The Winner's Curse in an Online Lending Market*

Don Carmichael†

First Version: October 23, 2017

Abstract

Using data from Lending Club and Prosper, the two largest peer-to-peer lenders in the U.S., we provide evidence of the winner's curse in the online personal lending market. Borrowers who were rejected by a competitor are twice as likely to default as borrowers who were not rejected, conditional on receiving the same contract. Borrowers are also more likely to default when offered higher interest rates or smaller loan amounts by a competitor. Surprisingly, the loan amount effect is larger than the interest rate effect, and loan amount is a more closely related to lender choice.

*For helpful comments and suggestions, I thank Kevin Roshak, Nisan Langberg, Paul Povel, Kris Jacobs, Praveen Kumar, Itzhak Ben-David, Tetyana Balyuk, and Rauli Susmel.

†University of Houston, C.T. Bauer College of Business; Houston, TX 77204-6021; phone: 713-306-3700; email: dbcarmichael@uh.edu

Introduction

The winner's curse is an important phenomenon in lending markets. The winner's curse in lending is a selection problem which occurs due to information asymmetries between lenders. The "curse" is that the winner of a common value auction will be the one who overvalued the product the most—in our application the one who overestimated the creditworthiness of the borrower the most. Bidders following an optimal bidding strategy must take this into account by adjusting their bids downward and/or refusing to participate (Broecker, 1990; Sharpe, 1990; Von Thadden, 2004), but the winner's curse should appear in lenders markets regardless of how lenders respond to it.

Thoroughly analyzing the winner's curse and its effects on markets requires data on competing bids (in the example of lending, this means contract offers or rejections). Such analysis has been performed for oil lease auctions (Capen et al., 1971), eBay auctions (Bajari and Hortacsu, 2003), and other kinds of auctions. For lending markets, however, no such analysis exists as losing bids (interest rate offers which are not accepted and rejections) are not generally observed. We make use of a novel data set, which includes multiple bids for a specific loan product for a specific borrower, as well as the subsequent performance of that borrower's loan. We identify borrowers who were rejected at one lender but given a loan by another. We also identify borrowers who had a pending loan application at one lender, with an offered interest rate and loan amount, but had that application terminated, and eventually received a loan from another lender. This allows us to consider the informational content of these offers and rejections; in other words, to investigate the relationship between competitors' offers/rejections and default. We expect that borrowers rejected by one lender, but accepted by a second lender should be worse credit risks than borrowers who were offered exactly the same contract by the second lender (interest rate, loan term, and loan amount) but were not rejected by the first lender. We also expect that borrowers who had their loan applications canceled by a competitor should be worse credit risks, and that higher competitor's interest rate offers and lower competitor's loan amount offers should imply higher default rates.

We investigate these hypotheses in an online personal lending market consisting of the two major peer-to-peer lenders in the United States, Lending Club and Prosper. We compare the performance of borrowers who were rejected by Lending Club, but accepted by Prosper to borrowers who were accepted by Prosper but not rejected by Lending Club. We use a number of matching procedures to ensure that our results are not driven by different appetites for risk between lenders. This includes comparing the Lending Club rejects to borrowers who received exactly the same contract (interest rate, loan term, and loan amount) but were not rejected by Lending Club. We find that borrowers rejected by Lending Club are twice as likely to default.

We also consider borrowers who had loan "listings" with Prosper, but eventually received a loan from Lending Club and not from Prosper. Prosper listings have been assigned an interest rate, but are awaiting the approval of the applicant's documentation and the funding of the loan (P2P loans are not issued unless investors have agreed to purchase the debt). Prosper listings can either be completed, in which case they become loans, or they can expire (not be funded by investors), be withdrawn by the applicant, or be canceled by Prosper. Lending Club loans to borrowers who had canceled listings at Prosper are 75% more likely to default. We also find that the interest rate assigned to expired/withdrawn/canceled listings is a significant predictor of default for Lending Club loans to the same individual. Surprisingly, competitor's loan amount is an even more significant predictor of default than competitor's interest rate. Consistent with this result, we find that loan amount is more closely related to lender choice than is interest rate. On the whole, we find that both lenders face a winner's curse stemming from the decisions of the other, and that the winner's curse occurs both due to the possibility of rejection and due to the possibility of an inferior contract offer (higher interest rate and/or loan amount) from the competitor.

Ours is not the first study to consider the winner's curse empirically. Shaffer (1998) uses bank-level data and finds evidence consistent with the winner's curse, namely that newer banks and banks in more competitive markets experience more charge-offs. In the same vein as Shaffer (1998), a number of studies consider situations in which one lender has superior information, because they are a repeat lender (Berger and Udell, 1995), or are located physically closer to the borrower (Agarwal and Hauswald, 2010). These studies consider the effects of superior information on pricing and loan performance.

However, our study is the first to consider the winner's curse in lending using data on competing bids. The need for such a study is evident from the literature. For example, Agarwal and Hauswald (2010) interpret the relationship between borrower-lender distance and price and loan performance as the consequence of informational frictions, whereas Degryse and Ongena (2005) interpret similar results as the consequence of transportation costs. Since we actually obtain the interest rate offers and internal credit evaluations of two competing lenders, however, our identification strategy is much cleaner than those in the existing literature. Our study is also the first to investigate the existence of the winner's curse in general, rather than in a situation where one lender is believed to have superior information.[1] While our finding that the winner's curse does indeed exist in lending markets is unsurprising, a bid-level analysis of the winner's curse in lending is long overdue.

There are two important differences between the market we study, and the corporate lending markets that have been the focus of the existing literature. Firstly, we study consumer lending. Secondly, and more importantly, the lenders in our sample are online, "arms-length" lenders, meaning that the information asymetries leading to adverse selection in these data are not "soft" information asymmetries arising from lenders knowing different facts about loan applicants, but are rather asymmetries in information processing techniques. These can stem from differences in algorithms, historical training data sets, methods of turning credit reports into variables, or choice of credit bureau. The informational competition between lenders is one between machines rather than between tradiational loan officers. This type of competition has been considered in the theoretical literature (Hauswald and Marquez, 2006; Hauswald and Marquez, 2003), but we are not aware of any empirical investigations into

[1] A negative relationship between concentration and charge-offs has been found in the literature (e.g., Shaffer (1998), but such a finding is consistent with a wide range of explanations besides the winner's curse.)

the effects of imperfect information in competitive lending markets where interest rates are set by algorithm. Credit evaluation by algorithm alone is becoming increasingly common (Berger, 2003). Credit card and personal lending markets have already made the transition to algorithmic rate setting. Mortgage markets are rapidly transitioning to algorithmic rate setting, and small-business lenders using algorithmic rate setting are beginning to appear.

Our research makes four main novel insights into the nature of winner's curse in lending. Firstly, we provide evidence that two competing lenders both face a winner's curse because of competition with each other, whereas the existing literature only provides evidence of a unilateral winner's curse. Secondly, we provide evidence using data on the same borrower from different lenders. Thirdly, we provide evidence of the winner's curse in a market where contract offers are determined purely by algorithm. Lastly, we show that loan amount offers, not just interest rate offers, are important in determining which contract a borrower chooses to accept, and in creating a winner's curse. To investigate the winner's curse in an online lending market, we first discuss this market's institutional details and available data (Section 1), then discuss results for borrowers rejected at Lending Club and accepted by Prosper (Section 2), and finally discuss results for borrowers with Prosper listings that eventually received Lending Club loans (Section 3).

1 Background and Data

1.1 Background

We use data from the two main P2P lenders in the U.S., Lending Club and Prosper. During the primary sample period (2009-2013),[2] these two lenders dominated the online personal lending market, and to a lesser extent, the entire personal lending market.[3] P2P lenders during this period (and today) functioned as arm's-length lenders, setting interst rates via

[2]We also use data from 2014–2017 for a limited number of tests. During this time, the Lending Club and Prosper face much more external competition.

[3]Traditional banks have always been hesitant to make unsecured personal loans, and were especially hesitant in the post-crisis period.

algorithm without the intervention of a credit officer.

Prior to our sample period, P2P lenders functioned much differently than they did during our sample period. The lending was truly peer-to-peer. Before being shut down by the SEC in 2008, Prosper, which was the first P2P lender in the U.S., was in many ways a social website. Borrowers would post profiles, complete with pictures and biographies. Lenders would form groups and make lending recommendations to each other. They would ask questions of prospective borrowers and make lending decisions in part based on the responses. Interest rates were set by Dutch auction. Banks, registered securities, and underwriting models were not a part of the process. Lending Club initially functioned similarly to Prosper. It was a Facebook application, rather than a standalone website, but was social, just like Prosper. Like Prosper, it allowed lenders to set rates by Dutch auction, though it did recommend an interest rate.

