our data, even after controlling for processing time. Together, these findings provide the first evidence that the ability to signal asset quality in secondary markets is related to stronger *ex ante* screening incentives on dimensions that are not observable to the final investors. In short, originators screen more and sell later, and this joint strategy is associated with better *ex post* loan outcomes.

Finally, we explore the broader implications of the screening–signaling mechanism. We calibrate our model to pre-2008 mortgage market data and find that the estimated parameters lie in a region of the model that admits multiple equilibria. In this region, market outcomes can endogenously shift between an efficient high-screening equilibrium and an inefficient low-screening one depending on participants' beliefs. The calibration suggests that, as securitization boomed in the mid-2000s, the market may have drifted toward a low-effort equilibrium in which both screening and signaling incentives were weak—a dynamic consistent with the deteriorating lending standards observed before the 2008 financial crisis. Moreover, the proximity of the economy to a fragile threshold in the multiple-equilibrium region means that even small shocks or changes in investor sentiment could have triggered a rapid collapse in secondary market liquidity. In our framework, this endogenous fragility offers a novel explanation for the sudden and severe freeze of the private-label securitization market during the crisis, beyond what traditional single-equilibrium models can capture.

Related Literature. This paper relates to two classic strands of the information economics literature. The first is the literature on adverse selection and signaling, pioneered by Akerlof (1970) and Spence (1973), and applied to asset sales in Leland and Pyle (1977), Myers and Majluf (1984), and Gorton and Pennacchi (1995). In these settings, uninformed buyers are concerned about the presence of low-quality assets ("lemons") that they cannot identify, forcing the informed seller to signal quality—either through retention of a stake in the asset (Leland and Pyle, 1977; DeMarzo and Duffie, 1999; DeMarzo, 2005) or by delaying sale (Janssen and Roy, 2002; Daley and Green, 2012)—in order to obtain higher prices. Such signaling is costly, however, because asset cash flows are not fully allocated to the highest-value party (the buyer in this case).[3]

The second strand is the literature on moral hazard in delegated decision-making, classically modeled by Holmström (1979) as a hidden-action problem in which optimal contracts trade off risk-sharing against incentives to exert effort. In asset markets, moral hazard arises when sellers' unobservable actions at the production or origination stage affect asset quality. For example, in the mortgage context, Keys et al. (2010) and Purnanandam (2010) show that weaker screening incentives as a result of loan sale lead to worse ex-post loan performance. These studies suggest that hidden effort at origination is a first-order determinant of asset quality.

Our model and empirical results connect these two strands by showing that moral hazard at origination and adverse selection in secondary markets are jointly determined. A related body of work examines the trade-off between incentives to originate good assets ("productive efficiency") and secondary market liquidity ("allocative efficiency"). Parlour and Plantin (2008) study the effect of loan sales on banks' origination decisions, while Chemla and Hennessy (2014) analyze how speculative information production or optimal regulation shapes this trade-off (see also Dell'Ariccia and Marquez, 2006; Malherbe, 2012). Vanasco (2017) shows that costly retention of cash flows is essential to implement *ex ante* asset screening. Our model is most closely related to Vanasco (2017), who studies a setting with ex-ante screening and asymmetric information in which originators can signal asset quality by retaining a larger share of future cash flows. In her framework, costly retention and screening are jointly determined: the ability to signal through retention sustains higher screening effort, and better screening increases the value of retention as a signal. We build on this insight but depart in two key ways. First, following AGH, we consider an environment in which the primary signaling instrument is the *timing* of sale rather than the share retained. Delaying

[3] Begley and Purnanandam (2017) study retention of equity tranches in residential mortgage-backed securities and find that higher retained tranches are associated with lower delinquency rates. Other work examines the role of ratings as a public signal that may alleviate some of these frictions, but potentially at the cost of reduced asset quality (Kremer and Skrzypacz, 2007; Daley, Green, and Vanasco, 2020).

securitization is costly because the originator bears default risk and discounting while the loan remains on the balance sheet, but it serves as a separating signal in equilibrium: bad loans are sold immediately, whereas good loans are sold with delay. Fuchs et al. (2025) show that when there is no commitment, trading is continuous, and trade histories are publicly observed, the timing of trade is used exclusively—rather than retention—as the signaling device in a dynamic adverse selection environment. Intuitively, delay of sale is a credible signal because it is inherently irreversible, whereas retention can, in principle, be undone by future trades unless such trades can be ruled out. Second, by explicitly modeling both the origination stage and the dynamic sale decision, our framework yields new comparative statics involving two observable time dimensions—processing time and time-to-sale—and links them to ex-post performance. These additional predictions allow us to take the model directly to loan-level data and provide empirical evidence on the joint determination of screening, signaling, and market liquidity.

The remainder of the paper is organized as follows. Section 2 presents the model setup and derives the equilibrium with endogenous screening and signaling. Section 3 tests the model's predictions using mortgage loan data. Section 4 discusses the calibration and implications for market stability. Section 5 concludes.

2 A Model of Screening and Signaling

We extend the signaling model of AGH by incorporating endogenous screening effort, following the approach of Vanasco (2017). In our framework, delayed sale, not loan retention as in Vanasco (2017), serves as the signaling device. While retention and delay of trade can serve similar signaling purposes, Fuchs et al. (2025) (discussed above) show that, in the absence of commitment, the timing of trade is used instead of retention in a dynamic adverse selection environment. Our framework further allows us to incorporate a notion of ease of securitization that depends on publicly observable information (e.g., credit scores). The model yields a set of testable implications, which we examine empirically in the next section.

2.1 Setup

The model proceeds in two stages: an *origination stage* followed by a *securitization stage*. An originator first screens the loan by exerting unobservable effort in the origination stage, and then sells the loan to outside investors in the securitization stage. At the end of the period, the state of the economy and the cash flows from the originated loan are realized.

Origination Stage. At the start of the period, the originator has access to a default origination channel that yields a risky, low-quality loan, denoted b-type ("bad"). A pool of such loans is always available, and originating from this pool requires negligible effort, which we normalize to zero. Each loan generates a constant cash flow of c dollars per unit of time until default. The default time τ_d follows a Poisson process with intensity $\overline{\lambda} > 0$.

Alternatively, the originator can exert positive effort $a \in (0, 1]$ using a private screening technology to identify higher-quality borrowers. Successful screening yields a high-quality loan, denoted type-g ("good"), characterized by a lower default risk. Specifically, the default intensity of a g-type loan is given by $\lambda(a) \in [\underline{\lambda}, \overline{\lambda}]$. We assume that $\lambda(0) = \overline{\lambda} = \lambda_b$, and $\lambda(a)$ is a monotonically decreasing convex function of effort a (analogous to the "endogenous quality" case in Vanasco (2017)).

Screening succeeds with probability a, in which case a g-type loan is originated. With probability $1 - a$, screening fails, in which case no high-quality borrower is identified and the exerted effort is entirely *sunk*.[4] Upon a failed screening, the originator reverts to the default channel and originates a b-type loan at zero effort. Hence, in equilibrium, every b-type loan is originated without any screening effort, and only g-type loans are associated with positive effort.

Exerting effort is costly, involving nonpecuniary cost $C(a)$, where $C : [0, 1] \to \mathbb{R}^+$, $C(0) = C'(0) = 0$, and $C'(\cdot)$, $C''(\cdot) > 0$ for $a \in (0, 1]$.

Importantly, neither the screening effort nor the resulting loan type is observable to investors at the point of sale, creating an information asymmetry (or more precisely, effort affects loan quality over and above the observable characteristics that are known to the investor). However, investors know that b-type loans are always originated effortlessly ($a_b = 0$) and have constant default intensity $\lambda_b = \overline{\lambda}$, while g-type loans are originated with positive effort ($a_g > 0$) and have default intensity $\lambda_g(a_g) = \lambda(a_g)$ as in equation (12). For notational simplicity, we drop type-specific subscripts whenever no confusion arises: a will denote the effort level for g-type loans (and zero effort for b-type loans), and $\lambda(a)$ will refer to the default intensity of g-type loans (and $\overline{\lambda}$ for b-type loans).

Securitization Stage. After origination, the originator decides when to sell the loan to competitive outside investors in the secondary market. Loan quality is not directly observable to buyers at the time of sale. However, sale timing can serve as a signal of quality: for

[4]More precisely, when screening fails, the originator does not identify a suitable high-quality borrower and therefore rejects the loan application. The effort expended in this failed attempt is fully sunk, and the originator subsequently turns to the default origination channel to originate a b-type loan at zero effort. In the empirical analysis, we observe only originated (i.e., accepted) loans, which correspond to b- and g-type loans in the model. Rejected applications—representing failed screening attempts—are not observed in our data and therefore play no direct role in our analysis.

instance, delaying sale may credibly convey that the loan is of higher quality.

Formally, with private information about loan type $z \in \{g, b\}$ and her hidden action a, the originator in the securitization stage solves the following problem for optimal time to sale:

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

where γ denotes the discount rate, $p(t) : \mathbb{R}^+ \to \mathbb{R}^+$ denotes the mapping from the time to sale to the price paid by investors, and $\mathbb{E}_a^z[\cdot]$ denotes the expectations operator over the cash flows of the z-type originator that exerted effort a. Recall that the default time τ_d follows a Poisson process with the default intensity equal to $\lambda(a)$ in (12) for a g-type loan, or $\overline{\lambda}$ for a b-type loan. When the expectation does not depend on effort a, we drop the subscript.

Both the originator and investors are risk-neutral with the same discount rate $\gamma > 0$. However, investors derive more utility from the cash flows: θc, where $\theta > 1$. The difference in utility from the same cash flows generates gains from trade. That the investors derive more utility from the mortgage cash flows (or have a higher discount rate, which would similarly result in gains from trade) may reflect credit constraints faced by the originator, who thus has an incentive to sell the loan, or other benefits from holding the loan (e.g., portfolio diversification or maturity matching).

Investors form beliefs about the originator's screening action, denoted by a^e, and about the type of the loan, denoted by $\mu : \mathbb{R}^+ \to [0, 1]$, where $\mu(t)$ is the probability of the loan being a g-type if she chooses to sell the loan at time t. Since the market is competitive, the market valuation for the loan is set so that investors make zero profits in expectations

$$\begin{aligned} p(t) &= \mathbb{E}_{a^e}^{\mu} \left[\int_t^\infty \theta c e^{-\gamma(u-t)} 1_{\tau_d \geq u} du \right] \\ &= \mu(t) \mathbb{E}_{a^e}^{g} \left[\int_t^\infty c e^{-\gamma(u-t)} 1_{\tau_d \geq u} du \right] + (1 - \mu(t)) \mathbb{E}^b \left[\int_t^\infty \theta c e^{-\gamma(u-t)} 1_{\tau_d \geq u} du \right] \\ &= \mu(t) \frac{\theta c}{\gamma + \lambda(a^e)} + (1 - \mu(t)) \frac{\theta c}{\gamma + \lambda_b}. \end{aligned} \tag{2}$$

As a benchmark, we first consider the first-best before we characterize the market equilibrium allocations in the following subsection. We assume the first-best has an interior solution for the optimal effort. That is, $\exists\, a^{FB} \in (0, 1)$ such that $a^{FB} = \arg\max_{a \in [0,1]} \theta \rho(a) - C(a)$, where $\rho(a)$ represents the expected payoff to the originator in the no-sale case, given by

$$\rho(a) \equiv a \frac{c}{\gamma + \lambda(a)} + (1 - a) \frac{c}{\gamma + \lambda_b}. \tag{3}$$

Proposition 1. *In the first-best, the originator sells the loan immediately,* $t_g^{FB} = t_b^{FB} = 0$, *and exerts effort $a^{FB} > 0$ at time 0 given by*

$$\theta\rho'\left(a^{FB}\right) - C'\left(a^{FB}\right) = 0. \tag{4}$$

In the first-best, the loan is sold to investors who value loan payments more, irrespective of loan type z. Consequently, the originator chooses screening effort such that the social benefit of asset screening equals its social marginal cost, which results in full productive efficiency. At the same time, she sells the loan immediately at time 0, thus rendering full allocative efficiency.

2.2 Equilibrium characterization

In this section, we solve the model by backward induction and characterize the market equilibrium allocations. We begin with the securitization stage: given the originator's screening effort a, investors' beliefs a^e and loan type $z \in \{g, b\}$, we solve for the originator's optimal time to sale. This yields the optimal securitization strategies: $t_z(a, a^e)$ for each loan type z. We then move to the origination stage. Taking the optimal securitization strategies as given, we determine the originator's optimal effort $a^*(a^e)$ as a function of investors' beliefs a^e. In equilibrium, the optimal effort chosen by the originator coincides with investors' beliefs, satisfying $a^*(a^e) = a^e$. Throughout, we index the originator's type by the loan type $z \in \{g, b\}$. A z-type originator is defined as an originator who produces a loan of type $z \in \{g, b\}$.

2.2.1 Securitization market equilibrium: Signaling through delay of sale

Given any level of effort a and market beliefs a^e, an equilibrium in the securitization market is given by a pricing function $p : \mathbb{R}^+ \to \mathbb{R}^+$, a z-type originator's strategy of when to sell t_z for $z \in \{g, b\}$, and belief function $\mu : \mathbb{R}^+ \to [0, 1]$ satisfying the following conditions:

- Originator's Optimality: Given $p(\cdot)$, t_z is the solution to (1) for $z \in \{g, b\}$.
- Belief Consistency: $\mu(\cdot)$ is derived from t_g and t_b using Bayes's rule when it applies.
- Zero Profit Condition: $p(t)$ is determined from (2).

To characterize the securitization market equilibrium, we distinguish between two cases: $a^e = 0$ and $a^e > 0$. In the first case with $a^e = 0$, from equation (2), we have $p(t) = \frac{\theta c}{\gamma + \lambda_b}$ because $\lambda(0) = \lambda_b$. Under the belief of zero effort, the loan is priced as if it were always of b-type. If gains from trade are sufficiently large—formally, $\lambda_b \leq (\theta - 1)\gamma + \theta\lambda(a)$—then the equilibrium is pooling: both b- and g-type loans are sold immediately. Otherwise, the equilibrium is separating: the originator sells b-type loans immediately but holds g-type loans indefinitely.

In the second case with $a^e > 0$, Lemma 1 in the appendix shows that pooling equilibria cannot survive D1-refinements; thus, the equilibrium must be separating. In this case, $\mu t_b = 0$ and $p(t_b) = \frac{\theta c}{\gamma+\lambda_b}$, so the b-type originator earns full-information payoff and optimally sells immediately ($t_b = 0$). On the other hand, if gains from trade are sufficiently large—specifically, $\lambda(a^e) \leq (\theta - 1)\gamma + \theta\lambda(a)$—then the g-type originator sells after the minimal delay necessary to satisfy the b-type's incentive compatibility constraint, denoted $t_g(a^e)$ in equation (5) below.

Proposition 2 formally characterizes the securitization market equilibrium.

Proposition 2. *Let $\{a, a^e\}$ be given. There exists a unique equilibrium in the securitization market in which the b-type originator sells immediately, $t_b = 0$, while the g-type originator sells after a delay $t_g > 0$, where*

$$t_g(a, a^e) = \begin{cases} t_g(a^e) \equiv -\frac{1}{\gamma+\lambda_b} \log\left(\frac{(\theta-1)(\gamma+\lambda(a^e))}{(\theta-1)\gamma+\theta\lambda_b-\lambda(a^e)}\right), & \textit{if } \lambda(a^e) \leq (\theta-1)\gamma+\theta\lambda(a), \\ \infty, & \textit{if } \lambda(a^e) > (\theta-1)\gamma+\theta\lambda(a). \end{cases} \tag{5}$$

When $a = a^e$, the equilibrium corresponds to the least costly separating equilibrium (LCSE), in which the g-type originator delays sale by the minimum amount necessary to deter imitation by the b-type, thereby revealing loan quality while incurring the lowest possible signaling cost. Specifically, in this case, the b-type originator sells the loan immediately, whereas the g-type originator delays sale just enough—given by $t_g(a^e)$—to satisfy the incentive compatibility constraint of the b-type originator.

When $a < a^e$ or when gains from trade are sufficiently large, the g-type originator continues to follow the LCSE strategy as if $a = a^e$, since in these cases selling the loan is still preferable. In contrast, when $a > a^e$ *and* gains from trade are not large enough (i.e., when the condition if $\lambda(a^e) > (\theta - 1)\gamma + \theta\lambda(a)$ holds), the g-type originator deviates from the LCSE strategy and prefers to hold the loan indefinitely rather than sell it. The reason is that investors undervalue the loan due to underestimating its quality; the equilibrium price in equation (2) reflects $\lambda(a^e)$ rather than $\lambda(a)$, leading to a valuation gap so severe that the originator prefers to hold rather than sell. This deviation is an *off-equilibrium-path* phenomenon: it requires $a > a^e$, which cannot occur in equilibrium because beliefs must be correct ($a = a^e$). Nevertheless, it is important for understanding the robustness of the separating structure—if beliefs were incorrect, sufficiently high actual effort could lead to a complete breakdown of trade.

2.2.2 Full equilibrium: Endogenous screening effort

Next, we solve for the optimal screening effort in the origination stage in characterizing the equilibrium of the full game.

First, note that a securitization market equilibrium outcome is a set of prices and time to sale per originator type, denoted by $\Phi(a, a^e) \equiv \{p_z, t_z\}_{z\in\{g,b\}}$, where t_z for $z \in \{g, b\}$ is determined in Proposition 2. In the case with $a^e > 0$, the securitization market equilibrium is separating, in which $p_b(a, a^e) = \frac{\theta c}{\gamma+\lambda_b}$ and $p_g(a, a^e) = \frac{\theta c}{\gamma+\lambda(a^e)}$. Given $\Phi(\cdot, \cdot)$, the value to the originator at time 0 is

$$\begin{aligned} V_0(a, a^e) &= a\mathbb{E}_a^g\left[\int_0^{t_g} ce^{-\gamma u}1_{\tau_d\geq u}du + e^{-\gamma t_g}1_{\tau_d\geq t_g}p_g\right] \\ &\quad +(1-a)\,\mathbb{E}^b\left[\int_0^{t_b} ce^{-\gamma u}1_{\tau_d\geq u}du + e^{-\gamma t_b}1_{\tau_d\geq t_b}p_b\right] \\ &\quad -C(a). \end{aligned} \tag{6}$$

An equilibrium of the full game is given by $\left\{a^e, a^*, p_g^*, p_b^*, t_g^*, t_b^*\right\} \in [0,1]^2 \times \mathbb{R}_+^4$ satisfying the following conditions:

- Originator's Optimality at time 0: Given a^e and $\Phi(\cdot, a^e)$, $a^*(a^e) = \arg\max_{a\in[0,1]} V_0(a, a^e)$.
- Securitization Market Equilibrium: $\{p_z^*, t_z^*\}_{z\in\{g,b\}} = \Phi(a^*, a^*)$.
- Belief Consistency: $a^e = a^* = a^*(a^*)$.

The following proposition characterizes the full equilibrium.

Proposition 3. *In any market equilibrium, effort and time to sale $\left\{a^*, t_g^*\right\}$ must satisfy the following two conditions:*

$$\begin{aligned} &\frac{c}{\gamma+\lambda(a^*)} - \frac{\theta c}{\gamma+\lambda_b} + e^{-(\gamma+\lambda(a^*))t_g^*}\frac{(\theta-1)\,c}{\gamma+\lambda(a^*)} \\ &+a^*\lambda'(a^*)\left(-\frac{c}{(\gamma+\lambda(a^*))^2}\left(1-e^{-(\gamma+\lambda(a^*))t_g^*}\right) - \frac{(\theta-1)\,ct_g^*}{\gamma+\lambda(a^*)}e^{-(\gamma+\lambda(a^*))t_g^*}\right) \\ = \;& C'(a^*), \end{aligned} \tag{7}$$

and

$$t_g^* = -\frac{1}{\gamma+\lambda_b}\log\left(\frac{(\theta-1)(\gamma+\lambda(a^*))}{(\theta-1)\gamma+\theta\lambda_b-\lambda(a^*)}\right). \tag{8}$$

There are at least two solutions to (7) and (8): one with $a^* > 0$ and $t_g^* > 0$, and another with $a^* = 0$ and $t_g^* = 0$. From now on, we denote the former solution with positive effort by $\left\{a^{ME}, t_g^{ME}\right\}$ and the latter with zero effort simply by $\{0, 0\}$.

The optimal effort in Proposition 3 is obtained by solving the following problem to maximize the originator's value at $t = 0$:

$$\max_{a\in[0,1]} a\left[\frac{c}{\gamma+\lambda(a)}\left(1-e^{-(\gamma+\lambda(a))t_g(a,a^e)}\right)+e^{-(\gamma+\lambda(a))t_g(a,a^e)}\frac{\theta c}{\gamma+\lambda(a^e)}\right]+(1-a)\frac{\theta c}{\gamma+\lambda_b}-C(a),$$

where $t_g(a, a^e)$ is given by (5). Differentiating the above objective with respect to a yields the first-order condition for the solution $a^*(a^e)$. Imposing further the condition $a^* = a^*(a^*) = a^e$ yields (7). Furthermore, substituting $a = a^e = a^*$ into $t_g(a, a^e)$ in (5) yields the expression for the optimal time to sale $t_g^* \equiv t_g(a^*, a^*)$ in (8).

When the condition $\lambda(a^e) \leq (\theta-1)\gamma + \theta\lambda(a)$ holds (which is the case in equilibrium), then the above objective of the originator at $t = 0$ can be simplified as

$$\max_{a\in[0,1]} a\frac{\theta c}{\gamma+\lambda(a^e)}+(1-a)\frac{\theta c}{\gamma+\lambda_b}-C_R(a;a^e)-C(a),$$

where

$$C_R(a;a^e)\equiv a\left(1-e^{-(\gamma+\lambda(a))t_g(a^e)}\right)\left(\frac{\theta c}{\gamma+\lambda(a^e)}-\frac{c}{\gamma+\lambda(a)}\right). \tag{9}$$

Intuitively, the originator chooses effort such that the private marginal benefit of exerting effort equals the marginal cost to the originator as shown in (7). Exerting additional effort enhances the value derived from an immediate sale of the loan (i.e., $\frac{\theta c}{\gamma+\lambda(a^e)} - \frac{\theta c}{\gamma+\lambda_b}$). Second, exerting more effort has two counteracting effects on the signaling cost $C_R(a; a^e)$. On the one hand, more effort increases the probability of originating a g-type loan, which leads to more signaling and tends to increase $C_R(a; a^e)$. On the other hand, by reducing default intensity, exerting more effort helps reduce the cost of signaling.

It is worthwhile to point out that the main result in AGH holds here: when the expected loan quality is lower (i.e., larger $\lambda(a)$), t_g 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 an unobservably better quality loan waits longer to trade and uses the delayed trade as a signal of better quality.

2.2.3 Multiple equilibria and market fragility

To show that an equilibrium without commitment exists, we must rule out any profitable deviation from the candidate equilibrium strategies. In particular, relevant deviations are *double deviations*: the originator may deviate from the prescribed effort at the origination stage and subsequently choose an off-path sale time at the securitization stage.

Note first that "downward deviations" to a lower effort level are not profitable. By Proposition 2, the LCSE condition continues to hold under such deviations, so remaining in the LCSE implies that a reduction in effort does not affect prices, which are pinned down by investors' beliefs a^e. With prices unchanged, lowering effort can only reduce the originator's payoff. Thus, we only need to rule out "upward" deviations in which the originator exerts higher effort, violating the LCSE condition in (5). In that case, Proposition 2 implies that the originator optimally holds the loan indefinitely and does not sell it. The best such double deviation, a^{NS}, solves

$$a^{NS} = \arg\max_{a \in [0,1]} \left\{ a \frac{c}{\gamma + \lambda(a)} + (1-a) \frac{\theta c}{\gamma + \lambda_b} - C(a) \right\}.$$

Moreover, from Proposition 3, any equilibrium of the full game, if it exists, must satisfy (7) and (8). These conditions pin down two candidate solutions: (i) a corner solution $\{0, 0\}$ and (ii) an interior solution $\{a^{ME}, t_g^{ME}\}$ with $a^{ME} > 0$ and $t_g^{ME} > 0$. The next result shows that the full game can feature multiple equilibria—or none at all, in which case the originator simply originates and holds the loan.

