Marketplace Lending: A New Banking Paradigm?

Boris Vallee  
Harvard Business School

Yao Zeng  
University of Washington

April, 2018

Abstract

Marketplace lending relies on large-scale loan screening and information production by investors, a major deviation from the traditional banking paradigm. Theoretically, the participation of sophisticated investors in marketplace lending increases volumes and improves screening outcomes, but also creates adverse selection to less sophisticated investors. In maximizing loan volume, the platform trades off these two forces. It may decrease information provision to investors and choose an intermediate level of platform pre-screening. We use novel investor-level data to test these predictions. We empirically show that more sophisticated investors 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, Fintech

JEL: G21, G23, D82

*We are grateful to Itay Goldstein, Wei Jiang, Andrew Karolyi (the Editors), Shimon Kogan (discussant), Chiara Farronato, Simon Gervais, David Scharfstein, Andrei Shleifer, Jeremy Stein, Xiaoyan Zhang, and three anonymous referees for helpful comments. We also thank seminar and conference participants at the RFS FinTech Workshops at Columbia University and Cornell Tech, NYU Stern FinTech Conference, Bentley University, Chinese University of Hong Kong Shenzhen, London School of Economics, Peking University, Tsinghua PBC School of Finance, the University of Oregon, and the University of Washington. The authors have received in-kind support from LendingRobot for this project in the form of a proprietary dataset of LendingRobot transactions. This data was provided without any requirements. We thank Yating Han, Anna Kruglova, Claire Lin, Kelly Lu, Lew Thorson, and in particular Botir Kobilov for excellent research assistance. All remaining errors are ours only.
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.\(^1\) 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 on the investor side. Although lending platforms pre-screen loan applications modestly and allocate them into a risk bucket, they provide additional information about loan applications to investors, which allows investors to further screen borrowers and actively invest in individual loans. This model significantly differs from the traditional banking paradigm where depositors buy a safe claim and are essentially isolated from the borrowers. Moreover, investor composition on lending platforms has been evolving significantly due to informationally sophisticated investors’ increasing participation. These diverse investors perform large-scale borrower screening, and the resulting joint information production between the platforms and investors challenges the traditional role of banks being the exclusive information producer on behalf of investors (Gorton and Pennacchi, 1990).

The joint information production between the platforms and investors of different levels of sophistication poses a series of research questions, which we address in this paper. First, are more sophisticated investors on lending platforms screening borrowers more intensively and thereby consistently outperforming less sophisticated investors? If so, then, how does sophisticated investors’ outperformance relate to the platform’s pre-screening and information provision? Finally, given the heterogeneity of investors, what is the optimal platform design in terms of loan

---

\(^1\)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. 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.
pre-screening and information provision to investors to maximize volumes? Answering these questions is essential for understanding how platform and investor information production interact with each other, which further speaks to the promises and potential pitfalls of this new banking business model.

Our study is also motivated by a puzzling event during the development of the marketplace lending industry. On November the 7th, 2014, Lending Club, the largest lending platform, removed 50 out of the more than 100 variables on borrower characteristics that they previously provided to investors. This change was unanticipated and surprised many market participants as it was the only investor-unfriendly move in Lending Club’s history.\(^3\) The whole marketplace lending model relies on transparency to allow active screening from investors, and to act as a substitute to a skin in the game from the intermediating entity. What is the economic rationale behind this reduction in the information set provided to investors?

In addressing the questions above, we develop a theoretical framework for marketplace lending and test 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 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. In contrast, less sophisticated investors do not screen; they invest in a loan passively if they can break even on average, or they do not invest at all. This informational advantage results in sophisticated investors’ outperformance relative to less sophisticated investors. 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 are reluctant to invest because the average pool of pre-screened borrowers is on average of poor quality. The platform can also learn from sophisticated investor screening, which in turns lowers its pre-screening costs.

However, the heterogeneity in investor sophistication creates an endogenous adverse selection problem among investors, which can hurt volume through two channels. Because sophisticated investors can identify and finance good loans, the easier it is for sophisticated investors to

\(^3\)We provide more discussion relevant institutional details and this specific event in Section 2.2.
acquire information, the lower the quality of an average loan facing a less sophisticated investor. Thus, when adversely selected, less sophisticated investors require a lower loan price to break even, resulting in a lower prevailing loan price on the platform. This lower price in turn lowers the amount of loan applications on the platform. 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.

Hence, to maximize volume, the platform optimally trades off these positive and negative effects of sophisticated investor participation. When platform pre-screening cost is initially high, the platform optimally chooses a low pre-screening intensity but distributes more information to investors. This investment environment encourages the participation of sophisticated investors, boosting volume 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, because doing that would screen out loans that could have been financed by less sophisticated investors (provided that these investors break even on average), and thus would reduce 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 characteristics and their loan portfolios are generally unavailable in the public domain. Fortunately, we obtain a rich dataset that includes portfolio composition for a large set of retail investors on two lending platforms, Lending Club and Prosper. 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

\[4\]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.
technologies provided by LendingRobot, an algorithmic third-party.

Our empirical analysis progresses in several steps. First, we show that more sophisticated investors rely on different loan characteristics to screen the loans they finance, which points to their informational advantage. Being selected by sophisticated investors predicts a significantly lower probability of default for a given loan, meaning that sophisticated investors consistently outperform unsophisticated investors over time and across all risk buckets. We find that loans selected by sophisticated investors have a default rate on average 3% 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.

Using the 2014 Lending Club episode, 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 at the time of the shock to the information set.\(^5\) We rationalize this unanticipated event, 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.

Finally, we find that platforms’ pre-screening intensity has been improving on the intensive margin: platform risk buckets are increasingly precise at predicting default. This increased precision also mitigates adverse selection by reducing the heterogeneity of loans within a given risk bucket.

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 retail investors do; while other institutional investors,\(^5\) This drop in performance could not be immediately observed by investors at the time of the change, as it takes time for loan performance to be revealed.
such as hedge funds, develop highly sophisticated investment strategies that are comparable to what LendingRobot offers to investors.\footnote{LendingRobot is actually partnering with some undisclosed hedge funds to execute their orders.}

To keep our paper focused, we leave a number of questions for future research. These include, for example, the welfare implications of marketplace lending, and whether marketplace lending poses any financial stability concerns due to investors’ potential safe asset creations.

**Related Literature.** Our paper mainly contributes to the burgeoning literature on marketplace lending by directly examining the sharing of information production between platforms and investors, which we believe is what makes marketplace lending special as a new banking model. So far, the literature of marketplace lending 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 and screening of sophisticated investors interact with 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 itself. On optimal platform design, our work also complements a few recent papers that study the motivation behind one lending platform’s switch from an auction-based pricing mechanism to platform direct pricing in its early stage (Franks, Serrano-Velarde and Sussman, 2016), and its effect on borrowers learning about their cost of credit (Liskovich and Shaton, 2017).

