Lending Platforms' Information Aggregation under Competition

Xin Xue*

Abstract

I examine how lending platforms use their information advantage to deter adverse selection in their competition for creditworthy borrowers. Using data from peer-to-peer lending platforms and a platform entry event which exogenously induces platform competition for both borrowers and lenders, I find that the incumbent platform possesses an information advantage by incorporating information that cannot be leveraged by the entrant into borrower screening and loan pricing. I show that, with better knowledge about the borrowers' quality, the incumbent platform attenuates its own adverse selection by aggressively undercutting loan prices for creditworthy borrowers, leaving a less favorable borrower composition and thus exacerbating entrant's adverse selection. Lenders on the incumbent platform free ride on its information advantage, while to compensate for information disadvantage, those on the entrant demand higher ex ante loan prices.

JEL Codes: G23, L13, D82

*xx3fe@virginia.edu, Economics and McIntire School of Commerce, University of Virginia.
I thank my dissertation committee advisors Zhaohui Chen, Andrew Kloosterman, Denis Nekipelov (Chair) and William Wilhelm for helpful advice and comments. I thank my former advisor Latchezar Popov; Eric Young, Simon Anderson, Karim Chalak, Gaurab Aryal, Stephen Karolyi, Mingfeng Lin, Jiasun Li, Ruomeng Liu, Mike Andrews, Don Carmichael and numerous seminar participants for their feedback and discussions. I am grateful for generous financial support from Bankard Fund and Kauffman Foundation. Errors are mine.

1 Introduction

Like the winner's curse in auctions (Milgrom and Weber 1982), a similar adverse selection problem arises in financial intermediation when intermediaries such as banks, credit rating agencies or platforms have asymmetric information about borrowers' quality.[1] Borrower adverse selection can cause financial market failure and is arguably exacerbated by competition intensity (Sharpe 1990, Marquez 2002, Von Thadden 2004). While several empirical studies have documented this type of adverse selection (Shaffer 1998,Carmichael 2017), none have investigated how intermediaries strategically attenuate it. I argue that intermediaries' lending standards lie at the heart of financial stability, because they centralize borrower screening and information aggregation for investors (lenders).

In this paper, I use novel datasets to examine how a lending platform uses its information advantage to strategically compete for creditworthy borrowers. I show that by incorporating information that cannot be used by other players into its lending standard, the platform deters its own borrower adverse selection and exacerbates that on other players. My empirical results show distinction from the symmetric Bertrand-Edgeworth equilibrium in Broecker (1990), in which pricing below the competitor's rate has the advantage of a more favorable borrower composition, while it can aggravate adverse selection under information asymmetry.[2] I discover that with better knowledge about borrowers' quality than its competitors, the platform leverages the information shared with its competitors to undercut prices for creditworthy borrowers. This paper provides new evidence

[1]A borrower accepting the offer at an intermediary indicates that the intermediary values the borrower higher than how much its competitors do and might have underestimated the loan price for the borrower.

[2]Without information advantages, intermediaries do not always undercut. If the competitor's signals of the borrowers are unknown, undercutting can exacerbate winner's curse. This mixed strategy Nash equilibrium is illustrated in Broecker (1990).

that prompts theory on asymmetric interbank competition with imperfect monitoring and without signaling devices such as collateralization.(Marquez 2002,Hauswald and Marquez 2006,Karapetyan and Stacescu 2013, Berger and Udell 1990)[3]

My study lies in the context of the peer-to-peer lending duopoly platforms Lending Club and Prosper in 2009. Peer-to-peer lending is a type of crowdfunding that matches individual borrowers and crowd investors, including individuals and institutions such as banks and pension funds. A peer-to-peer lending platform collects borrowers' information and posts it online to attract investments and profits from loan origination.

Compared with Prosper, Lending Club is able to possess an information advantage, because the two platforms used differentiated lending mechanisms during the period. Lending Club actively screened borrowers and posted loan prices. However, Prosper did not provide the borrower-screening service. Given borrowers' credit reports, interest rates were determined by the market-clearing prices of auctions, in which investors entered bids until the credit supply matched demand.[4] Although neither platform could disclose borrowers' private information, such as race, gender, education and exact residence location to investors, the difference in their lending mechanisms provided Lending Club (and its investors) with potential to obtain an information advantage.[5] That is, Lending Club could incorporate borrowers' private information into its lending standard, whereas Prosper could not. The winner's curse-alike adverse selection arises from the information

[3]See similar works on intercredit rating agency (CRA) competition. (Skreta and Veldkamp 2009,Bolton et al. 2012,Becker and Milbourn 2011,Flynn and Ghent 2017)

[4]Prosper set standards for the eligibility of a loan application, such as a FICO score above 600. Here screening refers as intermediaries using their private information including their private beliefs of the status quo of the economy and borrowers' quality, and information that cannot be public acquired.

[5]While they must disclose credit reports to investors, they must keep private information opaque by regulation. See the prospectuses of Lending Club and Prosper as well as the Equal Credit Opportunity Act and Regulation B Act.

asymmetry between Lending Club and (investors on) Prosper about borrowers' quality.[6]

Using the exogenous event of Prosper's entry where Lending Club's monopoly becomes a duopoly of the two, I show that to deter borrower adverse selection, Lending Club incorporates more private information into borrower screening and loan pricing following Prosper's entry. This additional information adds precision to Lending Club's knowledge about borrowers' quality, motivating its undercutting behavior to compete for creditworthy borrowers. I find that, compared with its loan pricing strategy preceding Prosper's entry, Lending Club provides creditworthy borrowers with better rating and lower interest rates. Conditional on the information shared with Prosper, it increases the loan prices for risky borrowers to deter borrower adverse selection, while lowering interest rates for creditworthy borrowers. By doing so, Lending Club leaves Prosper a pool of less creditworthy borrowers. Investors (lenders) on Lending Club free rides its additional information aggregation. Lacking screening devices, bidders on Prosper face a less favorable borrower composition and thus demand higher information rent. Loans on Prosper are even less likely to be funded due to credit rationing concerns (Stiglitz and Weiss 1981), leading to the demise of auctions in this market.(Wei and Lin 2016)

This paper provides policy implications for regulators regarding the information that should be disclosed to investors. Regulators intend to censor certain information such as race, gender and education to maintain equality, whereas it is arguable that borrowers intended to be protected might suffer the most.[7] They are plausibly the first screened out

[6]In addition to using different credit bureaus for borrowers' credit reports, Prosper targets itself as a social-financing platform, while Lending Club plays an "arms-length" game.While Lending Club uses TransUnion as its credit bureau, Prosper uses Experian. These credit bureaus' reporting standards vary, from which their information asymmetry may arise. As a social financing platform, borrowers and lenders may form a network for repeated borrowing (see Lin and Viswanathan (2015)). However, it is shown that the number of relationship accounts in this market is far less than that in commercial banking (see "Comparison to Banking" in the appendix).

[7]There are also counter arguments from the following empirical studies. Using data on Prosper

by Lending Club. Pooling with a greater number of risky borrowers, they are unlikely to obtain financing on Prosper either.[8] These negative externalities induced by information censorship plausibly engenders more inequality.

To pinpoint borrower adverse selection, I look for borrowers who shop for less expensive rates between Lending Club and Prosper.[9] Without observing borrower identities, I use unique data features such as borrowers' addresses, FICO scores and loan application dates to fuzzy match them across platforms.[10] However, adverse selection measured by matched borrowers bears both selection bias and the endogeneity issue. Data latency censors borrowers who withdraw applications from Lending Club for cheaper rates on Prosper. Additionally, Lending Club's lending standard arguably affects the severity of adverse selection. To identify how Lending Club reacts to adverse selection, I use Prosper's entry event, in which a platform monopoly became a duopoly, after which adverse selection emerges on both platforms. Its emergence does not come from variations in lending standards but rather from the exogenous formation of competition over borrowers. To study this event as a quasi-experiment, I first verify that the entry event is not anticipated by Lending Club.

I test whether Lending Club incorporates more private information into its lending standard after the entry event. Because I do not observe Lending Club's private infor-

1.0, Ravina (2007) finds that investors' investment decisions contain psychological components and bias towards borrowers' appearance. Lin and Viswanathan (2015) empirically document home bias in this market, where investors fund borrowers that live close to them. However, the short-run bias does not persist, if investors can learn in the long run to correct themselves (Freedman and Jin 2008).

[8]Without additional information, investors bid high to avoid the winner's curse on Prosper. Credit crunches can also occur due to investors' credit rationing. (see Iyer et al. 2015)

[9]Borrower rate shopping is a general term characterizing borrowers who apply for loans on both platforms and choose the less expensive option. (See Bolton et al. 2012) Here, due to credit rationing concerns (Stiglitz and Weiss 1981), rejected or unfunded borrowers would accept loans at very high interest rates.

[10]I thank Don Carmichael from the University of Houston for sharing his insights into cross-platform data differences and his understanding of institutional details. I attribute my revisions of the matching process to him and his work. See Carmichael (2017)

mation, I use borrowers' local economic status quo to measure platform-observed and investor-unobserved borrowers' quality. I use county-level bachelor's degree attainment from the U.S. Census and quarterly reported deposits of borrowers' local community banks from the FDIC, respectively, to measure borrowers' education and financial health. I use borrowers' loan description lengths to measure their efforts in communicating with investors, which is arguably correlated with borrowers' quality.[11] I first show that these measures of private information matter to borrowers' loan performance in logical directions. Then, I incorporate these measures into the estimations of Lending Club's lending standard. I find that these coefficient estimates show greater statistical significance and economic magnitude in the post-entry period. This finding indicates that, facing potential rate shoppers, Lending Club aggregates more private information.

To test whether aggregating private information helps Lending Club deter adverse selection, I focus on its rejected borrowers who obtained loans on Prosper. Had Lending Club not rejected these borrowers, would they have worsened or improved its average loan performance? I match interest rates of these Lending Club-rejected borrowers with the interest rates of Lending Club-accepted borrowers, and bilaterally compare their performance. I find that the loan performance of the Lending Club-rejected borrowers is significantly worse, indicating that Lending Club does attenuate adverse selection by screening them out. By comparing the performance between the Lending Club-rejected borrowers and other borrowers on Prosper, I find that these once-rejected borrowers' performance worsens. Since Prosper cannot convey private information to bidders or

[11]Note that I standardize the lengths of borrower descriptions by loan application month and loan rating for the following reasons. First, given a bad rating, a borrower might write more to signal his or her quality to investors. Second, during a time period when investors' credit supply is high, some borrowers choose to write little or nothing.

screen them out, the adverse selection problem is aggravated.

With superior knowledge of borrowers' quality than Prosper, Lending Club becomes more aggressive in competing for creditworthy borrowers using its information shared with Prosper. Using a sharp regression discontinuity (SRD) design, I show that, conditional on information shared with Prosper investors, the propensity of borrower acceptance by Lending Club increases after the entry event. Among the accepted borrowers, I use an ordinal general linear model to estimate Lending Club's credit rating scheme. Under the SRD setup, I show that "post-entry" borrowers are more likely to receive better credit ratings. I use a quantile treatment effect (QTE) model to measure interest rate changes with respect to borrowers of different creditworthiness. Specifically, I match propensity scores from borrowers in pre- and post-entry samples, and evaluate the difference in interest rates at various borrower creditworthiness. I find that creditworthy borrowers receive less expensive financing, indicating price undercutting, while Lending Club sets higher loan prices on risky borrowers, indicating deterrence of adverse selection. (Stango and Zinman 2015) Since the post-entry loans exhibit worse performance than the pre-entry loans, it is arguable that, without perfect information, adverse selection cannot be fully deterred. I show that as investors receive smaller revenue streams, their sentiment lowers, liquidity drops and credit crunches emerge. These outcomes illustrate the exact disincentive for Lending Club to bear such adverse selection.

This paper raises a larger question in investment bank security offerings: book building or auctions? Although auctions are rarely used in IPOs, they are perceived as a better mechanism, because they place shares among those who value them the most and arguably alleviate IPO underpricing. (Derrien and Womack 2003). However, if bookrunners have information superior to that of institutions or even the issuers themselves (Chen et al. 2014), to attenuate winner's curse, investors plausibly bid even less (Rock 1986).

2 Institutional Details and Data

2.1 Peer-to-Peer Lending Summary

Peer-to-peer (P2P) lending is a type of profit-seeking crowdfunding, serving as an online platform through which investors (lenders) and borrowers are matched.[12] Since its introduction in the U.S. in 2005, the accrued loan volume issued has reached $25 billion, leading the U.S. online alternative finance industry.[13] Two major platforms operating within the U.S., Lending Club and Prosper, account for 98% of the market.[14] Lending Club is currently the world's largest profit-seeking crowdfunding platform, followed by its largest competitor, Prosper.com.[15]

2.2 The Duopoly: Market Mechanism and Market Agents

Market mechanisms on peer-to-peer lending platforms change over time. Loans issued through peer-to-peer lending were not classified as securities that could be legally traded on a secondary market until October 15, 2008. Lending Club was required to suspend its activities and register with the SEC on April 8, 2008. One day after it returned on

[12] Other types of crowdfunding include donation (Donorschoose), product reward (Kickstarter), and equity venturing (AngelList).

[13] The Americas alternative finance benchmarking report, 2016. During 2014 alone, it accounted for $5.5 billion loans issued, and according to PwC's estimation, this figure could reach $150 billion annually. (See PwC 2015, Peer Pressure.)

[14] According to an Economist article in 2014, "Peer-to-Peer Lending: Banking without Banks."

[15] Lending Club currently accepts borrowers in all states but Iowa and West Virginia and investors in all but Ohio, Pennsylvania, New Mexico, North Carolina and Hawaii. Other platforms such as Academic Capital Exchange, CapAlly and GreenNote joined the market recently. By 2015, $22 billion in loans were issued among them.

October 15, 2008, Prosper 1.0, the other platform, left the market for to undergo same procedure. Prosper returned on July 13, 2009, as Prosper 2.0.[16] When returning to the market, Prosper 2.0 continued with its auction business model. On December 20, 2010, Prosper 2.0 changed to a platform-pricing mechanism.[17] However, the platform-pricing mechanism has been Lending Club's business model from the beginning. In this section, I describe the market and both the platform-posting and auction mechanisms between 2009 to 2011, as this is the period on which I focus.

Three types of agents participate in this market, borrowers, investors and platforms.[18] Any U.S. resident with a Social Security number is entitled to request an unsecuritized loan of less than $25,000 on a peer-to-peer lending platform for a 3-year or 5-year term.[19] An application requires the borrower's SSN, current employment and income, and homeownership status, address, the intended usage, the term and size of the loan and the borrower's personal financial status.[20] Although the usage of the loans is not monitored, more than 60% of the loans issued are claimed to be used for personal debt consolidation. [21] After obtaining borrowers' credit reports, the platforms use an algorithm or a strict rule to evaluate the borrowers' creditworthiness and screen out "unqualified" applicants.[22]

[16] Compared to Prosper 1.0, Prosper 2.0 set stricter rules to screen borrowers with poor credit history.

