Learning by Doing with Asymmetric Information: Evidence from Prosper.com[*]

Seth Freedman
University of Maryland

Ginger Zhe Jin
University of Maryland & NBER

February 1, 2010

Warning: This is an *academic* study using Prosper data from June 1, 2006 through July 31, 2008. Readers should *not* use it as an investment guide. Because none of the Prosper loans have reached their regular maturity, the loan performance reported in this paper is up to data availability as of August 1, 2008 (our data download date). Consequently, the estimated rate of return entails a number of assumptions. Any conclusion drawn from our study is subject to the validity of these assumptions.

JEL: D45, D53, D8, L81

---

[*]Seth Freedman, Department of Economics, University of Maryland, College Park. Phone: 301-405-3266, Email: freedman-s@econ.umd.edu. Ginger Zhe Jin, Department of Economics, University of Maryland, College Park. Phone: 301-405-3484, Fax: 301-405-3542, Email: jin@econ.umd.edu. We received constructive comments from Larry Ausubel, Robert Hampshire, John Haltiwanger, Anton Korinek, Phillip Leslie, Russel Cooper, Hongbin Cai, Jim Brickley and various seminar attendants at Rochester, Toronto, and Northwestern Kellogg. Special thanks to Chris Larsen, Kirk Inglis, Nancy Satoda, Reagan Murray, Anurag Malik, and other Prosper personnel for providing us detailed information about prosper.com, to Jim Bruene for helping us understand the lending industry in general, and to Adam Weyeneth and other prosper lenders for sharing with us their experience on prosper.com. We are grateful to the UMD Department of Economics, the Kauffman Foundation, and the Net Institute (www.netinst.org) for their generous financial support. An earlier draft has been circulated under the title "Dynamic Learning and Selection." In comparison, this draft has updated the data from December 2007 to July 2008 and corrected several mistakes in the data work. All the remaining errors are our own.

Abstract

This paper examines the nature of information asymmetry in online peer-to-peer (P2P) lending markets. These markets use the Internet to match individual borrowers and lenders of consumer loans without financial institutions as intermediaries. Like other anonymous interactions, P2P lending may face additional information asymmetry as compared to offline because P2P lenders have less access to "hard" information such as borrower credit history, income, or employment. However, the shortage of "hard" information could be mitigated by "soft" information via online social networks. We examine this tradeoff using data from all requested and funded loans between June 1, 2006 and July 31, 2008 on Prosper.com.

We have three main findings. First, Prosper lenders understand the ordinal difference across credit grades but the incomplete disclosure of a borrower's credit history leads to additional adverse selection relative to traditional markets. Second, some social networks help to mitigate information asymmetry and others do not, depending on the institutional incentives. Third, lenders, especially those who joined Prosper early, did not fully understand the market risk. We estimate, on average, lenders would have expected an annualized internal rate of return of -0.62% to -1.38% on a dollar invested if they had correctly understood the risk distribution of Prosper loans. However, lender learning is effective in reducing the risk of funded loans over time. As a result, the market has excluded more and more sub-prime borrowers and evolved towards the population served by traditional credit markets.

1 Introduction

The Internet as a platform for peer-to-peer (P2P) transactions has extended to job search, dating, social networks, and recently consumer lending. While search cost savings may explain part of the growth, it is puzzling how commodities that feature significant information asymmetry between buyers and sellers can flourish on the Internet. In particular, since many P2P platforms keep individual users anonymous to each other, the extent of information asymmetry is likely to be exaggerated on the Internet, and an Akerlof (1970) type unraveling of adverse selection could be more salient online than offline. Nevertheless, P2P lending has flourished on the Internet, despite a credit crisis. Does this suggest that P2P platforms are able to provide additional "soft" information about borrower risk to mitigate the lack of "hard" information due to anonymity? How do P2P lenders utilize this information? Answers to these questions are not only important for the long-run viability of P2P platforms, but will deepen the understanding of financial markets.

To address these questions, we examine transaction level data from Prosper.com between June 1, 2006 and July 31, 2008. As the first P2P lending website in the US,[1] Prosper.com was officially started in February 2006. As of July 31, 2008, Prosper had attracted 750,000 members and originated loans of over 160 million dollars.[2] Aside from operation style,[3] Prosper.com differs from a traditional lending market in two ways.

First, anonymous online interaction presents new challenges that do not apply to traditional banks. Although Prosper conducts identity checks and pulls a borrower's credit history from a major credit bureau (Experian), Prosper only posts a categorical credit grade (for privacy reasons) and the lender never observes the borrower's exact credit score. Furthermore, individual lenders, by definition smaller and less professional than financial institutions, may not have the expertise to screen risks ex ante and monitor loan performance ex post.[4] These features suggest

[1]Zopa.com (of UK) is the first peer-to-peer lending website world wide.

[2]The quick expansion of Prosper has coincided with a number of similar new P2P lending sites in the US. The best known examples are Kiva.org (incorporated November 05), Smava (launched in February 2007), Lending Club (open May 24, 2007 as part of Facebook), MyC4 (launched in May 2007), Globefunder (launched in October 2, 2007), and Zopa US (us.zopa.com, open December 4, 2007).

[3]Prosper.com automates the borrower-lender match via real-time auctions. With sufficient scale, this format may generate significant savings in operation costs, implying lower interest rates for individual borrowers and better returns for individual lenders. Unfortunately, we do not have sufficient data to measure the cost difference.

[4]Because most loans are funded by multiple lenders (for the purpose of risk diversification), each individual lender may not have the full incentive to gather information before funding and monitor performance after funding. In the recent subprime mortgage crisis, a similar argument may apply to those traditional lenders that initiate high risk loans, repackage them in securities, and spread the risk to the rest of the market. In this sense, it is unclear whether P2P lenders have more or less incentives to care about the loan performance as compared to

that Prosper lenders may be subject to a greater risk of adverse selection and moral hazard, as compared to traditional banks.

On the other hand, Prosper has gradually adopted features to facilitate the use of social networks. It encourages borrowers and lenders to form online groups and establish friendships with other members. It allows group leaders and friends to offer endorsements for a specific listing and highlights bids from these social connections. Given the potential for the Internet to facilitate information flow among borrowers and lenders, it is hoped that the soft information conveyed via social networks may alleviate the shortage of hard information in credit history and borrower identity. However, unlike the typical micro-finance groups, such as the Grameen Bank, Prosper does not impose joint liability within a group. Nor does an endorsing friend share any legal responsibility as a co-signer. The effectiveness of these Internet-based social networks remains an empirical question.

In short, there are three types of borrower information in a Prosper listing: the hard information included in credit history drawn from Experian, the soft information conveyed by Prosper-specific social network variables, and the self-reported non-verifiable information which we refer to as "cheap-talk." Ravina (2007) and Pope and Sydnor (2008) have examined the impact of cheap-talk information such as race, gender, age and beauty contents of the image that a borrower voluntarily posts in the listing. In comparison, we focus on the other two types of information while controlling for as many "cheap-talk" variables as we can. Iyer et al. (2009) is closer to our study as they examine the extent to which Prosper lenders infer borrower's credit risk in the absence of the exact credit score. We emphasize the other side of the coin, namely how the *lack* of credit score information drives adverse selection. We also examine the tradeoff between hard and soft information as well as how lenders utilize the information over time.

Focusing on the hard credit history information, we show that lenders understand the ordinal difference across disclosed credit grades but the imprecise disclosure of borrower credit history has introduced additional adverse selection as compared to traditional banks. There is also strong evidence that lenders, especially those who joined prosper early, did not fully understand the risk they faced on the Internet. From loan payment history, we find that, on average, lenders would have expected an annualized internal rate of return (IRR) of between -0.62% and -1.38% on a dollar invested if they had correctly understood the risk distribution of Prosper loans. After examining alternative explanations, we believe the low return of Prosper loans is likely driven by lender mistakes instead of charity lending, mean reversion or unexpected macro shocks.

Regarding the soft information via social networks, evidence suggests that some social networks help to mitigate information asymmetry and others do not, depending on the institutional

traditional lenders.

incentives. For example, loans with friend endorsements *and* friend bids tend to have less missed payments and yield significantly higher rates of return than other loans. But loans with friend endorsements *but no* friend bids generate a lower rate of return than the loans without endorsements. On average, Prosper groups do not identify high quality borrowers either, as the default rate of group loans is significantly higher and the estimated rate of return is systematically lower than non-group loans. Group rating adds little information, and there is evidence that some group leaders have gamed the system to extract group leader awards from Prosper without adequately screening the risk of borrowers in the group.

The large heterogeneity in expected IRR by loan suggests that individual lenders do not fully understand the meaning of hard and soft information. Many Prosper lenders made mistakes in their initial loan selections, but they learn vigorously over time. In particular, we show that a lender is more likely to stop funding any new loan if his existing loans are late. Conditional on funding new loans, the new loans shy away from the credit grade of the mis-performing loans in his portfolio. Lenders also learn to avoid group loans over time, especially when group loans in their own portfolio have missed payments. The movement towards effective soft information is gradual, which explains why we still observe better-than-average rates of return on loans with desirable social network features.

