Marketplace Lending: A New Banking Paradigm?*

Boris Vallée
Harvard Business School

Yao Zeng
University of Washington

May, 2018

Abstract

Marketplace lending relies on screening and information production by investors, a major deviation from the traditional banking paradigm. Theoretically, the participation of sophisticated investors improves screening outcomes but also creates adverse selection among investors. In maximizing loan volume, the platform trades off these two forces. As the platform develops, it optimally increases platform pre-screening intensity but decreases information provision to investors. Using novel investor-level data, we find that sophisticated investors systematically outperform, and this outperformance shrinks when the platform reduces information provision to investors. Our findings shed light on the optimal distribution of information production in this new lending model.

KEYWORDS: Marketplace lending, screening, information production, sophisticated investors, adverse selection, Fintech

JEL: G21, G23, D82

*We are grateful to Itay Goldstein, Wei Jiang, Andrew Karolyi (the Editors), Shimon Kogan, Gregor Matvos, Adair Morse (discussants), three anonymous referees, Chiara Farronato, Simon Gervais, David Scharfstein, Andrei Shleifer, Jeremy Stein, and Xiaoyan Zhang for helpful comments. We also thank seminar and conference participants at the RFS FinTech Workshops at Columbia University and Cornell Tech, NY Fed/NYU Stern Conference on Financial Intermediation, NYU Stern FinTech Conference, Texas Finance Festival, 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 Lending Club, Prosper, and Funding Circle, have been rapidly gaining market share in consumer and small business lending over the last decade.[1] This rapid development has important implications on the consumer lending market, and more broadly on banking.[2]

Designed as a two-sided platform that does not have skin in the game, a lending marketplace brings innovations to traditional banking on both the borrower and investor side. The innovation on the borrower side relies on streamlining an online application process that uses low-cost information technology to collect standardized information from dispersed individual borrowers on a large scale. With such information, lending platforms pre-screen loan applications, list some of them on the platform, and allocate them into a risk bucket.

The innovation on the investor side, which is the focus of this paper, is equally important and complementary. Instead of pooling loans and issuing safe claims, platforms provide detailed information on each pre-screened loan application to investors, relying on investors to further screen borrowers and to directly invest in individual loans. These investors include informationally sophisticated investors, both among retail and institutional investors, as well as passive and unsophisticated investors.[3] The diverse investor base on the platform performs large-scale borrower screening, and the resulting joint information production between the platform and investors challenges the traditional role of banks as being the exclusive information producer on behalf of investors (Gorton and Pennacchi, 1990, Dang, Gorton, Holmstrom and Ordonez, 2017). Notably, the feature of information production by investors who fully bear loan credit risks differs not only from the traditional banking model — where depositors hold a safe, information-insensitive claim — but also from modern securitization (i.e., shadow banking).

The joint information production between the platform and investors of varying sophisti-

[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]Currently, more than half of retail borrowers on lending platforms use the marketplace loans to refinance existing loans or pay off credit balances. Although the refinancing market represents an attractive entry point for platforms to achieve scale rapidly, it does not represent their whole addressable market, which include any form of consumer or small business loan.

[3]The platforms are typically segmented between retail and institutional investors, and allocate randomly loans between the two pool of investors. Platforms in general offer a passive investing feature as well.

cation poses several research questions, which we address in this paper. First, are more sophisticated investors on lending platforms screening borrowers more intensively and thereby systematically outperforming less sophisticated investors? If so, then, how does sophisticated investor 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 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 pitfalls of marketplace lending.

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 half of the 100 variables on borrower characteristics that it 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.[4] Given that informational transparency is crucial to substitute for the platform's skin in the game and to allow for investor screening, what is the economic rationale behind this reduction of the information set provided to investors?

In addressing the questions above, we develop a model for marketplace lending and test its predictions using a novel dataset that includes borrower and investor data.

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, unsophisticated investors do not screen; they invest in a listed loan passively if the platform pre-screens intensely enough so that they can break even on average, or they do not invest at all. Hence, sophisticated investors outperform unsophisticated ones. Because sophisticated investors can identify good loans and finance them, their participation helps boost the volume of loans financed on the platform when unsophisticated investors are reluctant to invest.

However, the heterogeneity in investor sophistication creates an endogenous adverse selection problem among investors, which can hurt volume. Because sophisticated investors can identify and finance good loans before unsophisticated investors invest, sophisticated investor partici-

[4] We provide more discussion relevant institutional details and this specific event in Section 2.2.

pation lowers the average quality of loans eventually facing unsophisticated investors. Being aware of this adverse selection problem, unsophisticated investors require a lower loan price to break even, resulting in a lower prevailing loan price on the platform. This lower price reduces the amount of loan applications on the platform, hurting volume. If adverse selection becomes too severe, unsophisticated investors may not break even and may 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 environment encourages sophisticated investor participation, boosting volume even if unsophisticated 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 unsophisticated investors are willing to invest, but at the same time distributes less information to mitigate the adverse selection caused by sophisticated investors.

Testing the model predictions crucially relies on data of investors of heterogenous sophistication. Although borrower-level data is made public by the platforms, data on investor characteristics and their loan portfolios is not publicly available.[5] Fortunately, we obtain a rich dataset provided by LendingRobot, an algorithmic third-party, which includes portfolio composition for a large set of retail investors on two largest lending platforms, Lending Club and Prosper. We can therefore study sophisticated investor screening and outperformance 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 and order-placing technologies offered by LendingRobot.

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

[5] 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.

lower probability of default for a given loan, meaning that sophisticated investors systematically 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. Sophisticated investors indeed produce information.

Using the 2014 Lending Club event described above, we then implement a difference-in-differences methodology to establish causal evidence of the impact of a large reduction in platform information provision on sophisticated investors' performance. We find that sophisticated investor outperformance drops by more than half at the time of the reduction. We rationalize this unanticipated event, corresponding to the platform "evening the playing field", by referring to the theoretical argument that platforms actively manage adverse selection. Under our rationalization, this event suggests that lending platforms value unsophisticated investors and therefore act as if they "protect" them, even in the absence of specific regulation.

Finally, we find that platforms' pre-screening intensity has also been improving in the sense that platform risk buckets are increasingly precise at predicting default. In addition to attracting unsophisticated investors directly, this increased precision is also likely to mitigate adverse selection by reducing the heterogeneity of loans within a given risk bucket. Consistent with platforms managing adverse selection, we also find robust time-series evidence that sophisticated investor outperformance has become lower in recent years.

Although our empirical tests mainly rely on the heterogeneity within the retail investor segment, our findings have external validity for the institutional investor segment of marketplace lending, as the heterogeneity in sophistication is comparable across investor segments. Many institutional investors, such as pension funds, only apply a little screening (for instance, only relying on a grade threshold) as retails investors do; while other institutional investors, such as hedge funds, develop highly sophisticated investment strategies that are comparable to what LendingRobot offers to investors.

We focus on the robust features in the development of marketplace lending so far while leave a number of interesting questions for future research. These include, for example, the overall welfare implications of marketplace lending, and whether marketplace lending poses any

financial stability concerns due to the increasing participation of institutional investors.

Related Literature. Our paper contributes to the burgeoning literature on marketplace lending by directly examining the sharing of information production between platforms and investors, one key factor that 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 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 complements a few recent papers that study the motivation behind Propser'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). Our work also complements Hildebrand, Puri and Rocholl (2016) who uses Prosper's earlier policy change of removing origination rewards to examine the role of investors' skin in the game.

In the context of marketplace lending competing with traditional banking, de Roure, Pelizzon and Thakor (2018) and Tang (2018) study whether these two types of lenders complement or substitute each other when providing credit to different borrower segments. Balyuk (2017) examines whether households improve their credit access from traditional banks when using marketplace lending. These papers support traditional banking as a benchmark to marketplace lending. Our work complements them by focusing on the investor side of marketplace lending and exploring how the role of these investors deviates from that of traditional bank depositors.

In addition, our paper also relates 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 that Fintech firms account for almost a third of shadow bank loan originations by 2015. Fuster, Plosser, Schnabl and Vickery (2018) also find that FinTech lenders process mortgage applications faster than traditional banks. In the context of shadow banking, however, investors produce little to no information as they do not screen loans individually, while marketplace lending in its original design does not conduct maturity or liquidity transformation. Consistent with this conceptual difference, Keys, Mukherjee, Seru and Vig (2010) find that securitization leads to loosened screening despite the skin in the game, while in marketplace lending this risk may be mitigated by the sharing of information production between the platforms and investors.

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 for marketplace lending, it sheds 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 the review of Morse (2015) for general institutional details of lending marketplaces.[6]

[6]Marketplace Lending received an important coverage in the year 2016, following some governance issues at the main platform, Lending Club. See "Inside the Final Days of Lending Club CEO Renaud Laplanche," the *Wall Street Journal*, May 16, 2016 for a detailed document. While these events revealed serious governance issue at Lending Club, they do not speak to the marketplace lending economic model.

2.1 Platform Information Collection and Pre-Screening

Information collection. By design, lending platforms only collect information on borrowers via online self-reporting, and through credit pulls. Thus, information collection itself is standardized and mostly costless.[7] 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.

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. On the intensive margin, they allocate the listed 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 risk buckets map into interest rates, i.e. loan prices. For a given loan maturity and at any given point in time, a risk bucket is associated with a single interest rate at loan issuance.

Platform pre-screening is scalable but costly. It is not equivalent to simply picking a FICO score threshold and listing loans above that threshold. Rather, it may potentially use many variables. The development of such a pre-screening model requires sophisticated data analysis, which involves a fixed cost. The more precise the allocation of loans into risk-buckets is, including the ones that do not get listed, the more costly 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.