Proposition 4. *There exist two thresholds for gains from trade, $\underline{\theta}$ and $\overline{\theta}$ with $1 < \underline{\theta} < \overline{\theta}$, such that: (i) for $\theta \geq \overline{\theta}$, there are two equilibria, $\{0, 0\}$ and $\{a^{ME}, t_g^{ME}\}$; (ii) for $\theta \in [\underline{\theta}, \overline{\theta})$, there is a unique equilibrium, $\{a^{ME}, t_g^{ME}\}$; and (iii) for $\theta < \underline{\theta}$, an equilibrium may fail to exist.*

Proposition 4 highlights a source of market fragility. When gains from trade are sufficiently large, both $\{0, 0\}$ and $\{a^{ME}, t_g^{ME}\}$ can be sustained as equilibria. If the market believes the originator exerted no effort, the originator's best response is to sell immediately, validating those beliefs; if instead the market expects effort a^{ME}, the originator optimally delays sale by t_g^{ME}, again validating beliefs. In this sense, beliefs are self-fulfilling. Notably, the "bad" equilibrium with zero effort and immediate sale is consistent with the observed deterioration in loan quality during the run-up to the subprime mortgage crisis.

When gains from trade lie in an intermediate range, the low-effort, immediate-sale profile $\{0, 0\}$ is no longer an equilibrium, because the originator can profitably deviate by exerting high effort and then holding the loan indefinitely rather than selling it at the undervalued market price.

Moreover, when gains from trade fall sharply and become sufficiently low, secondary markets can seize up: the originator cannot sell and must hold the loan indefinitely. This outcome is consistent with the post–subprime-crisis collapse of the PLS securitizations market.

2.3 Model implications and predictions

In this subsection, we discuss the main model implications and develop hypotheses for empirical tests in the next section.

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. That is, 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.

Corollary 1. *Compared to b-type loans, the originator exerts higher effort and waits longer to sell for the g-type loans that have a smaller default likelihood:*

$$\begin{aligned} a_g^* &> 0 = a_b^*, \\ t_g^* &> 0 = t_b^*, \\ \lambda\left(a_g^*\right) &< \lambda_b = \lambda\left(a_b^*\right). \end{aligned}$$

Corollary 1 highlights the tension between asset quality and market liquidity: screening improves asset quality but exacerbates information asymmetry, causing more delay in the 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.

In the model, both effort and time to sale are strictly positive for g-type loans, but zero for b-type loans. Therefore, if we regress time to sale on effort, our model implies the following coefficient for effort:

$$\beta_{TS\_PT} = \frac{t_g^*}{a^*} > 0.$$

If we regress default intensity on effort, the model-implied coefficient of effort in this regression is:

$$\beta_{DEF\_PT} = \frac{\lambda\left(a^*\right) - \lambda_b}{a^*} < 0.$$

Similarly, the model-implied coefficient of time to sale in a univariate regression of default intensity on time to sale is:

$$\beta_{DEF\_TS} = \frac{\lambda(a^*) - \lambda_b}{t_g^*} < 0.$$

We thus arrive at the following prediction.

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. If either time to sale or processing time is measured with error, or other factors affect either one of them (for example, inventory considerations for time to sale, or differences in lender technology for processing time), we would not expect time to sale to be a sufficient statistic for loan quality (and vice versa).

Securitization rule of thumb. A fundamental challenge in testing the central predictions of our and similar models 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 originated and, in the case of low-documentation loans, also more easily securitized during the time period of our sample (Keys et al., 2010, 2012). This rule of thumb implies a discrete positive increase in the ease of securitization for this subset of loans at the 620 threshold.

To capture the "rule of thumb" in the model, we assume that the securitization probability is a function of an observable characteristic s, and there exists a securitization threshold, denoted by s^*, such that there is a positive jump in the securitization probability at s^*. That is,

$$q_- \equiv \lim_{s \uparrow s^*} q(s) < q_+ \equiv \lim_{s \downarrow s^*} q(s).$$

One important implication from the existence of a rule of thumb in the market is that when gains from trade are not sufficiently large, 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 of loans right above the threshold compared to those right below.

Corollary 2. *Under the condition* $\frac{(\theta-1)c}{\gamma+\lambda(a^*)} + a^*\lambda'(a^*)\left(\frac{c}{(\gamma+\lambda(a^*))^2} - \frac{(\theta-1)ct_g(a^*)}{\gamma+\lambda(a^*)}\right) < 0$, *then a positive jump in the securitization probability at the threshold* s^* *leads to lower effort, a* *shorter time to sale, and lower loan quality:*

$$\begin{aligned} a^*_- &\equiv a^*(q_-) > a^*_+ \equiv a^*(q_+), \\ t^*_- &\equiv t^*_g(q_-) > t^*_+ \equiv t^*_g(q_+), \\ \lambda^*_- &\equiv \lambda\left(a^*_-\right) < \lambda\left(a^*_+\right) \equiv \lambda^*_+. \end{aligned}$$

Based on Corollary 2, there should be a discontinuity in processing time and loan quality around the securitization threshold, reminiscent of the findings in Keys et al. (2010). The lower processing time emerges endogenously in the model if there is a jump in the probability of securitization, and this is understood by both originators and investors (i.e., conditional on the existence of a threshold, there is no sense in which investors are "fooled" by the lower screening effort by the originators).

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

3 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 securitized loans originated between 2002 and 2007. These results serve the joint purpose of validating processing time as a measure of originator effort, as well as testing the main predictions of the model in Section 2.

3.1 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 2007.

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—the 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 these purposes because it only reports the year of the action date.

The CoreLogic LoanPerformance database provides loan performance information on whether a loan is current, delinquent, or in foreclosure for securitized residential mortgages.[5] 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 15 months of origination.[6]

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 details on the merge procedure). The merged data contain detailed information about borrower and loan characteristics. Specifically, we have information on borrower credit risk at 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. The merged data also includes information on loan characteristics, such as whether the loan rate is fixed or adjustable, the initial loan rate, the margin, and the first rate reset for adjustable-rate loans, as well as whether the loan has features like a prepayment penalty or a 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 macroeconomic variables such as local housing price appreciation, state-level unemployment rates, and local median household incomes 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 house price appreciation (HPA) from 36 months before loan origination using the house 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.

In Table 1, we report the summary statistics of our sample by origination year. The sample comprises about 8.5 million loans, including 2.9 million Alt-A loans and 5.7 million subprime loans. The number of loans increases during our sample period and peaks in 2005.

[5] 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.

[6] Alternatively, we have also considered 60 days past due or in foreclosure as the definition of default, and have also used 24 months as the horizon of delinquency, and obtain qualitatively similar results.

The average FICO score is 649, and the average CLTV ratio is close to 80% (reflecting the typical 20% down payment on first mortgages, Adelino, McCartney, and Schoar (2025)). Comparing Alt-A with subprime loans, we find that the FICO scores are lower, and the CLTV ratios are higher for subprime loans.

Turning to the delay of sale, the sample average time between origination and securitization is 14 weeks, similar to AGH. The sample average time to sale is generally longer for subprime loans, particularly after 2003. The sample average delinquency rate (defined as 60+ delinquency in the first 15 months after origination) is 6.7%, rising from 4.9% in 2002 to 13.5% in 2007. The average delinquency rate is significantly higher at 8.6% for subprime loans relative to Alt-A (2.9%).

Of particular interest to our paper, the average processing time is 29.2 days, 34.6 days for Alt-A loans, and 26.4 days for subprime 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 1 presents the distribution of processing time for the whole sample (Panel A) and for Alt-A and subprime loans separately (Panel B). The full distribution of processing time and time to sale is in Tables B.1 and B.2 of the Internet Appendix. Subprime loans are processed somewhat quicker than Alt-A (Panel B), and we observe a long right tail in the overall distribution and in both sub-samples well above eight weeks. In Figure A.1 of the Internet Appendix we show the residuals of processing time and time to sale after we include lender fixed effects. There is substantial variation left in these variables within lenders in our data.

3.2 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 (Corollary 2). Empirical evidence for such discontinuities in the data supports the model predictions and, more importantly, validates 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, particularly 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). By comparing loans on either side of the credit score threshold with otherwise nearly identical risk profiles, we can examine whether differential ease of origination led to changes in processing time, our measure of the screening behavior of lenders.

We apply a regression discontinuity design (RDD) (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:

$$PT_{i,t} = \alpha + \beta \times \mathbf{1}_{FICO \geq 620} + \gamma \times X_{i,t} + \delta_{i,t} + \epsilon_{i,t}, \quad (10)$$

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 and issuance year-quarter, state, and mortgage lender, labeled by $\delta_{i,t}$. The fixed effects for year-quarter of origination and issuance 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. Lender fixed effects mean that the results reflect only within-lender variation, which can help mitigate concerns that processing time would reflect differences in technological adiption between lenders. 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 the discontinuity.

Figure 2 presents the RDD plot of processing time, delay of sale, and loan delinquency without controlling for any covariates. We aggregate the variables of interest for each point in the FICO score, and generate the RDD plot for processing time, time to sale, and loan delinquency, for the sample of Alt-A and subprime low-doc loans.[7]

The patterns in Figure 2 show an upward jump in the fraction of loans in delinquency at

[7] The figure for the full-doc loans is in the Internet Appendix.

FICO score 620 for low documentation loans both for subprime and for Alt-A loans, consistent with Keys et al. (2010) and Keys et al. (2012). The differences in raw (unconditional) default rates are about 6 percentage points for Alt-A, and 1.6 percentage points for subprime (we will discuss conditional default differences in the context of Table 3 below).

Importantly, these patterns are largely mirrored by processing time. We find that processing time is higher by about three days for Alt-A loans below the 620 FICO threshold compared to loans above it for low-documentation loans. The pattern is similar but less pronounced for subprime loans, where processing time is approximately one day longer for loans below the threshold compared to those above it for low-documentation loans. These findings are consistent with our model's prediction and suggest that lenders engage in more intensive screening efforts for loans that are harder to securitize. The time to sale patterns also align with our framework. Loans below the 620 FICO threshold take longer to sell than those above it, consistent with investors' preference for loans above the credit cutoff.

Table 3 shows the point estimates for the regression discontinuity design results controlling for a large set of observable characteristics in the low documentation subsample (results for full-documentation loans are in Table B.3 in the Internet Appendix). We conduct our analysis separately for the Alt-A and subprime samples. In Panel A, our loan-level findings confirm a discontinuity at the 620 threshold. For Alt-A loans, processing time decreases by 2.4 days (t-statistic of -3.86) and 1.2 days (t-statistic of -2.03) without and with lender fixed effects, respectively. For subprime loans, processing time decreases by 0.5 days and 0.32 days, both also statistically significant. Panel B shows that there is also a jump in the time to sale of loans at the same threshold. For Alt-A loans, time to sale decreases by between 0.1 and 0.2 months, while for subprime loans, it decreases by about 0.01 to 0.02 months. Consistent with Figure 2, we find the magnitudes are larger for Alt-A than subprime for both processing time and time to sale.

The results in Panels A and B should be interpreted in combination with those on defaults. This confirms that we observe in our sample the same results as Keys et al. (2010) and Keys et al. (2012). In Panel C of Table 3, we report the results from estimating equation (10) with observed loan default as the dependent variable. The results confirm the existence of a positive jump in default intensity after accounting for all observables of about 1.4% for the low documentation loans immediately above the threshold of 620, for both the Alt-A and subprime samples.

The combined results in Panels A, B, and C 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.

3.3 Processing time, observable, and unobservable default risk

Next, we conduct a loan-level nonparametric analysis on the relationship between loan processing time and loan default. Loan default can depend on observable borrower and loan characteristics, as well as characteristics that are unobservable to buyers and the econometrician, but may be known by the seller, about the borrower's creditworthiness. Through their screening effort, lenders can learn about 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 effort, we expect it to be negatively correlated with the conditional default (or, put differently, to be positively associated with unobservable loan quality). As in AGH, we exploit the fact that, as econometricians, we can observe loan outcomes that were not available to market participants at the time loans were sold.

In order to identify whether unobservable quality is related to lender effort, we regress loan delinquency on processing time and control for observable loan and borrower characteristics, as well as fixed effects for origination and issuance year-quarter, state, and lender:

$$Default_{i,t} = \alpha + \beta \times ProcessingTime + \gamma \times X_{i,t} + \delta_{i,t} + \epsilon_{i,t}, \tag{11}$$

where variables are labeled as in the previous section for equation 10. We show a nonparametric 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 3, we plot the coefficients of processing time dummy variables using the loans with a processing time below one week as the omitted group. Given the extensive list of control variables, differences in default across loans are plausibly related to lender private information, as in AGH. 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 about 0.6 percentage points lower if they are processed in 4 weeks, and this effect bottoms out at about 1 percentage point for loans processed in 8 weeks or longer. As we discuss next, this is in stark contrast with the relationship between 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.[8] In the second step, we repeat the above regression (11), but replace realized default with predicted loan default probability as the dependent variable.

In Panel B of Figure 3, we plot the coefficients for predicted defaults (observable risk) on processing time dummy variables. In stark contrast to the results on excess defaults in Panel A (which measures unobservable risk), the processing time is significantly and positively correlated with the predicted default probability. The predicted default is approximately 0.15 percentage points higher for loans processed in 6 weeks compared 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.

3.4 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 (Corollar 1). Figure 4 presents the scatter plot between processing time and average delay of sale for each bin of processing time in the Alt-A and subprime samples, as well as subsamples divided by the level of documentation.

We observe a strong positive relation between processing time and delay of sale. This relationship is approximately linear for Alt-A full documentation loans, as well as for both subsamples of subprime loans; however, it is somewhat concave for the Alt-A low-documentation subsample.

Table 4 reports estimates from a regression of loan-level delay of sale on processing time, controlling for loan and borrower characteristics, local macroeconomic conditions, origination year-quarter, state, and issuance year-quarter fixed effects. The results are reported in columns (1) to (3) for the Alt-A sample and in columns (4) to (6) for the subprime sample.

We find a significant positive relationship between processing time and time to sale, consistent with the model prediction. Besides all the control variables and fixed effects

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

in columns (1) and (4), we add mortgage lender fixed effects in columns (2) and (5), and additionally mortgage lender by origination year-quarter fixed effects in columns (3) and (6) to account for any time-varying lender-level changes in screening technology or demand for loans. The coefficient estimate of processing time indicates that, for every additional month of processing time, delay of sale increases by 0.006 months for the Alt-A sample in column (2), and by 0.0084 months for the subprime loans in column (5). Further adding lender by origination year-quarter fixed effects in columns (3) and (6) yields estimates of 0.0051 for Alt-A loans and 0.0069 for subprime loans, suggesting that the model's predicted relationship is robust to accounting for time-varying lender characteristics. We note that the magnitudes of the estimates of the processing time are small because the included fixed effects greatly reduce the variations in time to sale, also leading to a high R^2. In Figure 4, the slopes of the curve between processing time and time to sale are much larger without the fixed effects included in the regression. We also find that loans with higher FICO scores, lower CLTV, lower loan rates, or low-doc loan features are sold more quickly.[9]

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 the delay by borrowers to close, rather than the lender's screening effort. One specific factor is that liquidity-constrained borrowers are more likely to close on a home purchase near the end of the month to reduce interest payments in closing costs or save on 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 (we experiment with alternative cutoff dates and find similar results). This directly addresses the issue discussed in (Bhutta and Ringo, 2021) regarding preferences for end-of-month closings that might be related to borrower constraints. Once we exclude these loans, we find that the relation between processing time and delay of sale is still significant both overall and in the Alt-A and subprime subsamples.[10]

3.5 Processing time, delay of sale, and mortgage default

We next test our model prediction on the relation between loan default, processing time, and delay of sale. In Corollary 1 of our model, unobservably better loans are associated with longer processing time, longer delay of sale, and lower default risk, controlling for observable

[9] We show all the coefficients for the controls in this regression in Table B.4 in the Internet Appendix.

[10] The results are reported in Table D.1 in the Internet Appendix.

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 one of the variables to fully absorb the other's explanatory power. 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 delinquency after origination, loan processing time, and delay of sale, controlling for loan and borrower characteristics, local housing price, and macro variables, as well as origination year-quarter, issuance year-quarter, and state fixed effects. We report the results for the Alt-A and subprime samples in Panels A and B, respectively.[11]

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 decreases by 0.18% (0.23%) on average when processing time (delay of sale) increases by one month in the Alt-A sample. The corresponding figures for subprime loans are 0.15% and 0.38%, all highly statistically significant. 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.

We repeat the regressions in columns (1) to (3), adding lender fixed effects in columns (4) to (6) and additionally lender by origination year-quarter fixed effects in columns (7) to (9). The specifications in columns (7)-(9) account for the possibility of time-varying technology adoption by lenders. Our findings are generally robust to the inclusion of these additional fixed effects across both the Alt-A and subprime samples. Comparing the estimates with and without lender fixed effects, we find that the estimates for the processing time variable are more sensitive than those for the delay of sale.

Following AGH, we address the potential selection effects of random delays, noting that loans that default before the sale are not included in our sample of securitized loans, which makes the loans sold later appear mechanically better than those sold immediately. Specifically, we drop all the loans that default within the first nine months of origination and

[11]The coefficients for all control variables are available in Table B.5.

repeat our analysis. The results are reported in Table 6. Our findings are qualitatively similar in this restricted sample, and the magnitudes of the estimates are moderately reduced. In our primary specifications, we use a default horizon of 15 months and define a mortgage as being in default if the borrower is two payments behind (i.e., 60+ days delinquent). In the Internet Appendix, we further show that our results remain robust across alternative specifications, including an alternative default definition of three payments behind (i.e., 90+ days delinquent, in Table C.1), and default horizons of 18 months (Table C.2) or 24 months (Table C.3). We also exclude loans that close after the 25th of each month to account for potentially constrained borrowers, as in the previous subsection (Table D.2 of the Internet Appendix).

In our final reduced-form empirical test, we separate the sample by loan documentation level. In Table 7 Panel A, we present the results of our analysis for the subsamples of low-doc and full-doc loans. Panel A shows that the coefficients for time to sale and processing time in the delinquency regressions are steeper for low-doc loans compared to full-doc loans across both the Alt-A and subprime samples. In contrast, the relationship between these two measures is quite similar for both low- and full-doc loans, as illustrated in Panel B. This indicates that each unit increase in time to sale or processing time reduces the expected delinquency rate more significantly for low-doc loans than for full-doc loans. In the next section, we calibrate the model, and from the lens of the calibrated model, we show in Figure 6 that the above empirical finding is consistent with the idea that low-doc loans exhibit poorer quality within the worst-performing segment.

3.6 Discussion of alternative explanations

Our finding of a positive relationship between mortgage processing time and the delay of sales supports the prediction of our model of dynamic signaling with endogenous screening effort. Nonetheless, since we cannot rely on random variation in either time to sale or processing time and instead trace out the equilibrium relationship between processing time and time to sale we observe in the data, it is possible that alternative explanations for this positive relationship exist outside the scope of our model. As we discuss below, any such explanation should reconcile not just the relationship between processing time and delay of sale, but also the evidence on observable risk (discussed in Section 3.3) and on conditional excess mortgage default, also addressed in Section 3.3.

The first alternative explanation is that some mortgages could emerge from more speedy origination and distribution channels, such as the "fast and easy" loan origination programs where mortgage loan buyers push for faster loan origination to meet the demand for MBS issuance. This could generate a positive relationship between processing time and time to sale, and the fast origination and issuance could also come at the expense of the loan screening effort. However, in order to speed up loan production, we would also expect mortgages with higher observable risks. We find the opposite: processing time is shorter for the loans with lower observable risks, as shown in Panel B of Figure 3. In addition, to the extent that this is understood by the buyers of the loans it will be reflected in prices, so this mechanism is isomorphic to our model and the mechanism we propose in the paper.

The second alternative explanation works through lender heterogeneity, specifically that lender size, technology, business model (for example, more efficient fintech lenders (Foote et al., 2019; Fuster et al., 2019)), and distribution channels could explain some of the patterns we uncover. If some lenders tend to originate low-quality loans and sell them quickly, and loan buyers understand this heterogeneity, we can still have our findings without the screening-signaling mechanism. Importantly though, lender fixed effects and lender-by-time fixed effects should account for most of this heterogeneity, making this explanation implausible.

A third alternative explanation, particularly regarding loans around the 620 credit score threshold, is that loans right above 620 benefit from additional demand. This increases gains from trade, thereby speeding up the sale and reducing the screening efforts. This aligns with the setting, featuring distinct mortgage distribution channels, such as the TBA and spec pools in the agency MBS markets (see, for example, Huh and Kim (2021)). Importantly, this mechanism would not apply beyond this specific threshold, and we conduct most of our analysis for the whole sample of loans. We find that our results hold in various subsamples, ranging from subprime loans with low documentation to Alt-A loans with full documentation, as shown in Table 7. This alleviates concerns that institutional variation on its own could explain the results.

4 Quantifying Market Fragility and Information Frictions

In this section, we calibrate our model based on the empirical findings in the previous section and use the calibrated model to quantify the loss of efficiency due to information frictions. To account for fundamental differences between Alt-A and subprime loans, we calibrate the model separately to these two market segments.

To better align the model with the data, we extend the baseline framework introduced in Section 2. Specifically, we incorporate the possibility of positive recovery in the event of default, assuming that the mortgage recovers a fraction $\alpha < 1$ of its original face value B.

Following AGH, we set the recovery rate $\alpha = 0.9$ to be consistent with the relatively high recovery rates and self-cure rates in the literature.

Next, we set the coupon payment such that $c = rB$, where r is the annualized percentage rate of the mortgage. Using the summary statistics reported in Table 2, we set $r = 6.22\%$ and $B = \$277,598$ for Alt-A loans, and 7.79% and $182,631, respectively, for subprime loans. For concreteness, we consider the following parametric form for the default intensity of the g-type:

$$\lambda(a) = \lambda_b + a^{\zeta}(\lambda_g - \lambda_b), \tag{12}$$

where $a \in (0, 1]$ and $\lambda_b > \lambda_g \geq 0$. The parameter $\zeta \geq 0$ governs the sensitivity of default intensity to screening effort: a higher ζ implies that small increases in effort yield larger reductions in default risk. Finally, we set the average expected default rate to be $\lambda_b = 0.0621$ as in AGH for both types of loans.

We estimate the remaining parameters by simulating the model and minimizing the difference between simulated data and actual data for a few key moments. Specifically, we assume that there is a continuum of "sub-markets." Each sub-market is characterized by two types of originators with default intensities λ_b and $\tilde{\lambda}_g$, the same as in the baseline model. For simplicity, we assume that λ_b is fixed and the same across all sub-markets, but $\tilde{\lambda}_g$ is drawn from the following beta distribution:

$$f(\lambda) = \frac{1}{(\lambda_b - \lambda_g)^{\alpha_1 + \alpha_2 - 1}} \frac{(\lambda - \lambda_g)^{\alpha_1 - 1}(\lambda_b - \lambda)^{\alpha_2 - 1}}{B(\alpha_1, \alpha_2)},$$

where $B(\alpha_1, \alpha_2)$ is the beta function. Following AGH, we set λ_g to match the unconditional average loan quality $E[\lambda] = 6.21\%$. That is, $\lambda_g = \frac{E[\lambda] - \frac{\alpha_1}{\alpha_1 + \alpha_2}\lambda_b}{1 - \frac{\alpha_1}{\alpha_1 + \alpha_2}}$. We also assume the following parametric forms for the cost function: $C(a) = \frac{1}{4}ka^4$, where the parameter $k = k_0 B$.

For the parameters $\Theta \equiv \{\gamma, \theta, \lambda_b, \lambda_g, \zeta, k_0, \alpha_1, \alpha_2\}$ to be calibrated, we simulate a sample of loans using 100,000 draws of $\tilde{\lambda}_g$. Each of such draws corresponds to a sub-market characterized by two types: λ_b and $\tilde{\lambda}_g$. In each sub-market, we solve for the optimal effort a^* as well as the optimal time to sale t_g^* and simulate a default time τ_d. Finally, we form a sample of g-type mortgages that are sold before their default times, i.e., $t_g^* \leq \tau_d$.

