Platform Screening, Investor Learning, and Default Risk in Marketplace Lending

Jin-Hyuk Kim* Peter Newberry† Nick Vikander‡

Abstract

We theoretically and empirically examine the role of platform screening and investor learning in an online platform where small businesses raise funding from the public by offering investment returns. The platform decides how much to screen potential borrowers, affecting the level of default risks for funded loans. The investor uses observed default rates as an imperfect signal of the platform's reputation when they make investment decisions. We show that, under certain conditions, the platform's screening incentive increases if the investors can observe outcomes of previously funded loans, which would lead to lower default rates. Using longitudinal data from a marketplace lender as well as expert evaluation on the posted fundraising campaigns, we show that platform-level outcomes simulated under different informational environment are consistent with our theoretical predictions.

Keywords: default; fintech; investment; monitoring; screening; transparency

*Department of Economics, University of Colorado at Boulder, 256 UCB, Boulder, Colorado 80309, USA. E-mail: jinhyuk.kim@colorado.edu.

†Department of Economics, University of Georgia, 620 S Lumpkin St, Athens, Georgia 30602, USA. E-mail: pwnewberry@uga.edu.

‡Department of Economics, University of Copenhagen, Øster Farimagsgade 5, 1353 Copenhagen, DNK. E-mail: nick.vikander@econ.ku.dk.

The expert survey data used in this paper were funded by NBER Digitization small grant. We have no additional items to disclose.

1 Introduction

Financial intermediation has been performed by traditional financial institutions such as commercial and investment banks and non-bank institutions such as private equity firms. Beginning in the 21st century, a new kind of financial intermediary emerged in the form of online platforms, directly inviting individuals (or institutions) to invest in various fundraising projects (or loans) for monetary returns. They are known as peer-to-peer or marketplace lending and have quickly captured a share of the credit market for consumers and small businesses. The estimated global market size of marketplace lending is $517 billion (Future Market Insights, 2024).

From the investor's standpoint, marketplace lending may offer higher returns than standard bank deposits do, but it can also come with higher or unknown risks. In particular, marketplace lenders often do not disclose how many borrowers applied and how many were subsequently screened out or allowed to launch their fundraising campaigns. Thus, investors have relatively little information to infer the quality of the campaigns or the platform's screening. This trust (or reputation) problem can be especially important for platforms facilitating small business loans, if businesses are more heterogeneous and harder to evaluate than individuals' creditworthiness.

We theorize and empirically examine a mechanism through which investors can infer the quality of platform screening—the default outcomes of previously funded projects. We model the repeated interaction between a platform and a representative investor and characterize the perfect public equilibrium in which the investor can punish the platform by walking away if a default outcome is observed. We show that the platform's screening incentive decreases under certain conditions if the investor cannot observe the outcome of the previously

funded projects. We then estimate equations that describe individuals' investment decisions and simulate fundraising outcomes at the platform level in the counterfactual informational environment in which investors do not observe outcomes of the previously funded projects. Our simulation reveals that the platform would plausibly lower the screening level in such an environment.

Notice that the platform's screening which determines the overall quality of the fundraising campaigns is different from the individual investor's screening of the posted campaigns based on the project description and financial terms. While the literature has often highlighted the latter, it implicitly assumes that platforms do some level of screening or at least it does not explicitly consider how the findings would depend on the platform screening. Larger platforms often have policy statements on screening; for instance, LendingClub (2024) says "The most common reasons for rejection include a low credit score or bad credit history, a high debt-to-income ratio, unstable employment history, too low of income for the desired loan amount, or missing important information or paperwork." However, visibility into screening practices can be poor for investors especially in smaller platforms dealing with small business loans.

Thus, we focus on the role of a marketplace lender in screening or certifying the fundraising campaigns, which in turn affects the investors' decisions. The traditional certification hypothesis in financial markets comes from the securities underwriters maintaining their reputational capital in marketing risky assets (e.g., Booth and Smith, 1986; Chemmanur and Fulghieri, 1994), which has received empirical support. For instance, Benzoni and Schenone (2010) find that Initial Public Offerings (IPOs) underwritten by relationship banks perform just as well as those underwritten by non-relationship underwriters. Duarte-Silva (2010) shows that an underwriter is in fact better able to certify an equity issue if it has a lending relationship with the issuer.[1]

[1]However, Sette (2011) finds that competition mitigates the incentives of underwriter-affiliated analysts to issue favorable recommendations. Additionally, Bolton et al. (2012) show that credit rating agencies tend to inflate ratings when more investors take the ratings at face value.

In a two-sided market extension, Bouvard and Levy (2017) consider the role of a certifier in a retail marketplace, where the certifier has a screening technology to imperfectly detect bad-quality goods. It is shown that the certifier's profit is an inverted U-shaped function of its reputation, because buyers are wary of the seller's goods at too low a reputation while too few sellers would get certified at too high a reputation. Hence, the certifier would want to increase the precision of their screening technology in the former and decrease it in the latter.[2] Our theoretical contribution is that we analyze the platform's incentives to monitor funded projects as a payment collector-enforcer as well as screen funding applications in a repeated interaction setting.

There is also a growing literature on screening. For instance, Cumming et al. (2019) survey Canadian crowdfunding platforms for factors that are associated with due diligence and find that the platforms applying due diligence tend to have more successful campaigns. Yannelis and Zhang (2023) show that lenders that have lower market shares tend to have lower incentives to invest in fixed screening technology. On the other hand, Daley et al. (2024) find that, in merger agreements, the acquirer takes more time and effort to learn about the seller, the higher the transaction price is; and this leads to a higher probability of deal failure. We add to this literature by examining how the platform's screening may depend on the information the investors have access to, using granular longitudinal data from an online platform.

There are papers analyzing the borrowers' asymmetric information and moral hazard problems in marketplace lending.[3] For instance, Kawai et al. (2022) consider borrowers' default costs and show that borrowers can signal their risk type through the interest rate they offer; Xin (2023) studies repeat borrowers with multiple loans and shows that the market may collapse without observing the borrower's history. While these two studies use loan-

[2]This tradeoff is similar to that of Beatty and Ritter (1986), in which an investment banker who prices an IPO issue too high will lose potential investors in the future IPOs while it will lose potential issuers if it underprices an IPO too much, forfeiting the value of its reputation capital.

[3]The literature has also examined the ability of individual investors to screen fundraising campaigns (e.g., Iyer et al., 2015; Mollick and Nanda, 2015) and whether investors exhibit biases in their decision-making (e.g., Zhang and Liu, 2012; Agrawal et al., 2015; Åstebro et al., 2024).

level data from a large marketplace lender, we use user-level data from a smaller marketplace lender, in which the number of funded loans is much smaller than that in Prosper, so default outcomes play a more salient role in determining the investor behavior. Hence, our findings on the implications of the information environment or platform disclosure policy may shed more light on startup platforms.

The platform we empirically study allows businesses to launch campaigns for what is known as purchase order or inventory financing, which finances the production of retail goods for specific buyers or general markets. Gindin (2022) analyzes a signaling mechanism in which firms propose a loan size and repayment terms based on the level of consumer demands and finds that credit-constrained firms may get larger loans and better terms with this type of financing compared to the traditional debt financing. While this mechanism influences the demand (borrower) side, our study focuses on the supply (investor) side specifically with respect to platform-intermediated loans.

Finally, our paper is related to the broader literature on reputation. For instance, Board and Meyer-ter-Vehn (2013) show that a firm may disinvest in product quality at high levels of reputation, while it invests at low levels of reputation. Bar-Isaac and Deb (2014) find that reputation concerns are more effective when players observe common outcomes rather than separate outcomes. Nosko and Tadelis (2015) show that buyers on eBay, a retail marketplace, may draw inferences about the overall quality of the platform from their transactions, creating a reputational externality among sellers. We contribute by examining how investors draw inferences about platform screening.

The remainder of this paper is organized as follows. Section 2 presents a game model between the platform and the investor, and Section 3 characterizes the equilibrium under different informational environment. Section 4 describes the dataset, and Section 5 contains the estimation of investment decisions. Section 6 shows the simulation outcomes based on the estimation results, and Section 7 concludes.

2 Model

There is a platform that lets borrowers raise funds from investors. We take the borrowers' funding applications exogenously given, and the platform's commissions are fixed at $R > 0$ per funded loan. Before the game starts, the platform chooses a screening level or technology $\alpha \in [0, 1]$ at a convex cost $C(\alpha)$.[4] We also assume that the screening technology remains constant throughout the game (or the time period under consideration). This greatly simplifies our analysis of the platform's moral hazard problem in monitoring funded loans à la Fudenberg et al. (1994).