[7] 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 or raising privacy issues, would increase the drop-off rate during the online application, which would effectively make information collection costly for the platform.

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 or Prosper.[8] This represents a deviation from the traditional banking model in which depositors hold a demandable debt contract issued by the bank without having any knowledge of the underlying loans (Gorton and Pennacchi, 1990, Dang, Gorton, Holmstrom and Ordonez, 2017).[9,10]

Information distribution. Platforms provide investors with a set of standardized information for each listed loan. This set is typically a subset of the information that the platform collects. Platforms distribute these information both on their websites and through their Application Programming Interfaces (APIs), which are customized programming protocols that allow more sophisticated investors to develop their own automated algorithms to place orders.

This provision of information has two main purposes. First, it allows investors to check the quality of the loan they purchase, 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. 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. 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.

Lending Club change in investor information set. On November 7th, 2014, Lending Club removed half of the 100 variables on borrowers' characteristics that it shared 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.

This change was unanticipated. Anecdotal evidence suggests that it was motivated by Lend-

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

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

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

ing Club's desire to "even the playing field between investors".[11] While Lending Club and other platforms regularly adjust the number of available variables, this negative change in information set is unique by its magnitude.

Role of investor screening. Investor screening plays an important role on lending platforms as it directly impacts whether a loan will eventually get funded.[12] In turn, this funding model allows the platforms to better calibrate their pre-screening intensity and loan pricing by monitoring investors' funding activities. Platforms also pre-screen and price listed loans in expectation of investors' ultimate investing decisions.[13] 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 clear the market.[14] These interest rates have evolved over time, as Figure 2 illustrates.

[Insert Figure 2]

2.3 Investor Composition

Lending platforms initially targeted retail investors only but 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 USD 25 notes that bear the credit risk of the loan. This process allows retail investors to diversify interest risks with a small portfolio size, encouraging their participation.

The past a few years have witnessed a significant upward trend of institutional and generally informationally sophisticated investor 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

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

[12] 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.

[13] 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.

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

of investors: banks and other institutional investors (55%), managed accounts (29%), and self-directed retail investors (13%).[15]

Within each segment of investors, the heterogeneity of investor sophistication has been increasing as well. For example, sophisticated third parties, such as LendingRobot, NSR Invest (serving retail investors), and Orchard (serving institutional investors) are increasingly providing investors with algorithmic tools to screen loan and automatically execute orders. A number of hedge funds have also publicly announced that they are investing on lending platforms.[16]

Within each segment, investor can therefore be further classified between two types: unsophisticated investors that buy loans passively or using self-directed rules, and sophisticated investors who actively produce information, screen loans using a proprietary screening model, and execute orders at high speed. Consistent with this trend of more informationally sophisticated investor participating, the speed for funding loans has increased dramatically, with the most popular loans being funded in seconds.[17]

3 Theoretical Framework

3.1 The Model

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 Ω of sophisticated investors, and 4) a competitive fringe of unsophisticated investors.[18] There are three dates, $t = 0, 1, 2$, with no time discount. The lending platform pre-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 further screen and subsequently buy these listed loans. Loan cash flows are realized at $t = 2$. We specify agents' objectives and strategies below.

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 π_0 ("good project") and

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

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

[17] See Figure D.1 in the appendix.

[18] 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.

$R = R_L$ with probability $1 - \pi_0$ ("bad project") at $t = 2$, where $\pi_0 \in (0, 1)$ and $R_L < I < R_H$. Following the banking literature that focuses on bank information production, we assume that applicants do not know their type ex-ante,[19] which allows us to focus on adverse selection on the investor side.

The number of loan applications x_0 is endogenous, and depends on the equilibrium price $p \geqslant I$ 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 π_0 does not depend on p as a benchmark.[20] We provide a 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 π_p of a listed loan being good at $t = 1$ before any further investor screening, where $\pi_p \in [\pi_0, 1]$. A higher π_p implies that the platform produces more information, and 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.[21] Accordingly, the number of loan

[19]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. Moreover, since our model features one risk bucket, we only need to assume that borrowers do not know their quality relative to the other loans of the same risk bucket.

[20]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. The interaction of two adverse selection problems on both the borrower and investor sides is also an open question itself in the theoretical literature. Since it is beyond the scope of this paper, we leave it for future research.

[21]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 all the $\pi_0 x_0$ good projects are screened in.

applications being screened in and listed on the platform is

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

Although we model only one risk bucket on the platform for simplicity, a recursive interpretation of our model covers platform pre-screening and subsequent allocation of listed loans into multiple risk buckets. A higher π_p implies that loan applications unfit for the safest risk bucket will be more accurately screened out. These screened-out applications are pre-screened again to see if they are fit for the second safest bucket, and again unfit loans will be more accurately screened out, and so forth.[22] Therefore, a higher π_p also implies that, empirically, the classification of loan applications into risk buckets is more predictive of loan default risks.

Platform pre-screening is costly. Since platform's pre-screening is implemented through scalable algorithms, we assume that pre-screening cost increases in π_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 $\kappa \geqslant 0$ to facilitate closed-form solutions. A higher κ implies a higher pre-screening cost.

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.[23]

Third, the platform distributes variables of listed loans to investors, which effectively determines sophisticated investors' information cost $\mu \geqslant 0$ in further screening the listed loans. Since these variables are already provided to the platform by the loan applicants, we assume that it is costless for the platform to change μ.

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 some listed loans may not be financed by investors in equilibrium.

[22] To formally model this recursive process requires at least three types of loan applications. Since it does not affect the qualitative predictions of our model, we abstract away from this aspect for theoretical clarify.

[23] 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.

Sophisticated investors. Sophisticated investors maximize expected profits. 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. An sophisticated investor has two potential advantages over an unsophisticated investor: information and speed, consistent with the discussion in Section 2 that sophisticated investors have an edge in both aspects. In Appendix C, we consider alternative specifications to illustrate that both advantages play roles in driving our results.

First, a sophisticated investor can purchase an information technology at cost μ, 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 $t = 1$.[24] We denote by ω the population of sophisticated investors who become informed, where $0 \leqslant \omega \leqslant \Omega$. When becoming informed, a sophisticated investor passes on a bad listed loan while offers a price p_i to buy a good one.

Second, sophisticated investors, when becoming informed, can screen and buy loans before uninformed investors do. This is consistent with 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 μ, she remains uninformed and is essentially identical to an unsophisticated investor in equilibrium.

Unsophisticated investors. Unsophisticated investors also maximize expected profits. Each unsophisticated investor may also buy at most one loan on the platform at $t = 1$. An unsophisticated investor 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 belief, which determines his financing decision. Especially, $p_u = 0$ suggests that unsophisticated investors cannot break even and thus do not participate on the platform. 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:

[24]For 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. 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.

Definition 1. *Given R_H, R_L, I and π_0, the sequential equilibrium is defined as a collection of $\{\pi_p, \mu, p, p_i, p_u, \omega\}$ such that:*[25]

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

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

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

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

v) 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, π_p, μ, as well as the population of informed sophisticated investors $0 \leqslant \omega \leqslant \Omega$.

Consider the pool of listed loans facing uninformed investors. Uninformed investors move after the ω informed investors screen ω listed loans and potentially finance $\pi_p\omega$ good loans. Hence, uninformed investors' updated posterior belief π'_p of a listed loan being good is

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

where $(\cdot)_+ = \max\{0, \cdot\}$. As $\pi'_p(\omega)$ is decreasing in ω, uninformed investors' posterior expected value of a listed loan in this pool is as follows and also decreases in ω:

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

[25] 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.

Since uninformed investors are competitive, they get zero profit in equilibrium. Therefore, they participate if and only if they can meet the investment requirement while still break even:

$$V'(\omega) \geqslant I\,, \tag{3.3}$$

and the price p_u they offer is

$$p_u(\omega) = \begin{cases} V'(\omega), & \text{if } V'(\omega) \geqslant I, \\ 0, & \text{if } V'(\omega) < I\,. \end{cases} \tag{3.4}$$

The fact that both $\pi_p'(\omega)$ and $V'(\omega)$ are decreasing in ω 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, buy good loans before uninformed investors can invest. Thus, informed investors' optimal offer price to an identified good loan is the loan applicant's outside option, depending on uninformed investors' participation decision:

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

where $V'(\omega)$ is determined in (3.2). Because the value of a good loan R_H is always higher than sophisticated investors' offer price $p_i(\omega)$, sophisticated investors enjoy a positive information rent. This implies that informed sophisticated investors outperform unsophisticated ones.

As long as there are good listed loans available, a sophisticated investor becomes informed if and only if the unconditional information rent exceeds the information cost set by the platform:

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

which is more (less) likely to be satisfied when the information cost μ is lower (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:

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

where $V'(\omega)$ is determined in (3.2) and thus $p(\omega)$ is a decreasing function of ω. Intuitively, unsophisticated investors are the marginal investor when they participate; otherwise, informed sophisticated investors are the marginal investor.