Overall, we conclude that P2P lending faces a significant tradeoff in information asymmetry. The anonymous nature of P2P lending implies a shortage of hard information on the borrower's credit history, but some social networks, if combined with the right incentives, can mitigate the problem by generating soft information. Overtime, as lenders realize the actual risk on the Internet, the P2P market has excluded more and more subprime borrowers and evolved towards the population served by traditional credit markets. This suggests that, unless P2P platforms can improve the power of social networks for borrowers with low credit scores, P2P lending is likely to compete head-to-head with traditional banks in the future and would not provide a viable alternative for those excluded from traditional credit markets.

Our work contributes to a number of literatures. Theorists have examined the impact of information asymmetry on market failure. For example, Akerlof (1970) predicts that a market may not exist at all if it is plagued by adverse selection. Focusing on the credit market, Stiglitz and Weiss (1981) point out that information asymmetry can leave some borrowers unfunded even if they are willing to pay a price higher than the market rate. This phenomenon is referred to as credit rationing. Our data cannot identify credit rationing from other explanations, but we show that the key assumption underlying credit rationing, namely the non-monotonic relationship between rate of return and interest rate, does appear in our data.[5]

[5]A number of empirical papers have used offline data to document evidence of credit rationing (e.g. Jaffee 1971,

A rich empirical literature has examined the existence of adverse selection and moral hazard in offline settings such as car insurance (Chiappori and Salanie 2000), annuities (Finkelstein and Poterba 2004), and life insurance (Cawley and Philipson 1999). Most studies test the Rothschild and Stiglitz (1976) prediction that high-risk buyers are likely to choose more comprehensive coverage because they expect more benefits from the insurer, and results depend on the context with some studies finding evidence of positive correlation between risk and coverage and others not (Cohen and Siegelman 2009). We use a different identification strategy: instead of inferring information asymmetry from borrower's contract choice, we test adverse selection directly by using credit score information that is not publicly available to lenders. In this sense, our test is similar to Finkelstein and McGarry (2006) and Finkelstein and Poterba (2006), who utilize buyer information that is either not available to or not used by insurers.

As the loan performance unravels, we also observe how the market evolves as the degree of information asymmetry changes over time. The learning patterns identified in our paper are similar to the reinforcement model in the learning literature[6] but different from most learning studies in the insurance context because our lenders learn the meaning of observables across borrowers instead of the specific unobservable within a borrower. This difference occurs as Prosper borrowers were not allowed to borrow two loans until the end of our sample.[7]

Our paper also relates to the literature of informal lending and micro-finance. Previous researchers have argued that informal lenders and micro-finance institutions have an information advantage over traditional banks because they utilize borrowers' social networks to ensure good risks (e.g. La Ferrara 2003, Udry 1994, Hoff and Stiglitz 1990). We show evidence both for and against this argument. It seems some social networks permitted on Prosper can clear information hurdles if they are combined with the right incentives.

Finally, as a fresh example of an online marketplace, the experience of Prosper highlights how the interaction of information and social relationships can contribute to the rise of e-commerce. Unlike previous studies that document the segmentation between online and offline markets (Jin and Kato 2007, Hendel, Nevo and Ortalo-Magne 2008), we show that Prosper is *converging* with

Cox and Japelli 1990, Berger and Udell 1992, Voridis 1993) and liquidity constraints (e.g. Souleles 1999, Parker 1999, Gross and Souleles 2002, Adam, Einav and Levin 2009). None of them can directly test the non-monotonic relationship between rate of return and interest rate because many low rate of return projects are excluded from observational data by definition. We have an opportunity to observe some of the low return investment, partly due to lender mistakes.

[6]See Salmon (2001) for a review of how economists identify reinforcement learning from belief learning in lab experiments.

[7]For this reason, there is little room for a lender to learn a borrower's type and then apply that information in future lending to the same borrower. However, the lender may associate the borrower's observable characteristics to the ex post performance and apply the updated meaning of the observables to other prospective borrowers.

the traditional market except for the positive signals contained in some social network variables.

The rest of the paper is organized as follows. Section 2 describes the background of Prosper.com and its major competitors in traditional lending. Section 3 describes the data, defines the sample, and provides a simple summary of the Prosper population over time. Sections 4 and 5 present evidence on the hard information of borrower credit history and soft information of social networks respectively. Section 6 uses the estimated internal rate of return to summarize lender understanding of information. Section 7 describes how individual lenders have learned to cope with the hard and soft information over time. A short conclusion is offered in Section 8.

2 Background

2.1 Market Setup

All Prosper loans are fixed rate, unsecured, three-year, and fully amortized with simple interest. Loan can range from $1,000 to $25,000. There is no penalty for early payment. By the end of our sample period (July 31, 2008), the loans are not tradable in any financial market,[8] which means a lender that funds a loan is tied up with the loan until full payment or default. Upon default Prosper hires collection agencies and any money retrieved in collections is returned to the loan's lenders.

Before a potential borrower lists a loan application on Prosper, Prosper authenticates the applicant's social security number, driver's license, and address. Prosper also pulls the borrower's credit history from Experian, which includes the borrower's credit score and historical credit information such as total number of delinquencies, current delinquencies, inquiries in the last six months, etc.[9] If the credit score falls into an allowable range, the borrower may post an eBay-style listing specifying the maximum interest rate she is willing to pay, the requested loan amount, the duration of the auction (3-10 days),[10] and whether she wants to close the listing immediately after it is fully funded (called autofunding). In the listing, the borrower may also describe herself, the purpose of the loan, the city of residence, how she intends to repay the loan, and any other information (including an image) that she feels may help fund the loan. In the same listing, Prosper will post the borrower's credit grade (computed based on credit score), home ownership status, debt-to-income ratio, and other credit history information.[11]

[8] In October 2008, Prosper began the process of registering with the SEC in order to offer a secondary market, which was approved in July 2009 and therefore is outside of our sample period.

[9] The credit score reported uses the Experian ScorePLUS model, which is different from a FICO score, because it intends to better predict risks for new accounts.

[10] As of April 15, 2008 all listings have a duration of 7 days.

[11] The debt information is available from the credit bureau, but income is self-reported. Therefore, the debt-to-income ratio reported in the listing is not fully objective.

Like borrowers, a potential lender must provide a social security number and bank information for identity confirmation. Lenders can browse listing pages which include all of the information described above, plus information about bids placed, the percent funded, and the listings current prevailing interest rate. To view historical market data, a lender can download a snapshot of all Prosper records from Prosper.com (updated daily), use a Prosper tool to query desired statistics, or visit a third party website that summarizes the data. Interviews conducted at the 2008 Prosper Days Conference suggest that there is enormous heterogeneity in lender awareness of the data, ability to process the data, and intent to track the data over time.

The auction process is similar to proxy bidding on eBay. A lender bids on a listing by specifying the lowest interest rate he will accept (so long as it is below the borrower's specified maximum rate) and the amount of dollars he would like to contribute (any amount above $50[12]). A listing is fully funded if the total amount bid exceeds the borrower's request. If the borrower chooses the autofunding option, the auction will end immediately and the borrower's maximum interest rate applies. Otherwise, the listing remains open and new bids will compete down the interest rate. Lenders with the lowest specified minimum interest rate will fund the loan and the prevailing rate is set as the minimum interest rate specified by the first lender excluded from funding the loan. We will refer to the resulting interest rate as the contract rate.

Prosper charges fees to both borrowers and lenders. These fees have changed over time, but in general borrowers pay a closing fee when their loan originates ranging from 1% to 3% depending on credit grade (there is no fee for posting a listing). If a borrower's monthly payment is 15 days late, a late fee is charged to the borrower and transferred to lenders in the full amount. Lenders are charged an annual servicing fee based on the current outstanding loan principal.[13] The lender fee has ranged from 0.5% to 1% depending on credit grade.

In legal terms, Prosper loans are first issued by Prosper and then sold to individual lenders. Prior to April 15, 2008, Prosper was subject to state usury laws which specify the maximum interest rate a lender can charge. The binding law was that of the borrower's state of residence, and within each state regulations depended on whether Prosper held a consumer loan license in that state. The interest rate caps varied from 6% to 36% across states. On April 15, 2008, Prosper became a partner of WebBank, a Utah-chartered industrial bank, which allows the site to circumvent most state usury laws. Following this partnership, the interest rate cap became a universal Prosper implemented 36% (except for Texas and South Dakota).

[12] After Prosper registered with SEC in July 2009, the minimum bid was reduced to $25.

[13] This fee is accrued the same way that regular interest is accrued on the loan.

2.2 Information Policies

As shown in Table 1, Prosper has continually changed the information that it provides lenders. At the beginning of our sample (June 2006), the credit information posted on Prosper includes debt-to-income ratio, credit grade, whether the borrower owns a home and some credit history information about delinquencies, credit lines, public records, and credit inquires. Throughout our sample time, credit grades are reported in categories, where grade AA is defined as 760 or above, A as 720-759, B as 680-719, C as 640-679, D as 600-639, E as 540-599, HR as less than 540, and NC if no credit score is available.[14] The numerical credit score is never available to lenders. On February 12, 2007, Prosper posted more detailed credit information plus self reported income, employment and occupation.[15] Additionally, Prosper tightened the definition of credit grade E from 540-599 to 560-599 and grade HR from less than 540 to 520-559 eliminating borrowers who do not have a credit score (NC) or have a score below 520 from borrowing on the site. On October 30, 2007, Prosper began to display a Prosper-estimated rate of return on the bidding page (bidder guidance). Before the change, a lender had to visit a separate page to look for the historical performance of similar loans.[16] As discussed in Section 4, these policy changes are likely to impact lender selection of loan risks on Prosper.