Once the game starts, time is discrete and infinite ($t = 1, 2, ...$). In each period t, a borrower arrives at the platform. The platform then decides whether to allow the borrower to launch a project on the platform, using the screening technology. Specifically, the borrower passes the screening with a probability $1-F(\alpha)$ while he fails the screening with a probability $F(\alpha)$, where $F(\alpha)$ is a distribution function defined on $[0, 1]$. Let $h(\alpha) \equiv f(\alpha)/(1 - F(\alpha))$ denote the hazard rate. For uniqueness of solution, we assume this hazard rate is increasing: $h'(\alpha) > 0$ for all $\alpha \in [0, 1]$.

A project that passes the screening is then presented to a representative investor, who decides whether or not to fund the project. (The representative investor can be interpreted as a continuum of investors rather than a single individual.) If the investor funds the project, the platform chooses whether or not to monitor, as described in more detail below. We denote the platform's cost of monitoring by $E > 0$. The project then either succeeds or fails. Success implies that the borrower repays the investor at the end of the period, while failure implies that the borrower defaults and the investor is not repaid.

The probability of repayment or default depends on potential liquidity shocks, platform screening, and platform monitoring. Specifically, a liquidity shock occurs to the borrower with probability $p > 0$, in which case the borrower always defaults. Conditional on a shock

[4]The cost function $C(\alpha)$ is not a fundamental part of our analysis, because it is an upfront fixed cost. It predicts, however, that the platform would choose an α in a segment on which the marginal revenue is still increasing in α, as we show in our simulation (see also Figure 1).

not occurring, repayment occurs with probability α, and default with probability $1-\alpha$, if the platform monitors. The interpretation is that tighter screening leads to a higher success probability after being funded. If the platform does not monitor, however, default occurs with probability 1.

Monitoring (or contract enforcement) is necessary to avoid default, even in projects that would otherwise be successful. In the absence of monitoring, borrowers may falsely claim that a shock occurred or simply fail to respond altogether. This issue is particularly relevant in our empirical context, whereby investors sometimes voice complaints (such as on social media or blog posts) about the platform's perceived lack of monitoring. While the platform asserts that it enforces contracts and actively monitors borrowers, investor concerns suggest that the credibility of enforcement efforts is important.

We assume the representative investor uses the following type of a trigger strategy: The investor begins by investing in the first period in which a project is posted on the platform. Thereafter, the investor's strategy depends on whether the borrower repays or defaults on the funded loan. If the borrower repays, the investor is willing to invest in the next period. If the borrower defaults, however, the investor may refuse to invest thereafter and walk away from the platform for good. Thus, default means that the platform risks being unable to fund projects in all future periods.

Let us denote the probability that the investor walks away for good after observing default by $r > 0$, the value of which will be determined in equilibrium. We focus on the constrained-efficient equilibrium, in which the platform monitors funded projects (unless the investor has walked away for good) and the investor's punishment probability r is the smallest one supporting the equilibrium. We assume a per-period discount factor $\delta \in (0,1)$ for the platform, or equivalently, an exogenous probability $1-\delta$ that the investor walks away at the end of each period.

We start by assuming that the investor cannot observe the platform's choice of screening technology α, and hence cannot condition their punishment probability r on this chosen α.

Instead, they base r, in part, on their conjecture of the chosen α. This would correspond to the case in which the platform does not disclosure past project outcomes. We then extend the model by assuming that the platform's choice of α is observable, so this choice can affect the punishment probability. Comparing results in these two environments allows us to examine how platform transparency affects equilibrium outcomes.

3 Analysis

Given the representative investor's trigger strategy, there are two relevant states of world when presented with a project in a given period t. One is where the borrower has defaulted in period $t-1$. The investor then walks away from the platform with probability r, and invests with probability $1-r$. The other is where there was no default in period $t-1$, so the investor will invest with probability 1.

Thus, investment occurs in period t if (i) the project passes the screening, which occurs with probability $1-F(\alpha)$, and if (ii) either there was no default at $t-1$, or there was default but the investor did not walk away. If the project does not pass the screening in period t, then there is no period t-investment or default, and play moves on to period $t+1$.

If the investor walks away from the platform for good after default in some period $t-1$, any future posted project cannot be funded, so it is clearly optimal for the platform not to monitor as of period t.

We now derive the platform's value function in a setting where the platform monitors all funded projects. Let V denote the present discounted value of the platform's profit streams starting from a period t where the investor is willing to invest. We have

$$(1) \qquad V = \Big[R - E + \big[(1-p)\alpha + [1-(1-p)\alpha](1-r)\big]\delta V\Big][1-F(\alpha)] + F(\alpha)\delta V,$$

where rearranging yields

$$V(\alpha, r) = \frac{R - E}{\frac{1-\delta F(\alpha)}{1-F(\alpha)} - \delta\Big[(1-p)\alpha + [1-(1-p)\alpha](1-r)\Big]}. \tag{2}$$

This value depends, in particular, on the screening technology α and the punishment probability r.

The value function $V(\alpha, r)$ is non-monotonic in α. Intuitively, there are two opposing forces at play. Tighter screening reduces the probability that projects are presented (and hence funded). However, conditional on being presented (and funded), tighter screening yields lower default probability, increasing the probability that the investor will fund future projects.

We now show $V(\alpha, r)$ is strictly unimodal in α, for any given r. The second derivative of the denominator in expression (2) for $V(\alpha, r)$, taken with respect to α, is

$$\frac{(1-\delta)[h'(\alpha) + h(\alpha)^2]}{1 - F(\alpha)},$$

where $h(\alpha) \equiv f(\alpha)/(1-F(\alpha))$. This second derivative is strictly positive by our assumption of an increasing hazard rate, $h'(\alpha) > 0$. This denominator is therefore convex in α and has a unique minimum. Thus, for any given r, $V(\alpha, r)$ is strictly unimodal in α, with a unique (local and global) maximizer which we denote by

$$\alpha_r \equiv \arg\max_{\alpha} \; V(\alpha, r).$$

Without loss of generality, we can restrict attention to candidate equilibria with $\alpha_r \in (0, 1)$. As we will see below, $\alpha = 0$ would violate the platform's incentive constraint to monitor, whereas $\alpha = 1$ would directly yield zero value (since no project would be presented).

The maximizer α_r is therefore defined by the first order condition

$$\frac{\partial}{\partial \alpha} V(\alpha, r) = 0,$$

or equivalently, from expression (2),

$$\frac{d}{d\alpha}\left[\frac{1-\delta F(\alpha)}{1-F(\alpha)}\right] - \delta(1-p)r = 0. \tag{3}$$

By unimodality, $V(\alpha, r)$ is increasing in α for all $\alpha < \alpha_r$ and decreasing in α for all $\alpha > \alpha_r$. Moreover, the derivative $\partial V(\alpha, r)/\partial \alpha$ is increasing in r, so a higher punishment probability is associated with a higher optimal screening level; that is, α_r is increasing in r. This is because tighter screening reduces expected default rates, which is particularly important for the platform when investors react more unfavorably to default. We also note from (3) that α_r will depend on parameters δ and p, but not on R and E.

On the other hand, $V(\alpha, r)$ is decreasing in r. Intuitively, for a given screening technology α, conditional on investors participating, a high punishment probability following default must hurt the platform, by reducing the expected number of periods with funded projects.

Since the platform's gross receipt R is generated when a project is funded, independent of whether the borrower repays or defaults on the loan, the platform may have an incentive not to monitor funded projects, to save on the monitoring cost E. This would lead to default, so the equilibrium must satisfy the following platform incentive constraint (IC):

$$V \geq \Big[R + (1-r)\delta V\Big][1 - F(\alpha)] + F(\alpha)\delta V. \tag{4}$$

If condition (4) is violated, then the platform will find it optimal not to monitor, and investors will therefore not participate.

Using expression (1) to substitute for V in condition (4) and simplifying yield

$$(1-p)\alpha r\big[\delta V(\alpha, r)\big] \geq E. \tag{5}$$

The intuition is that monitoring costs E for the platform, but not monitoring increases the probability of the investor walking away as of the next period by $(1-p)\alpha r$, where $(1-p)\alpha$ is the increase in default probability. The discounted expected loss for the platform if the investor walks away is $\delta V(\alpha, r)$.