[17] The auction mechanism is described in the following section. Lin and Wei studied this mechanism change in detail.

[18] Both Lending Club and Prosper partner with WebBank, an FDIC-insured Utah-chartered industrial bank. For both platforms, WebBank is the sole underwriter of all of the loans. It is believed to be favored due to the usury law and consumer finance code in Utah. See "The future of finance, the rise of the new shadow bank" from Goldman Sach's Research; Prosper Investor Registration Agreement and Lending Club Prospectus; "Where Peer-to-Peer Loans Are Born," Bloomberg.

[19] Currently, 1-year term loans are also available.

[20] Platforms do not require verification of employment or income. If a borrower's status is verified, it is posted on his or her profile.

[21] According to the Federal Reserve, the average interest rate on an issued credit card was between 13% and 14% (APR) in 2014. On Lending Club alone, conditional on a FICO score of 750 or greater, the average interest rate is 8.4%. Other frequently reported usages include car financing, housing and home purchases.

[22] Lending Club pulls credit reports from TransUnion, and Prosper pulls reports from Experian. Lending Club strictly rejects any borrower with a FICO score less than 640 while Prosper rejects those with

A borrower who passes the initial screening enters into a "scoring" system that rates his or her creditworthiness. For Lending Club, the rating spectrum ranges over 7 major categories, and within each, there are 5 finer categories.[23] For Prosper, within its 7 categories, the finer category indicates the term of the loan (3- or 5-year).[24]

2.2.1 Platform-Pricing Mechanism

Under the platform-pricing mechanism, each rating maps onto an interest rate. Each intended loan is split into notes of $25 as a type of security, and is posted on its platform's website along with the borrower's credit history to attract financing. The listing period can last for as long as 14 days. After observing the loan contracts and the borrowers' credit history, investors choose a number of notes to invest or set up portfolios for automated investments. An investment decision is a commitment and cannot be withdrawn unless the loan is canceled by the platform or the borrower.[25] A loan will be issued as soon as the investors' commitment reaches the requested loan size. In other cases, if the pledged amount exceeds 60% of the request at expiration, the borrower choose to issue the funded amount under the same interest rate or to reject the loan and refund the investors.[26] As previously mentioned, Lending Club has been using this mechanism since it was founded in 2007, whereas Prosper 2.0 changed to this mechanism from the auction business model on December 20, 2010.

---

a score less than 600.

[23]That is, a total of 35 categories from A1, A2—G5, where A1 is the most creditworthy and G5 the riskiest.

[24]Categories on Prosper 2.0 are denoted as AA, A, B,..., E, HR.

[25]During a loan's funding period, potential investors can observe the time remaining on the loan until its expiration and how many investors and total investments have been committed.

[26]It is 60% for Lending Club and 70% for Prosper. If the commitment is less than the threshold, the loans are be dropped, and the investors are refunded.

2.2.2 Auction Mechanism

Under Prosper's auction mechanism, regardless of loan rating, the contractual interest rate of a loan is a market-clearing price from a multiunit uniform price auction. After observing the borrower's quality and the auctioneer's rating, investors choose to enter the 7-day bidding process. First, the borrower sets a "reservation" interest rate as the opening bid, observable to all investors.[27] Investors submit their bids on the interest rate and unit of notes. [28] The bids are not completely sealed. As investors come in sequentially, they do not observe previous investors' interest rate bids, but they can observe their units and numbers of bidders.[29] At the listing expiration, all interest rate bids are ranked in ascending order, and the cutoff for the winners is where the pledged amount exceeds the requested amount. The loan's contractual interest rate is set as the first that is outbid. If the loan is not 100% funded, the loan is canceled, and the bidders are refunded.

Investors on the platform can be individual investors or institutions. Prosper once disclosed investors' composites on each loan and their complete portfolios. Information about individual investors was not disclosed. It was reported that institutional investors financed more than 80% of the total loan proceeds in the recent years.[30] According to the data and a former CEO, Lending Club as a platform, has partially funded some of the borrowers listed on its website.[31] The amount pledged by the platform on each loan is

[27]The reservation interest rate is the highest interest rate that the borrower accepts. However, borrowers may not reveal their true preferences.

[28]Their interest rate bids must be less than the opening bid. Additionally, remember that a loan is divided into notes of $25.

[29]Zhang and Liu (2012) study investors' herding behavior using the timestamps of the bids among the issued loans.

[30]According to Lending Club, institutions can be banks, pension funds, asset management companies, etc. See "The Evolving Nature Of P2P Lending Marketplaces," Techcrunch. Institutional investors tend to invest in whole loans rather than a fraction. See "Wall Street is hogging the peer-to-peer lending market," QUARTZ.

[31]However, Prosper has not played the role of a investor. See "A Look Back at the Lending Club and Prosper Quiet Periods," Lending Academy. However, even before and after the "quiet" period, Lending

not observable to the investors in the funding process, but it is a variable in the dataset.

Given an issued loan, the borrower pays an upfront origination fee from 0.5% to 5% depending on its rating.[32] With the origination fee, interest rate and loan terms, it becomes a take-it-or-leave-it contract offer back to the borrower.

The repayment structure on the loan contract amortizes monthly. Borrowers are expected to follow the payment schedule, and early payoffs are encouraged. If payments are delinquent more than 150 days, the loan is charged-off and sold to collection agencies for recovery. Defaults and delinquencies appear on borrowers' credit reports and limit their future borrowing ability. Other institutional details are described in the appendix.

3 Data and Summary

3.1 Data Sources and Sample

Both Lending Club and Prosper provide unsecured loan data to the general public. I obtain loan-level cross-section snapshot of all loans originating on and rejected by Lending Club from 2009 to 2011, and all loans listed and originated on Prosper during the same time frame.[33] Since my study focuses on the entry event, I describe the data based on the entry sample from February, 2009 to December, 2009. Summary statistics for Lending Club and Prosper are presented in Table 1 and 2. [34]

Club also lent to its borrowers. Former CEO also told the Wall Street Journal that Lending Club slowed its activity in the "quiet" period to use its own money to fund borrowers. "Peer-To-Peer Investors Get Into Secondary Market," WSJ.

[32]For a loan of $1,000 with a 5% fee, the platform immediately profits $50 of loan origination fee. The borrower receives $950 of capital, but the principal on the loan stays at $1,000. Fees on Prosper range from 0.5% to 4.5% and 1%-5% on Lending Club. See Table 3 for origination fee comparison.

[33]The datasets were obtained in 2017. Therefore, all loans originating in the sample period are fully paid or charged off. I only observe data on Prosper after its entry on July 13, 2009.

[34]In empirical analyses of the effect of private information on performance, I use a longer time frame from 2009-2011.

In the sample period, the total number of applications on Lending Club reached 58,945, compared to 27,855 since Prosper's re-entry. The average acceptance rate was 6.41% prior to Prosper's entry and 11.33% thereafter. The total number of listed on Prosper is 15,120, yielding a 14% funding rate. Among the post-entry funded loans, the average borrowers' funded loan size is \$10.34 K on Lending Club and \$ 4.38 K on Prosper.[35] The number of loans (loan proceeds) were issued between Lending Club and Prosper are 3,149 (\$32.563 million) and 2,141 (\$9.4 million), respectively, with a market share of 78% measured by proceeds. Loans (proceeds) issued on Lending Club amounted to 1,899 (\$17.32 million) preceding the entry event. Among issued loans, I observe 3 categories of information: the loan's funding and contract; borrower's characteristics and credit history; and the loan's ex post payments.[36]

3.2 Observable Characteristics

Using the loan size and the monthly installment in a loan contract, I compute the borrower's monthly interest rate using the internal rate of return method.[37] The average (standard deviation) IRR on Lending Club was 1.036% (0.21%) prior to Prosper's re-entry and 1.036% (0.23%) after it. The market clearing interest rate was 1.61% (1.52%) on Prosper. Other information includes loan ratings assigned by the platform, loan orig-

[35]Loan applications that were not funded to 70% or withdrawn by the borrowers are not observed on Lending Club.

[36]For loans rejected on Lending Club, each observation stands for a loan application declined by the platform. The data only contains 7 variables: the application date, the requested loan size, its intended purpose, the borrower's vintage FICO score at the time of the application, and the borrower's debt-to-income ratio, ZIP code and employment length.

[37]Internal rate of return (IRR) measures he break-even discount rate where the net present value (NPV) of the investment is 0. $K = \sum_{t=1}^{T} \frac{P_t}{(1+\mathrm{IRR})^t}$, where, K is the funded amount, P_t is the payment of the loan at time t, and T is the loan term of 36 or 60 months.

ination date and borrower-reported loan purposes.

A borrower's attributes show his or her address as a 3-digit ZIP code and state,home ownership, employment job title, length of current employment and a self-reported full loan description. The borrower's income is self-reported, but the platforms provides an indicator to investors regarding whether it is verified. The credit history consists of the range of a borrower's FICO score, debt-to-income ratio excluding mortgages or the current loan, number of credit lines, revolving balances and utilization, number of delinquencies over the preceding two years and date of the first credit line (month/year). As a brief comparison, borrowers' (applicants') FICO scores by TransUnion average 714 (578) on Lending Club and borrowers' FICO scores' upper bound by Experian is 701 (688) on Prosper.[38]

A loan's payment information includes 3 key variables. During the sample period, a loan could have a status of "fully paid" or "charged off." Among the loans issued on Lending Club, 685 were charged off, with 417 originated after Prosper's re-entry. The charged-off rate was 14%. Among the loans on Prosper, 330 were charged off, yielding a charged-off rate of 15.4%. With the total payments on each loan, the sum of the interest and the principal payments, I compute their return on investments and percentage of nonpayments.[39] The average ROI on Lending Club is 8.14% compared to 15.69% on Prosper. I also observe the last payment date on the loan.[40] Information on the number of investors and loan application date was not provided. I obtain these data features

[38]Note that Prosper only provide borrowers' FICO scores up to a range.

[39]Note that loans that are paid off early could have a lower ROI than that under the payment schedule. $\text{Percentage of Nonpayment} = \frac{\text{Contractual Payment} - \text{Actual Payments}}{\text{Contractual Payment}}$. For fully paid loans, I set its percentage of nonpayments as 0.

[40]Note that, for some fully paid loans, the total payment amount can exceed or be short of the amount listed in the contract. Recoveries and collection fees are also recorded but are not used in the paper and thus are not be reported.

from each loan's archived funding page on "Lendingclub.com" and match them back to the main data using loan ID. The post-entry average (median) number of investors per loan is 136 (130) on Lending Club and 147 (94) on Prosper.

3.3 Measuring Private Information

The platforms are able to observe borrowers' information that neither an average investor nor an econometrician can see. Under the platform pricing mechanism, the unobservable datum can be used by the platform on its loan pricing (see Blanchflower et al. 2003). However, under auction, it cannot be incorporated into investors' bidding prices, unless they have personal connection with the borrowers (see Freedman and Jin 2008). I measure the unobservable borrower characteristics from multiple dimensions by combining data uniqueness with several strands of the literature.

I observe applicants and borrowers' 3-digit ZIP codes. To measure platform-censored demographics, I obtain county-level education attainment measured by the percentage of population with a bachelor's degree (BA Attainment) from the Census. I aggregate the information from the county-level back to the 3-digit ZIP code level.[41] On Lending Club (Prosper), the average "BA Attainment" was 32.75 (30.99) among the borrowers during the post-entry period. Using a similar method, I collect quarterly data on "Community Banks'" deposits from the FDIC.[42] "Deposit" is the variable that measures local financial health at borrowers' residences, with an average of 21.66 $ million (12.95 $ million) on Lending Club (Prosper). Higher deposits indicates better financial stability.[43]

[41]For one-to-many mapping, I allow each ZIP code mapped onto it to share the same demographics.

[42]The reason that I do not use large commercial banks is because the loans and borrowers here are for personal use. The variation in commercial bank deposits at borrowers' addresses might not capture local financial stability.

[43]Other local variables are also experimented with, such as local housing prices, local bank charged-off

For borrowers who provide personal descriptions on their loans, I extract the length of the loan descriptions measured by their number of characters. The length of a loan description captures the cost of effort for the borrower to signal investors, independent of the loan rating.[44] This signal is arguably correlated with borrowers' types and is observable to the platforms.[45]

4 Identify Shopping Borrowers

A borrower that accepts an offer at one platform is possibly rejected or given a higher price by the other player, plausibly because the other player observes negative information.[46] Without observing borrower identity, I use several key criteria to construct fuzzy matches between borrowers on the two platforms.[47] The basic idea of the fuzzy matching method follows Carmichael (2017).[48] He provides a detailed cross-platform matching in this market using post-2011 data.[49] Rate shopping behavior between the two platforms attributes to the following 3 scenarios.

- A borrower is funded on both platforms but cancels the higher interest rate.

rates, and so on. However, these variables either maintain high correlations with either BA Attainment or Deposit or do not matter to borrower performance and thus are not described.

[44]To measure the cost of effort, I regress the lengths of loan descriptions on the ratings and monthly fixed effects. I use the residual to measure borrowers' cost of effort.

[45]I also experiment with "text-mining" typos, such as misspelling and grammar mistakes, within their loan descriptions, which could reflect borrowers' educational backgrounds. However, this information cannot be consistently tracked. For details about these descriptions, see the appendix. Additionally, uniquely on Prosper, conditional on an issued loan, I observe whether a the borrower is recommended by any investor.

[46]Additional adverse selection problems exist in this market. As borrowers appear in this market instead of local banks and credit unions, it may indicate that their characteristics do not favor them.

[47]If I am able to observe borrowers' identities, I can determine the exact rate shopping behavior in this market.

[48]Other such as Liu et al. (2013) cross-platform match application identities between two online app stores: iTunes and Android.

[49]Again, I thank Don Carmichael for sharing his insight into the data differences and understanding of institutional details. I attribute my later revisions of the matching process to him and his work.

- A borrower is rejected by one platform (or its investors) and funded on the other
- A borrower is rejected by both platforms (or their investors)

Both platforms share relatively unique information including 3-digit ZIP codes, application dates, FICO scores (from different agencies), opening year of first credit line and open accounts. I use this information to identify unique matches. That is, for each borrower (denoted by an "input") accepted and financed on one platform, I look for a "target" borrower on the other platform that was previously rejected or unfunded. Alternatively, for each "input" rejected or unfunded on one platform, I look for a "target" that is also rejected and unfunded. Note that, Prosper provides data on borrower-withdrawn loans, but Lending Club does not. Therefore, for the first scenario, the matching is one-sided.[50]

Step 1: Sample Selection and Refinement First, I assume that rate-shopping borrowers submit applications close to each other in time. I limit the "target" search pool to those who submit their application within two weeks of the "input's" application date. I use 2-week windows because the maximum listing period on a platforms is 14 days. [51] The input and the target must come from the same address or 3-digit ZIP code. The rest of the key variables add difficulties for exact matches, since the platforms use differentiated "hard" information generated by two credit rating agencies. I restrict the target borrowers' FICO scores to a ± 50 window centered at the "input's" FICO.[52] I match the borrowers' opening date of first credit lines at the year level.[53] For an "input,"

[50]In later years, banks partner with platforms to fund whole loans. Borrowers with unverified information are subject to cancelation by the platforms. I do not have to consider this situation before 2011.