Lending Club quickly transitioned into being a website rather than a Facebook application and began to set interest rates via its own underwriting model rather than facilitating an auction. Lending Club registered with the SEC in mid-2008. In November 2008, the SEC issued a cease and desist order to Prosper, who then changed their business model to mimic Lending Clubs and reopened to sell SEC-registered securities in July 2009. For our analysis, we will consider only loans issued in July 2009 and after. We wish to study competition and informational frictions between competing lenders, which occurs only from July 2009 and on. Before July 2009, Lending Club competed with individual investors using the Prosper platform, not with Prosper directly.

1.2 The Lending Process

On both Lending Club and Prosper, loans are now issued via the following process. A prospective borrower applies, reporting information about himself, his finances, and his need for a loan. The lender checks the borrower's score and report, then assigns him a risk subgrade and a corresponding interest rate. If the borrower accepts the interest rate, the loan is listed on the lender's website. Prospective investors can browse the loans listed on the website and agree to fund a portion of the loan. The loan is issued through a bank affiliated with the lender, so these loans are legally the same as traditional unsecured loans. However, the lender does not make loans until borrowers have agreed to fund the entirety of the loan. All issued loans are thus immediately securitized. The lender may choose to verify the borrower's income while the loan is in funding but does not delay the listing of the loan to verify income. If the borrower's income cannot be verified, the loan is removed from the site (canceled). If the loan becomes fully funded before the borrower's income can be verified, the loan is issued without verification. Some loans expire without being funded. This is relatively rare on Prosper. Lending Club claims this is essentially unheard of, but data on expired listings are not available from Lending Club. The borrower may also withdraw his listing from the site at any time until the loan is funded.

1.3 P2P Lending Literature

The existing P2P literature primarily makes use of data from before SEC regulation, primarily focusing on social aspects such as pictures (Duarte et al., 2012), text analysis (Michels, 2012; Herzenstein et al., 2011), and group membership (Lin et al., 2013; Everett, 2010). These social factors are not involved in the P2P data we use, which comes from 2009–2017, by which time the SEC had shut down the more social kind of P2P lending.

P2P lending after SEC regulation has been studied by a very limited number of papers, most of which focus on forecasting. However, Balyuk (2016) finds that P2P lending mitigates informational frictions in personal credit markets. This paper is closely related to ours in that it studies informational frictions and the P2P lending market. Balyuk (2016) shows that P2P lending reduces informational frictions, whereas we show that large informational frictions still remain within the P2P lending market.

1.4 Data

We use the following data sources:

1. Data on the borrower and loan for every loan issued via Prosper. These include the

credit score, many variables extracted from the credit history, borrower-reported variables, loan contract variables, and repayment information.[4]

2. Data on the applicant and loan for every listing on Prosper that is not eventually issued. These include canceled, withdrawn, and expired listings. The data available is identical to that available for issued loans, except of course, there is no loan performance data.[5]

3. Data on the borrower and loan for every loan issued via Lending Club. The specific data fields are somewhat different, but otherwise these data are very similar to those available for Prosper loans.[6]

4. Data on every applicant rejected by Lending Club. For rejected applicants, we have the following 8 data fields: amount requested, listing creation date, loan title, credit score, debt-to-income, city, state, and employment length.[7]

5. Interest rate offer data from online reviews from CreditKarma.com.[8]

We report summary statistics for all Lending Club loans in Table 1 and for all Prosper loans in Table 2. A key portion of our analysis, that on rejected borrowers, uses data on only loans issued in 2013 or before. We report summary statistics for these loans in Table 3. A typical P2P borrower borrows $10,000–15,000, and, relative to the U.S. population as a whole, has a high income (mean of $75,000 annually for Prosper and $77,000 for Lending Club) and a slightly below average credit score. Lending Club and Prosper borrowers have similar credit histories and financial situations in most ways, the main exception being that Lending Club borrowers have on average 0.51 fewer credit inquiries in the last 6

[4]These data are publicly available on Prosper.com. However, our application requires a legacy data set.

[5]These data are publicly available on Prosper.com

[6]These data are publicly available on LendingClub.com. However, our application requires a legacy data set.

[7]These data are publicly available on LendingClub.com

[8]Reviews are publicly available, but interest rate offers are gathered by skimming all reviews of both Lending Club and Prosper manually.

months. Although Lending Club and Prosper borrowers appear very similar based on summary statistics, Lending Club offers their borrowers much more favorable terms, on average lending $15,000 at 13% interest as opposed to $12,000 at 16% for Prosper borrowers. We show a histogram of interest rates at both lenders in Figure 1. The modal interest rate at both lenders is 13%, but Prosper has a much fatter right tail with sizable issuance even at 35%. Surprisingly, Prosper borrowers actually default 34% less frequently than Lending Club borrowers, despite paying higher interest rates. Prosper has historically had much more trouble attracting investors than Lending Club, with 19% of Prosper listings expiring without being funded.[9] The exact figure for Lending Club is not disclosed, but is very close to 0. Prosper's difficulty in attracting investor's and desire to build repuatational capital to attract investors in the future likely explains the outperformance of Prosper loans.

[**Place Table 1 about here**]

[**Place Table 2 about here**]

[**Place Table 3 about here**]

We use the Lending Club rejection data only for determining which loans in the Prosper data set were made to borrowers who were rejected by Lending Club. We describe the procedure for matching individuals between the Lending club rejection data and Prosper loan data in Section 2.1. We use data on Prosper listings that did not eventually become loans to determine which Lending Club borrowers received loan offers from Prosper and at what interest rate they received these offers. We describe the procedure for matching borrowers between the Prosper listing and Lending Club loan data sets in Section 3.1.

[9]The percentage of listings which eventually become loans is 52%. The other 29% are canceled or withdrawn.

2 Evidence for the Winner's Curse: Borrowers Rejected by a Competitor

Our first piece of evidence involves comparing the loan performance of Prosper borrowers rejected by Lending Club to borrowers not rejected by Lending Club. Our hypothesis is that borrowers rejected by Lending Club should perform worse, even when nothing in Prosper's data set suggests that this should be the case. This hypothesis is supported in the theoretical literature (Broecker, 1990; Hauswald and Marquez, 2006; Fishman and Parker, 2015) and is intuitive, but has not been tested empirically in the existing literature. We first match borrowers between Lending Club and Prosper to construct a dummy variable indicating whether the borrower was rejected. Then, we use a Cox proportional hazards model of default. We use a number of different techniques to ensure that we are comparing rejected borrowers to non-rejected borrowers that appear equally creditworthy to Prosper. These techniques include running the hazard model on a matched sample of rejected borrowers and non-rejected borrowers who received identical contracts.

2.1 Identifying Rejected Borrowers

Identifying the same applicant at multiple lenders is very difficult.[10] For privacy reasons, Lending Club and Prosper, like all other lenders, do not publish personally identifiable information on their borrowers or applicants. However, until 2013, Lending Club and Prosper both published the (mailing address) cities of their borrowers. Both lenders have since stopped providing such specific data on borrower location, but it leaves us with ample data that are surprisingly specific.

We identify a Prosper borrower as having been rejected by Lending Club if there is an applicant in the Lending Club rejection dataset that:

1. Applied on the same date as the Prosper borrower
2. Lives in the same city as the Prosper borrower

[10]We are aware of one other paper which identifies rejected borrowers who were subsequently accepted by another lender: Agarwal et al. (2016a). Their matching technique is very similar to our ours.

3. Has the same employment length as the Prosper borrower

This measure of rejection is very noisy if we include borrowers from larger cities. However, by restricting the sample to borrowers from small cities, we are able to determine which borrowers were rejected with a rate of false positives that is very close to 0. We consider a city to be small if there are fewer than 10 applications from that city in the rejection dataset. For robustness, we also use cutoffs other than 10. We also construct alternate samples using restrictions based on FICO score and debt to income, in addition to the three listed above.

At first glance, this matching technique may seem inaccurate. We can determine the accuracy of the match (that is, the rate of false positives) by using a date offset. If a rejection from the same city with the same employment length appears in the Lending Club data set after a borrower has already received a loan from Prosper, he is not the same borrower.[11] The matching technique described above yields 192 matches, whereas an identical technique with an offset of approximately 30 days yields on average 4 matches. This implies that the matches are 98% accurate. Measurement error exists, but is small enough to be ignored when considering the magnitude of the coefficient of interest.

The restriction that the Lending Club and Prosper applicants apply on the same day may seem stringent. But, in reality, very few applicants apply to Lending Club on one day, get rejected, and then go to Prosper the next. Using the above matching technique with a one-day offset (rejection before acceptance), we observe 15 matches. This implies a much lower matching accuracy of 73%. There is no need to introduce additional measurement error into the sample, so we use the same day restriction throughout this paper.

2.2 Rejected Borrowers and Default: Methods

We expect borrowers rejected by one lender to default more frequently when accepted by another lender. We must ensure, however, that we are comparing rejected and non-rejected borrowers who appear identical to the lender. In an attmept to do so, we will use four

[11] He could actually be the same individual, but even if he is, he is now shopping for an additional loan, not the same loan.

different methods to test our hypothesis. All four of these methods include a proportional hazards model, but we will use various matching and weighting schemes. We will use a standard proportional hazards model with no weighting or matching, an inverse probability of treatment weighting (IPTW), propensity score matching, and exact matching based on contract variables.