Broadly, our paper is related to the growing literature on shadow banking. In a recent review, Adrian and Ashcraft (2016) define shadows banks as non-bank financial institutions that conduct credit, maturity, and liquidity transformation. In the residential lending context, Buchak, Matvos, Piskorski and Seru (2017) find that the share of shadow banks in the mortgage market has tripled from 2007-2015, and Fintech firms accounted for almost a third of shadow bank loan originations by 2015. The two main differences between marketplace lending and other forms of shadow banking is that that in shadow banking investors do not typically produce information, and that marketplace lending in its original design is not involved in maturity
or liquidity transformation. Keys, Mukherjee, Seru and Vig (2010) find that securitization significantly leads to loosened screening despite the skin in the game, while we find that in the marketplace lending context it may not necessarily be the case due to the sharing of screening and information production between the platforms and investors.

In the context of marketplace lending competing with traditional banking, Balyuk (2017) examines whether marketplace lending eases borrowers’ cost of credit from traditional banks. Fuster, Plosser, Schnabl and Vickery (2018) find that FinTech lenders process mortgage applications faster than traditional banks. de Roure, Pelizzon and Thakor (2018) and Tang (2018) directly examine competition between traditional banks and marketplace lending by asking whether they complement or substitute each other in terms of providing credit to different borrower segments. Our work complement them by focusing on the role of investors in affecting borrower screening outcomes.

Our paper is also related to the theoretical literature on information acquisition by banks and investors (Hauswald and Marquez, 2003, 2006), in particular studies on the resulting endogenous adverse selection (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 a related trade-off: investor information acquisition helps guide efficient production but also introduces adverse selection that may 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.\textsuperscript{7}

2.1 Platform Information Collection and Pre-Screening

\textbf{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.\textsuperscript{8,9} 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.

\textbf{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 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.

Platform pre-screening is costly. Notably, platform pre-screening is not equivalent to simply picking a FICO score threshold and listing loans above that threshold. Rather, these screening

\textsuperscript{7}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 \textit{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.

\textsuperscript{8}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.

\textsuperscript{9}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.
decisions rely on the platform’s screening model, which may potentially use many more variables. 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.\(^{10}\) 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 (Diamond and Dybvig, 1983) without having any knowledge of the underlying loans (Gorton and Pennacchi, 1990).\(^{11,12}\)

**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 typically narrower than the information the platform collects. Platforms distribute these information both on their websites and through their Application Programming Interfaces (API), which are customized programming protocols that allow more sophisticated investors to develop their own automated algorithms to place orders.

---

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

11 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 tranch 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.

12 Lending Club also experimented with securitization, but it remains marginal in its funding model.
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. 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 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. 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. 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. 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. These interest rates have evolved over time, as Figure 2 illustrates.

---

13 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 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.
16 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.
17 See discussion on this point in Lending Club Investor Day 2017 Presentation, which is available at ir.lendingclub.com/Cache/1001230258.pdf.
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%).

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 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

\[18\] Managed accounts are passive investment vehicles that are distributed by conventional third-party marketers to high net worth individuals.
pass buying orders.\textsuperscript{19}

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

\textbf{Environment.} There are four types of risk-neutral agents: 1) a 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 $\Omega$ of sophisticated investors, and 4) a competitive fringe of unsophisticated investors.\textsuperscript{20} There are three dates, $t = 0, 1, 2$, with no time discount. The lending platform screens loan applications at $t = 0$. The applications that are screened in are listed on the platform at $t = 1$, and investors decide whether to screen and subsequently buy these listed loans. Loan cash flows are realized at $t = 2$. We specify agents’ objectives and strategies below.

\textbf{Loan applications.} Each penniless loan applicant has one project that requires an initial investment $I$. The project, if funded, pays $R = R_H$ with probability $\pi_0$ ("good project") and $R = R_L$ with probability $1 - \pi_0$ ("bad project") at $t = 2$, where $\pi_0 \in (0, 1)$ and $R_L < I < R_H$. As standard in the banking literature that focuses on bank information production, we assume that applicants do not know their type ex-ante, which allows us to focus on adverse selection on the investor side.\textsuperscript{21}

The number of loan applicants $x_0$ is endogenous, and depends on the equilibrium price $p \geq I$.

\textsuperscript{19}See Figure A.1 in the appendix.

\textsuperscript{20}In reality, investor level of informational sophistication is distributed across a wide spectrum. Our model represents a useful benchmark that captures the heterogeneity in informational sophistication.

\textsuperscript{21}Prominent examples are Hauswald and Marquez (2003, 2006), which are among the first to model bank information acquisition explicitly in a bank competition framework. In their models, borrowers can be good or bad, but they do not know their types ex-ante. More recent examples in the theoretical literature include Glode, Green and Lowery (2012) and Fishman and Parker (2015), in both of which a seller's asset can be good or bad, a buyer can acquire information about it, while the seller itself does not know its type ex-ante. Empirically, Liskovich and Shaton (2017) also find that many borrowers in the marketplace lending context do not know their cost of credit ex-ante. For our purpose, we only need to assume that borrowers do not know their quality relative to the other loans of the same risk-bucket.
on the platform. We assume that the supply curve of applications $x_0(p)$ is an increasing function of $p$, meaning that a higher price (i.e., a lower interest rate) attracts more applications. Because borrowers do not know their types ex-ante, it is also natural to assume that $\pi_0$ does not depend on $p$ as a benchmark.\(^{22}\) We provide one possible micro-foundation for this reduced-form supply curve in Appendix B, which follows standard specifications in the literature.

**Platform pre-screening, pricing, and information distribution.** The platform’s strategy is a triple \(\{\pi_p, p, \mu\}\), which we specify in order below.

First, the platform pre-screens the pool of loan applications and lists some of them at \(t = 0\). Specifically, the platform chooses the interim probability $\pi_p$ of a listed loan being good at \(t = 1\) before any further investor screening, where $\pi_p \in [\pi_0, 1]$. A higher $\pi_p$ implies that the average listed loan is more likely to be good and is indicative of a higher platform pre-screening intensity. We assume that the platform can always screen in a good project but may fail to screen out a bad project.\(^{23}\) Accordingly, the number of loan applications being screened in and listed on the platform is

\[
x_p = \frac{\pi_0 x_0}{\pi_p} \leq x_0.
\]

Platform pre-screening is costly. Since platform’s pre-screening is implemented through scalable algorithms, we assume that pre-screening cost increases in $\pi_p$ but not on the number of applications processed. We take a parametric form $C(\pi_p) = \frac{1}{2} \kappa (\pi_p - \pi_0)^2$ where $\kappa \geq 0$ to facilitate comparative statics.

Second, after pre-screening, the platform assigns an interest rate to listed loans, which is modeled by a price $p$. In equilibrium, the platform holds rational expectations and thus $p$ will be determined by the marginal investor’s offer price, which we detail below.\(^{24}\)

Third, the platform distributes information on listed loans to investors, which effectively

\(^{22}\)We are aware of the classic borrower adverse selection issues in the banking context, but whether a lower interest rate attracts relatively better or worse applicants is still an open question in the marketplace lending context. Since it is beyond the scope of this paper, we leave it for future research.

\(^{23}\)This modeling choice is a reduced-form approach to parsimoniously capture the outcome, rather than the detailed process, of platform pre-screening. It makes the model analytically tractable while is still general enough to capture all the possible pre-screening outcomes in reality: $\pi_p = \pi_0$ 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 and only the $\pi_0 x_0$ good projects are screened in.

