Marketplace Lending: A New Banking Paradigm?[*]

Boris Vallee

Harvard Business School

Yao Zeng

University of Washington

January, 2018

Abstract

Marketplace lending relies on large-scale loan screening by investors, a major deviation from the traditional banking paradigm. Theoretically, the participation of sophisticated investors in marketplace lending improves screening outcomes but also creates adverse selection. In maximizing loan volume, the platform trades off these two forces by choosing intermediate levels of platform pre-screening and information provision to investors. We use novel investor-level data to test these predictions. We empirically show that more sophisticated investors screen loans differently and systematically outperform less sophisticated investors. However, the outperformance shrinks when the platform reduces information provision to investors, consistent with platforms managing adverse selection.

KEYWORDS: Marketplace lending, screening, sophisticated investors, adverse selection

JEL: G21, G23, D82

1   Introduction

Lending marketplaces, also commonly referred to as peer-to-peer lending platforms, such as LendingClub and Prosper, have been rapidly gaining market share in consumer lending over the last decade. Loans issued by these platforms represent one third of unsecured consumer loans volume in the US in 2016, and their revenues are predicted to grow at a 20% yearly rate over the next five years.[1] This rapid development has important implications on the consumer lending market, and more broadly on retail banking.

Designed as a two-sided platform structure, marketplace lending brings innovations to traditional banking on both the borrower and the investor side. The innovation on the borrower side relies mainly on streamlining an online application process that uses low-cost information technology to collect standardized information from dispersed individual borrowers on a large scale.[2] However, the true breakthrough that marketplace lending creates lies in direct investor participation on platforms. Although lending platforms pre-screen loan applications modestly, they allow and rely on investors to screen borrowers directly, giving investors continuous and tailored exposure to an asset class which they could not access previously. This model significantly differs from the traditional banking paradigm where depositors are essentially isolated from the borrowers. Moreover, investor composition on lending platforms has been evolving significantly since the platforms' inception, due to informationally sophisticated investors' increasing participation. These diverse investors perform large-scale borrower screening, and the resulting allocation of information production between the platforms and investors challenges the traditional role of banks of exclusive information production on behalf of investors (Diamond and Dybvig, 1983, Gorton and Pennacchi, 1990).

The design of lending marketplaces and the composition of the investors active on these platforms jointly raise a series of research questions on this new banking business model. Are more sophisticated investors in lending marketplaces consistently better at screening borrowers and thereby outperform? If so, how does their screening outperformance relate to the design of the platforms, and how do their investment strategies interact with the ones from less sophisti-

---

[1] See *IBIS World Industry Report OD4736: Peer-to-Peer Lending Platforms in the US*, 2016.

[2] Traditional banks are also starting to follow this model, for example, Marcus by Goldman Sachs.

cated investors? Given the heterogeneity of investors, what is the optimal design of a platform? Specifically, what are the optimal levels of platform pre-screening and information provision to investors in order to maximize lending volumes? Answering these questions is essential to better understand the promises of this new banking business model, such as an informational synergy between the platform and its investors, as well as its potential pitfalls, such as adverse selection among investors.

Our study addresses these questions by developing a theoretical framework for marketplace lending and testing its predictions using a proprietary dataset that includes borrower and investor data. Thus, the contribution of our paper is both theoretical and empirical.

To start, we theoretically argue and find supportive empirical evidence that informationally sophisticated investors actively use information provided by the platform to screen listed loans (beyond platform pre-screening) and identify good loans to invest. Rather, less sophisticated investors do not screen; they invest in an average loan passively if they can break even, or they do not invest at all. This results in sophisticated investors' outperformance relative to less sophisticated investors. The more information the platform provides, the greater the outperformance sophisticated investors enjoy. Because sophisticated investors can identify good loans and finance them, their participation helps boost the volume of loans financed on the platform when less sophisticated investors do not invest.

However, the heterogeneity in investor sophistication creates an endogenous adverse selection problem, which can hurt volume. Adverse selection hurts volume through both a price effect and a quantity effect. On the price effect, because more sophisticated investors are better able to identify and finance good loans, the quality of an average loan facing a less sophisticated investor becomes lower. Thus, less sophisticated investors require a lower loan price to break even, resulting in a lower prevailing loan price on the platform. This eventually lowers the amount of loan applications in the first place. On the quantity effect, if adverse selection becomes too severe, less sophisticated investors may not break even and exit the market as a whole, leading to even lower volume.

To maximize volume, the platform optimally trades off these positive and negative effects of sophisticated investor participation when designing its optimal policies in terms of 1) pre-

screening intensity and 2) information provision to investors. Both policies impact sophisticated investors' screening and performance. When platform pre-screening cost is initially high, the platform optimally chooses a low pre-screening intensity but distributes more information to investors. This encourages the participation of sophisticated investors, boosting volume financed by them even if less sophisticated investors do not participate. When platform pre-screening cost becomes low as the platform develops, it optimally reverses the policies by choosing a high pre-screening intensity such that less sophisticated investors are willing to invest, but the platform also distributes less information to mitigate the adverse selection problem caused by sophisticated investors. However, the lending platform will never pre-screen loan applications too intensively, even if the physical cost of pre-screening is low. This is because too intensive pre-screening would screen out loans that could have been financed by less sophisticated investors (providing these investors breaking even) and thus reduce volume.

Our analysis relies on a theoretical model which is designed to rigorously capture the economic forces discussed above and to guide the empirical analysis. In the model, a platform pre-screens a pool of loan applications, lists some loans on the platform, and distributes some information to investors to facilitate their screening and investing decisions. Unsophisticated investors are competitive and cannot screen, and thus they are always uninformed about an individual loan's quality. They break even when investing in listed loans on the platform. Differently, sophisticated investors can pay a cost, which depends on how much information the platform provides, to become informed and identify good loans. Because sophisticated investors, once becoming informed, only finance good loans, they outperform uninformed, unsophisticated investors in equilibrium and impose adverse selection on them. The platform's optimal design accordingly trades off the positive and negative forces of sophisticated investor participation on volume.

Testing the model predictions crucially relies on data of investors of heterogenous sophistication. Although borrower- and loan-level data is made public by the platforms, data on investor

3

characteristics and their loan portfolios are generally unavailable in the public domain.[3,4] Also, pre-screening only varies at the platform level at a given time, which calls for data covering at least two platforms. Fortunately, the unique institutional context and novel data structure that we have allows us to mitigate these empirical challenges. First, we obtain a rich dataset that includes portfolio composition for a large set of retail investors on two lending platforms. We can therefore study sophisticated investor screening and its performance within the same investor segment, and across platforms. Importantly, our sample includes a significant source of heterogeneity in terms of sophistication: some investors invest by themselves, whereas others rely on the various screening technologies provided by LendingRobot, an algorithmic third-party. This setup allows us to test and quantify the potential gain in performance obtained by sophisticated investor screening, and build several relevant benchmarks.

Our empirical analysis progresses in several steps. First, we examine sophisticated investors' screening activities by running regression analyses that study what loan characteristics predict sophisticated investors' loan selection. We show that more sophisticated investors rely on significantly different loan characteristics in financing loans, which points to their screening advantage. We further test whether a loan being selected by more sophisticated investors predicts a lower probability of default and quantify the magnitude of more sophisticated investors' relative outperformance. We find that loans selected by sophisticated investors have a default rate on average 3 percentage points lower than the average loan, or loans picked by unsophisticated investors, which corresponds to a reduction of more than 20% of the average default risk.

We then implement a difference-in-differences methodology to establish causal evidence of the impact of a large reduction in the information provision to investors on sophisticated investors' performance. We find that the outperformance of sophisticated investors drops by more than half

---

[3]Traditionally, the breakdown between retail and institutional investors represents a natural source of heterogeneity in terms of sophistication, and platform public data allows to identify which loans are sold to retail (fractional loans) or institutional investors (whole loans). However, this distinction is not informative in marketplace lending, as the allocation between the retail and institutional investor segments is randomized by platforms, and each segment itself also has a large heterogeneity of investor sophistication. Study of the impact of investor sophistication need therefore to be conducted within these segments. In addition, it is not obvious what the right benchmark to compare sophisticated investors in lending marketplaces is.

[4]For example, we may view the overall platform lending as the relevant benchmark to sophisticated investors investing in lending platforms, or unsophisticated investors. Equilibrium effects potentially affect the first benchmark though, whereas the second one requires data covering both sophisticated and unsophisticated investors.

at the time of the shock to the information set.[5] We rationalize this reduction, corresponding to the platform "evening the playing field", by referring to the theoretical argument that platforms actively manage the potential adverse selection problem introduced by sophisticated investors. We also provide robust time-series evidence showing that sophisticated investors' outperformance has become lower in recent years, again consistent with platforms actively managing the potential adverse selection problem.

We finally document two additional stylized facts on platform pre-screening: on the extensive margin, platforms have been relaxing their pre-screening standards, while on the intensive margin their screening precision has been improving over time. This suggests that the platforms are constantly calibrating their pre-screening intensity depending on varying economic conditions, consistent with our theory prediction that an intermediate level of platform pre-screening is optimal.

Although our empirical tests mainly rely on heterogeneity within the retail investor segment, our findings have external validity for the institutional investor segment of marketplace lending as well, as the heterogeneity in screening sophistication is comparable across investor segments. Many institutional investors, such as pension funds, only apply a little screening (for instance, only relying on a grade threshold) as retails investors do; while other institutional investors, such as hedge funds, develop highly sophisticated investment strategies that are comparable to what LendingRobot offers to investors.[6] Our results thus represent a reasonably large range of performance gains by various types of sophisticated investors in lending marketplaces.[7] Also, both retail and institutional segments appear increasingly sophisticated as is suggested by the shortening of the average speed of loan sales.

To keep our paper focused on the asset side of intermediation, we leave a number of questions for future research. These include the impacts of sophisticated investor screening on borrower welfare, as well as the overall welfare implications of marketplace lending. Moreover, there are interesting open questions as to how marketplace lending competes with more traditional

---

[5]This drop in performance could not be immediately observed at the time of the change, as it takes time for loan performance to be revealed.

[6]LendingRobot is actually partnering with some undisclosed hedge funds to execute their orders.

[7]We acknowledge that some hedge funds are likely to possess even more sophisticated screening process than the most sophisticated investor who relies on LendingRobot's screening technology, while investors who only use LendingRobot to monitor their portfolios might also not be the least sophisticated.

retail banking and whether marketplace lending poses any financial stability concerns due to sophisticated investors' potential safe asset creations, which we plan to examine in future.

Related Literature. Our paper mainly contributes to the burgeoning literature on marketplace lending. This literature has mainly focused on how borrowers' soft information improves lending outcomes (for example, Duarte, Siegel and Young, 2012, Iyer, Khwaja, Luttmer and Shue, 2015, and see Morse, 2015 for a review). To the best of our knowledge, we are the first to study how investors' characteristics affect loan screening outcome and how the participation of informationally sophisticated investors affects optimal platform design. Paravisini, Rappoport and Ravina (2016) also use a sample of investor portfolio data from Lending Club in the 2007-2008 period, but they mainly test a classical asset pricing relationship between risk aversion and wealth rather than focusing on marketplace lending per se.

On optimal platform design, our work complements recent papers that study the motivation behind one lending platform's switch from an auction mechanism to posted loan prices in its early stage (Franks, Serrano-Velarde and Sussman, 2016), and its effect on high-risk borrowers (Liskovich and Shaton, 2017). By studying the informational design of lending marketplaces with a focus on investor screening, our work also speaks to the competition between lending marketplaces and the traditional banking sector in terms of loan screening (Mariotto, 2016, Balyuk, 2017, Buchak, Matvos, Piskorski and Seru, 2017).

Our paper is also related to the theoretical literature of endogenous adverse selection induced by investor information acquisition (Glode, Green and Lowery, 2012, Biais, Foucault and Moinas, 2015, for example). Closer to us are Fishman and Parker (2015), Bolton, Santos and Scheinkman (2016) and Yang and Zeng (2017) who consider various production settings, and also consider the two forces that investor information acquisition helps guide efficient production but also introduces adverse selection that may potentially lower gains from trade. The contribution of our paper is to embed this trade-off in an outer-level optimal platform design problem where the project supply is endogenous. Although our study is conducted in the context of marketplace lending, it sheds new light on other financial market design contexts in which adverse selection may be a concern, as suggested in Rochet and Tirole (2006).

2 Institutional Details of Marketplace Lending

In this section, we describe several key aspects of marketplace lending that are relevant to our study: platform information collection, platform pre-screening, the funding model (including both investor screening and platform information distribution), and changes in investor composition. We refer interested readers to existing papers (for example, Iyer, Khwaja, Luttmer and Shue, 2015) and in particular the review of Morse (2015) for general institutional details of lending marketplaces.[8]

2.1 Platform Information Collection and Pre-Screening

Information collection. By design, lending marketplaces only collect information on borrowers via online self-reporting, and through credit pulls. Thus, information collection itself is standardized and mostly costless.[9,10] A fraction of the self-reported information is verified by the platform by requiring supporting documents. Under the current practice, there is no personal interaction between the borrowers and platform employees, neither through physical meetings, phone interviews, or web chats.

 Pre-screening. Armed with the information they collect, platforms perform loan pre-screening on two margins. On the extensive margin, they decide to accept or reject the application; an accepted application is subsequently listed on the platform and made available to

---

[8]Marketplace Lending received an important coverage in the year 2016, following some governance issues at the main platform, Lending Club. First, it was discovered that Lending Club sold to an institutional investor loans that were not respecting the criteria the investor had set with Lending Club (some had a lower grade than the threshold set by the investor). Lending Club later repurchased these loans from the investor. Second, following inquiries into this event, it was found out that in 2014, Lending Club had made loans to the family of its CEO. Although the amount was relatively small, it helps Lending Club window dress its volume and allowed Lending Club to meet its issuance volume guidance at the time. See "Inside the Final Days of Lending Club CEO Renaud Laplanche," the *Wall Street Journal*, May 16, 2016 for a detailed document. While both these events revealed serious governance issue at Lending Club, they do not speak to the marketplace lending economic model. Media coverage of these two facts was largely misleading.

[9]At the inception of marketplace lending, platforms frequently collected soft information, i.e. non-standardized answers to questions, such as a description of the project to be financed, or even pictures, which were an important part of the "peer-to-peer" aspect that was initially supported. Since these early years, platforms have progressively stopped collecting soft information to standardize the information set and to streamline the application process. Collecting more information, especially if not standardized, would increase the drop-off rate during the online application, which would effectively make information collection costly for the platform.

[10] Some platforms currently offer to link applications to existing bank accounts or social media accounts to collect additional information, but the main players in retail lending marketplaces, Lending Club and Prosper, do not. Again, collecting this type of information is costly as it might reduce the pool of applicants willing to comply.

investors to potentially finance. On the intensive margin, they allocate the loan to a risk bucket, called a grade or a sub-grade. Currently, Prosper classifies its listed loans into 7 grades, while Lending Club uses a scoring system of 35 sub-grades. These grades map into interest rates, i.e. loan prices.

These screening decisions rely on the platform's screening model, and the development of such a model requires sophisticated data analysis, which requires a fixed cost. The more precise the allocation of loans into risk-buckets is, including the ones that do not get listed, the costlier the screening model is to develop.

The platform's screening model also evolves over time as the platform learns from the growing pool of loan applications that are listed and loans that are financed. The increasing data available to platforms suggests that the cost associated with pre-screening decreases over time. However, this learning process may be slowed by the absence of pre-existing screening expertise and data compared to traditional banks, by the relatively long maturity of the loans, typically 3 or 5 years, as well as by the lack of counterfactual for the extensive margin decision.

2.2 Marketplace Funding Model and Information Distribution

Funding model. The funding model of marketplace lending heavily relies on investors screening and investing in loans individually, as is the case for loans issued on Lending Club, Prosper or Funding Circle.[11] This funding model represents a fundamental difference from the traditional banking model in which depositors effectively hold a demandable debt contract issued by the bank without having any knowledge of the underlying loans (Diamond and Dybvig, 1983, Gorton and Pennacchi, 1990).[12,13]

Information distribution. To facilitate investor screening and purchase of listed loans, platforms provide investors with a set of standardized information for each listed loan, which is

---

[11]More precisely, individual investors purchase notes that are backed by loans. See Morse (2015) for a detailed description of this process.