In addition to providing hard information in the form of credit history, Prosper also facilitates the use of social networking through groups and friends. A non-borrowing individual may set up a group on Prosper and become a group leader. The group leader is responsible for setting up the group web page, recruiting new borrowers into the group, coaching the borrower members to construct a Prosper listing, and monitoring the performance of the listings and loans within the group. The group leader does not have any legal responsibility. Rather, the group leader is supposed to foster a "community" environment within the group so that the group members feel social pressure to pay the loan on time. Group leaders can also provide an "endorsement" on a member's listing and bids by group leaders and group members are highlighted on the listing page. Since October 19, 2006, Prosper has posted star ratings (one to five) in order to measure how well groups perform against expected (Experian historical) default rates.[17]

Prosper groups were initiated as a tool to expand the market, and thus Prosper initially rewarded a group leader roughly $12 when a group member had a loan funded (Mendelson

[14] Prosper has refined credit grade definition after its registration with the SEC in July 2009.

[15] On this date, lenders were also allowed to begin asking borrowers questions and the borrowers had the option to post the Q&A on the listing page.

[16] Prosper also introduced portfolio plans on October 30, 2007, which allow lenders to specify a criterion regarding what types of listings they would like to fund and Prosper will place their bids automatically. These portfolio plans simplified the previously existing standing orders.

[17] Groups must have at least 15 loan cycles billed before they are rated, otherwise they are "not yet rated."

2006). Given the fact that borrowing is immediate but payment does not occur until at least one month later, the group leader reward may have created a perverse incentive to recruit borrowers without careful screening of credit risk. To the extent that the group leader knows the borrower in other contexts (e.g. colleagues, college alumni, military affiliation), she could collect credit-related information via emails, interviews, house visits, employment checks, and other labor-intensive means.[18] However, when a group gets very large (some with over 10,000 members), it becomes difficult if not impossible to closely monitor each loan. The imbalance between member recruiting and performance monitoring prompted Prosper to discontinue the group leader reward on September 12, 2007.

Starting February 12, 2007, Prosper members could begin to invite their offline friends to join the website. The inviting friend receives a reward when the new member funds ($25) or borrows her first loan ($50). Existing Prosper members can become friends as well if they know each other's email address but the monetary reward does not apply. Friends can also provide endorsements on each other's listings and a bid by a friend is highlighted on the listing page. Beginning February 23, 2008 lenders could begin including aspects such as friend endorsements and bids from friends as criteria in their listing searches.

2.3 Offline Competitors

The main competition that Prosper face in the traditional market is credit card debt and unsecured personal loans.[19] Up to our observation date (August 1, 2008), 36% of all previous Prosper listings have mentioned credit card consolidation, which is higher than the mention of business (23%), mortgage (14%), education (21%), and family purposes (18%) such as weddings.[20] As of February 2008, Federal Reserve reports that the average interest rate is 13.71% for credit card accounts[21] and 11.40% for bank-issued unsecured personal loans. Most Prosper loans carry an interest rate much higher than these two numbers, but since we do not know the credit grade composition of credit card accounts, the comparison is not conditional on the same observable attributes. We will revisit this issue in Section 4.

Roughly 6% of Prosper listings mentions that the Prosper loan, if funded, will be used to pay off payday loans in the offline market. Compared to the APR of 528% that Caskey(2005)

[18] Group leaders do not have access to the borrower's credit report prior to listing.

[19] According to Federal Reserve G.19 Statistical Release as of April 7, 2008, the total consumer outstanding (excluding mortgages) was valued at $2.54 trillion in February 2008. Within this category, $0.95 trillion was revolving debts primarily borrowed in the form of credit cards. The rest ($1.58 trillion) were non-revolving debts including loans for cars, mobile homes, education, boats, trailers, vacations, etc.

[20] 69% of listings mention cars, but this is likely a result of borrowers listing their car payments as a monthly expense.

[21] Conditional on the accounts that have assessed interest.

reports for payday loans, one may argue Prosper could provide a much better alternative to payday loans, given the 3-year duration of Prosper loans and the interest rate cap no higher than 36%. However, lenders must consider the credit risk they face on Prosper. If a payday lender must charge an annual interest rate of 500% to survive competition (Skiba and Tobachman 2007), it is unclear why Prosper lenders would be willing to support this pool of borrowers with a much lower interest rate.

2.4 Macro Environment

As shown in Figure 1, consumer lending has undergone dramatic changes during our sample period, ranging from a calm market with stable monetary policies before August 2007 to the outbreak of the subprime mortgage crisis on August 9, 2007 and gradual spillovers to other types of lending and investment. In light of this, our analysis controls for a number of daily macroeconomic variables, including the bank prime rate,[22] the TED spread,[23] the yield difference between corporate bonds rated AAA and BAA, and S&P 500 closing quotes. According to Greenlaw et al. (2008), the middle two are the strongest indicators of the subprime mortgage crisis. Additionally, we include the unemployment rate reported by the Bureau of Labor Statistics (BLS) by state and month, the housing price index reported by the Office of Federal Housing and Enterprise Oversight (OFHEO) by state and quarter, and the quarterly percentage of senior loan officers that have eased or tightened credit standards for consumer loans, and the foreclosure rate reported by Realtytrac.com by state and month.

We also control for a number of daily Prosper-specific market characteristics, including the total value of active loan requests by credit grade, the total dollar amount of submitted bids by credit grade, and the percentage of funded loans that have ever been late by credit grade. The first two variables intend to capture the overall traffic on Prosper, which may vary by media coverage, word of mouth, or the mood of borrowers and lenders. The percent ever late intends to capture the ex-post performance of the Prosper market as a whole, so as to track the performance evolution that lenders may observe on Prosper over time. Because the financial turmoil observed in the macro environment is rooted in the subprime mortgage crisis, we control for the interaction of the OFHEO foreclosure rate and the borrower's home owner status and consumer loan easing and tightening with whether the borrower has a credit grade of E or HR. It is worth noting that most of the time-series variables, except for those specific to date, state or credit grade, will be absorbed in year-week fixed effects. Whenever possible, we estimate specifications with and without these fixed effects for robustness.

[22]Bank prime rate tracks the Fed funds rate with a 0.99 correlation.

[23]Defined as the difference between 3-month LIBOR and 3-month Treasury bills.

3 Data

In addition to macroeconomic indicators described above, our study utilizes data publicly available for download from Prosper's website and a private data set provided to us by Prosper. The main data set is downloaded on August 1, 2008. It includes all of the information available to borrowers and lenders on the website since Prosper's inception. For each listing it contains the credit variables extracted from Experian credit reports, the description and image information that the borrower posts, and a list of auction parameters chosen by the borrower. For those listings that become loans, we observe the full payment history up to the download date. For each Prosper member we observe their group affiliation and network of friends.[24] Finally, data on all Prosper bids allow us to construct each lender's portfolio on any given day.

We also utilize a private data set obtained from Prosper that includes the number of listings, number of loans, average contact interest rate, percent late at 6 months, and percent late at 12 months by state, month and credit score interval. These credit score intervals are finer than the publicly posted credit grades. For comparison, it also includes Experian data on historical loan performance in these finer credit intervals for offline consumer loans.

Our sample includes listings that began on or after June 1, 2006 and end on or before July 31, 2008 and the loans that originate from this set of listings. We exclude the few loans that were suspects of identity theft and as a result repurchased by Prosper. Table 2 summarizes listings and loans by quarter for this sample. This sample includes 293,808 listings and 25,008 loans for \$158.27 million. This implies an average funding rate of 8.51%, though this has varied over time ranging from 6.32% to 10.14%. Average listing size and average loan size both increased through the first half of 2007 and decreased since. Comparing listings and loans, the average listing requests \$7,592 and the average loan is worth \$6,329. This difference is preliminary evidence of credit rationing. It appears that lenders are wary of listings requesting larger loans and view this as a signal of higher risk. The average listing lists a maximum borrower rate of 19.19% while the average contract rate is 17.90%.[25]

In terms of social networks, Table 2 suggests that being a group member and having a friend endorsement increases the likelihood of funding. Both variables have larger representation in loans than in listings. However, it is striking that the proportion of listings and loans with group affiliation has decreased drastically from 60% and 70% to 7.5% and 10%, respectively. When

[24]The data dump reflects information about groups and friends as of the download date. Because these characteristics can change over time, we use monthly downloads beginning in January 2007 to identify these characteristics at the closest possible date to the actual listing.

[25]The sharp increase in borrower maximum rates between the first and second quarters of 2008 reflects the April 2008 removal of state specific interest rate caps.

friend and group leader endorsements became available, the percent of listings and loans with endorsements initially grew but have decreased since the middle of 2007. The only exception is the percent with friend endorsements plus bids. These patterns call into question the importance and effectiveness of social networks which we will explore in detail in Section 5.

Table 2 also summarizes the percent of default as observed on August 1, 2008. We define a loan as in "default" if it is four or more months late or labeled default by Prosper due to bankruptcy.[26] About 30% of loans originating in 2006 have defaulted by August 1, 2008, which makes it clear that lenders observe negative performance in their portfolios and the overall market. Note that the proportion of loans in the 2007 and 2008 cohorts that have defaulted are much lower than in the earlier cohorts as a result of the life cycle of loans.