\(^{24}\)This modeling choice of having both the platform price and the investor price captures 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 among investors for setting the interest rate.
determines sophisticated investors’ information cost $\mu \geq 0$ in further screening the listed loans. Since these characteristics are provided to the platform by the loan applicants, we assume that it is costless for the platform to change $\mu$.

Finally, the platform’s objective is to maximize the expected volume of eventually financed loans by investors on the platform, regardless of their type, minus its pre-screening cost. 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 by investors in equilibrium.

**Sophisticated investors.** Each sophisticated investor may screen and buy at most one listed loan on the platform at $t = 1$. Her strategy is to choose whether to become informed and to offer a (latent) loan price $p_i$. She has two potential advantages over an unsophisticated investor.

First, a sophisticated investor can purchase an information technology at cost $\mu$, which is set by the platform. This technology, capturing a screening algorithm in reality, allows her to become perfectly informed of one listed loan at $t = 1$. We denote by $\omega$ the population of sophisticated investors who become informed, where $0 \leq \omega \leq \Omega$. When becoming informed, a sophisticated investor passes on bad listed loans while offers a price $p_i$ to buy a good loan.

Second, sophisticated investors, when becoming informed, can screen and buy loans before uninformed investors do. This is consistent with the fact that informed investors use faster technology to place orders through the API. The screening outcome obtained by an informed investor is non-verifiable and non-transferrable.

If a sophisticated investor chooses not to pay $\mu$, she remains uninformed and is essentially identical to an unsophisticated investor in equilibrium. Both sophisticated and unsophisticated investors’ objective is to maximize expected profits.

**Unsophisticated investors.** Each unsophisticated may also buy at most one loan on the platform at $t = 1$. An unsophisticated cannot become informed and may buy a loan only after informed investors move. Thus, his strategy is to offer a (latent) loan price $p_u$ given his updated

---

25For tractability purpose we assume that an informed sophisticated investor is able to know the true type of one listed loan. This assumption can be easily motivated by the various capital constraints that the investors face. However, 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.
belief, which determines his financing decision. Especially, \( p_u = 0 \) suggests that unsophisticated investors cannot break even and thus do not participate at all. Because unsophisticated investors are competitive, they are the marginal investors on the platform if they participate. Otherwise, informed sophisticated investors are the marginal investors.

We formally define a sequential equilibrium as follows:

**Definition 1.** Given \( R_H, R_L, I \) and \( \pi_0 \), the sequential equilibrium is defined as a collection of \( \{\pi_p, \mu, p, p_i, p_u, \omega\} \) such that:\(^{26}\)

i) Given \( \pi_p, \mu \) and \( \omega \), the price \( p_u \) gives uninformed investors expected zero profit;

ii) Given \( \pi_p, \mu \) and \( \omega \), the price \( p_i \) maximizes the expected profit of informed investors;

iii) Given \( \pi_p \) and \( \mu \), the population \( \omega \) of sophisticated investors find it optimal to acquire the information technology and become informed;

iv) The platform’s choices of \( \pi_p \) and \( \mu \) maximize the expected volume of financed loans minus its pre-screening cost;

v) The platform’s choice of \( p \) satisfies rational expectation in the sense that it equals to the marginal investor’s offer price:

\[
p = \begin{cases} 
p_i, & \text{if } p_u = 0, \\
p_u, & \text{if } p_u > 0. \end{cases}
\]

vi) Agents use the Bayes’ rule to update their beliefs and follow sequential rationality.

**3.2 Equilibrium Analysis**

**Uninformed investors.** We derive the equilibrium by backward induction. We first consider uninformed investors’ financing decision on the platform, under any given generic \( x_p, \pi_p, \mu \), as well as the population of informed sophisticated investors \( 0 \leq \omega \leq \Omega \).

Consider the pool of listed loans facing uninformed investors. Uninformed investors move after the \( \omega \) informed investors screen \( \omega \) listed loans and potentially finance \( \pi_p \omega \) good loans.

\(^{26}\)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.
Hence, uninformed investors’ posterior belief $\pi'_p$ of a listed loan being good is
\[
\pi'_p(\omega) = \frac{\pi_p(x_p - \omega)_+}{(1 - \pi_p)\omega + (x_p - \omega)_+} \leq \pi_p,
\]
where $(\cdot)_+ = \max\{0, \cdot\}$. As $\pi'_p(\omega)$ is decreasing in $\omega$, their posterior expected value of a listed loan in this pool is
\[
V'(\omega) = \frac{\pi_p(x_p - \omega)_+ \cdot R_H + (1 - \pi_p)(\omega + (x_p - \omega)_+) \cdot R_L}{(1 - \pi_p)\omega + (x_p - \omega)_+},
\]
which is also decreasing in $\omega$.

Since uninformed investors are competitive, they enjoy zero profit in equilibrium. Therefore, they participate if they can meet the investment requirement while still break even:
\[
V'(\omega) \geq I,
\]
and the price $p_u$ they offer is
\[
p_u(\omega) = \begin{cases} 
V'(\omega), & \text{if } V'(\omega) \geq I, \\
0, & \text{if } V'(\omega) < I.
\end{cases}
\]

Importantly, the fact that both $\pi'_p(\omega)$ and $V'(\omega)$ are decreasing in $\omega$ indicates an important endogenous adverse selection problem introduced by more sophisticated investors becoming informed. As more informed investors pick up more good loans from the platform, the pool of listed loans facing uninformed investors becomes less valuable. Thus, uninformed investors may leave the market, hurting volume. As will be shown later, this adverse selection problem will also lead to fewer loan applications, hurting volume even if uninformed investors still participate.

**Informed investors.** Sophisticated investors, if becoming informed, only buy good loans, and their optimal offer price to an identified good loan is the loan applicant’s outside option:
\[
p_i(\omega) = \begin{cases} 
p_u(\omega) = V'(\omega), & \text{if } V'(\omega) \geq I, \\
I, & \text{if } V'(\omega) < I.
\end{cases}
\]
where $V'(\omega)$ is determined in (3.2). Intuitively, the informed investors offer the uninformed price $p_u(\omega)$ to a good loan when uninformed investors participate, while offer the investment requirement $I$ when uninformed investors do not participate. This implies that informed sophisticated investors outperform unsophisticated investors.

As long as there are good listed loans available, a sophisticated investor becomes informed if and only if the benefit of becoming informed and thus buying one good loan exceeds the cost set by the platform:

$$\pi_p(R_H - p_i(\omega)) \geq \mu,$$

which is more (less) likely to be satisfied when the information cost $\mu$ becomes lower (higher). Consequently, sophisticated investors’ outperformance (if any) is likely to become lower when the information cost becomes higher.

**Platform.** Under Definition 1, the platform price is $p = 0$ when no investor participates, and is pinned down by the marginal investor’s (latent) price when some investors participate:

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

where $V'(\omega)$ is determined in (3.2) and thus $p(\omega)$ is a decreasing function of $\omega$. Intuitively, unsophisticated investors are the marginal investor when they participate; while informed sophisticated investors are the marginal investor if uninformed investors do not participate.