[12]A few other FinTech firms follow a different model, with some players keeping the loans on their balance sheet while obtaining wholesale funding from another institution, as OnDeck is doing, or by implementing tranched securitization on a large pool of loans. These platforms, while doing online lending and collecting information as previously described, are not creating marketplaces per se, and are closer to the shadow banking sector. Some online lenders follow a hybrid model, such as Avant (both marketplace and balance sheet funding) or SoFi (both balance sheet funding and securitization). Lending Club and Prosper jointly captured around half of the lending activity by FinTech firms in 2016, while the third largest player, SoFi, only captured less than 7% of the market.

[13]Lending Club also experimented with securitization, but it remains marginal in its funding model.

typically narrower than the information the platform collects. This provision of information has two main purposes. First, it allows the investor to check the quality of the loan she purchases, as platforms do not have skin in the game.[14] Second, information provision makes it possible for investors to further screen loans if they choose to do so. Consistent with the innovative funding model, these two characteristics are also different from traditional banking or modern securitization as studied in the literature, where no information or only aggregate information is provided to end investors.

Lending Club change in investor information set. On November 7th, 2014, Lending Club removed 50 out of the 100 variables on borrowers' characteristics that they were sharing with investors previously. This removal affected new loan listings available on the website, listed loan information available through the application programming interface (API), as well as historical data available for download.

This change was unanticipated and anecdotal evidence suggests that it was motivated by the desire to even the playing field between investors.[15] While Lending Club and other platforms regularly adjust the number of available variables, this change in information set is unique by its magnitude and evidences how important this choice variable is for the platform.

Role of investor screening. Investor screening plays an important role on lending platforms as it directly impacts whether a loan will eventually get funded and be issued.[16] In turn, this funding model allows the platforms to better calibrate their pre-screening intensity and loan pricing by monitoring investors' funding activities. The platform pre-screens and prices listed loans in expectation of investors' ultimate investing decisions.[17] Investor screening therefore impacts the way the platform allocates loans to risk buckets, as well as the interest rates that the platform attributes to each risk bucket to have the market clear.[18] These interest rates have

---

[14]The role of information provision as a substitute for the originator's skin in the game is further backed by the fact that when the same platforms implement securitization, they provide less information and keep an equity position in the transaction.

[15]More details on this change are available at http://www.lendingmemo.com/cutting-open-data-50-lending-club-may-lose/.

[16]If total commitments to a loan by the expiration date are less than the requested amount but above a certain threshold, the borrower can accept that lesser amount or withdraw the loan request.

[17]This current setting is an evolution from the earlier practice of platforms to run auctions for setting the interest rate, which created liquidity issues. See Franks, Serrano-Velarde and Sussman (2016) and Liskovich and Shaton (2017) for more discussion on the auction model.

[18]See discussion on this point in Lending Club Investor Day 2017 Presentation, which is available at ir.lendingclub.com/Cache/1001230258.pdf.

evolved over time, as Figure 3 illustrates.

[Insert Figure 3]

2.3 Investor Composition

Marketplaces initially targeted retail investors only, but as marketplaces faced regulatory burden and challenges to scale up through the retail investor pool, the platforms increasingly opened up to institutional investors. Today, most lending platforms, including Lending Club, Prosper, and Funding Circle, target both retail and institutional investors. Typically, institutional investors purchase whole listed loans, while retail investors invest in a fractional loan, meaning that each loan is divided into small USD25 notes that bear the credit risk of the loan. The fractional process allows retail investors to access diversification of idiosyncratic risk even with a modest portfolio size, thus encouraging their participation. Morse (2015) suggests that the share of institutional investors on lending platforms has increased from less than 10% to more than 80% since 2012. Lending Club 2017 Investor Day Presentation also indicates the following breakdown of investors: banks and other institutional investors (55%), managed accounts (29%), and self-directed retail investors (13%).[19]

Within each segment of investors, the composition of investors has been evolving as well. While the initial investor base was made of households and traditional asset managers, sophisticated third parties, such as LendingRobot and NSR Invest for retail investors, and Orchard for institutional investors, are increasingly providing investors with algorithmic tools to automatically screen loan and execute orders. A number of hedge funds have also publicly announced that they are investing in this asset class, and banks are also increasingly participating. These latter types of investors potentially bring their own loan screening expertise, developed out of sample, to lending marketplaces. Within each segment, investor can therefore be further classified between two types: unsophisticated investors that buy loans with minimum screening, typically relying on an automated feature of the platform that allows passive investing, and sophisticated investors who actively screen loans using a proprietary screening model and execute orders at

---

[19]Managed accounts are passive investment vehicles that are distributed by conventional third-party marketers to high net worth individuals.

high speed through the platform API. The speed for funding loans has increased dramatically, with the most popular loans being funded in seconds, as platforms have opened API access to pass buying orders.[20]

This anecdotal evidence points towards an increase in the average informational sophistication of investors on lending marketplaces, on both retail and institutional segments. The presence and surge of informationally sophisticated investors have important implications regarding the design of lending platforms, which is the focus of this project.

3 Theoretical Framework

3.1 The Model

There are four types of risk-neutral agents: 1) a lending platform 2) a continuum of x_0 of loan applicants, each of which has a project and can submit a loan application, 3) a mass Ω of sophisticated investors, and 4) an competitive fringe of unsophisticated investors. There are three dates, $t = 0, 1, 2$, with no time discount. The lending platform screens loan applications at date 0. The applications that are successfully screened in are listed on the platform at date 1, and investors decide whether to buy these listed loans at that time. Loan cash flows are realized at date 2. Investors can only invest in loans that are listed on the platform. Figure 1 illustrates the model setting, with agents' objectives being specified subsequently.

Loan applications. Each loan applicant has no fund to start with and one project that requires an initial investment I. Funded projects generate a random cash flow R to the investor at date 2, which follows a Bernoulli distribution. The project pays $R = R_H$ with probability π_0 ("good project") and $R = R_L$ with probability $1 - \pi_0$ ("bad project"), where $\pi_0 \in (0, 1)$. We assume that $R_L < I < R_H$, which implies that ex post the project with $R = R_H$ is worth investing while the project with $R = R_L$ is not. The distribution of good and bad projects is common knowledge but the platform and investors do not know which project is good ex ante.

The number of initial loan applicants x_0 is endogenous, and depends on the interest rate offered by the platform. We model loan demand (or equivalently, project supply) in a reduced

---
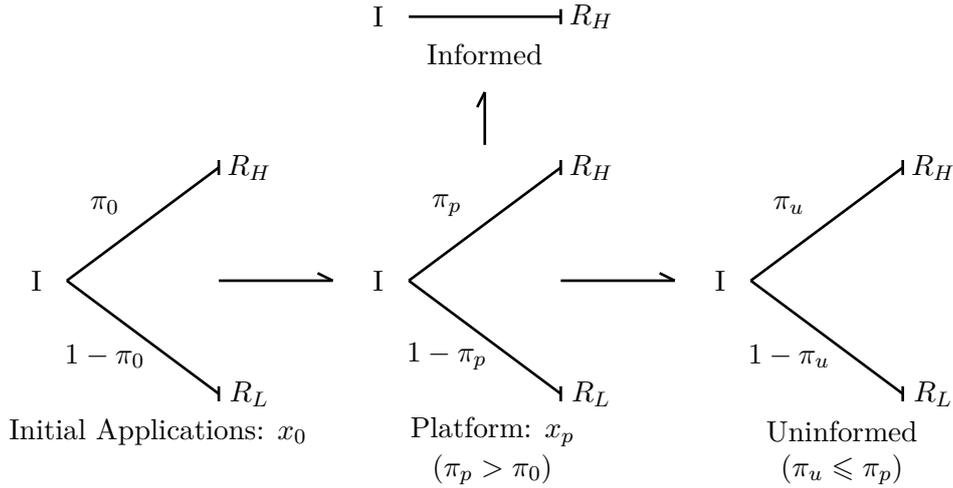
[20]See Figure A.1 in the appendix.

Figure 1: Model Illustration

but non-parametric form: $x_0(p)$ is an increasing function of p, which is in turn determined by the investors' offer price in equilibrium. When the price offered by the investors (and in equilibrium by the platform) is higher, that is, when the interest rate is lower, there are more loan applicants, resembling a higher supply of projects valuable to applicants, and in turn a higher demand for loans. We specify the pricing decision of the platform and investors shortly.

Although the number of initial loan applicants x_0 depends on the prevailing price p on the platform, the initial distribution of good and bad projects π_0 would not change. This assumption serves as a benchmark to help us isolate the key mechanism of the model, and is consistent with marketplace lending targeting and attracting the whole population of potential borrowers.

Platform pre-screening and information distribution. At $t = 0$, the platform first pre-screens the pool of loan applications and lists some of them, making them available for the investors to further screen and finance. For simplicity, we refer to listed loan applications as listed loans, but a listed loan may or may not be eventually financed by an investor. Without loss of generality, we assume that the platform can always screen in a good project but may fail to screen out a bad project. Therefore, the platform effectively chooses the interim probability π_p of a listed loan being good at $t = 1$ before any further investor screening, where $\pi_p \in [\pi_0, 1]$. A higher π_p means that the average listed loan is more likely to be good and is indicative of a higher platform pre-screening intensity. Accordingly, the number of loans being screened in and

12

listed on the platform is

$$x_p = \frac{\pi_0}{\pi_p} x_0 \leqslant x_0 \,.$$

Thus, $\pi_p = \pi_0$ means that the platform simply lists all the loan applications without any pre-screening, and therefore $x_p = x_0$, while $\pi_p = 1$ means that the platform perfectly screens out all the bad projects such that only the $\pi_0 x_0$ good projects are screened in.

We highlight that our current modeling of platform pre-screening is a reduced-form approach to capture a rather complex classification problem facing the platform. In reality, the platform pre-screens loans at both the extensive and intensive margin: it decides which loans to list and then allocates listed loans into different grades representing different levels of loan risk. Although we do not directly model all these aspects, we view a recursive interpretation of our model as consistent with platform pre-screening at both the extensive and intensive margin. For instance, a loan application being rejected as a grade-A listed loan can be further pre-screened as a potential grade-B listed loan, and so on. Since our model is ultimately qualitative, we leave this theoretical aspect to future research and use π_p as the sole summary statistic to capture the outcome of platform pre-screening in a reduced form. Platform pre-screening as we model it is not equivalent to (in fact, it is more than) simply picking a FICO score threshold and listing loans above that threshold; picking a FICO score threshold, however high it is, can never achieve the outcome of perfect pre-screening as described above.

Platform pre-screening is costly. Since platform's pre-screening is implemented through algorithms that are highly scalable, we assume that pre-screening cost only depends on the intensity of pre-screening but not on the number of applications processed. Specifically, the pre-screening cost $C(\pi_p)$ is weakly positive and weakly increasing in π_p, that is, $C(\pi_p) \geqslant 0$ and $C'(\pi_p) \geqslant 0$, which means that it is more costly for the platform to pre-screen loan applications more intensively. Although our results do not depend on the detailed specification of the cost function, we take a parametric form $C(\pi_p) = \frac{1}{2}\kappa(\pi_p - \pi_0)^2$ where $\kappa \geqslant 0$ to facilitate comparative statics.

When pre-screening, the platform assigns an interest rate to listed loans, which is modeled by a price p. We model the platform's loan pricing decision by assuming that the prevailing price

p on the platform is determined by the marginal investor's offer price in equilibrium. This is consistent with both the current practice that the platforms set the interest rate in expectation of investors' financing decisions and the earlier practice of platforms running auctions for setting the interest rate.

The platform also distributes information on listed loans to investors by choosing how many loan characteristics to provide to investors, which affects investors' information cost in evaluating the listed loan. Since these characteristics are provided to the platform by the loan applicants, we assume that it is costless for the platform to provide as few or as many of them to the investors.[21]

Sophisticated investors. Each sophisticated investor is a deep-pocket investor who may buy at most one listed loan on the platform and finance it. She is able to develop her own screening algorithm to further screen the listed loan she faces on the platform, based on the information made available by the platform. Specifically, she can purchase a technology at a cost of $\mu \geqslant 0$, which allows her to know the true type of one listed loan at date 1, and therefore become perfectly informed. We denote the population of sophisticated investors who become informed as ω, where $0 \leqslant \omega \leqslant \Omega$. When becoming informed, a sophisticated investor passes on bad listed loans while offers a price p_i to finance a good loan, which depends on ω in equilibrium.

We highlight that, while for tractability purpose we assume that an informed sophisticated investor is able to know the true type of one listed loan, our qualitative predictions would not change if we instead assumed that the informed sophisticated investor can screen all listed loans but only gets a noisy signal about each loan. All the model intuitions and predictions still hold under the latter interpretation of the model.

We interpret the screening cost μ as capturing both the level of sophistication of the sophisticated investor (a lower μ corresponds to a higher level of sophistication) and the easiness of screening (a lower μ means that it is easier for the sophisticated investors to screen listed loans on the platform, for instance, because the platform distributes more information on listed loans). Since how much information the platform distributes directly affects how costly it is

---

[21]Our model predictions would not change if we instead assume that it is costly for the platform to affect the investors' information cost.

for sophisticated investors to screen the listed loans, we consider μ as a choice variable of the platform in our model.[22] If a sophisticated investor chooses not to pay the technology cost, she remains uninformed.

We assume that informed sophisticated investors can screen and potentially finance good loans before uninformed sophisticated and unsophisticated investors are able to finance the listed loans. This is consistent with the sophisticated investors using faster technology placing orders through API. The screening outcome obtained by an informed sophisticated investor is non-verifiable and non-transferrable: other uninformed investors still do not know the type of a listed loan even if it has been found to be bad by an informed sophisticated investor.

Unsophisticated investors. Each unsophisticated investor is a deep-pocketed investor who may buy at most one loan on the platform. An unsophisticated investor makes a financing decision after informed sophisticated investors move, and he cannot further screen the listed loan that he faces. If a sophisticated investor chooses to remain uninformed, she is essentially identical to an unsophisticated investor in equilibrium.

Although unsophisticated investors cannot screen listed loans, they are still rational investors in the sense that they follow the Bayes' rule and sequential rationality when making their pricing and financing decisions. Since there are infinitely many unsophisticated investors, they are competitive. If they are willing to finance a loan, they have to offer the price p_u that delivers a zero expected profit. When unsophisticated investors only face negative expected profits, they do not finance any loan, in which case $p_u = 0$.

To summarize, the relevant assumptions for unsophisticated investors are that they cannot screen listed loans, and are competitive; the exact population of unsophisticated investors is not crucial. For sophisticated investors, the relevant assumptions are that they can acquire the information technology that allows them to further screen listed loans, and potentially finance them earlier than uninformed investors. In reality, there are not only two types of investors and their level of informational sophistication is distributed across a wide spectrum. Our model represents a useful benchmark that captures the equilibrium resulting from heterogeneity in informational

---

[22]We can effectively decompose μ as $\mu_i + \mu_p$, where μ_i captures sophisticated investors' level of sophistication while μ_p captures how much information the platform distributes.

sophistication, and could be generalized to a more granular cross-section of sophistication.

We could also relax the assumption that each investor can finance at most one loan. This assumption can be easily motivated by the various capital constraints that the investors face. Since our focus is on the interaction between the platform and investor screening, this assumption allows us to circumvent the investors' own portfolio choice problem, which is only tangential to our analysis.

Timeline. At date 0, the platform pre-screens loan applications by choosing π_p, lists the corresponding number x_p of loan applications on the platform, and choose the number of variables provided to investors on the platform that essentially determines μ. The platform also sets the prevailing price p on the platform, which is consistent with the investors' offer price. At date 1, sophisticated investors first decide whether to acquire the information technology and become informed. If a listed loan is found to be good by an informed sophisticated investor, she will finance it at the prevailing price. If the listed loan is found to be bad and thus passed, it remains on the platform and no one knows whether it has already been screened by an informed sophisticated investor. Both passed and unscreened loans listed on the platform then face the competitive and uninformed investors. These uninformed investors may or may not finance the loans based on their updated belief on the remaining listed loans on the platform. At date 2, all financed projects deliver their cash flows, agents consume, and unfinanced projects deteriorate.