The platform's program is to maximize the net present value $V(\alpha, r)$ by choosing a screening level α, given the investor punishment probability r. As we look for a constrained-efficient equilibrium, we assume that the investor punishment probability r makes the platform IC, given by condition (5), bind, given investor beliefs about the platform's chosen screening technology α. It follows from (5) that the equilibrium screening technology α must be strictly positive, since $\alpha = 0$ would give the platform no incentive to monitor.

We can show that, for any given $\alpha > 0$, a higher investor punishment probability r increases the platform's incentive to monitor, whereby we obtain the following result. All proofs can be found in the Appendix.

Lemma 1 *Fix E, p, δ, and α. There is a unique value $R(\alpha)$ such that, for any $R \geq R(\alpha)$, there is a unique value $r(\alpha) \in (0, 1]$ such that condition (5) holds if and only if $r \geq r(\alpha)$.*

We will focus on values of the gross platform receipt R that are sufficiently large to exceed $R(\alpha)$ for all relevant values of α. Then, in a constrained-efficient candidate equilibrium with screening technology α, the investor punishment probability will be $r(\alpha) \in (0, 1)$, which is the unique punishment probability that makes the platform IC bind. For this candidate to be an equilibrium, it remains to check whether α indeed maximizes platform value, given investor punishment behavior, that is, whether $\alpha = \alpha_r$ from condition (3).

The equilibrium screening value is the value of α that satisfies condition (3) evaluated at $r = r(\alpha)$. Denote this equilibrium screening level by α^{Un}, where the superscript 'Un' stands

for 'Unobservable'. We now show that equilibrium monitoring behavior depends on whether the gross platform receipt R exceeds $R(\alpha_1)$, the threshold value from Lemma 1 evaluated at $\alpha_1 = \arg\max_\alpha \ V(\alpha, 1)$.

Proposition 1 *If $R \geq R(\alpha_1)$, then there is a unique equilibrium screening probability α^{Un} and punishment probability $r(\alpha^{Un})$ in which the platform monitors. If instead $R < R(\alpha_1)$, then there is no equilibrium in which the platform monitors.*

Intuitively, if the gross platform receipt R is sufficiently high, then the constrained-efficient equilibrium involves monitoring. The platform sets the screening level that is optimal given equilibrium investor behavior. This behavior makes the platform IC bind, given the equilibrium screening level. If instead R is too close to E, then the platform has no incentive to monitor in equilibrium, even if investors would always walk away after default ($r = 1$).

To show how platform transparency in the choice of screening technology would affect equilibrium outcomes, we now modify our framework by allowing the investor to observe the platform's choice of α before he chooses a punishment probability r. This means that the punishment probability r can now be conditioned on the platform's chosen α.

We focus on parameter values such that $R \geq R(\alpha_1)$ from Lemma 1, so an equilibrium exists where the platform monitors. Denote the equilibrium screening level by α^{Ob}, where the superscript 'Ob' stands for 'observable'. Recall from Lemma 1 that $r(\alpha)$ is the unique punishment probability r that makes the platform IC bind, given α. Then we have

$$\alpha^{Ob} \equiv \arg\max_{\alpha} \ V(\alpha, r(\alpha)),$$

with corresponding equilibrium punishment probability $r^{Ob} \equiv r(\alpha^{Ob})$.

Thus, α^{Ob} must satisfy $\frac{d}{dr}V(\alpha, r(\alpha)) = 0$, or

$$\frac{\partial}{\partial \alpha}V(\alpha, r(\alpha)) + \left\{ \frac{\partial}{\partial r}V(\alpha, r)|_{r=r(\alpha)} \frac{dr(\alpha)}{d\alpha} \right\} = 0,$$

or equivalently,

$$\frac{d}{d\alpha}\left[\frac{1-\delta F(\alpha)}{1-F(\alpha)}\right] - \delta(1-p)r(\alpha) + \left\{\delta[1-(1-p)\alpha]\frac{dr(\alpha)}{d\alpha}\right\} = 0. \tag{6}$$

The last term in large brackets was absent from the corresponding expression with unobservable screening technology (i.e., compare conditions (6) and (3)). The difference is that, with observable screening technology, the platform takes into account how its choice of screening can influence the investor's punishment probability (i.e., $dr(\alpha)/d\alpha \neq 0$).

The following result shows that observability of the platform's chosen screening level results in tighter equilibrium screening and a lower equilibrium punishment probability, compared to those under unobservability.

Proposition 2 *Suppose that the platform's choice of screening technology is observable and that $R \geq R(\alpha_1)$. Then there exists an equilibrium with screening technology α^{Ob} and punishment probability α^{Ob}, in which the platform monitors. In any such equilibrium, $\alpha^{Ob} > \alpha^{Un}$ and $r^{Ob} < r^{Un}$.*

The intuition behind this result is that an increase in the screening technology makes it more costly for the platform to forego monitoring. This is because projects presented on the platform are likely to succeed and repay investors (conditional on monitoring) when screening is tight. These projects will instead fail if the platform fails to monitor. Thus, if investors were to observe tighter screening, they can employ a lower punishment probability r and still provide incentives for monitoring.

As a result, the platform will increase its screening value α above the level that would be optimal if screening was unobservable. The direct impact of this increase reduces the value function ($\partial V(\alpha, r(\alpha))/\partial\alpha < 0$). Doing so is nonetheless optimal, because it makes the investor less likely to walk away after default, which indirectly increases value ($\partial V(\alpha, r)/\partial r|_{r=r(\alpha)}$ $dr(\alpha)/d\alpha > 0$). The result is projects that default less frequently, and investors who walk away less often conditional on default.

In that case, the platform may find it profitable to disclose its choice of α, or facilitate its discovery by investors. The platform could always do just as well by setting $\alpha = \alpha^{Un}$, to earn the equilibrium (gross) profits associated with Proposition 1. But it could do strictly better by disclosing $\alpha = \alpha^{Ob}$, to earn the profits associated with Proposition 2, as long as the increase in the fixed cost of a more stringent screening technology is not high enough to offset the increase in profits.

Thus, the main empirical prediction of our theory is that the platform screening would be tighter in an environment where investors can observe the platform's screening level. In practice, investors may not directly observe the platform's screening. Some platforms may post statements on financial metrics to clarify the borrower's eligibility criteria, but much discretion is left to be exercised in how fledgling platforms select which projects to launch, in particular, if the platform is not actively regulated and the borrowers are small businesses which can be more difficult to vet.

Instead, investors may indirectly learn about the platform's screening level α if they observe the outcomes of previously funded projects. For instance, in a transparent environment, the platform can archive the outcomes of all previously funded projects on its website. Then, investors arriving at the platform can learn about the platform's screening level (or update belief about its type) by observing the platform's track records (e.g., the proportion of past project defaults). This would work through the usual mechanism in a repeated game setting, where the investors would be able to learn the platform's screening level accurately over time (Fudenberg and Levine, 1989, 1992).

Once their learning is complete, the investors can invest in a first project and then decide whether to invest in a second project depending on whether the first defaults, and so forth, which is what our model captures. Notice that in the transparent environment, we assume that new investors do not punish the platform in the present, by potentially walking away, for defaults that occurred in the past. Rather, past outcomes are just used to learn about the platform's screening level. Thus, in this information environment, the investors and the

platform essentially play the game with observable α.

In contrast, in an opaque environment, the platform may not disclose all or part of the outcomes of previously funded projects, so the investors cannot observe them. The investors and the platform would then play the game with unobservable α. With this interpretation, our theory predicts that the platform's choice of screening level (i.e., one that maximizes the platform's profit streams) would be lower in the environment where the past project outcomes are not observable, because the platform no longer faces a quantity-quality trade-off in screening projects. In what follows, we will examine whether observed data are consistent with our theoretical prediction.

4 Data

Our data come from an online marketplace lender for small retail businesses.[5] We were given direct access to anonymized data on all user accounts, including all transaction activities recorded during multiple (consecutive) quarters in the mid-2010s, public-facing data on all projects created on the platform during the same period of time, and then-current user and expert survey data collected by the platform and us, respectively.

On the platform, businesses apply for short-term debt-like financing to build inventories and/or fulfill purchase orders. The platform has staff who interview the businesses, review their track records, and decide whether to allow them to post their projects on the platform. A posted project includes an introduction of the business and its products and information on the terms of the financial arrangement such as a funding goal, an interest rate, and a repayment term (e.g., months).