With this simulated sample of mortgages, we calculate the following moments and then estimate the parameters in Θ by minimizing the distance between the simulated moments and the actual moments in the data. The first moment is the regression coefficient β_{DEF_TS}, which is the coefficient of the time to sale TS in the delinquency regression estimated using the simulated data. That is, we regress a dummy variable that equals one if the simulated mortgage has defaulted within the first 15 months, on its time to sale TS. We denote this estimate by $\hat{\beta}_{DEF\_TS}$. Second, we calculate the cumulative distribution of time to sale and denote this distribution by $\hat{\Phi}_{TS}(t)$, for the sale month $t = 1, 2, \cdots, 9$. Third, we calculate the average proceeds to the originator, given by

$$\hat{L} = \frac{1}{N}\sum_{n=1}^{N} \frac{B}{\gamma + \lambda_n}((r + \lambda_n \alpha)(1 - D_n) + (\theta r + \lambda_n \alpha) D_n),$$

where $D_n \equiv exp(-(\gamma + \lambda_n)t_g^*)$ and t_g^* is given in (8). Fourth, we calculate the average time to sale, denoted by $\widehat{TS}$.

In addition to the moments above, we also consider moments related to processing time. One challenge with mapping the calibrated model to the data is that we observe processing time PT in the data, but only the optimal effort a^* in the simulated data. This requires us to create a mapping from the observed processing time PT into the effort a^* implied in the model. To do this, we start with the observed cumulative distribution of processing time in the data, denoted by $\Phi_{PT}(t)$ for $t = 1, 2, \cdots, 9$ weeks. We then use the simulated distribution of the effort a^* and find the cutoff values a_i^*, $i = 1, 2, \cdots, 9$ such that the corresponding percentiles for those cutoff values are equal to $\Phi_{PT}(t)$, $t = 1, 2, \cdots, 9$ weeks. This establishes a mapping between the cutoff values $a_1^*, \cdots, a_9^*$ in effort and the cutoff values $1, \cdots, 9$ weeks in processing time. We then run a simple linear regression of the cutoff values in processing time on those in effort. We denote the regression coefficients as $\hat{\eta}_0$ and $\hat{\eta}_1$ such that, $\widehat{PT} = \hat{\eta}_0 + \hat{\eta}_1 a$. Using the estimated regression coefficients, we can thus map out the simulated efforts to the implied processing time. We then compute the average processing time based on this simulated data, denoted by $\widehat{PT}$, and the cumulative distribution $\hat{\Phi}_{PT}(t)$ for $t = 1, 2, \cdots, 9$ weeks.

Lastly, we estimate the remaining parameters $\Theta \equiv \{\gamma, \theta, \lambda_b, \lambda_g, \alpha_1, \alpha_2, \zeta, k_0\}$ by minimizing the equally weighted sum of squared differences between the moments of the simulated data and those of the actual data. Specifically, we solve the following problem:

$$min_\Theta \left[\begin{array}{l} ((\hat{\beta}_{DEF\_TS} - \beta_{DEF\_TS})/\beta_{DEF\_TS})^2 + \sum_{t=1}^{9}((\hat{\Phi}_{TS}(t) - \Phi_{TS}(t))/\Phi_{TS}(t))^2 + ((\hat{L} - L)/L)^2 \\ + ((\widehat{TS} - TS)/TS)^2 + ((\widehat{PT} - PT)/PT)^2 + \sum_{t=1}^{9}((\hat{\Phi}_{PT}(t) - \Phi_{PT}(t))/\Phi_{PT}(t))^2 \end{array} \right].$$

4.1 Calibration results

Table 8 presents the calibration results and the model's quantitative implications. Panel A reports the calibrated parameter values, which indicate that, relative to Alt-A loans, subprime loans are characterized by less patient originators, lower investor demand, poorer quality among extremely bad loans, and a larger spread between the qualities of good and bad loans. Panel B of Table 8 and Figure 5 show that the simulated distributions closely match their empirical counterparts. Panels C through E then explore the quantitative implications of the model.

Market Fragility. The calibration sheds light on market fragility as characterized in Proposition 4. In particular, it allows us not only to infer the gains from trade parameter θ, but also to estimate the associated thresholds $\underline{\theta}$ and $\overline{\theta}$. In light of Proposition 4, these estimates reveal how fragile the Alt-A and subprime market segments are, and how far each segment lies from a market collapse.

Panel C reports the average gains from trade θ and the thresholds computed from the simulated data. For Alt-A loans, the calibrated θ (about 1.134) lies well above the upper threshold $\overline{\theta} = 1.017$, placing this segment firmly in the multiple-equilibria region and indicating greater fragility. By contrast, for subprime loans the calibrated θ (about 1.077) lies much closer to the upper threshold $\overline{\theta} = 1.027$, suggesting comparatively less fragility at the upper end.

At the lower end, however, the estimated $\underline{\theta}$ is similar—around 1.01 for both Alt-A and subprime loans. Given that the calibrated θ is substantially smaller for subprime than for Alt-A, this implies that a sharp decline in gains from trade is more likely to push the subprime segment below $\underline{\theta}$ and trigger a market freeze in subprime than in Alt-A.

Efficiency Loss due to Information Frictions. The calibrated model allows us to evaluate the economic magnitude of efficiency losses, for example, due to information or commitment frictions. First, considering the ex-post efficiency, we use the simulated data to calculate the average proceeds to the originator after exerting effort. From the results reported in Panel D of Table 8, we can see that for Alt-A loans information friction reduces the average proceeds from \$296,238 in the first-best to \$295,561. This difference of \$677.5, or about 23 basis points, is the cost of signaling. As a benchmark, we can consider the average mortgage rate at origination in our sample of 7.3%. During this time period, average 10-year Swap rates were 4.9%, which means that a simple back-of-the-envelope estimate of mortgage spread was 2.4%, or approximately 240 basis points. The cost of signaling represents about 10 percent of average mortgage spreads in this period. Such cost of signaling is similar at around \$514.9 for subprime loans, or about 27 basis points relative to the first-best surplus.

If we further take into account the cost due to the lack of commitment (when we compare the market equilibrium outcome with the second best), the average proceeds is further reduced to $295,394, by an amount of $166.3 in the market equilibrium in the case of Alt-A loans. Similar findings apply to subprime loans.

Next, we consider the losses in *ex ante* efficiency, which is the expected value to the originator *before* she exerts effort. Once we take into account the possibility that screening is unsuccessful, the losses in the *ex ante* efficiency is much smaller as shown in Panel E, about $110 ($105) due to the cost of signaling, and $6 ($11) due to the lack of commitment for Alt-A (subprime) loans.

Low vs. Full Documentation. The calibrated model also helps interpret the results for low-doc and full-doc loans in Table 7. In the data, the coefficients on time to sale and processing time in the delinquency regressions are larger in magnitude for low-doc loans than for full-doc loans (Panel A), even though the relationship between time to sale and processing time is similar across the two groups (Panel B).

Our calibration rationalizes this pattern if low-doc loans are of poorer quality within the worst-performing segment. This is natural if borrowers select low documentation to conceal higher risk. To capture this, we raise the default intensity of bad loans, λ_b, by 0.5 percentage points for low-doc loans ($\lambda_b = 6.92\%$ for Alt-A and 7.19% for subprime), while keeping the calibrated full-doc values ($\lambda_b = 6.42\%$ and 6.69%, respectively). As shown in Figure 6, this shift steepens the relationship between expected default intensity and time to sale (Panels A1–A2), and between expected default intensity and processing time (Panels B1–B2), but has little effect on the relationship between time to sale and processing time (Panels C1–C2). Thus, the model indicates that worse underlying quality in the low-doc segment can account for the stronger link between timing variables and delinquency, despite similar timing comovement, consistent with the evidence in Table 7.

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) and AGH to show that increased screening at origination is related to more signaling, particularly through delayed sales. The model also provides a sharp distinction between the role of observable probability of securitization and that of soft information acquired through originator effort that is unobservable to the investors.

We use U.S. private-label securitized mortgage data from 2002 to 2007 and employ mortgage processing time as a measure of screening effort. We have three main empirical results. First, 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. Second, processing time is positively associated with observable credit risk but negatively correlated with conditional ex-post mortgage default (unobservable from the perspective of buyers of mortgages at the time), consistent with key predictions in the models. 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.

Quantitatively, we calibrate the model by estimating parameters to match simulated moments with actual data, focusing on delinquency, time to sale, proceeds, and processing time. Our results show that subprime loans have less patient originators, lower investor demand, and greater loan quality dispersion, leading to efficiency losses due to information and commitment frictions. These frictions reduce proceeds by up to \$677.5 (23 basis points) for Alt-A loans and \$514.9 (27 basis points) for subprime loans, with additional losses when accounting for commitment constraints.

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., B. McCartney, and A. Schoar (2025): "The Missing LTV Cycle," *American Economic Review: Insights*.

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.

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.

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.

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., P. Gottardi, and H. Moreira (2025): "Time Trumps Quantity in the Market for Lemons," Working Paper.

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.

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

Holmström, B. (1979): "Moral Hazard and Observability," *The Bell Journal of Economics*, 10, 74–91.

Huh, Y. and Y. S. Kim (2021): "The Real Effects of Secondary Market Trading Structure: Evidence from the Mortgage Market," *The Review of Financial Studies*, 35, 3574–3616.

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.

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

Purnanandam, A. (2010): "Originate-to-distribute Model and the Subprime Mortgage Crisis," *The Review of Financial Studies*, 24, 1881–1915.

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.

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

Appendix A: Proofs

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

$$\max_{t_z} \mathbb{E}_a^z \left[\int_0^{t_z} ce^{-\gamma u} 1_{\tau_d \geq u} du + e^{-\gamma t_z} 1_{\tau_d \geq t_z} p_z \right],$$

where $p_z = \mathbb{E}_a^z \left[\int_t^\infty \theta c e^{-\gamma(u-t)} 1_{\tau_d \geq u} du \right]$ since competitive investors will price the loan at its expected value. Because γ, it is straightforward to show that the solution is a corner one, meaning $t_g^{FB} = t_b^{FB} = 0$, independent of her z-type and her initial choice of effort a. 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:

$$\begin{aligned} V_0(a, a) &= ap_g + (1-a) p_b - C(a) \\ &= a\frac{\theta c}{\gamma + \lambda(a)} + (1-a)\frac{\theta c}{\gamma + \lambda_b} - C(a). \end{aligned}$$

The first-order condition is given by

$$\theta \left(\frac{c}{\gamma + \lambda(a)} - \frac{c}{\gamma + \lambda_b} \right) - \frac{a\theta c \lambda'(a)}{(\gamma + \lambda(a))^2} = C'(a).$$

By Assumption 1, the first-order condition given in (4) characterizes the solution to the problem $\max_a V_0(a, a)$ both in the first-best and in the full information equilibrium. □

We first introduce the D1 refinement.

Definition 1. *Given (a, a^e), we define by $b_z(t, v)$ the belief necessary to provide the z-type utility v if the time to sale is t, that is, $u_z(b_z(t, v), t) = v$, and by $B_z(t, v) \equiv (b_z(t, v), 1]$, the set of beliefs for which the z-type obtains strictly higher utility than v when the time to sale is t.*

In our model, the D1 refinement can be stated as follows. Fix an equilibrium endowing expected payoffs $\{u_b, u_g\}$. Consider a time-to-sale choice of t that is not in support of either type's strategy. Given (a, a^e), if $B_b(t, u_b) \subset B_g(t, u_g)$, then $\mu = 1$. If $B_g(t, u_g) \subset B_b(t, u_b)$, then $\mu = 0$.

Lemma 1. *With D1-Refinements, pooling equilibria in the securitization market do not exist for $a^e > 0$.*

Proof of Lemma 1. Let $a^e > 0$ and $\lambda(a^e) < (\theta - 1)\gamma + \theta\lambda(a)$, and suppose that there is a pooling equilibrium in the securitization market where both types choose to sell at time $t \geq 0$. Thus $\mu(t) = a^e$ and $p(t) = \mu(t)\frac{\theta c}{\gamma + \lambda(a^e)} + (1 - \mu(t))\frac{\theta c}{\gamma + \lambda_b}$ is given in (2).

Consider deviation to t' where the IC of the b-type is binding:

$$\begin{aligned} & \frac{c}{\gamma+\lambda_b}\left(1-e^{-(\gamma+\lambda_b)t'}\right)+e^{-(\gamma+\lambda_b)t'}\frac{\theta c}{\gamma+\lambda\left(a^e\right)} \\ = & \frac{c}{\gamma+\lambda_b}\left(1-e^{-(\gamma+\lambda_b)t}\right)+e^{-(\gamma+\lambda_b)t}p\left(t\right), \end{aligned}$$

or

$$e^{-(\gamma+\lambda_b)t'}=e^{-(\gamma+\lambda_b)t}\frac{p\left(t\right)-\frac{c}{\gamma+\lambda_b}}{\frac{\theta c}{\gamma+\lambda(a^e)}-\frac{c}{\gamma+\lambda_b}}.$$

Note that $p\left(t\right)<\frac{\theta c}{\gamma+\lambda(a^e)}$ as long as $\mu\left(t\right)<1$, and hence $t'<\infty$ and strictly greater than t when $\mu\left(t\right)<1$. Thus, the b-type originator is indifferent between deviating to sell at time t' and being identified as a g-type or continuing to sell at time t using the pooling strategy.

Finally, if the b-type is indifferent, then the g-type is strictly better off when the deviation is assigned belief $\mu=1$. The payoff to the g-type is $\frac{c}{\gamma+\lambda(a)}\left(1-e^{-(\gamma+\lambda(a))t}\right)+e^{-(\gamma+\lambda(a))t}p\left(t\right)$ from not deviating, and $\frac{c}{\gamma+\lambda(a)}\left(1-e^{-(\gamma+\lambda(a))t'}\right)+e^{-(\gamma+\lambda(a))t'}\frac{\theta c}{\gamma+\lambda(a^e)}$ from deviating and selling at time t', and $\frac{c}{\gamma+\lambda(a)}$ from holding the loan without selling it at all.

The proof is as follows. First, suppose $\lambda\left(a^e\right)<\left(\theta-1\right)\gamma+\theta\lambda\left(a\right)$, we have $\frac{\theta c}{\gamma+\lambda(a^e)}-\frac{c}{\gamma+\lambda(a)}>0$. In this case, $\left(\frac{\theta c}{\gamma+\lambda(a^e)}-\frac{c}{\gamma+\lambda(a)}\right)\left(p\left(t\right)-\frac{c}{\gamma+\lambda_b}\right)\geq\left(\frac{\theta c}{\gamma+\lambda(a^e)}-\frac{c}{\gamma+\lambda_b}\right)\left(p\left(t\right)-\frac{c}{\gamma+\lambda(a)}\right)$ (the equality holds when $a=0$), because $\frac{c}{\gamma+\lambda_b}<p\left(t\right)<\frac{\theta c}{\gamma+\lambda(a^e)}$. Furthermore, $e^{-(\gamma+\lambda(a))(t'-t)}>e^{-(\gamma+\lambda_b)(t'-t)}$ (recall $t'>t$) and $e^{-(\gamma+\lambda_b)(t'-t)}=\frac{p(t)-\frac{c}{\gamma+\lambda_b}}{\frac{\theta c}{\gamma+\lambda(a^e)}-\frac{c}{\gamma+\lambda_b}}$ (from the binding IC condition of the b-type). Therefore, we have

$$e^{-(\gamma+\lambda(a))(t'-t)}>e^{-(\gamma+\lambda_b)(t'-t)}=\frac{p\left(t\right)-\frac{c}{\gamma+\lambda_b}}{\frac{\theta c}{\gamma+\lambda(a^e)}-\frac{c}{\gamma+\lambda_b}}>\frac{p\left(t\right)-\frac{c}{\gamma+\lambda(a)}}{\frac{\theta c}{\gamma+\lambda(a^e)}-\frac{c}{\gamma+\lambda(a)}},$$

or

$$\begin{aligned} & \frac{c}{\gamma+\lambda\left(a\right)}\left(1-e^{-(\gamma+\lambda(a))t'}\right)+e^{-(\gamma+\lambda(a))t'}\frac{\theta c}{\gamma+\lambda\left(a^e\right)} \\ > & \frac{c}{\gamma+\lambda\left(a\right)}\left(1-e^{-(\gamma+\lambda(a))t}\right)+e^{-(\gamma+\lambda(a))t}p\left(t\right). \end{aligned}$$

Thus, while the set of beliefs for which the b-type is strictly better off by deviating is empty, the g-type is strictly better off for $\mu\in(\overline{\mu},1]$. By D1-Refinements, this deviation is assigned belief $\mu=1$. Therefore, it is profitable for the g-type to deviate. Contradiction.

Second, suppose $\lambda\left(a^e\right)<\left(\theta-1\right)\gamma+\theta\lambda\left(a\right)$, we have $\frac{\theta c}{\gamma+\lambda(a^e)}-\frac{c}{\gamma+\lambda(a)}=0$. In this case, the value for the g-type from deviating and selling at t' is $\frac{\theta c}{\gamma+\lambda(a^e)}=\frac{c}{\gamma+\lambda(a)}$, which is strictly greater than the value from not deviating: $\frac{c}{\gamma+\lambda(a)}\left(1-e^{-(\gamma+\lambda(a))t}\right)+e^{-(\gamma+\lambda(a))t}p\left(t\right)$, because $p\left(t\right)<\frac{\theta c}{\gamma+\lambda(a^e)}$ as long as $\mu<1$. By D1-Refinements, this deviation is assigned belief $\mu=1$. Therefore, it is profitable for the g-type to deviate. Contradiction.

Third, suppose $\lambda\left(a^e\right)>\left(\theta-1\right)\gamma+\theta\lambda\left(a\right)$, we have $\frac{\theta c}{\gamma+\lambda(a^e)}-\frac{c}{\gamma+\lambda(a)}<0$. In this case, the value for the g-type from deviating and holding the loan without selling it is $\frac{c}{\gamma+\lambda(a)}$, which is strictly greater than the value from not deviating: $\frac{c}{\gamma+\lambda(a)}\left(1-e^{-(\gamma+\lambda(a))t}\right)+e^{-(\gamma+\lambda(a))t}p\left(t\right)$, because $p\left(t\right)\leq\frac{\theta c}{\gamma+\lambda(a^e)}<\frac{c}{\gamma+\lambda(a)}$. □

Lemma 2. *Let $\{a, a^e\}$ be given, and let $\{t_b, t_g\} \in [0,\infty]^2$ be the time to sale in a separating equilibrium in secondary markets. Then t_g is given by the solution to the problem*

$$\max_{t\in[0,\infty]} \frac{\theta c}{\gamma+\lambda(a^e)} e^{-(\gamma+\lambda(a))t} + \frac{c}{\gamma+\lambda(a)}\left(1-e^{-(\gamma+\lambda(a))t}\right),$$
$$s.t. \ \frac{\theta c}{\gamma+\lambda(a^e)} e^{-(\gamma+\lambda_b)t} + \frac{c}{\gamma+\lambda_b}\left(1-e^{-(\gamma+\lambda_b)t}\right) = u_b^*, \tag{13}$$

where $u_b^* \equiv \frac{\theta c}{\gamma+\lambda_b} e^{-(\gamma+\lambda_b)t_b} + \frac{c}{\gamma+\lambda_b}\left(1-e^{-(\gamma+\lambda_b)t_b}\right)$ is the value of the b-type originator in this equilibrium.

Proof of Lemma 2. We prove this lemma by contradiction. Suppose that t_g is not the solution to the above problem. Consider the deviation to some $t_g^\dagger$ that solves (13).

First, consider payoffs for the b-type. If the b-type deviates to $t_g^\dagger$, the induced payoff depends on market beliefs μ:

- If $\mu = 1$ (the market is certain the seller is g-type), then by the constraint in (13) the b-type obtains exactly u_b^*, i.e. he is indifferent.
- If $\mu \in [0,1)$, then the price reflects some probability of facing a b-type, and thus the b-type obtains strictly less than u_b^*.

Hence, the b-type never strictly benefits from deviating to $t_g^\dagger$.

Next, consider payoffs for the g-type. By construction, $t_g^\dagger$ maximizes the g-type's expected payoff subject to keeping the b-type at u_b^*. Furthermore, because t_g is assumed not to be the solution to the above problem, it follows that the g-type strictly prefers $t_g^\dagger$ for $\mu \in (\underline{\mu}, 1]$.

Lastly, under the D1 refinement, when a deviation is observed, beliefs must assign probability one to the type that can profit most from the deviation. Here:

- The b-type never profits from $t_g^\dagger$.
- The g-type strictly benefits from $t_g^\dagger$ for $\mu \in (\underline{\mu}, 1]$.

Therefore, the market must assign $\mu = 1$ (certainty that the deviator is the g-type). Therefore, the original equilibrium with t_g can not hold. Contradiction. It follows that in equilibrium t_g must solve problem (13). □

Proof of Proposition 2. To characterize the securitization market equilibrium, we distinguish between two cases: $a^e = 0$ and $a^e > 0$. In the first case with $a^e = 0$, from equation (2), we have $p(t) = \frac{\theta c}{\gamma+\lambda_b}$ because $\lambda(0) = \lambda_b$. Under the belief of zero effort, the loan is priced as if it were always of b-type. If gains from trade are sufficiently large—formally, $\lambda_b \le (\theta-1)\gamma + \theta\lambda(a)$—then the equilibrium is pooling: both b- and g-type loans are sold immediately. Otherwise, the equilibrium is separating: the originator sells b-type loans immediately but holds g-type loans indefinitely.

In the second case with $a^e > 0$, Lemma 1 in the appendix shows that pooling equilibria cannot survive D1-refinements; thus, the equilibrium must be separating. In this case,

$\mu t_b = 0$ and $p(t_b) = \frac{\theta c}{\gamma+\lambda_b}$, so the b-type originator earns full-information payoff and optimally sells immediately ($t_b = 0$). On the other hand, the g-type's optimal time to sale t_g^* has to solve the following problem (Lemma 2):

$$\max_{t\in[0,\infty]} \frac{c}{\gamma+\lambda(a)}\left(1-e^{-(\gamma+\lambda(a))t}\right)+e^{-(\gamma+\lambda(a))t}\frac{\theta c}{\gamma+\lambda(a^e)}, \tag{14}$$
$$s.t., \frac{c}{\gamma+\lambda_b}\left(1-e^{-(\gamma+\lambda_b)t}\right)+e^{-(\gamma+\lambda_b)t}\frac{\theta c}{\gamma+\lambda(a^e)} \le \frac{\theta c}{\gamma+\lambda_b}.$$

Let t_g^{IC} be given by the binding IC of the b-type originator:

$$e^{-(\gamma+\lambda_b)t_g^{IC}}\left(\frac{\theta c}{\gamma+\lambda(a^e)}-\frac{c}{\gamma+\lambda_b}\right)=\frac{(\theta-1)c}{\gamma+\lambda_b},$$

or

$$t_g^{IC} = -\frac{1}{\gamma+\lambda_b}\log\left(\frac{(\theta-1)(\gamma+\lambda(a^e))}{(\theta-1)\gamma+\theta\lambda_b-\lambda(a^e)}\right)>0. \tag{15}$$

Consider $t_g < t_g^{IC}$, then t_g violates the IC of the b-type. Therefore, there is pooling in the securitization market, which by Lemma 1 cannot be an equilibrium.

Consider $t_g > t_g^{IC}$.

(a) When $\lambda(a^e) < (\theta-1)\gamma+\theta\lambda(a)$, then we have $\frac{\theta c}{\gamma+\lambda(a^e)} > \frac{c}{\gamma+\lambda(a)} > \frac{c}{\gamma+\lambda_b}$. Then there is a profitable deviation to delay the sale at time t_g': $t_g > t_g' > t_g^{IC}$. To see this, note that from the IC of the b-type, the set of beliefs for which the b-type benefits from deviating to t_g' is empty, since the IC is slack for $t_g' > t_g^{IC}$ for $\mu \in [0,1]$. The g-type's extra payoff from deviating is given by $\left(e^{-(\gamma+\lambda(a))t_g'}-e^{-(\gamma+\lambda(a))t_g}\right)\left(\frac{\theta c}{\gamma+\lambda(a^e)}-\frac{c}{\gamma+\lambda(a)}\right)$ and this deviation is strictly profitable for $\mu = 1$ or close to 1 when $\frac{\theta c}{\gamma+\lambda(a^e)}-\frac{c}{\gamma+\lambda(a)} > 0$. As a result, the LCSE with $t_g = t_g^{IC}$ and $t_b = 0$ is the unique equilibrium in the securitization market.

(b) When $\lambda(a^e) = (\theta-1)\gamma+\theta\lambda(a)$, it is without loss of generality to assume that $t_g = t_g^{IC}$ since the g-type is indifferent between selling at any time $t_g > 0$.