Having characterized investors’ participation and pricing decisions in the sub-game, we finally solve for the platform’s optimal pre-screening and information distribution policies $\{\pi_p, \mu\}$.

**Theorem 1.** There exists two thresholds of platform pre-screening cost $0 < \kappa \leq \pi$ such that:

i). If $\kappa \geq \pi$, the platform optimally chooses a low $\pi_p$ and a low $\mu$. In this case, sophisticated investors become informed and they invest, while unsophisticated investors do not participate. The volume is $\min\{\pi_0x_0(I), \pi_p\Omega\}$.

ii). If $\kappa \leq \kappa$, the platform optimally chooses a high $\pi_p$ and a high $\mu$, where $\pi_p > \pi_p$ and $\mu > \mu$. In this case, sophisticated investors do not become informed, while all uninformed investors participate. The volume is $\frac{\pi_0x_0(p(0))}{\pi_p}$. 

16
iii). If \( \kappa < \kappa < \pi \), the platform optimally choose either \( \{\pi_p, \mu\} \) or \( \{\pi, \pi\} \), depending on which case gives the platform a higher expected payoff.

The mathematical expressions of the thresholds \( \kappa, \pi \), and the optimal policies \( \pi_p, \mu, \pi, \pi \) are given in Appendix A.

The following proposition is a direct time-series implication of Theorem 1:

**Proposition 1.** As platform pre-screening cost \( \kappa \) goes from high to low over time, the platform optimally:

a). increases pre-screening intensity \( \pi_p \), but \( \pi_p < 1 \) even if \( \kappa = 0 \);

b). increases investor information cost \( \mu \), that is, distributing less information.

Theorem 1 and in particular Proposition 1 imply that as the platform’s pre-screening cost becomes lower over its life-cycle, the platform optimally increases its pre-screening intensity while provides strictly less information to investors.\(^\text{27}\) Providing less information to investors may seem surprising at first.

To better understand the results and the economic trade-offs facing the platform, we discuss the first two generic equilibrium cases of Theorem 1 in detail by considering what happens if the platform deviate from its optimal strategies prescribed by Theorem 1. To do this, we formally show in the proof of Theorem 1 that under any generic platform policy \( \{\pi_p, \mu\} \), the economy must end up in one of the following four types of sub-game equilibrium. The respective volume in each type of sub-game equilibrium is given as follows:

<table>
<thead>
<tr>
<th>Equilibrium Investor Participation and Platform Volume</th>
<th>High ( \mu )</th>
<th>Low ( \mu )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low ( \pi_p )</td>
<td>0</td>
<td>( \min{\pi_0x_0(I), \pi_p\Omega} )</td>
</tr>
<tr>
<td>High ( \pi_p )</td>
<td>( \frac{\pi_0x_0(p(0))}{\pi_p} )</td>
<td>( \frac{\pi_0x_0(p(\Omega))}{\pi_p} )</td>
</tr>
</tbody>
</table>

\(^{27}\)Platform pre-screening cost decreases over time because of improvements in information technologies and screening algorithms. See Lending Club Investor Day 2017 Presentation available at ir.lendingclub.com/Cache/1001230258.pdf.
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

\[ 0 < \min\{\pi_0 x_0(I), \pi_p \Omega\} < \frac{\pi_0 x_0(p(\Omega))}{\pi_p} < \frac{\pi_0 x_0(p(0))}{\pi_p}. \]  

(3.8)

According to Theorem 1, a type-2 (type-4) equilibrium, where both the optimal \( \pi_p \) and \( \mu \) are low (high) happens when platform pre-screening cost is high (low).

First, we consider the trade-off underlying case i) in Theorem 1 when platform pre-screening cost \( \kappa \) is high. This immediately implies that the platform may find it optimal to choose a low pre-screening intensity \( \pi_p \) to save its cost. Because the quality of an average listed loan is low, unsophisticated investors cannot break even and thus will not participate. If the platform would choose a high \( \mu \), that is, provide little information to sophisticated investors, sophisticated investors find it too hard to become informed and thus stay out of the market as well. Hence, the economy would end up in a type-1 equilibrium where no investor participates at all and the volume is accordingly 0. In other words, screening is completely inefficient. To improve screening efficiency, it is optimal for the platform to instead choose a low \( \mu \) to attract sophisticated investors to become informed, according to their participation condition (3.6). In this case, informed sophisticated investors become the marginal investor and the platform price is \( p = I \). They identify and subsequently buy good loans up to their capacity, helping the platform boost the volume financed to either \( \pi_0 x_0(I) \) or \( \pi_p \Omega \), whichever is larger.\(^{28}\)

To summarize, in case i), the screening efficiency concern dominates due to the relatively high platform pre-screening cost. To invite sophisticated investors to help improve the screening and thus the financing outcomes, a low pre-screening intensity \( \pi_p \) and more information provision, captured by a low \( \mu \), are optimal. The economy ends up in a type-2 equilibrium.

Next, we consider case ii) in Theorem 1 when platform pre-screening cost \( \kappa \) is low. The platform may find it optimal to choose a high pre-screening intensity \( \pi_p \) to attract uninformed,\(^{28}\) Although our baseline model takes a reduced-form supply curve of loan applications regardless of the type, this result is robust in a possible extension, in which only good loan applicants apply when the marginal investor is informed (because bad loan applications know that their applications will be identified and rejected for sure). In this case, the platform does not pre-screen, and the platform price is still \( p = I \), only \( \min\{\pi_0 x_0(I), \pi_p \Omega\} \) good loan applicants will apply and then be listed, while the volume eventually financed is still \( \min\{\pi_0 x_0(I), \pi_p \Omega\} \).
unsophisticated investors, according to their participation condition (3.3). Accordingly, uninformed investors become the marginal investor. If the platform would choose a low \( \mu \), that is, provide much information to sophisticated investors, they would become informed and thus buy good loans earlier than uninformed investors.

This strategy would introduce an adverse selection problem: the quality of an average loan facing uninformed (and marginal) investors would become lower, leading to a lower platform price \( p(\Omega) \). This would in turn lead to a lower amount of loan applications \( x_0(p(\Omega)) \) in the first place, driving the economy into a type-3 equilibrium and hurting volume. If adverse selection is severe enough, it would deter uninformed investors from participating at all. Hence, to eliminate this adverse selection problem, the platform optimally chooses a high \( \mu \) so that sophisticated investors do not find it profitable to become informed. In this case, all the investors are uninformed as well as marginal investors, leading to a higher platform price \( p(0) \). This implies a higher amount of loan applications \( x_0(p(0)) \). Ultimately, all the listed loans \( \frac{\pi_0 x_0(p(0))}{\pi p} \) will be financed on the platform, yielding the highest possible volume for the platform.

To summarize, in case ii), the screening efficiency concern is no longer relevant because uninformed investors becoming willing to participate. Rather, the adverse selection concern dominates due to the risk of informed investors driving out uninformed investors and pushing down the platform price. To invite uninformed investor participation while mitigating adverse selection, a high pre-screening intensity \( \pi_0 \) and less information provision, captured by a high \( \mu \), are optimal. The economy ends up in a type-4 equilibrium.