Objectives. The platform maximizes the expected volume of eventually financed loans on the platform, regardless of their type, minus its pre-screening cost. In what follows, we call the expected volume of eventually financed loans on the platform simply as volume. This objective function is motivated by the compensation scheme of platforms, which are typically a percentage of volume. Notably, the volume may be different from the amount of listed loans x_p, because all the listed loans may not be financed in equilibrium. The investors maximize expected profits.

We have the following definition of an equilibrium in this economy:

DEFINITION 1. *Given R_H, R_L, I and π_0, the sequential equilibrium is defined as a collection of $\{\pi_p, \mu, p_i, p_u, \omega\}$ such that:*

i) Given π_p, μ and ω, the price p_u gives uninformed investors expected zero profit;

ii) Given π_p, μ and ω, the price p_i maximizes the expected profit of informed investors;

16

iii) Given π_p and μ, the population ω of sophisticated investors find it optimal to acquire the information technology and become informed;

iv) The platform's choices of π_p and μ maximize the expected volume of financed loans minus its pre-screening cost;

v) Agents use the Bayes' rule to update their beliefs and follow sequential rationality.[23]

Based on this equilibrium definition, the prevailing price p set by the platform will be determined by either p_u or p_i depending on who the marginal investor is, as we show later.

3.2 Equilibrium Analysis and Empirical Predictions

We derive the equilibrium by backward induction, which allows us to clarify the main driving forces of the model and progressively derive testable predictions.

Uninformed investors. We first consider uninformed investors' pricing and financing decision on the platform, under any given generic number of listed loans x_p, the corresponding interim belief π_p, information cost μ, as well as the population of informed sophisticated investors $0 \leqslant \omega \leqslant \Omega$. This allows us to illustrate how the participation of uninformed investors affects the volume, and how this participation is, in turn, affected by the participation of informed sophisticated investors.

Since informed investors screen and potentially finance good loans before uninformed investors do so, the number of loans screened but passed by these ω informed sophisticated investors is $(1 - \pi_p)\omega$, and the number of unscreened listed loans (if any) is $(x_p - \omega)_+$, where $f(\cdot)_+ = \max\{0, f(\cdot)\}$. As information is non-transferrable and loans are anonymous, the number of listed loans facing uninformed investors is the sum of the two, $(1 - \pi_p)\omega + (x_p - \omega)_+$. On the other hand, in this pool, only $\pi_p(x_p - \omega)_+$ listed loans are good. Hence, the posterior probability π_p' of a listed loan being good in this pool is

$$\pi_p'(\omega) = \frac{\pi_p(x_p - \omega)_+}{(1 - \pi_p)\omega + (x_p - \omega)_+} \leqslant \pi_p \,, \tag{3.1}$$

which is decreasing in ω. Consequently, the posterior expected value of a listed loan in this pool

---

[23]More formally, each element of the collection is defined on its corresponding information set; we omit the detailed specification of the information sets for simplicity.

is

$$V'(\omega) = \frac{\pi_p(x_p - \omega)_+ \cdot R_H + (1 - \pi_p)(\omega + (x_p - \omega)_+) \cdot R_L}{(1 - \pi_p)\omega + (x_p - \omega)_+} \,, \tag{3.2}$$

which is also decreasing in ω.

Since uninformed investors are competitive, they only enjoy zero profits in equilibrium. Therefore, the price p_u that they offer is

$$p_u(\omega) = \begin{cases} V'(\omega), & \text{if } V'(\omega) \geqslant I, \\ 0, & \text{if } V'(\omega) < I, \end{cases} \tag{3.3}$$

where $V'(\omega)$ is determined in (3.2). If the posterior expected value of a listed loan is higher than the investment requirement I, it will be financed by an uninformed investor. Otherwise, this listed loan will be rejected.

This result has two implications. First, because all the uninformed investors are identical, the remaining listed loans (after informed investors' screening and potential financing of identified good loans) will be either financed as a whole (when $V'(\omega) \geqslant I$) or rejected as a whole (when $V'(\omega) < I$) by the uninformed investors, depending on the interim belief of a listed loan being good or bad. To summarize, the uninformed investors participate on the platform if and only if

$$V'(\omega) \geqslant I \,. \tag{3.4}$$

Second and more importantly, condition (3.2) illustrates an endogenous information asymmetry between informed sophisticated investors and other uninformed investors (including all unsophisticated investors), which leads to an adverse selection problem. From (3.1) and (3.2), we see that both $\pi_p'(\omega)$ and $V'(\omega)$ are decreasing in ω. When more sophisticated investors acquire the information technology and thus successfully pick up more good loans from the platform, the posterior expected value of a listed loan facing the uninformed investors becomes lower. This lowers the price that they would like to offer in equilibrium, potentially driving uninformed investors out of the platform. As will be shown later, this adverse selection problem will also lead to a lower supply of loan applications, hurting the volume even if uninformed investors still participate.

Informed investors. Informed sophisticated investors screen listed loans and finance good loans only. Because they can potentially finance a good loan before uninformed investors do so, their optimal offer price to buy good loans is the uninformed price p_u as long as the uninformed investors participate on the platform. If instead the uninformed investors are out of the platform, informed investors are able to buy good loans at the investment requirement I. Hence, the price for a known good loan by an informed sophisticated investor is determined by

$$p_i(\omega) = \begin{cases} p_u(\omega) = V'(\omega), & \text{if } V'(\omega) \geqslant I, \\ I, & \text{if } V'(\omega) < I, \end{cases} \tag{3.5}$$

where $V'(\omega)$ is determined in (3.2).

An important implication of this result is that sophisticated investors, if becoming informed, outperform unsophisticated investors in any equilibrium where unsophisticated investors participate. While unsophisticated investors always get zero expected profit, an informed sophisticated investor's expected profit of financing an identified good loan is

$$R_H - V'(\omega) = R_H - \frac{\pi_p(x_p - \omega)_+ \cdot R_H + (1 - \pi_p)(\omega + (x_p - \omega)_+) \cdot R_L}{(1 - \pi_p)\omega + (x_p - \omega)_+} \geqslant 0,$$

and the inequality takes a strict form when $\pi_p < 1$. This gives our first empirical prediction:

EMPIRICAL PREDICTION **1.** *Informed sophisticated investors outperform uninformed and unsophisticated investors at any loan price.*

A natural question is when do sophisticated investors choose to acquire the information technology and become informed. In our model, it is straightforward to see that a sophisticated investor acquires the information technology if and only if

$$\pi_p(R_H - p_i(\omega)) \geqslant \mu \tag{3.6}$$

given that ω sophisticated investors become informed. This condition is more (less) likely to be satisfied when the information cost μ becomes lower (higher). This gives our second empirical prediction regarding how sophisticated investors respond to their information cost, which is in

19

turn affected by the platform's policy:

EMPIRICAL PREDICTION **2.** *When their information cost becomes higher, sophisticated investors are less likely to become informed and thus less likely to outperform unsophisticated investors, at any loan price.*

Platform. Given at least some investors participate on the platform and their (latent) pricing decisions, the prevailing price on the platform is given by:

$$p(\omega) = \begin{cases} p_i(\omega) = p_u(\omega) = V'(\omega), & \text{if } V'(\omega) \geqslant I, \\ p_i(\omega) = I, & \text{if } V'(\omega) < I, \end{cases} \tag{3.7}$$

where $V'(\omega)$ is determined in (3.2) and thus $p(\omega)$ is a decreasing function of ω. This suggests that depending on 1) whether sophisticated investors choose to become informed and screen listed loans and 2) whether uninformed investors participate on the platform, there are four types of potential sub-game equilibria given the platform's choices of π_p and μ:

1. Sophisticated investors stay uninformed and none of the uninformed investors participate. In this equilibrium, the information cost is high such that sophisticated investors do not find it profitable to become informed or screen listed loans. At the same time, platform pre-screening is not intensive enough such that the quality of an average listed loan is low, and hence uninformed investors do not find it profitable to invest. Hence, no investor participates on the platform at all, leading to zero volume.

2. Sophisticated investors become informed and screen listed loans, but uninformed investors do not participate. In this equilibrium, the information cost is low such that sophisticated investors find it profitable to become informed. Thus, they screen listed loans and finance any good loans that they identify. But at the same time, due to a low pre-screening intensity and adverse selection caused by informed investor participation, the quality of an average listed loan is low and thus uninformed investors do not find it profitable to invest. In this case, the prevailing price on the platform is the investment requirement I offered by informed sophisticated investors, and the equilibrium volume depends on the population of sophisticated investors: Ω sophisticated investors can jointly identify and

finance $\pi_p \Omega$ good loans up to all the good loans listed on the platform, that is, $\pi_0 x_0(I)$. Notably, all the loans financed are good loans.

3. Sophisticated investors become informed and screen listed loans, and uninformed investors also invest. In this equilibrium, the information cost is low such that sophisticated investors find it profitable to become informed. Thus, they screen listed loans and finance any good loans that they identify. At the same time, the pre-screening intensity is high enough such that the quality of an average listed loan is high enough for uninformed investors to find it profitable to invest, and they finance all the remaining listed loans on the platform. Adverse selection is still present in this equilibrium, leading to a relatively lower prevailing price on the platform, but it is not severe enough to discourage uninformed investors from investing. In this case, the prevailing price on the platform is $p(\Omega)$, resulting in $x_0(p(\Omega))$ initial applications. As a result, the equilibrium volume is $\frac{\pi_0 x_0(p(\Omega))}{\pi_p}$, which includes all the listed loans regardless of their type.

4. Sophisticated investors stay uninformed and do not screen loans, but uninformed investors invest and finance all the listed loans. In this equilibrium, the information cost is high enough such that sophisticated investors do not find it profitable to become informed or screen listed loans. At the same time, the pre-screening intensity is high enough such that the quality of an average listed loan is high enough. Consequently, uninformed investors find it profitable to invest, and they finance all the listed loans on the platform. In this case, there is no adverse selection, leading to a relatively higher prevailing price on the platform $p(0)$. Under this price, the amount of initial applications is $x_0(p(0))$, and the equilibrium volume is $\frac{\pi_0 x_0(p(0))}{\pi_p}$, which includes all the listed loans regardless of their type.

In the description above, we order the four types of sub-game equilibria by the equilibrium volume: the volume increases from type-1 equilibrium to type-4 equilibrium. In the appendix, we formally show that the economy must end up in one of these four potential types of sub-game equilibria and provide detailed mathematical derivations of the four types of sub-game equilibria. This analysis can be formally summarized by the following Proposition 1:

PROPOSITION **1.** *The economy admits four types of sub-game equilibrium. The respective volume in each type of sub-game equilibrium is given as follows:*

Equilibrium Volume of Loans Financed

	High μ	Low μ
Low π_p	0	$\min\{\pi_0 x_0(I), \pi_p \Omega\}$
High π_p	$\dfrac{\pi_0 x_0(p(0))}{\pi_p}$	$\dfrac{\pi_0 x_0(p(\Omega))}{\pi_p}$

where we call the top-left equilibrium type-1 equilibrium, top-right type-2 equilibrium, bottom-right type-3 equilibrium, and bottom-left type-4 equilibrium. The equilibrium volumes satisfy

$$0 < \min\{\pi_0 x_0(I), \pi_p \Omega\} < \frac{\pi_0 x_0(p(\Omega))}{\pi_p} < \frac{\pi_0 x_0(p(0))}{\pi_p}. \tag{3.8}$$

Given the volume in each type of sub-game equilibrium, we follow a backward induction to solve for the platform's optimal choice of prescreening intensity π_p and investor information cost μ, and characterize the full equilibrium. The following Proposition 2 presents the result, where the exact mathematical expressions of the optimal π_p and μ are given in the appendix.

PROPOSITION **2.** *In any full equilibrium where the platform optimally chooses pre-screening intensity π_p and investor information cost μ, only type-2 and type-4 sub-game equilibrium can happen. Specifically, there exists two thresholds $0 < \underline{\kappa} < \overline{\kappa}$ such that:*

i). If platform pre-screening cost is high in the sense that $\kappa \geqslant \overline{\kappa}$, the platform optimally chooses a low π_p and a low μ, ending up in a type-2 sub-game equilibrium where sophisticated investors become informed and only they invest, and the volume is $\min\{\pi_0 x_0(I), \pi_p \Omega\}$.

ii). If platform pre-screening cost is low in the sense that $\kappa \leqslant \underline{\kappa}$, the platform optimally chooses a high π_p and a high μ, ending up in a type-4 sub-game equilibrium where sophisticated stay uninformed but all uninformed investors invest, and the volume is $\frac{\pi_0 x_0(p(0))}{\pi_p}$.

iii). If platform pre-screening cost is intermediate in the sense that $\underline{\kappa} < \kappa < \overline{\kappa}$, the platform optimally ends up in either a type-2 or type-4 sub-game equilibrium, depending on which gives the platform a higher expected payoff.

22

Proposition 2 illustrates the key trade-offs facing the platform's optimal design. We elaborate them as follows.

When platform pre-screening cost is relatively high, it is hard for the platform to attract uninformed unsophisticated investors, because the quality of an average loan listed on the platform is not high enough. Thus, the only way to get some listed loans financed is to rely on informed sophisticated investors' participation. In this sense, sophisticated investors help the platform boost its volume, because they always finance good loans if they become informed. To ensure this, the platform optimally chooses a low π_p but also more importantly a low μ so that sophisticated investors find it profitable to become informed and screen the listed loans.

Instead, when platform pre-screening cost is relatively low, it becomes relatively easy for the platform to attract uninformed investors, because the quality of an average loan listed on the platform may become high. In this case, sophisticated investors, if they become informed, no longer help boost volume but rather introduce an adverse selection problem. Because their screening lowers the quality of an average loan facing uninformed investors, the prevailing price on the platform becomes lower, leading to a lower amount of loan applications in the first place. If adverse selection is severe enough, it may even completely deter uninformed investors from investing at all. Hence, to manage this adverse selection problem, the platform optimally chooses a high π_p but also more importantly a high μ so that sophisticated investors do not find it profitable to become informed.

Theoretically, the different types of sub-game equilibria are reminiscent of the different types of equilibria in Fishman and Parker (2015) and Bolton, Santos and Scheinkman (2016), which also feature the two competing forces that informed investors' expertise guides more efficient production decisions but at creates adverse selection simultaneously. Our contribution is to endogenize the initial supply of projects and also embed these two forces to a realistic platform design problem in which the platform can use various policies to actively trade off these two forces. The platform's pre-screening further introduces a new trade-off that a more intensive pre-screening encourages uninformed investor participation but at the same time screens out more applications that would have been potentially financed.

Proposition 2 can further generate predictions, many of which are testable and helpful in

terms of understanding some seemingly conflicting facts in the recent development of marketplace lending.

One important aspect to consider is what happens when the platform pre-screening cost κ decreases as the platform develops along its life cycle. Physically, platform pre-screening cost decreases over time because of the newly available information technologies and screening algorithms.[24] Moreover, because platforms actively observe sophisticated investors' screening and investing behavior, a feedback effect helps the platforms to further improve their pre-screening models, lowering the pre-screening cost as well. This feedback effect, although not directly modeled in our framework, is likely to be particularly strong when the presence of informationally sophisticated investors becomes high.

COROLLARY **1.** *As the platform pre-screening cost κ goes from high to low, the platform optimally increases investor screening cost μ.*

Corollary 1 is a direct implication of Proposition 2 when the equilibrium switches from type-2 to type-4; it reflects the platform's active management of the adverse selection problem introduced by sophisticated investors. Because the information cost for sophisticated investors becomes higher, they are less likely to become informed and therefore less likely to outperform. Thus, Corollary 1 leads to the following two testable predictions, which we test and discuss in the empirical analysis.