4 Hard Information of Credit Grade

To a traditional loan officer, one of the most important pieces of hard information is credit score. In contrast, Prosper lenders observe categorical credit grades instead of the exact score. This has two implications: first, Prosper lenders can differentiate borrowers across credit grades. Second, two borrowers with same grade but different scores will look identical to Prosper lenders if the other observable information is the same. According to Akerlof (1970), this will trigger adverse selection within each grade. This section presents evidence for both implications.

4.1 Differentiation Across Grades

Table 3 presents the funding rate, interest rate, the percent late, and the percent 3-months late or worse (as of August 1, 2008) by the 8 credit grades observable to Prosper lenders. As expected, a better grade means a higher funding rate, lower interest rate, and better loan performance. The last two columns attempt to compare Prosper loan performance to all the Experian accounts that have a new credit line approved in September 2003. Since the performance of Experian accounts are observed as of September 2005, we summarize the observed 2-year performance for the Prosper loans that were originated in July 2006. While the time horizon of Prosper and Experian loans are not exactly the same, it is clear that Prosper loans perform much worse than the traditional Experian accounts.[27] One potential explanation is that Pros-

[26]Once a loan is four months late, Prosper considers it eligible for debt sale, and once it is sold it is considered to be in default. However, because debt buyers only purchase packages of loans, a four month late loan will not be considered "default" immediately. Loans can thus be labeled "4+ months late" for long periods of time. Our default definition overcomes this mechanical ambiguity.

[27]According to the Federal Reserve, the credit card charge-off rate has increased from 4.3% in the third quarter of 2005 to 5.5% in the second quarter of 2008. If we had grade-specific performance data in 2008, the comparison between traditional and Prosper loans would be less stark.

per loan composition is worse than the Experian accounts *within* each credit grade because Prosper attracts more borrowers towards the lower end of the grade.

The ordinal difference across grades remains salient after we control for other listing attributes. In particular, we run three descriptive regressions that correlate observable listing attributes to the probability of being funded (1_{funded}), the interest rate if funded ($InterestRate$), and whether the loan is default or late as of August 1, 2008 ($1_{defaultorlate}$). If a certain listing attribute (say credit grade) is a well-understood indicator of credit risk, we should see a greater funding probability, a lower interest rate, and better ex-post performance for listings of better risk. The regressions equations are as follows:

$$1_{funded,i} = f_1(ListingAttributes_i, macro, FE_{yw}) + e_{1it} \tag{1}$$

$$InterestRate_i = f_2(ListingAttributes, macro, FE_{yw}) + e_{2it} \tag{2}$$

$$1_{defaultorlate,it} = f_3(ListingAttributes, macroFE_{yw}, FE_a) + e_{3it} \tag{3}$$

In all three regressions, we include year-week fixed effects (FE_{yw}) to control for the changing environment on and off Prosper. In the third equation, we also control for a full set of monthly loan age dummies (FE_a) to control for the life cycle of loan performance. As summarized in Table 4, the listing attributes we use in these regressions include Experian-verified credit history information, borrower-specified loan terms (e.g. amount request and maximum interest rate), borrower self-reported information (e.g. loan purpose, image, description) and social network variables (e.g. where the borrower belongs to a Prosper group, whether the listing is endorsed and/or bid by group leaders and friends). The funding rate and performance regressions are estimated by probits and the interest rate regression is estimated by OLS.

According to the regression results in Table 5, the probability of being default or late increases by credit grade, and in response, interest rate increases and the funding probability decreases. This suggests that credit grade is an important measure of borrower risk and lenders understand their ordinal differences. Similarly, lenders understand that the more a borrower requests to borrow, the higher the risk of mis-performance, and therefore deserves a lower funding probability and a higher interest rate.[28] Lenders also foresee the higher risk of autofunded loans and adjust funding rate and interest rate accordingly.

In contrast, the consistency between funding rate, interest rate and loan performance fail to hold for some of the self-reported attributes. For instance, having an image increases the funding rate and conditional on being funded, a listing with an image on average enjoys a lower interest rate (0.1 percentage points) than a seemingly equivalent listing without an image. However, the two listings do not show any significant difference in loan performance. The lack of performance

[28]Loan size is a typical method of credit rationing.

difference is not surprising, given the fact that Prosper does not verify the authenticity or content of the image. What is surprising is the systematically higher funding rate and lower interest rate for listings with an image, which combined with the performance data implies that the expected rate of return will be lower for listings with an image. Similar inconsistency appears in some variables that indicate the length and content of listing description. These regressions also include social network variables, which we will discuss in Section 5.

4.2 Adverse Selection within Grade

To detect adverse selection within a grade, we use Prosper-provided confidential data to summarize listings and loans by "half grade." Except for the two ends of the score distribution, half grades are defined as a 20-point interval of credit scores, for instance, 600-619 (referred to as D-) and 620-639 (D+). In total, we have 20 half grades, which is much more detailed than the 8 credit grades posted on Prosper.com during our sample period.[29]

Figure 2 compares the c.d.f. of Prosper listings, Prosper loans, the Experian population, and Experian new accounts across the 20 half grades. By Experian population, we mean all the accounts that have a score by the Experian ScorexPLUS model in December 2003.[30] Note that a person may have a record in Experian without demanding credit. In this sense, Experian population is an imperfect comparison for Prosper listings. The Experian new accounts are defined as above, where the credit could be secured (such as a mortgage) or unsecured (such as a credit card). Even though the Prosper vs. Experian comparison is imperfect,[31] there is no doubt that Prosper listings have much greater concentration on lower credit intervals. Prosper lenders are able to select better risks from the listing pool, but the overall distribution of Prosper loans is still worse than that of Experian accounts.

Figures 3 and 4 present the p.d.f. of Prosper listings and Prosper loans by the 20 half grades and across time. The loan distribution is also compared with the p.d.f. of Experian new accounts as defined above. Not surprisingly, Prosper attracts listings towards the lowest end of the credit score distribution (Figure 3) while the traditional lenders tend to focus on the highest end (Figure 4). These facts are probably linked: because traditional lenders cannot satisfy the credit demand of near or subprime risk (e.g. due to credit rationing), these risks find

[29] The precise definition of the 20 half grades are 300-479, 480-499, 500-519, 520-539 (HR-), 540-559 (HR+), 560-579 (E-), 580-599 (E+), 600-619 (D-), 620-639 (D+), 640-659 (C-), 660-679 (C+), 680-699 (B-), 700-719 (B+), 720-739 (A-), 740-759 (A+), 760-779 (AA-), 780-799 (AA+), 800-819, 820-839, 840-900.

[30] "Redeveloped Experian/Fair, Issac Risk Model" (December 2003) accessed at www.chasecredit.com/news/expficov2.pdf on September 5, 2008.

[31] Given the stability of credit markets before the subprime crisis and the credit crunch after August 2007, the Experian distribution is likely to overestimate the traditional credit access in 2006-2008 and therefore constitutes a conservative comparison group against Prosper.

Prosper an attractive alternative. More interestingly, the Prosper loan distribution is much less smooth than the Experian new accounts. From Figure 4 we see a higher frequency at D- than D+, C- than C+, etc. in the Prosper loans, but not in the Experian accounts. Given the fact that the Prosper listing distribution leans toward the very low tail of credit scores, lower listing frequency at the lower end of each grade could be due to adverse selection or the skewness itself. We will examine this alternative explanation below. Also note that the jumpiness of Prosper loans does not disappear over time. The listing and loan distributions are both moving towards the right, which could be due to the credit crunch forcing near prime and prime risks to seek credit on Prosper, Prosper revealing more information hence discouraging subprime risks, or Prosper lenders learning to avoid subprime risks.

To further explore the systematic difference between minus and plus grades, we examine the population of Prosper listings and loans by half-grade, census division and month. For a half-grade interval i in census division[32] c and month t, we estimate the following two specifications:

$$\begin{aligned} Y_{ict} =& 1_{minusgrade} \cdot \beta_1 + MonthCount_t \cdot 1_{minusgrade} \cdot \beta_2 \\ &+ MonthCount_t \cdot 1_{nearprime} \cdot \beta_3 + MonthCount_t \cdot 1_{subprime} \cdot \beta_4 \\ &+ \mu_c + \mu_t + \mu_{grade} + \epsilon_{ict} \end{aligned} \tag{4}$$

$$\begin{aligned} Y_{ict} =& 1_{minusgrade} \cdot \beta_1 + Macro_{ct} \cdot 1_{minusgrade} \cdot \beta_{2m} \\ &+ Macro_{ct} \cdot 1_{nearprime} \cdot \beta_{3m} + Macro_{ct} \cdot 1_{subprime} \cdot \beta_{4m} \\ &+ ProsperPolicy_t \cdot 1_{minusgrade} \cdot \beta_{2p} + ProsperPolicy_t \cdot 1_{nearprime} \cdot \beta_{3p} \\ &+ ProsperPolicy_t \cdot 1_{subprime} \cdot \beta_{4p} + \mu_c + \mu_t + \mu_{grade} + \epsilon_{ict} \end{aligned} \tag{5}$$

In both specifications, we adopt six dependent variables: (1) the number of Prosper listings, (2) the number of Prosper loans, (3) the funding rate,[33] (4) the average interest rate of loans, (5) the percent late after 6 months, and (6) the percent late after 12 months. In principle, all six outcomes can be driven by borrower behavior, lender behavior, or both. Both specifications control for a full set of year-month dummies, a set of credit grade dummies (i.e. one dummy for AA, one for A, etc.), and a set of dummies for census division. The coefficient on the dummy of minus grade tells us how minus grades differs from plus grades *within* the same grade. Standard errors are clustered by census division.