Specifically, once the project is funded, periodic payments are due from the borrower on pre-specified dates. For those projects in arrears, some may make late payments bringing the balance to zero, but more often borrowers just stop paying. We observe the date and

[5]Except upon request, we cannot disclose the identity of the platform to protect their sensitive information. Also, we may need to suppress some identifying information in return for making the data available for our study. The authors retained full independence in this research.

the amount of each investment flowing from, and repayment flowing into, each user account along with a project ID, so we can construct the share of the projects that have a payment overdue for, say, two weeks or one month for each user as well as at the platform level for each day in our sample.

Thus, we observe whether and how much a given user invests in each of the projects that were launched during that user's tenure, where tenure is defined as the time between the user's account creation and the last login date. We do not observe if the user browsed each project, but we assume that the user considered every project that was open during his tenure.[6] Since projects became available only periodically and there were at most few projects at the same time, this assumption may be reasonable.

After cleaning, our dataset comprises 6101 users, 312 projects, 248 borrowers, and around 350,000 user-project observations. Table 1 shows project-level summary statistics. The average project is considered by about 1100 users and receives funding from just over 60 of them for a total of about $21,000. The distribution of total investment amounts is skewed to the right with a median around $9500.

There are times when a borrower fails to meet the promised payment schedule (similar to an installment payment), which we refer to as the payment being late. The last two rows of Table 1 show that 22% of the projects have been at least 2 weeks late at some point during their payment schedule, while 19% have been at least 1 month late. In such cases, investment returns often fall short of the promised rate.

Table 2 shows user-level summary statistics. The average user is presented with 56 projects, invests in three of them for a total of over $1000 during their tenure, which is about 67 days. The users may use the observed late payment events to form an expectation of future investment risks. We thus construct two signals: the share of projects across the platform that have late payments (a 'public' signal) and the share of projects each user has

[6]Over 97% of the platform users have opted in for notification on new funding opportunities, so they were at least notified of the new projects via their preferred contact method. We dropped the 3% opt-out users from the sample; this did not change the estimation results much.

invested in that have late payments (a 'private' signal).[7]

The last four rows of Table 2 show the average values of the public and private signals during the users' tenure. The average share of projects the user has invested in that have payments at least two weeks late is 1.54%. The average is 1.1% for those a month late. The average public signals, on the other hand, are 11.1% and 8.6% for two weeks and a month late, respectively. Notice that the private signal is on average smaller than the public signal because the private signal is 0 when a user first invests, but the platform may already have a non-zero share of troubled projects.

We supplement the data with an expert evaluation on the ex-ante quality of the 312 posted projects. To do this, we hired an underwriting manager from a direct lender (not a marketplace lender) in the same line of business providing inventory and purchase order financing. The expert worked as a credit analyst and underwriter in two previous jobs before joining this financial company, where the person had worked for several years as the lead underwriter when we did this survey in the mid-2010s.[8]

Table 3 show the descriptive statistics for six survey questions. The answers range from 1 to 5 with 0.5 increments, where higher numbers indicate higher likelihoods. For the second and the third questions, the higher numbers indicate more attractive payment terms and riskier investments, respectively. The last question refers to the likelihood of the expert's financial firm lending under the observed payment terms.

5 Estimation

We specify and estimate two investment equations for the users at the extensive and intensive margins, respectively, with a focus on the extent to which public and private signals impact

[7]The public signal is available to users as long as the platform updates and archives the full history of the payment status during the payment period for each project, which was the case for the sample period. The private signal is based on one's own previous investments.

[8]The expert was provided the same information observable by investors on the platform, except that the platform identity was suppressed. The expert was paid at a market rate, and from our communications, it seemed that the expert gave sufficient care and attention to the task.

user behavior. The first equation is a binary choice model of whether or not user i invests in project j on day t, which we denote D_{ijt}, given the value of the private signal, s_{ijt}^{pr}, and the public signal, s_{jt}^{pu}. The hypothesis is that both signals are a function of the level of platform screening and monitoring, so the user may use them to form an expectation of the return on their investment in project j on day t.

Because the private signal is based only on a subset of projects, the public signal may serve as a stronger predictor of quality of any given project and thus impact user behavior to a greater degree. However, because the private signal is based on personal experience, a user may put more weight on it to determine his expectation of project j's riskiness. That is, even though no additional information is learned from a private signal as long as the public signal is available, the private signal could matter if the user's preferences exhibit a bias towards own experience or discount others' experience.

We use the following model to describe the user's investment decision:[9]

$$D_{ijt} = \beta_0 + \beta_1 s_{jt}^{pu} + \beta_2 s_{ijt}^{pr} + \beta_3 Z_{it} + \xi_i + \zeta_j + \chi_t + \epsilon_{ijt}, \tag{7}$$

where Z_{it} includes time-varying characteristics of user i: the log of the number of previous investments (plus one), a dummy for zero previous investment, and the user tenure (number of days on the platform) as of day t. These are included to account for different levels of user experience on the platform. The dummy for zero prior investment on the platform accounts for non-linear effects of no investment. ξ_i, ζ_j, and χ_t are the project-level, user-level, and day-of-week fixed-effects, respectively; and ϵ_{ijt} is an i.i.d. error.

Table 4 displays the estimation results, where the effect of the signals is identified by within-project and within-user variations in investment decisions as the signals change. The result in column (1) indicates that the public signal negatively impacts the user's investment decision. Specifically, if the share of all funded projects on the platform that are late increases

[9]Notice that we cannot estimate the model using nonlinear functional forms such as probit and logit because of the large number of fixed effects. Instead, the linear model for the investment decision will be used for the sample selection model below following Olsen (1980).

by one percentage point, the likelihood that a user invests in a given project decreases on average by about 2.5%, which is relatively large considering that the unconditional average of the investment decision in the sample is around 6%.

Our model does not have a specific prediction on the sign of the control variables. The fact that the number of investment is negatively correlated with investment in a given project suggests that the average investor does not keep reinvesting on the platform, which could be due to the default risk. It seems intuitive that no investment experience is negatively correlated with investment because the user may experience a hurdle prior to start investing. In contrast, the user tenure is positively correlated with investment as it may represent the user's general level of interest in the platform.

Column (2) of Table 4 shows the estimation results including the private signal, which has a negative and statistically significant effect on the investment decision. However, the magnitude of the coefficient is much smaller than that of public signal's. The fact that the public signal dominates implies that the signal (or reputation) at the platform level relative to the user's own experience is an important determinant of the user's investment decision. This also suggests that the average user is 'rational' in that he puts almost the same weight on others' investment experience as he does on his own.

Columns (3) and (4) replicate the two previous columns using the 1-month late signals instead of the 2-week late signals. The question is whether the average user would consider the 2-week signals to be sufficient to form an expectation of the quality of projects on the platform, or wait until a payment is overdue for more than a month to form such an expectation. The estimation results reveal that the 1-month public signal is not a statistically significant factor for the investment decision, while the 1-month private signal is. This implies that 2-week public signals are sufficiently informative to the user, so a repayment after two weeks may not significantly change the user's expectation.

The second equation we estimate is a model of the log investment amount, which we denote I_{ijt}. Because an amount is only observed if the user decides to invest, we need to

correct for the selection bias. We do so by including the expected likelihood of investment, $\hat{D}_{ijt}$, as predicted by equation (7), as an additional covariate. The instrument excluded from regressions of the investment amount is user tenure, where the assumption is that the length of time a user has been on the platform correlates with the investment decision, as we have shown in the above, but not how much the user invests.

The latter assumption can be supported by an observation from the platform's user survey. Closer to the end of our sample period, the platform independently fielded an anonymous survey to the registered users, which included how likely the user is going to use the platform again and their risk tolerance when it comes to alternative investments, both on a 1-to-5 scale. Table 5 shows bivariate correlations between the answers to these two questions and user attributes. For instance, it is intuitive that the years of investment experience are positively correlated with levels of risk tolerance.

Among others, Table 5 indicates that the correlation between household income and the likelihood of using the platform again in the future is not statistically significant, while the correlation between the income and risk tolerance is positive and statistically significant. This implies that a user with a higher income would invest more because he is more risk tolerant, but he is not more or less likely to use the platform again in the future. This also implies that income may not be significantly correlated with user tenure, and thus the investment amount may not be correlated with the user tenure.

We use the following model to describe the amount invested in a project:

$$I_{ijt} = \beta_0 + \beta_1 s_{jt}^{pu} + \beta_2 s_{ijt}^{pr} + \beta_3 \tilde{Z}_{it} + \beta_4 \hat{D}_{ijt} + \xi_i + \zeta_j + \chi_t + \epsilon_{ijt}, \tag{8}$$

where I_{ijt} and $\hat{D}_{ijt}$ are explained above, and $\tilde{Z}_{it}$ does not include the user tenure for the identification purpose. Everything else is the same as that explained after equation (7).