Having characterized investors' participation and pricing decisions in the sub-game, we 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* $0 < \underline{\kappa} \leqslant \overline{\kappa}$ *such that:*

i). If $\kappa \geqslant \overline{\kappa}$*, the platform optimally chooses a low* $\underline{\pi_p}$ *and a low* $\underline{\mu}$*. In this case, sophisticated investors become informed and they invest, while unsophisticated investors do not participate. The volume is* $\min\{\pi_0 x_0(I), \pi_p\Omega\}$.

ii). If $\kappa \leqslant \underline{\kappa}$*, the platform optimally chooses a high* $\overline{\pi_p}$ *and a high* $\overline{\mu}$*, where* $\overline{\pi_p} > \underline{\pi_p}$ *and* $\overline{\mu} > \underline{\mu}$*. In this case, sophisticated investors do not become informed, while all uninformed investors participate. The volume is* $\frac{\pi_0 x_0(p(0))}{\pi_p}$.

iii). If $\underline{\kappa} < \kappa < \overline{\kappa}$*, the platform optimally choose either* $\{\underline{\pi_p}, \underline{\mu}\}$ *or* $\{\overline{\pi_p}, \overline{\mu}\}$*, depending on which case gives the platform a higher expected payoff.*

The mathematical expressions of the thresholds $\underline{\kappa}$, $\overline{\kappa}$, and the optimal policies $\underline{\pi_p}$, $\underline{\mu}$, $\overline{\pi_p}$, and $\overline{\mu}$ are given in Appendix A.

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

PROPOSITION **1.** *As pre-screening cost* κ *goes from high to low over time, the platform optimally:*

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

b). increases investor information cost μ*, that is, distributes less information.*

Theorem 1 and in particular Proposition 1 speaks to the optimal platform design as the platform's pre-screening cost becomes lower over its life-cycle.[26] Providing less information to investors may seem surprising at first given that marketplace lending relies on investor screening.

[26]Platform pre-screening cost decreases over time because of improvements in information technologies

To better understand the results and the underlying economic trade-offs, 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 (in Appendix A) 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:

Equilibrium Investor Participation and Platform Volume

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

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

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

According to Theorem 1, a type-2 (type-4) equilibrium, where both the optimal π_p and μ 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 κ is high. This implies that the platform optimally chooses a low pre-screening intensity π_p to save its cost. Because the resulting quality of an average listed loan is low, unsophisticated investors cannot break even and thus will not participate according to condition (3.3). If the platform chose a high μ, that is, provided little information to sophisticated investors, sophisticated investors would 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 and the volume is accordingly 0. To improve screening efficiency, the platform optimally chooses a low μ 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 $p = I$.

and screening algorithms. See Lending Club Investor Day 2017 Presentation available at ir.Lending Club.com/Cache/1001230258.pdf.

They identify and buy good loans up to their capacity, helping the platform boost the volume to either $\pi_0 x_0(I)$ or $\pi_p \Omega$, whichever is larger.[27] The economy ends up in a type-2 equilibrium, where the screening efficiency concern dominates.

Next, we consider case ii) in Theorem 1 when κ is low. By condition (3.3), the platform optimally chooses a high pre-screening intensity π_p to attract uninformed, unsophisticated investors. Accordingly, uninformed investors become the marginal investor. If the platform chose a low μ, that is, provided much information to sophisticated investors, they would become informed and thus buy good loans earlier than uninformed investors. Notice that this strategy would introduce adverse selection: 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 $x_0(p(\Omega))$, driving the economy into a type-3 equilibrium and hurting volume. If adverse selection were severe enough, it would deter uninformed investors from participating at all. Hence, to eliminate adverse selection, the platform optimally chooses a high μ so that sophisticated investors do not become informed, according to condition (3.6). In this case, all the investors are uninformed, 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. The economy ends up in a type-4 equilibrium, where the adverse selection concern dominates.

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.[28]

[27] 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 $\min\{\pi_0 x_0(I), \pi_p \Omega\}$.

[28] 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 ω.

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. They speak to our first two research questions of sophisticated investor information production and its response to platform policies. Specifically, the two predictions come from conditions (3.5) and (3.6) of the model. Although they are theoretically straightforward, it is not obvious empirically whether there is any room for sophisticated investors to screen and outperform, and if so, to what extent the outperformance 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, empirically observing 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.*

Predictions 3 and 4 focus on optimal platform design and derive from the full equilibrium. They stem from Theorem 1 and Proposition 1, and speak to our third research question regarding the optimally joint information production. 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 π_p only hurts volume.

4 Data and Investor Sophistication

4.1 Data

Our investor-level data is provided by LendingRobot, a leading robo-advisor for retail investors on lending marketplaces.[29] 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, Lending Club and Prosper, as well as all historic transactions from portfolios monitored by LendingRobot. We combine this data with the loan-level data from Lending Club and Prosper, which is publicly available.

LendingRobot provides an automated investment tool for its clients which relies on a sophisticated screening model calibrated on historical data from the platforms.[30] This tool also allows to execute orders at high speed through an API. Independently, LendingRobot also offers a free monitoring tool that helps investors monitor they portfolios on Lending Club and Prosper without automated investing. Thus, we observe both portfolios invested under the help of LendingRobot's investment tool, as well as portfolios built by investors themselves.[31]

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.
- **Account**: A user can have one or several accounts. An account represents a portfolio of notes that an investor holds on a single lending platform: Lending Club or Prosper. There are three types of accounts in our data. A *monitor-only* account is an account where LendingRobot only monitors the portfolio but do not screen or buy notes through its technology. In a *robot* account, notes screening and purchase are executed by the LendingRobot investment tool. In an *advanced* account, investors can further implement their own screening criteria, combined or not with LendingRobot's screening model, and rely on LendingRobot to automatically execute the orders when relevant loans appear on

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

[30] For more details on LendingRobot screening model, refer to http://blog.lendingrobot.com/research/predicting-the-number-of-payments-in-peer-lending/.

[31] Investors who use the monitoring tool only may rely on Lending Club's and Prosper's passive investment feature, or on their own discretion.

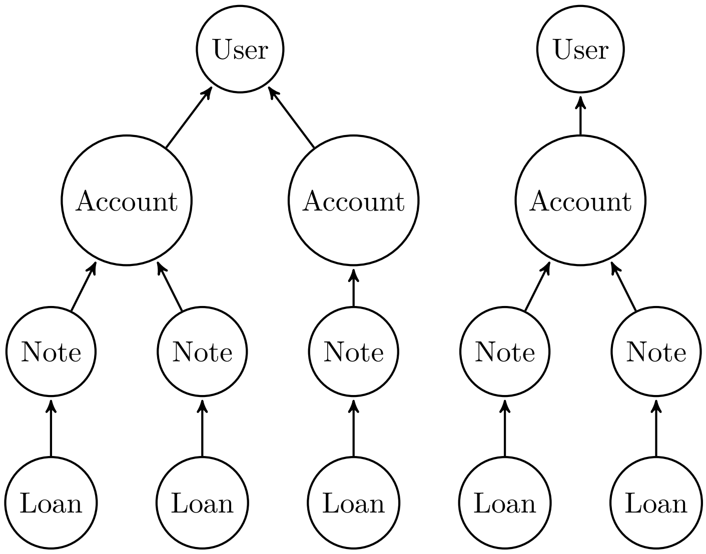


Figure 1: LendingRobot Data Structure

the platform.

- **Note**: Retail investors invest in notes, each of which is backed by a single loan. 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 standard in the literature. We observe these variables for all the loans issued by Lending Club and Prosper.

4.2 Investor Sophistication

Our data provides us with portfolios of retail investors with heterogenous levels of sophistication, which is key to our study.

Within our data, *monitor-only* accounts are the less sophisticated ones as they do not implement LendingRobot's investment tool with information and speed advantages.[32] We also

[32]Monitored-only investors are selected on having registered to LendingRobot, which might suggest a higher sophistication than fully naive investors. However, this potential selection effect can only bias against finding

consider the average retail investor on a given platform, which we can observe by collapsing the whole platform borrowers' data, as unsophisticated.

Robot accounts represent sophisticated investors, as they invest through the LendingRobot investment tool which provides both screening and automated execution at high speed.

Advanced accounts also represent sophisticated investors. While some *advanced* accounts might be potentially even more sophisticated than *robot* accounts, *advanced* accounts are however at risk of making mistakes compared to the *robot* benchmark, as they rely at least partly on their own screening criteria, which are automatically executed by LendingRobot.[33] For the purpose of our analysis, we consider *advanced* accounts as a different type of sophisticated investors without specifying whether they are more or less sophisticated than *robot* accounts.

While our data focuses on the segment of retail investors, the heterogeneity in sophistication within this segment is arguably comparable to the heterogeneity of sophistication within the institutional investor segment. As discussed in the institutional background, traditional mutual and pension funds usually buy loans on the platforms passively while hedge funds screen loans actively before investing. In addition, LendingRobot indeed performs hedge-fund-like order execution for its *robot* and *advanced* accounts.

4.3 Summary Statistics

Table 1 provides the summary statistics. Our sample is representative: the difference between Lending Club and Prosper accounts on LendingRobot, in terms of both the number of accounts and the amount invested, is comparable to the difference between the two platforms overall.

On portfolio size, while *robot* accounts are the most representative 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. Portfolio size should not be interpreted as a proxy for sophistication, however, as any *robot* account invests automatically along the same criteria, independently of its size.

differences between portfolios of monitored-only and robot investors in our empirical analysis. Such a bias would only weaken but not strengthen our outperformance results.

[33]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.

On portfolio risk, 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.

[Insert Table 1]

5 Empirical Analysis

5.1 Investor Screening

Our theoretical analysis suggests that informationally sophisticated investors may 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 restrict our analysis to fractional loans on these platforms, that is, loans that are offered to retail investors only.