EMPIRICAL PREDICTION **3.** *If both sophisticated investors and unsophisticated investors participate and finance listed loans, the platform may optimally increase the information cost of sophisticated investors by distributing few variables to investors.*

EMPIRICAL PREDICTION **4.** *If both sophisticated investors and unsophisticated investors participate and finance listed loans, sophisticated investors' outperformance decreases as the platform develops over time.*

Another natural question is what happens when platform pre-screening cost becomes negligible, as some observers argue today. Would a maximal pre-screening intensity π_p be optimal

---

[24]See Lending Club Investor Day 2017 Presentation available at ir.lendingclub.com/Cache/1001230258.pdf.

for the platform when pre-screening is costless? The following corollary shows that the answer is no.

COROLLARY **2.** *Even if $\kappa = 0$, the platform optimally chooses $\pi_p < 1$.*

Corollary 2 stems from the fact that when type-4 sub-game equilibrium is achieved, further increasing π_p only hurts volume. This effect results from the platform's objective to maximize volume rather than the quality of loans that get financed, which we view as realistic. As we suggested earlier, this objective leads to a new trade-off: a high pre-screening intensity encourages uninformed investor participation but if pre-screening intensity is too high, many loan applications that would have been financed end up being not listed. In reality, the platform optimally calibrates its pre-screening intensity to varying economic conditions as captured by R_H, R_L, and π_0 in our model. This may be reflected by a fluctuation of pre-screening intensity over time, which speaks to the following empirical prediction:

EMPIRICAL PREDICTION **5.** *Even if the physical cost of platform pre-screening is low, the platform optimally chooses an intermediate level of pre-screening intensity, which may fluctuate over time.*

Our model delivers other important implications regarding the life cycle of a platform. As discussed before, condition (3.8) suggests that the volume on a platform always increases as the platform's pre-screening cost decreases. In addition, it is also straightforward that the equilibrium volumes of good loans from sub-game equilibrium 1 to 4 satisfy

$$0 < \min\{\pi_0 x_0(I), \pi_p \Omega\} < \pi_0 x_0(p(\Omega)) < \pi_0 x_0(p(0)), \tag{3.9}$$

suggesting that the volume of good loans on a platform also increases as the platform's pre-screening cost decreases. Conditions (3.8) and (3.9) jointly imply that the ex-post quality of an average financed loan may decrease as the platform develops. These implications fundamentally stem from the platform's objective to maximize total volume but not the percentage of good loans. To directly test these predictions is challenging because it is hard to empirically isolate the forces described in our model from other un-modeled economic conditions. However, these

implications are consistent with the general trends that the volume in lending marketplaces has been growing at a rapid pace but the ex-post delinquency rate has been also rising in the past several years.

4 Data and Investor Sophistication

Our empirical analysis relies on a proprietary dataset, which combines investor data covering both sophisticated and non-sophisticated retail investors, with borrowers' data published by the platforms. This dataset allows us to empirically test the predictions derived in the previous section.

4.1 Data

Our investor-level data is provided by LendingRobot, a leading robo-advisor for retail investors on lending marketplaces.[25] The data covers all transactions executed by LendingRobot users between January 2014 and February 2017, which represents more than $120 million invested on the two major lending platforms, LendingClub and Prosper, as well as all historic transactions from portfolios monitored by the company. LendingRobot provides an automated investment tool for its clients which relies on a sophisticated screening model calibrated on historical data from the platforms.[26] This tool allows to execute purchase orders at high speed through an API, which is key for accessing loans in high demand. LendingRobot also offers a free monitoring tool that can be linked with current portfolios on Lending Club and Prosper. The use of the monitoring tool and the investment tool are independent, which means that we observe portfolios invested with the help of LendingRobot investment tool, as well as portfolios built by investors themselves.[27]

The LendingRobot data is organized at the investor level, as shown in Figure 2. We access a set of variables at each level of this data structure.

---

[25]LendingRobot was acquired by NSR Invest, its main competitor, in 2017.

[26]For more detail on LendingRobot credit model, refer tohttp://blog.lendingrobot.com/research/predicting-the-number-of-payments-in-peer-lending/ .

[27]Monitored-only accounts might be invested with the platform automatic investment strategy, for investment Lending Club *automated investing* feature, or at the investor's discretion.
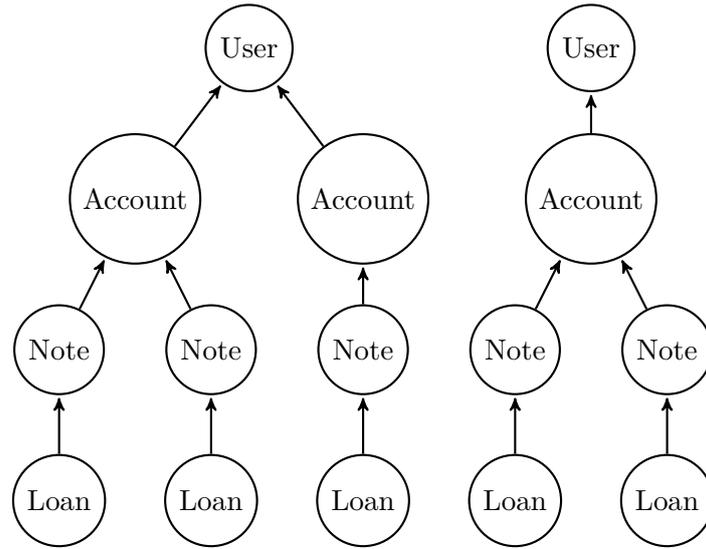
Figure 2: LendingRobot Data Structure

- **User**: Each user represents a distinct physical investor. A user can have one or several accounts.

- **Account**: An account represents a portfolio of notes an investor holds on a single lending platform: Lending Club or Prosper. A *monitor-only* account is an account where LendingRobot only monitors the portfolio but do not execute notes purchase through its technology. In a *robot* account, notes purchase are executed by the LendingRobot investment tool, which combines a screening model and automatically places orders through an API. In an *advanced* account, investors are implementing their own screening criteria, combined or not with LendingRobot investment credit model, and rely on LendingRobot to automatically execute the orders when relevant loans appear on the platform.

- **Note**: Retail investors invest in notes, each of which is backed by a single consumer loan. This form of securitization was developed to comply with SEC regulation, and to allow small investment in loans that have amounts over 10,000 USD. The information available at the note level is its nominal value, which is 25 USD or 50 USD, and the underlying loan identifier.

- **Loan**: Each loan is associated with a large set of financial characteristics of the borrower at the loan issuance, made public by lending platforms. These variables include loan

amount, FICO score, debt-to-income ratio, employment length, three-digit zip-code, and many others. For brevity, we refer interested readers to Morse (2015) for the description of loan-level data, which is standardly used in the literature. We can observe these variables for all loans issued by Lending Club and Prosper, including the ones in which LendingRobot investors participate.

4.2 Investor Sophistication

Our data provide us with portfolios of retail investors with heterogenous levels of sophistication, which is key to our study. While we focus on the retail segment, the heterogeneity in sophistication within this segment is arguably comparable to the heterogeneity of sophistication within the institutional investor segment, for example, traditional mutual and pension funds usually buy loans from the platform passively while hedge funds screen loans actively before investing,[28] and LendingRobot indeed performs hedge-fund-like order execution for its *robot* accounts.

Within our data, *monitor-only* accounts are the less sophisticated ones as they do not implement automated trading, and have therefore not internalized the disadvantage they face versus more sophisticated investors that place order at high speed.[29] *Robot* accounts represent sophisticated investors, as they are invested through the LendingRobot algorithmic tool. While some *advanced* accounts might be potentially even more sophisticated than *robot* accounts, because *advanced* accounts rely mostly on their own screening criteria, they are at risk of making mistakes compared to the *robot* benchmark.[30]

4.3 Summary Statistics

Table 1 provides summary statistics on our data set. Lending Club represents the largest platform for the investors in our sample both in terms of the number of accounts and the amount invested. The relative size of the Prosper universe on LendingRobot is broadly consistent with

---

[28]See "Hedge Funds Pursue Alternative Lending," the *Financial Times*, November 2, 2014.

[29]Monitored-only investors are selected on having registered to LendingRobot, which suggests a higher sophistication than fully naive investors. However, this potential selection effect can only bias against finding differences between the two types of portfolios in our empirical analysis.

[30]Per discussion with LendingRobot, a few of the investors with an advanced account are institutional investors, such as family offices, but the majority are individual investors. LendingRobot displays a warning to investors that they proceed at their own risk, when they select an advanced account.

the difference in size between the two platforms. While *robot* accounts are the most represented type of accounts, *advanced* accounts are, on average, larger. The distribution of portfolio size is skewed, with a few investors having invested more than one million dollars, driving the average amount invested significantly above the median amount invested. Accounts are on average modestly tilted towards riskier loans compared to the overall platform average, as exhibited by a higher average interest rate of portfolios across the board.

[Insert Table 1]

5   Empirical Analysis

In this section, we formally test the empirical predictions derived from our theoretical framework.

5.1   Investor Screening

Our theoretical analysis relies on the premise that some informationally sophisticated investors actively screen listed loans on the platforms. Thus, the first step of our empirical analysis is to study whether more sophisticated investors indeed screen differently, which is implied by Empirical Prediction 1.

For this purpose, we focus on loans invested by different LendingRobot account types and conduct the following empirical test at the loan level, separately for Lending Club and Prosper loans over our sample period from 2014 to 2016.

We define three indicator variables $TypeAccount_i$ equal to one if 1) at least two *robot* accounts are invested in this loan 2) at least two *advanced* accounts are invested in this loan 3) at least two *monitored-only* accounts are invested in the loan. We use these indicator variables as left-hand-side variables, and run linear probability regressions on borrower characteristics.[31] We include interest rate fixed effects to control for loan price, and therefore focus on screening that aims at picking the loans with the lower default risk within each grade on Prosper, and sub-grades on

---

[31]The results are similar to logit regressions.

Lending Club.[32]

$$Prob(TypeAccount_i = 1) = \beta \times BorrowerCharacteristics + IR_i + m_t + \epsilon_i, \qquad (5.1)$$

where $BorrowerCharacteristics$ are provided by the lending platform, IR_i are fixed effects for each level of interest rate paid to the investor, i.e. the loan price at issuance, m_t are month fixed effects, and ϵ_i is the error term. Table 2 displays the regression coefficients.

[Insert Table 2]

There are several key takeaways from this analysis. First, a large number of specific borrower characteristics strongly predict an investment by *robot* investors. This suggests that LendingRobot's screening model considers that the risk these characteristics are associated with are either mis-estimated, or not fully incorporated by Lending Club or Prosper when listing the loans at a given price.[33]

For variables positively (negatively) correlated with risk, a positive coefficient in column 1 suggests that LendingRobot's screening model considers that the platform over-penalizes (under-penalizes) this borrower characteristic. Conversely, a negative coefficient points out to characteristics that are under-penalized (over-penalized) when the platform lists the loan.

For instance, the positive coefficients on loan amount or revolving utilization in column 1 suggest that LendingRobot's screening model considers that Lending Club penalizes excessively borrowers that take a large loan or have large revolving balance. On the other hand, the positive coefficients on annual income and FICO score indicates that LendingRobot value more these positive attributes than Lending Club does. The purpose of the loans, although ambiguous in terms of risk, also appears to be an important screening criterion by LendingRobot investors.

When comparing column 1 to column 4, we observe that the regression coefficients are different for the two platforms. For some characteristics, they have opposite signs. This result stresses that investor screening can only be interpreted relative to platform pre-screening at the

---

[32]We abstract from the question on whether loans from the same grade or sub-grades are on average fairly priced.

[33]For instance, the granularity of the interest rate scale mechanically creates heterogeneity in risk within the same price.

intensive margin, and reveals that LendingRobot believes that the pre-screening models differs significantly between Lending Club and Prosper.

To further shed light on how investors with different levels of informational sophistication screen differently, we also look at whether screening behavior varies by type of investor, thanks to the segmentation offered by our data.

First, *advanced* investors have screening criteria that are largely consistent with the screening criteria of LendingRobot's screening model: the coefficients in column 2 are comparable to the ones in column 1, which suggests that some of these investors combine LendingRobot screening model with their own screening criteria.

However, *monitored-only* investors appear to screen significantly differently from *robot* investors, as many characteristics have opposite prediction in terms of participation of *monitored-only* vs. *robot* investors on both platforms. The correlation of investment by these investors with certain characteristics is not necessarily causal, since it might result from the loans with the opposite characteristics having been already been picked up and financed by more sophisticated investors. For instance, if more sophisticated investors massively invest in loans with high FICO scores within a sub-grade, less sophisticated investors will be mechanically more likely to invest in loans with a relatively lower FICO score. This is also consistent with our view that *monitored-only* investors represent the less sophisticated segment of investors who eventually pick up the loans being passed by more sophisticated investors.

In Table IA.1 in the appendix, we run a similar analysis as a robustness check using the log of the number of accounts, for each type of accounts, as the dependent variable. We find consistent coefficients.

5.2 Investor Performance

Having documented that sophisticated investors actively screen loans, and do so differently than less sophisticated investors, we now test whether active screening allows more sophisticated investors to consistently outperform average and unsophisticated investors on the platform.

Specifically, we first test Empirical Prediction 1 by investigating whether, controlling for price, having sophisticated investors participate in a loan predicts a lower likelihood of default.

We measure performance at the loan level, as is currently done in the literature (Iyer, Khwaja, Luttmer and Shue, 2015, for example), by using an indicator variable for the loan being in default or charged-off. A loan enters default status when it is 121+ days past due and enters charge-off status after 150 days.[34]

We first plot the share of defaulted loans by sub-grades as of December 2017 for the entire Lending Club platform, as well as the subsets of loans where at least two *robot*, *advanced* and *monitor-only* accounts are invested for loans issued between 2014 and 2016.

[Insert Figure 4]

This figure documents the consistent gain in performance achieved by *robot* and *advanced* investors across the whole spectrum of risk. *Monitor-only* performance slightly outperforms the whole platform.

We perform regressions to further dig into this outperformance. We regress the performance indicator on indicator variables for participation by different types of investors, while including both interest rate fixed effects and month of issue fixed effect. This specification allows us to measure to what extent investment by LendingRobot investors predicts a lower likelihood of default of a loan, controlling for its price and its monthly vintage. We run these regressions as OLS due to the high number of fixed effects and to interpret the economic magnitude of the coefficients.[35]

We run the following specification:

$$Prob(ChargedOff = 1)_i = \beta_1 \times \mathbb{1}_{TypeAccount} + IR_i + m_t + \epsilon_i, \tag{5.2}$$

where $ChargedOff$ is an indicator variable for the loan to be in default or charged-off, $\mathbb{1}_{TypeAccount}$ is an indicator variable equal to one if at least two accounts of a given type are invested in the loan. Loan maturity (either 3 or 5 years) is implicitly controlled for by interest rate fixed effects, as interest rate varies with maturity within the same sub-grade. Table 4 displays

---

[34]We rely on loan status data from Lending Club as of December 2017. While some of the loans have not yet matured, the majority of loan defaults happens in their first year, as documented on the platform websites and in the literature (Morse, 2015). We also control for monthly vintage in our regressions.

[35]The results are robust under a logit specification.

the coefficient of these regressions. The type of account is *robot* accounts in column 1, *advanced* accounts in column 2, and *monitored-only* accounts in column 3. The following columns interact $\mathbb{1}_{TypeAccount}$ with year fixed effects, and with loan grade fixed effects.

[Insert Table 4]

Column 1 documents that the default rate of loans in which *robot* accounts invest is significantly lower than for the whole Lending Club population over the 2014-2016 period, while controlling for loan prices. This difference in default rate is measured within sub-grade, and therefore sophisticated investor out-performance cannot result from composition effect across the sub-grades. The economic magnitudes are particularly large, as the OLS specification suggests a reduction by more than three percentage point of the default rate, to compare to a 14% average default rate for the whole sample. The LendingRobot screening model therefore translates into a reduction in average default rate of more than 20%. Column 2 shows that over the period, *advanced* investors' outperformance is even larger than the one of *robot* accounts, when controlling for monthly vintage. By contrast, column 3 shows that participation by *monitored-only* investors only weakly predicts a lower default rate, with an economic magnitude four times smaller.