With no explicit control for Prosper policy changes or the macro environments, Specification (4) intends to describe how Prosper populations have changed across credit intervals and over time. In Specification (5), we relate the over-time population change with various macro variables

[32] We have state level data but some states have too few observations in the count of listings or loans. Aggregation into census division alleviates this problem. We have also tried aggregation into census regions, results are similar.

[33] Which is literally the number of loans divided by the number of listings in each cell.

and the major Prosper policies. Since most macro and Prosper policy variables are simple time-series, it is difficult to tell them apart from the general time trend. As a result, we interact the macro/policy variables with whether the credit interval refers to a minus grade, whether the interval falls into the near prime range defined by Experian (600-679), or whether the interval belongs to the subprime range (below 600). These interactions capture the co-movement of the Prosper population and the overall environment, but do not necessarily represent causal effects.

Table 6-1 presents the regression results for Specification (4), with two columns for each dependent variable. Because the February 2007 Prosper policy disallowed any listing with credit score below 520, to facilitate comparison the regression sample excludes credit scores below 520.[34] The odd numbered columns suggest significant adverse selection: compared to plus grades, minus grades have on average 11 more listings and 2 more loans per division-grade-month. Both numbers imply a significant density towards minus grades as there are only 30 listings and 6 loans in each cell on average. As expected, the minus grade loans perform significantly worse. The fact that Prosper lenders do not observe credit scores explains why the funding rate is no different between minus and plus grades after we control for the fixed effects of month, grade and division. However conditional on funding, lenders do charge 0.4 percentage point higher interest rates on the minus grades, which suggests that they may have some clue as to which loans are minus grades and which are not. The even numbered columns include the interactions of month count (since June 2006) with minus grade, near prime, and subprime. Over time we observe more near and subprime listings relative to prime, but less subprime loans.[35] While the overall risk composition has improved, the adverse selection towards minus grades increases over time at a speed of 0.55 more minus grade listings and 0.18 more minus grade loans per month.

Table 6-2 replaces the month count interactions with those that involve macro variables and Prosper policy changes. For both the number of loans and listings, the concentration towards minus grades has increased after February 2007 and April 2008. Interestingly, Prosper lenders do not demand a higher interest rate on a minus grade loan until February 2007. We suspect this occurs because the extra credit information that Prosper provided since February 2007 has helped lenders distinguish risks within a grade. One may argue that Prosper introduced friend endorsement in February 2007 as well and that could contribute to lenders charging higher interest rates. We cannot test this explicitly because the half-grade data is aggregated. But as shown in Table 2, friend endorsements account for only 20-30% of the Prosper population and

[34] Results using all the "half-grade" intervals are very similar to the presented results except for the chop-off of the scores below 520 after February 2007.

[35] This is slightly different from what is seen in Figures 3 and 4, because these regressions describe the absolute number of listings and loans in each interval as opposed to the relative distribution.

these percentages are declining sharply over time. These facts do not explain why the increased interest rate for the minus grade loans appears abruptly in February 2007 and stays stable afterwards. Turning to the interaction of Prosper policies and subprime, the results suggest that Prosper policies, especially the bidder guidance introduced in October 2007, may have helped lenders better understand the true meaning of a subprime grade, and therefore motivate them to shift towards better grades. This result will be further examined in Section 7.

To check whether the data is truly driven by adverse selection towards minus grades, we conduct several robustness checks. First of all, we run Specification (4) on the counts of Experian new accounts by half grade[36] and find that the coefficient for the minus grade dummy is statistically zero (with t=0.25). Second, we control for the mid-point of each credit interval, its square, and its interaction with the count of months from June 2006 to the sampled month. These controls intend to capture a continuous skewness of the listing/loan distribution by credit score. With these controls, the coefficient on the minus-grade dummy remains positive and significant. Lastly, we perform a placebo test by shifting up the definition of credit grades (and the corresponding minus and plus grades) by 20 credit score points. If the minus-grade phenomenon is primarily driven by a distribution frequency that continually declines from low scores to high scores, the density of minus grades should remain higher than that of plus grades under the new half grade definition.

This alternative explanation has some support in the number of listings but not in the number of loans, funding rate, or contract rate. Focusing on the number of listings, Table 6-3 column 1 shows that the coefficient for the minus grade dummy is 5.775 under the placebo definition, which is statistically significant but much lower than what we find in the true data (11.381). This suggests that the higher density of minus grades may be partially due to the declining density in Prosper listings, and partially due to adverse selection. The same argument does not apply to Prosper loans. Conditional on loans, column 2 shows that the coefficient of the minus grade dummy has changed from 1.908 in the true data to -0.752 in the placebo test. Consistently, the minus grade dummy under the placebo test has a negative coefficient in the funding rate regression and much higher positive coefficient in the contract rate regression (as compared to zero and small positive in the true data). These changes are not surprising: under the placebo definition, lenders observe the difference between D- and D+ because these grades correspond to D+ and E- in the real data. Overall, Table 6-3 suggests that distribution skewness cannot explain all of the observed differences between minus and plus grades.

To summarize, we find that Prosper lenders understand the ordinal differences across credit grades but the crude definition of credit grade has resulted in adverse selection towards the lower

[36]But not by census division or time because we only have one cross-section of Experian new accounts

end of each grade. While the macro environment and the Prosper policies may have contributed to the shift towards better grades, the adverse selection towards minus grades has increased over time. The market is aware of the problem, as the adverse selection is partially offset by higher interest rates when Prosper began posting more credit information in February 2007.

5 Soft Information via Social Networks

Given the lack of legal responsibility in Prosper groups and friends, the effect of social networks is not straightforward. On the one hand, social networks may help alleviate the information asymmetry between borrowers and lenders because friends and group leaders may certify a borrower's history of on-time payment, some social connections (say education and employment) may be a good signal of the borrower's payment ability, and the concern of losing face in front of friends and group members may motivate a borrower to pay back. On the other hand, some social networks may fail to solve or even aggravate the information problem. For example, a friend that endorses and bids may intend for charity lending. Alternatively, a group leader may endorse a risky borrower because she wants to earn the group leader award and therefore does not perform adequate risk screening. A borrower can also bribe a friend to endorse and bid on her listing. Bearing all these possibilities in mind, we aim to evaluate the overall effectiveness of social networks on Prosper, while leaving the specific mechanism to a companion paper (Freedman and Jin 2008).

Table 7 presents a detailed summary of social variables. As we saw in Table 2, 28.8% of listings have some group affiliation, 19.1% have friends, 3.2% have an endorsement from a group leader (2.2% with a leader bid), and 3.2% receives friend endorsement (1% with a friend bid). All of these fractions increase substantially in the loan sample, indicating that social loans are more likely to be funded than the listings that have no group affiliation or friend ties. The later part of Table 7 shows a large dispersion in group size, group type (e.g. groups by alumni, military, etc.), group rating and whether a group requires borrower review (by the group leader) before listing. Table 7 also shows that, while social ties contribute positively to the final funding of social loans, most funding comes from stranger lenders. Due to space limit, the analysis presented in this paper does not discuss group size and group type in detail. These variables are extensively examined in Freedman and Jin (2008).

Table 8 reports what types of borrowers are more likely to have social ties. Given the importance of credit grade in summarizing borrower risk, we present the grade composition conditional on group borrowers, endorsed borrowers, and the whole Prosper market. At the listing level, it is clear that group borrowers are more likely to have lower credit grades, especially E and HR. This observation does not hold at the loan level, though group affiliation does help funding within each credit grade at or below D. For the endorsed loans, we focus on the borrower population after February 12, 2007 because friends are not allowed to make endorsement until this time. Unlike group listings, the grade composition of endorsed listings is very similar to that of the whole market. Judging from the grade composition, friend endorsement seems to help funding in grades at or below D, a phenomenon similar to that of group loans.

We now discuss the coefficients of social network variables in Specifications (1), (2), and (3) from above.[37] As shown in Table 5, we observe both consistency and inconsistency for the social network variables. Compared with others, borrowers that belong to a group are 0.4 percentage points more likely to get funded, enjoy a 0.4 percentage point lower interest rate, but are 0.5 percentage points more likely (though statistically insignificant) to be default or late. Similar inconsistency occurs for a group loan that receives an endorsement and bid from the group leader. These results imply that group loans, especially those that receive an endorsement and bid from the group leader, may generate lower returns than the non-group loans. How much lower the rate of return is and why lenders are willing to support lower-return loans will be examined below. In comparison, having a friend endorsing and bidding on the listing shows more consistency: it has a large effect on the funding rate (9.6 percentage points), and conditional on being funded, the interest rate is 0.7 percentage points lower and the probability of default or late is 4.1 percentage points lower. Whether the favorable interest rate has over- or under-compensated the better loan performance is an empirical question discussed below.

An interesting question is whether social network variables work differently for different credit grades. It is possible that lenders lay more emphasis on social network variables for lower credit grades because it is more difficult to differentiate borrower quality in these grades. To check this, we rerun Specifications (1), (2) and (3) with social network variables specific to A-AA, B-D and E-HR. As shown in Table 9-1, results suggest that friend endorsement plus bid, the only social network variable that seems to capture better loan performance in Table 5, is more effective for A-AA than for B-D and not effective at all for E-HR in predicting loan performance. In contrast, funding rate and interest rate are more favorable for listings with social networks in B-D and E-HR than for A-AA. This suggests that either lenders do not realize the true meaning of social networks for low grades, or they intentionally extend charity lending to low grade social loans.