We define three loan-level indicator variables, $RobotAccount_i$, $AdvancedAccount_i$, $MonitoredOnlyAccount_i$, equal to one if at least two *robot*, *advanced* or *advanced* accounts invested in a given loan i. We use these variables as left-hand-side variables, and run linear probability regressions on borrower characteristics.[34] We include interest rate fixed effects, and therefore identify screening that occurs among loans with the same price. Such screening aims at picking the loans with lower default risk, keeping interest rate constant.[35]

$$Prob(TypeAccount_i = 1) = \beta \times BorrowerCharacteristics + \sum_{j=1}^{n} IR_j + \sum_{t=1}^{T} m_t + \epsilon_i, \qquad (5.1)$$

where $BorrowerCharacteristics$ are provided by the lending platform, IR_j are fixed effects for the n distinct levels of interest rate paid to investors (i.e. the loan price at issuance) observed in our data, m_t are month of issuance fixed effects, and ϵ_i is the error term. Note that loan maturity (either 3 or 5 years) is implicitly controlled for by interest rate fixed effects, as within the same risk bucket interest rate varies with maturity only. Table 2 displays the results.

[34] The results are similar when conducting logit regressions.

[35] We abstract from the question as to whether loans with a given interest rate are on average fairly priced at a given time.

[Insert Table 2]

There are two key takeaways from this analysis. First, a large number of specific borrower characteristics strongly predict an investment by *robot* accounts, suggesting that *robot* accounts indeed actively screen loans within a risk bucket.[36] This also suggests that LendingRobot's screening model considers the risk associated with these characteristics to be mis-estimated or not fully incorporated by the platform when listing the loans. 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 coefficient on loan amount or revolving utilization in column 1 suggests that LendingRobot's screening model considers that Lending Club excessively penalizes borrowers that take a large loan or have large revolving balance.[37]

Second, screening behavior varies by the type of investors. First, although *advanced* accounts screen differently from *robot* accounts, their screening criteria are largely consistent with each other: the coefficients in column 2 are largely comparable to the ones in column 1. This again suggests that both types of accounts represent sophisticated investors. Consistently with our view that *monitored-only* accounts are less sophisticated, they appear to screen significantly differently from *robot* accounts as well as from *advanced* accounts do.

We note that the observed correlation between investment decisions and borrower characteristics is not necessarily causal. It might result from the loans with some characteristics having been first picked up and financed by more sophisticated investors, and then less sophisticated investors invest in loans with the opposite characteristics. For instance, if more sophisticated investors massively invest in loans with high FICO scores within a risk bucket, less sophisticated investors will be mechanically more likely to invest in loans with a relatively lower FICO score. This interpretation is consistent with the view that *monitored-only* accounts represent the less

[36] We note that this is different from the usual perception that robot-advisors are a form of passive investors.

[37] 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.

sophisticated segment of investors who eventually pick up the loans being passed on by more sophisticated investors.

In Table D.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 regression coefficients.

5.2 Investor Performance

Having documented that more sophisticated investors actively screen loans and do so differently from less sophisticated investors, we test whether active screening allows more sophisticated investors to systematically outperform average and unsophisticated investors on the platform. Specifically, we test Prediction 1 by investigating whether, controlling for interest rates, having sophisticated investors participate in a loan predicts a lower likelihood of default.

Following the literature, we measure performance at the loan level, by using an indicator variable for the loan being in default or charged-off. A loan enters default status when it is more than 121 days past due and enters charge-off status after 150 days.[38]

We first plot the share of defaulted loans by sub-grades as of December 2017 for every fractional loan on the Lending Club platform, as well as for the subsets of loans in which at least two *robot*, *advanced* and *monitor-only* accounts invested. The results in Figure 3 suggest systematic outperformance achieved by *robot* and *advanced* accounts over *monitor-only* accounts and the average Lending Club loans, across the whole spectrum of loan risk.

[Insert Figure 3]

We conduct regressions to more precisely explore investor outperformance. We regress the performance indicator on indicator variables for participation by different types of investors. We use OLS specifications due to the many fixed effects and to help interpret the economic magnitude of the coefficients.[39]

[38]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.

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

Specifically, we run the following specification:

$$Prob(ChargedOff = 1)_i = \beta_1 \times \mathbb{1}_{TypeAccount} + \sum_{j=1}^{n} IR_j + \sum_{t=1}^{T} m_t + \epsilon_i, \qquad (5.2)$$

where $ChargedOff$ is an indicator variable for the loan to be in default or charged-off, $\mathbb{1}_{TypeAccount}$ is an indicator variable equal to one if at least two accounts of a given type are invested in the loan, and IR_j and m_t are the same interest rate and month fixed effects as those in specification (5.1). Table 3 shows the results. *Robot* accounts are shown in column 1, *advanced* accounts in column 2, and *monitored-only* accounts in column 3. The following columns interact $\mathbb{1}_{TypeAccount}$ with year fixed effects, and with loan grade fixed effects.

[Insert Table 3]

In Table 3, column 1 documents that the default rate of loans in which *robot* accounts invest is significantly lower than that for the whole Lending Club loan population over the 2014-2016 period, controlling for interest rates. This outperformance is measured within each sub-grade, and therefore cannot result from any compositional effect across the sub-grades. The economic magnitudes are particularly large: the OLS specification suggests a reduction by more than 3% of the default rate, comparing to a 14% average default rate for the Lending Club loan population. This outperformance therefore translates into a reduction in average default rate by more than 20%. Column 2 shows that over the full sample period, *advanced* accounts' outperformance is slightly larger than that of *robot* accounts. By contrast, column 3 shows that participation by *monitored-only* accounts only weakly predicts a lower default rate with an economic magnitude four times smaller, consistent with *monitored-only* accounts being less sophisticated.

Moreover, column 7 and 8 illustrate that the reduction in default rate obtained by *robot* and *advanced* accounts is broadly increasing in loan risk, but this larger outperformance is partly driven by riskier loans having a higher average default rate.

This analysis empirically establishes that informationally sophisticated investors systematically outperform less sophisticated ones. It suggests that sophisticated investors produce additional information about the underlying loans. It also suggests that platform pre-screening leaves significant room for further investor screening, representing a deviation from the tradi-

tional banking model where banks screen loans exclusively on behalf of depositors.

5.3 Screening Cost and Investor Performance

Given the outperformance of sophisticated investors, we follow Predictions 2 and 3 to investigate the relationship between platform information provision and investor screening performance.

For this purpose, we exploit the Lending Club 2014 event, that is, the unexpected shock to investor information set described in Section 2.2 with a difference-in-differences setting. While the shock affects all investors, it affects investors differentially, as unsophisticated investors are less likely to use the removed variables than sophisticated investors do.

We run the following specification:

$$\begin{aligned} Prob(ChargedOff = 1)_i = {} & \beta_1 \times \mathbb{1}_{robot} + \beta_2 \times \mathbb{1}_{robot} \times Post \\ & + \beta_3 \times \mathbb{1}_{advanced} + \beta_4 \times \mathbb{1}_{advanced} \times Post \\ & + \beta_5 \times \mathbb{1}_{monitor} + \beta_6 \times \mathbb{1}_{monitor} \times Post + \sum_{j=1}^{n} IR_j + \sum_{t=1}^{T} m_t + \epsilon_i \end{aligned} \tag{5.3}$$

where $\mathbb{1}_{robot}$ is an indicator variable equal to one if at least two *robot* accounts are invested in the loan i, $\mathbb{1}_{advanced}$ is an indicator variable equal to one if at least two *advanced* accounts are invested in the loan, $\mathbb{1}_{monitor}$ is an indicator variable equal to one if at least two *monitor-only* accounts are invested in the loan, $Post$ is an indicator variable for being in the period after the shock to the information set, and IR_j and m_t are the same interest rate and month fixed effects as those in specification (5.1).

We find empirical evidence consistent with Prediction 2 that when investors' screening cost increases, sophisticated investor outperformance is reduced. The results are displayed in Table 4. Column 1 implements this specification for all Lending Club loans issued in the period spanning three months before the event month and three months after. Column 2 restrict the loan sample to loans with grade C to G. Column 3 restricts the period to two months before the event month 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.

[Insert Table 4]

This analysis reveals that the increase in investor screening cost significantly impacts the screening performance of *robot* accounts, as well as that of *advanced* accounts to a smaller extent. On the other hand, *monitor-only* accounts are not affected at all. 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* accounts in the period immediately preceding the shock drops by more than a half, compared to that at the time of the shock.[40] When comparing the effect between *robot* and *advanced* accounts, we observe that *advanced* accounts are less affected by the shock. But note that this smaller effect for *advanced* accounts is partly driven by *advanced* accounts' relatively lower outperformance before the time of the shock.

To further pin-down the causal impact of the information shock on sophisticated investor 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 platform pre-screening intensity, may drive our result.

For this purpose, we implement two regressions with two different samples: all fractional Lending Club loans, the control group, and all the Lending Club loans in which at least two *robot* accounts invested, 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 still control for interest rate fixed effects to alleviate concerns over potential compositional changes during the sample period. The reference period for the month fixed effects is month 5 and 6 after the event month.

Figure 4 shows the result, which illustrates that the overall fractional loan performance on the Lending Club platform is unaffected by the shock, while the performance of the loans invested by *robot* accounts sharply deteriorates at the time of the shock.[41] The sharpness in the change of performance, as well as its synchronicity with the time of the shock, are supportive of

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

[41] 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.

a causal interpretation.

[Insert Figure 4]

We ensure that our results do not come from a possibly sharp change in the composition of listed loans that may not be captured by the interest rate fixed effects. For this purpose, Table D.2 in the appendix illustrates that the pool of listed loans indeed remains unchanged after the shock in terms of major observable characteristics of the borrowers.

Finally, consistent with Prediction 3, this difference-in-differences analysis provides an economic rationale for Lending Club's reduce information provision to investors. It is consistent with the platform "evening the playing field" to actively manage the potential adverse selection problem introduced by sophisticated investors.[42]

5.4 Trends in Platform Pre-screening Intensity and Investor Performance

According to Prediction 4, we study the time-series evolution of platform pre-screening and investor relative performance as a final step of our empirical analysis.