Column 7 and 8 illustrate how the reduction in default rate obtained by *robot* and *advanced* accounts is broadly increasing in the risk of the loan, as should be expected as riskier loans have higher baseline default rates, and are potentially harder to screen/more information sensitive.

This analysis empirically establishes that platform pre-screening leaves significant room for screening by investors to generate over-performance. The screening model developed by LendingRobot, which only relies on hard information provided by the platform to all investors and fully drives the investment decision for its *robot* accounts, has statistically and economically significant predictive power over loan default rates beyond the Lending Club risk ratings.

5.3 Screening Cost and Investor Performance

The outperformance of sophisticated investors raises the question of whether and how the platform manages the adverse selection that it generates. The choice of the information set provided

to investors has a direct impact on their screening cost, as investors have to exert more effort to identify good loans when the information they access is more restricted.[36] We therefore investigate the relationship between the information set available on the platform and investor screening performance. We find empirical evidence consistent with Empirical Prediction 2 that when investors' screening cost increases, the performance of sophisticated investors is reduced.

For this purpose, we exploit the unexpected shock to investor information set described in Section 2.2 with a difference-in-difference setting. While the shock affects all investors, it affects investors differentially, as unsophisticated investors are unlikely to be using the 50 variables that get removed, or at least are doing so to a smaller extent that sophisticated investors.

We run the following specification:

$$Prob(ChargedOff = 1)_i = \beta_1 \times \mathbb{1}_{robot} + \beta_2 \times \mathbb{1}_{robot} \times Post$$
$$+ \beta_3 \times \mathbb{1}_{advance} + \beta_4 \times \mathbb{1}_{advance} \times Post$$
$$+ \beta_5 \times \mathbb{1}_{monitor} + \beta_6 \times \mathbb{1}_{monitor} \times Post + IR_i + m_t + \epsilon_i \quad (5.3)$$

where $\mathbb{1}_{robot}$ is an indicator variable equal to one if at least two *robot* accounts are invested in the loan i, $\mathbb{1}_{advanced}$ is an indicator variable equal to one if at least two *advanced* accounts are invested in the loan, $\mathbb{1}_{monitor}$ is an indicator variable equal to one if at least two *monitor-only* accounts are invested in the loan, $Post$ is an indicator variable for being in the period after the shock to the information set, IR_i are interest rate fixed effects, m_t are month fixed effects and ϵ_i is the error term. The choice of a linear probability model is again motivated by the large number of fixed effects we use, and also facilitates interpreting the economic magnitude.

Results are also displayed in Table 4. Column 1 implements this specification for all Lending Club loans issued in the period spanning three months before the month of the change and three months after. Column 2 restrict the loan universe to loans with grade C or below. Column 3 restricts the period to two months before the month of the change and two months after. Column 4 restricts the sample to loans that have either *robot* accounts or *monitor-only* accounts invested

---

[36]This cost goes to infinity as the information is reduced, as in the absence of information it is impossible to further screen.

in, thereby implementing a difference-in-differences between the two groups.

This analysis reveals that the increase in investor screening cost significantly impacts the screening performance of *robot* accounts, as well as *advanced* accounts to a lower extent. On the other hand monitor-only accounts are not significantly affected. This result is robust to all specifications, and is more pronounced for lower grade loans. The magnitude of the effect is large: the outperformance of *robot* investors in the period immediately preceding the change in the information set drops by more than half at the time of the shock to the information set.[37] We rationalize this reduction in sophisticated investors' performance by the platform "evening the playing field" to actively manage the potential adverse selection problem introduced by sophisticated investors. When comparing the effect between *robot* and *advanced*, we observe that *advanced* accounts are less affected than *robot* accounts, which is consistent with their lower screening ability over the previous period.

To further pin-down the causal impact of the information set change on sophisticated investors' performance, we implement an event-study type analysis to investigate the exact timing of the change in performance. This zooming-in is important to rule out that an underlying trend on the platform, for instance a gradual improvement in the platform pre-screening ability, may drive our result. For this purpose, we implement two regressions with two different samples: all fractional Lending Club of the period, the control group, and all the Lending Club loans where at least two *robot* account participated in, the treatment group. For both regressions, the dependent variable is an indicator variable equal to one if the loan is charged-off as of December 2017, and the explanatory variables are month fixed effects. We control for interest rate fixed effects to alleviate concerns over potential composition changes during the sample period. The reference period for the month fixed effects is month 5 and 6 after the change in information set.

Because the constant of each regression is absorbed by the interest rate fixed effects, each line only speaks to the relative evolution of default rate in each sample over that period, and not to the initial level of charged-off in each loan population before the shock. Results are displayed

---

[37]This drop in performance could not be immediately observed at the time of the change, as it takes time for loan performance to be revealed to investors.

in Figure 5.

This figure illustrates how the overall loan performance on the Lending Club platform is unaffected, while the performance of the loans screened by *robot* accounts sharply deteriorates at the time of the shock. The sharpness in the change of performance, as well as its synchronicity with the change in the information set, are supportive of a causal interpretation.

[Insert Figure 5]

Last, we ensure that the change in performance does not result from a sharp change in the composition of listed loans that investor face on the platform that would not be captured by the interest rate fixed effects. For this purpose, we compare the main characteristics of the listed loans before and after the change in the information set. Table IA.2 in the appendix illustrates how the pool of listed loans remains unchanged after the shock on borrowers' observable characteristics.

We conclude this difference-in-difference analysis by pointing out that the decision for Lending Club to reduce information provision to investors is consistent with our Empirical Prediction 3, suggesting the platform is actively managing the potential adverse selection problem introduced by sophisticated investors.

5.4 Trends in Platform Pre-screening Intensity and Investor Performance

As a final step of our empirical analysis, we study the time-series evolution of platform pre-screening and investor relative performance. We find results consistent with Empirical Predictions 4 and 5 in the model.

5.4.1 Platform Pre-screening

We first explore whether the platforms have been changing the quality thresholds for borrowers to get listed. We plot the evolution of the share of borrowers on the Lending Club platform whose FICO score is below 670 and 660, and whose debt-to-income ratio is above 30% and 35%. Results are displayed in Figure 6. This figure is consistent with Lending Club relaxing

its acceptance standards over time, which is in line with the overall increase in loan prices from Figure 3.

[Insert Figure 6]

On the intensive margin, we study whether the explanatory power of platform grades over loan default has been evolving over time. For this purpose, we build ROC curves - which graph the true positive rate against the false positive rate - obtained when using Lending Club sub-grades as a predictor of loans being charged-off as of December 2017. The larger the area below a ROC curve, the more precise the predictor is. The test is computed separately for loans issued and graded in 2014, 2015 and 2016. Results are displayed in Figure 7. This figure documents how Lending Club sub-grades and Prosper Grades have been gaining explanatory power towards default over time. This evolution suggests an improvement in the accuracy of platform pre-screening. Lending Club sub-grades also appear to better predict default than Prosper grade, as the areas below the ROC curves are larger for Lending Club.[38]

[Insert Figure 7]

Overall, these two figures support the view that platforms appeared to loose the minimum standard for a loan application being listed, while constantly improve the accuracy of the classification within listed loans. These trends are consistent with Empirical Prediction 5 in the sense that platforms never choose too high pre-screening intensity but are constantly calibrating it, as long as less sophisticated investors still find it profitable to stay on the platforms.

5.4.2 Evolution of Screening Performance over 2014-2016

We finally explore the evolution of investor screening performance over our sample period, and find that sophisticated investor out-performance has been decreasing over time. While the change in investor screening cost we investigate in section 5.3 plays an important role in the evolution between 2014 and 2015, it cannot account for the whole trend, which likely results mostly from the improvement in pre-screening by the platforms we previously document.

---

[38]Using grades instead of sub-grades for Lending Club still yields a better predictor than Prosper grades.

For each year, we plot the share of defaulted loans by sub-grades (as of December 2017) for the entire Lending Club platform, as well as the subsets of loans where at least two *robot*, *advanced* and *monitor-only* accounts are invested for loans.

[Insert Figure 8]

The figure shows that the out-performance of *robot* and *advanced* investors appear to decrease with time,[39] consistent with Empirical Prediction 4. Again, we view these patterns consistent with the platforms actively managing the potential adverse selection problem introduced by more sophisticated investors.

Columns 4 to 6 of Table 4 also explore the time-series of investor type performance by interacting the indicator for having a given type of investors participating in a loan with year fixed effects. For both *robot* and *advanced* investors, outperformance is significantly stronger in 2014 than in 2015, and even more so than in 2016. *Robot* accounts outperform *advanced* accounts in 2014, but get overcome in the two following years. On the other end, there is no time trend for *monitor-only* accounts, and the interaction terms make their out-performance statistically insignificant for any given year.

6 Conclusion

Different from the conventional banking paradigm, one prominent feature of the burgeoning marketplace lending (i.e., peer-to-peer lending) is that investors conduct tasks traditionally performed by banks. Lending platforms pre-screen loan applications moderately, while investors, heterogeneous in their level of sophistication, further screen and decide whether or not to finance the loans.

In this paper, we theoretically argue that the participation of informationally sophisticated investors improves lending outcomes but creates an endogenous adverse selection problem. In maximizing loan volume, the platform trades off these two forces. Thus, intermediate levels of platform pre-screening intensity and information provision to investors are optimal. Using

---

[39]The 2016 chart should however be interpreted with a grain of salt as a large share of the loans from that year have not matured as of December 2017.

novel investor-level data, we empirically show that despite facing the same information set, more sophisticated investors screen loans differently from less sophisticated ones and significantly outperform. However, the outperformance shrinks when the platforms reduce the information set available to investors. These empirical facts are consistent with platforms dynamically managing adverse selection through platform design and screening intensity. Since the platform maximizes total volume of loans financed, a lower adverse selection by sophisticated investors always implies higher volume but not necessarily a higher quality for an average loan being financed.

We leave a number of interesting questions for future research, including the welfare and financial stability implication of marketplace lending. Our study represents a significant step forward to tackle these broader questions.

References

BALYUK, TETYANA (2017). "Financial Innovation and Borrowers: Evidence from Peer-to-Peer Lending." Working Paper.

BIAIS, BRUNO, THIERRY FOUCAULT and SOPHIE MOINAS (2015). "Equilibrium Fast Trading." *Journal of Financial Economics*, 116:292-313.

BOLTON, PATRICK, TANO SANTOS and JOSE SCHEINKMAN (2016). "Cream Skimming in Financial Markets." *Journal of Finance*, 71:709-736.

BUCHAK, GREG, GREGOR MATVOS, TOMASZ PISKORSKI, and AMIT SERU (2017). "Fintech, Regulatory Arbitrage, and the Rise of Shadow Banks." Working Paper.

DIAMOND, DOUGLAS and PHILIP DYBVIG (1983). "Bank Runs, Deposit Insurance, and Liquidity." *Journal of Political Economy*, 91: 401-419.

DUARTE, JEFFERSON, STEPHAN SIEGEL, and LANCE YOUNG (2012). "Trust and Credit: the Role of Appearance in Peer-to-Peer Lending." *Review of Financial Studies*, 25: 2455-2484.

FISHMAN, MICHAEL and JONATHAN PARKER (2015). "Valuation, Adverse Selection, and Market Collapses." *Review of Financial Studies*, 28: 2575-2607

FRANKS, JULIAN, NICOLAS SERRANO-VELARDE, and OREN SUSSMAN (2016). "Marketplace Lending, Information Aggregation, and Liquidity." Working Paper.

GLODE, VINCENT, RICHARD GREEN and RICHARD LOWERY (2012). "Financial Expertise as an Arms Race." *Journal of Finance*, 67: 1723-1759.

GORTON, GARY and GEORGE PENNACCHI (1990). "Financial Intermediaries and Liquidity Creation." *Journal of Finance*, 45: 49-71.

IYER, RAJKAMAL, ASIM KHWAJA, ERZO LUTTMER, and KELLY SHUE (2015). "Screening Peers Softly: Inferring the Quality of Small Borrowers." *Management Science*, 62: 1554-1577.

LISKOVICH, INESSA and MAYA SHATON (2017). "Borrowers in Search of Feedback: Evidence from Consumer Credit Markets." Working Paper.

MARIOTTO, CARLOTTA (2016). "Competition for lending in the Internet era: The case of Peer-to-Peer Lending Marketplaces in the USA." Working Paper.

Morse, Adair (2015). "Peer-to-Peer Crowdfunding: Information and the Potential for Disruption in Consumer Lending." *Annual Review of Finance and Economics*, 7: 463-482.

Paravisini, Daniel, Veronica Rappoport, and Enrichetta Ravina (2016). "Risk Aversion and Wealth: Evidence from Person-to-Person Lending Portfolios." *Management Science*, 63: 279-297.

Rochet, Jean-Charles and Jean Tirole (2006). "Two-Sided Markets: A Progress Report." *The RAND Journal of Economics*, 37: 645-667.

Yang, Ming and Yao Zeng (2017). "Financing Entrepreneurial Production: Security Design with Flexible Information Acquisition." Working paper.
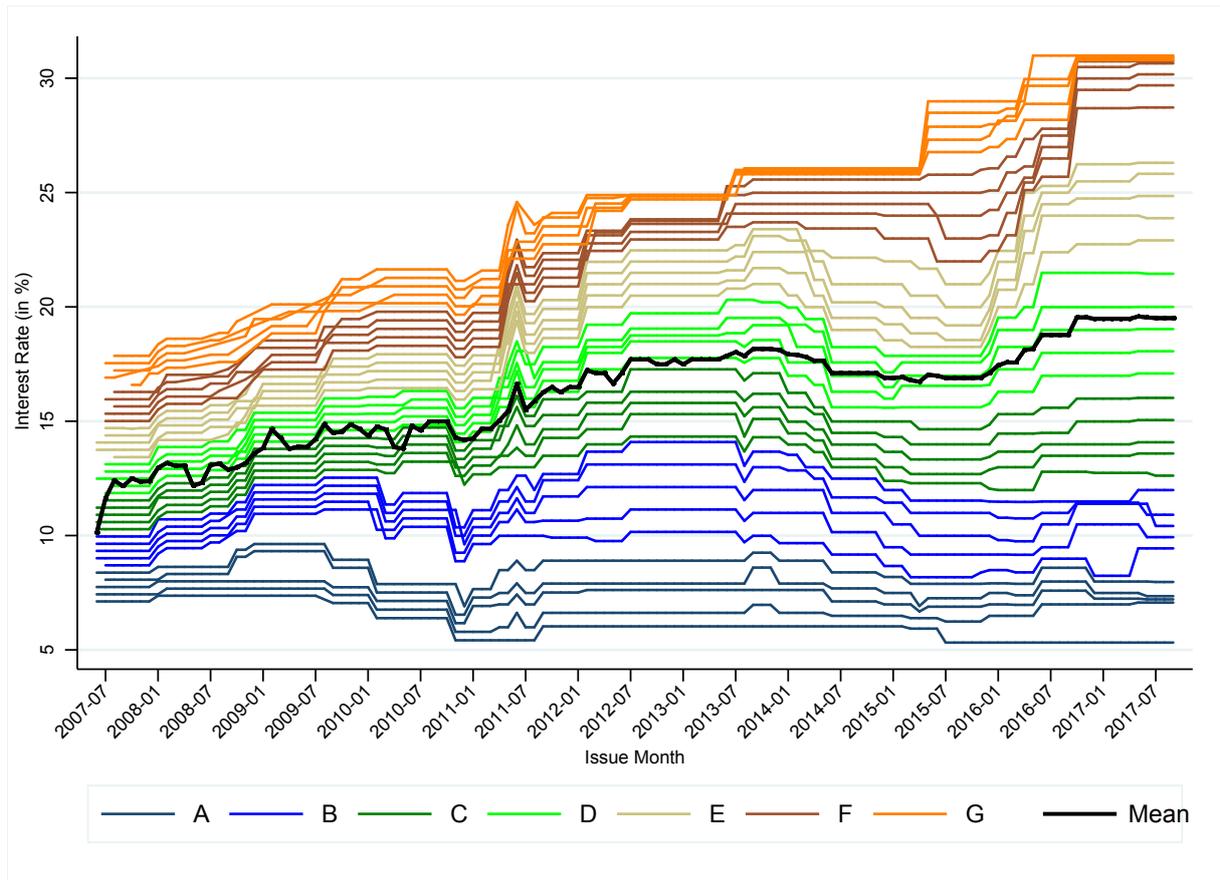
Figure 3: Evolution of Interest Rates on Lending Club by Sub-Grades