(c) When $\lambda(a^e) > (\theta-1)\gamma+\theta\lambda(a)$, then the discount received in the market by the g-type is large enough that she prefers to hold the loan without selling it: $t_g = \infty$, while the IC of the b-type is slack when $t_g = \infty$: the b-type is strictly better off selling immediately. Therefore, the unique equilibrium in the securitization market has $t_b = 0$ and $t_g = \infty$. □

Proof of Proposition 3. In any equilibrium, $a = a^e$. Using the results from Proposition 2, the problem of the g-type originator at time 0 for market beliefs $a^e \in [0,1]$ can be written as

$$\max_{a\in[0,1]} a\left[\frac{c}{\gamma+\lambda(a)}\left(1-e^{-(\gamma+\lambda(a))t_g(a,a^e)}\right)+e^{-(\gamma+\lambda(a))t_g(a,a^e)}\frac{\theta c}{\gamma+\lambda(a^e)}\right]+\frac{(1-a)\theta c}{\gamma+\lambda_b}-C(a)$$
$$= \max_{a\in[0,1]} ae^{-(\gamma+\lambda(a))t_g(a^e)}\max\left\{\frac{\theta c}{\gamma+\lambda(a^e)}-\frac{c}{\gamma+\lambda(a)},0\right\}+a\frac{c}{\gamma+\lambda(a)}+\frac{(1-a)\theta c}{\gamma+\lambda_b}-C(a),$$

where $t_g(a^e) \equiv -\frac{1}{\gamma+\lambda_b}\log\left(\frac{(\theta-1)(\gamma+\lambda(a^e))}{(\theta-1)\gamma+\theta\lambda_b-\lambda(a^e)}\right)$ and $t_g(a,a^e)$ is given by (5) in Proposition 2. Note that the objective is differentiable with respect to a at $a=a^e$, since $\frac{\theta c}{\gamma+\lambda(a^e)} > \frac{c}{\gamma+\lambda(a)} = \frac{c}{\gamma+\lambda(a^e)}$. Thus, in any equilibrium, a^* has to satisfy the first-order condition

$$\begin{aligned}
& \frac{c}{\gamma+\lambda(a^*)} - \frac{\theta c}{\gamma+\lambda_b} + e^{-(\gamma+\lambda(a^*))t_g(a^*)}\frac{(\theta-1)c}{\gamma+\lambda(a^*)} \\
& + a^*\lambda'(a^*)\left[-\frac{c}{(\gamma+\lambda(a^*))^2}\left(1-e^{-(\gamma+\lambda(a^*))t_g(a^*)}\right) - \frac{(\theta-1)ct_g(a^*)}{\gamma+\lambda(a^*)}e^{-(\gamma+\lambda(a^*))t_g(a^*)}\right] \\
= \; & C'(a^*).
\end{aligned}$$

Replacing a^e by a and rearranging terms yields (7) and (8).

Note that $a=0$ and $t_g=0$ is one possible solution, because substituting $a=0$ into (8) yields $t_g=0$, and we can verify $a=0$ and $t_g=0$ satisfy (7).

□

Proof of Proposition 4. To establish the existence of a full-game equilibrium, we need to rule out double-deviations. We can first rule out downward deviations $a < a^{ME}$. This is because such deviations keep the game in LCSE since the secondary market prices are pinned down by beliefs a^e not the actual effort a. Therefore, they cannot improve the value function at $t=0$ for the originator and can be ruled out.

Thus only upward deviations $a > a^{ME}$ can matter. An upward deviation can switch the secondary market out of LCSE to the originate-to-hold channel by the g-type if $\lambda(a^e) > (\theta-1)\gamma+\theta\lambda(a)$; in that region, the unique outcome is $t_g=\infty$. Hence define the best such double-deviation as a^{NS}.

We consider three cases below. In the first case where $(\theta-1)\gamma+\theta\lambda(a^{NS}) \geq \lambda_b$, it follows that $(\theta-1)\gamma+\theta\lambda(a^{NS}) > \lambda(a^{ME})$. Starting from beliefs $a^e=0$ or $a^e=a^{ME}$, a deviation to a^{NS} satisfies $\lambda(a^e) \leq (\theta-1)\gamma+\theta\lambda(a)$, so the secondary market stays in LCSE and it is never optimal ex post to hold the loan without selling it. Given LCSE pricing depends on a^e (not on a), the optimal response to the LCSE strategies is, by construction, $a^*=0$ when $a^e=0$ and $a^*=a^{ME}$ when $a^e=a^{ME}$; hence the deviation is unprofitable. Both candidates are equilibria.

In the second case where $\lambda_b > (\theta-1)\gamma+\theta\lambda(a^{NS}) \geq \lambda(a^{ME})$, the same LCSE logic as in Case 1 rules out profitable deviations when $a^e=a^{ME}$, so $\{a^{ME}, t_g^{ME}\}$ is an equilibrium. From the corner $\{0,0\}$ (i.e., $a^e=0$), $\lambda_b > (\theta-1)\gamma+\theta\lambda(a^{NS})$ implies the g-type holds the loan indefinitely after deviating to a^{NS}, delivering strictly higher value because $V^{NS} \equiv \frac{a^{NS}c}{\gamma+\lambda(a^{NS})} + \frac{(1-a^{NS})\theta c}{\gamma+\lambda_b} - C(a^{NS}) \geq \frac{\theta c}{\gamma+\lambda_b}$. Thus $\{0,0\}$ is not an equilibrium; $\{a^{ME}, t_g^{ME}\}$ is unique.

In the last case where $(\theta-1)\gamma+\theta\lambda(a^{NS}) < \lambda(a^{ME})$, the deviation to $\{a^{NS},\infty\}$ is available from both candidates. The positive-effort candidate $\{a^{ME}, t_g^{ME}\}$ survives iff the value function under this candidate no less than the value function in the no-sale case V^{NS}. If not, no equilibrium exists.

□

Proof of Corollary 2. Similar to Lemma 2, we can show that the g-type's optimal time to sale t_g^* remains the same as in (8).

Similarly as Proposition 2, the problem of the g-type originator at time 0 for market beliefs $a^e \in [0,1]$ can be written as

$$\max_{a\in[0,1]} \left\{ \begin{array}{c} q(s)\left[a\left(\frac{c}{\gamma+\lambda(a)}\left(1-e^{-(\gamma+\lambda(a))t_g(a^e)}\right)+e^{-(\gamma+\lambda(a))t_g(a^e)}\frac{\theta c}{\gamma+\lambda(a^e)}\right)+(1-a)\frac{c}{\gamma+\lambda_b}\right] \\ +(1-q(s))\left[a\frac{c}{\gamma+\lambda(a)}+(1-a)\frac{\theta c}{\gamma+\lambda_b}\right]-C(a) \end{array} \right\}$$

$$= \max_{a\in[0,1]} q(s)\left[ae^{-(\gamma+\lambda(a))t_g^*(a^e)}\max\left\{\frac{\theta c}{\gamma+\lambda(a^e)}-\frac{c}{\gamma+\lambda(a)},0\right\}+a\frac{c}{\gamma+\lambda(a)}+(1-a)\frac{\theta c}{\gamma+\lambda_b}\right]$$
$$+(1-q(s))\left[a\frac{c}{\gamma+\lambda(a)}+(1-a)\frac{\theta c}{\gamma+\lambda_b}\right]-C(a),$$

Therefore, it is straightforward to show that the first-order condition is given by

$$\rho'(a^*)+q(s)e^{-(\gamma+\lambda(a^*))t_g(a^*)}\left[\frac{(\theta-1)c}{\gamma+\lambda(a^*)}+a^*\lambda'(a^*)\left(\frac{c}{(\gamma+\lambda(a^*))^2}-\frac{(\theta-1)ct_g(a^*)}{\gamma+\lambda(a^*)}\right)\right]=C'(a^*).$$

When the term in the brackets is negative, a positive jump in the securitization probability above the threshold s^* leads to a lower level of effort, and a higher default intensity. From (8), a higher default intensity implies a shorter time to sale, all else equal.

□

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;[12] (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.

[12]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.

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.

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	An indicator variable: 1 if the loan has an adjustable rate
Delinquency	An indicator variable: 1 if 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)
Low Doc	An indicator variable: 1 if the borrower's income and assets are not fully documented in the underwriting process and 0 otherwise.
FICO	The credit score of the borrower at origination. All models include both the continuous FICO variable and a set of indicator variables corresponding to 5 FICO intervals: FICO $<$ 580, 580 $\leq$ FICO $<$ 620, 620 $\leq$ FICO $<$ 660, 660 $\leq$ FICO $<$ 700, and FICO $\geq$ 700.
Initial Rate	Initial or original interest rate as of the loan's first payment date
Jumbo	An indicator variable: 1 if the loan amount at origination exceeds the conforming loan limit set by statute that limits the size of mortgages eligible to be insured by the GSEs and 0 otherwise.
Lien Type	Lien position (e.g., first lien)
Loan Amount	Loan origination amount
Purchase Loan	An indicator variable: 1 if the purpose of the loan is used to purchase property and 0 otherwise
Refinance (traditional)	An indicator variable: 1 if the loan is used to refinance previous mortgage debt without converting any equity into cash and 0 otherwise
Refinance (cashout)	An indicator variable: 1 if the loan is used to refinance previous mortgage debt with a portion of the equity converted to cash and 0 otherwise
Loan Type	Type of the loan (e.g., conventional)
LTV	Combined loan-to-value ratio (including first and second liens)
Balloon	Indicator variable: 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.
interest rate	
IOflag	An indicator variable: 1 if the loan has an interest-only feature.
Occupancy	An indicator variable for whether owner-occupied or not
Prepay Penalty	An indicator variable: 1 when the loan has a prepayment penalty and/or is an option ARM or negative amortization loan.
Property Type	Type of the property (i.e., single-family residence (SFR))
TS	The period between loan origination and MBS closing
ZIP Code	ZIP code of the property
Term	The maturity length of the mortgage in months

Table B2: Variables from the Confidential HMDA Database

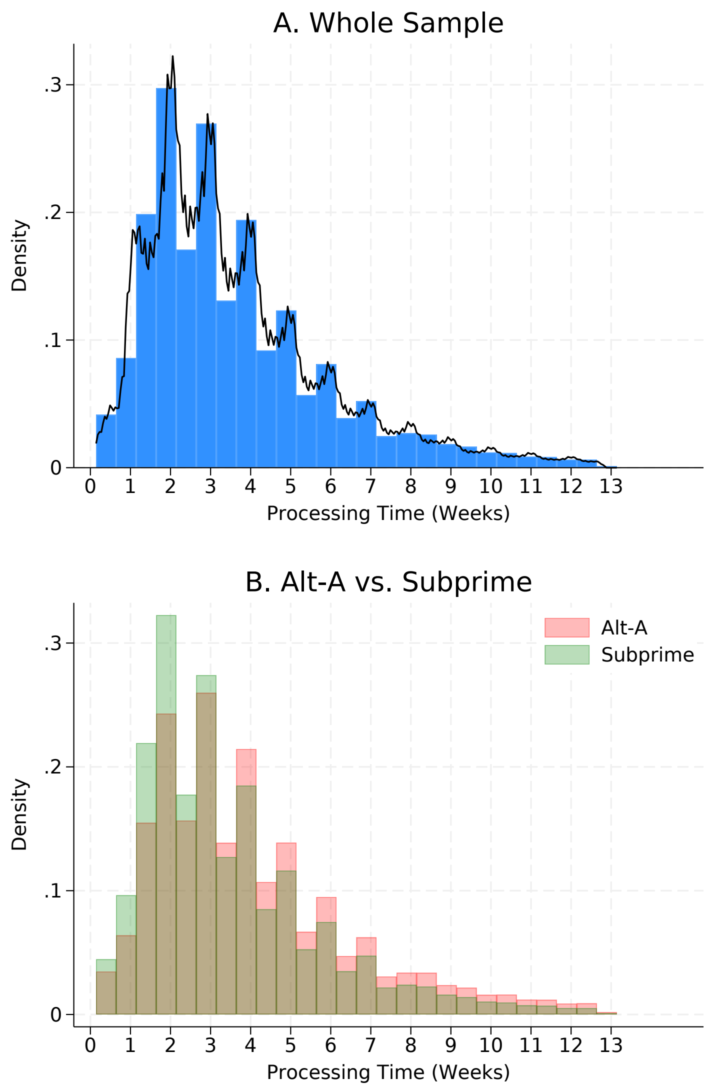
Variable List	Definition
Action Date	Date of action was taken on application
Applicant Race	Indicator variable for the race of the loan applicant (e.g., White)
Applicant Sex	Indicator variable to classify male or female
Applicant Income	Total gross annual income of applicant in thousands of dollars
Application 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 and county-level unemployment rate 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

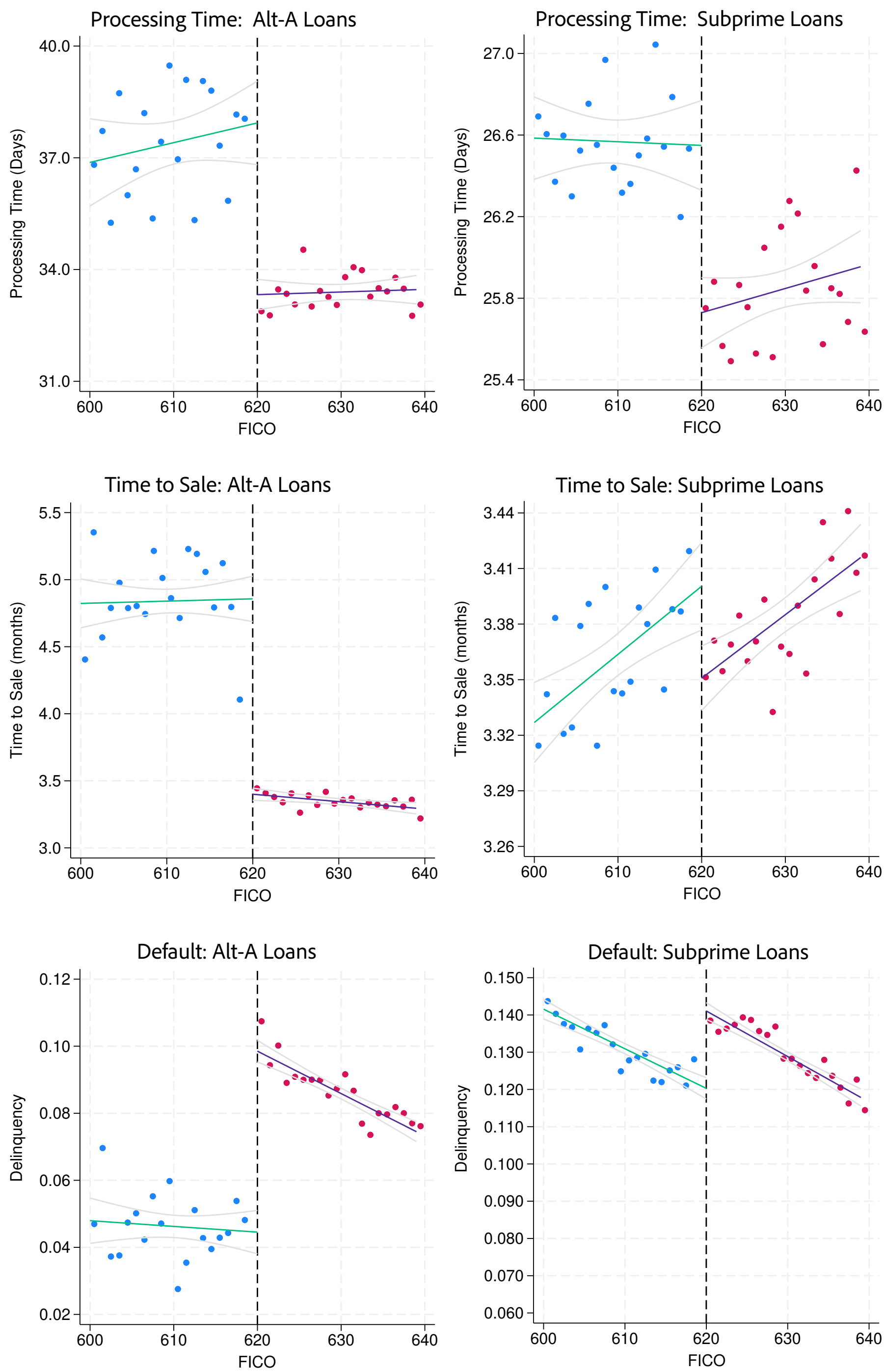
Variable List	Definition
HPA	The 36-month change in the housing price index for the borrower's county prior to loan origination
Unemployment	County-level unemployment rates from the Bureau of Labor Services (BLS). Both the unemployment rates in the county in the origination month and the 36-month cumulative growth in the rates from the month of mortgage origination are included in our analyses.

Figure 1: Histogram of Processing Time between 2002 and 2007



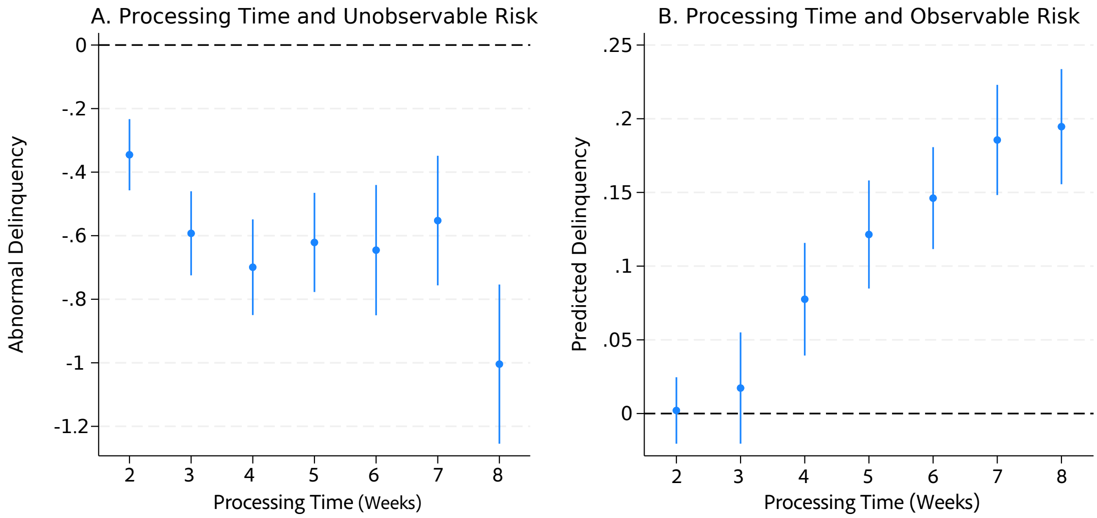
Note: This figure shows the histogram of mortgage processing time for the whole sample, 2002-2007, in Panel A and the Alt-A and subprime subsamples in Panel B. We add the scaled kernel density estimate of the density estimated with the Epanechnikov kernel and asymptotically optimal bandwidth in Panel A. The sample is the merged confidential HMDA and CoreLogic ABS database.

Figure 2: RDD Regression of Processing Time, Delay of Sale, and Delinquency (Low-Doc Loans)



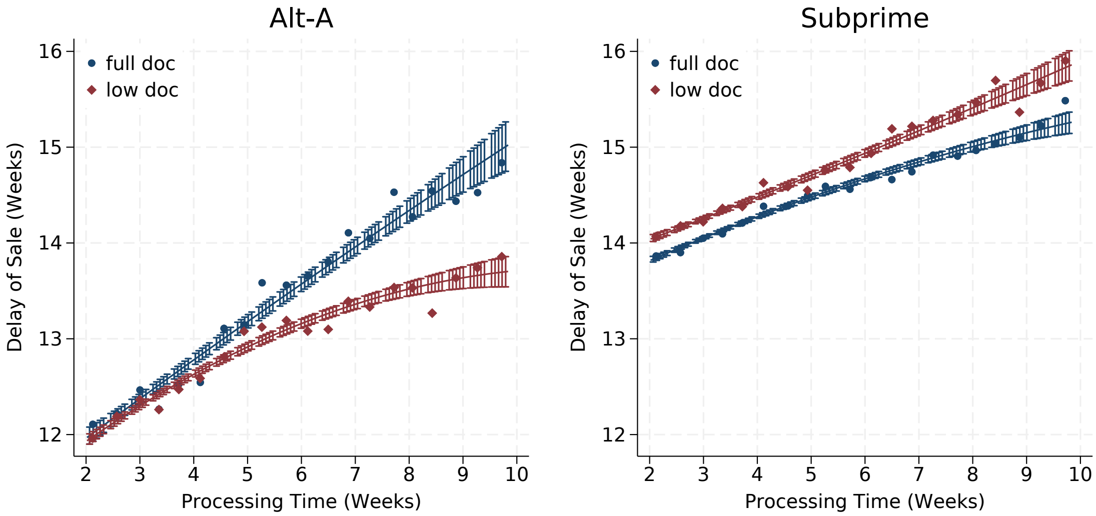
Note: This figure shows regression discontinuity plots of processing time, delay of sale, and delinquency for the merged confidential HMDA and CoreLogic ABS database for Alt-A and subprime low-doc loans. We compute average processing time, average time to sale, and average delinquency rate for each one-point FICO bin 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. Y-axis scale of the plots in the third row is in decimal points ("0.1" represents 10%).

Figure 3: Processing Time, Unobservable, and Observable Mortgage Default Risk



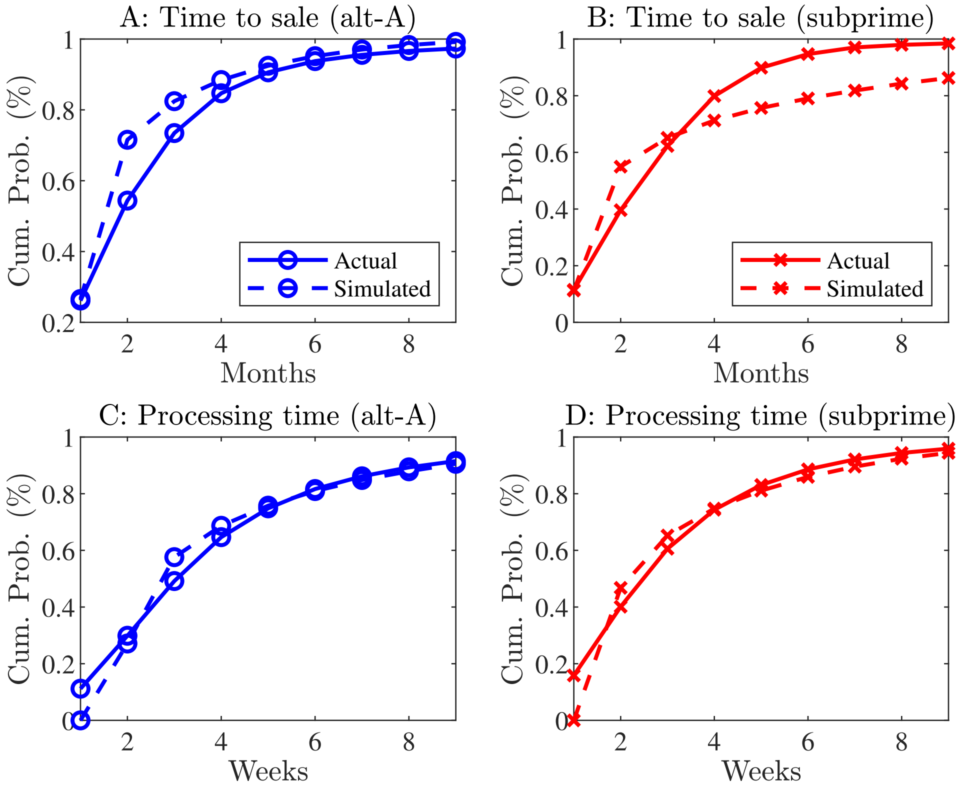
Note: This figure shows the coefficient estimates of processing time grouped into dummy variables from one week to 8+ weeks in a regression of loan delinquency (Panel A) or predicted default (Panel B), borrower and loan characteristics, origination year-quarter, state, and issuance year-quarter fixed effects. This figure shows point estimates and 95% confidence intervals. The predicted default probability in Panel B is estimated with a logistic model using borrower and loan characteristics with a 2-year rolling window. Y-axis scale is in percentage points, so that "0.1" represents a 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 4: Processing Time and Delay of Sale



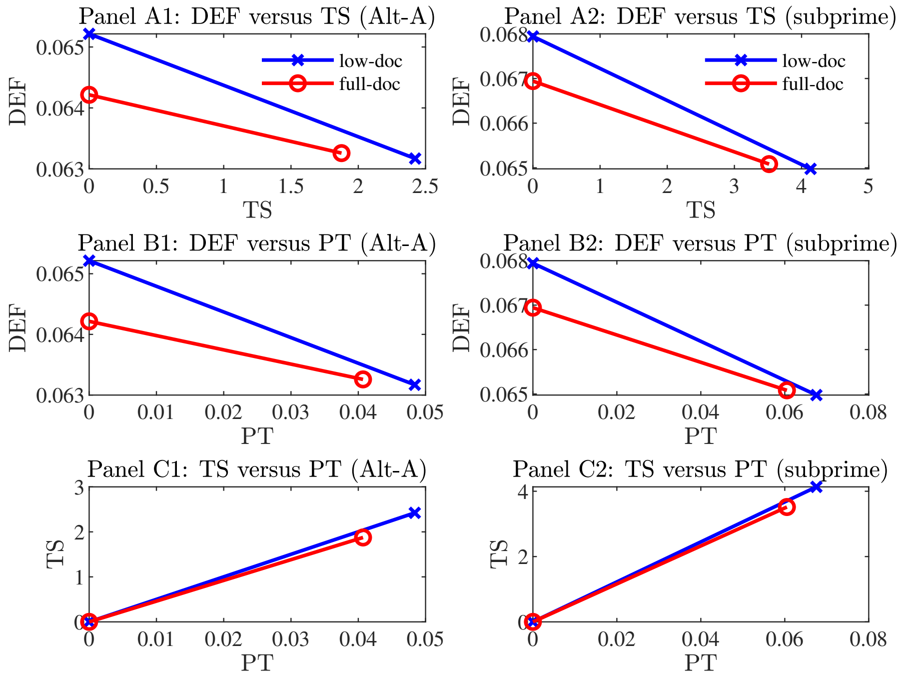
NOTE: This figure shows the scatter plot of processing time and delay of sale for the Alt-A loans (the left panel) and the subprime loans (the right panel), with the low-doc and full-doc loans separated in each panel. We average the delay of sale for each processing time bin and plot a quadratic fit, along with the 95% confidence interval. The sample is the merged confidential HMDA and CoreLogic ABS database.