[51]I also have experimented with 10-day and 7-day windows.

[52]Although the platforms obtain credit reports from different agencies, I contend that their information is highly correlated.

[53]My personal first credit line is 08/2008 on the TransUnion Report, but it is showns as 11/2008 on Experian. The total number of accounts is 8 on TransUnion and 9 on Experian.

the first step can generate no or many matches. For cases with no matches, I do not further explore them. For one-to-many matches, I proceed to step 2 to construct a metric to refine the match.

Step 2. Fuzzy Matching Due to differentiated reporting scheme by the platforms and their credit agencies, I construct a metric on the difference between the input and each target in the refined sample from step 1. The metric is a product of two differences: the difference between the input and targets' FICO scores and the difference between their application dates. I choose the borrower with the highest matching metric. [54]

This cross-platform matching method is mechanical, and noise may be generated through the following channels. Borrowers' addresses are maintained at the city/county level by Prosper and manually mapped onto a 3-digit ZIP code to match with Lending Club.[55] A large chunk of credit information on rejected borrowers is censored by Lending Club. To avoid false positive matches, I follow Carmichael (2017) to conduct placebo tests. I subtract 6 months from the borrower application dates on one platform to match them using the same method. I find a significant reduction in the matching rate. To further validate the matching method, I also use borrowers' loan descriptions to verify their previous applications on the opposing platform.[56]

[54]I also do not use information that the borrowers self-report and that is later verified such as annual income. The metric can serve as a reference to weed out matches with large disparities.

[55]I also experiment with 2-digit ZIP codes and do not find significant difference.

[56]One borrower wrote "I've made a very successful move to Springfield, MO, after I made a Prosper loan to do so. I've been here in state for a year and a half now, and love the quality of life. I found that Prosper classified me as HR, but I don't quite understand why." It is not a large sample; thus, I do not conduct statistical inference examinations.

4.0.1 Rate Shopper Summary

I identify 37,021 matches from 2009 to 2011, with 7,312 in the sample period. These matches are one-to-one or many-to-one matches from Prosper to Lending Club, among which many-to-one matches identify shoppers listed on Prosper multiple times within 30 days.[57] These matches boil down to 4,566 rate shoppers, with 10% (477) having obtained loans from Prosper and 1.5% (70) from Lending Club (Table 5).

This uneven distribution of rate shoppers between the platforms is partially attributed to data latency. I do not observe borrowers who withdraw their loans at Lending Club for less expensive loans on Prosper. (See Figure 1) This data latency causes selection bias, since these censored observations likely obtain less expensive financing through Prosper's auctions.[58] With very few shoppers receiving financing on Lending Club, I do not conduct statistical inference using these shoppers. Instead, I focus on borrowers rejected by Lending Club but financed by Prosper.

5 Identification Strategy

In addition to the aforementioned selection bias, reverse causality concern exists since any unobserved variation in platforms' lending standards can result in changes in borrower adverse selection. Therefore, I must explore exogenous channels that induces borrower adverse selection. Prosper's entry event marks the commencement of borrower rate shopping within this market. As long as I can show that this event was not anticipated by

[57]Many of these repeated listings are later canceled due to expiration or withdrawals by borrowers.

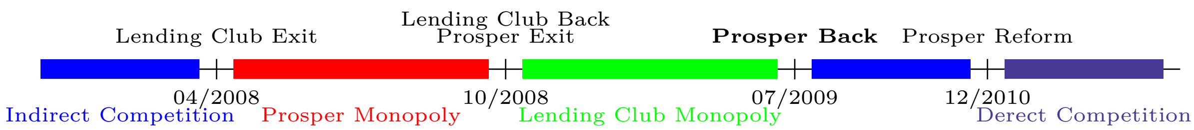
[58]These shoppers plausibly have "investor connections", and thus prefer auctions. Since Prosper has established itself as a social financing platform, it attracts highly creditworthy borrowers with "investor connections." See Freedman and Jin (2008)

Lending Club, its lending standard changes at the time of the event should be attributed to adverse selection and platform competition.

5.1 Prosper's (Re-)Entry Event

Prior to April 8, 2008, loans issued on Lending Club or Prosper were not classified as securities, and both platforms were asked by the SEC to undergo review and registration. On April 8, 2008, Lending Club first underwent its SEC registration and entered a "quiet" period. Lending Club discontinued new investor registrations, stopped advertising to borrowers and slowed its platform activities. On October 14, 2008, Lending Club returned, and all of its loans issued henceforth have been divided into notes and could be traded on a secondary market. One day later, Prosper (then 1.0) exited the market to begin its registration. From this point, Lending Club monopolized the peer-to-peer lending market until July 13, 2009, when Prosper announced its immediate return. Prosper 2.0 started setting stricter guidelines to screen borrower credit backgrounds, where an eligible borrower must have a FICO score of greater than 600. Still using the auction mechanism, Prosper 2.0 also began to rate borrowers' creditworthiness. As the monopoly became the duopoly, I start to identify rate shoppers, and the monthly number of applications on Lending Club decreased by nearly 2,000. (See Figure 1) To study this event under a quasi-experiment setting, I show that the event is indeed exogenous to Lending Club.[59]

Market Structure Timeline



[59] I show this finding in the Robustness section of the appendix.

6 Hypothesis Development and Estimation

The commencement of the duopoly introduces their competition for both borrowers and investors, distorting Lending Club's loan pricing strategy. Loan price undercutting may result in a more favorable borrower composition. Conversely, not observing borrowers' signals on Prosper, loan price undercutting may aggravate adverse selection. These two-sided arguments yield the Bertrand-Edgeworth mixed-strategy equilibrium in Broecker (1990). By incorporating additional information that cannot be conveyed by Prosper to its bidders, Lending Club gains better knowledge about borrowers' quality, attenuates borrower adverse selection and becomes more aggressive to undercut loan prices.

Measures of both observable borrowers' credit histories and private information are multidimensional. I hereby introduce some notations. I denote borrowers' observable characteristics as X and private information as Z. For the platform, upon these observables, the platform chooses a message for investors. Empirically, the message "m" subsumes borrowers' acceptance, credit ratings and interest rates.

6.1 Private Information Aggregation

To show the entire mechanism, I first must show that the variables that I measure with private information matter to borrowers' loan performance. I use borrowers' local bachelor's degree attainment and borrowers' local community banking deposits to measure borrowers' education and financial health, respectively. I use the borrowers' loan description length to measure their cost of effort and willingness to provide information.[60]

[60]Note that I am not arguing that BA Attainment, Deposit or Description Length per se are the exact factors considered by the platform. Without observing information at borrowers' individual level, borrower's local characteristics can measure an average borrower's quality as well as exposure to local financial health. Other variables are also experimented such as local housing price and local charged-

Hypothesis 1. *Better BA Attainment improves borrowers' performance, as do Deposit and Loan Description Length.*

If the previous hypothesis is valid, I test whether Lending Club uses this information in borrower screening and credit rating to deter the adverse selection problem. I hypothesize that during post-entry, private information matters more in terms of its statistical significance and economic magnitude.

Hypothesis 2. *Higher BA Attainment, Deposits and Description Length improve borrowers' likelihood of acceptance and credit rating. Moreover, their effects become more significant following the entry event.*

I do not observe loan description for rejected loans. Therefore, Description Length is omitted out of the borrower-screening estimation. Given that Next, I must show that Lending Club benefited from the additional information. To show this phenomenon, I use data on both platforms to examine these shoppers' performance and test the following two hypotheses.

Hypothesis 3.
- *The deterrence of adverse selection: Had these Lending Club-rejected shoppers been issued on Lending Club, their performance would worsen the average performance on Lending Club.*
- *The exacerbation of the opponent's adverse selection: These Lending Club-rejected shoppers' performance is worse than the average performance on Prosper*

These two hypotheses address different issues. The first compares the performance of rejected shoppers' with borrowers on the incumbent platform. Hypothetically, these

off rates. However, these variables preserve high correlations (absolute value around 0.7) with BA Attainment or Deposit, and thus are omitted from estimation.

shoppers would have below-average loan performance if they were not rejected. The second hypothesis compares the shoppers' performance with that of other nonshoppers on Prosper. That is, if these shoppers' performance is inferior, it indicates that Prosper's adverse selection problem is exacerbated.

6.1.1 Empirical Strategy

I first test whether the measures of private information, Z, matter to loan performance, Y. I use ex post loan performance measured by Return on Investments, Nonpayment and Default as 3 dependent variables. In OLS specifications, I control for borrowers' observables and contract terms and include monthly fixed-effects at issuance, monthly fixed effects at loan closure, and address fixed effects at the state level.[61]

$$Y_i = X_i\beta + Z_i\delta + \varepsilon_i \tag{1}$$

To illustrate their economic magnitude, I standardized all of the coefficients for easier interpretation. Table 6 shows that these measures of private information do matter to borrowers' performance, particularly BA Attainment.[62] A one standard deviation increase in BA Attainment corresponds to a 0.042 standard deviation decrease in default propensity (or 1.4% decrease in default propensity), a 0.03 standard deviation increase in ROI and a 0.037 standard deviation decrease in borrowers' Nonpayment. The coefficient estimates are consistent with the hypothesis that the measures of borrower private

[61]Here I use full sample period from 2009 to 2011 since "Deposit" in 2009 may not be a norm.
[62]I do not observe loan descriptions for rejected loans. Therefore, Description Length is not incorporated into borrower screening.

information correlate with loan performance in the logical direction.[63]

Table 6 Here

Next, I incorporate these measures to estimate (Pre-Entry Screen), (Post-Entry Screen) and (Credit Rating). I compare their economic magnitudes and statistical significance before and after the entry event. Table 7 compares how those measures matter differently to borrower screening. In a pooled regression (Column 1), I use the entry dummy to interact with each covariate. I also separately estimate borrower screening using only pre- and post-entry samples, and pairwise compare the coefficients. Column (1) shows that a one standard deviation increase in BA Attainment corresponds to a 1.33% increase in acceptance propensity prior to entry. Its magnitude increases by 0.56% after the entry event. Similarly, a one standard deviation increase in borrowers' local deposit correspond to 0.46% additional acceptance propensity following the entry event. A pairwise comparison between Columns 2 and 3 shows similar results.

To better interpret the odds ratios in the ordinal linear models, I predict the average marginal effects of the coefficients with respect to each outcome. These marginal effects corresponds to the average effects measured in probabilities. Table 8 shows that for a one standard deviation increase in Deposit, the propensity of the creditworthy rating outcomes (Grades A and B) increases by 0.28% (0.2%+0.008%) pre-entry and 0.56% (0.37%+0.19%) post-entry. Relatively, an increase in Deposit corresponds to a decrease in the probability for risky ratings. In addition, the statistical power and economic magnitudes of the results show that the private information holds more explanatory

[63]Note that the higher Nonpayment that a loan has, the worse that it performs. BA Attainment represents borrowers' local bachelor's degree attainment to the population ratio; deposit indicates their local total deposits in "Community Banks", from the FDIC; Description Length represents the borrower-specific loan description length, standardized by loan Grade and origination month.

power following the entry event.[64]

Table 7 Here

I finally test the last hypothesis of the section. Is Lending Club able to deter adverse selection and exacerbate that of Prosper? Measuring the first part of the hypothesis, it requires data on borrowers who are rejected by Lending Club and financed on Prosper, denoted by shoppers. Since I do not observe many characteristics on shoppers rejected on Lending Club, I borrow their loans and characteristics from Prosper to compare with other loans on Lending Club. By doing so, I bear the following limitations. I assume that shoppers would not have different performance, had their loans been issued on Lending Club. Second, shopper identification is not an accurate procedure. Without observing borrower identities, there exists false-positive matches. Most importantly, in equilibrium, the accepted borrower quality of the platforms are likely to be fundamentally different. Pooling the data together likely introduces an outlier-problem which biases the average effects. Therefore, I use a propensity score matching (PSM) model to compare very similar borrowers across the platforms.

Controlling for borrowers' observables and contracts, I expect that Lending Club-rejected shoppers would obtain lower ROI. That is, had Lending Club not rejected the shoppers, Lending Club's performance would have become worse. The estimated results show that $\hat{\gamma}_{lc} = -2.32$ with a p-value 0.081, which is interpreted as, compared to a loan with the same contract term, rejected shoppers yield 2.32% less than accepted borrowers on Lending Club.

To test the latter part of the hypothesis, I only use Prosper's data and run the following

[64]BA Attainment does not show significant effects on any outcome, and thus is omitted from reporting.

regression. I use a dummy 1{Shopper} to indicate whether the loan was previously rejected by Lending Club.

$$Y_i = X_i\beta + \gamma_p 1\{\text{Shopper}\} + \varepsilon_i \qquad \text{(Adverse Selection)}$$

, where Y, loan performance, is measured with default, nonpayment and ROI. γ_p measures the average difference in performance between the identified shoppers and nonshoppers on Prosper, controlling for observables. In other words, the more negative that γ_p becomes, the more adverse selection Lending Club pushes towards Prosper. Table 9 shows that compared to the borrowers who do not show up on Lending Club, the borrowers who were previously rejected by Lending Club on average yield a 6% higher default rate, 4% less ROI and 3.4 % more nonpayment. This finding shows significant deterioration of loan performance on Prosper induced by adverse selection deterred from Lending Club.

Table 9 Here

6.2 Competition for Creditworthy Borrowers and Loan Pricing

In the previous section, I show that by aggregating private information, Lending Club attenuates its own borrower adverse selection and exacerbates that on Prosper. Holding the information advantage, Lending Club has better knowledge about borrowers' quality than Prosper. I expect that Lending Club becomes more aggressive to compete for creditworthy borrowers following the entry event.

Hypothesis 4. *Following Prosper's entry, conditional on the information shared with Prosper, Lending Club accepts more borrowers and gives creditworthy borrowers with bet-*

ter ratings.

Regardless of its information advantage, the information that Lending Club observes remains imperfect. The tension between undercutting and the adverse selection deterrence persists on Lending Club. I hypothesize that interest rates undercutting is contingent upon the noisiness in borrower information. If borrower's information is highly precise, their interest rates should exhibit undercutting. Otherwise, their interest rates might increase to deter adverse selection. This argument is similar to Cutler and Reber (1998) in examining insurance markets and to Stango and Zinman (2015) in studying credit card markets.