Note: This figure plots the evolution of interest rates for the different risk buckets (sub-grades) for Lending Club.
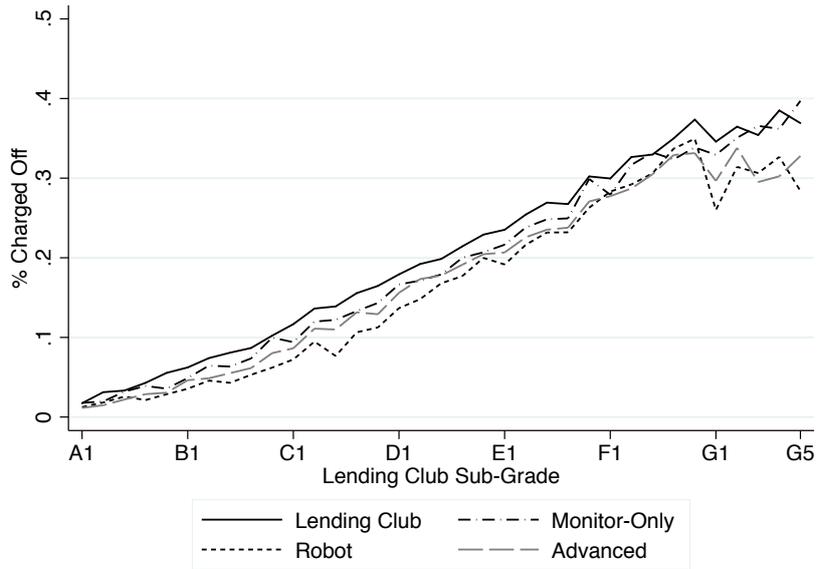
Figure 4: Charged Off Loans

Note: This figure displays the share of fractional loans issued on Lending Club in the 2014 to 2016 that are in default or charged-off as of December 2017. These shares are plotted over Lending Club sub-grades, which map into a given interest rate at a given time. This share is calculated for the whole Lending Club platform, as well as for the restricted samples of loans which have at least two *robot* accounts invested in this loan, at least two *advanced* accounts invested in this loan, and at least two *monitored-only* investors invested in the loan. A loan enters default status when it is 121+ days past due and enters charge-off status after 150 days.
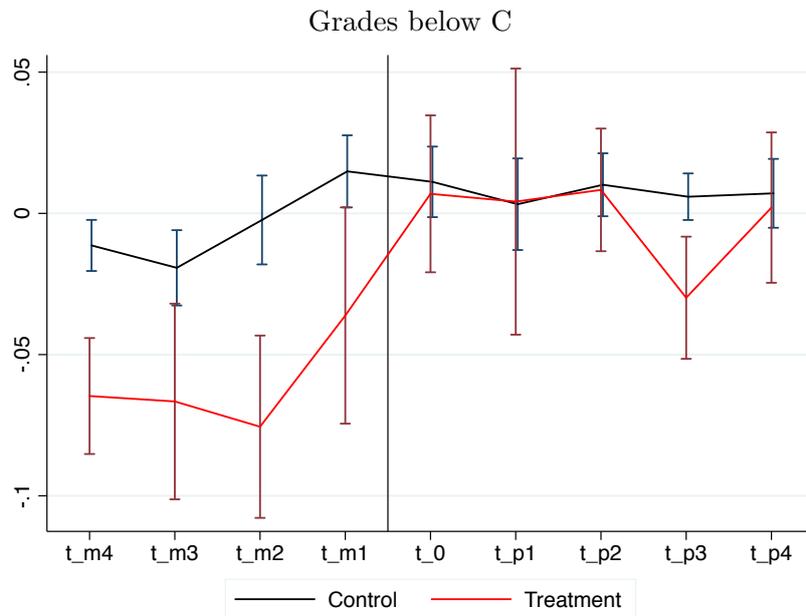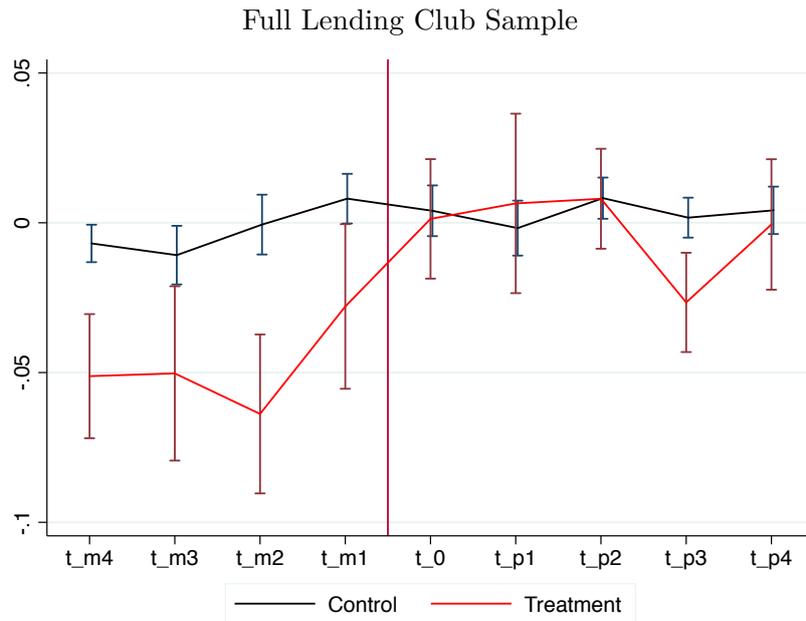
Figure 5: Change to Investor Screening Cost: Difference-in-differences analysis

Note: This figure plots regression coefficients from a difference-in-differences analysis. The left hand-side variable of the regression is an indicator variable equal to one if the loan is charged-off as of December 2017. The explanatory variables are month fixed effects, and the regression includes interest rate fixed effects. The reference period is month 5 and 6 after the change in information set. Each line results from a distinct regression, where the sample is all fractional Lending Club loans for the blue line (control), and all Lending Club where are at least two *robot* accounts are invested in (treatment). Because the constant of each regression is absorbed by the interest rate fixed effects, each line only speaks about the evolution, and not the absolute level of charged-off in each loan population. Segments represents confidence intervals at 10%, where standard errors are clustered at the interest rate level. The bottom panel restricts the sample to loans with grade C or below.
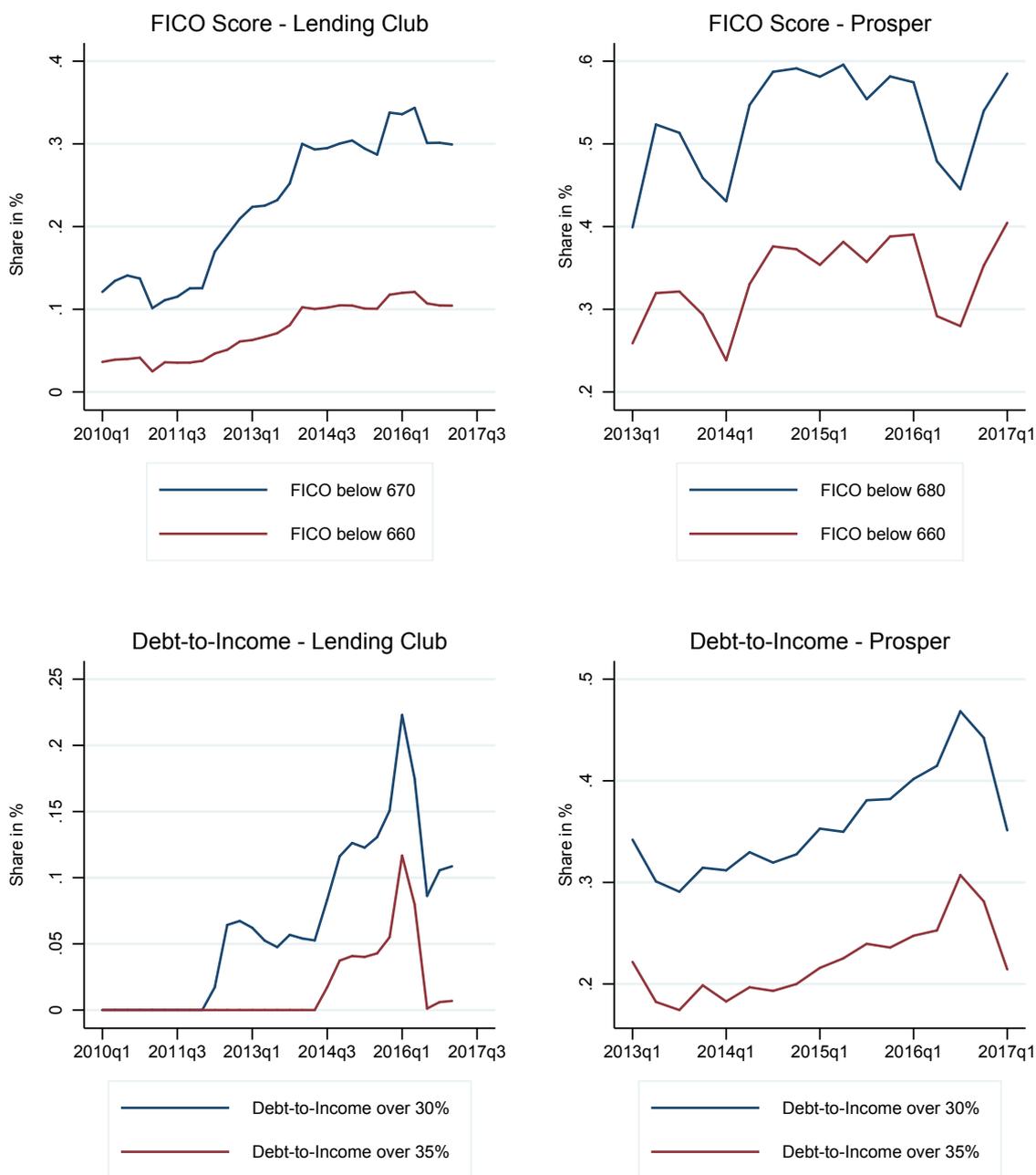
Figure 6: Extensive Margin of Platform Pre-Screening: Evolution of the Share of Borrowers above FICO and Debt-to-Income Thresholds

Note: This figure plots the evolution of the share of borrowers on the Lending Club platform whose FICO score is below 670 (680) and 660, and whose debt-to-income ration is above 30% and 35%, for both Lending Club and Prosper.
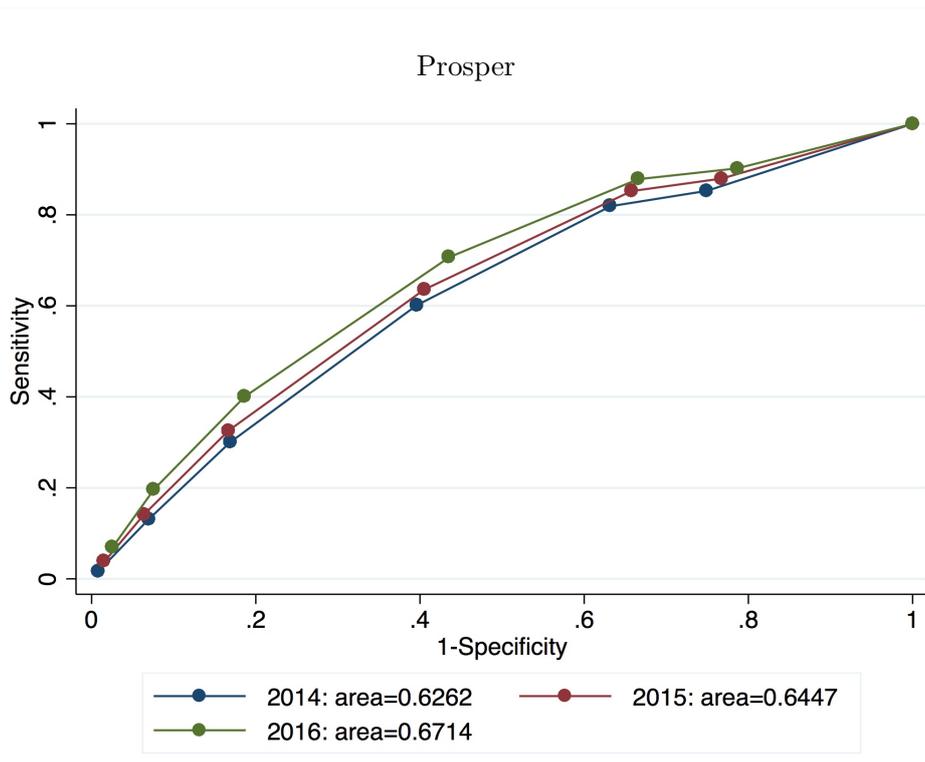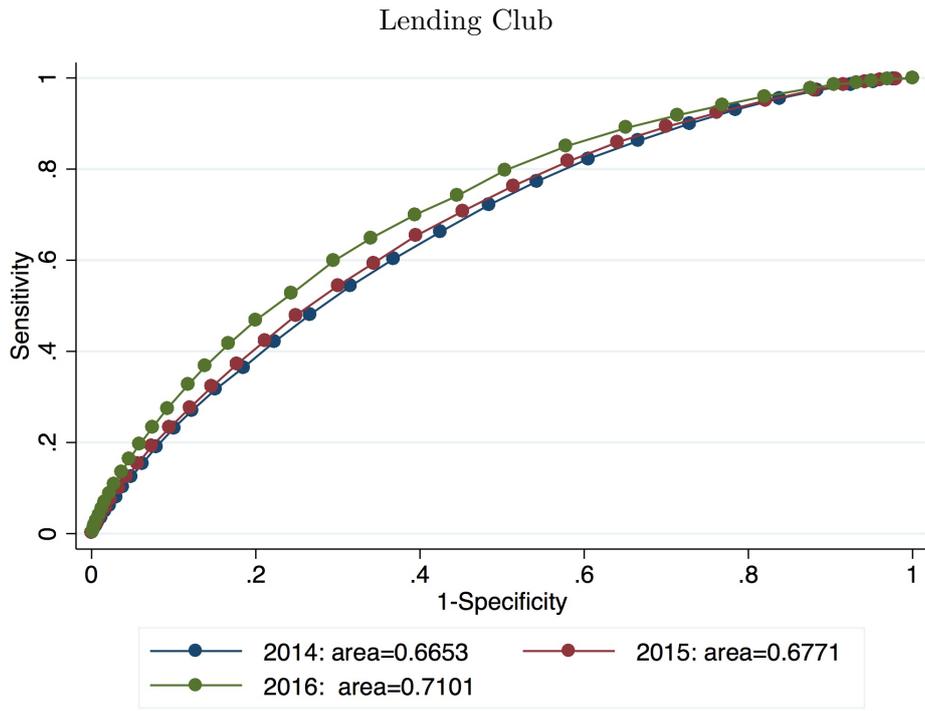
Figure 7: Intensive Margin of Platform Pre-Screening: ROC Curve of Lending Club Sub-grades and Prosper Grades on Charged-Off