The first of these methods is a straightforward proportioanl hazards model, without any weighting or matching. The specification is the following:

$$\lambda_i(t) = \lambda_0(t)\exp(X_i'\beta + \gamma Rejected_i)$$

where $\lambda_i(t)$ is the instaneous default hazard rate for loan i at time t, $\lambda_0(t)$ is the baseline hazard rate at time t, X_i is a vector of borrower and contract variables, and *Rejected* is an indicator variable taking a value of 1 if the borrower was rejected by Lending Club. The coefficient of interest is γ, and $\gamma > 0$ is our hypothesis.

The second method, IPTW, consists of weighting each observation by the inverse probability of receiving the treatment that was received (reciprocal of probability of rejection for rejected borrowers and reciprocal of probability of not having been rejected for non-rejected borrowers). We calclulate the treatment probabilities using a logistic regression with the following specification:

$$Rejected'_{it} = x_{it}^{\top}\beta + \epsilon_{it} \tag{1}$$

where rejected is the latent variable corresponding to *Rejected*, and is greater than 0 where the borrower was rejected by Prosper. This logit yields probabilities of the following form:

$$\frac{p(Rejected'_{it})}{1 - p(Rejected'_{it})} = e^{x_{it}^{\top}\beta} \tag{2}$$

We then use the extracted treatment probabilities, to assign weightings of $1/p(Rejected'_{it})$ for rejected borrowers, and $1 - 1/p(Rejected'_{it})$ for non-rejected borrowers. These weights are then used in a weighted proportional hazards model, which is identical to the one used in method 1, except for the weighting scheme.

The third method, propensity score matching, makes use of the probabilities from the same logit, but instead of weighting all the observations, we create a matched sample of observations with similar rejection probabilities (propensity scores). Since there are many more non-rejections than rejections, we match each rejection to the three non-rejections with the nearest propensity scores, weighting each match with one third the weight of each rejected borrower. We run the same proportional hazards model as in method 1, but only on the matched sample, not on the full sample.

The last method consists of exact matching on all of the contract terms (loan term, interest rate, and amount). We match each rejected loan-month observation to 3 observations with the same term, interest rate, amount, and time to maturity. Where an exact amount match is not available, we use the closest available amount. As in method 3, we run the same proportional hazrds model as in method 1, but on the matched sample rather than the full sample.

2.3 Rejected Borrowers and Default: Empirical Results

We report regression statistics for a Cox proportional hazards model without any matching or weighting in Table 4. We are interested in the coefficient on the dummy variable indicating whether or not the borrower was rejected by Lending Club. These coefficients of interest can be interpreted as follows: rejected borrowers are e raised to the coefficient size times more likely to default at an given instant. Since default is a relatively rare event, it is in general a good approximation that lifetime probability of default is affected by the same percentage as the instantaneous probability of default. Rejected borrowers are 2.2 times more likely to default. This reduces the expected return on a typical 19% interest loan from 7% to −6%, or a difference in total payments received of about $5900 for a $15000, 3-year maturity loan ($10000 for a 5-year loan).

[Place Table 4 about here]

In Table 5, we report the regression statistics for a proportional hazards model using IPTW (column 1) and propensity score matching (column 2). Both of these methods are based on a first stage propensity scoring regression (not reported). This first stage regression is a logistic regression modeling whether or not a borrower was rejected. For IPTW, we calculate the probability of rejection for every borrower using the first stage logit. We then weight every observation in the second stage by the inverse probability of receiving the treatment that was actually received. So if $p(X)$ is the probability of rejection for a given borrower with covariates X, rejected borrowers are given weight $1/p(X)$, and non-rejected borrowers are given weight $1/(1-p(X))$. For propensity score matching, we use the same first state regression as for IPTW to calculate $p(X)$ for every borrower. In the second stage hazard model, however, we use a sample that consists of every rejected borrower and the 3 non-rejected borrowers with the closest values for p. Other non-rejected borrowers are not included in the sample. Using these two propensity-score-based methods, we estimate that rejected borrowers are 2.2 (IPTW) and 1.8 (matching) times more likely to default, very much consistent with the 2.2 we calculate for the hazard model without matching or weighting.

[Place Table 5 about here]

In Table **??**, we report the regressions statistics for a Cox proportional hazards using exact matching based on all contract variables. This is a 3-to-1 match, just like the propensity score matching, but instead of finding the closest 3 propensity score matches, we find 3 exact matches for all contract terms (interest rate, maturity, and loan amount). Sometimes, exact amount matches are not available, and so we use the closes available amount. Exact matching on all contract variables is the most effective method for confirming that we are comparing borrowers who appear to have the same creditworthiness to Prosper. There is no rational explanation for why Prosper would see a difference in creditworthiness between two borrowers but offer them the same contract.[12] According to the exactly matched hazard model, rejected borrowers are 2.0 times more likely to default.

[Place Table 6 about here]

[12]We also know that borrowers are only given the same contract when they are assigned the same risk grade

We use several alternative methods of determining whether borrowers were rejected. As mentioned previously, we use different small city size thresholds and also use FICO score and debt to income as additional criterion for mapping between the Lending Club rejections and Prosper acceptances. For brevity, we do not report summary or regression statistics using any of these alternative samples. Use of FICO score and/or debt to income restrictions does not change the results, except that they slightly decrease the sample size (and therefore the t-stats). Decreasing the city size threshold also merely reduces the sample size. Increasing the city size threshold decreases the magnitude of the coefficients because it introduces measurement error, but increases the t-stats because it increases the sample size. Altogether, the robustness checks cast no doubt on the results.

2.4 Economic Magnitude

These results represent an clear case of the winner's curse. Borrowers rejected by Lending Club should not be given loans by Prosper. However, just because a rejection by a competitor is a very negative signal when it occurs does not mean that the possibility of a rejection has a large distorting effect on the market as a whole. What is the overall impact of the winner's curse from the possibility of rejections on this market?

The first question we need to answer is what fraction of borrowers were rejected by a competitor. To answer this question, we must use a matching technique that eliminates the possibility of false negatives. With this objective in mind, we construct 300 different matches between the Lending Club rejection data set and the Prosper loan data set. For loan and rejection to be considered matched, they must be from the same city and occur n days apart, we let n vary between -100 and 200.

We graph these matches in Figure 2. At 0 on the x-axis are same-day matches. At 1 on the x-axis are matches where the Prosper loan application comes 1 day after the Lending Club rejection, and so on. In the figure, we clearly see a baseline of false positives, with a signal for the same day matches, as well as the matches. We can see just over $4,000$ genuine matches for the same day, and perhaps $5,000$ total genuine matches including the same day matches and the matches where the rejection occurs one day before or after the acceptance at the competitor. The total number of loans is $180,000$, meaning that 2–3% of Prosper borrowers were rejected by Lending Club.

[Place Figure 2 about here]

This means that the possibility of rejection by a competitor impacts expected annual returns by 0.25%. For high interest rate (30%) loans, the probability of rejection is higher, as is the effect of rejection on returns. High interest rate loans suffer a 2.5% decrease in returns due to the possibility of rejection by a competitor. Because this effect is concentrated among high interest rate borrowers, it likely creates a credit rationing problem, though identifying the exact amount of credit rationing requires estimating a counter-factual and is beyond the scope of our paper. Loans to borrower rejected by Lending Club experience $10 million dollars more charge-offs annual than similar loans to non-rejected borrowers. It is important to keep in mind that this is the figure for only one lender based on rejections from only one other lender. All P2P lenders likely face a similar winner's curse because of the possibility of rejections at any of it's competitors. The aggregate effect on the market as a whole is thus several times higher than the $10 million quoted above.

3 Evidence for the Winner's Curse: Applicants with Multiple Offers

In this section, we will consider borrowers who were conditionally accepted by Prosper and offered a loan contract, but who eventually received a loan from Lending Club rather than from Prosper. The theoretical predictions for the case in which an applicant receives two different interest rate offers are less clear than the case where the borrower is rejected. In some theoretical models of competition in bank lending with imperfect information, rejections are informative, but interest rates are not (Broecker (1990), Hauswald and Marquez (2006)). In situations with more than two possible borrower types (as is the case in the theoretical models cited above), interest rate offers will likely be informative, however.

Applicants who are conditionally accepted and have their loans listed on Prosper's website may have their listing terminated for three reasons. The first is withdrawal. For borrowers who withdraw their applications and then receive loans from Lending Club, the rationale is quite clear: they received a more attractive contract from Lending Club (likely at a lower interest rate, but it is also possible that loan amount or term could play a role as well) and took the more attractive offer. For applicants whose listings expired, the rationale is also clear: these applicants passed the initial screening by Prosper, but were rejected by investors. For applicants whose listings were canceled there are two possible cases. Applications are canceled either when the applicant fails to provide documentation (for income, employment, etc.) or when his documentation cannot be verified. The first case, a failure to provide documentation, is analogous to a withdrawal, whereas the second case, verification failure, is essentially a rejection from Prosper.