Table 6 displays the estimation results. Across the columns, neither the public nor the private signal impacts the amount invested by users in a statistically significant manner.

This suggests that the default outcomes of the projects impact whether or not users invest, but not how much they invest conditional on deciding to invest in a given project. The negative and statistically significant coefficients on the predicted probability of investment indicate that those who are less likely to invest, if observed to have invested, would have higher investment amounts than the population, so there is a positive sample selection.

6 Simulation

We quantify the effects of different signals and screening stringencies on fundraising outcomes by simulating user behavior over time under different informational environment. To do so, we first introduce our measures of project quality and screening based on the expert's evaluation. The goal is to have such measurements from outside of the platform.

Table 7 shows bivariate correlations between the expert's evaluations and the late payment outcomes.[10] The correlations among the evaluation questions are strong, so it is difficult to isolate the impact of individual questions. However, the degree of riskiness the expert perceived (Q3) is a strong predictor of whether or not there is a late payment.

The likelihood that the business raising the capital will be successful (Q1) is also negatively correlated with the late payment outcomes, but its statistical significance is less than that of Q3. Further, project risk is narrower in scope than business success. We will thus refer Q1 as a broader ex-ante measure of the project quality in our simulations that is not solely determined by the ex-post default outcomes.

We then use Q3 to adjust the screening stringency by allowing only projects with the expert's Q3 rating below a certain threshold value to launch on the platform in our simulations. For instance, allowing projects with a Q3 rating below 5 would lead all but those sample projects with a Q3 rating of 5 to be included in our simulations. Thus, a lower cutoff value would imply more stringent screening.

[10]These correlations can also serve as measures of how well the expert's evaluations match up with the realized outcomes. For instance, the expert's evaluation of how likely the loan will be overdue for more than a month is positively correlated with the actual late outcomes.

To perform the simulation for a given signaling environment and a screening level, we start with the projects on the platform on the first day ($t = 1$) of our sample and the users who are active on that day. We simulate their investment decisions both at the extensive and intensive margins. The investment probability is predicted from specification (2) in Table 4, where we simulate the binary decision based on this probability. The amount of investment is assumed to be at the observed level if a user had invested, or the predicted level of investment from specification (2) in Table 6 if he had not.

We then update the signals based on the users' decisions in $t = 1$ and the observed late payments and move onto the users and the projects that are active on the platform on the second day ($t = 2$). We perform this simulation until $t = T$ and calculate the total investment and investors for each project as well as aggregate investment at the platform level. We run this full simulation 100 times and calculate the average project-level and aggregate investment across simulation runs and compare these outcomes across different environments. We also calculate the aggregate investment for the projects below the average Q1 rating and refer them as investment in low-quality projects.

The simulation results are presented in Table 8. Table 8 contains summary statistics for the simulations in three panels varying in informational environment: i) both signals are observed, ii) no private signal is observed, and iii) no public signal is observed. On the other hand, the first column displays the statistics observed in the data; the second column displays them for the simulation in which all projects in the data are allowed; and the remaining columns display them for different levels of screening, whereby the screening becomes more stringent as the threshold Q3 rating becomes lower.

We first compare column (1) to column (2) in the first panel in order to assess the model fit. Some discrepancies between the model and the data are expected because we are simulating the users' choices based on the models and different choices will lead to different signals, which in turn lead to different choices, and so forth. All projects in the data that reached at least 80% of their funding goal were funded, so we make the assumption that a

project is funded in the simulation if it reaches 80% of the goal. Overall, the model seems to do a good job in matching the data.

We now examine the impact of screening in the first panel. Going from left to right, both the average number of investors and the amount of investment per project increase as the screening gets more stringent. This is because the projects are less likely to default, hence the signals are more positive, as screening eliminates those that tend to have late outcomes. One exception is when the Q3 rating changes from 4.0 to 3.5, in which case both measures decrease. This is because there are a number of projects with Q3 ratings between 3.5 and 4.0 that investors funded and made payments on time.

As we increase the screening stringency, the quality of projects that get funded increases, with the exception of the last screening level (Q3<3.5). Figure 1 displays the aggregate total platform-level investment (indicated with light shaded bars) as well as the investment going to low-quality projects (dark shaded bars) over the different screening levels. The graph suggests that the platform faces a trade-off as they screen out projects. On the one hand, there are fewer projects, but on the other hand, there are higher investment per project and a smaller share invested in low-quality projects. That is, both the amount and the share of investment going to low-quality projects decrease.

In particular, screening out only those projects with the highest Q3 rating (Q3<5.0) would increase the total aggregate investment compared to the simulation using all projects that were launched on the platform because the signals become more positive. Doing so also reduces the amount and the share of investment going to low-quality projects. However, increasing the screening stringency further leads to lower aggregate investment, indicating that the effect of reducing the number of projects outweighs the effect of more positive signals, though the quality of the projects increases.

The second panel of Table 8 shows the simulation outcomes when we remove the user's private signal (i.e., we set the value of the coefficient on the private signal equal to zero). The statistics in all columns are fairly similar to those of the first panel, as can also be seen

in Figure 2. For instance, in column (2), only two more projects are funded and investors and investment increase slightly without the private signal. This indicates that the investors may consider private signals from own investment in addition to the public signal in making their investment decisions but the former is not a significant factor at the platform level. That is, investors appear to be more rational than behavioral.

The third panel of Table 8 shows the simulation outcomes when we remove the public signal instead. Both the number of investors and the aggregate investment increase compared to those in the first panel for almost all levels of screening. Specifically, the aggregate investment increases by a few million dollars across columns. This is because, without the public signal, the likelihood of investment increases. However, some of the increase in investment goes to low-quality projects, although the share of investment going to low-quality projects remains at a similar level to those in the other two cases.

Figure 3 shows that the platform no longer faces the trade-off between the number of projects and the signals as the screening stringency increases. That is, the levels of aggregate investment are higher than those in Figure 1, and they are monotonically decreasing in the screening level, as the screening only serves to reduce the number of projects. Thus, the platform may have an incentive to lower the screening level further under this informational environment because the optimal screening level occurs where the marginal benefit is zero or positive. This stands in contrast to the implication of Figure 1 where the screening level is likely to be optimal or close to being optimal.

7 Conclusion

We have investigated theoretically and empirically a platform's choice of screening and the user's choice of investment under different informational environment. The platform decides whether to monitor the funded projects, and different screening levels lead to different probabilities of late payments or defaults, which generates public and private signals to the

investors. We showed that under a plausible set of assumptions, the platform would decrease the screening stringency if the public signal is not observed by the investors.

Empirically, we found some support for this prediction by estimating and simulating the investor's decisions to fund projects that were launched on a marketplace lending platform. While the private signal from own investment does not have a large impact on investment decisions, the public signal generated at the platform level plays an important role in providing incentives for the platform to set a high enough screening level to approve applications, so the users' investments are not used to fund too many low-quality projects.

From a policy standpoint, it would be important that the marketplace lenders disclose the outcomes of all funded projects to potential investors in an easily accessible manner. Such a policy would be easier to implement and enforce compared to regulatory laws (though it is not a substitute for them). The implications of our findings may apply to the platform businesses more generally, where reviews and ratings may play an important role for the user's level of trust in the platform. Future research may examine other contexts.

References

[1] Agrawal, A., C. Catalini, and A. Goldfarb. 2015. "Crowdfunding: Geography, Social Networks, and the Timing of Investment Decisions." *Journal of Economics & Management Strategy* 24: 253–274.

[2] Åstebro, T., M. Fernández, S. Lovo, and N. Vulkan. 2024. "Herding in Equity Crowdfunding." *RAND Journal of Economics* 55: 403–441.

[3] Bar-Isaac, H. and J. Deb. 2014. "(Good and Bad) Reputation for a Servant of Two Masters." *American Economic Journal: Microeconomics* 6: 293–325.

[4] Beatty, R. and J. Ritter. 1986. "Investment Banking, Reputation, and the Underpricing of Initial Public Offerings." *Journal of Financial Economics* 15: 213–232.

[5] Benzoni, L. and C. Schenone. 2010. "Conflict of Interest and Certification in the U.S. IPO Market." *Journal of Financial Intermediation* 19: 235–254.

[6] Board, S. and M. Meyer-ter-Vehn. 2013. "Reputation for Quality." *Econometrica* 81: 2381–2462.