5.4.1 Platform Pre-screening

Based on the recursive interpretation of platform pre-screening in our model, a higher pre-screening intensity π_p means good (bad) loans are more precisely screened-in (-out) in any risk bucket, that is, a higher accuracy of the classification of loan applications into different risk buckets. We thus explore how this accuracy, the empirical counterpart of π_p, evolves over time.

For this purpose, we follow the literature (Iyer, Khwaja, Luttmer and Shue, 2015, for example) to build receiver operating characteristic (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 is, the more precise the predictor is. The test is computed separately for loans issued and graded in 2014, 2015 and 2016 to understand how platform pre-screening evolves over time.

[42] An alternative explanation for the reduction in the information set would be that it reduces noise, making it easier for certain investors, typically the unsophisticated ones, to pick good loans. However, such an hypothesis would predict that these investors are better-off after the change, which we do not observe in our data.

Results are displayed in Figure 5. This figure documents how Lending Club sub-grades and Prosper grades have been improving explanatory power for 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 grades, as the areas below the ROC curves are larger for Lending Club.[43]

[Insert Figure 5]

As a complementary test, we also explore whether the platforms have been changing the risk thresholds for a loan application to get listed (before allocating them into a risk bucket). We plot the evolution of the share of borrowers on Lending Club whose FICO score is below 670 and 660, and whose debt-to-income ratio is above 30% and 35%. Results are displayed in Figure 6, showing an increasing share of ex-ante high-risk loans being listed. This figure is consistent with the overall increase in loan prices on Lending Club from Figure 2. Listing an increasing share of ex-ante high-risk or short-credit-history applications, while still improving the accuracy of their risk classification, further supports the hypothesis that platforms are increasing their pre-screening intensity.

[Insert Figure 6]

Overall, these two figures illustrate that platforms improve over time the accuracy of the classification within listed loans, which corresponds to a higher but still intermediate level of platform information production. These trends are consistent with Prediction 4.

5.4.2 Evolution of Investor Screening Performance over 2014-2016

We finally explore the evolution of investor screening performance over our sample period. 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 in which at least two *robot*, *advanced* and *monitor-only* accounts invested. Figure 7 shows that the outperformance of *robot* and *advanced* investors appear to decrease over time.[44]

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

[44] 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.

[Insert Figure 7]

Using regressions, columns 4 to 6 of Table 3 also explore the time-series of outperformance by investor type, by interacting the indicator of having a given type of accounts participating in a loan with year fixed effects. For both *robot* and *advanced* accounts, outperformance is significantly stronger in 2014 than in 2015, and even more so than in 2016. Between the two more sophisticated types, *robot* accounts outperform *advanced* accounts in 2014, but the pattern get reversed in the two following years. Consistently with *monitor-only* accounts being less sophisticated, we find no time trend for them.

We view these patterns consistent with the platforms actively managing the potential adverse selection problem introduced by more sophisticated investors. We also note that, while the Lending Club information shock we investigate above plays an important role in the evolution between 2014 and 2015, it cannot account for the whole trend, which likely results as well from the improvement in pre-screening by the platforms we previously document.

6 Conclusion

Different from the conventional banking paradigm, one prominent feature of the burgeoning marketplace lending is the joint information production by the platform and investors. In maximizing loan volume, the platform trades off the improving screening outcome and an adverse selection problem, both resulted from the participation of informationally sophisticated investors. Thus, intermediate levels of platform pre-screening intensity and information provision to investors are optimal. Using novel investor-level data, we empirically show that 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, and shed light on the optimal distribution of information production between the platform and investors.

References

Adrian, Tobias and Adam Ashcraft (2016). "Shadow Banking: A Review of the Literature." in *Banking Crises*, 282-315. Palgrave Macmillan UK.

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

Biais, Bruno, Thierry Foucault and Sophie Moinas (2015). "Equilibrium Fast Trading." *Journal of Financial Economics*, 116: 292-313.

Bolton, Patrick, Tano Santos and Jose Scheinkman (2016). "Cream Skimming in Financial Markets." *Journal of Finance*, 71: 709-736.

Buchak, Greg, Gregor Matvos, Tomasz Piskorski, and Amit Seru (2017). "Fintech, Regulatory Arbitrage, and the Rise of Shadow Banks." *Journal of Financial Economics*, forthcoming.

Dang, Tri Vi, Gary Gorton, Bengt Holmstrom, Guillermo Ordonez (2017). "Banks as Secret Keepers." *American Economic Review*, 107: 1005-1029.

de Roure, Calebe, Loriana Pelizzon, Anjan Thakor (2018). "P2P Lenders versus Banks: Cream Skimming or Bottom Fishing?" Working Paper.

Duarte, Jefferson, Stephan Siegel, and Lance Young (2012). "Trust and Credit: the Role of Appearance in Peer-to-Peer Lending." *Review of Financial Studies*, 25: 2455-2484.

Fishman, Michael and Jonathan Parker (2015). "Valuation, Adverse Selection, and Market Collapses." *Review of Financial Studies*, 28: 2575-2607.

Franks, Julian, Nicolas Serrano-Velarde, and Oren Sussman (2016). "Marketplace Lending, Information Aggregation, and Liquidity." Working Paper.

Fuster, Andreas, Matthew Plosser, Philipp Schnabl, and James Vickery (2018). "The Role of Technology in Mortgage Lending." Working Paper.

Glode, Vincent, Richard Green and Richard Lowery (2012). "Financial Expertise as an Arms Race." *Journal of Finance*, 67: 1723-1759.

Gorton, Gary and George Pennacchi (1990). "Financial Intermediaries and Liquidity Creation." *Journal of Finance*, 45: 49-71.

Hauswald, Robert and Robert Marquez (2003). "Information Technology and Financial Services Competition." *Review of Financial Studies*, 16: 921-948.

Hauswald, Robert and Robert Marquez (2006). "Competition and Strategic Information Acquisition in Credit Markets." *Review of Financial Studies*, 19: 967-1000.

Hildebrand, Thomas, Manju Puri, and Jorg Rocholl (2016). "Adverse Incentives in Crowdfunding." *Management Science*, 63: 587-608.

Iyer, Rajkamal, Asim Khwaja, Erzo Luttmer, and Kelly Shue (2015). "Screening Peers Softly: Inferring the Quality of Small Borrowers." *Management Science*, 62: 1554-1577.

Keys, Benjamin, Tanmoy Mukherjee, Amit Seru, and Vikrant Vig (2010). "Did Securitization Lead to Lax Screening? Evidence from Subprime Loans." *Quarterly Journal of Economics*, 125: 307-362.

Liskovich, Inessa and Maya Shaton (2017). "Borrowers in Search of Feedback: Evidence from Consumer Credit Markets." Working Paper.

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

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

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

Tang, Huan (2018). "Peer-to-Peer Lenders versus Banks: Substitutes or Complements?" Working Paper.

Yang, Ming and Yao Zeng (2017). "Financing Entrepreneurial Production: Security Design with Flexible Information Acquisition." *Review of Financial Studies*, forthcoming.