For borrowers who had canceled listings with Prosper, we expect higher default rates, analogous to the higher default rates of rejected borrowers found in the previous section. However, since some canceled listings are canceled because borrowers decide to not finish the application process and go elsewhere, we would expect the relationship between cancellation and default to be weaker than the relathionship between rejection and default. For borrowers with withdrawn or expired listings, we may see a positive relationship between default and withdrawal/expiration that is related to the winner's curse, but it will likely be much weaker than that observed for cancellations. Simply the fact the borrower applied to a competitor implies as stronger winner's curse than if he did not (the winner's curse is stronger the more bidders are involved). However, if most applicants apply to both Lending Club and Prosper, this effect will be weak. For expired listings, the winner's curse may play a role if investors have the expertise needed to screen borrowers.[13]

We might also expect borrowers whose interest offers were higher on Prosper to default more frequently. Although in many lending-specific theoretical models, interest rate offers convey no information about borrower quality, this is unlikely to be the case in reality. The winner's curse is based on the idea that competitor's valuations are lower than one's own

[13] Iyer et al. (2015) suggest that they do.

valuation, conditional on winning an auction. For the winner's curse to be a reality, the actual bid that the competitor makes must be correlated with default, conditional on one's own bid. Thus, we test for the winner's curse in lending, based on the interest rate offers made by two competing lenders, by testing whether competitors' interest rate offers can predict default, when combined with a lender's internal pricing model.

3.1 Identifying Borrowers with Withdrawn, Canceled, or Expired Listings

To analyze the relationship between Prosper listing status and Lending Club loan performance, we must first match borrowers between the Prosper listing and Lending Club loan data sets. This process is similar to, but slightly different from, that used in Section 2.1 to identify borrower rejected by Lending Club, but accepted by Prosper. The matching procedure needs to be different because there is no reason to expect the Prosper and Lending Club applications to take place on the same day. In fact, the Prosper and Lending Club applications will likely not take place on the same day. If the borrower had applied simultaneously to Lending Club and Prosper, and the Lending Club terms were more favorable, the borrower would have accepted the Lending Club offer and never have had a Prosper listing. Thus, the borrower must first apply to Prosper and initiate a listing, and *then* apply to Lending Club.

Although we cannot use the same day restriction like we do for Lending Club rejects, the Prosper listing data set is much richer than the Lending Club rejection data set in terms of credit and employment variables, which allows for an accurate match based on month of application and credit variables, rather than date of application and employment length alone. Because we do not need the exact date of application for the Lending Club loans, we can use a more updated data set. More recent Lending Club data sets include three-digit zip code and month of application. We match this data set to the Prosper listings data set using the following procedure: a borrower in the Lending Club data set is considered to have had a withdrawn (canceled, expired) Prosper listing if the Prosper listing data includes a listing with the following characteristics:

1. The same month of application as the Lending club loan
2. The same three-digit zip code
3. The same annual income (rounded down to nearest $100)
4. The same number of years (not exact months) of credit history
5. The same number of open accounts
6. The same homeownership status

By design, these matching criteria are highly restrictive. This means that many loans to borrowers that actually had Prosper withdrawn/canceled/expired listings are not classified as such. Lending Club and Prosper pull credit information from different bureaus, so it is very likely that data fields on the same individual will be observed as different by the two lenders. However, given that having a Prosper withdrawn/canceled/expired listing is a relatively rare event, the high rate of our algorithm's failure to identify matching Prosper listings will not cause significant attenuation bias.[14] However, misclassifying borrowers without withdrawn/canceled/expired listings as having them can result in significant attenuation bias. This is why we design our match to be highly restrictive.

We cannot estimate the rate of false negatives. However, the rate of false positives can be estimated by constructing a placebo sample. To construct the placebo sample we add (subtract) 365 days to the Lending Club application date for loans accepted before (after) January 2014.[15] We then perform our matching algorithm using these modified data. The

[14] According to Aigner (1973), $\hat{\beta} = \hat{\beta}_{OLS}(1 - \eta - \nu)$ is an unbiased consistent estimator of β, where β is the true effect of a binary variable (in our case withdrawal, cancellation, or expiration) on an outcome variable (in our case, default), $\hat{\beta}_{OLS}$ is the OLS estimator of β, η is the rate of false negatives (fraction borrowers marked as not having withdrawn/canceled/expired listings which in reality do), and ν is the rate of false positives (borrowers marked as having withdrawn/canceled/expired listings which in reality do not). In the case where the treatment is a rare event (which withdrawal, cancellation, and expiration all are), η will usually be low. It is much more important to construct a matching algorithm that results in a high ν.

[15] We must add to the earlier dates and subtract from the later ones to ensure that all modified dates correspond to dates in the Prosper listings data

algorithm yields 203 placebo matches, as compared to 1007 matches when using the correct dates. Thus, we estimate the rate of false positives (ν) to be 203/1007, or 20%.

3.2 Empirical Results: Borrowers with Canceled Listings

We expect Lending Club borrowers with canceled Prosper listings to perform worse than other Lending Club borrowers. A cancellation is very similar to a rejection in that loans are canceled when the lender decides not to lend to a particular borrow. A cancellation is, however, a weaker signal than a rejection. This is because a cancellation is a more ambiguous signal than a rejection. Loans are generally canceled when there is a problem with the documentation process. Typically, this we be because the lender has caught the applicant providing misleading information, but a cancellation may also occur when the applicant chooses not to provide documentation because he has decided not to take the loan, likely because he received a better offer elsewhere. Because of the ambiguity associated with cancellations, we believe the results offered in previous section provide a better assesment of the magnitude of the relationship between rejection and default. However, since we lack data on applicants initially rejected by Prosper, considering applicants whose listings were canceled by Prosper can help us determine whether both lenders face a winner's curse problem.

We assess the relationship between cancellation and default using the same techniques we used to assess the relationship between rejection and default. We use a Cox proportional hazards model of default, employing both equal weighting and inverse probability of treatment weighting, as well as propensity score matching and exact matching based on contract variables. In Table 7, we report the results of our Cox proportional hazards using all four of the weighting/matching techniques.

[Place Table 7 about here]

The coefficients on cancellation are all between 0.42 and 0.50, suggesting that borrowers who had canceled listings are about 60% more likely to default. Adjusting for attenuation bias, this figure is 75%. This result is consistent with each lender facing a similar winner's curse problem. The coefficient on rejection for Prosper loans a third larger, compared to the coefficient on cancellation for Lending Club loans. This is to be expected given that a cancellation is an ambiguous signal, whereas a rejection is not. Taken together, our results for rejected and canceled applications suggest that both Lending Club and Prosper face a large winner's curse problem in that their pool of applicants includes borrower rejected by their competitor.

3.3 The Winner's Curse and Competitors' Contract Offers

We have tested whether rejections and cancellations by a competitor are indicative of greater credit risk. But what about the more common case in which both lenders offer contracts? Does the competitor's offer provide information about credit risk in this case? In a common value auction with uncertain valuation, the competitor's offer should provide information about credit risk.

Before discussing default results, we first establish that the difference in contract terms between lenders determines lender choice. For this analysis, we focus only on withdrawn applications. We do not have data on borrowers who initially turn down the lenders' contract offers, and we do not include canceled and expired listings because cancellation and expiration are outside the borrowers' control, and thus the contract terms on these loans should not have an impact on lender choice.

For all Lending Club loans which match to a withdrawn Prosper listing, we show a histogram of interest rate differential (offered Prosper rate minus the rate on the Lending Club loan) in Figure 3. 73% of borrowers who withdrew their loans from Prosper received a lower interest rate from Lending Club. This is consistent with the idea that borrowers withdraw their loans from Prosper to take out loans from Lending Club. However, the magnitude is much too low for interest rate to be borrowers' only reason for switching lenders.[16]

[16] We would expect 90% in this case: 100% for the 80% of matches that are true positives and 50% for the

[Place Figure 3 about here]

Because interest rate difference does not seem to fully explain why borrowers withdraw their Prosper listings and take Lending Club loans, we also consider loan amount. Perhaps some lenders are choosing to withdraw listings in order to receive a larger loan rather than a lower interest rate loan. In Figure 4, we show a histogram of loan amount difference (offered Prosper amount minus Lending Club loan amount). The figure shows that 98% of borrowers who withdrew from Prosper to take out Lending Club loans received larger loans from Lending Club. Loan amount thus seems to be a more important consideration than interest rate to borrowers choosing which loan offer to accept.