We note that, the type-3 equilibrium, where sophisticated investors become informed and participate while uninformed unsophisticated investors also participate, reflects a transition phase from a type-2 to a type-4 equilibrium. In this transition phase, the platform still provides some information to sophisticated investors, which is inherited from a type-2 equilibrium. All the listed loans get financed, but adverse selection is active. This transition phase may capture the current developing stage of some platforms in which platforms start to discourage sophisticated investor screening but both informed and uninformed investors still co-exist.\textsuperscript{29}

\textsuperscript{29}Theoretically, this transition phase can be also sustained as a full equilibrium in a potential model extension, where the platform actively learn from sophisticated investors’ screening criteria to further reduce its own pre-screening cost, that is, where \( \kappa(\omega) \) is decreasing in \( \omega \).
3.3 Empirical Predictions

Our model generates a set of empirical predictions, which we subsequently bring to the data.

Prediction 1. Sophisticated (and informed) investors outperform unsophisticated (and uninformed) investors at any loan price.

Prediction 2. When their information cost becomes higher, sophisticated investors are less likely to become informed and thus their outperformance may shrink, at any loan price.

Predictions 1 and 2 focus on investors and derive from the sub-game equilibrium where the platform’s strategy is considered as a parameter from the investors’ perspective. Specifically, the two predictions come from conditions (3.5) and (3.6) of the model. Although they are theoretically straightforward, empirically it is not obvious whether there is any room for more sophisticated investors to screen and outperform, and if so, to what extent performance responds to platform policy changes. Moreover, in the traditional banking model or in the shadow banking model, depositors and investors, however sophisticated they are, are unlikely to outperform others. Thus, observing empirically sophisticated investor outperformance serves as a smoking gun of marketplace lending being a different banking model.

Prediction 3. The platform may increase the information cost of sophisticated investors as it develops, by distributing fewer variables to investors.

Prediction 4. The platform may increase its pre-screening intensity as it develops, but it will eventually keep an intermediate pre-screening intensity, and it may fluctuate over time.

Predictions 3 and 4 focus on the platform design and derive from the full equilibrium. They stem from Theorem 1 and Proposition 1. Beyond the key intuition of managing investor adverse selection, which we have articulated above, an optimally intermediate pre-screening intensity is justified by the fact that as long as no investor becomes informed and all uninformed investors participate, further increasing $\pi_p$ only hurts volume. This suggests that in reality, the platform optimally calibrates its pre-screening intensity to varying economic conditions, which may be reflected by a fluctuation of pre-screening intensity over time.
4 Data and Investor Sophistication

Our empirical analysis relies on a proprietary dataset, which combines investor data covering both sophisticated and unsophisticated retail investors, with borrowers’ data from the platforms.

4.1 Data

Our investor-level data is provided by LendingRobot, a leading robo-advisor for retail investors on lending marketplaces.\footnote{LendingRobot was acquired by NSR Invest, its main competitor, in 2017.} 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.\footnote{For more detail on LendingRobot credit model, refer to\url{http://blog.lendingrobot.com/research/predicting-the-number-of-payments-in-peer-lending/}.} 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.\footnote{Monitored-only accounts might be invested with the platform automatic investment strategy, for investment Lending Club \textit{automated investing} feature, or at the investor’s discretion.}

The LendingRobot data is organized at the investor level, as shown in Figure 1. We access a set of variables at each level of this 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 invest-
ment 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, 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. 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.

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 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.

---

33 See “Hedge Funds Pursue Alternative Lending,” the Financial Times, November 2, 2014.
34 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.
35 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.
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 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.\(^{36}\) 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 Lending Club.\(^{37}\)

\[
\text{Prob}(TypeAccount_i = 1) = \beta \times BorrowerCharacteristics + IR_i + m_t + \epsilon_i, \quad (5.1)
\]

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 $\epsilon_i$ is the error term. Table 2 displays the regression coefficients.

\(^{36}\)The results are similar to logit regressions.

\(^{37}\)We abstract from the question on whether loans from the same grade or sub-grades are on average fairly priced.
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.\footnote{For instance, the granularity of the interest rate scale mechanically creates heterogeneity in risk within the same price.}

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 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 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.\footnote{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.}

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.
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.\textsuperscript{40}

We run the following specification:

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

where ChargedOff is an indicator variable for the loan to be in default or charged-off, \(\mathbb{1}_{\text{TypeAccount}}\) 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 3 displays 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 \(\mathbb{1}_{\text{TypeAccount}}\) with year fixed effects, and with loan grade fixed effects.

\[\text{[Insert Table 3]}\]

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

\textsuperscript{40}The results are robust under a logit specification.
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.\footnote{This cost goes to infinity as the information is reduced, as in the absence of information it is impossible to further screen.} We therefore investigate the relationship between the information set available on the platform and investor screening performance. We find empirical evidence consistent with 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:

\[
Prob(\text{ChargedOff} = 1)_i = \beta_1 \times 1_{\text{robot}} + \beta_2 \times 1_{\text{robot}} \times Post \\
+ \beta_3 \times 1_{\text{advance}} + \beta_4 \times 1_{\text{advance}} \times Post \\
+ \beta_5 \times 1_{\text{monitor}} + \beta_6 \times 1_{\text{monitor}} \times Post + IR_i + m_t + \epsilon_i
\]  

(5.3)

where \(1_{\text{robot}}\) is an indicator variable equal to one if at least two robot accounts are invested in the loan \(i\), \(1_{\text{advanced}}\) is an indicator variable equal to one if at least two advanced accounts are invested in the loan, \(1_{\text{monitor}}\) 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 \(\epsilon_i\) 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 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.\(^{42}\)

We rationalize this reduction in sophisticated investors’ performance by the platform “evening

\(^{42}\)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.
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 in Figure 4.

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 4]

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 1A.2 in the appendix illus-
trates 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 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 Prediction 4 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 5. 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 2.

[Insert Figure 5]

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 6. 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.\footnote{Using grades instead of sub-grades for Lending Club still yields a better predictor than Prosper grades.}

[Insert Figure 6]

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 Prediction 4 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.