[7] Bolton, P., X. Freixas, and J. Shapiro. 2012. "The Credit Ratings Game." *Journal of Finance* 67: 85–111.

[8] Booth, J. and R. Smith. 1986. "Capital Raising, Underwriting and the Certification Hypothesis." *Journal of Financial Economics* 15: 261–281.

[9] Bouvard, M. and R. Levy. 2017. "Two-Sided Reputation in Certification Markets." *Management Science* 64: 4755–4774.

[10] Chemmanur, T. and P. Fulghieri. 1994. "Investment Bank Reputation, Information Production, and Financial Intermediation." *Journal of Finance* 49: 57–79.

[11] Cumming, D., S. Johan, and Y. Zhang. 2019. "The Role of Due Diligence in Crowdfunding Platforms." *Journal of Banking and Finance* 108: 105661.

[12] Daley, B., T. Geelen, and B. Green. 2024. "Due Diligence." *Journal of Finance* 79: 2115–2161.

[13] Duarte-Silva, T. 2010. "The Market for Certification by External Parties: Evidence from Underwriting and Banking Relationships." *Journal of Financial Economics* 98: 568–582.

[14] Fudenberg, D. and D. Levine. 1989. "Reputation and Equilibrium Selection in Games with a Patient Player." *Econometrica* 57: 759–778.

[15] Fudenberg, D. and D. Levine. 1992. "Maintaining a Reputation when Strategies are Imperfectly Observed." *Review of Economic Studies* 59: 561–579.

[16] Fudenberg, D., D. Levine, and E. Maskin. 1994. "The Folk Theorem with Imperfect Public Information." *Econometrica* 62: 997–1039.

[17] Future Market Insights. 2024. "Peer-to-Peer Lending Market." April 12, 2024. `https://www.futuremarketinsights.com/reports/peer-to-peer-lending-market`.

[18] Gindin, A. 2022. "Purchase Order Financing: A Signalling Approach." Middlebury College, Working Paper.

[19] Iyer, R., A. Khwaja, E. Luttmer, and K. Shue. 2015. "Screening Peers Softly: Inferring the Quality of Small Borrowers." *Management Science* 62: 1554–1577.

[20] Kawai, K., K. Onishi, and K. Uetake. 2022. "Signaling in Online Credit Markets." *Journal of Political Economy* 130: 1585–1629

[21] LendingClub. 2024. "What to Do If You're Declined a Personal Loan." December 16, 2024. `https://www.lendingclub.com/resource-center/personal-loan/what-to-do-if-you-are-declined-a-personal-loan`.

[22] Mollick, E. and R. Nanda. 2015. "Wisdom or Madness? Comparing Crowds with Expert Evaluation in Funding the Arts." *Management Science* 62: 1533–1553.

[23] Nosko, C. and S. Tadelis. 2015. "The Limits of Reputation in Platform Markets: An Empirical Analysis and Field Experiment." NBER Working Paper No. 20830.

[24] Olsen, R., 1980, "A Least Squares Correction for Selectivity Bias." *Econometrica* 48: 1815–1820.

[25] Sette, E. 2011. "Competition and Optimistic Advice of Financial Analysts: Evidence from IPOs." *Journal of Financial Intermediation* 20: 441–457.

[26] Xin, Y. 2023. "Asymmetric Information, Reputation, and Welfare in Online Credit Markets." Caltech, Working Paper.

[27] Zhang, J. and P. Liu. 2012. "Rational Herding in Microloan Markets." *Management Science* 58: 892–912.

[28] Yannelis, C. and A. Zhang. 2023. "Competition and Selection in Credit Markets." *Journal of Financial Economics* 150: 103710.

Appendix

Proof of Lemma 1. We show that the left-hand side of condition (5) is increasing in r. Using expression (2) to substitute for $V(\alpha, r)$, this left-hand side is

$$\frac{(1-p)\alpha r\delta(R-E)}{\frac{1-\delta F(\alpha)}{1-F(\alpha)} - \delta\big[(1-p)\alpha + (1-(1-p)\alpha)(1-r)\big]}.$$

Differentiating with respect to r yields

$$(R-E)(1-p)\alpha\delta\frac{\frac{1-\delta F(\alpha)}{1-F(\alpha)} - \delta\big[(1-p)\alpha + (1-(1-p)\alpha)(1-r)\big] - \delta(1-(1-p)\alpha)r}{\left[\frac{1-F(\alpha)}{1-F(\alpha)} - \delta\big[(1-p)\alpha + (1-(1-p)\alpha)(1-r)\big]\right]^2},$$

which simplifies to

$$(R-E)(1-p)\alpha\delta\frac{\frac{1-\delta F(\alpha)}{1-F(\alpha)} - \delta}{\left[\frac{1-F(\alpha)}{1-F(\alpha)} - \delta\big[(1-p)\alpha + (1-(1-p)\alpha)(1-r)\big]\right]^2} > 0.$$

Condition (5) is violated at $r = 0$, since the right-hand side of condition is a strictly positive constant, and then left-hand side then evaluates to zero. Thus, since the left-hand side is increasing in r, there exists $r(\alpha) \in (0, 1]$ such that condition (5) holds for $r \in [r(\alpha), 1]$ but is violated for $r \in [0, r(\alpha))$, if the condition hold when $r = 1$. If instead the condition is violated at $r = 1$, then it is also violated for all $r \in [0, 1]$.

To complete the proof, we derive an expression for $R(\alpha)$, such that condition (5) holds at $r = 1$ if and only if $R(\alpha) \geq R$. Evaluating condition (5) at $r = 1$ yields

$$\frac{(1-p)\alpha\delta(R-E)}{\frac{1-\delta F(\alpha)}{1-F(\alpha)} - \delta(1-p)\alpha} \geq E,$$

where the left-hand side is increasing in R. Given any $\alpha \in (0, 1)$, let $R(\alpha)$ denote the value

of R for which this condition binds, i.e.

$$R(\alpha) \equiv \frac{E}{(1-p)\alpha\delta}\left[\frac{1-\delta F(\alpha)}{1-F(\alpha)}\right].$$

Thus, condition (5) holds at $r = 1$ if and only if $R(\alpha)$, as required. □

Proof of Proposition 1. Consider a candidate equilibrium with punishment probability r. This candidate must involve screening level $\alpha_r \equiv \arg\max_\alpha \ V(\alpha, r)$. Evaluating the left-hand side of condition (5) at $\alpha = \alpha_r$ yields $(1-p)\delta\alpha_r r V(\alpha_r, r)$, where differentiating with respect to r gives

$$(9) \qquad (1-p)\delta\left[\alpha_r\left(V(\alpha_r,r) + r\frac{\partial V(\alpha_r,r)}{\partial r}\right) + r\frac{d\alpha_r}{dr}\left(V(\alpha_r,r) + \alpha_r\frac{\partial V(\alpha,r)}{\partial \alpha}|_{\alpha=\alpha_r}\right)\right].$$

The proof of Lemma 1 showed that the left-hand side of condition (5) was increasing in r, holding α constant at any value, including $\alpha = \alpha_r$. Hence, $\alpha_r\left(V(\alpha_r, r) + r\frac{\partial V(\alpha_r,r)}{\partial r}\right) > 0$. The main text showed that α_r was increasing in r, thus, $\frac{d\alpha_r}{dr} > 0$. Moreover, $\frac{\partial V(\alpha,r)}{\partial \alpha}|_{\alpha=\alpha_r} = 0$, by the definition of α_r. It follows that expression (9) is strictly positive, and so $(1-p)\delta\alpha_r r V(\alpha_r, r)$ is increasing in r.

Suppose that $R < R(\alpha_1)$, so that condition (5) is violated at $r = 1$ and $\alpha = \alpha_1$: $(1-p)\delta\alpha_1 V(\alpha_1, 1) < E$. Since $(1-p)\delta\alpha_r r V(\alpha_r, r)$ is increasing in r, it follows that $(1-p)\delta\alpha_r r V(\alpha_r, r) < E$, and hence (5) is violated, for all $r \in [0, 1]$. Thus, no equilibrium with monitoring can exist when $R < R(\alpha_1)$.