[Place Figure 4 about here]

The winner's curse may occur from differences in loan amount and not just differences in interest rate. The mechanism is slightly more complex than with interest rate offers. For interest rate offers, lower interest rate offers result in lower returns mechanically. Higher loan amounts do not mechanically result in lower returns, however, the additional debt burden associated with higher loan amounts can result in lower returns.[17] Adverse selection, where the worst borrowers choose to take out the largest loans, may also be a factor. In any case, higher loan amounts are associated with higher default rates, and borrowers may experience a winner's curse from loan amount as well as from interest rate.

We test whether the interest rates and loan amounts on canceled/withdrawn/expired Prosper listings have explanatory power for the performance of Lending Club loans made to the same individual. As in the previous sections, we use a Cox proportional hazards model of default. In this case, the difference between the competitor's interest rate and the true interest rate on the loan (*Interest Rate Diff*) is the variable of interest. To avoid multicollinearity issues, we use interest rate dummy variables instead of a continuous measure of interest rate. All loans are included in the sample, not just canceled/withdrawn/expired

20% that are false positives.

[17] See Fuster and Willen (2017) and Tracy and Wright (2016) for evidence on payment size and default.

loans. Loans without corresponding Prosper listings are coded as having an *Interest Rate Diff* and *Amount Diff* of 0. The choice of this value has no effect on the value or standard error of the coefficients of interest, as the choice is absorbed by the canceled/withdrawn/expired coefficients. In our regressions specification, the hazard rate at time t for individual i is specified as:

$$\lambda_i(t) = \lambda_0(t)\exp(X_i'\beta_i + \mathit{Interest\ Rate}'\beta_2 + \gamma \mathit{Interest\ Rate\ Diff}_i + \delta \mathit{Amount\ Diff}_i)$$

where $\lambda_0(t)$ is the baseline hazard rate at time t, X_i is a vector of controls and $\mathit{Interest\ Rate}_i$ is a vector of interest rate indicators. γ and δ are the coefficients of interest. If $\gamma > 0$, then the competitor's interest rate offer provides credit risk information not present in Lending Club's own model. If $\delta < 0$, then the competitor's amount offer provides such information.

We report regression statistics in Table 8. In column 1, we include only interest rate fixed effects, expired/canceled/withdrawn indicators, and the difference in interest rate offers. In column 2, we add state, and issue time indicators; in column 3, we add other contract variables besides interest rate (as well as the differences between the contract offers at Prosper versus Lending Club); in column 4, we add credit variables.

[**Place Table 8 about here**]

The coefficient on *Interest Rate Diff* is significantly positive in all specifications. The coefficient size is 2.3 in the fullest specification and very close to that value in the other specifications. This size of the coefficient can be interpreted as follows: for every percentage point increase in interest rate difference, the instantaneous probability of default rises by 2.3%. This value should be interpreted as a lower bound. Although our matching algorithm has a rate of false positives of about 20%, the coefficient of interest is likely attenuated much more than this. This is because the largest interest rate differences are the most likely to be false positives. We cannot, however, estimate the exact level of attenuation bias.

The coefficient on *Amount Diff* is significantly negative. For every $\$1,000$ decrease in the competitor's amount offer, borrowers are 2% less likely to default. The average borrower in the withdrawn sample is 20% more likely to default than a borrower who received equal loan amount offers from both lenders. The average borrower in the withdrawn sample is 10% more likely to default than an individual who received equal interest rate offers from both lenders. This means that, in terms of default magnitudes, loan amount is responsible for winner's curse effect about twice that of interest rate.

These results demonstrate that Lending Club faces a winner's curse problem because of the possibility of higher interest rate offers from Prosper, as well as because of the possibility of smaller loan amount offers from Prosper. When the Prosper interest rate is higher, the default rate is higher, and when the Prosper loan amount is lower, the default rate is also higher. For winning bids, the Prosper interest rate is necessarily higher than the Lending Club rate (in our sample by an average of 3.5 p.p.), and so Lending Club suffers a winner's curse due to this interest rate difference. For loan amount differences, the effect is even more clear, with the average borrower receiving a $9,950 larger loan with Lending Club than would have received from Prosper. It is unsurprising that the winner's curse occurs on both interest rate and loan amount dimensions. However, in light of the strong focus on interest rates in both the theoretical and empirical literatures on informational frictions in lending markets, it is surprising that the winner's curse is actually more evident when considering loan amount than when considering interest rate.

4 Conclusion

We provide evidence of the winner's curse in an online personal lending market. Borrowers who were rejected by a competing lender are twice as likley to default as borrowers who were not rejected by the competing lender, conditional on receiving the same contract. Borrowers who had their loan listings with a competitor canceled are 75% more likely to default. These two results show that both of the lenders in our sample face a winner's curse due to the possibility of rejection by their competitor. Among borrowers who were offered contracts by a competitor, borrowers who were offered higher rates or smaller loan amounts are more likely to default. This shows that the lender faces a winner's curse because, for any borrower who accepts its offer, the competitor likely offered a higher interest rate and/or lower loan amount, which we show empirically. Surprisingly, the winner's curse is much stronger along the loan amount dimension than the interest rate dimension. Borrowers are often willing to take a higher interest rate loan if it is larger, but are rarely willing to take a smaller loan, even at a lower interest rate. The observed relationship between competitor's offer and default is also much stronger for loan amount than for interest rate.

Our results make four important contributions to our understanding of the winner's curse in lending. Firstly, we provide evidence of the winner's curse in general, rather than of the winner's curse when one lender has superior information. We study two competing lenders, showing that both face a winner's curse problem due to competition with each other. Secondly, we show that lender choice, and correspondingly, the winner's curse, is dependent not loan amount as well as interest rate. Thirdly, we provide evidence of the winner's curse in lending using data on how two competing lenders treated the same borrowers application, which has never been done before and helps establish the winner's curse much more definitively than can be done without such data. Lastly, we provide evidence of the winner's curse in market where interest rates are set purely by algorithm rather than by credit officers. Informational frictions occur in such markets, just is in more often discussed traditional lending markets, even in the absence of soft information.

References

Agarwal, S., Amromin, G., Ben-David, I., and Evanoff, D. D. Loan product steering in mortgage markets. *NBER Working Paper No. w22696*, 2016a.

Agarwal, S., Chomsisengphet, S., and Liu, C. The importance of adverse selection in the credit card market: Evidence from randomized trials of credit card solicitations. *Journal of Money, Credit and Banking*, 42(4):743–754, 2010.

Agarwal, S., Chomsisengphet, S., and Liu, C. An empirical analysis of information asymmetry in home equity lending. *Journal of Financial Services Research*, 49(1):101–119, 2016b.

Agarwal, S., Chomsisengphet, S., Liu, C., and Souleles, N. S. Benefits of relationship banking: evidence from consumer credit markets. 2009.

Agarwal, S. and Hauswald, R. Distance and private information in lending. *Review of Financial Studies*, 23(7):2757–2788, 2010.

Aigner, D. J. Regression with a binary independent variable subject to errors of observation. *Journal of Econometrics*, 1(1):49–59, 1973.

Akerlof, G. A. The market for lemons: Quality uncertainty and the market mechanism. *The Quarterly Journal of Economics*, 84(3):488–500, 1970.

Ausubel, L. M. Adverse selection in the credit card market. Technical report, working paper, University of Maryland, 1999.

Bajari, P. and Hortacsu, A. The winner's curse, reserve prices, and endogenous entry: Empirical insights from ebay auctions. *RAND Journal of Economics*, pages 329–355, 2003.

Balyuk, T. Financial innovation and borrowers: Evidence from peer-to-peer lending. 2016.

Berger, A. N. The economic effects of technological progress: Evidence from the banking industry. *Journal of Money, Credit, and Banking*, 35(2):141–176, 2003.

Berger, A. N. and Hannan, T. H. The price-concentration relationship in banking. *The Review of Economics and Statistics*, pages 291–299, 1989.

Berger, A. N. and Udell, G. F. Relationship lending and lines of credit in small firm finance. *Journal of Business*, pages 351–381, 1995.

Broecker, T. Credit-worthiness tests and interbank competition. *Econometrica: Journal of the Econometric Society*, pages 429–452, 1990.

Bubb, R. and Kaufman, A. Securitization and moral hazard: Evidence from credit score cutoff rules. *Journal of Monetary Economics*, 63:1–18, 2014.

Capen, E. C., Clapp, R. V., Campbell, W. M., et al. Competitive bidding in high-risk situations. *Journal of petroleum technology*, 23(06):641–653, 1971.

Cerqueiro, G., Degryse, H., and Ongena, S. Rules versus discretion in loan rate setting. *Journal of Financial Intermediation*, 20(4):503–529, 2011.

Cressy, R. and Toivanen, O. Is there adverse selection in the credit market? *Venture Capital: An International Journal of Entrepreneurial Finance*, 3(3):215–238, 2001.

Degryse, H. and Ongena, S. Distance, lending relationships, and competition. *The Journal of Finance*, 60(1):231–266, 2005.