Hypothesis 5. *For borrowers with less noisy information, their interest rates decrease. For those with more uncertainty, their interest rates increase.*

6.2.1 Empirical Strategy

I apply a sharp regression discontinuity (SRD) design to study the event as a quasi-experiment. (see Lee and Lemieux (2010)) SRD is a cleaner and more robust strategy than average treatment effect (ATE) models. The quality of the applicant pool changes over time which contaminates the ATE estimates. (see Table 1) Applicant quality shows improvement over time; thus, the ATE overestimates the effect of entry. I take advantage of the exogeneity of the event by focusing on the discontinuity of screening policies. I use Probit models to estimate borrowers' acceptance, $\mathbb{1}\{\text{Accept}\}$, separately before and after the entry event.

$$\mathbb{E}\{\mathbb{1}_{\{\text{Accept}\}}|X_i, \text{Entry} = 0\} = \Phi(X\beta_b)_{Pre-entry} \qquad \text{(Pre-Entry Screen)}$$

$$\mathbb{E}\{\mathbb{1}_{\{\text{Accept}\}}|X_i, \text{Entry} = 1\} = \Phi(X\beta_a)_{Post-entry} \qquad \text{(Post-Entry Screen)}$$

where X contains borrowers' public observables, monthly fixed-effects at issuance, borrowers' address fixed-effects at the state level and loan purpose fixed effects. I conduct in-sample predictions with respect to the point estimates, $\hat{\beta}_a$ and $\hat{\beta}_b$, separated by the entry event. Within each predicted sample, I fit a local Epanechnikov polynomial of borrower acceptance propensities against days until the event (or after the event). [65] To render it "local," I use 15-day windows on both sides of the entry event.

Figure 4 shows the discontinuity plot at the time of the entry event, represented by the vertical red line. Each "gray" dot represents a predicted propensity of acceptance for a loan applicant. The bandwidth of each polynomial represents its 95% confidence interval. The discontinuity is statistically significant, and the increment of accepting propensity is more than 4%, compared to a 6% average acceptance rate prior to the entry event.

Figure 4 Here

However, the SRD method implicitly assumes that the event itself does not induce discontinuity on borrower quality. Next, I verify that borrowers' creditworthiness does not show discontinuity at the event. I use borrower screening policy prior to the entry event, $\hat{\beta}_b$, as a standard measure of creditworthiness and predict borrowers' acceptance propensity in both samples. With a similar discontinuity setup, I plot local polynomial fits of

[65] For example, for loans submitted 5 days prior to the entry, the value on the timeline is -5. For those submitted 10 days after the entry, the value is +10.

average applicant's creditworthiness against the number of days until and after the event. In Figure 5, I do not observe a significant discontinuity of borrowers' creditworthiness, indicating that the estimated result above is robust.

Figure 5 Here

The platform classifies the accepted borrowers into 35 categories, A1, A2, ..., G5. To ensure adequate statistical power, I reduce the number of bins to 7, ordered by creditworthiness from A to G. I adapt the methodology applied by most credit rating studies (see Becker and Milbourn (2011)). Rating categories are monotonically ranked in borrowers' creditworthiness, and thus I use an ordinal linear model to estimate borrower classification:

$$\text{Grade} = \begin{cases} A & \underline{x_A} \geq X\beta + \mu_i \\ B & \underline{x_B} \geq X\beta + \mu_i > \underline{x_A} \\ C & \underline{x_C} \geq X\beta + \mu_i > \underline{x_B} \\ D & \underline{x_D} \geq X\beta + \mu_i > \underline{x_C} \\ E & \underline{x_E} \geq X\beta + \mu_i > \underline{x_D} \\ F & \underline{x_F} \geq X\beta + \mu_i > \underline{x_E} \\ G & X\beta + \mu_i > \underline{x_F} \end{cases} \tag{Credit Rating}$$

Note that in general, θs can be heterogeneous for different outcomes.[66] I respectively estimate ordinal linear models over pre-entry and post-entry accepted borrowers, and compare the cut off points, $\underline{x_i}$, where i={A,B,C,D,E,F}.[67]

[66]Other approaches including ordered probit, OLS or ordered logistic Models are also experimented with for robustness.

[67]Note that 7 categories yield only 6 cutoff points. Second, preliminary plot 7 shows that the interest rate jump occurred on August 1, 2010, compared to the event date, July 13, 2009. For similar reason,

Figure 7 Here

Table 10 compares pre- versus post-entry point estimates of the cutoff thresholds. The first 2 columns are estimates from the pre-entry sample, and the next 2 columns are those using post-entry data. A/B's threshold, or in the equation above $\underline{\hat{x_A}}$, marks at -73 for pre-entry and -64.8 for post-entry. If a pre-entry borrower's quality $X\beta$ is less than -73, she would be categorized as a Grade A loan. For a post-entry borrower, her rating would be A if the value of $X\beta$ is less than -64.8.[68] I pairwise compare pre- and post-entry thresholds using one-sided t-tests. I show that all of the post-entry cutoff thresholds become more relaxed, indicating credit inflation.

Table 10 and Figure 6 Here

To apply SRD, I use the point estimates to conduct in-sample predictions of the probabilities pre- and post-entry. With 7 outcomes, each borrower is associated with 7 predicted probabilities, with sum equal to 1. For each observation, I calculate the sum the predicted probabilities of outcomes A and B, denoted by "Creditworthy Rating Propensity." The sum of the remaining probabilities is denoted as "Risky Rating Propensity." The goal here is to show a discontinuity on "Creditworthy Rating Propensity" between pre- and post-entry.[69] Similar to the procedure described above, I fit two local polynomials with 30-day windows before and after the event.[70] Figure 6 shows that the "Creditworthy Rating Propensity" appears to experience a jump at the entry event, with a magnitude

I experiment with the discontinuity design with two different dates and choose the most salient result. Result salience provides robustness, since discontinuities should only appear on the date when it occurred. Otherwise, it would raise concerns about the results' robustness.

[68]Note that borrowers' quality is in a inverse relationship with $X\beta$ because borrowers' creditworthiness here is ranked in an ascending order.

[69]This is similar if I replicate it with "Risky Rating Propensity," since the sum of the two equals 1.

[70]Note that the reason that 15-day windows are no longer sufficient is that the number of accepted loans is much smaller than that of applications.

close to 10%. This observation indicates that, relative to "Risky Rating Propensity," a borrower is 10% more likely to be categorized as "creditworthy."[71]

Since I do not observe borrower information precision, I use their ex post loan performance to recover the information precision. In particular, I run a regression using equation (1), controlling for all public observables and fixed effects. Using the point estimates, I obtain the residuals, $\hat{\varepsilon}_i$. I argue that information precision can be measured by the heteroskedasticity of the regression. That is, given borrowers' information X and Z, if X and Z are more precise, they have higher explanatory power on loan performance Y. Therefore, the residuals have less variance. Note that the measure of precision is not about good signal versus bad signal. It is about the information content of a signal, or how much it reflects the borrower's type.

Figure 3 shows the distributions of residuals separated by loan Grade. The upper (lower) adjacent value for each Grade represents the 95th (5th) percentile. For each Grade, the box between the upper adjacent value and lower adjacent value represents the range from the 25th percentile to the 75th percentile, with the median inside the box. These graphs show that the residual distributions for creditworthy ratings are much more centered than those for risky ratings. Therefore, better ratings also contains less noisy information. Based on Figure 3 and the last hypothesis, I hypothesize that the post-entry interest rates on creditworthy ratings decrease, while those on risky ones increase. A preliminary plot 2 shows that post-entry creditworthy (risky) borrowers obtain less (more) expensive financing with respect to loan ratings. However, the accepted borrower pool has become more heterogeneous and the rating scheme differs. To show this outcome with

[71] I do not show the case for "Risky Rating Propensity" since these dual measures are relative to each other.

rigor, I borrower a semi-parametric approach from Firpo (2007) and estimate the quantile treatment effects (QTEs) of entry on the interest rates. Similar to a the propensity score matching method, this approach evaluates the borrower's propensity scores at given quantiles of interest rates and compares the "same" types of borrowers' interest rates before and after the entry event. To show the heterogeneous effects, I estimate QTE at each decile of borrower interest rates as a measure of borrower creditworthiness.[72]

Table 11 shows the estimated treatment effects on interest rates measured by investors' monthly internal rate of returns. The result shows that the most creditworthy borrowers obtain a 0.06% less expensive monthly interest rate, while the riskiest borrowers obtain loans that are 0.05% more expensive. Borrowers of average creditworthiness do not seem to obtain different interest rates in terms of their economic magnitude or statistical significance.

Table 11 and Figure 3 Here

6.3 Ex Post Loan Performance

In equilibrium, if undercutting and adverse selection deterrence attracts a sufficient number of good borrowers, they should be able to maintain loan performance, compared to loans originated pre-entry. Knowing that borrowers' information is imprecise for risky borrowers, the platform raises interest rates because it expects a lower yield from risky borrowers.

Hypothesis 6. *Loan performance is maintained for creditworthy borrowers, but deteriorates for risky borrowers.*

[72]For robustness, I also experiment with an out-of-sample validation approach, in which I use the estimated pre-entry pricing mechanism to forecast out-of-sample interest rates on post-entry borrowers.

6.3.1 Empirical Strategy

To recover the possible heterogeneous effects, I again separate the samples into creditworthy and risky borrowers by their loan ratings, where creditworthy borrowers have loan Grade A and B. The remainder are risky borrowers. I apply quantile treatment effect models separately for each group. I bilaterally match propensity scores generated from characteristics and loan contracts between the pre-entry and post-entry samples. I evaluate the difference in performance at the first, second and third quartiles, i.e. the 25th, 50th and 75th percentiles, respectively.

Table 12 shows that loan performance measured by ROI does not show significant reduction within the creditworthy ratings. However, within the risky ratings, their payoffs to investors are significantly less at the 1st quartile and at the median. This finding indicates that, compared to pre-entry, bad borrowers are less likely to repay, resulting in a 5% reduction at the 25th percentile level and a 2% reduction at the median, which provides evidence that although private information aggregation deters adverse selection, the platform cannot screen perfectly, bearing adverse selection under borrower rate shopping.

7 Conclusion and Extension

In this paper, I study how the platform incorporates private information that cannot be conveyed explicitly to investors and gains a competitive advantage against the auctioneer. I find that by doing so, the platform deters adverse selection and becomes less conflicted in competing for creditworthy borrowers.

There are questions in multiple directions that cannot be fully addressed in this paper.

I am addressing them in this section. First, this paper does not study the platform's interest misalignment with its investors. As investors play the most important role in market clearing, I must show that the adverse selection deterrence benefits investors. In the robustness section, I show that investors can "punish" the platform by plausibly exiting the market. As investors receive smaller revenue streams from the loans, credit crunches emerge, and it takes longer for loans to be cleared off the market. (see appendix and Table 13)

Second, this paper argues that rate shoppers choose the lower interest rates of the two lending platforms. In addition, borrowers also care whether their loans will be financed. Additionally, the platform gains its competitive advantage by committing to financing unfunded loans. This action provides signals to both investors and borrowers. By getting its "skin in the game," the platform can maintain trust from investors. By committing to unfunded loans, it signals to borrowers that their loans are guaranteed to be cleared off the market as long as they can pass the initial screening.

8 Tables and Graphs

8.1 Summary of Estimation Sample

Panel A: Summary Statistics of Loan Applicants

	Pre-entry			Post-Entry		
	Mean	SD	Median	Mean	SD	Median
Loan Size (1K)	9.12	5.57	8.00	10.34	6.35	9.00
Funding Duration	10.70	3.91	11.00	8.80	4.02	8.00
Platform Percentage	15.32	21.95	3.80	2.07	5.36	0.66
Number of Investors	102.61	57.75	92.00	136.71	71.19	130.00
DTI	3.61	17.85	0.12	3.47	17.52	0.17
Employment Length	3.88	3.44	3.00	4.14	3.56	3.00
Interest Rate (IRR%)	1.04	0.21	1.04	1.04	0.23	1.04
FICO	713.57	35.32	710.00	716.38	36.95	710.00
Revolving Utilization	0.03	0.13	0.00	0.05	0.18	0.00
Revolving Balance (1K)	15.81	33.65	8.65	16.99	31.09	8.69
Open Accounts	9.28	4.29	8.00	9.25	4.63	8.00
Total Accounts	20.87	11.38	19.00	21.69	11.90	20.00
Public Record	0.05	0.24	0.00	0.05	0.22	0.00
Annual Income (1K)	65.16	48.33	55.00	72.06	70.78	58.80
Delinquency	0.12	0.39	0.00	0.14	0.47	0.00
1(Borrower's Description)	1.00	0.02	1.00	0.95	0.22	1.00
Description Length (Standardized)	1.21	446.15	-135.65	-0.73	483.10	-145.66
BA Attainment	33.28	10.43	31.08	32.75	10.16	31.02
Deposit ($ Million)	20.00	52.73	2.09	21.66	58.19	1.96
% Nonpayment	7.95	22.20	0.00	7.30	21.29	0.00
ROI%	8.14	26.28	15.87	8.59	25.27	16.01
Default	0.14	0.35	0.00	0.13	0.34	0.00

Panel B: Summary Statistics of Loan Applicants

	Pre-entry			Post-Entry		
	Mean	SD	Median	Mean	SD	Median
Accepted	0.06	0.24	0.00	0.11	0.32	0.00
Requested Size (1K)	9.88	7.19	8.00	10.88	7.92	9.60
Employment Length	3.09	3.39	2.00	3.50	3.51	2.00
FICO	578.12	174.64	624.50	604.43	172.26	658.00
BA Attainment	29.81	9.39	29.64	30.52	9.51	30.60
Deposit ($ Million)	14.68	41.61	1.74	16.61	48.05	1.73

Table 1: Lending Club: Summary Statistics of Entry-Estimation Sample

This table shows summary statistics of loan borrowers (Panel A) and applicants (Panel B) on Lending Club in the sample period around the entry event. In particular, a 5.5-month window is restricted on each side of the entry event, yielding a time span from February 2009- December 2009. Statistics are separated by the entry event and include sample mean, standard deviation, median. The total number of applications before (after) the entry event is 31,090 (27,855). The number of accepted loans is 1,899 (3,149) before (after) the entry event. Note that I only observe limited information on loan applicants.