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 7]

The figure shows that the out-performance of robot and advanced investors appear to decrease with time.\footnote{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.} 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 3 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 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


Figure 2: 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 3: 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 4: 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 5: 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 6: 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 7: 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 monitor-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

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

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%.
Table 2: Investor Screening - 2014-2016

<table>
<thead>
<tr>
<th></th>
<th>Lending Club</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Robot (1)</td>
<td>Advanced (2)</td>
<td>Monitored (3)</td>
<td>Robot (4)</td>
<td>Advanced (5)</td>
<td>Monitored (6)</td>
</tr>
<tr>
<td>Loan amount</td>
<td>0.005***</td>
<td>0.008***</td>
<td>0.015***</td>
<td>0.015***</td>
<td>0.012***</td>
<td>0.018***</td>
</tr>
<tr>
<td></td>
<td>(18.89)</td>
<td>(27.97)</td>
<td>(36.16)</td>
<td>(25.83)</td>
<td>(19.00)</td>
<td>(21.22)</td>
</tr>
<tr>
<td>FICO Score</td>
<td>0.000</td>
<td>0.001***</td>
<td>-0.001***</td>
<td>-0.000</td>
<td>0.000***</td>
<td>-0.000***</td>
</tr>
<tr>
<td></td>
<td>(1.44)</td>
<td>(11.62)</td>
<td>(-10.56)</td>
<td>(-0.01)</td>
<td>(1.76)</td>
<td>(-2.95)</td>
</tr>
<tr>
<td>Annual Income</td>
<td>0.001***</td>
<td>0.001***</td>
<td>0.000***</td>
<td>-0.000***</td>
<td>0.001***</td>
<td>-0.000*</td>
</tr>
<tr>
<td></td>
<td>(7.18)</td>
<td>(13.42)</td>
<td>(9.83)</td>
<td>(-2.90)</td>
<td>(5.67)</td>
<td>(-1.78)</td>
</tr>
<tr>
<td>Employment Length</td>
<td>0.002***</td>
<td>0.007***</td>
<td>0.001***</td>
<td>0.000</td>
<td>0.002***</td>
<td>0.002***</td>
</tr>
<tr>
<td></td>
<td>(8.96)</td>
<td>(19.42)</td>
<td>(5.47)</td>
<td>(1.43)</td>
<td>(6.27)</td>
<td>(4.01)</td>
</tr>
<tr>
<td>Debt to Income</td>
<td>-0.001***</td>
<td>-0.002***</td>
<td>0.001***</td>
<td>0.041</td>
<td>-0.108*</td>
<td>0.137***</td>
</tr>
<tr>
<td></td>
<td>(-4.67)</td>
<td>(-10.36)</td>
<td>(8.71)</td>
<td>(1.37)</td>
<td>(-1.74)</td>
<td>(3.95)</td>
</tr>
<tr>
<td>Own Home Ownership</td>
<td>0.033***</td>
<td>0.054***</td>
<td>0.006**</td>
<td>-0.017**</td>
<td>0.024**</td>
<td>0.006</td>
</tr>
<tr>
<td></td>
<td>(8.96)</td>
<td>(14.33)</td>
<td>(2.53)</td>
<td>(-2.71)</td>
<td>(2.53)</td>
<td>(1.27)</td>
</tr>
<tr>
<td>Open Accounts</td>
<td>0.002***</td>
<td>0.001***</td>
<td>0.000</td>
<td>0.001***</td>
<td>0.002***</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>(7.04)</td>
<td>(5.73)</td>
<td>(0.89)</td>
<td>(3.30)</td>
<td>(2.50)</td>
<td>(0.03)</td>
</tr>
<tr>
<td>First Credit Line</td>
<td>-0.000</td>
<td>-0.001**</td>
<td>-0.001***</td>
<td>-0.000</td>
<td>-0.001***</td>
<td>-0.001**</td>
</tr>
<tr>
<td></td>
<td>(-1.56)</td>
<td>(-2.50)</td>
<td>(-9.17)</td>
<td>(-0.62)</td>
<td>(-5.19)</td>
<td>(-2.50)</td>
</tr>
<tr>
<td>Delinquency</td>
<td>-0.005***</td>
<td>-0.019***</td>
<td>-0.006***</td>
<td>-0.000</td>
<td>-0.002***</td>
<td>-0.001***</td>
</tr>
<tr>
<td></td>
<td>(-6.70)</td>
<td>(-18.68)</td>
<td>(-6.73)</td>
<td>(-0.24)</td>
<td>(-4.34)</td>
<td>(-3.78)</td>
</tr>
<tr>
<td>Term</td>
<td>-0.012</td>
<td>-0.066***</td>
<td>0.045***</td>
<td>-0.000</td>
<td>-0.004***</td>
<td>0.004***</td>
</tr>
<tr>
<td></td>
<td>(-1.59)</td>
<td>(-7.65)</td>
<td>(6.89)</td>
<td>(-0.42)</td>
<td>(-5.31)</td>
<td>(8.40)</td>
</tr>
<tr>
<td>Inquiries, last 6 months</td>
<td>-0.038***</td>
<td>-0.060**</td>
<td>-0.003*</td>
<td>-0.008***</td>
<td>-0.045***</td>
<td>-0.001</td>
</tr>
<tr>
<td></td>
<td>(-14.47)</td>
<td>(-28.10)</td>
<td>(-2.00)</td>
<td>(-3.59)</td>
<td>(-11.45)</td>
<td>(-0.45)</td>
</tr>
<tr>
<td>Revolving Utilization Rate</td>
<td>0.000***</td>
<td>0.000***</td>
<td>0.000***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(3.39)</td>
<td>(5.36)</td>
<td>(2.93)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Derogatory public records</td>
<td>-0.042***</td>
<td>-0.044***</td>
<td>0.003*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-5.07)</td>
<td>(-6.65)</td>
<td>(1.92)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mortgage</td>
<td>0.045***</td>
<td>0.075***</td>
<td>0.017***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(9.20)</td>
<td>(15.44)</td>
<td>(8.09)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Loan Purpose</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Car</td>
<td>0.051***</td>
<td>0.021*</td>
<td>0.007</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(6.33)</td>
<td>(1.89)</td>
<td>(1.08)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Credit Card</td>
<td>0.063***</td>
<td>0.072***</td>
<td>-0.010***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(8.37)</td>
<td>(13.71)</td>
<td>(-4.53)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Home Improvement</td>
<td>-0.005</td>
<td>-0.027***</td>
<td>-0.004</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-1.52)</td>
<td>(-6.34)</td>
<td>(-1.11)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medical</td>
<td>-0.103***</td>
<td>-0.189***</td>
<td>0.006</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-7.23)</td>
<td>(-16.90)</td>
<td>(0.95)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| Interest Rate FE         | Yes         | Yes       | Yes       | Yes       | Yes       | Yes       |
| Month FE 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 R²                | 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

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Prob(Charged-Off)</td>
<td></td>
<td></td>
<td>Prob(Charged-Off)</td>
<td></td>
<td></td>
<td>Prob(Charged-Off)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Account Type</td>
<td>-0.031***</td>
<td>-0.044***</td>
<td>-0.008***</td>
<td>-0.084***</td>
<td>-0.070***</td>
<td>-0.005</td>
<td>0.012*</td>
<td>-0.015***</td>
<td>0.007***</td>
</tr>
<tr>
<td></td>
<td>(-10.84)</td>
<td>(-18.04)</td>
<td>(-4.68)</td>
<td>(-20.56)</td>
<td>(-19.86)</td>
<td>(-1.27)</td>
<td>(1.66)</td>
<td>(-3.64)</td>
<td>(2.21)</td>
</tr>
<tr>
<td>Account Type x 2015</td>
<td>0.051***</td>
<td>0.029***</td>
<td>-0.006</td>
<td>(10.38)</td>
<td>(7.11)</td>
<td>(-1.27)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Account Type x 2016</td>
<td>0.075***</td>
<td>0.050***</td>
<td>-0.002</td>
<td>(13.66)</td>
<td>(12.42)</td>
<td>(-0.45)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Account Type x Grade B</td>
<td>-0.041***</td>
<td>-0.019***</td>
<td>-0.009**</td>
<td>(-3.72)</td>
<td>(-3.36)</td>
<td>(-2.11)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Account Type x Grade C</td>
<td>-0.058***</td>
<td>-0.030***</td>
<td>-0.015***</td>
<td>(-6.36)</td>
<td>(-5.28)</td>
<td>(-3.07)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Account Type x Grade D</td>
<td>-0.052***</td>
<td>-0.037***</td>
<td>-0.027***</td>
<td>(-5.97)</td>
<td>(-6.06)</td>
<td>(-4.58)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Account Type x Grade E</td>
<td>-0.049***</td>
<td>-0.047***</td>
<td>-0.019**</td>
<td>(-4.62)</td>
<td>(-4.58)</td>
<td>(-2.22)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Account Type x Grade F</td>
<td>-0.026**</td>
<td>-0.039***</td>
<td>-0.005</td>
<td>(-2.43)</td>
<td>(-3.19)</td>
<td>(-0.48)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Account Type x Grade G</td>
<td>-0.089***</td>
<td>-0.081***</td>
<td>-0.006</td>
<td>(-4.31)</td>
<td>(-3.66)</td>
<td>(-0.31)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