Another concern is that some social variables may reflect borrower gaming instead of real

[37]These coefficients do not necessarily represent the causal effect of social network variables on funding and repayment probabilities, given the likelihood of omitted variables that are correlated with social network characteristics. If there are such unobservable characteristics, these coefficients capture the signaling effect of social networks in the sense that they represent the average differences in outcomes for borrowers by social network status including all of the omitted variables.

networks. For example, two stranger borrowers may agree to endorse each other or even bid on each other's listings with no money exchange. If lenders do not pay attention to such mutual endorsement, this could increase the funding rate (or the lower interest rate) for both listings. To address this concern, we generate a dummy equal to one if a listing involves mutual endorsement without bid and another dummy for mutual endorsement with bid. On average, 11.46% of listings and 16% of loans with a friend endorsement are involved in a mutual endorsement and 6% of listings and loans with an endorsement and bid are involved in a mutual endorsement and bid. For both mutual endorsements with and without bids, the median number of days between the two endorsements is around 30 days.

We then add these two dummies to Specifications (1), (2), (3). As shown in Table 9-2, results suggest that mutual endorsement without bids increases the funding probability by 0.4% but has little effect on loan performance. This is consistent with the gaming hypothesis. In contrast, mutual endorsement with bids seems to reduce the probability of default or late by 5.9% but has little impact on funding rate and interest rate. This suggests that mutual endorsements with bids may reflect the real network effects instead of gaming.

6 Expected Internal Rate of Return

To better understand lender perception of borrower risk, we follow two principles to compute the internal rate of return (IRR) that a sophisticated lender should expect from a Prosper loan: the lender considers all of the information at the time of listing, and he projects the risk of late payments and default throughout the 36-month loan life. To obtain IRR, we first calculate the annual discount factor (call it R) that equalizes the loan amount to the present value of all the predicted monthly payments, and then compound R monthly so that IRR reflects the annual percentage yield from the loan (i.e. $IRR = (1 + R/12)^{12} - 1$). We believe this method reflects the rate of return that a lender *expects* to earn at the start of the loan if he can perfectly predict the statistical distribution of loan performance. The step-by-step algorithm is described in the Appendix. The Appendix also explains why our method is more comprehensive than the ones used by Prosper and LendingStats.[38]

We use three dummies to measure loan performance: default, default or late, and missed payment. Located between the most optimistic (default) and the most pessimistic (default or late), the dummy of missed payment is defined as one if the loan's payment history indicates that the borrower has missed the payment in a specific month. If the borrower misses the payment at month t but makes it up in a later month, we count it as not missing the payment. Early paid off is treated as a bulk of cash flow in the actual month of payment and zero afterwards.

[38] LendingStats is a popular independent website that tracks Prosper activities in real time.

This implicitly assumes that the early payoff is reinvested into a loan that is identical to the loan under study.[39]

The first step of our algorithm is predicting the probability of default, missed payment, default or late, and early payoff for each loan and each month using the full history from June 1, 2006 to July 31, 2008. Since Prosper added new credit information in February 2007, our full-sample prediction only uses the variables that are always available.[40] Note that our prediction regressions do *not* include year or month dummies. This way, the regressions only capture the average relationship between loan performance and observable loan attributes throughout the whole sample period. If two loans were originated at different dates but have the same observable attributes, they always have the same predicted performance. In other words, any difference in the predicted loan performance is due to different selection of observable loan attributes, not when the loan was originated.

Because all Prosper loans are three years, the majority of them are still ongoing at our observation date unless they were paid early or have already defaulted. As a result, we do not have information for loan performance in months 25-36 if we use the full sample to predict risk. Instead of using arbitrary roll over rates, we report two sets of IRR estimates: one assumes that the cumulative mis-performance remains constant after month 24 (referred to as "flat IRR"), and the other assumes that the mis-performance rate will follow a linear projection after month 24 ("linear IRR").[41] Our prediction regressions suggest that mis-performance does not have a statistically significant increase after month 18. While this concavity may be driven by fewer observations towards the later loan life, the raw Prosper performance depicted in Figure 5 does confirm that mis-performance is more likely to occur in the earlier part of the loan life, as typically observed in the industry. Based on this observation, we believe the truth is somewhere between the flat and linear IRRs. More details of the potential bias in our IRR calculation are discussed in the Appendix.

Table 10 summarizes six versions of IRR estimates depending on which performance measure

[39]The main drawback of IRR is assuming that early positive cashflows will be reinvested into projects that are identical to the project under study. This could lead to an over- or under-estimate for the return of investment in a given period. One way to overcome this problem is assuming the reinvestment rate equal to a specific average cost of capital. Here we do not use modified IRR, partly because any choice of the reinvestment rate is arbitrary, partly because this study is not meant to be an investment guide. We emphasize performance comparison across loans, not whether Prosper loans financially worth investing in a fixed time period.

[40]As a robustness check, we rerun the prediction regression for the post-Feb-2007 sample only. For this sample, including or excluding the new credit information makes very little difference (less than 0.2%) in the final IRR estimates. From this we conclude that not using the new credit information does not bias our IRR estimates.

[41]More specifically, linear projection for the full sample means that the predicted mis-performance rate at month x (where $x \geq 25$) is equal to [predicted risk at month 24 + (x-24)*(predicted risk at month 24-predicted risk at month 23)].

we use and whether we assume a flat or linear projection in the unobserved loan life. In theory, the present value formula is monotone and should have a unique solution of R that is bounded between -1200% and the contract rate. After monthly compounding, IRR is bounded between -100% and $(1 + \text{contract rate}/12)^{12} - 1$. In practice, we do achieve over 99% of convergence if we do not impose any constraint on IRR. However, since we predict the likelihood of early and regular on-time payment separately,[42] there is a small chance (less than 10%) that the sum of the estimated likelihood is over one in at least one of the 36 months, hence the converged IRR can exceed the theoretical bounds. To address this problem, the IRRs reported in Table 10 are estimated with the imposed constraint that they cannot lie outside their theoretical bounds. All the results reported below are robust if we focus on the loans whose unconstrained IRR does not exceed the bounds.

Our algorithm calculates IRR per loan but the loan-to-lender match allows us to compute average IRR for each lender portfolio. Weighting each loan by dollar leads to average IRR by dollar. At all three levels – loan, lender or dollar – the comparison of IRR1 to IRR6 (conditional on converged IRR) is consistent with expectation. Flat IRRs are 0.5-1 percentage points higher than linear IRRs because the latter is more pessimistic about the unobserved loan life. Within flat IRRs, the average return based on missed payments (IRR2) is bounded between those based on default (IRR1) and default or late (IRR3). Average IRR2 does not differ much across loan (-0.48%), lender (-0.17%) and dollar (-0.62%) levels, suggesting that big lenders and big loans do not earn significantly higher returns than small lenders or small loans.

For comparison, the average annual yields of 3-year Treasury Bill and S&P 500 are 3.97% and -0.66% in the same period (June 1, 2006 to July 31, 2008). One explanation for the seemingly low average IRR of Prosper loans is that lenders expect to have positive returns but the unexpected macro shock after the subprime crisis led to negative returns. To address this concern, Figure 6 plots real, predicted and unpredicted risk of default or late by the calendar month of payment. If the sub-prime crisis and its aftermath are the main reason driving bad performance, the unpredicted risk should be systematically higher after than before August 2007. Figure 6 does not support this prediction.

In all six versions of IRR estimates, we find large heterogeneity across loans, ranging from -99.9% to +39.8%. The lender-level IRRs are less dispersed than loan-level IRRs because most lenders diversify investment in more than one loan. To describe how the IRRs differ across different types of loans, the tables and figures discussed below utilize IRR2, which is computed based on the risk of missed payment and assumes a flat projection into the unobserved loan life. Using the other IRR estimates generates very similar comparison.

[42] The joint estimation takes an extremely long time and does not yield stable results.

Figure 7 presents the kernel density of loan-specific IRR2 by credit grade. On average, grades AA-A have the highest average rate of return (5.25%) as compared to the other categories (1.53% B-D and -10.52% E-HR). AA-A also has a tighter distribution and less variability than B-D and E-HR. As one would expect, E-HR has the longest left tail and the lowest return on average. The negative IRRs suggest that some lenders are not experienced enough to foresee a negative return, or they have specific incentives to fund lower-quality loans. As shown in Section 7, both explanations hold to some extent but the first is more dominant.

The second IRR heterogeneity we explore is testing the key assumption underlying the credit rationing theory of Stiglitz and Weiss (1981): because lenders have imperfect information about borrower risk (due to either adverse selection or moral hazard), the willingness to pay for a higher interest rate may signal higher risk, which in turn drives a non-monotonic relationship between rate of return and interest rate. To test this assumption, Figure 8 plots IRR2 against the contract rate.[43] The curve suggests that IRR2 first increases and then decreases after the interest rate reaches 7-8%. The final uptick suggests that an increase of interest rate does not fully compensate the increased risk until 30-31%. According to the 5 and 95 percentiles shown in the same figure, there is more noise at the two ends due to smaller numbers of observations.