Panel A: Summary Statistics of Loans Issued					
	Mean	SD	Min	Max	Median
Loan Size (1K)	4.38	4.09	1.00	25.00	3.00
Number of Investors	147.03	153.59	1	1189	94
DTI	9.83	29.43	0.01	100	0.23
Employment Length	4.99	3.61	0	10	4
Interest Rate (IRR%)	1.61	0.76	0.35	2.92	1.52
FICO	700.85	53.15	600	778	702
Revolving Utilization	0.5	0.3	0	1	0.49
Revolving Balance (1K)	20.2	35.03	0	526.58	9.97
Open Accounts	8.86	4.8	0	33	8
Total Accounts	6.72	4.37	0	31	6
Public Record	0.19	0.53	0	6	0
Annual Income (1K)	61.28	38.6	0	425	54
Delinquency	0.69	2.16	0	28.29	0
Description Length (Standardized)	47.66	464.28	-364.76	2439.81	-138.89
BA Attainment	30.99	9.65	10.28	60.69	30.72
Deposit ($ Million)	12.95	42.12	0.03	421.19	1.81
% Nonpayment	7.89	22.48	-38.07	100	0
ROI%	15.69	31.37	-100	143.26	17.1
Default	0.15	0.36	0	1	0

Panel B: Summary Statistics of Loan Applicants					
	Mean	SD	Min	Max	Median
Accepted	0.14	0.35	0	1	0
Requested Size (1K)	7.39	6.18	1	25	5
DTI	21.87	41.04	0	100	0.32
Employment Length	5.01	3.64	0	10	4
FICO	688.21	48.54	600	778	665
Revolving Utilization	0.55	0.31	0	1	0.58
Revolving Balance (1K)	26.08	48.42	0	985.3	10.71
Open Accounts	8.79	5.05	0	40	8
Total Accounts	6.73	4.61	0	42	6
Public Record	0.26	0.64	0	10	0
Annual Income (1K)	65.28	415.12	0	50100.15	50.2
Delinquency	0.96	2.48	0	28.29	0
Description Length (Standardized)	0	441.06	-364.76	2514.87	-183.01
BA Attainment	30.42	9.43	10.17	60.69	30.64
Deposit ($ Million)	12.79	40.98	0.03	421.19	1.71

Table 2: Prosper: Summary Statistics of Entry-Estimation Sample

This table shows summary statistics of borrowers financed (Panel A) and borrowers listed (Panel B) on Prosper in the sample period after the entry event. Statistics include sample mean, standard deviation, minimum, maximum and median for the "entry sample period" between July 2009 to December 2009. In this period, Prosper receives a total of 15,120 listings and 2,141 issued loans.

	Loan Grade	A	B	C	D	E	F	G
Lending Club	Interest Rate (APR)	∼7.37%	∼11.13%	∼13.69%	∼15.93%	∼17.94%	∼20.32%	∼23.52%
	Return on Investment (ROI)	6.82%	9.50%	10.49%	11.09%	13.33%	14.03%	12.16%
	Origination Fee	1∼3%	4∼5%	5%	5%	5%	5%	5%
	Loan Grade	AA	A	B	C	D	E	HR
Prosper 2.0	Interest Rate 1 (APR)	8.37 %	10.30 %	14.61%	20.29%	26.18%	31.95%	32.38%
	Interest Rate 2 (APR)	7.82%	11.47 %	16.43%	19.84%	26.04%	31.22%	31.80%
	Origination Fee	0.5%	3%	3%	4.5%	4.5%	4.5%	4.5%
	Return on Investment (ROI)	7.64%	8.68%	12.22%	17.77%	17.34%	17.96%	20.88%

Table 3: Annual Interest Rates and Origination Fees from 2009-2011

In this table, I compare interest rates and origination fees between Lending Club and Prosper in the sample period. Note that the interest rates from Prosper stem out of two mechanisms. From July 13, 2009—December 20, 2010, the interest rates for Prosper are market-clearing prices under the auction mechanism. Since December 20, 2010, interest rates are under the platform-pricing mechanism.

	A	B	C	D	E	F	G
2009	1204	1456	1353	817	313	105	57
	8723.2	15914.9	13282.0	8732.4	3728.4	1326.8	704.2
2010	2931	3708	2748	1899	978	317	132
	23707.4	38610.0	26750.9	20530.6	11441.6	4492.6	2283.4
2011	5606	6248	3712	2603	1595	664	188
	50128.2	68531.0	43849.5	33748.1	27457.2	12767.5	3978.8
	AA	A	B	C	D	E	HR
2009	319	504	107	422	359	205	225
	2061.01	2495.31	654.23	1680.29	1312.14	498.00	687.20
2010	646	962	556	588	1340	650	898
	4404.60	5605.15	4216.04	2409.15	5394.89	2341.58	2736.14
2011	545	1454	1597	875	3263	2303	1155
	5615.28	13484.54	10633.70	6815.83	19898.19	11855.79	3995.12

Table 4: Loans (Proceeds in $1K) Issued between Lending Club and Prosper, 2009-2011

This table shows a comparison between the two platforms on the number of issuance and loan proceeds from 2009-2011. The table above represents Lending Club separated by its loan Grades and the table below represents Prosper separated by Prosper's loan Ratings.

	Prosper	2009/07-2009/12		2010/01-2010/12		2011/01-2011/12
Lending Club	Unfinanced	Financed	Unfinanced	Financed	Unfinanced	Financed
Reject	4,019	477	8,322	1,707	5,312	6,137
Accept	70		181		97	

Table 5: Rate Shopper Summary

The table shows "fuzzy-matched" borrowers that applied on both platforms separated by year. I only document 348 shoppers who successfully obtained financing from Lending Club. The average monthly interest rate is 1.61% and return on investment is 13.72% on Prosper. Among the 343 borrowers who are accepted by Lending Club and not fully funded by investors on Prosper, their default rate is 18.8%, 4% higher than the platform average. For the borrowers who are rejected by Lending Club and receive a loan contract on Prosper, their default rate is 21%, 2% higher than the platform average.

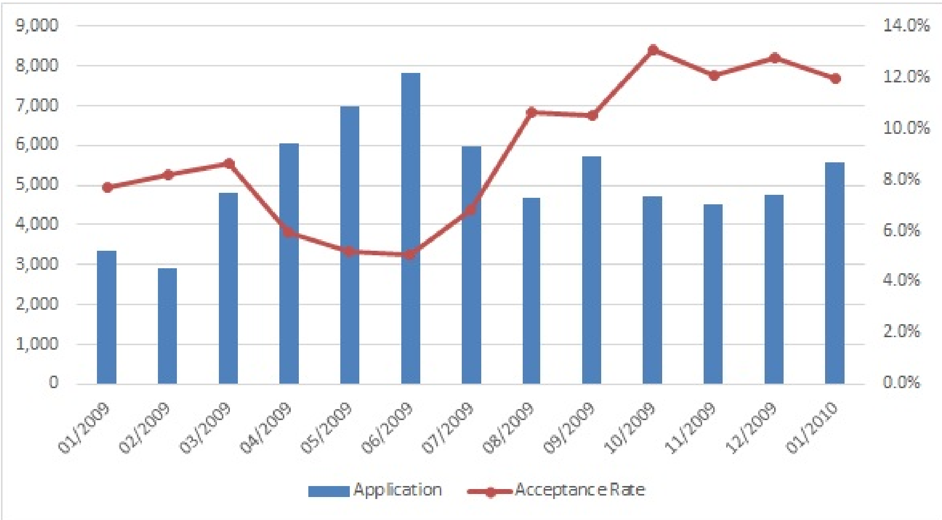


Figure 1: Entry, Applications and Acceptance

This figure shows Lending Club's number of applications (on the left Y-axis) and borrower acceptance rate (on the right Y-axis) from January, 2009- January, 2010, including the estimation sample period. This figure provides two purposes. First, it shows that the number of applications drop significantly after the entry event (See the difference between June and July.). However, the acceptance rate increased from below 6% to above 10% shortly after the entry event. This indicates that the lending standard has changed. Second, the declining number of applications in data does not reflect the big picture, because it does not show borrowers who have withdrawn loans from Lending Club for cheaper loans on Prosper. This means that the number of applications is underrepresented in the data.

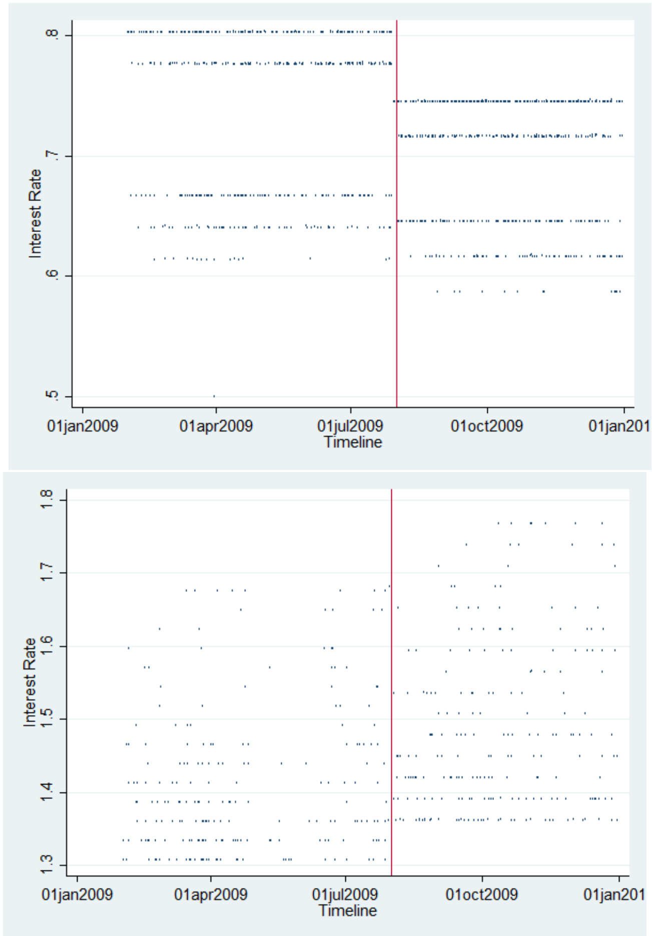


Figure 2: Entry and Interest Rates Visualization

The upper graph shows interest rates over time for Grade A loans and the lower plot for Grade E, F & G (riskiest borrowers). The vertical lines represent the discontinuity happening on August 1, 2009, two weeks after the entry event. Judged by ratings only, creditworthy borrowers obtain cheaper financing after the entry event, and risky borrowers obtain more expensive interest rates.

8.2 Estimation Results

	OLS		
	(1) Default	(2) ROI%	(3) Nonpayment%
BA Attainment	-0.0419^{***}	0.0312^{***}	-0.0376^{***}
	(0.0003)	(0.0241)	(0.0215)
Deposit	-0.0083^{**}	0.0078^{**}	-0.0101^{***}
	(0.0000)	(0.0000)	(0.0000)
Description Length	-0.0068^{**}	0.0078^{**}	-0.0072^{*}
	(0.0000)	(0.0002)	(0.0002)
Observations	37884	37884	37884
R^2	0.161	0.404	0.239
Borrower Characteristics	Yes	Yes	Yes
Contract Terms	Yes	Yes	Yes
Issued Monthly FE	Yes	Yes	Yes
Closed Monthly FE	Yes	Yes	Yes
Address State FE	Yes	Yes	Yes

Standardized beta coefficients; Standard errors in parentheses

Standard errors are clustered at borrower state level.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 6: Private Information and Performance

The table shows that these measures of private information matter to borrowers' performance to some extent. BA Attainment represents borrowers' local bachelor's degree attainment to the population ratio; Deposit indicates their local total deposits from community bank, according to the FDIC; Description Length represents borrower-specific loan description length, standardized by loan Grade and origination month.

Borrowers' performance is measured by Default (binary outcome), Return On Investments and Nonpayment (percentage of the loan not paid). To show their economic magnitude, I standardized all the coefficients. To interpret the coefficients of BA Attainment in all columns, a one standard deviation increase in BA Attainment corresponds to a 0.042 standard deviation drop in default propensity (or 1.4% drop in default propensity), a 0.03 standard deviation increase in ROI and a 0.037 standard deviation decrease in borrowers' nonpayment.

	OLS		
	(1) Interaction	(2) Pre-Entry	(3) Post-Entry
BA Attainment	0.0133***	0.0141***	0.0180***
	(0.0016)	(0.0018)	(0.0024)
Deposit	0.0006	0.0011	0.0048**
	(0.0020)	(0.0020)	(0.0023)
Entry=1 × BA Attainment	0.0056***		
	(0.0017)		
Entry=1 × Deposit	0.0046*		
	(0.0027)		
Observations	52372	26740	25632
R^2	0.195	0.264	0.174
Borrower Characteristics	Yes	Yes	Yes
Issued Monthly FE	Yes	Yes	Yes
Address State FE	Yes	Yes	Yes

Standard errors in parentheses

Standard errors are robust to heteroskedasticity.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 7: Private Information and Borrower Screening

This table compares how those measures matter differently in borrower screening. Column (1) shows a pooled regression in which I use the entry dummy to interact with each of the covariates and compare the magnitudes. Columns (2) and (3) show separate estimates from pre- and post-entry samples, and pairwise compare the coefficients. To better interpret the Linear Probability Model (LPM) estimates, I standardize independent variables. For example, in Column (1), a one standard deviation increase in BA Attainment corresponds to 1.33% increase in acceptance propensity prior to entry. Its magnitude increases by 0.56% after the entry event.

	Pre-Entry	Post-Entry
Deposit		
A	0.0020	0.0037***
	(0.0029)	(0.0010)
B	0.0008	0.0019***
	(0.0011)	(0.0005)
C	−0.0001	−0.0010***
	(0.0001)	(0.0003)
D	−0.0011	−0.0027***
	(0.0016)	(0.0007)
E	−0.0009	−0.0012***
	(0.0013)	(0.0004)
F	−0.0004	−0.0004***
	(0.0005)	(0.0001)
G	−0.0003	−0.0003***
	(0.0004)	(0.0001)
Description Length		
A	0.0048	0.0048***
	(0.0044)	(0.0018)
B	0.0019	0.0025***
	(0.0017)	(0.0009)
C	−0.0002	−0.0012***
	(0.0002)	(0.0005)
D	−0.0027	−0.0035***
	(0.0024)	(0.0012)
E	−0.0022	−0.0016***
	(0.0020)	(0.0006)
F	−0.0009	−0.0005***
	(0.0008)	(0.0002)
G	−0.0007	−0.0004**
	(0.0006)	(0.0002)

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 8: Private Information and Credit Rating, Marginal Effect

This table shows the model-predicted average marginal effect of Description Length and Deposit associated with the Ordinal Linear Models. BA Attainment does not show significant effects on any outcome, and thus is omitted from reporting. For example, for a one standard deviation increase in Deposit, I would expect that the propensity of the rating outcome of Grade A increases by 0.2% pre-entry but 0.37% post-entry. Since the probability measures are relative, if the marginal effects on creditworthy ratings are positive, those on risky ratings must be negative.