Now suppose instead that $R \geq R(\alpha_1)$. Condition (5) evaluated at $\alpha = \alpha_r$ is then violated at $r = 0$ and satisfied at $r = 1$. Since $(1-p)\delta\alpha_r r V(\alpha_r, r)$ is increasing in r, it follows that there exists a unique $r \in (0, 1]$ such that (r, α_r) makes condition (5) bind. □

Proof of Proposition 2. We solve for an equilibrium with monitoring, given observable screening, that makes condition (5) bind. Since $R \geq R(\alpha_1)$, an equilibrium with monitoring exists

under unobservable screening, where $r^{Un} \in (0,1]$ and $\alpha^{Un} = \alpha_{r^{Un}}$ make condition (5) bind.

As condition (5) involves a weak inequality, the set of α, for which there exists $r(\alpha) \in [0,1]$ that makes (5) bind, must be closed. The set is also non-empty, as it contains $\alpha_{r^{Un}}$.

Under observability, the equilibrium screening level satisfies

$$\alpha^{Ob} \equiv \arg\max_{\alpha} \; V(\alpha, r(\alpha)),$$

where we allow α to range over this closed set. It is immediate that such an α^{Ob} exists. Moreover, $\alpha^{Ob} \in (0,1)$, since both $\alpha \in \{0,1\}$ violate the monitoring condition (5).

Consider a candidate equilibrium $(\alpha, r(\alpha))$. Suppose first that $r = r^{Un}$, hence $(\alpha, r(\alpha)) = (\alpha^{Un}, r^{Un})$, corresponding to the equilibrium values under unobservable screening. Note that

$$\frac{\partial V(\alpha, r(\alpha))}{\partial \alpha} + \left[\frac{\partial}{\partial r} V(\alpha, r)|_{r=r(\alpha)} \frac{dr(\alpha)}{d\alpha}\right] = 0 \tag{10}$$

must hold in equilibrium, by the definition of α^{Ob}.

To determine the sign of $\frac{dr(\alpha)}{d\alpha}$, recall that the left-hand side of (5) is strictly increasing in r. Thus, $\frac{dr(\alpha)}{d\alpha} < 0$ if and only if the left-hand side of (5) is strictly increasing in α: $V(\alpha, r(\alpha)) + \alpha \frac{\partial V(\alpha, r(\alpha))}{\partial \alpha} > 0$. We have $\frac{\partial V(\alpha, r(\alpha))}{\partial \alpha}|_{\alpha=\alpha^{Un}} = 0$. Hence, $\frac{dr(\alpha)}{d\alpha} < 0$ when evaluated at α^{Un} and r^{Un}. Combined with $\frac{\partial V(\alpha, r)}{\partial r} < 0$,the left-hand size of condition (10) therefore reduces to $\frac{\partial}{\partial r} V(\alpha, r)|_{r=r(\alpha)} \frac{dr(\alpha)}{d\alpha} > 0$. Thus, (α^{Un}, r^{Un}) cannot be an equilibrium.

Since all terms are continuous in parameters α and r, $V(\alpha, r(\alpha)) + \alpha \frac{\partial V(\alpha, r(\alpha))}{\partial \alpha} > 0$ must hold in a neighborhood of (α^{Un}, r^{Un}). Thus, $\frac{dr(\alpha)}{d\alpha} < 0$ in this neighborhood. That is, a reduction in α from α^{Un} to $\alpha^{Un} - \epsilon$, for $\epsilon > 0$ sufficiently small, will induce an increase in the punishment probability: $r(\alpha^{Un} - \epsilon) > r(\alpha^{Un})$.

We now consider whether $(\alpha^{Un} - \epsilon, r(\alpha^{Un} - \epsilon))$ can be an equilibrium with observable screening. Recall that for any given r, α_r is increasing in r, where α_r is the screening level for which $\frac{\partial V(\alpha, r)}{\partial \alpha}|_{\alpha=\alpha_r} = 0$. Thus, since $V(\alpha, r)$ is strictly unimodal in α, $\frac{\partial V(\alpha, r(\alpha))}{\partial \alpha} > 0$ holds when evaluated at any $\alpha \in [\alpha^{Un} - \epsilon, \alpha^{Un})$. This in turn implies $\frac{dr(\alpha)}{d\alpha} < 0$ when evaluated at

$\alpha = \alpha^{Un} - \epsilon$. Thus, both terms on the left-hand side of condition (10) are strictly positive. We conclude that condition (10) is violated at $(\alpha^{Un} - \epsilon, r(\alpha^{Un} - \epsilon))$, and hence this cannot be an equilibrium. Repeatedly applying this argument implies that (10) is strictly positive for all $(\alpha, r(\alpha))$ with $\alpha \leq \alpha^{Un}$. Thus, condition (10) will be violated for all such pairs.

We have shown that the equilibrium with observable screening involves $\alpha^{Ob} > \alpha^{Un}$. To complete the proof, we show that $r^{Ob} \equiv r(\alpha^{Ob}) < r^{Un}$. Suppose to the contrary that $r^{Ob} \geq r^{Un}$. For any $r \geq r^{Un}$, we have

$$V(\alpha^{Ob}, r) \leq V(\alpha^{Ob}, r^{Un}) < V(\alpha^{Un}, r^{Un}),$$

where the weak inequality follows from $\frac{\partial V(\alpha, r)}{\partial r} < 0$, and the strict inequality follows from $\alpha^{Un} = \arg\max_{\alpha} \ V(\alpha, r^{Un})$. Thus, $V(\alpha^{Ob}, r^{Ob}) < V(\alpha^{Un}, r^{Un})$. But this implies that under observable screening, the platform prefers setting α^{Un}, which yields value $V(\alpha^{Un}, r^{Un})$, to setting α^{Ob}, which yields value $V(\alpha^{Ob}, r^{Ob})$. This contradicts $\alpha^{Ob} = \arg\max_{\alpha} \ V(\alpha, r(\alpha))$. We therefore conclude that $r^{Ob} < r^{Un}$. $\square$

Table 1: Project Summary Statistics

	Mean	StDev	Median	Min	Max
Total registered users	1100	399.3	1110	42	1912
Total users investing	61.10	48.59	46	3	284
Total amount invested	21,063	29,916	9456	550	250,538
Ever 2 weeks late	0.22	0.42	0	0	1
Ever 1 month late	0.19	0.40	0	0	1

Statistics for 312 projects that were posted during the sample period. For the last two rows, 38 projects were excluded that have not begun the repayment period.

Table 2: User Descriptive Statistics

	Mean	StDev	Median	Min	Max
Total projects available	56.26	85.68	10	1	312
Total projects invested in	3.12	11.75	0	0	251
Total amounts invested	1077	6074	0	0	154,830
Mean amount invested	497.7	1778	125	20	37,500
User tenure (days)	67.64	119.6	0	0	587
Private 2-weeks late (%)	1.54	6.61	0	0	94
Private 1-month late (%)	1.15	5.56	0	0	92
Public 2-weeks late (%)	11.10	4.11	12	0	21
Public 1-month late (%)	8.62	3.73	9	0	14

Statistics for 6101 users who created an account from the start of the platform until the sample period ends. For those who kept logging on, activities are right-censored.

Table 3: Expert Evaluation Descriptive Statistics

	Mean	StDev	Median	Min	Max
Q1: likelihood of business success	3.21	0.97	3.5	1.0	5.0
Q2: attractiveness of payment terms	3.48	1.38	3.5	1.0	5.0
Q3: riskiness of investment	3.47	1.02	4.0	1.0	5.0
Q4: likelihood of late payments (2 weeks)	3.14	0.95	3.0	1.0	5.0
Q5: likelihood of late payments (1 month)	2.60	0.92	3.0	1.0	5.0
Q6: likelihood of institution lending	3.08	1.13	3.0	1.0	5.0

Statistics from a survey questionnaire fielded shortly after data access, completed by a lead underwriter of a capital firm in a similar line of business on the 312 projects in the sample.

Table 4: Probability of Investment

	(1)	(2)	(3)	(4)
Public signal (2 wk)	-2.522	-2.510		
	(1.072)**	(1.072)**		
Private signal (2 wk)		-0.021		
		(0.004)***		
Public signal (1 mo)			0.974	0.977
			(3.111)	(3.112)
Private signal (1 mo)				-0.020
				(0.004)***
Log(investments+1)	-0.018	-0.018	-0.017	-0.018
	(0.004)***	(0.004)***	(0.004)***	(0.004)***
Zero investment	-0.043	-0.044	-0.043	-0.043
	(0.009)***	(0.009)***	(0.008)***	(0.008)***
User tenure	0.020	0.020	0.019	0.019
	(0.006)***	(0.006)***	(0.006)***	(0.006)***
Offer FE	Y	Y	Y	Y
User FE	Y	Y	Y	Y
Day of Week FE	Y	Y	Y	Y
Adj. R^2	0.25	0.25	0.25	0.25
Observations	342,937	342,937	342,937	342,937

The dependent variable is 1 for investment and 0 otherwise. Zero investment is 1 if a user has not made any investment and 0 otherwise. Standard errors are in parentheses. Statistical significance is denoted as *** $p < .01$, ** $p < .05$, * $p < .10$.