Figure 5: Cumulative Distributions of TS and PT



Note: This figure shows simulated versus actual cumulative distributions of time to sale (in months) and processing time (in weeks) for both Alt-A and subprime loans.

Figure 6: Low-doc vs. Full-doc Loans



Note: This figure provides a numerical illustration regarding low- and full-doc loans for Alt-A and subprime loans. The parameters are set using the calibrated values reported in Table 8, except that λ_b is increased by 0.5% for low-doc loans.

Table 1: Summary Statistics (By Year)

Year	PT (day)	TS (wk)	FICO	LTV	Def. (%)	LD (%)	SFR (%)	Pur. (%)	Prim. (%)	N
2002	32.4	22.8	629.2	79.6	4.9	36.7	78.5	32.3	89.7	260,825
Alt-A	40.8	33.4	707.8	78.3	2.7	58.0	69.7	48.6	77.1	51,375
Subprime	30.3	20.2	609.9	80.0	5.5	31.4	80.7	28.3	92.7	209,450
2003	34.0	15.0	636.9	79.5	3.9	38.7	76.9	32.6	87.5	1,238,653
Alt-A	42.4	15.8	712.1	75.2	0.9	64.4	67.5	44.0	69.1	269,065
Subprime	31.6	14.7	616.0	80.7	4.7	31.6	79.5	29.5	92.6	969,588
2004	29.3	13.4	644.9	79.7	4.3	42.1	74.0	44.0	86.4	2,038,305
Alt-A	35.5	12.6	708.3	76.6	1.0	62.7	64.2	58.4	74.0	609,518
Subprime	26.6	13.7	617.9	81.1	5.7	33.3	78.1	37.8	91.7	1,428,787
2005	28.4	13.6	653.4	79.1	5.6	47.1	71.5	49.7	85.2	2,561,939
Alt-A	34.7	11.8	711.0	76.0	1.6	67.3	61.3	58.7	73.9	914,880
Subprime	24.9	14.6	621.4	80.8	7.9	35.9	77.2	44.7	91.5	1,647,059
2006	26.9	13.3	653.4	79.2	10.9	53.5	70.9	50.0	84.6	1,998,011
Alt-A	31.7	12.3	706.6	76.3	5.0	78.2	61.3	54.3	74.8	794,926
Subprime	23.7	14.0	618.3	81.1	14.9	37.2	77.3	47.1	91.2	1,203,085
2007	28.4	10.0	662.4	78.4	13.5	55.0	71.4	38.4	83.2	440,364
Alt-A	31.4	8.4	712.5	76.0	8.6	79.0	63.8	42.8	74.6	216,248
Subprime	25.4	11.5	614.2	80.7	18.3	31.9	78.8	34.2	91.4	224,116
Total	29.2	13.8	648.7	79.3	6.7	46.3	73.0	44.8	85.7	8,538,097
Alt-A	34.6	12.6	709.3	76.2	2.9	69.8	62.8	54.7	73.8	2,856,012
Subprime	26.4	14.3	618.2	80.9	8.6	34.4	78.0	39.9	91.7	5,682,085

Note: This table reports summary statistics for all privately securitized (PLS) loans and for loans backing Alt-A and subprime PLS in the merged confidential HMDA–CoreLogic ABS dataset. "PT" denotes processing time (days); "TS" denotes time to sale (weeks); "FICO" is the FICO score; "LTV" is the loan-to-value ratio; "Def." is the delinquency rate; "LD" indicates low documentation; "SFR" indicates single-family residence; "Pur." indicates purchase loans; "Prim." indicates primary residence; and "N" is the number of observations. The sample includes only first-lien mortgages originated between January 2002 and December 2007.

Table 2: Summary Statistics (By Type, All Years)

	All PLS		Alt-A		Subprime	
	Mean	Std. dev.	Mean	Std. dev.	Mean	Std. dev.
Term	356	37	356	43	356	33
Orig. rate	7.26	1.59	6.22	1.56	7.79	1.33
Orig. amt	214,398	169,731	277,598	227,124	182,631	119,766
LTV	79.3	12.6	76.2	12.8	80.9	12.2
FICO	648.7	71.0	709.3	48.8	618.2	60.1
Purchase (%)	44.8	49.7	54.7	49.8	39.9	49.0
Cash-out refi (%)	44.8	49.7	30.3	46.0	52.1	50.0
ARM (%)	47.8	50.0	26.9	44.3	58.4	49.3
Balloon (%)	5.1	22.1	0.8	8.8	7.3	26.1
Interest only (%)	20.6	40.4	38.8	48.7	11.4	31.8
Jumbo (%)	9.5	29.3	18.4	38.7	5.1	21.9
Low-doc (%)	46.3	49.9	69.8	45.9	34.4	47.5
Prepay pnlty (%)	59.4	49.1	35.9	48.0	71.3	45.3
Primary (%)	85.7	35.0	73.8	44.0	91.7	27.6
SFR (%)	73.0	44.4	62.8	48.3	78.0	41.4
Unemployment	5.11	1.56	4.83	1.49	5.25	1.58
36-mth unem. chg	1.39	2.91	1.97	3.05	1.11	2.80
36-mth HPA	4.64	26.98	-0.62	28.38	7.29	25.84
Delinq. (%)	6.68	24.97	2.91	16.80	8.58	28.01

Note: This table reports summary statistics for all privately securitized securities (PLS) and loans backing Alt-A and subprime PLS in the merged confidential HMDA and CoreLogic ABS dataset. The merged dataset includes only first-lien mortgages originated between January 2002 and December 2007.

Table 3: Loan-level Regression Discontinuity Around 620 FICO Threshold (Low-doc only)

	Alt-A		Subprime	
	(1)	(2)	(3)	(4)
Panel A: Processing Time				
$\mathbf{1}[FICO \geq 620]$	-2.40	-1.19	-0.52	-0.32
	(-3.86)	(-2.03)	(-3.50)	(-2.48)
Adjusted R^2	0.030	0.146	0.025	0.193
N	119,479	118,627	474,844	473,548
Panel B: Time to Sale				
$\mathbf{1}[FICO \geq 620]$	-0.14	-0.17	-0.02	-0.01
	(-4.48)	(-7.07)	(-3.67)	(-1.75)
Adjusted R^2	0.925	0.927	0.840	0.846
N	119,479	118,627	474,844	473,548
Panel C: Delinquency				
$\mathbf{1}[FICO \geq 620]$	0.0141	0.0108	0.0143	0.0142
	(3.71)	(2.85)	(7.90)	(7.78)
Adjusted R^2	0.067	0.071	0.089	0.093
N	119,479	118,627	474,844	473,548
Orig YQ FE	Y	Y	Y	Y
State FE	Y	Y	Y	Y
Issue YQ FE	Y	Y	Y	Y
Lender FE	N	Y	N	Y
Other cntrls	Y	Y	Y	Y

Note: This table reports the results of the loan-level regression of discontinuity based on the merged ABS and HMDA dataset for low-documentation loans with FICO between 600 and 640. Results for the full-documentation loans are in Table B.3 in the Internet Appendix. $\mathbf{1}[FICO \geq 620]$ 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 state and origination quarter, and t-statistics are reported in parentheses.

Table 4: Processing Time and Time to Sale

	Alt-A			Subprime		
	(1)	(2)	(3)	(4)	(5)	(6)
PT	0.0039	0.0060	0.0051	0.0240	0.0084	0.0069
	(2.49)	(4.12)	(3.63)	(10.60)	(2.91)	(2.60)
Adj. R^2	0.919	0.920	0.924	0.872	0.876	0.886
Obs.	2,842,335	2,840,865	2,828,260	5,650,633	5,649,131	5,631,026
Orig YQ FE	Y	Y	Y	Y	Y	Y
State FE	Y	Y	Y	Y	Y	Y
Issue YQ FE	Y	Y	Y	Y	Y	Y
Lender FE	N	Y	Y	N	Y	Y
Lender × Orig-YQ FE	N	N	Y	N	N	Y
Other cntrls	Y	Y	Y	Y	Y	Y

Note: This table reports the results of a loan-level regression of time to sale on processing time based on the merged ABS and HMDA dataset. Both time to sale and processing time are expressed in months in the regressions. The control variables are defined in Appendix B. Standard errors are clustered by state and origination quarter, and t-statistics are reported in parentheses.

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

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A: Alt-A									
PT	-0.0018		-0.0018	-0.0009		-0.0009	-0.0009		-0.0009
	(-11.69)		(-11.77)	(-5.83)		(-5.79)	(-6.35)		(-6.32)
TS		-0.0023	-0.0023		-0.0021	-0.0021		-0.0022	-0.0022
		(-11.67)	(-11.74)		(-11.42)	(-11.44)		(-11.65)	(-11.65)
Adj. R^2	0.055	0.055	0.055	0.059	0.06	0.06	0.066	0.066	0.066
Obs.	2,842,335	2,842,335	2,842,335	2,840,865	2,840,865	2,840,865	2,828,260	2,828,260	2,828,260
Panel B: Subprime									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
PT	-0.0015		-0.0015	-0.0006		-0.0006	-0.0006		-0.0006
	(-5.48)		(-5.21)	(-1.60)		(-1.50)	(-1.73)		(-1.66)
TS		-0.0038	-0.0038		-0.0035	-0.0035		-0.0033	-0.0033
		(-17.31)	(-17.32)		(-16.19)	(-16.21)		(-20.65)	(-20.67)
Adj. R^2	0.075	0.075	0.075	0.077	0.077	0.077	0.079	0.079	0.079
Obs.	5,650,633	5,650,633	5,650,633	5,649,131	5,649,131	5,649,131	5,631,026	5,631,026	5,631,026
Orig YQ FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
State FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
Issue YQ FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
Lender FE	N	N	N	Y	Y	Y	Y	Y	Y
Lender × Orig-YQ FE	N	N	N	N	N	N	Y	Y	Y
Other cntrls	Y	Y	Y	Y	Y	Y	Y	Y	Y

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 dataset. The dependent variable is loan delinquency within 15 months of loan origination. Both time to sale and processing time are expressed in months in the regressions. The control variables are defined in Appendix B. Standard errors are clustered by state and origination quarter, and t-statistics are reported in parentheses.

Table 6: Processing Time, Time to Sale, and Loan Default: Robustness

Panel A: Alt-A									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
PT	-0.0009		-0.0009	-0.0006		-0.0006	-0.0006		-0.0006
	(-11.23)		(-11.26)	(-5.89)		(-5.85)	(-6.19)		(-6.15)
TS		-0.0011	-0.0011		-0.0010	-0.0010		-0.0010	-0.0010
		(-9.05)	(-9.03)		(-8.60)	(-8.57)		(-8.83)	(-8.80)
Adj. R^2	0.036	0.036	0.036	0.039	0.039	0.039	0.042	0.042	0.042
Obs.	2,803,741	2,803,741	2,803,741	2,802,268	2,802,268	2,802,268	2,789,696	2,789,696	2,789,696
Panel B: Subprime									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
PT	-0.0009		-0.0009	-0.0006		-0.0006	-0.0006		-0.0006
	(-5.99)		(-5.78)	(-3.52)		(-3.46)	(-3.84)		(-3.80)
TS		-0.0015	-0.0015		-0.0013	-0.0013		-0.0012	-0.0012
		(-7.21)	(-7.11)		(-6.12)	(-6.10)		(-5.94)	(-5.91)
Adj. R^2	0.043	0.043	0.043	0.044	0.044	0.044	0.045	0.045	0.045
Obs.	5,375,868	5,375,868	5,375,868	5,374,350	5,374,350	5,374,350	5,356,496	5,356,496	5,356,496
Orig YQ FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
State FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
Issue YQ FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
Lender FE	N	N	N	Y	Y	Y	Y	Y	Y
Lender × Orig-YQ FE	N	N	N	N	N	N	Y	Y	Y
Other cntrls	Y	Y	Y	Y	Y	Y	Y	Y	Y

Note: This table reports the results from robustness analyses of a loan-level regression of loan default on time to sale and processing time based on the merged ABS and HMDA dataset. In the restricted sample analysis, we exclude loans that defaulted within the first 9 months of loan origination. The dependent variable is loan delinquency within the first 15 months of loan origination. Both time to sale and processing time are expressed in months in the regressions. In the analysis with alternative default definitions, the dependent variable is loan delinquency within the first 18 and 24 months of loan origination for the sample of all PLS loans. The control variables are defined in Appendix B. Standard errors are clustered by state and origination quarter, and t-statistics are reported in parentheses.

Table 7: Documentation Results

Panel A: Loan Default	Alt-A		Subprime	
	Low Doc	Full Doc	Low Doc	Full Doc
PT	-0.0009	-0.0003	-0.0012	-0.0000
	(-5.17)	(-2.58)	(-2.09)	(-0.34)
TS	-0.0027	-0.0010	-0.0042	-0.0032
	(-11.55)	(-6.95)	(-10.68)	(-15.59)
Adj. R^2	0.065	0.034	0.096	0.068
Obs.	1,983,761	855,822	1,946,642	3,700,841
Panel B: Time to sale and processing time	Alt-A		Subprime	
	Low Doc	Full Doc	Low Doc	Full Doc
PT	0.0048	0.0066	0.0078	0.0090
	(2.99)	(5.01)	(2.27)	(3.17)
Adj. R^2	0.918	0.928	0.870	0.881
Obs.	1,983,761	855,822	1,946,642	3,700,841
Orig YQ FE	Y	Y	Y	Y
State FE	Y	Y	Y	Y
Issue YQ FE	Y	Y	Y	Y
Lender FE	Y	Y	Y	Y
Other cntrls	Y	Y	Y	Y

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 dataset. The dependent variable is loan delinquency within 15 months of loan origination. Both time to sale and processing time are expressed in months in the regressions. The control variables are defined in Appendix B. Standard errors are clustered by state and origination quarter, and t-statistics are reported in parentheses.

TABLE 8: Calibration Results

Panel A: Parameter Values

Alt-A Preset		Alt-A Calibrated		Subprime Preset		Subprime Calibrated	
B	\$277,598	γ	5.64%	B	\$182,631	γ	7.27%
r	6.22%	θ	1.134	r	7.79%	θ	1.077
α	0.9	λ_b	6.42%	α	0.9	λ_b	6.69%
$E[\lambda]$	6.21%	λ_g	4.79%	$E[\lambda]$	6.21%	λ_g	3.32%
α_1	1.91	ζ	0.51	α_1	1.78	ζ	0.51
α_2	0.28	k_0	0.55	α_2	0.29	k_0	0.48

Panel B: Moments

	Alt-A Simulated	Alt-A Actual		Subprime Simulated	Subprime Actual
β_{DEF_TS}	-0.0006	-0.0021	β_{DEF_TS}	-0.0008	-0.0035
Orig. Amount	\$295,395	\$277,598	Orig. Amount	\$188,594	\$182,631
Avg. TS (mon)	1.48	3.15	Avg. TS (mon)	3.59	3.57
Avg. PT (week)	3.40	4.94	Avg. PT (week)	2.60	3.77

Panel C: Market Fragility

	Alt-A $\underline{\theta}$	Alt-A $\overline{\theta}$	Alt-A θ	Subprime $\underline{\theta}$	Subprime $\overline{\theta}$	Subprime θ
Estimate	1.008	1.017	1.134	1.010	1.027	1.077

Panel D: Loss in *ex post* efficiency

	Alt-A FB	Alt-A SB	Alt-A ME	Subprime FB	Subprime SB	Subprime ME
Ex post surplus	\$296,238	\$295,561	\$295,394	\$189,254	\$188,739	\$188,595
Loss		\$677.5	\$166.3		\$514.9	\$144.2

Panel E: Loss in *ex ante* efficiency

	Alt-A FB	Alt-A SB	Alt-A ME	Subprime FB	Subprime SB	Subprime ME
Ex ante Surplus	\$295,499	\$295,389	\$295,382	\$188,669	\$188,564	\$188,553
Loss		\$110.5	\$6.1		\$104.7	\$10.8

NOTE: This table presents the results of our quantitative analysis of the loss in efficiency. Panel A presents our parameter estimates for both preset and calibrated parameters. Panel B shows moments in the simulated and actual data. Panel C reports the estimates of the gains from trade θ as well as the thresholds $\underline{\theta}$ and $\overline{\theta}$ for both Alt-A and subprime loans. Panel D presents ex post surplus conditional on the optimal effort has been exerted. The columns "FB", "SB", and "ME" correspond to the cases of first-best, second-best, and market equilibrium. Panel E presents ex ante surplus, which is V_0 in (6), before choosing the effort. The loss in Column "SB" in Panels D and E reports the loss in (ex post or ex ante) surplus in the second-best, compared to the first-best, which estimates the loss in efficiency due to the cost of signaling (information friction). The loss in Column "ME" in Panels D and E reports the loss in surplus in the market equilibrium, compared to the second-best, which estimates the loss in efficiency due to the lack of commitment.

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

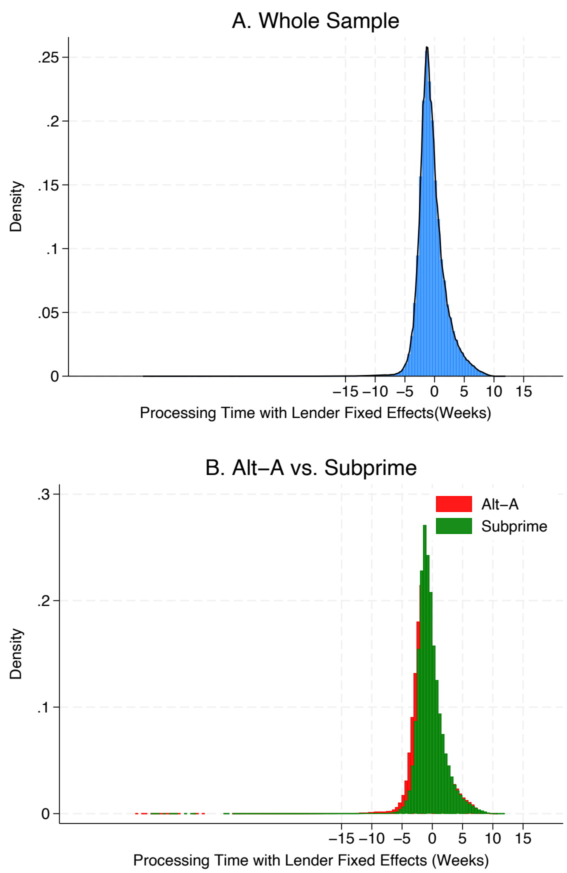
Manuel Adelino, Bin Wei, Feng Zhao

Contents of Internet Appendix

A Additional Figures

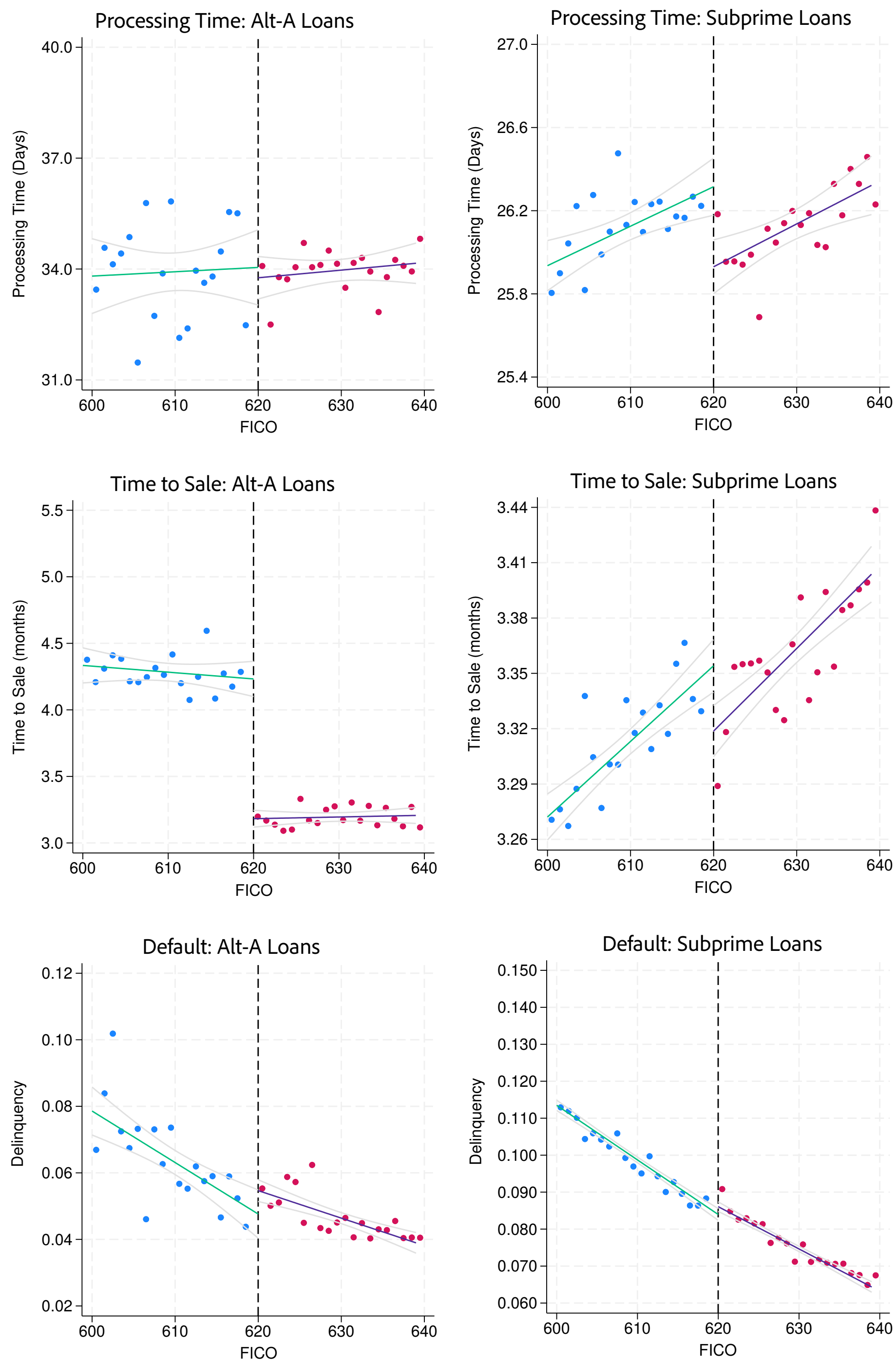
This section provides the additional figures referenced in the main text. Figure A.1 shows the histogram of the residuals from the regression of mortgage processing time on lender and origination year fixed effects for the whole sample, 2002-2007. Figure A.2 presents the RDD plots for the sample of full-doc loans.