Turning to social loans, we compare the kernel density of IRR2 by group status in Figure 9 and endorsement status in Figure 10. As shown in Figure 9, group loans perform worse than the non-group loans. This result is against the intuition that group members may have better "soft" information to signal a "good type" all else equal. In contrast, friend endorsements show a different pattern. On average, loans that have friend endorsements and friend bids perform better than the loans without friend endorsements. However, loans with friend endorsements only perform worse. Combined with the descriptive regressions reported in Table 4, this suggests that friend endorsements may not provide any positive signal about borrower risk until the endorsing friend is willing to certify the "soft" information with a bid of their own.

Figure 11 plots the average IRR2, contract rate, and the predicted risk of missed payment as a function of loan origination month. It is clear that the estimated IRR2 increases steadily over time, which is attributed to a significant decline in missed payments and a relatively smaller decline in interest rate. These trends are consistent with the fact that the market is moving towards better credit grades over time, partly a result of lender learning as shown below. The two vertical lines drawn on this picture represent Prosper information policies in February 2007 and October 2007. While we cannot identify whether the policy changes have a causal effect on the increase of IRR2, they definitely coincide with the trend towards improving returns: the average IRR2s before February 2007, between February and October 2007, and after October

[43] Interest rates are rounded by percentage points in Figure 8.

2007 are -4.29%, -1.40%, and 2.48% respectively. As the average IRR2 improves over time, the overall distribution of IRR2 becomes less dispersed, mostly due to less density in really low IRRs (Figure 12).

7 Lender Learning

The estimated IRR distribution presents two puzzles. First, why do some lenders choose loans whose observable characteristics predict significantly poorer performance than a safe alternative such as treasury bills?[44] Second, why do we observe a great deal of heterogeneity in the expected IRR? If lenders aim for the best financial return, competition should equalize the expected return across all funded loans.

In this section we try to explain these two puzzles by examining if and how lenders learn from their past actions on Prosper. In particular, we estimate how lenders' choices change as a function of the performance of their previous Prosper loans, both in terms of hard and soft information. Strictly speaking, lenders may learn from not only their own experience but also market-wide performance. We choose to focus on the former because it is difficult to distangle market-wide performance from other unobservable time series that affect the Prosper market at the same time. In this sense, the evidence documented below describes the extra learning that lenders obtain from their own experience *in addition to* their learning from the market. We will revisit market-wide learning later when we compare different lender cohorts.

Patterns of learning can help us untangle three potential explanations of the first puzzle that lenders choose low return loans. First, lenders may simply lack the expertise to properly evaluate the risk of loans on Prosper. If this is the case, we should see that lenders learn over time to choose higher return loans. Since we calculate IRR based on observables, the learning pattern identified in this paper focuses on lenders learning how observable attributes signal the unobserved trustworthiness. Second, lenders may choose low return loans because of a charity motive or risk-loving preferences. If lenders choose loans out of charity, we predict that they will not learn from past mistakes. They may or may not continue to make low return loans throughout their life on Prosper, depending on their budget for charity, but this decision should be insensitive to the performance of their existing charity loans. The same argument would apply if lenders are risk-loving. Third, lenders may have expected higher returns ex ante, but the unexpected macroeconomic shock associated with the subprime credit crisis may have led to lower ex post returns. This story does not explain lender learning before the macro downturn.

A detailed analysis of lender learning can also help us decipher the second puzzle. We suspect

[44] If lenders are risk averse, they may even discount the returns they receive online by a greater degree. If this is the case, the disparity between returns on Prosper and T-bills may even be greater.

that the heterogeneity of returns is a result of lenders differing in their ability to evaluate the risk of loans on Prosper. To this end, we examine differences in lender learning by lender cohort and by the types of loans they choose in their first month on Prosper. As lenders learn to better interpret the observable information in listings, we expect the less sophisticated lenders to learn more and the heterogeneity of IRRs to decline over time. At the end of this section, we will also address alternative explanations of IRR heterogeneity, including mean reversion, the perverse incentives of group leader rewards and bidder sniping.

7.1 Lender Mistakes vs. Charity

To identify the extent to which lenders learn from their own mistakes, we estimate a series of regressions describing how a lender's choices to fund, amount to fund, and type of loans to fund respond to the performance of their previous loans, with lender and time fixed effects:

$$FundedALoan_{it} = g_1(PortChar_{it-1}, PortLate_{it-1}) + a_{1it} + \mu_{1i} + \gamma_{1t} + \epsilon_{1it} \tag{6}$$

$$AmountFunded_{it} = g_2(PortChar_{it-1}, PortLate_{it-1}) + a_{2it} + \mu_{2i} + \gamma_{2t} + \epsilon_{2it} \tag{7}$$

$$\begin{aligned} PortComp_{it} =& g_3(PortChar_{it-1}, AtoAALate_{it-1}, BtoDLate_{it-1}, EtoHRLate_{it-1}, \\ & NCLate_{it-1}) + a_{3it} + \mu_{3i} + \gamma_{3t} + \epsilon_{3it} \end{aligned} \tag{8}$$

$$\begin{aligned} IRR2_{it} =& g_4(PortChar_{it-1}, AtoAALate_{it-1}, BtoDLate_{it-1}, EtoHRLate_{it-1}, \\ & NCLate_{it-1}) + a_{4it} + \mu_{4i} + \gamma_{4t} + \epsilon_{4it} \end{aligned} \tag{9}$$

These regressions describe a particular behavior of lender i in week t as a function of characteristics and performance of the lender's portfolio up through week $t-1$. Equation 6 is a linear probability model[45] of an indicator that a lender funded at least one loan in a given week. The other three equations only include the sample of lenders who funded at least one loan in week t. In Equation 7, $AmountFunded_{it}$ is the dollar amount invested by an active lender in week t. Equation 8 is estimated for various $PortComp_{it}$ variables which specify the percentage of an active lender's investment in AA to A, B to D, and E to HR loans in week t. Equation 9 captures the average IRR2 of all the loans that lender i started to fund in week t.

$PortChar_{it-1}$ includes lender i's portfolio HHI and portfolio size through the previous week to control for time varying lender characteristics. $PortLate_{it-1}$ reflects the percentage of lender i's portfolio that has ever been late as of the previous week. $AtoAALate_{it-1}$, $BtoDLate_{it-1}$,

[45] Because we will use a large number of fixed effects, we choose a linear probability model over a probit model for this set of regressions

$EtoHRLate_{it-1}$ are the percentage of lender i's portfolio through the previous week that has ever been late in each of the three respective credit grade categories.[46]

All regressions include lender, time and lender age fixed effects: γ_{jt} is a set of year-week fixed effects allowing us to control for changes in the macro environment and the Prosper market.[47] μ_{ji} is a set of lender fixed effects. With these fixed effects the coefficients on the ever late variables are identified by *within* lender changes in portfolio performance and investment decisions. a_{jit} are monthly lender age fixed effects for lender i.[48] These dummies will capture any general pattern in lenders' choices as they age. For example, if lenders have a strategy to invest a certain amount when they first join prosper and less in subsequent months, these age dummies will allow us to capture lenders' responses to the performance of their portfolios, controlling for the general aging pattern. In all of the results presented here, we cluster standard errors at the lender level.

The results of regressions (6)-(9) are reported in Table 11. Lenders show strong responses to poorly performing loans in their portfolios. On average, a ten percentage point increase in the proportion of their portfolio that has ever been late decreases their probability of funding a loan by 0.78 percentage points and decreases the amount they invest in an active week by $79.5. The second column shows that the probability of funding a new loan is sensitive to lateness in all credit grades, and the sensitivity is greater for better grades.

One may argue that there is a mechanical relationship between portfolio performance and new investment because bad past performance implies less money available for new investment. However, this mechanical relationship does not explain why a lender changes his portfolio composition in response to past performance. Columns 5-7 display the coefficients from the different versions of the $PortComp$ regressions. As lenders observe late loans, they tend to decrease their funding of loans in the grade with the adverse shock and increase their funding of higher quality grades.[49] These results indicate strong evidence of learning. The high late and default rates of E and HR loans have driven lenders away from these loans and toward higher credit grades as lenders have learned about the dangers of investing in these lower credit grades. In results not shown here, coefficients from regressions describing the propensity to fund loans in other categories including autofunded loans and loans of various sizes as a function of late loans in these categories show similar patterns.

[46] We have also tried specifications using the percent of a lender's portfolio (in total or in various categories) that is currently late or in default and the results are very similar.

[47] Results of identical regressions with controls for macro variables and Prosper supply, demand, and market performance instead of week fixed effects are very similar.

[48] We count a lender as joining Prosper when he funds his first loan

[49] Note that when lenders observe late AA to A loans, they do show slight substitution towards the lower credit grade loans.

We also directly test whether lenders shift toward loans with higher rates of return in response to late loans in their portfolio. In Column 8 of Table 11 we present results of a regression as above but with the average IRR2 of loans funded by a lender in a given week as the dependent variable and the percent of the lender's portfolio that has ever been late as the key explanatory variable. As lenders see more late loans in their portfolios, subsequent loans that they fund do in fact have a higher rate of return. The coefficient implies that when the average lender sees a ten percentage point increase in the portion of his portfolio that has been late, his newly funded loans have a 2.17 percentage point higher rate of return. A more careful look into grade-specific performance suggests that the improvement of IRR is more sensitive to bad performance in B to D loans that that of A-AA or E-HR.