Note: This figure plots the ROC curve - which graphs the true positive rate against the false positive rate - obtained when using Lending Club sub-grades and Prosper Grades as a predictor of charged-off. The larger the area below a ROC curve, the more precise the predictor is. The test is computed separately for loans issued and graded by each platform in 2014, 2015 and 2016.
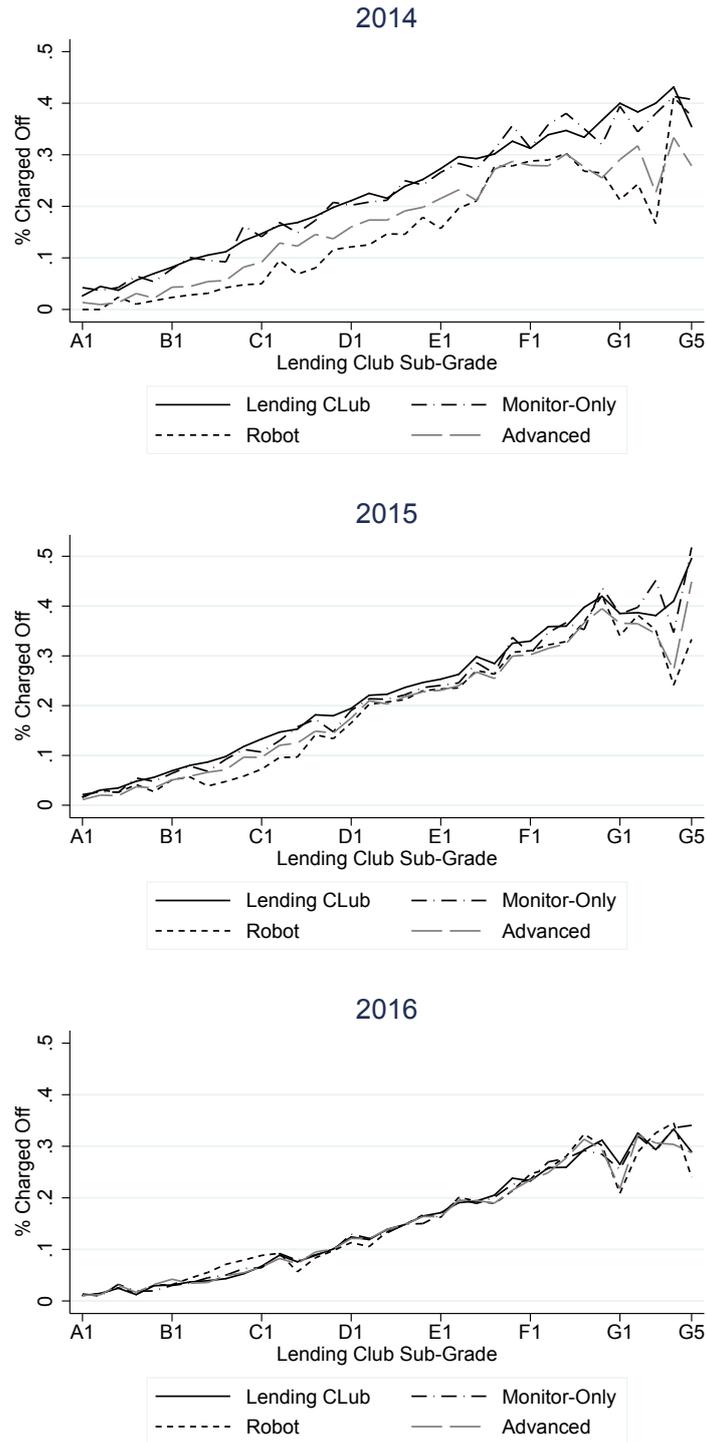
Figure 8: Charged Off Loans

Note: These figures display the share of fractional loans issued on Lending Club in 2014, 2015, and 2016 that are in default or charged-off as of December 2017. These shares are plotted over Lending Club sub-grades, which map into a given interest rate at a given time. This share is calculated for the whole Lending Club platform, as well as for the restricted samples of loans which have at least two *robot* accounts invested in this loan, at least two *advanced* accounts invested in this loan, and at least two *monitored-only* investors invested in the loan. A loan enters default status when it is 121+ days past due and enters charge-off status after 150 days.

Table 1: Summary Statistics

	Number (1)	Total Amount Invested (2)	Median Amount Invested (3)	Mean Amount Invested (4)	Max Amount Invested (5)	Avg. Int. Rate (6)	Platform Avg. Int. Rate (7)	Risk Tolerance (8)
Lending Club							15.76%	
Total	7,368	138,633,952	3,050	18,815.7	3,712,900	18.98%		-
Regular	4,435	56,692,279	1,600	12,783.6	2,102,925	19.34%		7.96%
Advanced	2,933	81,703,628	5,925	27,936.8	3,712,900	18.83%		-
Monitor-Only	636	13,309,525	4,650	20,926.9	722,750	19.20%		-
Prosper							16.32%	
Total	1,616	21,039,794	2,425	13,019.7	658,639	19.84%		-
Regular	1,095	13,421,524	1,900	12,257.1	630,937	19.86%		8.01%
Advanced	521	7,618,145	3525	14,622.4	658,639	19.80%		-
Monitor-Only	126	1,699,350	1,925	13,486.9	155,575	16.54%		-

Note: This table provides summary statistics on the proprietary dataset used for the empirical analysis. The data covers all transactions executed by LendingRobot users between January 2014 and February 2017, which represents more than $120 million invested on the two major lending platforms, LendingClub and Prosper, as well as all historic transactions from portfolios monitored by the company. *Robot* accounts are invested automatically based on LendingRobot credit model and investors risk tolerance. Trades are executed through the lending platforms' API. *Advanced* accounts combine LendingRobot screening tool with tailored screening criteria and are executed through the platform APIs. *Monitor-only* accounts are managed by the investor themselves, and LendingRobot only tracks the portfolio performance, without performing any trades. Risk tolerance corresponds to the target return the investor select, between between 6.3% and 8.5%, and and associated maximum losses between 7% and 12%.

	Lending Club			Prosper		
	Robot (1)	Advanced (2)	Monitored (3)	Robot (4)	Advanced (5)	Monitored (6)
Loan amount	0.005*** (18.89)	0.008*** (27.97)	0.015*** (36.16)	0.015*** (25.83)	0.012*** (19.00)	0.018*** (21.22)
FICO Score	0.000 (1.44)	0.001*** (11.62)	-0.001*** (-10.56)	-0.000 (-0.01)	0.000* (1.76)	-0.000*** (-2.95)
Annual Income	0.001*** (7.18)	0.001*** (13.42)	0.000*** (9.83)	-0.000*** (-2.90)	0.001*** (5.67)	-0.000* (-1.78)
Employment Length	0.002*** (8.96)	0.007*** (19.42)	0.001*** (5.47)	0.000 (1.43)	0.002*** (6.27)	0.002*** (4.01)
Debt to Income	-0.001*** (-4.67)	-0.002*** (-10.36)	0.001*** (8.71)	0.041 (1.37)	-0.108* (-1.74)	0.137*** (3.95)
Own Home Ownership	0.033*** (8.96)	0.054*** (14.33)	0.006** (2.53)	-0.017** (-2.71)	0.024** (2.53)	0.006 (1.27)
Open Accounts	0.002*** (7.04)	0.001*** (5.73)	0.000 (0.89)	0.001*** (3.30)	0.002** (2.50)	0.000 (0.03)
First Credit Line	-0.000 (-1.56)	-0.001** (-2.50)	-0.001*** (-9.17)	-0.000 (-0.62)	-0.001*** (-5.19)	-0.001** (-2.50)
Delinquency	-0.005*** (-6.70)	-0.019*** (-18.68)	-0.006*** (-6.73)	-0.000 (-0.24)	-0.002*** (-4.34)	-0.001*** (-3.78)
Term	-0.012 (-1.59)	-0.066*** (-7.65)	0.045*** (6.89)	-0.000 (-0.42)	-0.004*** (-5.31)	0.004*** (8.40)
Inquiries, last 6 months	-0.038*** (-14.47)	-0.068*** (-28.10)	-0.003** (-2.00)	-0.008*** (-3.59)	-0.045*** (-11.45)	-0.001 (-0.45)
Revolving Utilization Rate	0.000*** (3.39)	0.000*** (5.36)	0.000*** (2.93)			
Derogatory public records	-0.042*** (-5.07)	-0.044*** (-6.65)	0.003* (1.92)			
Mortgage	0.045*** (9.20)	0.075*** (15.44)	0.017*** (8.09)			
Loan Purpose						
Car	0.051*** (6.33)	0.021* (1.89)	0.007 (1.08)			
Credit Card	0.063*** (8.37)	0.077*** (13.71)	-0.010*** (-4.53)			
Home Improvement	-0.005 (-1.52)	-0.027*** (-6.34)	-0.004 (-1.11)			
Medical	-0.103*** (-7.23)	-0.189*** (-16.90)	0.006 (0.95)			
Interest Rate FE	Yes	Yes	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes	Yes	Yes
Cluster	Interest Rate	Interest Rate	Interest Rate	Interest Rate	Interest Rate	Interest Rate
Observations	365,685	365,685	365,685	38,047	38,047	38,047
Pseudo R^2	0.284	0.222	0.222	0.115	0.173	0.215

Note: This table displays coefficients from OLS regressions where the dependent variable is an indicator variable equal to one if at least two *robot* accounts invested in this loan (column 1 and 4), at least two *advanced* accounts invested in this loan (column 2 and 5), and at least two *monitored-only* investor invested in the loan (column 3 and 6). *Robot* accounts are invested automatically based on LendingRobot credit model and investors risk tolerance. Trades are executed through the lending platforms' API. *Advanced* accounts combine LendingRobot screening tool with tailored screening criteria and are executed through the platform APIs. *Monitor-only* accounts are managed by the investor themselves, and LendingRobot only tracks the portfolio performance, without performing any trades. The sample includes all Lending Club fractional loans issued between 2014 and 2016 for columns 1 to 3, and all Prosper fractional loans issued between 2014 and 2016 for columns 4 to 6. Standard errors of the coefficients are clustered by platform interest rate, and t-statistics are reported below the coefficients. *, **, and *** represent statistical significance at the 10%, 5%, and 1% confidence levels, respectively.

Table 3: Screening Performance

	Prob(Charged-Off)								
Account Type	Robot (1)	Advanced (2)	Monitor (3)	Robot (4)	Advanced (5)	Monitor (6)	Robot (7)	Advanced (8)	Monitor (9)
Account Type	-0.031*** (-10.84)	-0.044*** (-18.04)	-0.008*** (-4.68)	-0.084*** (-20.56)	-0.070*** (-19.86)	-0.005 (-1.27)	0.012* (1.66)	-0.015*** (-3.64)	0.007** (2.21)
Account Type x 2015				0.051*** (10.38)	0.029*** (7.11)	-0.006 (-1.27)			
Account Type x 2016				0.075*** (13.66)	0.050*** (12.42)	-0.002 (-0.45)			
Account Type x Grade B							-0.041*** (-3.72)	-0.019*** (-3.36)	-0.009** (-2.11)
Account Type x Grade C							-0.058*** (-6.36)	-0.030*** (-5.28)	-0.015*** (-3.07)
Account Type x Grade D							-0.052*** (-5.97)	-0.037*** (-6.06)	-0.027*** (-4.58)
Account Type x Grade E							-0.049*** (-4.62)	-0.047*** (-4.58)	-0.019** (-2.22)
Account Type x Grade F							-0.026** (-2.43)	-0.039*** (-3.19)	-0.005 (-0.48)
Account Type x Grade G							-0.089*** (-4.31)	-0.081*** (-3.66)	-0.006 (-0.31)
Month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Interest Rate FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Cluster	Int. Rate	Int. Rate	Int. Rate	Int. Rate	Int. Rate	Int. Rate	Int. Rate	Int. Rate	Int. Rate
Observations	365,691	365,691	365,691	365,691	365,691	365,691	365,691	365,691	365,691
Pseudo R^2	0.062	0.064	0.061	0.062	0.065	0.061	0.062	0.064	0.061

Note: This table contains the OLS regression coefficients for fractional loans originated on the Lending Club platform for the period 2014-2016. The dependent variable is an indicator variable for the loan being charged off or in default status as of December 2017. Explanatory variables are indicator variables equal to one if at least two *robot*, *advanced*, and *monitor-only* accounts are invested in this loan. *Robot* accounts are invested automatically based on LendingRobot credit model and investors risk tolerance. Trades are executed through the lending platforms' API. *Advanced* accounts combine LendingRobot screening tool with tailored screening criteria and are executed through the platform APIs. *Monitor-only* accounts are managed by the investor themselves, and LendingRobot only tracks the portfolio performance, without performing any trades. Grades A to G are as per Lending Club typology. Standard errors of the coefficients are clustered by interest rate, and t-statistics are reported below the coefficients. *, **, and *** represent statistical significance at the 10%, 5%, and 1% confidence levels, respectively.

Table 4: Difference in Difference Analysis

	-3/+3 months Window (1)	Grade below C (2)	-2/+2 months Window (3)	Control Group: Monitor (4)
Robot Account	-0.072***	-0.076***	-0.074***	-0.098***
	(-7.00)	(-5.34)	(-6.98)	(-10.85)
Robot Account x Post	0.040***	0.049***	0.037**	0.043***
	(3.20)	(3.01)	(2.68)	(3.65)
Advanced Account	-0.057***	-0.064***	-0.053***	
	(-8.03)	(-6.20)	(-6.14)	
Advanced Account x Post	0.013*	0.008	0.015	
	(1.73)	(0.71)	(1.42)	
Monitor-Only Account	0.013*	0.020**	0.001	
	(1.88)	(2.15)	(0.16)	
Monitor-Only Account x Post	-0.001	-0.002	0.016	
	(-0.09)	(-0.19)	(1.71)	
Month FE	Yes	Yes	Yes	Yes
Interest Rate FE	Yes	Yes	Yes	Yes
Cluster	Int. Rate	Int. Rate	Int. Rate	Int. Rate
Observations	65,859	35,880	37,615	11,283
Pseudo R^2	0.059	0.030	0.060	0.071

Note: This table displays the regression coefficients from linear probability regressions. The dependent variable is an indicator variable for a given loan to be charged-off as of December 2017. $\mathbb{1}_{robot}$ is an indicator variable equal to one if at least two robot accounts are invested in the loan, $\mathbb{1}_{advance}$ is an indicator variable equal to one if at least two advance accounts are invested in the loan, $\mathbb{1}_{monitor}$ is an indicator variable equal to one if at least two monitor-only account are invested in the loan, *Post* is an indicator variable for being in the period after the shock to the information set. All regressions include interest rate fixed effects and month fixed effects. Standard errors of the coefficients are clustered by interest rate, and t-statistics are reported below the coefficients. *, **, and *** represent statistical significance at the 10%, 5%, and 1% confidence levels, respectively.

Appendix

A   Equilibrium Derivation

This appendix gives the mathematical derivation of the sub-game and then full equilibria of the model.

A.1   Proof of Proposition 1

This proof repeatedly uses the two participation conditions (3.4) and (3.6). To recap, condition (3.4) specifies that uninformed investors invest on the platform if and only if the expected value of a listed loan, after potential screening by informed sophisticated investors, is higher than the investment requirement. Condition (3.6) specifies that sophisticated investors become informed and screen list loans if and only if their expected profit from screening is higher than the information cost.

By construction, the determination of the four types of sub-game equilibria is governed by whether neither, at least one, or both of the two participations are satisfied. Denote type-j sub-game equilibrium profile by SE_j, $j \in \{1, 2, 3, 4\}$, which is a subset of the (π_p, μ)-space denoted by $S = \{(\pi_p, \mu, \cdot) | \pi_0 \leqslant \pi_p \leqslant 1, \mu \geqslant 0\}$.

In a type-1 sub-game equilibrium, none of the investors participates. This implies that $\omega = 0$ and condition (3.5) further implies that $p_i(0) = I$. Thus, conditions (3.4) and (3.6) give that

$$SE_1 = \{(\pi_p, \mu, \cdot) | V'(0) < I, \pi_p(R_H - I) < \mu\} . \tag{A.1}$$

In a type-2 sub-game equilibrium, sophisticated investors become informed but uninformed investors do not participate. This implies that $\omega = \min\{\Omega, x_p\}$ and condition (3.5) further implies that $p_i(\min\{\Omega, x_p\}) = I$. Thus, conditions (3.4) and (3.6) give that

$$SE_2 = \{(\pi_p, \mu, \cdot) | V'(\Omega) < I, \pi_p(R_H - I) \geqslant \mu\} , \tag{A.2}$$

where $V'(\Omega) = V'(\min\{\Omega, x_p\})$ by construction.