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

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

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. $I_{\text{robot}}$ is an indicator variable equal to one if at least two robot accounts are invested in the loan, $I_{\text{advance}}$ is an indicator variable equal to one if at least two advance accounts are invested in the loan, $I_{\text{monitor}}$ 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 Proof of Theorem 1

The proof proceeds in three steps. First, we characterize different types of sub-game equilibria as described in the main text. Second, 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. Third, we show that a type-2 (type-4) sub-game equilibrium happens in a full equilibrium when the platform’s information cost $\kappa$ is larger (smaller) than a threshold.

Step 1. This step repeatedly uses the two participation conditions (3.3) and (3.6). To recap, condition (3.3) 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 $(\pi_p, \mu)$-space denoted by $S = \{(\pi_p, \mu, \cdot) | \pi_0 \leq \pi_p \leq 1, \mu \geq 0\}$.

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 (3.3) and (3.6) give that

$$SE_1 = \{ (\pi_p, \mu, \cdot) | V'(0) < I, \pi_p(R_H - I) < \mu \}.$$  \hfill (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 (3.3) and (3.6) give that

$$SE_2 = \{ (\pi_p, \mu, \cdot) | V'(\Omega) < I, \pi_p(R_H - I) \geq \mu \},$$  \hfill (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 (3.3) and (3.6) give that

$$SE_3 = \{ (\pi p, \mu, \cdot) | V'(\Omega) \geq I, \pi p(R_H - V'(\Omega)) \geq \mu \},$$  

(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 (3.3) and (3.6) give that

$$SE_4 = \{ (\pi p, \mu, \cdot) | V'(0) \geq I, \pi p(R_H - V'(0)) < \mu \}.$$  

(A.4)

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

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

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 $(\pi p, \mu)$. However, as we show later in Step 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 this step.

**Step 2.** 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 $\pi p(R_H - I) \geq \mu'$. Notice that $\mu'$ must exist because $0 < \pi_0 \leq \pi p$. Because $V'(0) \geq V'(\Omega)$, we have that $(\pi p, \mu', \cdot) \in SE_2$ as characterized in (A.2). Because changing $\mu$ 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 \( \mu \) to \( \mu' \) without changing \( \pi_p \), 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 \( \pi_0 x_0(p(\Omega)) \pi_p \).

Consider an alternative equilibrium profile \( (\pi_p, \mu', \cdot), \mu' > \mu \) such that \( \pi_p(R_H - V'(0)) < \mu' \). Notice that \( \mu' \) must exist because \( 0 < \pi_0 \leq \pi_p \) and \( I \leq V'(0) < R_H \). Again because \( V'(0) \geq V'(\Omega) \), we have that \( (\pi_p, \mu', \cdot) \in SE_4 \) as characterized in (A.4). Because changing \( \mu \) 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 \( \mu \) to \( \mu' \) without changing \( \pi_p \), 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 3.** 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

\[
S_H = \{ (\pi_p, \mu, \cdot) | V'(0) \geq I \}.
\]

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

\[
\hat{\pi}_p = \frac{I - R_L}{R_H - R_L} \in (0,1),
\]

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

\[
(SE_1 \cup SE_4) \cap \{ (\pi_p, \mu, \cdot) | \mu > \hat{\mu} \} = S \cap \{ (\pi_p, \mu, \cdot) | \mu > \hat{\mu} \}.
\]

Hence, if

\[
C(\pi_p) = \frac{1}{2} \kappa((\pi_p - \pi_0)_+)^2 \geq \min\{\pi_0 x_0(I), \pi_p \Omega\},
\]

48
by the argument in Step 2, 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

$$\kappa = 2 \min\{\pi_0 x_0(I), \bar{\pi}_p \Omega\} > 0.$$

Otherwise if

$$C(\bar{\pi}_p) = \frac{1}{2} \kappa((\bar{\pi}_p - \pi_0)_+)^2 \leq \frac{\pi_0 x_0(p(\Omega))}{\bar{\pi}_p},$$

a type-4 sub-game equilibrium is achievable by choosing \(\bar{\pi}_p\) such that it solves

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

and choosing \(\bar{\pi} > \hat{\pi}\), in which case we define

$$\bar{\pi} = \frac{2\pi_0 x_0(p(\Omega))}{\bar{\pi}_p((\bar{\pi}_p - \pi_0)_+)^2} > 0.$$ 

Because \(\pi > \kappa\), this concludes the proof.

**B  The Supply Curve of Loan Applications**

This appendix provides one simple microfoundation for the upward-sloping supply curve of loan applications \(x_0(p)\).

Keeping all the initial setting in the baseline model, we further assume that a loan applicant, if and only if she successfully establishes a loan from a lending marketplace, may incur an adoption cost. In other words, if a loan applicant goes to the platform, applies, but does not get funded, there will be no adoption cost. This adoption cost reflects the fact that the concept of marketplace lending is still relatively new, and thus a loan applicant may incur physical or mental costs to understand its rules and trust them, to become familiar with its monthly repayment requirements, as well as to actually set up the monthly repayments with the platform. Different
applicants may have different adoption costs, and it is also natural that some applicants, who are already familiar with marketplace lending, may not incur any adoption cost at all.

This notion of adoption cost follows the classic idea of the Hotelling model and has been commonly used in the literature of bank competition. For example, Hauswald and Marquez (2006) use it to model how different banks may specialize in different borrower characteristics. Thus a borrower who borrows from a bank that does not specialize in her characteristics incurs a cost. In a marketplace lending context similar to ours, de Roure, Pelizzon and Thakor (2018) introduce a notion of “acquisition cost” to reflect that it is costly for a borrower to leave her bank, with which she already has a relationship with, to join a lending platform.