Duarte, J., Siegel, S., and Young, L. Trust and credit: The role of appearance in peer-to-peer lending. *Review of Financial Studies*, 25(8):2455–2484, 2012.

Everett, C. R. Group membership, relationship banking and loan default risk: the case of online social lending. *Relationship Banking and Loan Default Risk: The Case of Online Social Lending (March 15, 2010)*, 2010.

Fishman, M. J. and Parker, J. A. Valuation, adverse selection, and market collapses. *Review of Financial Studies*, 28(9):2575–2607, 2015.

Fuster, A. and Willen, P. S. Payment size, negative equity, and mortgage default. *American Economic Journal: Economic Policy (forthcoming)*, 2017.

Greenbaum, S. I., Kanatas, G., and Venezia, I. Equilibrium loan pricing under the bank-client relationship. *Journal of Banking & Finance*, 13(2):221–235, 1989.

Hauswald, R. and Marquez, R. Information technology and financial services competition. *Review of Financial Studies*, 16(3):921–948, 2003.

Hauswald, R. and Marquez, R. Competition and strategic information acquisition in credit markets. *Review of Financial Studies*, 19(3):967–1000, 2006.

Herzenstein, M., Sonenshein, S., and Dholakia, U. M. Tell me a good story and I may lend you money: the role of narratives in peer-to-peer lending decisions. *Journal of Marketing Research*, 48(SPL):S138–S149, 2011.

Iyer, R., Khwaja, A. I., Luttmer, E. F., and Shue, K. Screening peers softly: Inferring the quality of small borrowers. *Management Science*, 62(6):1554–1577, 2015.

Jacobson, T. and Roszbach, K. Bank lending policy, credit scoring and value-at-risk. *Journal of Banking & Finance*, 27(4):615–633, 2003.

Karlan, D. and Zinman, J. Observing unobservables: Identifying information asymmetries with a consumer credit field experiment. *Econometrica*, 77(6):1993–2008, 2009.

Keys, B. J., Mukherjee, T., Seru, A., and Vig, V. Did securitization lead to lax screening? evidence from subprime loans. *The Quarterly Journal of Economics*, 125(1):307–362, 2010.

Lin, M., Prabhala, N. R., and Viswanathan, S. Judging borrowers by the company they keep: Friendship networks and information asymmetry in online peer-to-peer lending. *Management Science*, 59(1):17–35, 2013.

Michels, J. Do unverifiable disclosures matter? Evidence from peer-to-peer lending. *The Accounting Review*, 87(4):1385–1413, 2012.

Puri, M. and Rocholl, J. On the importance of retail banking relationships. *Journal of Financial Economics*, 89(2):253–267, 2008.

Rajan, R. G. Insiders and outsiders: The choice between informed and arm's-length debt. *The Journal of Finance*, 47(4):1367–1400, 1992.

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

Sharpe, S. A. Asymmetric information, bank lending, and implicit contracts: A stylized model of customer relationships. *The Journal of Finance*, 45(4):1069–1087, 1990.

Stiglitz, J. E. and Weiss, A. Credit rationing in markets with imperfect information. *The American Economic Review*, 71(3):393–410, 1981.

Tracy, J. and Wright, J. Payment changes and default risk: The impact of refinancing on expected credit losses. *Journal of Urban Economics*, 93:60–70, 2016.

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

Table 1: Lending Club Loans

Statistic	Mean	St. Dev.	Pctl(25)	Median	Pctl(75)	N
Income	76,909.080	86,918.010	46,000.000	65,000.000	91,000.000	1,415,299
Credit History Length	199.830	91.876	137	181	246	1,321,818
Debt-to-Income	18.369	8.546	12.050	17.840	24.240	1,415,134
Employment Length	34.972	30.205	12	12	60	1,335,710
Home Ownership	0.603	0.489	0	1	1	1,414,576
Incomce Verification	0.696	0.460	0	1	1	1,415,303
Inquiries	0.641	0.949	0	0	1	1,415,273
Interest Rate	0.132	0.046	0.098	0.128	0.159	1,415,303
Loan Amount	14,756.880	8,675.654	8,000	12,800	20,000	1,415,303
Term	42.796	10.813	36	36	60	1,415,303
Percent Funded	1.000	0.013	1.000	1.000	1.000	1,415,303
Loan Age	25.666	17.235	13	22	35	1,415,303
Default	0.091	0.287	0	0	0	1,415,303
Open Credit Lines	11.666	5.488	8	11	14	1,415,274

Notes to Table: Summary statistics for all loans issued by Lending Club are reported above. Credit history length is in months. Debt-to-Income is the ratio of monthly debt payments (including the P2P loan to monthly income). Employment length is in months, but is censored at 10 years. Credit inquiries are hard inquiries in last 6 months. Loan age is months since issue. Term is an indicator variable taking a value of 1 if the loan is a 60-month maturity loan and 0 if a 36-month maturity loan.

Table 2: Prosper Loans

Statistic	Mean	St. Dev.	Pctl(25)	Median	Pctl(75)	N
Income	74,532.320	1,019,969.000	41,600.000	60,000.000	85,000.000	384,965
Credit History Length	217.616	102.248	148	205	275	367,452
Debt-to-Income	25.850	12.475	17.000	24.000	33.000	372,324
Employment Length	67.707	39.977	28	76	108	376,879
FICO	5.685	1.977	4	5	7	313,336
Home Ownership	0.436	0.496	0	0	1	384,965
Incomce Verification	0.968	0.176	1	1	1	384,965
Inquiries	1.150	1.728	0	1	2	367,465
Interest Rate	0.156	0.063	0.108	0.146	0.193	384,967
Loan Amount	11,828.230	8,150.527	5,300.000	10,000.000	16,000.000	384,967
Term	43.452	11.105	36	36	60	384,967
Percent Funded	0.999	0.014	1.000	1.000	1.000	384,965
Loan Age	26.290	29.856	8	18	32	384,967
Default	0.060	0.237	0	0	0	384,967
Open Credit Lines	10.147	5.032	7	9	13	360,889

Notes to Table: Summary statistics for all loans issued by Prosper are reported above. Credit history length is in months. Debt-to-Incomeis the ratio of monthly debt payments (including the P2P loan to monthly income). Employment length is in months, but is censored at 10 years. Credit inquiries are hard inquiries in last 6 months. Loan age is months since issue. Termis an indicator variable taking a value of 1 if the loan is a 60-month maturity loan and 0 if a 36-month maturity loan.

Table 3: Summary Statistics

	Prosper	Lending Club	LC Reject
Income	71227	72218	65764
Inquiries	1.33543	0.816393	0.9171271
FICO	696.5957	700.7903	676.5
Loan Amount	9678.785	13657.75	8738.807
Credit History Length	17.69986	14.52355	19.4054
Interest Rate	0.1800368	0.1394538	0.1954779

Notes to Table: Summary statistics for all loans issued by Prosper (column 1) and Lending Club (column 2) are reported above. Summary statistics for all Prosper loans made to borrowers rejected by Lending Club are reported in column 3.

Table 4: Rejected Borrowers and Default: Hazard Model

	Dependent variable:
	Default
Interest Rate	5.919*** (0.688)
Rejected	0.798*** (0.240)
City Size	−0.046*** (0.016)
Term	0.170* (0.103)
Issue Month	−0.007 (0.005)
FICO	0.001 (0.001)
Income	−0.0001*** (0.00001)
Inquiries	0.135*** (0.021)
Credit History Length	−0.006 (0.005)
Open Credit Lines	−0.008** (0.003)
Observations	3,386
R^2	0.067

Notes to Table: Regression statistics for a Cox proportional hazards with default as the dependent variable. *p<0.1; **p<0.05; ***p<0.01.

Table 5: Rejected Borrowers and Default: Propensity Scoring

	Dependent variable:	
	Default	
	(IPTW)	(Propensity Matching)
Interest Rate	5.920*** (0.969)	9.472*** (2.134)
Rejected	0.797** (0.341)	0.572** (0.234)
City Size	−0.046** (0.023)	−0.087* (0.047)
Term	0.170 (0.145)	−0.473 (0.320)
Issue Month	−0.007 (0.007)	−0.025* (0.014)
FICO	0.001 (0.001)	0.008** (0.003)
Income	−0.0001** (0.00002)	−0.00002 (0.00004)
Inquiries	0.135*** (0.030)	0.128 (0.099)
Credit History Length	−0.006 (0.008)	−0.004 (0.018)
Open Credit Lines	−0.008* (0.005)	−0.007 (0.010)
Observations	3,386	200
R^2	0.034	0.245

Notes to Table: Regression statistics for a Cox proportional hazards with default as the dependent variable reported above. Column 1 is an inverse probability of treatment weighted logit. Column 2 is a logit on a propensity score matched sample in which each rejected borrower is matched to 3 non-rejected borrowers. *p<0.1; **p<0.05; ***p<0.01.