In a type-3 sub-game equilibrium, sophisticated investors become informed and uninformed investors also participate. This implies that $\omega = \min\{\Omega, x_p\}$ and condition (3.5) further implies that $p_i(\min\{\Omega, x_p\}) = V'(\min\{\Omega, x_p\})$. Thus, conditions (3.4) and (3.6) give that

$$SE_3 = \{(\pi_p, \mu, \cdot) | V'(\Omega) \geqslant I, \pi_p(R_H - V'(\Omega)) \geqslant \mu\}, \tag{A.3}$$

where $V'(\Omega) = V'(\min\{\Omega, x_p\})$ by construction.

In a type-4 sub-game equilibrium, sophisticated investors do not become informed but all the uninformed investors participate. This implies that $\omega = 0$ and condition (3.5) further implies that $p_i(0) = V'(0)$. Thus, conditions (3.4) and (3.6) give that

$$SE_4 = \{(\pi_p, \mu, \cdot) | V'(0) \geqslant I, \pi_p(R_H - V'(0)) < \mu\}. \tag{A.4}$$

Because $V'(\omega)$ is decreasing in ω, it immediately follows that

$$\cup_{j \in \{1,2,3,4\}} SE = S,$$

implying that one of the four types of sub-game equilibrium must happen.

We note that, as common in games with strategic complementarity, multiple equilibria may happen in the sub-game for a given pair of (π_p, μ). However, as we show later in the proof of Proposition 2, it is not a concern in the full equilibrium analysis and no equilibrium selection mechanism is needed for the sub-game.

Finally, direct calculation of the equilibrium volume in each sub-game equilibrium concludes the proof.

A.2 Proof of Proposition 2

This proof proceeds in two steps. First, we show that either type-1 or type-3 sub-game equilibrium can not sustain a full equilibrium; in other words, the full equilibrium only admits either a type-2 or a type-4 sub-game equilibrium. Second, we show that a type-2 (type-4) sub-game equilibrium happens in a full equilibrium when the platform's information cost κ is larger (smaller)

than a threshold.

Step 1. Suppose $(\pi_p, \mu, \cdot) \in SE_1$, that is, the full equilibrium admits a type-1 sub-game equilibrium as characterized in (A.1). The equilibrium volume is 0.

Consider an alternative equilibrium profile $(\pi_p, \mu', \cdot)$, $\mu' < \mu$ such that $\pi_p(R_H - I) \geqslant \mu'$. Notice that μ' must exist because $0 < \pi_0 \leqslant \pi_p$. Because $V'(0) \geqslant V'(\Omega)$, we have that $(\pi_p, \mu', \cdot) \in SE_2$ as characterized in (A.2). Because changing μ is costless for the platform, this implies that the platform will then find it profitable to deviate to a type-2 sub-game equilibrium by decreasing μ to μ' without changing π_p, enjoying a higher volume $\min\{\pi_0 x_0(I), \pi_p \Omega\} > 0$. This implies that a full equilibrium can only admit a type-2 sub-game equilibrium but not a type-1 sub-game equilibrium.

Similarly, suppose $(\pi_p, \mu, \cdot) \in SE_3$, that is, the full equilibrium admits a type-3 sub-game equilibrium as characterized in (A.3). The equilibrium volume is $\frac{\pi_0 x_0(p(\Omega))}{\pi_p}$.

Consider an alternative equilibrium profile $(\pi_p, \mu', \cdot)$, $\mu' > \mu$ such that $\pi_p(R_H - V'(0)) < \mu'$. Notice that μ' must exist because $0 < \pi_0 \leqslant \pi_p$ and $I \leqslant V'(0) < R_H$. Again because $V'(0) \geqslant V'(\Omega)$, we have that $(\pi_p, \mu', \cdot) \in SE_4$ as characterized in (A.4). Because changing μ is costless for the platform and $x_0(p(0)) > x_0(p(\Omega))$, this implies that the platform will then find it profitable to deviate to a type-4 sub-game equilibrium by increasing μ to μ' without changing π_p, enjoying a higher volume $\frac{\pi_0 x_0(p(0))}{\pi_p} > \frac{\pi_0 x_0(p(\Omega))}{\pi_p}$. This implies that a full equilibrium can only admit a type-4 sub-game equilibrium but not a type-3 sub-game equilibrium.

Step 2. Consider the hyperplane $V'(0) = I$ that decomposes the $(\pi_p, \mu, \cdot)$-space into two half-spaces:

$$S_L = \{(\pi_p, \mu, \cdot)|V'(0) < I\},$$

and

$$S_H = \{(\pi_p, \mu, \cdot)|V'(0) \geqslant I\}.$$

Notice that the hyperplane $V'(0) = I$ gives a unique platform pre-screening intensity

$$\widehat{\pi_p} = \frac{I - R_L}{R_H - R_L} \in (0, 1),$$

and by construction, $SE_1 \subset S_L$ and $SE_4 \subset S_H$ according to conditions (A.1) and (A.4). Thus, $SE_1 \cap SE_4 = \varnothing$. In particular, conditions (A.1) and (A.4) imply that there must exist a $\widehat{\mu} > 0$ such that

$$(SE_1 \cup SE_4) \cap \{(\pi_p, \mu, \cdot)|\mu > \widehat{\mu}\} = S \cap \{(\pi_p, \mu, \cdot)|\mu > \widehat{\mu}\} \,.$$

Hence, if

$$C(\widehat{\pi_p}) = \frac{1}{2}\kappa((\widehat{\pi_p} - \pi_0)_+)^2 \geqslant \min\{\pi_0 x_0(I), \widehat{\pi_p}\Omega\} \,,$$

by the argument in Step 1, a type-2 sub-game equilibrium is achievable by choosing

$$\pi_p = \frac{\Omega}{\kappa} + \pi_0 \text{ and } \mu = 0 \,,$$

in which case we define

$$\underline{\kappa} = \frac{2\min\{\pi_0 x_0(I), \widehat{\pi_p}\Omega\}}{((\widehat{\pi_p} - \pi_0)_+)^2} > 0 \,.$$

Otherwise if

$$C(\widehat{\pi_p}) = \frac{1}{2}\kappa((\widehat{\pi_p} - \pi_0)_+)^2 \leqslant \frac{\pi_0 x_0(p(\Omega))}{\widehat{\pi_p}} \,,$$

a type-4 sub-game equilibrium is achievable by choosing π_p such that it solves

$$\frac{\partial\left(\frac{\pi_0 x_0(p(\Omega))}{\pi_p} - C(\pi_p)\right)}{\partial \pi_p} = 0 \,,$$

and choosing $\mu > \widehat{\mu}$, in which case we define

$$\overline{\kappa} = \frac{2\pi_0 x_0(p(\Omega))}{\widehat{\pi_p}((\widehat{\pi_p} - \pi_0)_+)^2} > 0 \,.$$

Because $\overline{\kappa} > \underline{\kappa}$, this concludes the proof.

Table IA.1: Investor Screening - 2014-2016 - Robustness

	Lending Club			Prosper		
	Robot (1)	Advanced (2)	Monitored (3)	Robot (4)	Advanced (5)	Monitored (6)
Loan amount	0.011*** (16.82)	0.018*** (26.95)	0.023*** (36.81)	0.051*** (19.85)	0.027*** (13.60)	0.024*** (19.09)
FICO Score	0.000*** (2.71)	0.002*** (12.32)	-0.001*** (-10.21)	0.001*** (3.54)	0.001*** (4.67)	-0.000*** (-3.21)
Annual Income	0.002*** (7.63)	0.003*** (12.20)	0.001*** (9.20)	-0.001*** (-4.38)	0.001*** (10.78)	-0.000 (-1.54)
Employment Length	0.006*** (9.30)	0.015*** (20.99)	0.002*** (7.31)	0.003*** (3.96)	0.005*** (10.80)	0.002*** (3.39)
Debt to Income	-0.001*** (-4.00)	-0.004*** (-14.26)	0.002*** (9.81)	-0.394*** (-3.87)	-0.406*** (-4.95)	0.171*** (3.98)
Own Home Ownership	0.086*** (8.98)	0.137*** (14.31)	0.015*** (4.03)	-0.228*** (-4.71)	-0.017 (-0.63)	0.012* (1.78)
Open Accounts	0.005*** (7.51)	0.004*** (6.99)	0.000 (1.06)	0.010*** (7.42)	0.008*** (6.18)	-0.001 (-1.07)
First Credit Line	-0.002** (-2.51)	-0.003*** (-4.63)	-0.002*** (-9.86)	-0.000 (-0.65)	-0.003*** (-7.11)	-0.001* (-1.87)
Delinquency	-0.012*** (-7.13)	-0.043*** (-22.24)	-0.011*** (-10.14)	0.001 (1.63)	-0.004*** (-8.97)	-0.002*** (-4.25)
Term	-0.054*** (-2.94)	-0.170*** (-11.68)	0.057*** (6.13)	-0.006*** (-4.11)	-0.012*** (-16.17)	0.005*** (9.88)
Inquiries, last 6 months	-0.096*** (-14.95)	-0.162*** (-22.42)	-0.005** (-2.02)	-0.067*** (-7.70)	-0.133*** (-18.18)	-0.003 (-0.76)
Revolving Utilization Rate	0.001*** (3.09)	0.001*** (5.82)	0.000*** (4.26)			
Derogatory public records	-0.113*** (-5.07)	-0.117*** (-6.48)	0.007*** (2.93)			
Mortgage	0.118*** (9.74)	0.193*** (14.01)	0.028*** (9.90)			
Car	0.130*** (5.56)	0.073*** (3.43)	0.015 (1.64)			
Credit Card	0.171*** (8.42)	0.199*** (11.15)	-0.019*** (-6.23)			
Home Improvement	-0.010 (-1.38)	-0.078*** (-10.49)	-0.010** (-2.14)			
Medical	-0.287*** (-7.19)	-0.407*** (-14.88)	0.007 (0.94)			
Interest Rate FE	Yes	Yes	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes	Yes	Yes
Cluster	Int.Rate	Int.Rate	Int.Rate	Int.Rate	Int.Rate	Int.Rate
Observations	365,685	365,685	365,685	38,047	38,047	38,047
Pseudo R^2	0.355	0.362	0.290	0.282	0.386	0.280

Note: This table displays coefficients from OLS regressions where the dependent variable is the log of 1 + the number of *robot* accounts invested in this loan (column 1 and 4), *advanced* accounts invested in this loan (column 2 and 5), and *monitored-only* investor invested in the loan (column 3 and 6). *Robot* accounts are invested automatically based on LendingRobot credit model and investors risk tolerance. Trades are executed through the lending platforms' API. *Advanced* accounts combine LendingRobot screening tool with tailored screening criteria and are executed through the platform APIs. *Monitor-only* accounts are managed by the investor themselves, and LendingRobot only tracks the portfolio performance, without performing any trades. The sample includes all Lending Club fractional loans issued between 2014 and 2016 for columns 1 to 3, and all Prosper fractional loans issued between 2014 and 2016 for columns 4 to 6. Standard errors of the coefficients are clustered by platform interest rate, and t-statistics are reported below the coefficients. *, **, and *** represent statistical significance at the 10%, 5%, and 1% confidence levels, respectively.

Table IA.2: Borrower Characteristics Before and After the Change in Information Set

	Two Months Before	Two Months After
Loan Amount (in USDk)	14.85	15.29
FICO Score	696	697
Annual Income (in USDk)	72.94	74.38
Employment Length	5.6	5.5
Debt to Income	18.7	18.6
Own Home Ownership	11.5%	11.8%
Open Accounts	11.8	11.6
First Credit Line (in years)	18.0	17.9
Delinquency	0.345	0.333
Observations	58,566	56,335

Note: This table displays the averages for key borrowers' characteristics for the two-month period before and after the change in the information set provided to investors by Lending Club.
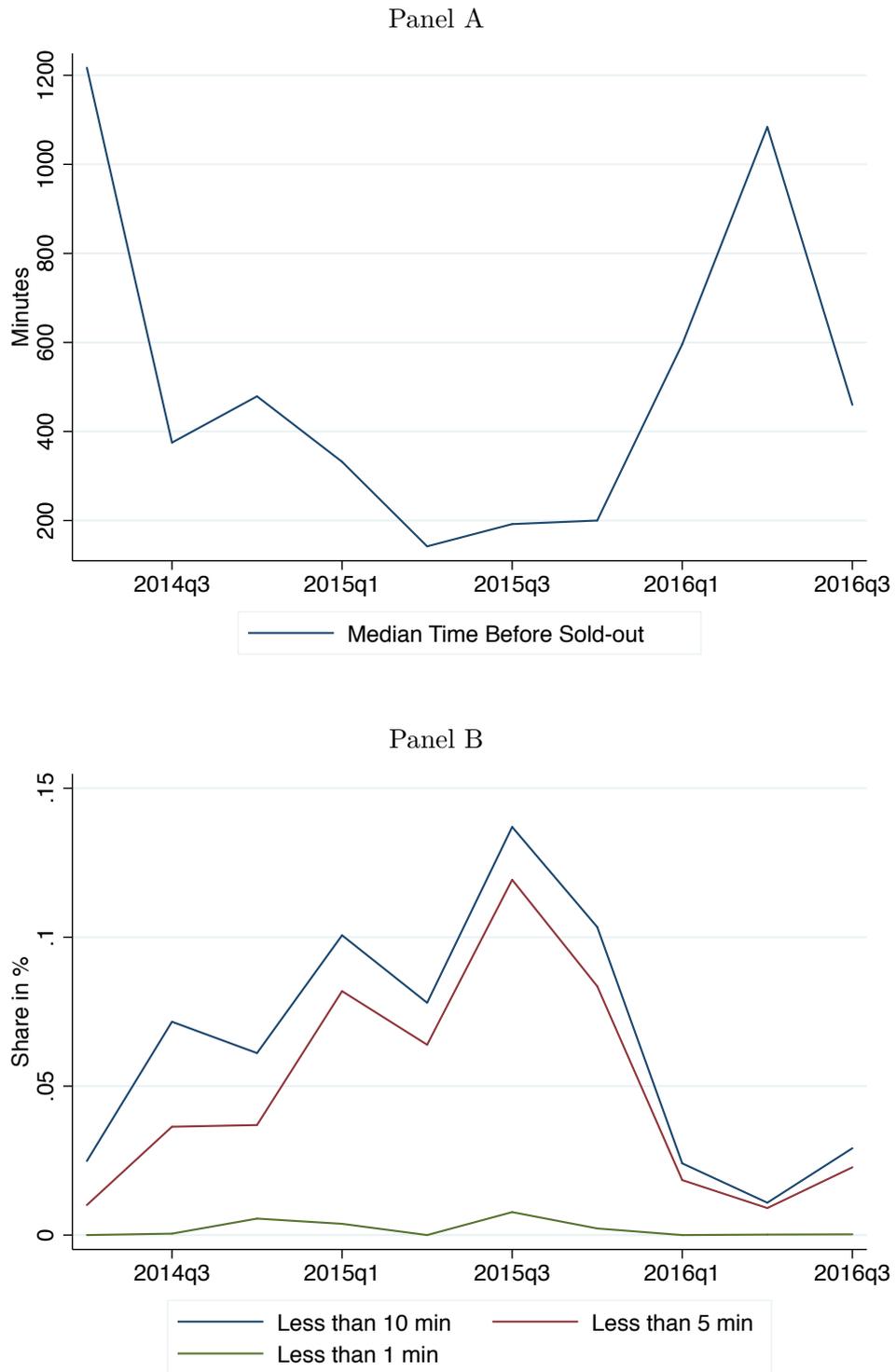
Panel A

Panel B

Figure IA.1: Time for Loans to Sell Out

Note: Panel A plots the median time for loans to sell out. Panel B plots the share of loans that sell out faster than a 10 minutes, 5 minutes and 1 minute threshold. The sample only covers Lending Club loans in which Lending Robot assisted investors participated in.