	OLS		
	(1) Default	(2) ROI%	(3) Nonpayment%
1{Shopper}	0.0654^{***}	-4.0036^{***}	3.3958^{***}
	(0.0154)	(1.1299)	(0.9199)
Observations	2141	2141	2141
Borrower Characteristics	Yes	Yes	Yes
Contract Terms	Yes	Yes	Yes
Issued Monthly FE	Yes	Yes	Yes
Closed Monthly FE	Yes	Yes	Yes
Address State FE	Yes	Yes	Yes
R-Squared	0.2006	0.4268	0.2606

Standard errors in parentheses

Standard errors are robust to heteroskedasticity.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 9: Prosper: Adverse Selection and Performance

This table tests shopping-induced adverse selection using data from Prosper in the sample period. Compared to the borrowers who do not show up on Lending Club, the borrowers who were previously rejected by Lending Club on average yield a 6% higher default rate, 4% less ROI and 3.4 % more nonpayment. This finding shows significant deterioration of loan performance on Prosper induced by adverse selection deterred from Lending Club.

	Pre-Entry Threshold	SE	Post-Entry Threshold	SE	Pre − Post Difference	T Score
A/B	-73.00	2.71	-64.80	2.15	-8.20***	2.37
B/C	-70.04	2.68	-60.99	2.09	-9.05***	2.66
C/D	-67.13	2.62	-58.36	2.02	-8.77***	2.65
D/E	-64.90	2.55	-55.92	1.98	-8.98***	2.78
E/F	-63.10	2.54	-53.91	1.92	-9.19***	2.89
F/G	-61.84	2.60	-52.11	1.84	-9.73***	3.05

Table 10: Credit Rating Cutoff Points

This table compares pre- versus post-entry point estimates of the cutoff thresholds. The first 2 columns are estimates from the pre-entry sample, and the next 2 columns are those using post-entry data. A/B's threshold, or in the equation above $\underline{\hat{x_A}}$, marks at -73 for pre-entry and -64.8 for post-entry. If a pre-entry borrower's quality $X\beta$ is less than -73, she would be categorized as a Grade A loan. For a post-entry borrower, her rating would be A if her $X\beta$ is less than -64.8. Note that borrowers' quality is in a inverse relationship with $X\beta$ because borrowers' creditworthiness is ranked in an ascending order. In the last 2 columns, I pairwise compare pre- and post-entry thresholds using one-sided t-tests. I show that all of the post-entry cutoff thresholds become more relaxed, indicating credit inflation.

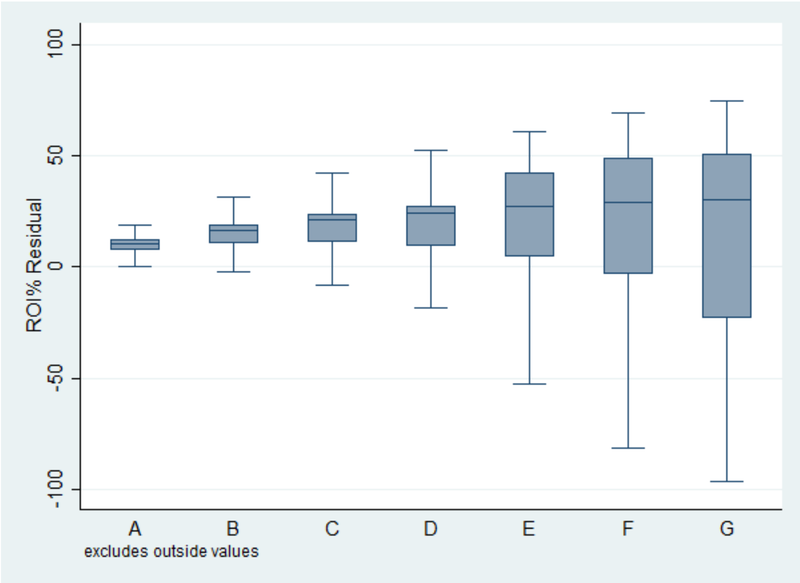


Figure 3: Public Information Precision and Creditworthiness

The purpose of this plot is to show that the precision of borrowers' public information varies with their creditworthiness. Note that the measure of precision is not about good signal versus bad signal. It is about the information content of a signal, or how much it reflects the borrower's type. The plot shows the distributions of residuals of regression (1) separated by loan Grade. The upper (lower) adjacent value within each Grade represents the 95th (5th) percentile of the distribution of the residuals for the corresponding Grade. The box between the upper adjacent value and lower adjacent value represents the range from 25th percentile to 75th percentile, with the median line inside the box. This graph shows that the residual distributions for creditworthy borrowers are much more centered than those for risky borrowers. In other words, the precision of creditworthy borrowers' information is higher.

Decile	QTE	Z-score
1	-0.0602^{***}	-7.58
2	-0.0575^{***}	-5.07
3	-0.0079	-0.72
4	-0.0026	-0.29
5	0.0001	0.01
6	0.0053	0.52
7	0.0366^{***}	3.42
8	0.0447^{***}	3.72
9	0.0525^{***}	3.42

Table 11: Entry and Interest Rate: Quantile Treatment Effects

The table shows the estimated results of the Quantile Treatment Effect model (QTE). QTE matches borrowers' propensity scores at each decile of their creditworthiness (measured by interest rates) conditional on all observable characteristics. At each decile, I compare the difference interest rates resulting from the entry event. The event study satisfies the exogenous treatment requirement by QTE. The result shows the estimated treatment effects on interest rates measured by investors' monthly internal rate of returns. The result shows that the most creditworthy borrowers obtain a 0.06% less expensive monthly interest rate, while the riskiest borrowers obtain loans that are 0.05% more expensive. Borrowers of average creditworthiness do not seem to obtain different interest rates in terms of their economic magnitude or statistical significance.

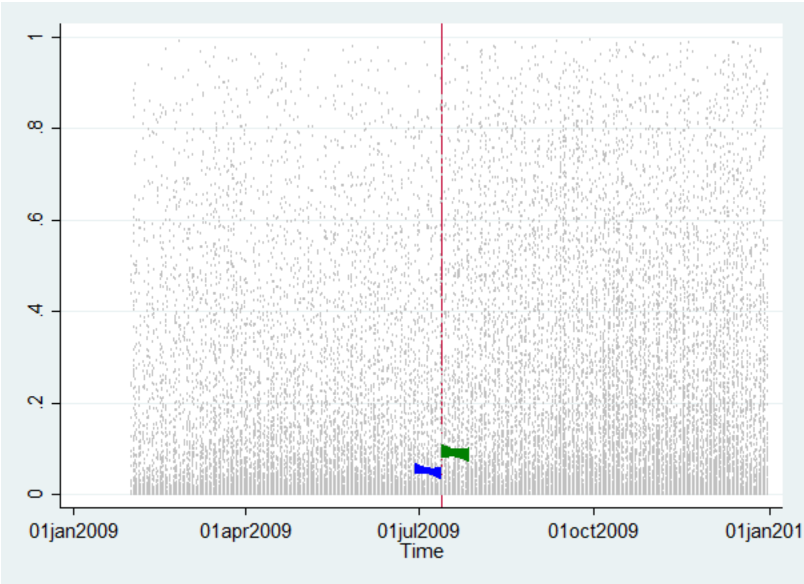


Figure 4: RDD: Loan Selection conditional on Public Information

The graph shows the discontinuity plot at the time of the entry event, represented by the vertical red line. Each "gray" dot represents a predicted propensity of acceptance for a loan applicant, only conditional on her public information. The bandwidth of each polynomial represents its 95% confidence interval. The discontinuity is statistically significant, and the increment of accepting propensity increases by more than 4%, compared to a 6% average acceptance rate prior to the entry event.

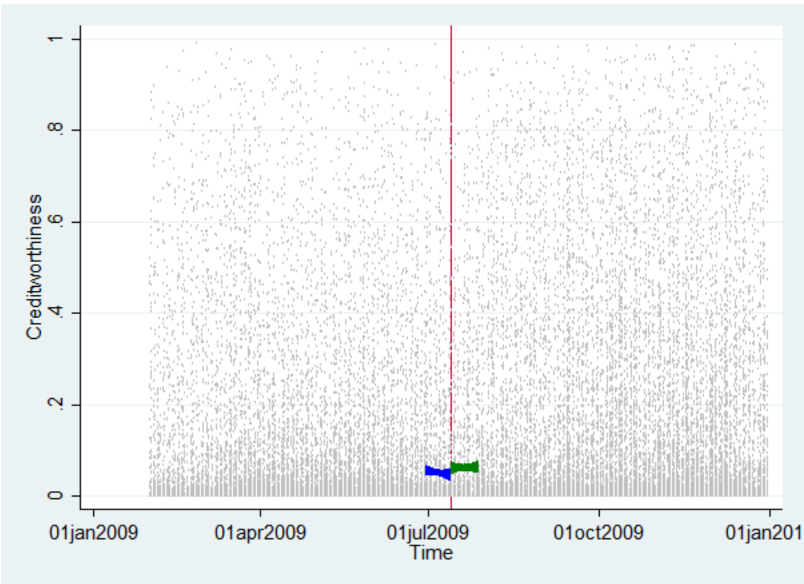


Figure 5: Robustness

As a robustness check, this graph is to show if discontinuity also appears on borrowers' creditworthiness. Using ex ante acceptance policy $\hat{\beta}_b$, I conduct in-sample predictions on loans issued before the entry event and out-of-sample predictions on those after the entry event. I do not observe significant discontinuity of borrowers' creditworthiness.

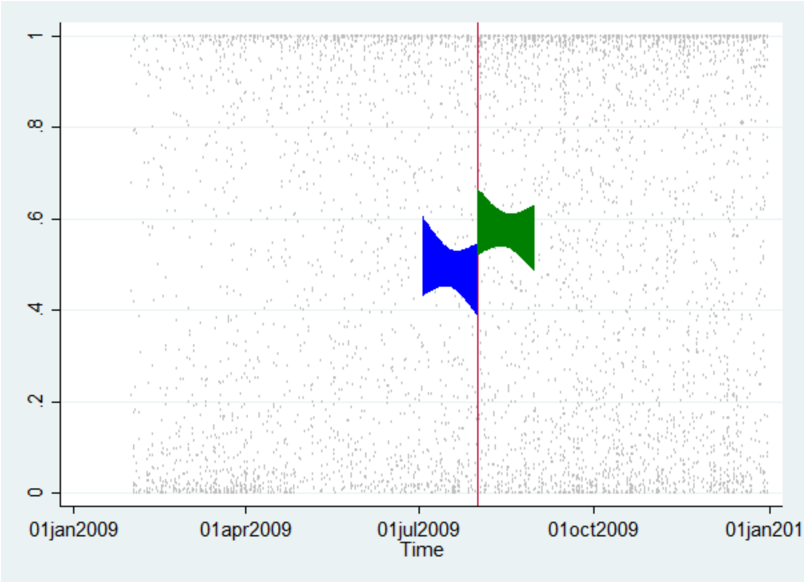


Figure 6: RDD: Credit Rating conditional on Public Information

This graph shows local polynomial fits of "Creditworthy Rating Propensity" against days until and after the entry event, with a 30-day window on each side. Note that here I no longer use 15-day windows because observations on issued loans are fewer than the number of applicants. I do not show the case for "Risky Rating Propensity," since these dual measures are relative to each other.

	Creditworthy Ratings		Risky Ratings	
	$ROI_{post} - ROI_{pre}$	Z-Score	$ROI_{post} - ROI_{pre}$	Z-Score
1st Quartile	-0.2773	-0.19	-5.7399^{***}	-2.18
2nd Quartile	-0.1956	-0.31	-2.1266^{***}	-2.63
3rd Quartile	0.3338	0.49	-0.4544	-0.8

Table 12: Quantile Treatment Effect, Performance

This table shows Quantile Treatment Effect separated by groups of Creditworthy Ratings (Grade A and B) and Risky Ratings (Grade C—G). I bilaterally match propensity scores generated from characteristics and loan contracts between the pre-entry and post-entry samples, and I evaluate the difference in performance at the first, second and third quartiles, i.e. 25th, 50th and 75th percentiles. It shows that loan performance measured by ROI does not show significant reduction within the Creditworthy Ratings. However, within the Risky Ratings, their payoffs to investors are significantly less at the 1st quartile and the median. This finding indicates that compared to pre-entry, bad borrowers are less likely to repay, resulting in a 5% reduction at the 25th percentile level and a 2% reduction at the median, which provides evidence that although private information aggregation deters adverse selection, the platform cannot screen perfectly, bearing adverse selection under borrower rate shopping.

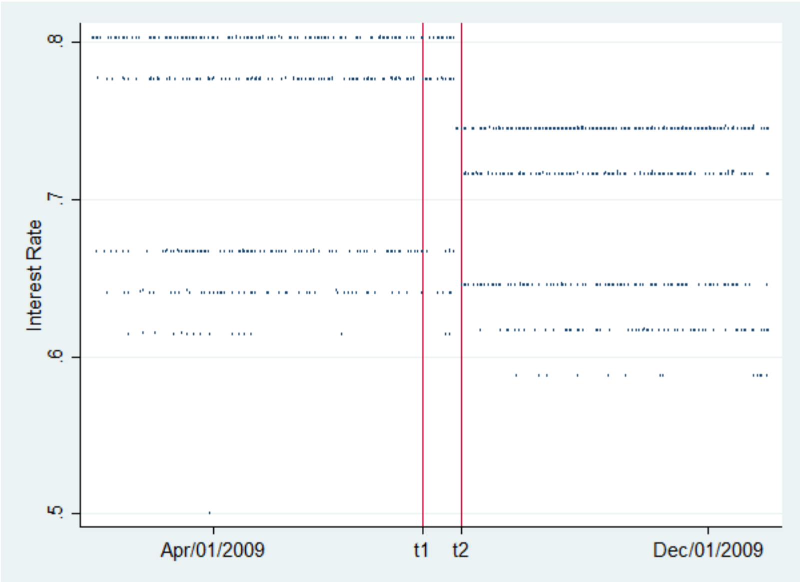


Figure 7: Entry and Interest Rates of Grade A

This figure demonstrates the interest rate with Grade A loans from early 2009 to early 2010. The first vertical line "t1" maps onto the date 7/13/2009 and the second, "t2," onto 8/1/2009. Remember that the entry announcement is on 7/13/2009, but the interest rate jump is on 8/1/2009.

	OLS	
	(1) Funding Duration	(2) Platform Pct
Observed ROI	-0.1285^{***}	-0.1166^{***}
	(0.0140)	(0.0591)
Observations	38634	38634
R^2	0.369	0.305
Borrower Characteristics	*Yes*	*Yes*
Issued Monthly FE	*Yes*	*Yes*
Address State FE	*Yes*	*Yes*

Standardized beta coefficients; Standard errors in parentheses
Standard errors are robust to heteroskedasticity.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 13: Entry and Investor's Punishment

This table shows that, observing higher performance measured by Observed ROI, investors fill the loans more quickly. In the meantime, it requires less capital committed from the platform to make the loans issue. In particular, as Observed ROI increases by 1% for the given month, the Funding Duration is expected to decrease by 0.128 days and the platform commits 0.17% less on capital investment of loans on average. This table also shows that worse "Observed ROI" reduces market-clearing efficiency: 1. It takes longer for loans to clear off the market. 2. Credit crunches emerge, and Lending Club's capital commitment increases.