Table 6: Rejected Borrowers and Default: Exact Matching

	Dependent variable:
	Default
Interest Rate	8.671*** (2.239)
Rejected	0.679*** (0.233)
City Size	−0.068 (0.047)
Term	−0.339 (0.307)
Issue Month	−0.019 (0.015)
FICO	0.006* (0.003)
Income	0.00002 (0.00004)
Inquiries	0.078 (0.091)
Credit History Length	0.006 (0.018)
Open Credit Lines	−0.00001 (0.009)
Observations	193
R^2	0.186

Notes to Table: Regression statistics for a Cox proportional hazards on a matched sample with default as the dependent variable reported above. The sample is exactly matched on all contract terms (so that treated borrowers are being compared to untreated borrowers who received exactly the same contract). Each rejected borrower is matched to 3 non-rejected borrowers. *p<0.1; **p<0.05; ***p<0.01.

Table 7: Canceled Listings and Default: Cox Regressions

	Dependent variable:			
	Default			
	(Standard Cox)	(IPTW)	(Propensity Matching)	(Exact Matching)
Interest Rate	11.469*** (0.066)	11.469*** (0.094)	10.925*** (0.926)	10.112*** (0.894)
Rejected	0.498*** (0.088)	0.498*** (0.125)	0.447*** (0.082)	0.425*** (0.093)
City Size	0.00000*** (0.00000)	0.00000* (0.00000)	−0.00000 (0.00001)	−0.00000 (0.00001)
Term	0.0003 (0.0003)	0.0003 (0.0004)	0.001 (0.004)	0.004 (0.004)
Issue Month	−0.008*** (0.0002)	−0.008*** (0.0002)	−0.017*** (0.003)	−0.007** (0.003)
FICO	(0.000)	(0.000)	(0.000)	(0.000)
Income	−0.00000*** (0.00000)	−0.00000*** (0.00000)	−0.00000 (0.00000)	−0.00000* (0.00000)
Inquiries	0.104*** (0.002)	0.104*** (0.003)	0.023 (0.041)	0.035 (0.040)
Credit History Length	−0.0002*** (0.00001)	−0.0002*** (0.00001)	−0.001** (0.001)	−0.001* (0.001)
Open Credit Lines	0.010*** (0.001)	0.010*** (0.001)	0.010 (0.010)	0.012 (0.008)
Observations	1,415,273	1,415,273	3,360	3,360
R^2	0.036	0.018	0.081	0.069

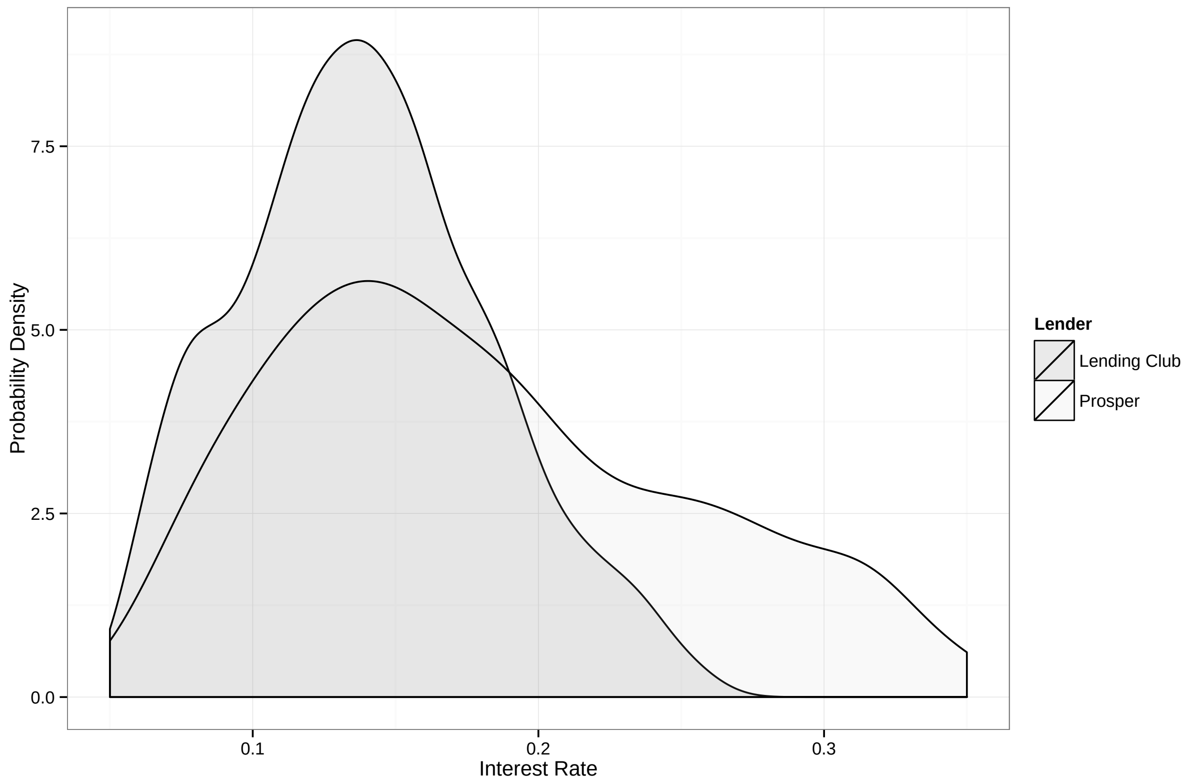
Notes to Table: Regression statistics for a Cox proportional hazards with default as the dependent variable. Column 1 is a standard Cox. Column 2 uses inverse probability of treatment weighting (IPTW). Column 3 uses propensity score matching. Column 4 uses exact matching based on all contract variables (interest rate, term, and loan amount). Columns 3 and 4 use 3-to-1 matching (3 non-canceled loans matched to every canceled loan). *p<0.1; **p<0.05; ***p<0.01.

Table 8: Interest Rate Difference and Default

	Dependent variable:			
	default			
	(1)	(2)	(3)	(4)
Expired	0.021	−0.582*	−0.761**	−0.647**
	(0.311)	(0.312)	(0.321)	(0.321)
Withdrawn	0.471**	0.107	−0.120	−0.071
	(0.185)	(0.186)	(0.208)	(0.208)
Canceled	0.424***	0.404***	0.357***	0.374***
	(0.087)	(0.087)	(0.091)	(0.091)
Interest Rate Diff	2.256**	2.508**	2.552**	2.298**
	(1.105)	(1.104)	(1.107)	(1.105)
Term			−0.003***	−0.003***
			(0.0003)	(0.0003)
Term Diff			−0.003	−0.002
			(0.006)	(0.006)
Loan Amount			−0.00000***	0.00001***
			(0.00000)	(0.00000)
Amount Diff			−0.00002**	−0.00002**
			(0.00001)	(0.00001)
Income				−0.00000***
				(0.00000)
Inquiries				0.095***
				(0.002)
Credit History Length				−0.0001***
				(0.00001)
Open Credit Lines				0.009***
				(0.001)
Interest Rate	Yes	Yes	Yes	Yes
State	No	Yes	Yes	Yes
Half-Year Indicators	No	Yes	Yes	Yes
Observations	1,415,364	1,415,364	1,415,364	1,415,334
R^2	0.032	0.040	0.040	0.042
Note:			*p<0.1; **p<0.05; ***p<0.01	

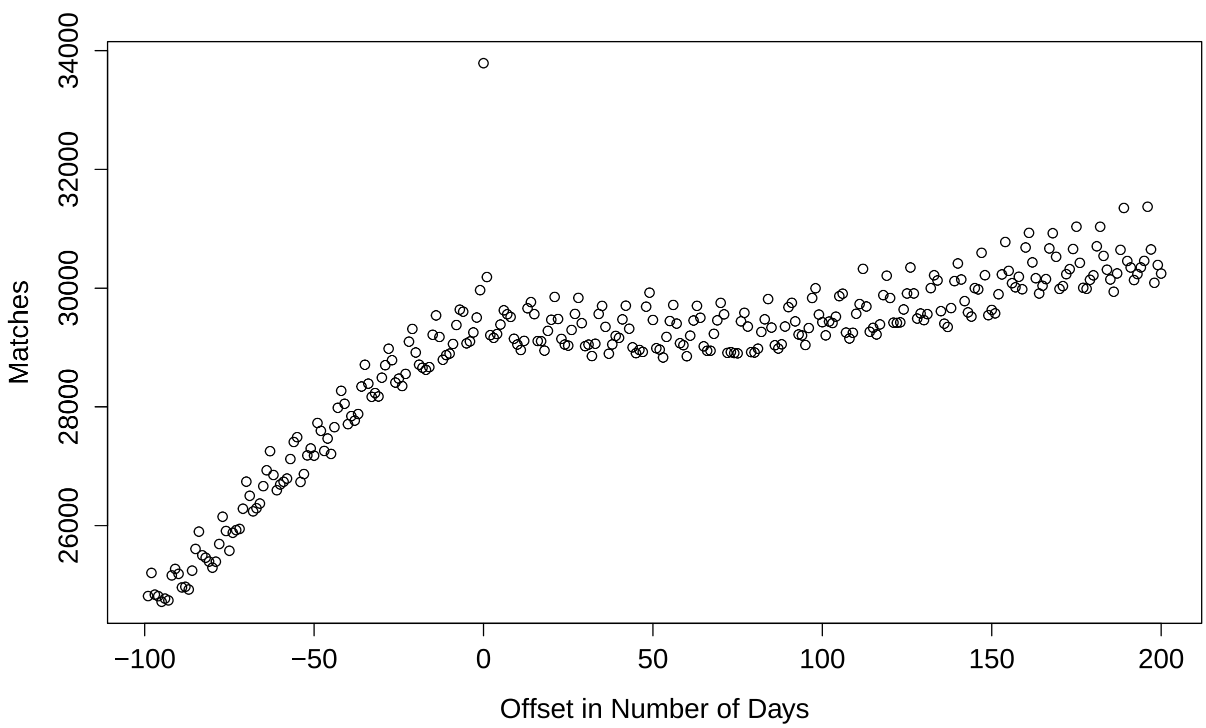
Notes to Table: Regression statistics for a Cox proportional hazards with default as the dependent variable. All columns are a standard equally-weighted Cox proportional hazards model, but with different specifications. *p<0.1; **p<0.05; ***p<0.01.