Figure A.1: Histogram of Processing Time Residuals Between 2002 and 2007



Note: This figure shows the histogram of the residuals from the regression of mortgage processing time on lender fixed effects for the whole sample, 2002-2007. The sample is the merged confidential HMDA and CoreLogic ABS database.

Figure A.2: RDD Regression of Processing Time, Delay of Sale, and Delinquency (Full-Doc Loans)



Note: This figure shows regression discontinuity plots of processing time, delay of sale, and delinquency for the merged confidential HMDA and CoreLogic ABS database for Alt-A and subprime full-doc loans. We compute average processing time, average time to sale, and average delinquency rate for each one-point FICO bin 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. Y-axis scale of the plots in the third row is in decimal points ("0.1" represents 10%).

B Summary Statistics and Coefficient Estimates of Control Variables

In this section, we provide distribution of processing time and time to sale, additional robustness tests, and show the coefficient estimates for the control variables for baseline regressions in the main manuscript. The distribution of time to sale and processing time are reported in Table B.1 and Table B.2, respectively.

TABLE B.1: Distribution of time to sale in the CoreLogic PLS sample

	Whole Sample			Alt-A			Subprime		
TS	Freq.	Percent	Cum.	Freq.	Percent	Cum.	Freq.	Percent	Cum.
0/1	1,400,164	16.4	16.4	747,154	26.16	26.16	653,010	11.49	11.49
2	2,403,705	28.15	44.55	805,845	28.22	54.38	1,597,860	28.12	39.61
3	1,831,218	21.45	66	544,780	19.07	73.45	1,286,438	22.64	62.25
4	1,322,791	15.49	81.49	321,076	11.24	84.69	1,001,715	17.63	79.88
5	735,893	8.62	90.11	168,789	5.91	90.6	567,104	9.98	89.86
6	364,473	4.27	94.38	90,394	3.17	93.77	274,079	4.82	94.69
7	183,533	2.15	96.53	50,231	1.76	95.53	133,302	2.35	97.03
8	83,505	0.98	97.51	31,358	1.1	96.63	52,147	0.92	97.95
9	47,979	0.56	98.07	19,986	0.7	97.32	27,993	0.49	98.44
≥10	164836	1.90	100	76,399	2.66	100	88437	1.5	100
Total	8,538,097	100		2,856,012	100		5,682,085	100	

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

TABLE B.2: Distribution of processing time in the CoreLogic PLS sample

	Whole Sample			Alt-A			Subprime		
PT	Freq.	Percent	Cum.	Freq.	Percent	Cum.	Freq.	Percent	Cum.
0/1	1,219,986	14.29	14.29	319,431	11.18	11.18	900,555	15.85	15.85
2	1,911,931	22.39	36.68	533,246	18.67	29.86	1,378,685	24.26	40.11
3	1,713,934	20.07	56.76	551,543	19.31	49.17	1,162,391	20.46	60.57
4	1,225,327	14.35	71.11	442,671	15.5	64.67	782,656	13.77	74.34
5	787,016	9.22	80.32	290,654	10.18	74.84	496,362	8.74	83.08
6	507,429	5.94	86.27	194,284	6.8	81.65	313,145	5.51	88.59
7	327,638	3.84	90.11	128,419	4.5	86.14	199,219	3.51	92.1
8	220,209	2.58	92.68	90,120	3.16	89.3	130,089	2.29	94.39
9	148,159	1.74	94.42	62,935	2.2	91.5	85,224	1.5	95.89
≥10	476468	5.58	100	242,709	8.5	100	76,399	4.11	100
Total	8,538,097	100		2,856,012	100		5,682,085	100	

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

Table B.3 reports the results from the loan-level regression discontinuity around 620 FICO Threshold for full-doc loans. The coefficient estimates for Table 4 in the main draft are reported in Table B.4. The coefficient estimates for Table 5A in the main draft are reported in Table B.5.

Table B.3: Loan-level Regression Discontinuity Around 620 FICO Threshold (Full-doc only)

	Alt-A		Subprime	
	(1)	(2)	(3)	(4)
Panel A: Processing Time				
$\mathbf{1}[FICO \geq 620]$	-0.75	0.67	-0.25	0.07
	(-1.13)	(1.06)	(-2.84)	(0.84)
Adjusted R^2	0.038	0.187	0.025	0.180
N	69,942	69,063	1,011,830	1,010,280
Panel B: Time to Sale				
$\mathbf{1}[FICO \geq 620]$	-0.13	-0.12	-0.01	-0.01
	(-5.22)	(-5.49)	(-2.69)	(-2.44)
Adjusted R^2	0.920	0.921	0.840	0.845
N	69,942	69,063	1,011,830	1,010,280
Panel C: Delinquency				
$\mathbf{1}[FICO \geq 620]$	-0.0001	-0.0025	0.0027	0.0026
	(-0.02)	(-0.62)	(2.36)	(2.36)
Adjusted R^2	0.038	0.043	0.057	0.059
N	69,942	69,063	1,011,830	1,010,280
Orig YQ FE	Y	Y	Y	Y
State FE	Y	Y	Y	Y
Issue YQ FE	Y	Y	Y	Y
Lender FE	N	Y	N	Y
Other cntrls	Y	Y	Y	Y

Note: This table reports the results of the loan-level regression of discontinuity based on the merged ABS and HMDA dataset for full-documentation loans with FICO between 600 and 640. Results for the full documentation are in Figure 2. $\mathbf{1}[FICO \geq 620]$ 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 state and origination quarter, and t-statistics are reported in parentheses.

Table B.4: Processing Time and Time to Sale

	All PLS			Alt-A			Subprime		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
PT	0.0093	0.0054	0.0039	0.0039	0.006	0.0051	0.024	0.0084	0.0069
	(4.53)	(2.65)	(1.98)	(2.49)	(4.12)	(3.63)	(10.60)	(2.91)	(2.60)
Term	-0.0008	-0.0006	-0.0006	-0.0009	-0.0009	-0.001	-0.0008	-0.0004	-0.0003
	(-16.54)	(-17.38)	(-19.91)	(-12.57)	(-15.97)	(-19.13)	(-10.08)	(-7.16)	(-7.79)
Initial Rate	0.0248	0.0324	0.0352	0.0626	0.0562	0.0587	-0.0201	-0.0136	-0.0105
	(5.90)	(7.71)	(9.62)	(17.84)	(15.43)	(17.13)	(-3.48)	(-2.53)	(-2.09)
Loan Amount	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000
	(-3.97)	(-3.78)	(-2.51)	(-0.46)	(4.23)	(2.79)	(-9.34)	(-7.29)	(-6.06)
LTV	0.0004	0.0004	0.0005	0.0004	0.0002	0.0002	0.0004	0.0004	0.0005
	(3.55)	(3.28)	(4.06)	(2.19)	(1.24)	(1.21)	(3.29)	(3.38)	(4.53)
FICO	-0.0000	-0.0001	-0.0001	0.0003	0.0002	0.0002	-0.0004	-0.0006	-0.0005
	(-0.50)	(-2.11)	(-1.86)	(6.57)	(5.77)	(5.13)	(-4.87)	(-6.68)	(-6.23)
Purchase	-0.0325	-0.0381	-0.0379	-0.0695	-0.0629	-0.0576	-0.0133	-0.0266	-0.0226
	(-9.57)	(-11.86)	(-11.05)	(-13.96)	(-13.38)	(-12.45)	(-2.61)	(-5.99)	(-5.74)
Refi_cashout	-0.0298	-0.0337	-0.0331	-0.0231	-0.043	-0.042	-0.0391	-0.046	-0.0424
	(-10.62)	(-12.69)	(-12.14)	(-4.32)	(-10.74)	(-10.03)	(-11.88)	(-14.58)	(-12.78)
ARM	-0.0378	-0.0153	-0.0131	0.0593	0.0653	0.0636	-0.0719	-0.0515	-0.0535
	(-7.24)	(-3.25)	(-3.74)	(5.76)	(7.16)	(7.25)	(-15.92)	(-13.31)	(-20.07)
Balloon	-0.0384	-0.0269	-0.0451	-0.1339	-0.1268	-0.0901	-0.1119	-0.0935	-0.0935
	(-4.04)	(-2.83)	(-5.84)	(-3.89)	(-4.28)	(-3.25)	(-15.42)	(-9.54)	(-13.43)
IOflag	-0.0354	-0.0307	-0.0309	0.0306	0.0286	0.0268	-0.049	-0.0638	-0.0714
	(-7.67)	(-9.47)	(-8.61)	(3.86)	(4.74)	(4.05)	(-7.01)	(-9.26)	(-14.08)
Jumbo	-0.0139	-0.0107	-0.0083	-0.0015	0.0092	0.012	0.014	0.0208	0.0197
	(-1.75)	(-1.42)	(-1.37)	(-0.20)	(1.21)	(1.99)	(1.61)	(2.63)	(2.93)
Lowdoc	-0.0413	-0.0461	-0.0414	-0.0593	-0.0539	-0.0533	0.007	0.0072	0.0084
	(-9.57)	(-12.81)	(-11.06)	(-12.52)	(-11.38)	(-12.68)	(1.40)	(1.64)	(1.88)
Prepay_Penalty	0.0172	0.0072	0.0166	0.0194	-0.0041	0.0056	-0.0312	-0.0513	-0.0492
	(2.61)	(1.09)	(2.99)	(2.88)	(-0.73)	(1.10)	(-2.35)	(-4.39)	(-4.50)
Primary_Occupancy	0.0609	0.038	0.0349	0.0531	0.0312	0.0286	0.0274	-0.0165	-0.0192
	(20.80)	(12.38)	(15.51)	(14.07)	(8.07)	(9.12)	(6.15)	(-3.68)	(-4.85)
SFR	0.0067	0.0056	0.0042	0.0101	0.0039	0.0064	-0.0071	-0.0049	-0.0054
	(3.82)	(3.41)	(2.72)	(3.79)	(1.73)	(3.06)	(-3.03)	(-2.37)	(-3.02)
Unemp. Rate	-0.0021	-0.0005	0.0001	0.0024	0.0012	0.0007	-0.0047	-0.0015	-0.0004
	(-2.45)	(-0.57)	(0.16)	(1.96)	(0.97)	(0.73)	(-4.33)	(-1.60)	(-0.47)
Unemp. Rate Chg.	-0.001	-0.0018	-0.0028	0	-0.0013	-0.0037	-0.0014	-0.0008	-0.0005
(36m)	(-0.82)	(-1.53)	(-3.12)	(-0.01)	(-0.95)	(-2.58)	(-0.85)	(-0.61)	(-0.51)
hpi36	-0.0004	-0.0003	-0.0003	0.0003	0.0002	0.0002	-0.0009	-0.0008	-0.0006
	(-2.31)	(-2.03)	(-4.31)	(1.55)	(0.97)	(1.99)	(-6.17)	(-5.11)	(-7.07)
FICO $<$580	-0.0008	-0.0148	-0.0147	0.2974	0.2416	0.2516	-0.0595	-0.0961	-0.0902
	(-0.10)	(-1.95)	(-2.11)	(7.11)	(5.86)	(6.84)	(-6.68)	(-10.37)	(-10.19)
580 $\leq$ FICO $<$620	0.0044	-0.0015	-0.0007	0.1555	0.1423	0.1629	-0.08	-0.1039	-0.0956
	(0.72)	(-0.24)	(-0.11)	(6.90)	(7.01)	(7.84)	(-10.85)	(-13.85)	(-12.87)
620 $\leq$ FICO $<$660	0.0073	0.0012	0.0034	0.0147	0.0096	0.0093	-0.0728	-0.0899	-0.0827
	(1.73)	(0.28)	(0.90)	(3.98)	(2.71)	(2.92)	(-12.86)	(-15.62)	(-14.21)
660 $\leq$ FICO $<$700	0.0021	-0.0012	0	0.0026	0	-0.0015	-0.0494	-0.0582	-0.0533
	(0.68)	(-0.40)	(-0.02)	(0.91)	(-0.00)	(-0.65)	(-12.67)	(-15.04)	(-13.81)
LTV $<$70	-0.07	-0.0365	-0.0459	-0.1629	-0.1833	-0.1886	-0.0696	-0.0145	-0.0228
	(-10.17)	(-4.99)	(-7.53)	(-13.16)	(-15.54)	(-15.02)	(-8.46)	(-1.77)	(-3.04)
70 $\leq$ LTV $<$80	-0.0842	-0.0512	-0.0575	-0.1789	-0.2005	-0.2045	-0.067	-0.0112	-0.0194
	(-12.62)	(-7.32)	(-9.57)	(-14.32)	(-16.28)	(-15.65)	(-8.52)	(-1.42)	(-2.67)
80 $\leq$ LTV $<$90	-0.0973	-0.0557	-0.0623	-0.2096	-0.2225	-0.2269	-0.0793	-0.0173	-0.0257
	(-18.23)	(-9.83)	(-11.78)	(-18.25)	(-19.51)	(-18.11)	(-13.24)	(-2.73)	(-4.17)
90 $\leq$ LTV $<$100	-0.0899	-0.0584	-0.0683	-0.1597	-0.1942	-0.2013	-0.0748	-0.0177	-0.0273
	(-20.61)	(-11.85)	(-13.40)	(-13.00)	(-16.81)	(-16.49)	(-14.94)	(-3.18)	(-4.89)
Constant	3.5249	3.4493	3.4035	3.0052	3.1314	3.165	4.3391	4.2299	4.1388
	(57.58)	(56.18)	(63.82)	(78.72)	(72.98)	(86.46)	(43.17)	(44.34)	(43.88)
Adj. R^2	0.892	0.895	0.901	0.919	0.920	0.924	0.872	0.876	0.886
Obs.	8492968	8491567	8472185	2842335	2840865	2828260	5650633	5649131	5631026
Orig YQ FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
State FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
Issue YQ FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
Lender FE	N	Y	Y	N	Y	Y	N	Y	Y
Lender × Orig-YQ FE	N	N	Y	N	N	Y	N	N	Y
Other cntrls	Y	Y	Y	Y	Y	Y	Y	Y	Y

Note: This table reports the results of a loan-level regression of time to sale on processing time based on the merged ABS and HMDA dataset. Both time to sale and processing time are expressed in months in the regressions. The control variables are defined in Appendix C. Standard errors are clustered by state and origination quarter, and t-statistics are reported in parentheses.

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

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
PT		-0.0021	-0.0021		-0.0009	-0.0009		-0.0009	-0.0009
		(-9.81)	(-9.78)		(-3.38)	(-3.33)		(-4.71)	(-4.68)
TS	-0.0029		-0.0029	-0.0026		-0.0026	-0.0028		-0.0028
	(-14.88)		(-14.86)	(-14.79)		(-14.80)	(-22.41)		(-22.39)
Term	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	(-1.28)	(-1.14)	(-1.36)	(-2.18)	(-2.05)	(-2.22)	(-1.60)	(-1.44)	(-1.63)
Initial Rate	0.0200	0.0199	0.0199	0.0198	0.0197	0.0198	0.0201	0.0200	0.0201
	(27.89)	(27.47)	(27.84)	(26.90)	(26.56)	(26.85)	(29.18)	(28.75)	(29.13)
Loan Amount	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	(6.77)	(6.80)	(6.80)	(7.04)	(7.03)	(7.03)	(7.54)	(7.54)	(7.54)
LTV	-0.0003	-0.0003	-0.0003	-0.0002	-0.0002	-0.0002	-0.0002	-0.0002	-0.0002
	(-8.09)	(-8.07)	(-8.03)	(-7.15)	(-7.18)	(-7.14)	(-7.10)	(-7.14)	(-7.10)
FICO	-0.0004	-0.0004	-0.0004	-0.0004	-0.0004	-0.0004	-0.0004	-0.0004	-0.0004
	(-41.71)	(-42.14)	(-41.79)	(-42.68)	(-43.06)	(-42.69)	(-38.86)	(-39.27)	(-38.90)
Purchase	0.0200	0.0200	0.0199	0.0185	0.0186	0.0185	0.0185	0.0186	0.0185
	(10.79)	(10.84)	(10.81)	(10.45)	(10.49)	(10.45)	(10.28)	(10.33)	(10.27)
Refi cashout	-0.0167	-0.0169	-0.0169	-0.0159	-0.0159	-0.0160	-0.0153	-0.0153	-0.0154
	(-37.60)	(-38.34)	(-38.65)	(-36.72)	(-36.87)	(-37.23)	(-32.74)	(-32.80)	(-33.07)
ARM	0.0232	0.0231	0.0230	0.0219	0.0220	0.0219	0.0229	0.0229	0.0229
	(18.69)	(18.80)	(18.78)	(19.50)	(19.59)	(19.54)	(21.14)	(21.21)	(21.17)
Balloon	0.0611	0.0609	0.0608	0.0573	0.0573	0.0573	0.0489	0.0489	0.0488
	(15.92)	(15.80)	(15.75)	(15.63)	(15.61)	(15.59)	(15.76)	(15.78)	(15.72)
IOflag	0.0133	0.0133	0.0132	0.0149	0.0150	0.0149	0.0159	0.0160	0.0159
	(15.49)	(15.62)	(15.43)	(19.16)	(19.27)	(19.14)	(19.80)	(19.88)	(19.79)
Jumbo	0.0100	0.0100	0.0100	0.0101	0.0102	0.0101	0.0104	0.0104	0.0104
	(7.44)	(7.50)	(7.51)	(7.52)	(7.54)	(7.55)	(8.25)	(8.29)	(8.30)
Lowdoc	0.0216	0.0217	0.0216	0.0208	0.0209	0.0208	0.0213	0.0214	0.0213
	(23.67)	(23.98)	(23.76)	(24.82)	(25.10)	(24.89)	(25.75)	(26.09)	(25.81)
Prepay Penalty	0.0089	0.0086	0.0086	0.0070	0.0069	0.0070	0.0082	0.0081	0.0082
	(10.37)	(10.09)	(10.16)	(8.63)	(8.60)	(8.62)	(11.25)	(11.15)	(11.21)
Primary Occupancy	-0.0063	-0.0064	-0.0062	-0.0048	-0.0049	-0.0048	-0.0043	-0.0044	-0.0043
	(-3.11)	(-3.18)	(-3.10)	(-2.60)	(-2.65)	(-2.60)	(-2.29)	(-2.34)	(-2.28)
SFR	0.0032	0.0030	0.0030	0.0030	0.0030	0.0030	0.0028	0.0027	0.0027
	(3.13)	(2.94)	(2.96)	(3.27)	(3.23)	(3.24)	(3.04)	(2.99)	(3.00)
Unemp. Rate	0.0038	0.0038	0.0038	0.0037	0.0038	0.0038	0.0037	0.0037	0.0037
	(3.90)	(3.95)	(3.94)	(4.01)	(4.02)	(4.01)	(3.99)	(4.01)	(4.00)
Unemp. Rate Chg.	-0.0005	-0.0005	-0.0005	-0.0003	-0.0003	-0.0003	-0.0003	-0.0003	-0.0003
(36m)	(-0.85)	(-0.75)	(-0.76)	(-0.56)	(-0.52)	(-0.53)	(-0.49)	(-0.44)	(-0.45)
hpi36	-0.0008	-0.0008	-0.0008	-0.0008	-0.0008	-0.0008	-0.0008	-0.0008	-0.0008
	(-14.84)	(-14.80)	(-14.79)	(-14.65)	(-14.64)	(-14.64)	(-16.47)	(-16.44)	(-16.45)
FICO<580	0.0240	0.0241	0.0240	0.0222	0.0223	0.0222	0.0216	0.0216	0.0216
	(19.58)	(19.68)	(19.69)	(18.22)	(18.30)	(18.26)	(17.03)	(17.13)	(17.06)
580≤FICO<620	0.0170	0.0169	0.0169	0.0143	0.0143	0.0143	0.0132	0.0132	0.0132
	(14.43)	(14.38)	(14.35)	(12.33)	(12.36)	(12.32)	(11.06)	(11.10)	(11.06)
620≤FICO<660	0.0057	0.0055	0.0056	0.0034	0.0034	0.0034	0.0027	0.0027	0.0027
	(4.51)	(4.39)	(4.39)	(2.77)	(2.76)	(2.76)	(2.14)	(2.13)	(2.13)
660≤FICO<700	-0.0046	-0.0047	-0.0047	-0.0053	-0.0053	-0.0053	-0.0054	-0.0054	-0.0054
	(-6.86)	(-6.95)	(-6.94)	(-8.02)	(-8.02)	(-8.02)	(-7.82)	(-7.84)	(-7.83)
LTV<70	-0.0341	-0.0334	-0.0336	-0.0341	-0.0339	-0.0340	-0.0328	-0.0326	-0.0327
	(-16.78)	(-16.29)	(-16.46)	(-17.39)	(-17.20)	(-17.28)	(-18.34)	(-18.18)	(-18.26)
70≤LTV<80	-0.0173	-0.0167	-0.0169	-0.0179	-0.0177	-0.0178	-0.0163	-0.0160	-0.0161
	(-8.64)	(-8.16)	(-8.34)	(-9.52)	(-9.29)	(-9.40)	(-9.56)	(-9.33)	(-9.44)
80≤LTV<90	0.0022	0.0027	0.0025	-0.0004	-0.0002	-0.0003	0.0012	0.0014	0.0013
	(1.50)	(1.81)	(1.63)	(-0.31)	(-0.15)	(-0.26)	(0.99)	(1.20)	(1.05)
90≤LTV<100	0.0026	0.0031	0.0028	-0.0006	-0.0004	-0.0005	0.0003	0.0005	0.0004
	(1.34)	(1.57)	(1.44)	(-0.33)	(-0.21)	(-0.30)	(0.18)	(0.34)	(0.22)
Constant	0.1679	0.1600	0.1701	0.1613	0.1532	0.1621	0.1512	0.1425	0.1521
	(15.27)	(14.03)	(15.58)	(14.11)	(13.07)	(14.34)	(15.37)	(14.27)	(15.59)
Adj. R^2	0.077	0.077	0.078	0.081	0.081	0.081	0.084	0.084	0.084
Obs.	8492968	8492968	8492968	8491567	8491567	8491567	8472185	8472185	8472185
Orig YQ FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
State FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
Issue YQ FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
Lender FE	N	N	N	Y	Y	Y	Y	Y	Y
Lender × Orig-YQ FE	N	N	N	N	N	N	Y	Y	Y
Other cntrls	Y	Y	Y	Y	Y	Y	Y	Y	Y

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 dataset. The dependent variable is loan delinquency within 15 months of loan origination. Both time to sale and processing time are expressed in months in the regressions. The control variables are defined in Appendix C. Standard errors are clustered by state and origination quarter, and t-statistics are reported in parentheses.

C Alternative Definition of Default

This section provides a robustness check of our main findings using alternative definitions of default. The results corresponding to Table 5 in the main draft are reported in Table C.1 using 90+ days delinquency within 15 months of origination for default, Table C.2 using 60+ days delinquency within 18 months of origination for default, and Table C.3 using 60+ days delinquency within 24 months of origination for default.