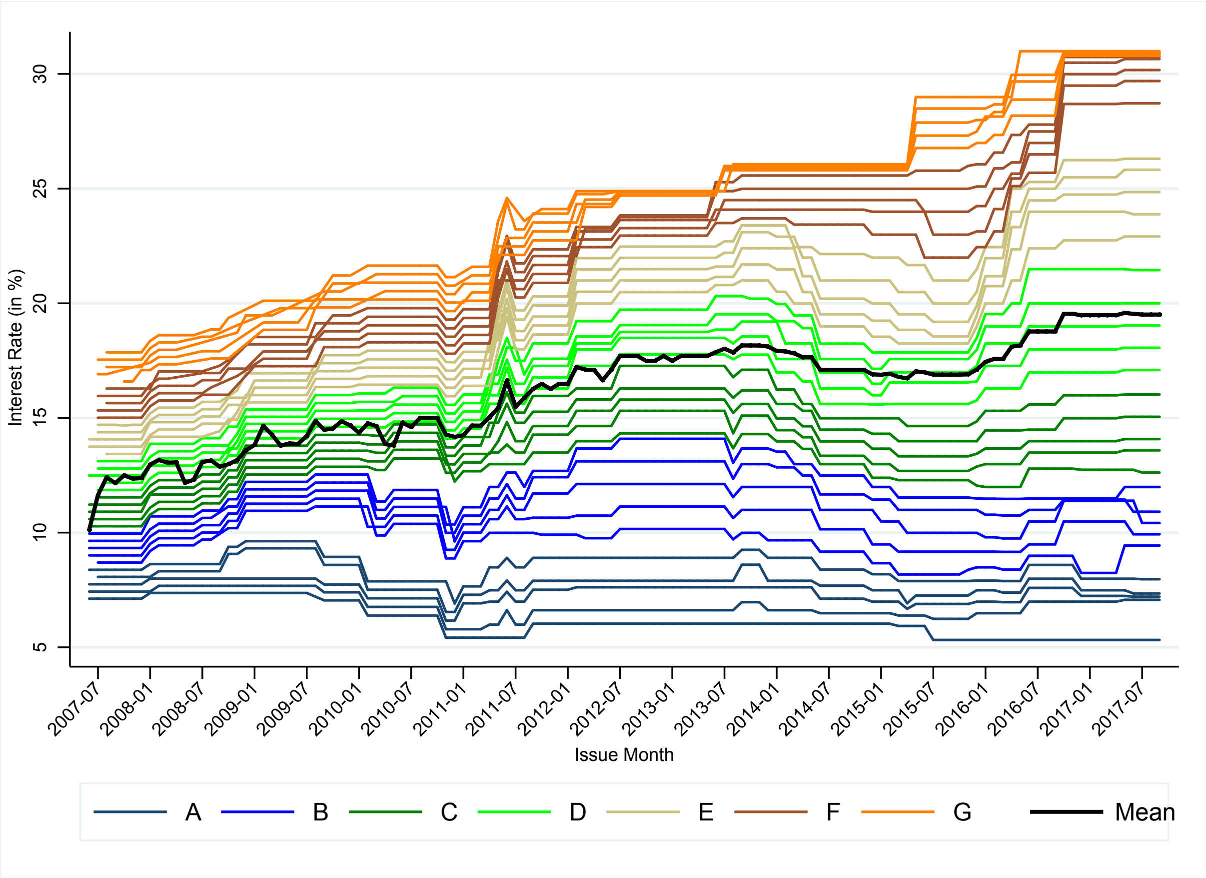


Figure 2: Evolution of Interest Rates on Lending Club by Sub-Grades

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

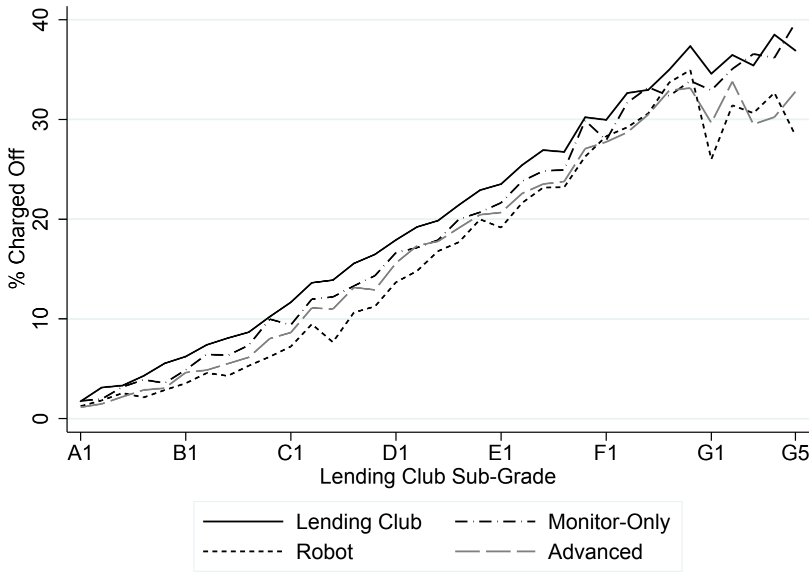
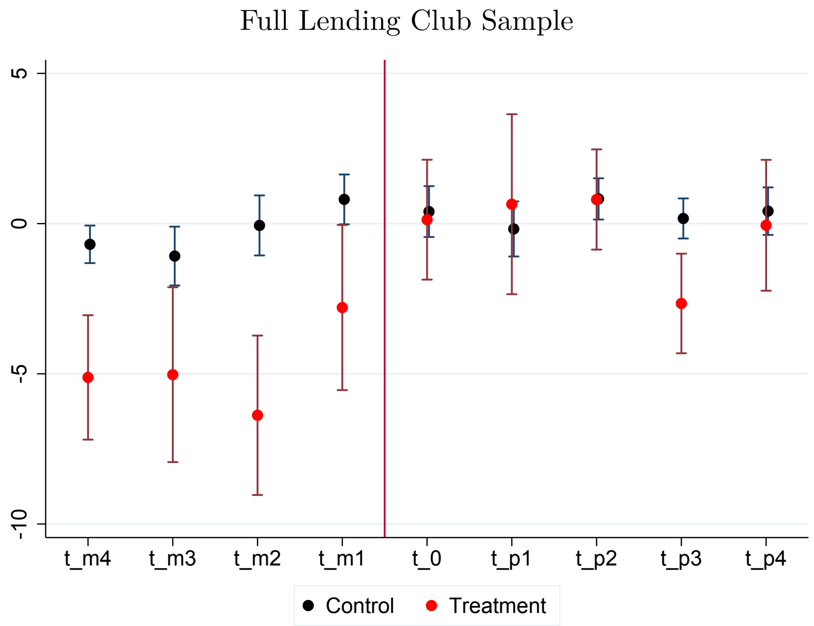


Figure 3: Charged Off Loans

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.



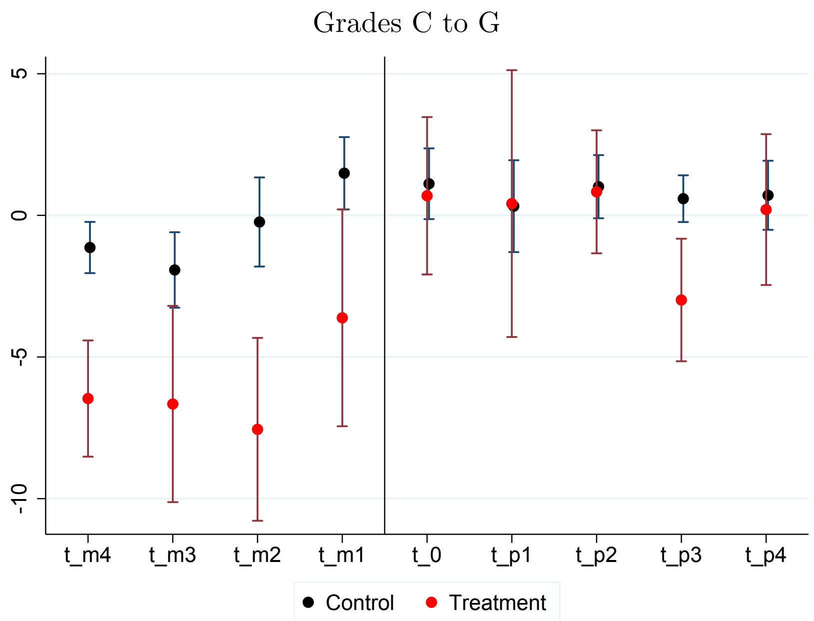
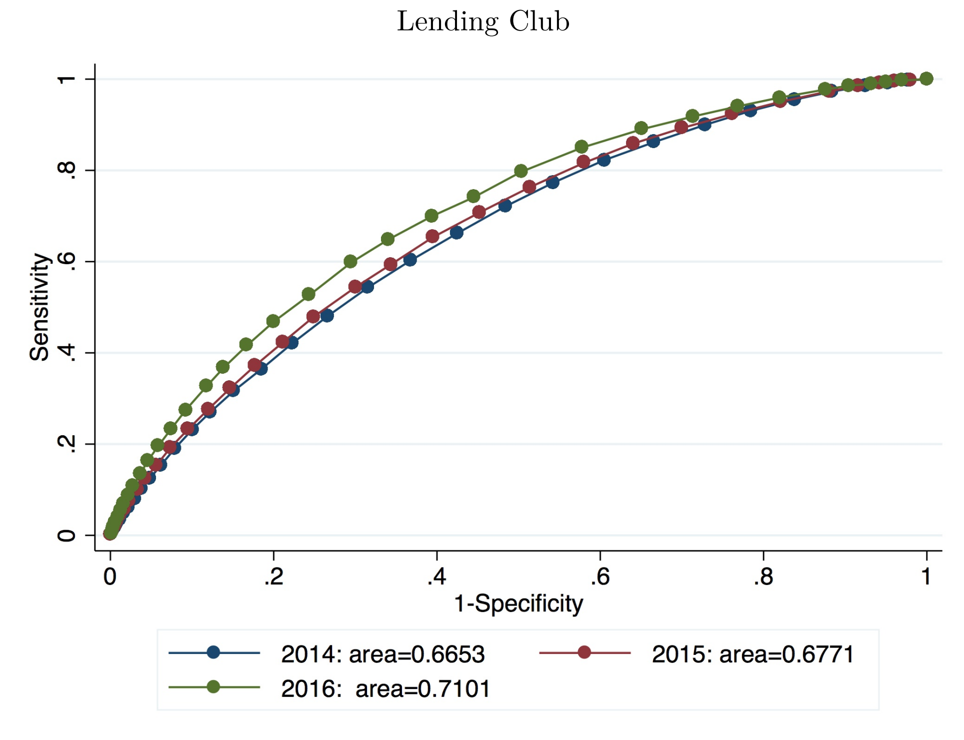


Figure 4: Change to Investor Screening Cost: Difference-in-differences analysis

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 to G.