Specifically, we assume that the population of potential loan applicants who do not have any adoption costs is \( q(0) > 0 \), and the population of potential loan applicants who have a positive adoption cost \( z \) is \( q(z) \geq 0 \), where \( z \in (0, +\infty) \) and follows an arbitrary distribution \( F(z) \). Consistent with our baseline model, we still assume that loan applicants do not know their type ex-ante. Following the literature such as Hauswald and Marquez (2006) and de Roure, Pelizzon and Thakor (2018), we also assume that the adoption cost is independent to the type of loan applicants.

Under this setting, any platform price \( p \) that attracts a potential loan applicant to apply has to be at least as high as the investment requirement \( I \) plus the adoption cost of that specific potential applicant. Thus, the population of actual loan applications at any platform price \( p \) is given by

\[
x_0(p) = \begin{cases} 
q(0) + \int_0^{p-I} q(z) dF(z), & \text{if } p > I, \\
q(0), & \text{if } p = I,
\end{cases}
\]

which is increasing in \( p \). Note that, because loan applicants do not know their type ex-ante, the above expression does not depend on the distribution of their types. This micro-foundation provides us with the reduced-form supply curve of loan applications we use in the baseline model.

C Additional Tables and Figures
## Table IA.1: Investor Screening - 2014-2016 - Robustness

<table>
<thead>
<tr>
<th></th>
<th>Lending Club</th>
<th></th>
<th>Prosper</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Robot (1)</td>
<td>Advanced (2)</td>
<td>Monitored (3)</td>
<td>Robot (4)</td>
<td>Advanced (5)</td>
</tr>
<tr>
<td>Loan amount</td>
<td>0.011***</td>
<td>0.018***</td>
<td>0.023***</td>
<td>0.051***</td>
<td>0.027***</td>
</tr>
<tr>
<td></td>
<td>(16.82)</td>
<td>(26.95)</td>
<td>(36.81)</td>
<td>(19.85)</td>
<td>(13.60)</td>
</tr>
<tr>
<td>FICO Score</td>
<td>0.000***</td>
<td>0.002***</td>
<td>-0.001***</td>
<td>0.001***</td>
<td>0.001***</td>
</tr>
<tr>
<td></td>
<td>(2.71)</td>
<td>(12.32)</td>
<td>(-10.21)</td>
<td>(3.54)</td>
<td>(4.67)</td>
</tr>
<tr>
<td>Annual Income</td>
<td>0.002***</td>
<td>0.003***</td>
<td>0.001***</td>
<td>-0.001***</td>
<td>0.001***</td>
</tr>
<tr>
<td></td>
<td>(7.63)</td>
<td>(12.20)</td>
<td>(9.20)</td>
<td>(-4.38)</td>
<td>(10.78)</td>
</tr>
<tr>
<td>Employment Length</td>
<td>0.006***</td>
<td>0.015***</td>
<td>0.002***</td>
<td>0.003***</td>
<td>0.005***</td>
</tr>
<tr>
<td></td>
<td>(9.30)</td>
<td>(20.99)</td>
<td>(7.31)</td>
<td>(3.96)</td>
<td>(10.80)</td>
</tr>
<tr>
<td>Debt to Income</td>
<td>-0.001***</td>
<td>-0.004***</td>
<td>0.002***</td>
<td>-0.394***</td>
<td>-0.406***</td>
</tr>
<tr>
<td></td>
<td>(-4.00)</td>
<td>(-14.26)</td>
<td>(9.81)</td>
<td>(-3.87)</td>
<td>(-9.45)</td>
</tr>
<tr>
<td>Own Home Ownership</td>
<td>0.086***</td>
<td>0.137***</td>
<td>0.015***</td>
<td>-0.228***</td>
<td>0.012*</td>
</tr>
<tr>
<td></td>
<td>(8.98)</td>
<td>(14.31)</td>
<td>(4.03)</td>
<td>(-4.71)</td>
<td>(1.78)</td>
</tr>
<tr>
<td>Open Accounts</td>
<td>0.005***</td>
<td>0.004***</td>
<td>0.000</td>
<td>0.010***</td>
<td>0.008*</td>
</tr>
<tr>
<td></td>
<td>(7.51)</td>
<td>(6.99)</td>
<td>(1.06)</td>
<td>(7.42)</td>
<td>(6.18)</td>
</tr>
<tr>
<td>First Credit Line</td>
<td>-0.002**</td>
<td>-0.003***</td>
<td>-0.002***</td>
<td>-0.000</td>
<td>-0.003***</td>
</tr>
<tr>
<td></td>
<td>(-2.51)</td>
<td>(-4.63)</td>
<td>(-9.86)</td>
<td>(-6.65)</td>
<td>(-7.11)</td>
</tr>
<tr>
<td>Delinquency</td>
<td>-0.012***</td>
<td>-0.043***</td>
<td>-0.011***</td>
<td>0.001</td>
<td>-0.004***</td>
</tr>
<tr>
<td></td>
<td>(-7.13)</td>
<td>(-22.24)</td>
<td>(-10.14)</td>
<td>(1.63)</td>
<td>(-8.97)</td>
</tr>
<tr>
<td>Term</td>
<td>-0.054***</td>
<td>-0.170***</td>
<td>0.057***</td>
<td>-0.006***</td>
<td>-0.012**</td>
</tr>
<tr>
<td></td>
<td>(-2.94)</td>
<td>(-11.68)</td>
<td>(6.13)</td>
<td>(-4.11)</td>
<td>(-16.17)</td>
</tr>
<tr>
<td>Inquiries, last 6 months</td>
<td>-0.096***</td>
<td>-0.162***</td>
<td>-0.005**</td>
<td>-0.067***</td>
<td>-0.133***</td>
</tr>
<tr>
<td></td>
<td>(-14.95)</td>
<td>(-22.42)</td>
<td>(-2.02)</td>
<td>(-7.70)</td>
<td>(-18.18)</td>
</tr>
<tr>
<td>Revolving Utilization Rate</td>
<td>0.001***</td>
<td>0.001***</td>
<td>0.000***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(3.09)</td>
<td>(5.82)</td>
<td>(4.26)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Derogatory public records</td>
<td>-0.113***</td>
<td>-0.117***</td>
<td>0.007***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-5.67)</td>
<td>(-6.48)</td>
<td>(2.93)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mortgage</td>
<td>0.118***</td>
<td>0.193***</td>
<td>0.028***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(9.74)</td>
<td>(14.01)</td>
<td>(9.90)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Car</td>
<td>0.130***</td>
<td>0.073***</td>
<td>0.015</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(5.56)</td>
<td>(3.43)</td>
<td>(1.64)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Credit Card</td>
<td>0.171***</td>
<td>0.199***</td>
<td>-0.019***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(8.42)</td>
<td>(11.15)</td>
<td>(-6.23)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Home Improvement</td>
<td>-0.010</td>
<td>-0.078***</td>
<td>-0.010**</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-1.38)</td>
<td>(-10.49)</td>
<td>(-2.14)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medical</td>
<td>-0.287***</td>
<td>-0.407***</td>
<td>0.007</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-7.19)</td>
<td>(-14.88)</td>
<td>(0.94)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Interest Rate FE**
- Yes

**Month FE**
- Yes

**Cluster**
- Int.Rate

**Observations**
- 365,685

**Pseudo $R^2$**
- 0.355

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

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

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.
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.