Table C.1: Processing Time, Time to Sale, and Loan Default
(90+ days delinquency within 15 months of origination)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A: All PLS									
PT		-0.0018	-0.0018		-0.0006	-0.0006		-0.0009	-0.0009
		(-9.79)	(-9.78)		(-3.72)	(-3.68)		(-5.12)	(-5.10)
TS	-0.0025		-0.0024	-0.0022		-0.0022	-0.0024		-0.0024
	(-14.28)		(-14.27)	(-14.13)		(-14.13)	(-21.61)		(-21.57)
Adj. R^2	0.066	0.066	0.066	0.069	0.069	0.069	0.072	0.072	0.072
Obs.	8492968	8492968	8492968	8491567	8491567	8491567	8472185	8472185	8472185
Panel B: Alt-A									
PT		-0.0015	-0.0015		-0.0009	-0.0009		-0.0009	-0.0009
		(-12.26)	(-12.34)		(-6.39)	(-6.35)		(-6.98)	(-6.95)
TS	-0.0020		-0.0020	-0.0019		-0.0019	-0.0019		-0.0019
	(-11.38)		(-11.44)	(-10.90)		(-10.91)	(-11.44)		(-11.43)
Adj. R^2	0.048	0.048	0.048	0.052	0.052	0.052	0.059	0.059	0.059
Obs.	2842335	2842335	2842335	2840865	2840865	2840865	2828260	2828260	2828260
Panel C: Subprime									
PT		-0.0015	-0.0012		-0.0006	-0.0003		-0.0006	-0.0006
		(-5.17)	(-4.92)		(-1.72)	(-1.63)		(-1.92)	(-1.85)
TS	-0.0032		-0.0032	-0.0030		-0.0030	-0.0029		-0.0029
	(-16.44)		(-16.45)	(-15.47)		(-15.49)	(-19.42)		(-19.43)
Adj. R^2	0.065	0.064	0.065	0.067	0.067	0.067	0.069	0.069	0.069
Obs.	5650633	5650633	5650633	5649131	5649131	5649131	5631026	5631026	5631026
Orig YQ FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
State FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
Issue YQ FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
Lender FE	N	N	N	Y	Y	Y	Y	Y	Y
Lender × Orig-YQ FE	N	N	N	N	N	N	Y	Y	Y
Other cntrls	Y	Y	Y	Y	Y	Y	Y	Y	Y

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 dataset. The dependent variable is loan delinquency using 90+ days delinquency within 15 months of loan origination. Both time to sale and processing time are expressed in months in the regressions. The control variables are defined in Appendix C. Standard errors are clustered by state and origination quarter, and t-statistics are reported in parentheses.

Table C.2: Processing Time, Time to Sale, and Loan Default
(60+ days delinquency within 18 months of origination)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A: All PLS									
PT		-0.0027	-0.0027		-0.0009	-0.0009		-0.0012	-0.0012
		(-10.88)	(-10.85)		(-3.75)	(-3.70)		(-5.20)	(-5.17)
TS	-0.00328		-0.00326	-0.00289		-0.00289	-0.00319		-0.00318
	(-13.98)		(-13.92)	(-13.40)		(-13.40)	(-19.74)		(-19.71)
Adj. R^2	0.096	0.096	0.096	0.100	0.100	0.100	0.104	0.104	0.104
Obs.	8492968	8492968	8492968	8491567	8491567	8491567	8472185	8472185	8472185
Panel B: Alt-A									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
PT		-0.0021	-0.0021		-0.0012	-0.0012		-0.0012	-0.0012
		(-12.77)	(-12.85)		(-6.98)	(-6.94)		(-7.37)	(-7.33)
TS	-0.00287		-0.00286	-0.00266		-0.00265	-0.00271		-0.00270
	(-12.18)		(-12.22)	(-11.74)		(-11.74)	(-12.28)		(-12.26)
Adj. R^2	0.077	0.077	0.077	0.082	0.082	0.082	0.089	0.088	0.089
Obs.	2842335	2842335	2842335	2840865	2840865	2840865	2828260	2828260	2828260
Panel C: Subprime									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
PT		-0.0018	-0.0018		-0.0006	-0.0006		-0.0006	-0.0006
		(-5.64)	(-5.36)		(-1.46)	(-1.37)		(-1.62)	(-1.54)
TS	-0.00422		-0.00420	-0.00385		-0.00385	-0.00363		-0.00363
	(-15.51)		(-15.48)	(-13.85)		(-13.87)	(-17.17)		(-17.17)
Adj. R^2	0.092	0.092	0.092	0.095	0.095	0.095	0.098	0.097	0.098
Obs.	5650633	5650633	5650633	5649131	5649131	5649131	5631026	5631026	5631026
Orig YQ FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
State FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
Issue YQ FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
Lender FE	N	N	N	Y	Y	Y	Y	Y	Y
Lender × Orig-YQ FE	N	N	N	N	N	N	Y	Y	Y
Other cntrls	Y	Y	Y	Y	Y	Y	Y	Y	Y

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 dataset. The dependent variable is loan delinquency using 60+ days delinquency within 18 months of loan origination. Both time to sale and processing time are expressed in months in the regressions. The control variables are defined in Appendix C. Standard errors are clustered by state and origination quarter, and t-statistics are reported in parentheses.

Table C.3: Processing Time, Time to Sale, and Loan Default (60+ days delinquency within 24 months of origination

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A: All PLS									
PT		-0.0036	-0.0033		-0.0015	-0.0015		-0.0015	-0.0015
		(-12.26)	(-12.23)		(-4.56)	(-4.51)		(-6.28)	(-6.26)
TS	-0.00407		-0.00404	-0.00355		-0.00355	-0.00387		-0.00387
	(-17.23)		(-17.14)	(-16.43)		(-16.41)	(-23.15)		(-23.10)
Adj. R^2	0.135	0.135	0.135	0.140	0.139	0.140	0.144	0.144	0.144
Obs.	8492968	8492968	8492968	8491567	8491567	8491567	8472185	8472185	8472185
Panel B: Alt-A	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
PT		-0.0033	-0.0033		-0.0018	-0.0018		-0.0018	-0.0018
		(-14.93)	(-15.04)		(-9.48)	(-9.47)		(-10.11)	(-10.10)
TS	-0.00397		-0.00395	-0.00362		-0.00361	-0.00361		-0.00360
	(-14.77)		(-14.79)	(-13.08)		(-13.07)	(-13.08)		(-13.05)
Adj. R^2	0.128	0.128	0.128	0.135	0.135	0.135	0.142	0.142	0.142
Obs.	2842335	2842335	2842335	2840865	2840865	2840865	2828260	2828260	2828260
Panel C: Subprime	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
PT		-0.0021	-0.0018		-0.0003	-0.0003		-0.0003	-0.0003
		(-5.26)	(-4.96)		(-1.08)	(-0.98)		(-1.16)	(-1.08)
TS	-0.00509		-0.00507	-0.00462		-0.00462	-0.00432		-0.00432
	(-18.72)		(-18.68)	(-16.75)		(-16.77)	(-21.51)		(-21.52)
Adj. R^2	0.128	0.128	0.128	0.132	0.132	0.132	0.134	0.134	0.134
Obs.	5650633	5650633	5650633	5649131	5649131	5649131	5631026	5631026	5631026
Orig YQ FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
State FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
Issue YQ FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
Lender FE	N	N	N	Y	Y	Y	Y	Y	Y
Lender × Orig-YQ FE	N	N	N	N	N	N	Y	Y	Y
Other cntrls	Y	Y	Y	Y	Y	Y	Y	Y	Y

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 dataset. The dependent variable is loan delinquency using 60+ days delinquency within 24 months of loan origination. Both time to sale and processing time are expressed in months in the regressions. The control variables are defined in Appendix C. Standard errors are clustered by state and origination quarter, and t-statistics are reported in parentheses.

D Excluding Loans with Month-End Closing Dates

This section provides a robustness check 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 3 in the main draft are reported in Table D.1. The results corresponding to Table 4 in the main draft are reported in Table D.2.

TABLE D.1: Processing Time and Time to Sale
(excluding loans with month-end closing dates)

	All PLS			Alt-A			Subprime		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
PT	0.0105	0.0072	0.0057	0.0036	0.0069	0.0063	0.027	0.0108	0.0093
	(4.48)	(3.03)	(2.48)	(2.06)	(4.19)	(3.82)	(10.55)	(3.24)	(2.97)
Adj. R^2	0.894	0.896	0.903	0.918	0.920	0.924	0.875	0.879	0.889
Obs.	5932483	5931039	5913099	2020410	2019007	2007342	3912073	3910513	3894005
Orig YQ FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
State FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
Issue YQ FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
Lender FE	N	Y	Y	N	Y	Y	N	Y	Y
Lender × Orig-YQ FE	N	N	Y	N	N	Y	N	N	Y
Other cntrls	Y	Y	Y	Y	Y	Y	Y	Y	Y

NOTE: This table reports the results of a loan-level regression of time to sale on processing time based on the merged ABS and HMDA dataset. Both time to sale and processing time are expressed in months in the regressions. The control variables are defined in Appendix C. Standard errors are clustered by state and origination quarter, and t-statistics are reported in parentheses.

Table D.2: Processing Time, Time to Sale, and Loan Default
(excluding loans with month-end closing dates)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A: All PLS									
PT		-0.0021	-0.0021		-0.0009	-0.0006		-0.0009	-0.0009
		(-8.86)	(-8.84)		(-3.00)	(-2.94)		(-4.06)	(-4.01)
TS	-0.00273		-0.00271	-0.00245		-0.00245	-0.00264		-0.00264
	(-14.22)		(-14.24)	(-14.21)		(-14.25)	(-20.89)		(-20.94)
Adj. R^2	0.078	0.078	0.078	0.082	0.082	0.082	0.085	0.085	0.085
Obs.	5932483	5932483	5932483	5931039	5931039	5931039	5913099	5913099	5913099
Panel B: Alt-A									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
PT		-0.0018	-0.0018		-0.0009	-0.0009		-0.0009	-0.0009
		(-12.03)	(-12.12)		(-5.82)	(-5.79)		(-6.29)	(-6.25)
TS	-0.00224		-0.00223	-0.00210		-0.00209	-0.00210		-0.00209
	(-10.22)		(-10.30)	(-9.85)		(-9.87)	(-10.19)		(-10.20)
Adj. R^2	0.056	0.056	0.056	0.061	0.061	0.061	0.067	0.067	0.067
Obs.	2020410	2020410	2020410	2019007	2019007	2019007	2007342	2007342	2007342
Panel C: Subprime									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
PT		-0.0015	-0.0015		-0.0003	-0.0003		-0.0003	-0.0003
		(-4.55)	(-4.32)		(-0.99)	(-0.89)		(-1.03)	(-0.96)
TS	-0.00356		-0.00353	-0.00333		-0.00333	-0.00303		-0.00303
	(-15.91)		(-15.96)	(-15.02)		(-15.06)	(-18.10)		(-18.14)
Adj. R^2	0.075	0.075	0.075	0.078	0.078	0.078	0.080	0.080	0.080
Obs.	3912073	3912073	3912073	3910513	3910513	3910513	3894005	3894005	3894005
Orig YQ FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
State FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
Issue YQ FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
Lender FE	N	N	N	Y	Y	Y	Y	Y	Y
Lender × Orig-YQ FE	N	N	N	N	N	N	Y	Y	Y
Other cntrls	Y	Y	Y	Y	Y	Y	Y	Y	Y

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 dataset. The dependent variable is loan delinquency within 15 months of loan origination. Both time to sale and processing time are expressed in months in the regressions. The control variables are defined in Appendix C. Standard errors are clustered by state and origination quarter, and t-statistics are reported in parentheses.

E Proofs: The Second-best

We focus on direct revelation mechanisms that stipulate a time to sale $t_{\hat{z}}$ and price $p_{\hat{z}}$ contingent on reported type $\hat{z} \in \{g, b\}$.

Define

$$\begin{aligned} u_a^z(t) &= \mathbb{E}_a^z\left[\int_0^t c e^{-\gamma u} 1_{\tau_d \geq u} du\right] = \frac{c}{\gamma + \lambda(a)}\left(1 - e^{-(\gamma+\lambda(a))t}\right), \\ v_a^z &= \mathbb{E}_a^z\left[\int_t^\infty \theta c e^{-\gamma(u-t)} 1_{\tau_d \geq u} du\right] = \frac{\theta c}{\gamma + \lambda(a)}. \end{aligned}$$

Similarly, we define $u^b(t)$ and v^b.

Definition 2. *The optimal mechanism is given by an implementation effort level and transfers* $\{a, t_z, p_z\}_{z\in\{g,b\}}$ *that maximizes the value for the originator at* $t = 0$

$$\max_{\{a,t_z,p_z\}} a\left(u_a^g(t_g) + e^{-(\gamma+\lambda(a))t_g} p_g\right) + (1-a)\left(u^b(t_b) + e^{-(\gamma+\lambda_b)t_b} p_b\right) - C(a), \tag{16}$$

subject to:

1. *Incentive Compatibility for Type Revelation:*

$$u^b(t_g) + e^{-(\gamma+\lambda_b)t_g} p_g \leq u^b(t_b) + e^{-(\gamma+\lambda_b)t_b} p_b, \tag{17}$$

$$u_a^g(t_b) + e^{-(\gamma+\lambda(a))t_b} p_b \leq u_a^g(t_g) + e^{-(\gamma+\lambda(a))t_g} p_g. \tag{18}$$

2. *Investors' Participation Constraint:*

$$\begin{aligned} & a\mathbb{E}_a^g\left(e^{-\gamma t_g} 1_{\tau_d \geq t_g}(v_a^g - p_g)\right) + (1-a)\mathbb{E}^b\left(e^{-\gamma t_b} 1_{\tau_d \geq t_b}\left(v^b - p_b\right)\right) \\ = \;& a e^{-(\gamma+\lambda(a))t_g}(v_a^g - p_g) + (1-a) e^{-(\gamma+\lambda_b)t_b}\left(v^b - p_b\right) \geq 0. \end{aligned} \tag{19}$$

3. *Incentive Compatibility for Effort Choice:*

$$\begin{aligned} a = \;& \arg\max_{\hat{a}} \hat{a} \max\left\{u_{\hat{a}}^g(t_g) + e^{-(\gamma+\lambda(\hat{a}))t_g} p_g, u_{\hat{a}}^g(t_b) + e^{-(\gamma+\lambda(\hat{a}))t_b} p_b\right\} \\ & + (1-\hat{a})\left(u^b(t_b) + e^{-(\gamma+\lambda_b)t_b} p_b\right) - C(\hat{a}). \end{aligned} \tag{20}$$

4. *Feasibility:*

$$p_g - \frac{c}{\gamma + \lambda(a)} \geq p_b - \frac{c}{\gamma + \lambda_b} \geq 0. \tag{21}$$

The following proposition characterizes the equilibrium in the second-best.

Proposition 5. *In the second-best, the optimal effort and time-to-sale* $\left\{a^{SB}, t_g^{SB}\right\}$ *are the interior solution to the following optimization problem:*

$$\max_{a\in[0,1], t_g \geq 0} \theta\rho(a) - C_R(a; t_g) - C(a), \tag{22}$$

subject to

$$\begin{aligned} & u_a^g(t_g) - u^b(t_g) + \left(e^{-(\gamma+\lambda(a))t_g} - e^{-(\gamma+\lambda_b)t_g}\right) p_g \\ & + a\lambda'(a)\left[-\frac{c}{(\gamma+\lambda(a))^2}\left(1 - e^{-(\gamma+\lambda(a))t_g}\right) + \frac{c t_g}{\gamma + \lambda(a)} e^{-(\gamma+\lambda(a))t_g} - e^{-(\gamma+\lambda(a))t_g} t_g p_g\right] \\ = \;& C'(a), \end{aligned} \tag{23}$$

where p_g is given in (29) in Appendix A.2, and $C_R(a;t_g)$ denotes the cost of signaling via delayed sales, given by

$$C_R(a;t_g) = \frac{(\theta-1)\,ac}{\gamma+\lambda(a)}\left(1-e^{-(\gamma+\lambda(a))t_g}\right). \tag{24}$$

Global Deviations. Constraint (20) controls for the possibility of the originator choosing to deviate on her effort at the origination stage and then misreporting her z-type at the securitization stage. To address this, we proceed as follows. We replace the incentive compatibility constraint for effort (20) with the first-order condition for effort choice, obtained when the incentive compatibility for type revelation of the g-type (18) holds:

$$\begin{aligned}&\underbrace{u_a^g(t_g) - u^b(t_b) + \left(e^{-(\gamma+\lambda(a))t_g}p_g - e^{-(\gamma+\lambda_b)t_b}p_b\right)}_{\text{Difference in Payoff between } g\text{- and } b\text{-types}} + \underbrace{a\left[\left.\frac{\partial u_{\widehat{a}}^g(t_g)}{\partial \widehat{a}}\right|_a - e^{-(\gamma+\lambda(a))t_g}\lambda'(a)\,t_g p_g\right]}_{\text{Marginal Change in Quality of Delayed Sale}}\\
=\;&\underbrace{C'(a)}_{\text{Marginal Cost}},\end{aligned} \tag{25}$$

where

$$\left.\frac{\partial u_{\widehat{a}}^g(t_g)}{\partial \widehat{a}}\right|_a = -\frac{c\lambda'(a)}{(\gamma+\lambda(a))^2}\left(1-e^{-(\gamma+\lambda(a))t_g}\right) + \frac{c\lambda'(a)\,t_g}{\gamma+\lambda(a)}e^{-(\gamma+\lambda(a))t_g}.$$

Later, we verify that the allocations obtained under the first-order approach satisfy global incentive compatibility.

The following lemma presents the first important result: only delayed sale of the g-type manager is desired in the optimal mechanism.

Lemma 3. *Under the optimal mechanism, the bad-type originator does not delay any sale, $t_b = 0$, while the good-type originator does delay sale: $t_g > 0$ if effort is strictly positive (i.e., $a > 0$).*

Proof of Lemma 3. Lemmas 10 and 11 show that under the optimal mechanism of the participation constraint (PC) of investors and the incentive compatibility constraint (IC) of the b-type bind. By plugging in the binding PC of investors to the value for the originator at $t = 0$, we obtain

$$\begin{aligned}
V_0 &= a\left(u_a^g(t_g) + e^{-(\gamma+\lambda(a))t_g}p_g\right) + (1-a)\left(u^b(t_b) + e^{-(\gamma+\lambda_b)t_b}p_b\right) - C(a) \\
&= a\left(u_a^g(t_g) + e^{-(\gamma+\lambda(a))t_g}v_a^g\right) + (1-a)\left(u^b(t_b) + e^{-(\gamma+\lambda_b)t_b}v^b\right) - C(a) \\
&= a\left(\frac{c}{\gamma+\lambda(a)}\left(1-e^{-(\gamma+\lambda(a))t_g}\right) + e^{-(\gamma+\lambda(a))t_g}\frac{\theta c}{\gamma+\lambda(a)}\right) \\
&\quad + (1-a)\left(\frac{c}{\gamma+\lambda_b}\left(1-e^{-(\gamma+\lambda_b)t_b}\right) + e^{-(\gamma+\lambda_b)t_b}\frac{\theta c}{\gamma+\lambda_b}\right) - C(a) \\
&= \left[av_a^g + (1-a)\,v^b\right] - \left(\begin{array}{c} a\left(\frac{\theta c}{\gamma+\lambda(a)}\left(1-e^{-(\gamma+\lambda(a))t_g}\right) - \frac{c}{\gamma+\lambda(a)}\left(1-e^{-(\gamma+\lambda(a))t_g}\right)\right) \\ +(1-a)\left(\frac{\theta c}{\gamma+\lambda_b}\left(1-e^{-(\gamma+\lambda_b)t_b}\right) - \frac{c}{\gamma+\lambda_b}\left(1-e^{-(\gamma+\lambda_b)t_b}\right)\right)\end{array}\right) - C(a) \\
&= \left[av_a^g + (1-a)\,v^b\right] - (\theta-1)\left(\begin{array}{c} a\frac{c}{\gamma+\lambda(a)}\left(1-e^{-(\gamma+\lambda(a))t_g}\right) \\ +(1-a)\frac{c}{\gamma+\lambda_b}\left(1-e^{-(\gamma+\lambda_b)t_b}\right)\end{array}\right) - C(a).
\end{aligned} \tag{26}$$

And by plugging in the binding IC for the b-type into the IC for effort, we obtain (see Lemma 12)

$$\begin{aligned}
&u_a^g(t_g) - u^b(t_g) + \left(e^{-(\gamma+\lambda(a))t_g} - e^{-(\gamma+\lambda_b)t_g}\right)p_g \\
&+a\lambda'(a)\left[-\frac{c}{(\gamma+\lambda(a))^2}\left(1-e^{-(\gamma+\lambda(a))t_g}\right) + \frac{ct_g}{\gamma+\lambda(a)}e^{-(\gamma+\lambda(a))t_g} - e^{-(\gamma+\lambda(a))t_g}t_g p_g\right] \\
=\;& C'(a).
\end{aligned} \tag{27}$$

Therefore, an optimal mechanism $\{a^*, p_z, t_z\}$ maximizes (26) subject to (27) and to the ICs for type revelation and the PC of investors.

We next prove that $t_b = 0$ by contradiction. Assume $t_b > 0$. Consider an alternative $t_b' = t_b - \epsilon$, where $\epsilon > 0$ is sufficiently small. We next choose p_b' so that the IC for the b-type is unaffected: $u^b(t_b') + e^{-(\gamma+\lambda_b)t_b'}p_b' = u^b(t_b) + e^{-(\gamma+\lambda_b)t_b}p_b$. That is, $p_b' = \frac{c}{\gamma+\lambda_b}\left(1 - e^{-(\gamma+\lambda_b)\epsilon}\right) + e^{-(\gamma+\lambda_b)\epsilon}p_b$.

We next prove that the IC for the g-type is relaxed by proving $u_a^g(t_b') + e^{-(\gamma+\lambda(a))t_b'}p_b' \leq u_a^g(t_b) + e^{-(\gamma+\lambda(a))t_b}p_b$. In fact

$$\begin{aligned}
& \left(u_a^g(t_b') + e^{-(\gamma+\lambda(a))t_b'}p_b'\right) - \left(u_a^g(t_b) + e^{-(\gamma+\lambda(a))t_b}p_b\right) \\
= \; & \frac{c}{\gamma+\lambda(a)}\left(1 - e^{-(\gamma+\lambda(a))t_b'}\right) + e^{-(\gamma+\lambda(a))t_b'}\left(\frac{c}{\gamma+\lambda_b}\left(1 - e^{-(\gamma+\lambda_b)\epsilon}\right) + e^{-(\gamma+\lambda_b)\epsilon}p_b\right) \\
& - \left(\frac{c}{\gamma+\lambda(a)}\left(1 - e^{-(\gamma+\lambda(a))t_b}\right) + e^{-(\gamma+\lambda(a))t_b}p_b\right) \\
= \; & -\frac{c}{\gamma+\lambda(a)}\left(e^{-(\gamma+\lambda(a))t_b'} - e^{-(\gamma+\lambda(a))t_b}\right) + e^{-(\gamma+\lambda(a))t_b'}\frac{c}{\gamma+\lambda_b}\left(1 - e^{-(\gamma+\lambda_b)\epsilon}\right) \\
& + e^{-(\gamma+\lambda(a))t_b}p_b\left(e^{(\gamma+\lambda(a))\epsilon-(\gamma+\lambda_b)\epsilon} - 1\right) \\
= \; & e^{-(\gamma+\lambda(a))t_b'}\left[-\frac{c}{\gamma+\lambda(a)}\left(1 - e^{-(\gamma+\lambda(a))\epsilon}\right) + \frac{c}{\gamma+\lambda_b}\left(1 - e^{-(\gamma+\lambda_b)\epsilon}\right)\right] \\
& + e^{-(\gamma+\lambda(a))t_b}p_b\left(e^{(\lambda(a)-\lambda_b)\epsilon} - 1\right) \\
\leq \; & 0,
\end{aligned}$$

where we have used the following results: $e^{(\lambda(a)-\lambda_b)\epsilon} \leq 1$ (because $\lambda(a) \leq \lambda_b$) and

$$-\frac{c}{\gamma+\lambda(a)}\left(1 - e^{-(\gamma+\lambda(a))\epsilon}\right) + \frac{c}{\gamma+\lambda_b}\left(1 - e^{-(\gamma+\lambda_b)\epsilon}\right) = -u_a^g(\epsilon) + u^b(\epsilon) \leq 0, \text{ (Lemma 7(ii))}$$

Note that under the new policy, the IC for effort (27) is unaffected. However, we prove next that the PC for investors is relaxed. Note that

$$\begin{aligned}
& e^{-(\gamma+\lambda_b)t_b'}\left(v^b - p_b'\right) - e^{-(\gamma+\lambda_b)t_b}\left(v^b - p_b\right) \\
= \; & e^{-(\gamma+\lambda_b)t_b'}\left(\frac{\theta c}{\gamma+\lambda_b} - p_b' - e^{-(\gamma+\lambda_b)\epsilon}\left(\frac{\theta c}{\gamma+\lambda_b} - p_b\right)\right) \\
= \; & e^{-(\gamma+\lambda_b)t_b'}\left(\frac{\theta c}{\gamma+\lambda_b}\left(1 - e^{-(\gamma+\lambda_b)\epsilon}\right) - \left(p_b' - e^{-(\gamma+\lambda_b)\epsilon}p_b\right)\right) \\
= \; & e^{-(\gamma+\lambda_b)t_b'}(\theta - 1)\frac{c}{\gamma+\lambda_b}\left(1 - e^{-(\gamma+\lambda_b)\epsilon}\right) \\
> \; & 0.
\end{aligned}$$

Next, let $p_b'' = p_b' + e^{(\gamma+\lambda_b)t_b'}\epsilon'$, $p_g'' = p_g' + e^{(\gamma+\lambda_b)t_g}\epsilon'$, where $\epsilon' > 0$ is small enough that the investors' PC continues to hold. Similarly, as in the proof of Lemma 10, it is easy to check that the IC for the b-type is unaffected. The IC for the g-type continues to hold. However, these new prices increase the value to the manager and do not affect the other constraints. Contradiction.