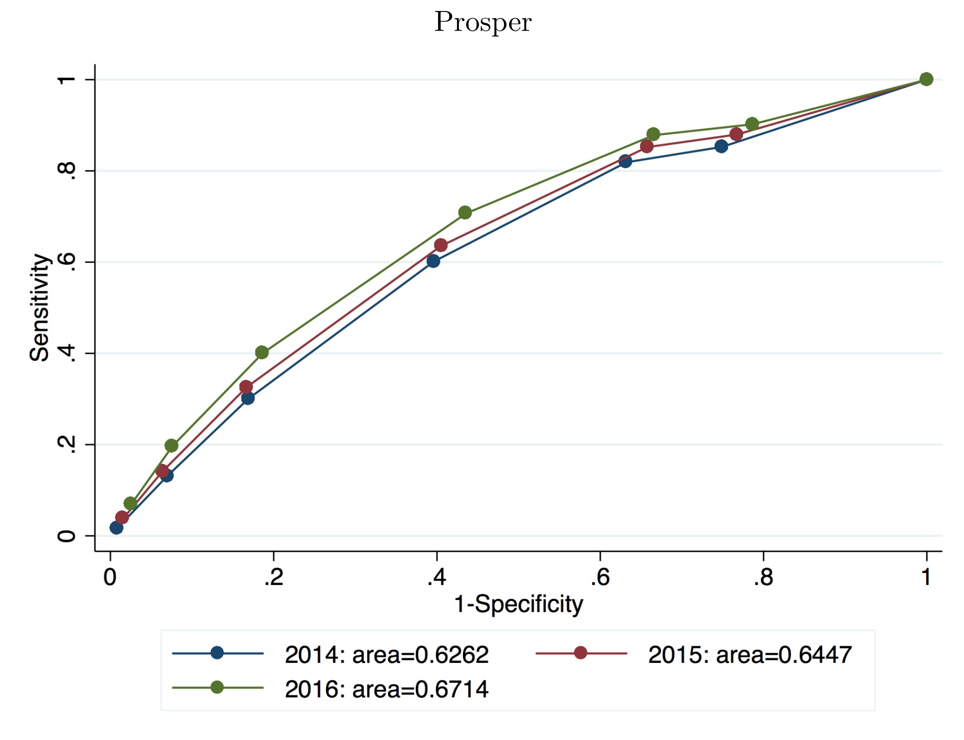


Figure 5: Platform Pre-Screening: ROC Curve of Lending Club Sub-grades and Prosper Grades on Charged-Off

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.

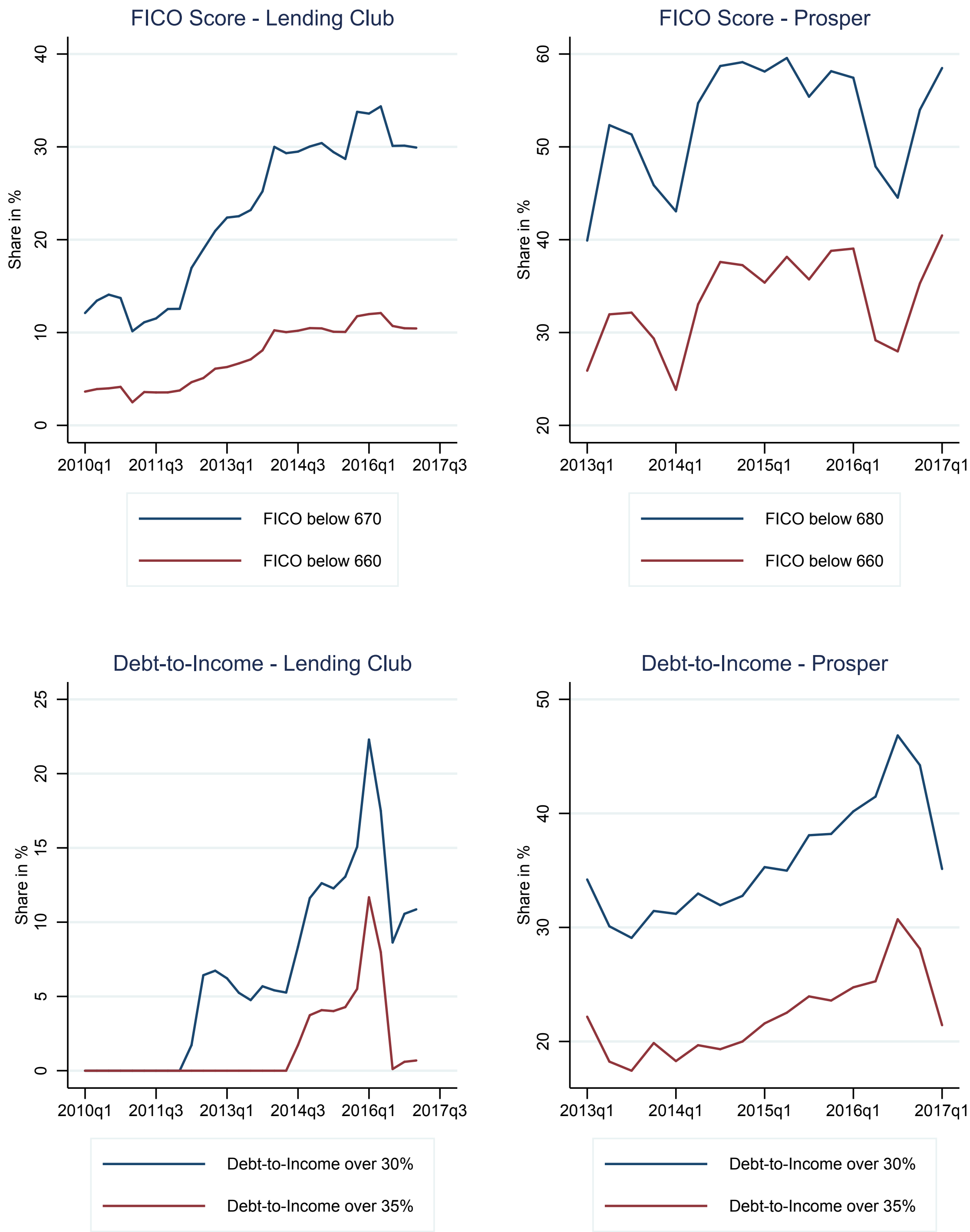


Figure 6: Platform Pre-Screening: Evolution of the Share of Borrowers above FICO and Debt-to-Income Thresholds

Note: This figure plots the evolution of the share of borrowers on the Lending Club platform whose FICO score is below 670 (680) and 660, and whose debt-to-income ration is above 30% and 35%, for both Lending Club and Prosper.

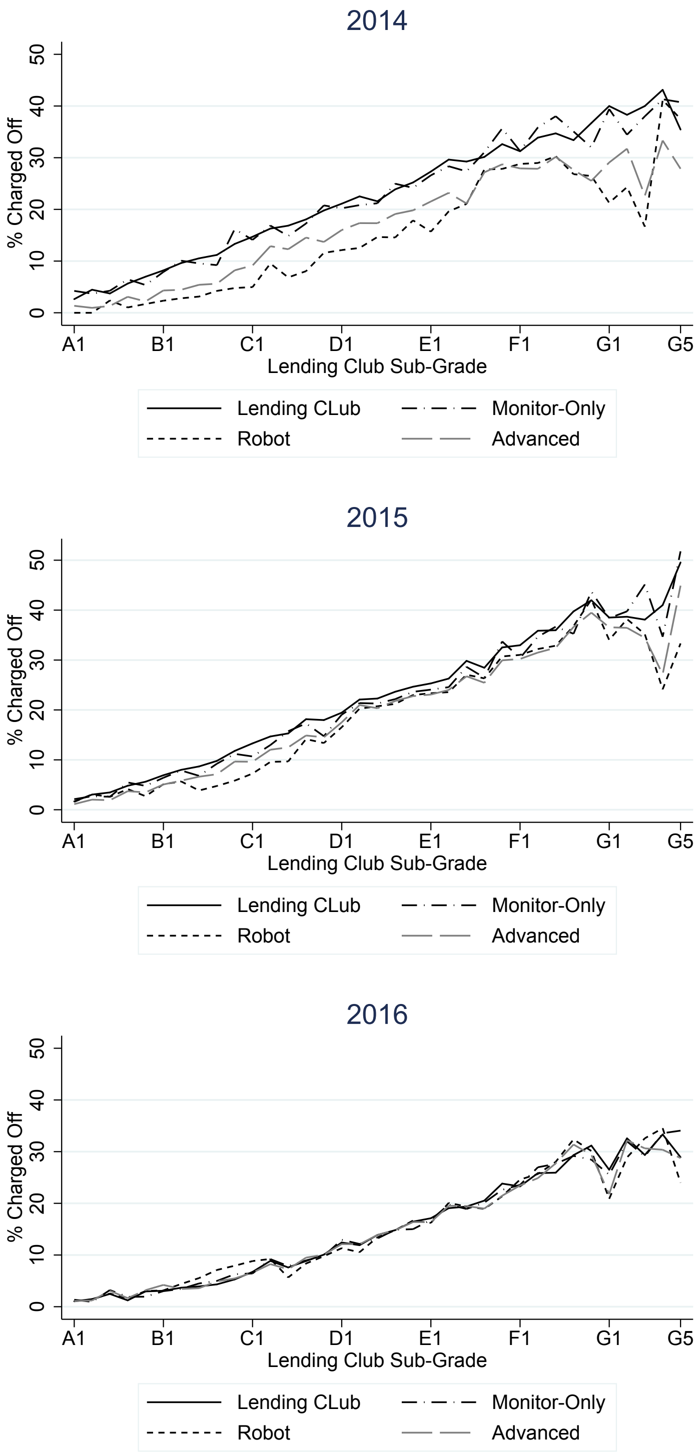


Figure 7: Charged Off Loans

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

Table 1: Summary Statistics

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

Note: This table provides summary statistics on the proprietary dataset used for the empirical analysis. The data covers all transactions executed by LendingRobot users between January 2014 and February 2017, which represents more than $120 million invested on the two major lending platforms, Lending Club 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

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

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

Table 3: Screening Performance

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

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

Table 4: Difference in Difference Analysis

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

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

Appendix

A 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 κ 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 (π_p, μ)-space denoted by $S = \{(\pi_p, \mu, \cdot) | \pi_0 \leqslant \pi_p \leqslant 1, \mu \geqslant 0\}$.