Appendix A Robustness and Extension

A.1 Event Exogeneity

I examine information asymmetry between market agents on the entry event. Prosper "re-entry" may have been well anticipated by Lending Club, and thus decisions that are unobservable to borrowers and investors would have been made prior to the event. By not clocking the event discontinuity correctly, I may attenuate the result. To test if Prosper's re-entry is anticipated, I plot the incumbent's interest rate changes at the time of the event. (See Figure 7) The vertical lines pin down the actual timing of Prosper's entry on July 13, 2009 and the jump of the interest rates happens on August 1, 2009. First, I conclude that the incumbent, Lending Club, may or may not anticipate Prosper's re-entry, but the timing of the event is exogenous and unforeseen. Second, its interest rate changes did not happen until 15 days after Prosper's re-entry announcement. Consistent with the interest rate change, I set the time of the event at 8/1/2009.

A.2 Investors' Punishment

Investors clear loans off the market. It is hard to study investors' decision without observing their identities. However, we know that in a one shot game between 3 players, a borrower, a platform and an investor. If the investor has a trusting nature (see Bolton et al. (2012)), then the platform's rating or interest rate doesn't contain any information. A strong disincentive for a platform to be imprudent and inflate borrowers' credit comes from the fact that investors can punish the platform by exiting the market. To show this hypothesis, I examine market-clearing efficiency when investors observe deteriorated performance from matured loans. In particular, I study the effect of underperformance from previous borrowers on the market clearing of new loans' origination. Here, I measure market-clearing efficiency using the time duration for a loan to be funded and the Lending Club's capital provision as credit crunches emerge.[73] I measure investors' exposure to matured loans' underperformance as follow. As I observe loans' last payment dates, origination dates and performance measured by ROI, I then get to observe the revenue stream for each loan. At any given month, I compute the average ROI across loans within the same Grade who are still under payments. That is, for each month, I have a measure on the current state of return for each grade. These values are observable to repeated investors and are denoted as "Observed ROI." As an investor is exposed to worse return from failing loans, I hypothesize that the market-clearing efficiency for new loans' origination drops.

Hypothesis 7. *Worse "Observed ROI" reduces market-clearing efficiency: 1. It takes longer for loans to clear off the market. 2. Credit crunches emerge, and Lending Club's capital provision increases*

[73]The less time a loan takes to issue, the more efficient the market is.

$$\text{Funding Duration} = \text{Loan Issuance Date} - \text{Loan Application Date}$$

The less funding provided by the platform, the more efficient the market clears.

$$\text{Pct Platform} = \frac{\text{Lending Amount by Lending Club}}{\text{Total Lending Amount}}$$
$$\text{Pct Investors} = 1 - \text{Pct Platform}$$

Note that one does not observe how the platform selects to fund the borrowers. Using the ex post loan performance, I test the two opposing theories: the platform getting its "skin in the game" by making investments using own capital Vs. it signaling its market-clearing competency. With a linear probability model, I estimate the equation below:

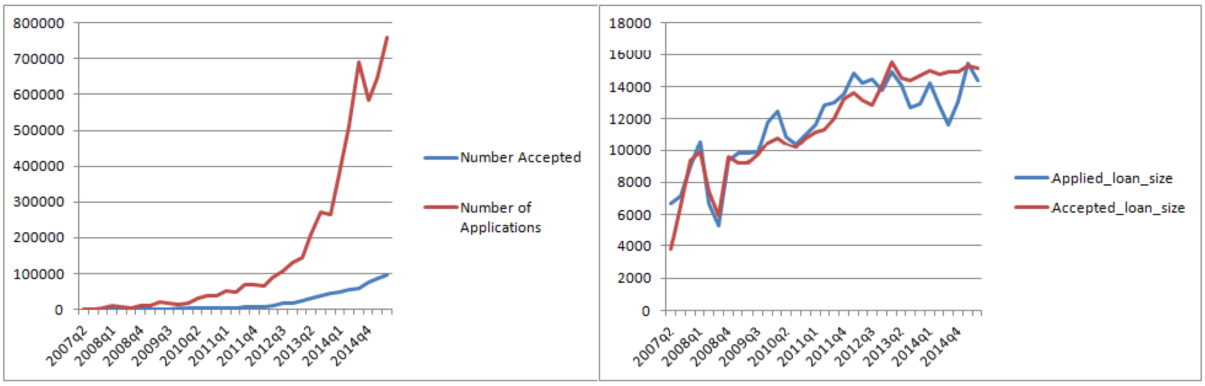
$$\mathbb{1}_{\{\text{Default}\}} = X_i \cdot \theta + \delta \text{Pct Platform} + \varepsilon \tag{2}$$

If $\delta > 0$, we can conclude that the correct theory is the latter one.

Appendix B Other Institutional Details

B.1 Lending Club

Founded in 2007, Lending Club issued 646,389 among 5,317,010 loan applications by the end of 2014. Without collateral, Lending Club targets 'prime' borrowers with FICO scores above 640. An algorithm-based screener inputs several categories of information. Some are self-reported such as age, address, income, employment length and homeownership, and may require verification by additional documents. Some are prompted for the intended loan such as loan size request, loan purposes and loan term. Others are pulled from a credit reporting agency that include FICO scores, total account, first credit line, revolving utilization and balance, total debt, credit history and public record such as delinquencies and defaults. Then, for the intended loan, it outputs a rating and an interest rate (or rejection).[74] On a loan listing, an individual investor observes features



The number of loan applications on Lending Club grew since founded in 2007. Once it reached 100,000 per quarter in early 2012, the number jumps to 700,000 per quarter in 2014. However, the number of loans being accepted has been growing at a much steadier rate and reached about 100,000.
The average requested and originated loan sizes appear to grow in the same pattern to $15K.

Figure 8: Left: applications vs origination. Right: Loan size requested and issued (in $)

including contract details, borrower credit history and instantaneous information on the

[74] A loan can either have a term of 3 or 5 years. Short-term loans are available on Prosper.com (see Figure 8)

funding status of the loan, which is captured by the current funded amount and the instantaneous number of investors who pleaded on the loan. Contract description shows the borrower's requested loan size, the loan term, its listing expiration date, intended purpose, rating and interest rate.[75] Observable borrower characteristics include her 3-digit ZIP code, employment length, employment title and annual income. Borrower credit history features her debt-to-income ratio, recent FICO score range, delinquency record within the last two years, credit card revolver balance and utilization and default history. For loans issued before Sep, 2009 investors observe some descriptions entailing the borrower's usage of the loan, her current financial situation and a Q&A between the borrower and other investors.[76] Data show that absences of loan descriptions or descriptions with 10 characters or less largely emerge in the 4th quarter of 2009. Anecdotal research also shows this phenomenon (see Figure 9).[77] For loans with vintage prior to 2010, the net annualized return stays between 5% to 7% across all the loan grades. Figure 10 illustrates the contractual vs. actual returns measured by the internal rate of return (IRR) for loans with vintage between 2008 and 2010 aggregated at monthly level.

[75] All loans that were issued before 2010 only had 3-year maturity.

[76] For example, one borrower elaborated "I am applying for this loan because I am trying to lower my credit card so I can start saving up some money. I graduated college two years ago and have had my current job for about a year and a half. I just moved home, (so no rent/bills- thanks mom and dad!) and I don't have a whole lot of other expenses. With some frugal months I could have this paid off, but I am getting married in about ten months and have been slammed with deposits, and a big dental bill for $3,000, virtually eliminating my savings. My parents are paying for the "big stuff" for the wedding, but I have been picking up the deposits. So, I am not in any way concerned with having to pay a few hundred dollars a month, I just would like to not be paying that high interest rate and would like to be saving some money on my end. No credit card debt with a steady amount at a lower interest rate is what I am hoping for. A monthly payment would be easily managed." See https://www.lendingclub.com/browse/loanDetail.action?loan_id=364451

[77] The reason for this event was not disclosed by Lending Club. I contend this incident is related to Prosper's re-entry. Data shows that the propensity of a borrower's comments or describe her loan went down to nearly 50% in 2010.

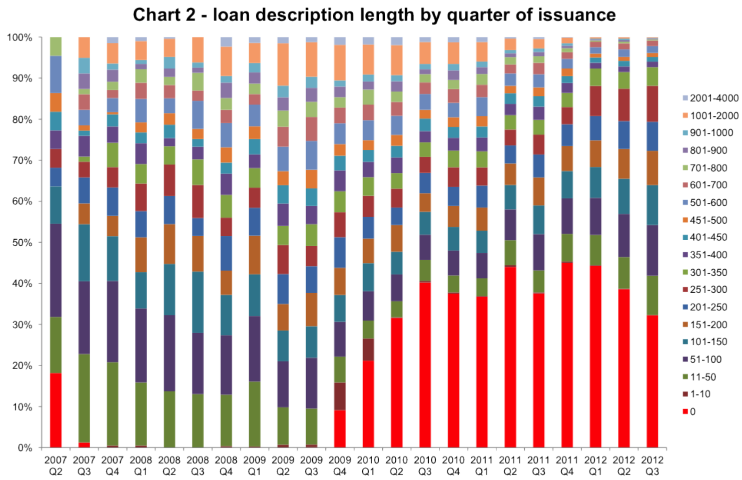


Figure 9: Loan Descriptions Length (in characteristics) with Vintage 2007-2012

Figure 9 is borrowed from an article describing research conducted by Sam Kramer on Lending Academy. http://www.lendacademy.com/lending-club-loan-descriptions-1/. While every loan originated from early 2008 to the 3rd quarter of 2009 is required to carry some description about the loan from the borrower, this requirement disappears in 2009. In late 2010, 40% of the loans do not contain any information directly from the borrower.

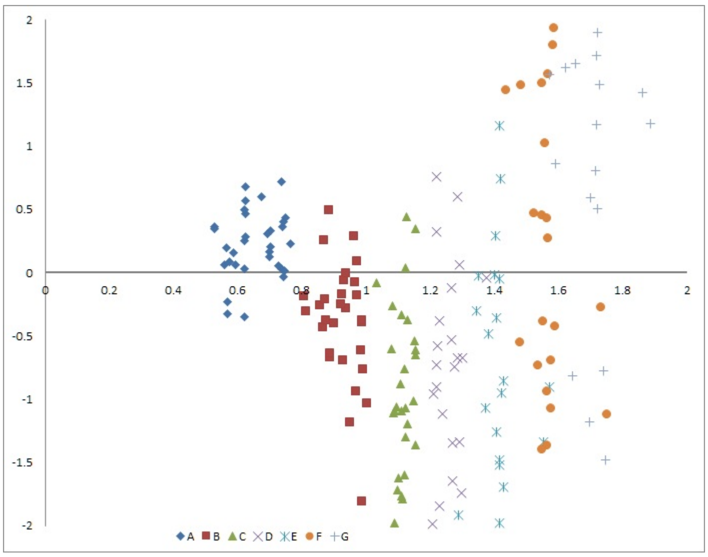


Figure 10: Lending Club: IRR for Loans Vintaged from 2008-2010

The X-axis: interest rates (IRR) of loans averaged within each rating within each origination month.
The Y-axis: their performance measured by IRR.
Note that for loans with no or little repayment, the IRR are highly left skewed. Therefore, the average for loans with risky ratings can still be negative.

B.2 Prosper 1.0

Founded two years before Lending Club in 2005, Prosper debuts peer-to-peer lending in U.S. under an auction model, while accepting borrowers with any credit background. Prior to Oct, 2008, Prosper had granted 28,936 loans with 18,480 fully paid off and 10,456 loans defaulted, consisting of total loan volume of \$178K, \$47K of which was written off, implying a loss rate of 26.1%. The auction mechanism works as follows. Borrowers put down reservation interest rates. Prosper acquired the borrower's credit reports and posted them online for a 7-day open-bid multi-unit uniform-price auction with reservation price. Investors (bidders) specified the amount and the interest rate bids. Lending position are ranked in a descending order by their interest rate bids. Once the pledged amount exceeds the requested loan size, the lowest winning bid is the ongoing interest rate for the loan. If the loan is not fully funded by expiration, the ongoing interest rate is the borrower's reservation price. .[78] Prosper 2.0 announced that it would only accept borrowers with

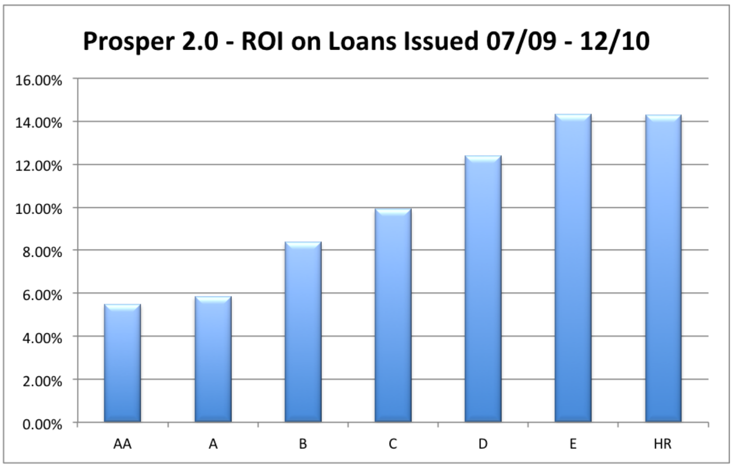


Figure 11: Return on Investments Prosper 2.0

FICO above 600 and started classifying borrowers into different risk ratings ranging from

[78]Prosper 2.0 is much improved compared to Prosper 1.0. (See figure 11, from A Look Back at Prosper 1.0? How Relevant are the Numbers? Lending Academy)

AA to HR by its evaluation of borrowers' creditworthiness, so that investors can better understanding the default risk.[79] Similar to Lending Club, debt-consolidation is the main reason for loan request. Other purposes including home improvement and small business are also quite popular.

At the time of its re-entry, Prosper and Lending Club were almost identical except several key discrepancies as follows. Foremost, still under the auction business model, investors on Prosper placed bids on the interest rates, and thus did not observe the final interest rate until the loan was issued.[80] Both borrowers and investors were and had always been price takers on Lending Club. Second, Prosper's address information was at city level whereas Lending Club was at county level. Third, as aforementioned, the FICO scores on Prosper and Lending Club came from different agencies. More than what was observable on Lending Club, a investor can observe if the borrower was a repeated borrower on Prosper.[81] Regardless of the differences between them, preliminary results show that Prosper's market re-entry tightens the competition with Lending Club.

B.3 Prosper 2.0 vs Lending Club

While both Lending Club and Prosper screen borrowers using their credit reports, they have very different borrower selection and pricing mechanisms on the interest rates. Lending Club in general accepts borrower with higher FICO scores and lower debt-to-income ratios. After accepting the borrowers, Prosper does not price loan interest rates, but it uses an auction business model where borrowers provide reservation interest rates and

[79]'P2P lender Prosper is back and better than ever', AOL Finance

[80]In Dec, 2010, Prosper got rid of the auction business model and switched to the posting-interest-rate business model as Lending Club, and the listing expiration for a loan increased from 7 days to 14 days. This event was studied by Wei and Lin (2016).