Table 5: Correlations between User Responses and User Characteristics

	(1) Platform Usage	(2) Risk Tolerance
Age	-0.055	0.049
Female	-0.106**	-0.189***
Education (four standard categories)	-0.158***	0.073
Household income (in $25K brackets)	-0.034	0.286***
Years of investment experience	0.052	0.277***
Percent of alternative investment	0.084	0.067

Each cell displays the correlation between the survey responses on (1) and (2), ranging from 1 to 5, and the self-reported characteristics; the number of survey respondents is suppressed. Statistical significance is denoted as *** $p < .01$, ** $p < .05$, * $p < .10$.

Table 6: Amount of Investment

	(1)	(2)	(3)	(4)
Public signal (2 wk)	-0.490	-0.440		
	(1.316)	(1.317)		
Private signal (2 wk)		-0.079		
		(0.103)		
Public signal (1 mo)			2.154	2.162
			(1.884)	(1.885)
Private signal (1 mo)				-0.049
				(0.129)
Log(investments+1)	-0.039	-0.040	-0.039	-0.040
	(0.015)***	(0.015)***	(0.015)***	(0.015)***
Zero investment	-0.018	-0.020	-0.020	-0.020
	(0.030)	(0.030)	(0.029)	(0.030)
Pr(invest)	-0.885	-0.883	-0.926	-0.925
	(0.157)***	(0.157)***	(0.168)***	(0.168)***
Offer FE	Y	Y	Y	Y
User FE	Y	Y	Y	Y
Day of Week FE	Y	Y	Y	Y
Adj. R^2	0.72	0.72	0.72	0.72
Observations	18,497	18,497	18,497	18,497

The dependent variable is the amount of investment. The predicted probability of investment from the matching columns of Table 4 is included to correct for selection bias. Statistical significance is denoted as *** $p < .01$, ** $p < .05$, * $p < .10$.

Table 7: Correlations between Project Outcomes and Expert Evaluations

	(1) 2 Weeks Late	(2) 1 Month Late
Q1: likelihood of business success	-0.138**	-0.117*
Q2: attractiveness of payment terms	-0.110*	-0.123**
Q3: riskiness of investment	0.172***	0.145**
Q4: likelihood of late payments (2 weeks)	0.080	0.056
Q5: likelihood of late payments (1 month)	0.131**	0.103*
Q6: likelihood of institution lending	-0.079	-0.063

Each cell displays the correlation between the outcomes and the expert evaluations of 274 projects; 38 projects are excluded because they did not begin the payment period. Statistical significance is denoted as *** $p < .01$, ** $p < .05$, * $p < .10$.

Table 8: Platform-level Outcomes by Screening Stringency

	(1) Data	(2) All	(3) Q3<5.0	(4) Q3<4.5	(5) Q3<4.0	(6) Q3<3.5
			Both Signals Observed			
Funded projects	278	277	262	219	124	89
Mean investors	43.48	40.72	47.81	48.51	61.34	53.99
StDev investors	38.22	40.38	48.80	50.14	61.83	65.41
Mean investment	15.22	14.07	16.32	17.26	21.71	19.66
StDev investment	21.64	19.22	22.22	23.58	26.31	27.10
Mean quality (Q1)	2.90	2.87	3.01	3.15	3.48	3.17
			No Private Signal			
Funded projects		279	262	219	124	89
Mean investors		41.53	47.85	48.58	61.39	54.02
StDev investors		40.96	48.88	50.27	61.90	65.50
Mean investment		14.27	16.32	17.26	21.71	19.66
StDev investment		19.25	22.22	23.58	26.31	27.10
Mean quality (Q1)		2.89	3.01	3.15	3.48	3.17
			No Public Signal			
Funded projects		305	283	239	138	112
Mean investors		65.32	66.10	68.98	76.30	69.95
StDev investors		65.18	65.85	68.75	71.44	67.85
Mean investment		18.68	19.03	20.41	24.21	22.27
StDev investment		23.22	23.67	25.09	27.47	26.19
Mean quality (Q1)		3.17	3.26	3.45	3.90	3.99

Each column, except for the first one, describes results of simulating investor behavior during the sample period based on the estimates in column (2) of Table 4 and Table 6, while changing the required level of project risk as proxied by the expert's rating on Q3.

Figure 1: Total Investments by Screening Level (Both Signals Observed)

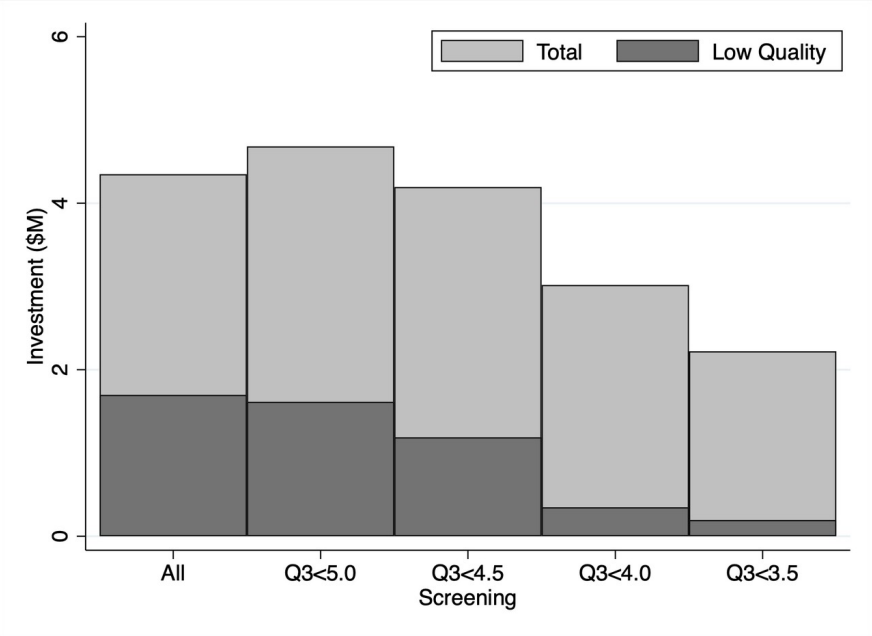


Figure 2: Total Investments by Screening Level (No Private Signal)

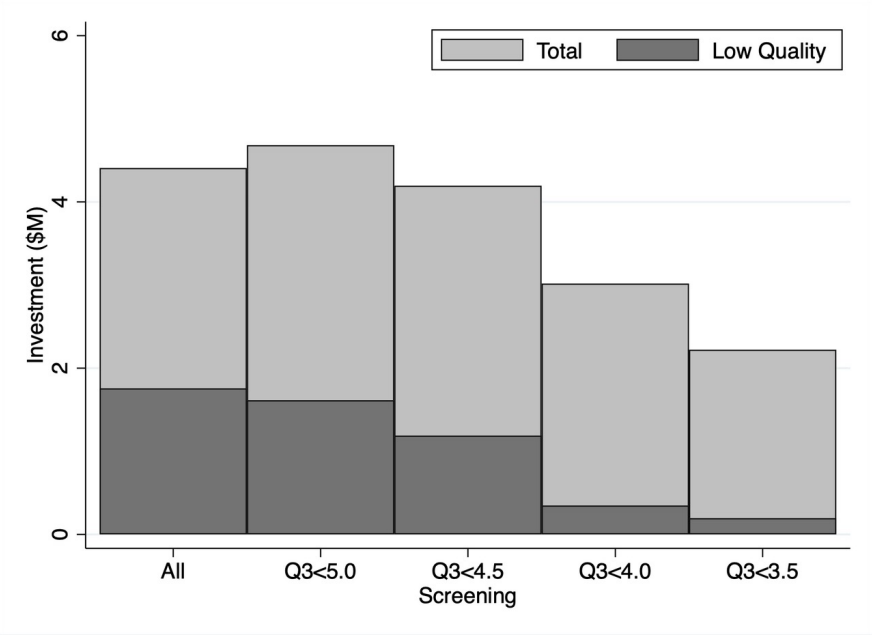


Figure 3: Total Investments by Screening Level (No Public Signal)