Lastly, from the proof of Lemma 10, we have $t_g > t_b = 0$ for $a > 0$. □

Lemma 4. *Under the optimal mechanism, for given effort and time-to-sale $\{a, t_g\}$, the prices are given by*

$$p_b = \frac{ae^{-(\gamma+\lambda(a))t_g}\left(e^{-(\gamma+\lambda_b)t_g}v_a^g + u^b(t_g)\right) + (1-a)\,e^{-(\gamma+\lambda_b)t_g}v^b}{ae^{-(\gamma+\lambda(a))t_g} + (1-a)\,e^{-(\gamma+\lambda_b)t_g}}, \tag{28}$$

$$p_g = \frac{ae^{-(\gamma+\lambda(a))t_g}v_a^g + (1-a)\left(v^b - u^b(t_g)\right)}{ae^{-(\gamma+\lambda(a))t_g} + (1-a)\,e^{-(\gamma+\lambda_b)t_g}}. \tag{29}$$

Proof of Lemma 4. Lemmas 10 and 11 show that under the optimal mechanism, the PC of investors, and the b-type IC bind. The binding constraints together with $t_b = 0$ in Lemma 3 imply

$$\begin{aligned} p_b &= u^b(t_g) + e^{-(\gamma+\lambda_b)t_g}p_g, \\ ae^{-(\gamma+\lambda(a))t_g}p_g + (1-a)\,p_b &= ae^{-(\gamma+\lambda(a))t_g}v_a^g + (1-a)\,v^b. \end{aligned}$$

By solving this system of equations for $\{p_b, p_g\}$, the expressions in (28) and (29) are obtained. □

Lemma 5. *In the optimal mechanism, effort and time to sale $\{a^*, t_g^*\}$ solve:*

$$\max_{a\in[0,1],t_g\geq 0} \left(av_a^g + (1-a)\,v^b\right) - a(\theta-1)\frac{c}{\gamma+\lambda(a)}\left(1 - e^{-(\gamma+\lambda(a))t_g}\right) - C(a), \tag{30}$$

subject to

$$\begin{aligned} & u_a^g(t_g) - u^b(t_g) + \left(e^{-(\gamma+\lambda(a))t_g} - e^{-(\gamma+\lambda_b)t_g}\right)p_g \\ & +a\lambda'(a)\left[-\frac{c}{(\gamma+\lambda(a))^2}\left(1 - e^{-(\gamma+\lambda(a))t_g}\right) + \frac{ct_g}{\gamma+\lambda(a)}e^{-(\gamma+\lambda(a))t_g} - e^{-(\gamma+\lambda(a))t_g}t_g p_g\right] \\ = \; & C'(a), \end{aligned} \tag{31}$$

where p_g is given in (29). Screening effort is always below first-best: $a^* < a^{FB}$.

Proof of Lemma 5. Plugging the binding IC of the b-type (Lemma 11) into the IC for effort resulting from the first-order approach (25), we obtain the equation (23).

$$\begin{aligned} & u_a^g(t_g) - u^b(t_g) + e^{-(\gamma+\lambda(a))t_g}p_g - e^{-(\gamma+\lambda_b)t_g}p_g \\ & +a\lambda'(a)\left[-\frac{c}{(\gamma+\lambda(a))^2}\left(1 - e^{-(\gamma+\lambda(a))t_g}\right) + \frac{ct_g}{\gamma+\lambda(a)}e^{-(\gamma+\lambda(a))t_g} - e^{-(\gamma+\lambda(a))t_g}p_g\right] \\ = \; & C'(a) \end{aligned}$$

Next, plugging the binding IC of the b-type and the expression of p_g in (29), the value to the originator at time 0 for a given effort choice a and time-to-sale t_g is

$$\begin{aligned}
& a\left(u_a^g\left(t_g\right)+e^{-(\gamma+\lambda(a)) t_g} p_g\right)+(1-a)\left(u^b\left(t_b\right)+e^{-\left(\gamma+\lambda_b\right) t_b} p_b\right)-C(a) \\
= \; & a u_a^g\left(t_g\right)+a e^{-(\gamma+\lambda(a)) t_g} p_g+(1-a) p_b-C(a) \\
= \; & a u_a^g\left(t_g\right)+a e^{-(\gamma+\lambda(a)) t_g} p_g+(1-a)\left(u^b\left(t_g\right)+e^{-\left(\gamma+\lambda_b\right) t_g} p_g\right)-C(a) \quad \text{(IC-b)} \\
= \; & a u_a^g\left(t_g\right)+(1-a) u^b\left(t_g\right)+\left(a e^{-(\gamma+\lambda(a)) t_g}+(1-a) e^{-\left(\gamma+\lambda_b\right) t_g}\right) p_g-C(a) \\
= \; & a u_a^g\left(t_g\right)+(1-a) u^b\left(t_g\right)+a e^{-(\gamma+\lambda(a)) t_g} v_a^g+(1-a)\left(v^b-u^b\left(t_g\right)\right)-C(a) \\
= \; & a\left(u_a^g\left(t_g\right)+e^{-(\gamma+\lambda(a)) t_g} v_a^g\right)+(1-a) v^b-C(a) \\
= \; & \left(a v_a^g+(1-a) v^b\right)-a\left(\left(1-e^{-(\gamma+\lambda(a)) t_g}\right) v_a^g-u_a^g\left(t_g\right)\right)-C(a) \\
\equiv \; & \left(a v_a^g+(1-a) v^b\right)-C_R\left(a ; t_g\right)-C(a),
\end{aligned}$$

where

$$C_R\left(a ; t_g\right) \equiv a\left(\left(1-e^{-(\gamma+\lambda(a)) t_g}\right) v_a^g-u_a^g\left(t_g\right)\right)=\frac{(\theta-1) a c}{\gamma+\lambda(a)}\left(1-e^{-(\gamma+\lambda(a)) t_g}\right)>0.$$

Therefore, the mechanism chooses $\{a, t_g\} \in [0,1] \times \mathbb{R}^{+}$ to maximize (22) subject to (23), since $\{p_b^*, p_g^*\}$ ensures that the IC for type revelation and the PC of investors hold for any given pair $\{a, t_g\}$, and where a binding IC for the b-type ensures a slack IC for the g-type. □

Lemma 5 shows that effort is chosen to maximize the originator's $t=0$ value, which is lower than in the first-best when there is delayed sale. The additional term (relative to the value function in the first-best), $a\left(\left(1-e^{-(\gamma+\lambda(a)) t_g}\right) v_a^g-u_a^g\left(t_g\right)\right)=a(\theta-1) \frac{c}{\gamma+\lambda(a)}\left(1-e^{-(\gamma+\lambda(a)) t_g}\right)$, captures the indirect cost of effort given by the delay of sale required to implement it.

Let $\bar{a} \in (0,1]$ denote the maximum effort level that can be implemented under the optimal mechanism, given by

$$\frac{c}{\gamma+\lambda(\bar{a})}-\frac{c}{\gamma+\lambda_b}-\frac{\bar{a} c \lambda^{\prime}(\bar{a})}{(\gamma+\lambda(\bar{a}))^2}=C^{\prime}(\bar{a}).$$

By comparison with (4), it follows that the level of effort under the optimal mechanism is always below the first-best, $\bar{a}<a_{FB}^*$.

The following condition is necessary and sufficient for positive effort to be implemented under the optimal mechanism. The condition states that there exists a positive effort level that gives the manager a higher $t=0$ payoff than exerting zero effort.

Lemma 6. *In the endogenous quality case, the IC for effort (20) can be replaced by the following two constraints:*

$$a=\arg \max_{\widehat{a}} \widehat{a}\left(u_{\widehat{a}}^g\left(t_g\right)+e^{-(\gamma+\lambda(\widehat{a})) t_g} p_g\right)+(1-\widehat{a})\left(u^b\left(t_b\right)+e^{-\left(\gamma+\lambda_b\right) t_b} p_b\right)-C(\widehat{a}), \tag{32}$$

and

$$\begin{aligned}
& \max_{\widehat{a}} \widehat{a}\left(u_{\widehat{a}}^g\left(t_g\right)+e^{-(\gamma+\lambda(\widehat{a})) t_g} p_g\right)+(1-\widehat{a})\left(u^b\left(t_b\right)+e^{-\left(\gamma+\lambda_b\right) t_b} p_b\right)-C(\widehat{a}) \\
\geq \; & \max_{\widehat{a}} \widehat{a}\left(u_{\widehat{a}}^g\left(t_b\right)+e^{-(\gamma+\lambda(\widehat{a})) t_b} p_b\right)+(1-\widehat{a})\left(u^b\left(t_b\right)+e^{-\left(\gamma+\lambda_b\right) t_b} p_b\right)-C(\widehat{a}).
\end{aligned} \tag{33}$$

Proof. First, note that the IC for the b-type is independent of a, and thus holds for all $a \in [0,1]$. When effort is endogenous, the IC for the g-type does depend on a. Therefore, to ensure that the mechanism is robust to global deviations, the constraint (33) is imposed. The constraint ensures that reporting a g-type truthfully and exerting the corresponding best-response effort gives the manager at least as much value as deviating both on her effort choice and on her ex post report of type. □

Complementary lemmas for the second-best

Lemma 7. *The following rudimentary results hold:*

(i) Fixing $t > 0$, $u_a^g(t) = \frac{c}{\gamma+\lambda(a)}\left(1 - e^{-(\gamma+\lambda(a))t}\right)$ is a strictly increasing function of effort a.

(ii) Fixing $t > 0$, $u_a^g(t) - u^b(t) > 0$ for $a > 0$.

(iii) Fixing $t > 0$, $\frac{c}{r+\lambda(a)}\left(1 - e^{-(r+\lambda(a))t}\right) > \frac{c}{\gamma+\lambda(a)}\left(1 - e^{-(\gamma+\lambda(a))t}\right)$.

(iv) Fixing $a \geq 0$, $u_a^g(t) + e^{-(\gamma+\lambda(a))t}p_g$ is strictly decreasing in t if $p_g > \frac{c}{\gamma+\lambda(a)}$.

Proof of Lemma 7. (i) Let $f(a;t) \equiv \frac{c}{\gamma+\lambda(a)}\left(1 - e^{-(\gamma+\lambda(a))t}\right)$ and $t > 0$. It is straightforward to show $f'(a;t) > 0$ if $t > 0$, because

$$\begin{aligned} f'(a;t) &= -\frac{c\lambda'(a)}{(\gamma+\lambda(a))^2}\left(1 - e^{-(\gamma+\lambda(a))t}\right) + \frac{c\lambda'(a)\,t}{\gamma+\lambda(a)}e^{-(\gamma+\lambda(a))t} \\ &= -\frac{c\lambda'(a)}{(\gamma+\lambda(a))^2}e^{-(\gamma+\lambda(a))t}\left(e^{(\gamma+\lambda(a))t} - 1 - (\gamma+\lambda(a))\,t\right) \\ &> 0, \end{aligned}$$

where the last inequality is obtained because $\lambda'(a) = \underline{\lambda} - \lambda_b < 0$ and $e^x - 1 - x > 0$ for any $x > 0$.

(ii) Fixing $t > 0$, we have shown that $f(a;t) > f(0;t) = u^b(t)$ for any $a > 0$.

(iii) The proof is similar as in (1) except that we denote $f(x;t) \equiv \frac{c}{x+\lambda(a)}\left(1 - e^{-(x+\lambda(a))t}\right)$, fixing $t > 0$. Then $f'(x;t) = -\frac{c}{(x+\lambda(a))^2}e^{-(x+\lambda(a))t}\left(e^{(x+\lambda(a))t} - 1 - (x+\lambda(a))\,t\right) < 0$. Therefore, $f(r;t) > f(\gamma;t)$ for $\gamma > r > 0$.

(iv) Denote $f(t;a) = \frac{c}{\gamma+\lambda(a)}\left(1 - e^{-(\gamma+\lambda(a))t}\right) + e^{-(\gamma+\lambda(a))t}p_g$. Then $f'(t;a) = e^{-(\gamma+\lambda(a))t}(\gamma+\lambda(a))\left(\frac{c}{\gamma+\lambda(a)} - p_g\right) < 0$ if $p_g > \frac{c}{\gamma+\lambda(a)}$. □

Lemma 8. *Under the optimal mechanism, then $t_g \geq t_b$. If the optimal effort is positive, then $t_g > t_b$.*

Proof. From the IC for the b-type in (17), we have

$$\frac{c}{\gamma+\lambda_b}\left(1 - e^{-(\gamma+\lambda_b)t_g}\right) + e^{-(\gamma+\lambda_b)t_g}p_g \leq \frac{c}{\gamma+\lambda_b}\left(1 - e^{-(\gamma+\lambda_b)t_b}\right) + e^{-(\gamma+\lambda_b)t_b}p_b,$$

which implies:

$$p_g - \frac{c}{\gamma+\lambda_b} \leq e^{(\gamma+\lambda_b)(t_g-t_b)}\left(p_b - \frac{c}{\gamma+\lambda_b}\right).$$

From feasibility constraint (21) in Definition 2, we have $p_g - \frac{c}{\gamma+\lambda_b} \geq p_g - \frac{c}{\gamma+\lambda(a)} \geq p_b - \frac{c}{\gamma+\lambda_b} \geq 0$. If $p_b = \frac{c}{\gamma+\lambda_b}$, then the above inequality and the feasibility constraint together imply $p_g = \frac{c}{\gamma+\lambda_b}$. Otherwise, if $p_b > \frac{c}{\gamma+\lambda_b}$, then the above inequality implies $e^{(\gamma+\lambda_b)(t_g-t_b)} \geq \frac{p_g - \frac{c}{\gamma+\lambda_b}}{p_b - \frac{c}{\gamma+\lambda_b}} \geq 1$ or $t_g \geq t_b$.

If $a > 0$, then the derivation implies $e^{(\gamma+\lambda_b)(t_g-t_b)} > 1$ or $t_g > t_b$. □

Lemma 9. *Under the optimal mechanism with positive effort (i.e., $a^* > 0$), then the following statements hold:*

(i) $p_b > \frac{c}{\gamma+\lambda_b}$ *and* $p_g > \frac{c}{\gamma+\lambda(a)}$.

(ii) The IC for g-type revelation cannot bind.

Proof. (i) From the feasibility constraint (21) in Definition 2, we have $p_b \geq \frac{c}{\gamma+\lambda_b}$. Suppose $p_b = \frac{c}{\gamma+\lambda_b}$ instead. Then from the IC for the b-type in (17), we have

$$\frac{c}{\gamma+\lambda_b}\left(1-e^{-(\gamma+\lambda_b)t_g}\right)+e^{-(\gamma+\lambda_b)t_g}p_g \leq \frac{c}{\gamma+\lambda_b}\left(1-e^{-(\gamma+\lambda_b)t_b}\right)+e^{-(\gamma+\lambda_b)t_b}p_b = \frac{c}{\gamma+\lambda_b},$$

implying $p_g \leq \frac{c}{\gamma+\lambda_b}$, which contradicts with the feasibility constraint (21): $p_g \geq \frac{c}{\gamma+\lambda(a)} > \frac{c}{\gamma+\lambda_b}$ if $a > 0$.

As a result, $p_b > \frac{c}{\gamma+\lambda_b}$, which together with the feasibility constraint (21) implyg $p_g - \frac{c}{\gamma+\lambda(a)} \geq p_b - \frac{c}{\gamma+\lambda_b} > 0$ or $p_g > \frac{c}{\gamma+\lambda(a)}$.

(ii) Suppose the IC for g-type binds. From the binding IC for the g-type, we have

$$u_a^g(t_b) + e^{-(\gamma+\lambda(a))t_b}p_b = u_a^g(t_g) + e^{-(\gamma+\lambda(a))t_g}p_g$$

From Lemma 7(iv) and $t_g > t_b$ proved in Lemma 8, we have

$$u_a^g(t_g) + e^{-(\gamma+\lambda(a))t_g}p_g > u_a^g(t_b) + e^{-(\gamma+\lambda(a))t_b}p_g.$$

Thus the above two inequalities imply $p_g < p_b$. However, from the feasibility constraint (21), $p_g \geq p_b - \frac{c}{\gamma+\lambda_b} + \frac{c}{\gamma+\lambda(a)} \geq p_b$. Contradiction. □

Lemma 10. *Under the optimal mechanism, the PC of investors binds.*

Proof. From Lemma 8, we have $t_g \geq t_b$. We now prove that the PC of investors in equation (19) must bind. Assume not. If the PC of investors is slack, we can increase prices as follows: $p_b' = p_b + e^{(\gamma+\lambda_b)t_b}\epsilon$, $p_g' = p_g + \frac{e^{(\gamma+\lambda(a))t_g}\epsilon}{1-a\lambda'(a)t_g} \leq p_g + e^{(\gamma+\lambda(a))t_g}\epsilon$, where $\epsilon > 0$ is small enough that the investors' PC continues to hold. It is easy to check that the IC for effort (25) holds at the optimal mechanism effort level a^* when evaluated with the new policies.

Furthermore, the IC for the b-type is relaxed, because $\epsilon - e^{-(\gamma+\lambda_b)t_g}\frac{e^{(\gamma+\lambda(a))t_g}}{1-a\lambda'(a)t_g}\epsilon \geq \epsilon - e^{(\lambda(a)-\lambda_b)t_g}\epsilon \geq 0$. Also, from Lemma 9(ii), the IC for the g-type is slack, and it continues to be slack for small enough ϵ. These new prices increase the value to the manager and do not affect the other constraints. Contradiction. □

Lemma 11. *Under the optimal mechanism, the IC constraint of the b-type binds.*

Proof. Assume that, under the optimal mechanism $\{a^*, t_g, t_b, p_g, p_b\}$, the IC for the b-type does not bind.

First, note that if $t_g = t_b = 0$ or if $a^* = 0$, then both ICs for type revelation bind. Thus, it must be that $a^* > 0$ and that either $t_g \neq 0$ or $t_b \neq 0$. From the proof of Lemma 10, we know $t_g > t_b$ if $a^* > 0$. Thus it must be true that $t_g > 0$.

Given the slack IC for the b-type, let

$$\begin{aligned} e^{-(\gamma+\lambda_b)t_b}p_b' &= e^{-(\gamma+\lambda_b)t_b}p_b - a\epsilon, \\ e^{-(\gamma+\lambda(a))t_g'}p_g' &= e^{-(\gamma+\lambda(a))t_g}p_g + (1-a)\,\epsilon, \\ t_g' &= t_g - \epsilon'. \end{aligned}$$

It implies that $p_g' = p_g - (\gamma+\lambda(a))\,p_g\epsilon' + (1-a)\,e^{(\gamma+\lambda(a))t_g}\epsilon$.

We choose the pair $\{\epsilon, \epsilon'\}$ such that the IC for effort (25) holds at the optimal mechanism effort level a^* when evaluated with the new policies:

$$\left\{ \begin{array}{c} u^g_{a^*}(t'_g) - u^b(t_b) + \left(e^{-(\gamma+\lambda(a^*))t'_g} p'_g - e^{-(\gamma+\lambda_b)t_b} p'_b\right) + a^* \times \\ \left[-\frac{c\lambda'(a^*)}{(\gamma+\lambda(a^*))^2}\left(1 - e^{-(\gamma+\lambda(a^*))t'_g}\right) + \frac{c\lambda'(a^*)t'_g}{\gamma+\lambda(a^*)} e^{-(\gamma+\lambda(a^*))t'_g} - e^{-(\gamma+\lambda(a^*))t'_g} \lambda'(a^*) t'_g p'_g \right] \end{array} \right\} = C'(a^*).$$

Therefore,

$$-ce^{-(\gamma+\lambda(a^*))t_g}\epsilon' + \epsilon + a^* \left[\begin{array}{c} \frac{c\lambda'(a^*)}{\gamma+\lambda(a^*)} e^{-(\gamma+\lambda(a^*))t_g}\epsilon' \\ -\frac{c\lambda'(a^*)}{\gamma+\lambda(a^*)} e^{-(\gamma+\lambda(a^*))t_g}\epsilon' + \frac{c\lambda'(a^*)t_g}{\gamma+\lambda(a^*)} e^{-(\gamma+\lambda(a^*))t_g}(\gamma+\lambda(a^*))\epsilon' \\ -\lambda'(a^*)\left[-e^{-(\gamma+\lambda(a^*))t_g} p_g \epsilon' + (1-a^*) t_g e^{-(\gamma+\lambda(a^*))t_g}\epsilon\right] \end{array} \right] = 0,$$

or

$$\begin{aligned} & \left[c - a^*\lambda'(a^*)\left(\frac{c}{\gamma+\lambda(a^*)} + ct_g + \left(p_g - \frac{c}{\gamma+\lambda(a^*)}\right)\right)\right] e^{-(\gamma+\lambda(a^*))t_g}\epsilon' \\ = & \left[1 - a^*(1-a^*)\lambda'(a^*) t_g e^{-(\gamma+\lambda(a^*))t_g}\right]\epsilon. \end{aligned}$$

Since $\lambda'(a^*) \leq 0$, the coefficients of ϵ' and ϵ in the equation above are both positive, implying: $\epsilon' > 0$.

Note that we can keep ϵ and ϵ' sufficiently small so that the IC for the b-type remains relaxed, and the IC for the g-type is further relaxed (Lemma 7(iii)). Furthermore, the new policies also make the PC for investors slack now as long as $v^g_a \geq p_g$, and strictly increase the objective function, while the constraints remain satisfied. Contradiction. □

Lemma 12. *The IC for effort under the optimal mechanism can be rewritten as the equation (23).*

Proof. The binding IC for the b-type (by Lemma 11) imposes a condition on the difference between the present value of payments at $t = 0$:

$$e^{-(\gamma+\lambda_b)t_g} p_g - e^{-(\gamma+\lambda_b)t_b} p_b = u^b(t_b) - u^b(t_g).$$

Plugging the binding IC of the b-type into the IC for effort resulting from the first-order approach (25), we obtain the equation (23):

$$\begin{aligned} 0 &= u^g_a(t_g) - u^b(t_b) + \left(e^{-(\gamma+\lambda(a))t_g} p_g - e^{-(\gamma+\lambda_b)t_b} p_b\right) \\ &\quad + a\left[\left.\frac{\partial u^g_{\widehat{a}}(t_g)}{\partial \widehat{a}}\right|_a - e^{-(\gamma+\lambda(a))t_g}\lambda'(a) t_g p_g\right] - C'(a) \\ &= u^g_a(t_g) - u^b(t_g) + \left(e^{-(\gamma+\lambda(a))t_g} - e^{-(\gamma+\lambda_b)t_g}\right) p_g \\ &\quad - a\lambda'(a)\left[\frac{c}{(\gamma+\lambda(a))^2}\left(1 - e^{-(\gamma+\lambda(a))t_g}\right) - \frac{c\lambda'(a) t_g}{\gamma+\lambda(a)} e^{-(\gamma+\lambda(a))t_g} + e^{-(\gamma+\lambda(a))t_g} t_g p_g\right] - C'(a), \end{aligned}$$

where $u^g_{\widehat{a}}(t_g) = \frac{c}{\gamma+\lambda(\widehat{a})}\left(1 - e^{-(\gamma+\lambda(\widehat{a}))t_g}\right)$ and

$$\left.\frac{\partial u^g_{\widehat{a}}(t_g)}{\partial \widehat{a}}\right|_a = -\frac{c\lambda'(a)}{(\gamma+\lambda(a))^2}\left(1 - e^{-(\gamma+\lambda(a))t_g}\right) + \frac{c\lambda'(a) t_g}{\gamma+\lambda(a)} e^{-(\gamma+\lambda(a))t_g}.$$

□