Figure 1: Histogram of Interest Rates



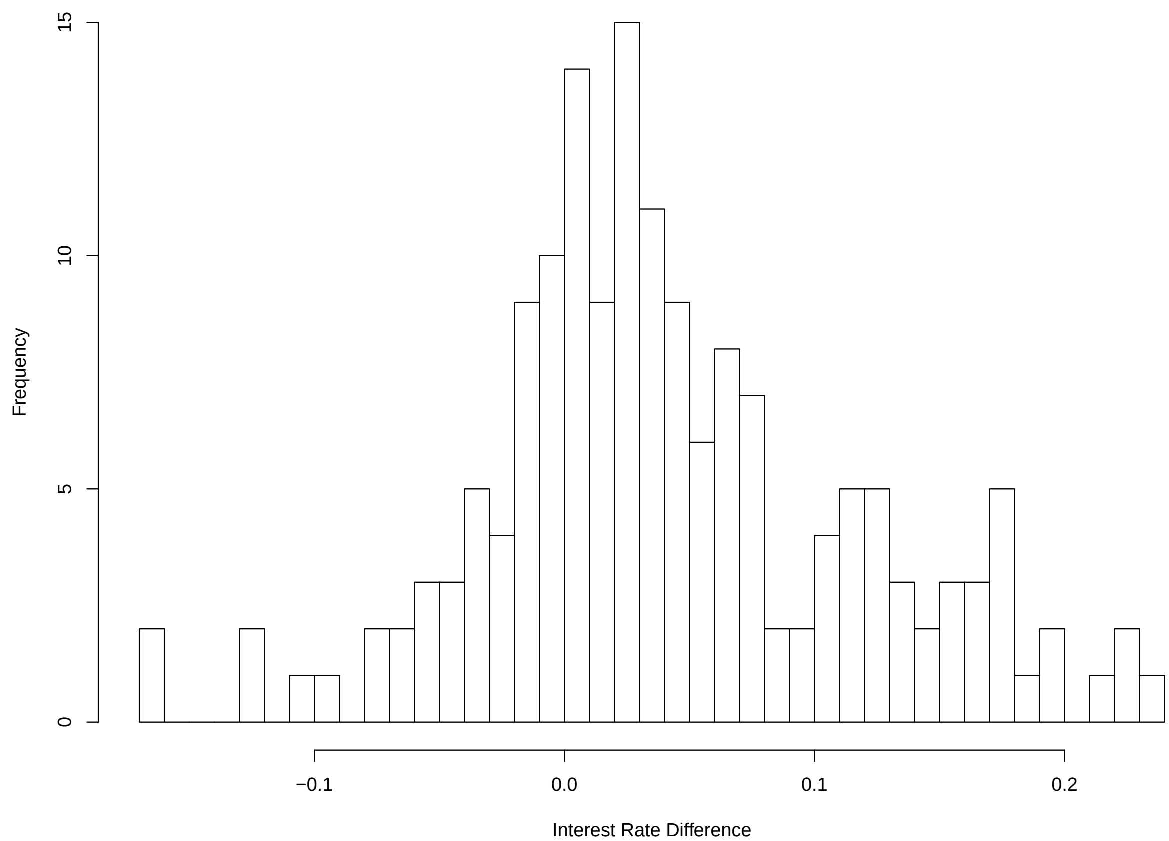
Notes to Figure: Histogram of interest rates for all loans issued through Lending Club and Prosper.

Figure 2: Placebo Matches



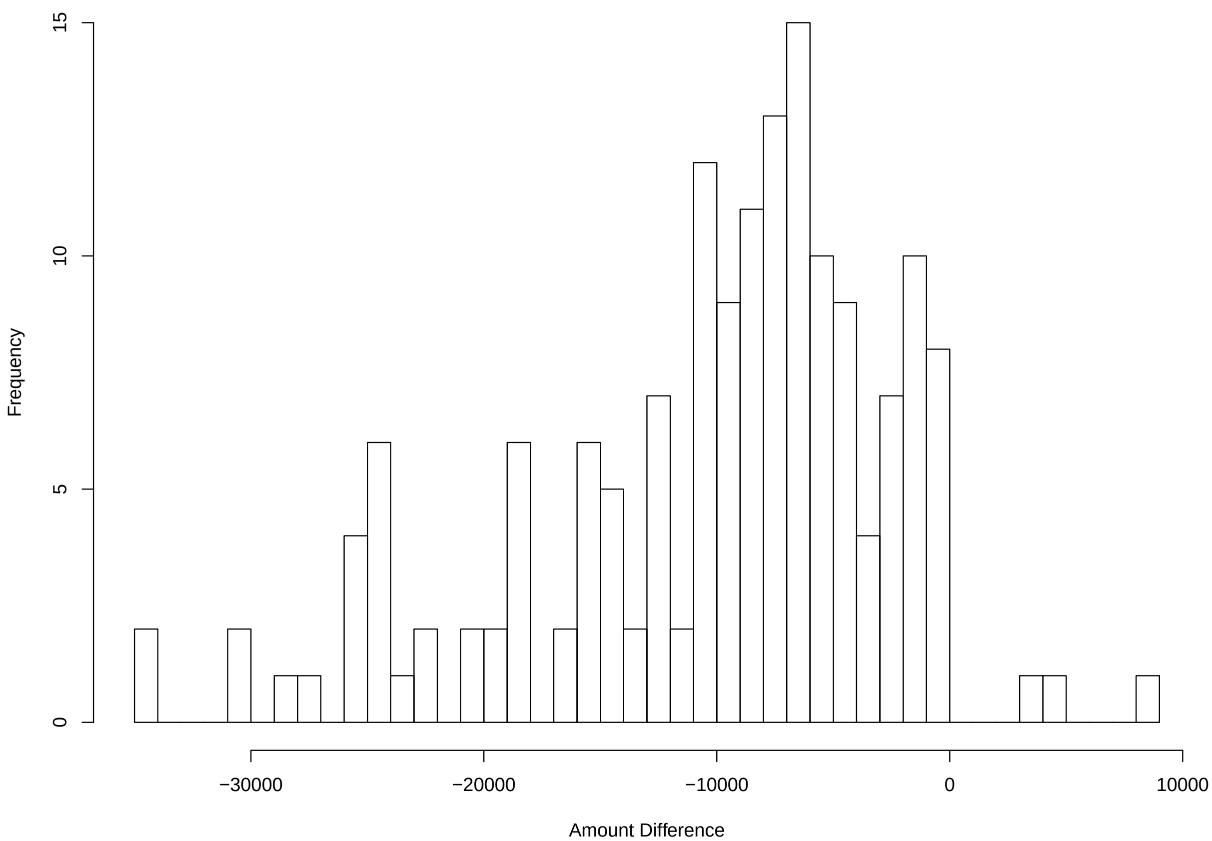
Notes to Figure: Graph shows the number of Prosper loans (out of $180,000$ total) which match to an application in the Lending Club rejection data set. To match to a Lending Club loan, the Prosper loan must come X days after the Lending Club rejection from the same city (where X is the x-axis in the graph).

Figure 3: Interest Rate Difference Between Lenders



Notes to Figure: Histogram of the interest rate difference (Prosper minus Lending Club) for borrowers with withdrawn Prosper listings and Lending Club loans.

Figure 4: Loan Amount Difference Between Lenders



Notes to Figure: Histogram of the loan amount difference (Prosper minus Lending Club) for borrowers with withdrawn Prosper listings and Lending Club loans.