In a type-1 sub-game equilibrium, none of the investors participates. This implies that $\omega = 0$ and condition (3.5) further implies that $p_i(0) = I$. Thus, conditions (3.3) and (3.6) give that

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

In a type-2 sub-game equilibrium, sophisticated investors become informed but uninformed investors do not participate. This implies that $\omega = \min\{\Omega, x_p\}$ and condition (3.5) further implies that $p_i(\min\{\Omega, x_p\}) = I$. Thus, conditions (3.3) and (3.6) give that

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

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

In a type-3 sub-game equilibrium, sophisticated investors become informed and uninformed investors also participate. This implies that $\omega = \min\{\Omega, x_p\}$ and condition (3.5) further implies that $p_i(\min\{\Omega, x_p\}) = V'(\min\{\Omega, x_p\})$. Thus, conditions (3.3) and (3.6) give that

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

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

In a type-4 sub-game equilibrium, sophisticated investors do not become informed but all the uninformed investors participate. This implies that $\omega = 0$ and condition (3.5) further implies that $p_i(0) = V'(0)$. Thus, conditions (3.3) and (3.6) give that

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

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

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

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

Similarly, suppose $(\pi_p, \mu, \cdot) \in SE_3$, that is, the full equilibrium admits a type-3 sub-game

equilibrium as characterized in (A.3). The equilibrium volume is $\frac{\pi_0 x_0(p(\Omega))}{\pi_p}$.

Consider an alternative equilibrium profile $(\pi_p, \mu', \cdot)$, $\mu' > \mu$ such that $\pi_p(R_H - V'(0)) < \mu'$. Notice that μ' must exist because $0 < \pi_0 \leqslant \pi_p$ and $I \leqslant V'(0) < R_H$. Again because $V'(0) \geqslant V'(\Omega)$, we have that $(\pi_p, \mu', \cdot) \in SE_4$ as characterized in (A.4). Because changing μ is costless for the platform and $x_0(p(0)) > x_0(p(\Omega))$, this implies that the platform will then find it profitable to deviate to a type-4 sub-game equilibrium by increasing μ to μ' without changing π_p, enjoying a higher volume $\frac{\pi_0 x_0(p(0))}{\pi_p} > \frac{\pi_0 x_0(p(\Omega))}{\pi_p}$. Hence, 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

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

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

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

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

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

Hence, if

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

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

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

in which case we define

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

Otherwise if

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

a type-4 sub-game equilibrium is achievable by choosing $\overline{\pi_p}$ such that it solves

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

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

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

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

B The Supply Curve of Loan Applications

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

Keeping the initial setting in the baseline model, we further assume that a loan applicant, when successfully establishes a loan from a lending marketplace, may incur an adoption cost.[45,46] This adoption cost reflects the fact that marketplace lending is still relatively new, and thus a loan applicant may incur physical or cognitive costs to understand and follow its rules and trust them. Different applicants may have different adoption costs, and it is natural that some applicants, who are already familiar with marketplace lending, may not incur any adoption cost.

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

[45] If a loan applicant goes to the platform, applies, but does not get funded, there will be no adoption cost.

[46] This notion of adoption cost follows the classic Hotelling model and has been commonly used in the theoretical literature of bank competition. For example, Hauswald and Marquez (2006) use it to model that different banks may specialize in different borrower characteristics. In a marketplace lending context similar to ours, de Roure, Pelizzon and Thakor (2018) also 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, to join a lending platform.

positive adoption cost z is $q(z) \geqslant 0$, where $z \in (0, +\infty)$ and follows an arbitrary distribution $F(z)$. Consistent with our baseline model, the loan applicants do not know their type ex-ante. Following the literature (Hauswald and Marquez, 2006, de Roure, Pelizzon and Thakor, 2018, for example), 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 the investment requirement I plus the adoption cost of that 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 Alternative Specifications of Investor Sophistication

Without fully solving the model, this appendix discusses three alternative specifications of investor sophistication. The aim is to qualitatively illustrate that in our baseline model, both information and speed advantages are important in driving sophisticated investor outperformance and adverse selection among investors. We keep other model setting and assumptions unchanged.

Alternative specification 1. Both types of investors are always uninformed, but sophisticated investors can acquire a technology to become faster in the sense that it allows them to fund a listed loan before other investors do so. Because faster investors are uninformed, they fund a listed loan if and only if $\pi_p R_H + (1 - \pi_p) R_L \geqslant I$. In addition, because faster investors do not screen loans, slower investors' interim posterior belief of a listed loan being good is still π_p. It follows that slower investors use the same funding criterion as faster investors. Therefore, faster investors cannot outperform slower ones and sophisticated investors never choose to become faster. Subsequently, there is no adverse selection among investors.

Alternative specification 2. Both types of investors are always informed about a listed

loan facing them, but sophisticated investors can acquire a technology to become faster. Because all investors are informed, they use the same criterion to screen listed loans, always offer the same price I to an identified good loan and pass on an identified bad loan, regardless of their speed. Thus, bad loans are never funded, and conditional on an investment on the platform, faster investors cannot outperform slower ones.[47]

Alternative specification 3. Both types of investors are equally fast, but sophisticated investors can acquire a technology to become informed about a listed loan. An informed investor offers price I to an identified good loan and passes on an identified bad loan. However, because uninformed investors are equally fast, an identified bad loan (passed on by an informed investor) will never reach them. Rather, they fund a listed loan if and only if $\pi_p R_H + (1 - \pi_p) R_L \geqslant I$ under any given pre-screening intensity π_p. Thus, there is no adverse selection. Moreover, the bad loans that are passed on by informed, sophisticated investors are never funded.

D Additional Tables and Figures

[47] Note that these results may change if we instead assume at least three types of projects, in particular, more than one type of positive NPV project. Then, faster investors may take higher-value projects, leaving lower-value projects for slower investors. But such an assumption would require to change the model setting and add more assumptions, making the predictions not directly comparable to the ones of our baseline model.

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

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

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

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

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

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

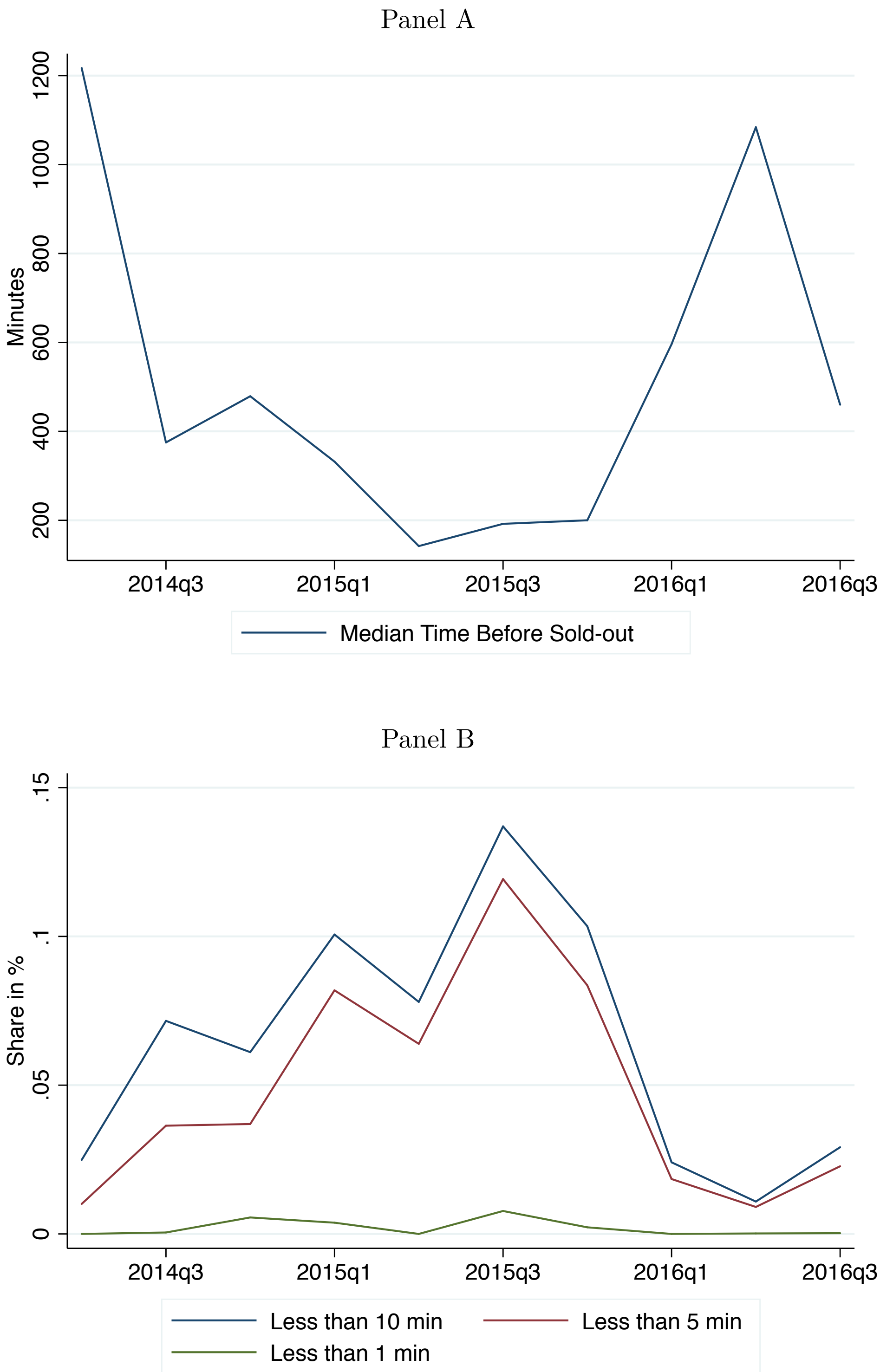


Figure D.1: Time for Loans to Sell Out

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.