[81]Later on, Lending Club also added this feature.

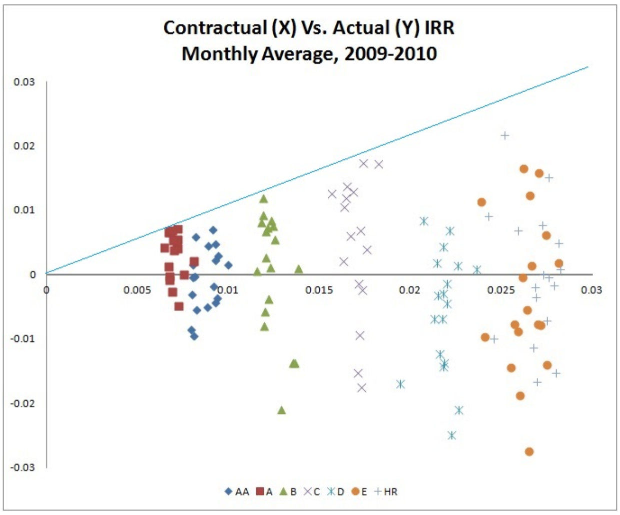


Figure 12: Prosper 2.0: IRR for Loans Vintaged from 2009-2010

investors make bid offers.[82]

The funded rate on Prosper is much lower than that on Lending Club for two reasons. First, Lending Club did provide capital to some borrowers, making loans with credit crunches issued. Second, historically, Prosper was not able to maintain strong a reputation among investors. Prior to Oct, 2008, Prosper 1.0 granted 28,936 loans with 18,480 fully paid off and 10,456 loans defaulted, consisting of total loan volume of \$178K, \$47K of which was written off, implying a loss rate of 26.1%. In order to maintain enough supply from the lending side, Prosper sets higher interest rates than Lending Club.[83]

[82] This business model was replaced with platform pricing in late 2010.

[83] In a Reuters article on January 19, 2010, Renaud Laplanche, the former CEO of Lending Club, wrote "Lending Club approves 10% of the loan applications. That's an underwriting decision. These 10% most creditworthy loans are made available on the platform for investors to invest in, and all loan listings get fully funded. Currently, the platform is "demand constrained," meaning that I have more investors willing to invest in these loans than loans available." The article continues "Prosper's 10% is very different in nature: most loan applications received by Prosper get listed on their platform, and only 10% actually get funded, either because of insufficient supply of investors funds, or just because investors don't want to fund the other 90% of the loans. The question here is whether the 10% that get funded are "the right 10%?" See, "http://blogs.reuters.com/felix-salmon/2010/01/19/the-problem-with-peer-to-peer-lending/"

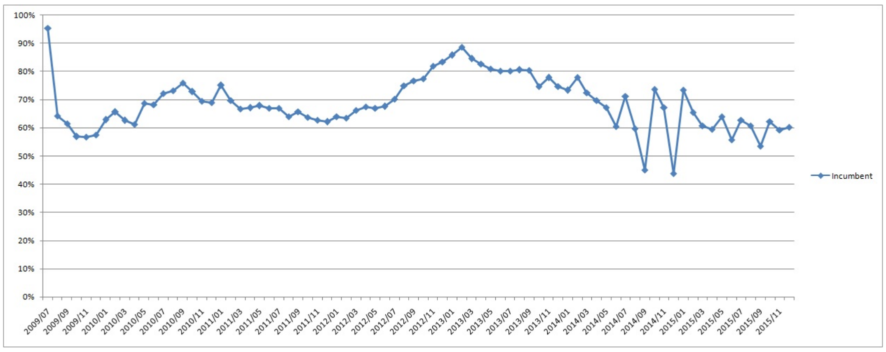


Figure 13: Lending Club's Market Share since Prosper's Entry

It shows the market share of the incumbent, Lending Club. The market share for the entrant, Prosper, is simply 1 minus the incumbent's market share. Lending Club has always been a dominant player in this market and once peaked at 90% in 2013.

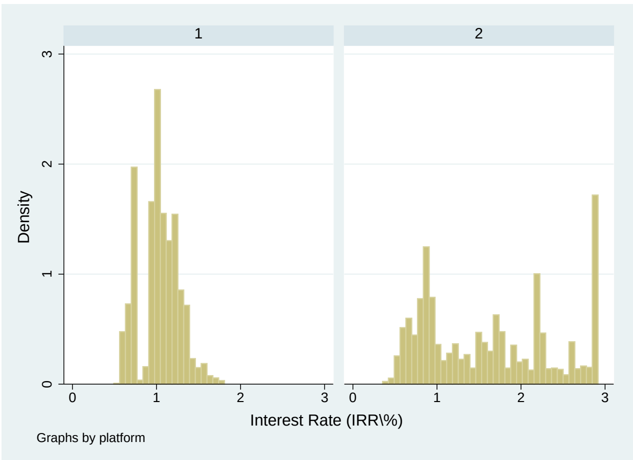


Figure 14: Interest Rate Comparison: Left, Lending Club; Right, Prosper

Left: In sample monthly interest rate distribution measured by IRR in percentage on Lending Club. Right: upper bounds of FICO score ranges of financed borrowers. Lending Club has monthly interest rates mostly below 1.5% and does not go beyond 1.8%. The lowest interest rate on Lending Club does not go below 0.5%. However, in comparison, Prosper's interest rates can go as high as 3% monthly and as low as 0.4% in IRR, and are more evenly distributed.

B.3.1 Prosper's Mechanism Change

Prior to December 20, 2010, interest rates on Prosper are decided by the market-clearing prices from multi-unit uniform price auctions.[84] On December 20, 2010, Prosper unexpectedly changed from its auction mechanism to the platform-pricing mechanism, and set interest rates on all the loans.[85]

B.4 Comparison to Banking

Similar to commercial banking and credit rating agencies (CRA), peer-to-peer lending produces and aggregates information on borrowers, prices interest rates and makes a market between borrowers and investors. However, unlike banks, a peer-to-peer lending platform typically does not take positions in securities for investors and thus does not need capital requirements or deposit insurance.[86] Unlike CRAs, peer-to-peer lending platforms are responsible for market clearance between investors and borrowers, in addition to information aggregation.

Opinions are dispersed on the future of peer-to-peer lending.[87] Not being a perfect

[84]After observing the borrower's quality and platform's rating, investors choose to enter the 7-day bidding process. First, the borrower sets a "reservation" interest rate as the opening bid, observable to all investors. The reservation interest rate is the highest interest rate the borrower accepts. However, borrowers may not reveal their true preferences. Investors submit their bids on the interest rate and unit of notes. Their interest rate bids must be below the opening bid. Also, remember a loan is divided into notes of $25. The bids are not completely sealed. As investors come in sequentially, they do not observe previous investors' interest rate bids, but they can observe their units and number of bidders.Zhang and Liu (2012) study investors' herding behavior using the timestamps of the bids among the issued loans. At the listing expiration, all interest rate bids are ranked in an ascending order, and the cutoff for the winners is where the pledged amount exceeds the requested amount. The loan's contractual interest rate is set at the first who is outbid. If the loan is not 100% funded, the loan is canceled and bidders get refunded.

[85]See Wei and Lin (2016). However, rumors were out prior to the organizational change. See "Prosper.com Ending Their Auction Process Dec 19th," Lending Academy. Aligned with Lending Club, it extended the listing period for the new borrowers from 7 days to 14.

[86]Due to this feature, liquidity shocks induced by agency costs between banks and investors are not applied here. See Hellmann et al. (2000), Allen and Gale (2000) and Diamond and Dybvig (1983)

[87]Some suggest that peer to peer lending isn't a threat to the banking industry, while others claim it may be the future of banking and the credit market. See "Peer-to-peer lenders will never challenge

substitute for banking, it exists due to several comparative advantages, and is able to undercut banks on both borrowing and lending. Banks and borrowers benefit heavily from relationship formation, where borrowers have access to cheaper credit and banks in return get lower credit risk. Agarwal et al. (2018) show that 56% of accounts in their sample are "Relationship Accounts." Data provided by Prosper show that although once peaked to 45% in 2011, the number stays around 20%-30%.[88] First time borrowers who are screened by banks get undercut by P2P platforms.

Second, the claimed purposes of most personal loans are for debt consolidation and refinancing. Facing higher interest rates from credit card debt, borrowers may receive lower interest on a peer-to-peer lending platform. According to the Fed, the average interest rate on an issued credit card is between 13-14% (APR) in 2014. On Lending Club alone, conditional on a FICO score 750 and above, the average interest rate is 8.4%. [89] P2P also caters to borrowers with lower creditworthiness, where the interest rate can go as high as 32%, 10% more than the highest legal rate among states with regulations.[90]

On the lending side, as deposit institutions, banks provide deposit insurance, and thus guarantee "risk-free" returns. First, P2P platforms cater to agents with heterogeneous risk preferences. The average deposit interest rate is less than 0.5% on a 3-year CD, whereas, the adjusted annual return for loans on Lending Club in 2016 averages between 4.9% and 8.3%. Second, institutions such as banks and funds account for more than

---

the banks, says Deloitte," The Telegraph. See "Lending Club Can Be a Better Bank Than the Banks," Bloomberg.

[88]'Is the Surge of Repeat Borrowers at Prosper Over?', Lendacademy

[89]Also, controlled for loan contract terms and borrower credit history, a loan with claimed purpose as 'debt consolidation' is 2.4% cheaper than "small business." Securitized loans from commercial bank such as auto-loan are generally cheaper (4% APR) than peer-to-peer lending, while incomparable since P2P lending does not require collateral.

[90]It is still debatable on what the credit rationing interest rate is. Some states are not regulated, such as ME, NH, NV, UT, SC and NM.

80% of the loan volume, which indicates that institutions are essentially undercutting each other using P2P lending platforms.[91] Banks compete locally due to geographical limitation and regulation. P2P lending provides the means for banks to undercut each other at the national level.

References

Sumit Agarwal, Souphala Chomsisengphet, Chunlin Liu, Changcheng Song, and Nicholas S Souleles. Benefits of relationship banking: Evidence from consumer credit markets. Journal of Monetary Economics, 2018.

Franklin Allen and Douglas Gale. Financial contagion. Journal of Political Economy, 2000.

Bo Becker and Todd Milbourn. How did increased competition affect credit ratings? Journal of Financial Economics, 2011.

Allen N Berger and Gregory F Udell. Collateral, loan quality and bank risk. Journal of Monetary Economics, 1990.

David G Blanchflower, Phillip B Levine, and David J Zimmerman. Discrimination in the small-business credit market. Review of Economics and Statistics, 2003.

Patrick Bolton, Xavier Freixas, and Joel Shapiro. The credit ratings game. Journal of Finance, 2012.

[91]Citigroup announces their partnership with Lending Club as an institutional investor. See "This Huge Bank Is Coming to Lending Club's Rescue." Fortune

Thorsten Broecker. Credit-worthiness tests and interbank competition. Econometrica: Journal of the Econometric Society, 1990.

Don Carmichael. The winner's curse in an online lending market. Working Paper, 2017.

Zhaohui Chen, Alan D Morrison, and William J Wilhelm Jr. Traders vs. relationship managers: Reputational conflicts in full-service investment banks. Review of Financial Studies, 28(4):1153–1198, 2014.

David M Cutler and Sarah J Reber. Paying for health insurance: the trade-off between competition and adverse selection. Quarterly Journal of Economics, 1998.

Francois Derrien and Kent L Womack. Auctions vs. bookbuilding and the control of underpricing in hot ipo markets. Review of Financial Studies, 2003.

Douglas W Diamond and Philip H Dybvig. Bank runs, deposit insurance, and liquidity. Journal of Political Economy, 1983.

Sergio Firpo. Efficient semiparametric estimation of quantile treatment effects. Econometrica, 2007.

Sean Flynn and Andra Ghent. Competition and credit ratings after the fall. Management Science, 2017.

Seth Freedman and Ginger Zhe Jin. Do social networks solve information problems for peer-to-peer lending? evidence from prosper.com. Working Paper, 2008.

Robert Hauswald and Robert Marquez. Competition and strategic information acquisition in credit markets. Review of Financial Studies, 2006.

Thomas F Hellmann, Kevin C Murdock, and Joseph E Stiglitz. Liberalization, moral hazard in banking, and prudential regulation: Are capital requirements enough? American Economic Review, 2000.

Rajkamal Iyer, Asim Ijaz Khwaja, Erzo FP Luttmer, and Kelly Shue. Screening peers softly: Inferring the quality of small borrowers. Management Science, 2015.

Artashes Karapetyan and Bogdan Stacescu. Information sharing and information acquisition in credit markets. Review of Finance, 18(4):1583–1615, 2013.

David S Lee and Thomas Lemieux. Regression discontinuity designs in economics. Journal of Economic Literature, 2010.

Mingfeng Lin and Siva Viswanathan. Home bias in online investments: An empirical study of an online crowdfunding market. Management Science, 62(5):1393–1414, 2015.

Yongdong Liu, Denis Nekipelov, and Minjung Park. Timely versus quality innovation: The case of mobile applications on itunes and google play. Working Paper, 2013.

Robert Marquez. Competition, adverse selection, and information dispersion in the banking industry. Review of Financial Studies, 2002.

Paul R Milgrom and Robert J Weber. A theory of auctions and competitive bidding. Econometrica: Journal of the Econometric Society, pages 1089–1122, 1982.

Enrichetta Ravina. Beauty, personal characteristics, and trust in credit markets. Working Paper, 2007.

Kevin Rock. Why new issues are underpriced. Journal of Financial Economics, 15(1-2): 187–212, 1986.

Sherrill Shaffer. The winner's curse in banking. Journal of Financial Intermediation, 7 (4):359–392, 1998.

Steven A Sharpe. Asymmetric information, bank lending, and implicit contracts: A stylized model of customer relationships. Journal of Finance, 1990.

Vasiliki Skreta and Laura Veldkamp. Ratings shopping and asset complexity: A theory of ratings inflation. Journal of Monetary Economics, 2009.

Victor Stango and Jonathan Zinman. Borrowing high versus borrowing higher: price dispersion and shopping behavior in the us credit card market. Review of Financial Studies, 2015.

Joseph E Stiglitz and Andrew Weiss. Credit rationing in markets with imperfect information. American Economic Review, 1981.

Ernst-Ludwig Von Thadden. Asymmetric information, bank lending and implicit contracts: the winner's curse. Finance Research Letters, 1(1):11–23, 2004.

Zaiyan Wei and Mingfeng Lin. Market mechanisms in online peer-to-peer lending. Management Science, 2016.

Juanjuan Zhang and Peng Liu. Rational herding in microloan markets. Management science, 2012.