Lower returns to some types of social network related loans may represent another form of charity. We look to distinguish between lender mistakes and lender charity by testing if lenders learn from poor performing social loans as well. Similar regressions to Specifications (6)-(8) that focus on social network portfolio composition instead of credit grade composition are shown in Table 12. Lenders show expected substitution patterns between loans with friend endorsement no bids, friend endorsements and bids, and no friend endorsements (Panel A), and between group and non-group loans (Panel B). These results indicate that the market has reacted to the performance of these types of loans at least on the funding margin even if interest rates have not adjusted accordingly.

One may argue that charity lending is more likely to occur if lenders and borrowers belong to the same network (e.g. Harvard alumni help Harvard students). To check this, we run versions of Specifications (6)-(8) to test whether group lenders also substitute away from own group loans when they observe late own group loans in their portfolios. As reported in Panel C of Table 12, lenders who belong to groups fund less own group loans when previous own group loans in their portfolios have been late. This suggests within group charity is not a large factor.

An alternative explanation for the IRR fluctuation is mean reversion: if a lender's first investment receives a bad draw, the next investment is likely to perform better even if the investment is random. Similarly, if the first investment happens to receive a good draw, a random second investment is likely to be worse. To test this explanation, we separate lenders into two groups, depending on whether the average expected IRR of their first-month investment is above or below the market-wide average IRR of the corresponding month. Figure 13 plots the average expected IRR of later loans funded by these two groups by a lender's Prosper age. By law of large numbers, if mean reversion is the main explanation, the two groups should have their average IRRs converged to the same level immediately after the first month. In contrast, if the two groups differ systematically in their initial ability of risk evaluation, we would expect the below-mean group to catch up through learning and the IRR gap between the two groups to shrink gradually over time. While we cannot rule out some degree of mean reversion, the gradual pattern shown in Figure 13 suggests learning is important, and mean reversion is not the sole explanation for the coefficients discussed above. In addition, we note that mean reversion cannot explain why lenders try to avoid funding future listings that have the same observables as the bad loans in their earlier portfolios.

7.2 Lender Mistakes vs. Macro Shock

To this point, our results suggest that lenders do in fact learn to choose higher return loans when they observe late loans in their portfolio, and they substitute from particular types of loans when these types of loans perform poorly. These results suggest that the low returns of Prosper loans are better explained by lender "mistakes" as opposed to lender "charity." However, they do not rule out the macroeconomic recession as an explanation for low ex post returns. Here, we examine how learning varies by time to see if lenders exibited learning both before and after the subprime crisis.

We run similar regressions as above, but we interact the ever late variables with dummies indicating different 6 month periods, with the first six month period of our sample beginning in June 2006 as the excluded group. Results are presented in the Panel A of Table 13. During the six month period beginning in June 2006, a ten percentage point increase in the portion of a lender's portfolio that had ever been late leads to a 0.54 percentage point decreases in the probability of funding a loan, a $75 decrease in the amount funded in weeks with any investment, and a 2.28 percentage point higher average return on funded loans. The interaction coefficients reveal that the magnitude of these coefficients have changed over time. Except for the funding decision in the first half of 2008, all the implied learning effects are no smaller (in absolute value) in June 2006 through June 2007 than in later periods. This result suggests that lenders made mistakes and learned from them even before the recession.

7.3 Explaining Heterogeneous Returns

Finally, we test the hypothesis that the large heterogeneity of IRRs is likely driven by differences in lender's sophistication in risk screening. If lenders differ in their ability to screen loans and their ability to learn from their mistakes, competition will not lead to equalized expected returns across funded loans. The first evidence in support of this argument is that we get very different results if we rerun specifications (6) to (9) *without* lender fixed effects. For example, the coefficient of portfolio late in the IRR regression is be -0.264 instead of 0.217 and remains statistically significant at the 1% level. This finding suggests that lender characteristics are correlated with portfolio performance, and therefore lenders differ in their ability to identify borrower risk. However, because lender heterogeneity is confounded with many macro factors over time, we turn to examining the differences in how lenders learn from their own portfolio performance.

Panel B of Table 13 tests whether different lender cohorts learn at different rates. We define a cohort as the group of lenders who funded their first Prosper loan during each six month period. While our sample only includes lending activity after June 2006, it still includes lenders who joined Prosper prior to this date. Therefore, the first cohort, which is the reference group, contains those lenders who began lending on Prosper during the first half of 2006. The coefficients on the cohort interaction terms are all of opposite signs to the main effect. This pattern implies that lenders in latter cohorts respond less to the late loans in their portfolios as compared to the first cohort on both the extensive and intensive funding margin and on the quality of their chosen loans. One caveat in interpreting these results is that the latest cohorts have had less opportunity to observe late loans in their portfolios, which could lead to the decreased relationship between late loans and lending behavior, particularly the large negative coefficient on the interaction between percent late and the last cohort (the second half of 2008) in the IRR regression.

One reason the lenders in later cohorts may learn at a lower speed than those in earlier cohorts is that they choose higher return loans when they initially join Prosper. Figure 14 plots the average lender's IRR2 for loans he funds in a given week by lender cohort, where this figure defines cohorts by the quarter in which they funded their first loan. As lenders age, they clearly fund loans with a higher rate of return. Interestingly, new cohorts pick up the market trend, perhaps responding to information revealed by the market that was not available when older lenders joined Prosper.

To look at another aspect of lender heterogeneity we investigate whether less sophisticated lenders learn faster *within* the same lender cohort. As described above, we separate lenders into "above-mean" and "below-mean" groups according to whether the average IRR of a lender's first month's portfolio is above or below the market-wide average of that month. If the two groups are fundamentally different in the ability to evaluate risk, the "above-mean" group should be more sophisticated at the beginning but the "below-mean" group should catch up in the degree of sophistication over time. To formally test this idea, we first define the unit of observation as cohort c in group g at lender age a, where cohort and age are both measured in months. We then regress the average IRR of all loans funded by c, g at age a on cohort fixed effects, age fixed effects, a dummy indicating the "above-mean" group ($1_{abovemean}$) and interactions of $1_{abovemean}$

with a linear cohort term and with a linear age term:

$$\begin{aligned} AverageIRR2_{cga} =& \mu_{cohort} + \gamma_{age} + \beta_1 \cdot 1_{abovemean} \\ & + \beta_2 \cdot 1_{abovemean} \cdot cohort + \beta_3 \cdot 1_{abovemean} \cdot age \\ & + \beta_4 \cdot 1_{abovemean} \cdot cohort \cdot age + \epsilon_{cga}. \end{aligned} \tag{10}$$

As shown in Table 14, we find $\beta_1 = 0.073$, which indicates that the first-month portfolio of "above-mean" lenders has an average IRR 7.3 percentage points higher than the "below-mean" lenders. While this difference could be due to lender heterogeneity in sophistication or pure luck, the estimates of β_2 (-0.004) and β_3 (-0.002) suggest that later cohorts are systematically more homogeneous and the IRR difference between above- and below-mean lenders declines steadily within each cohort as "below-mean" lenders learn. Taken together, these results seem to suggest heterogeneity in the initial level of lender sophistication. The estimate of β_4 (0.0001) is positive, suggesting that the relatively less sophisticated lenders in the later cohorts learn almost as much as the earlier cohorts, even if they are more similar to the sophisticated lenders when they started lending on Prosper.

In addition to heterogeneity in learning, we also examine two alternative explanations of IRR heterogeneity. First, group leader incentives may have led to the lower returns of loans to borrowers who were group members. Recall that before September 12, 2007, a group leader could earn monetary rewards for every loan funded in her group, which generates an incentive for the group leader to relax risk evaluation, endorse and bid on a group member's listing, and get as many listings funded as possible. Such incentive should have been reduced after Prosper eliminated the group leader rewards in September 2007. Additionally, the introduction of group ratings should lead to better risk screening if a group leader cares about group reputation.

In Table 15, we statistically test how the IRR gap between group and non-group loans changed with the introduction of group ratings and the removal of group leader rewards. The initial gap between group and non-group loans was a statistically significant 4.2 percentage points. This gap increased slightly when ratings were adopted, but decreased by a statistically significant 1.3 percentage points after leader rewards were eliminated. While this regression does not necessarily estimate the causal effect, it appears that the incentives of group leaders rewards may have partially contributed to the lower returns to group loans.

Another explanation for IRR heterogeneity is bidder sniping. Roth and Ockenfels (2002) suggest that, in a common value auction, bidders may submit their bids towards the end of a hard-close auction because informed bidders want to avoid giving information to others through their own early bids. Since uninformed bidders have less chance to compete with the sniping bidders, this could generate abnormal positive returns for the auctions that end with sniping.

In our data, we observe some sniping for the loans without auto-funding.[50] The median difference between the ending time of a listing and the last bid is 5.07 minutes, and 49.4% of loans have at least one bid within the last 5 minutes. Conditional on having a bid within the last 5 minutes, the average number of bids is 3 and the median is 2. These statistics may over-estimate bidder sniping for two reasons. First, because of Prosper's use of proxy bidding, additional bids will be placed in response to a new bid if the new bid is lower than the prevailing interest rate but higher than the specified interest rate of previous bidders. Second, a fraction of Prosper bids were submitted automatically via so-called portfolio plans. In a portfolio plan, a lender specifies the types of listing attributes he likes and then delegates Prosper to choose which listing to bid on. This automatic bidding would not necessarily predict abnormal returns.

If bidder sniping is a significant factor explaining the heterogeneity of returns, we should observe higher returns for the loans that end with sniping bids. To examine this, we regress the estimated IRR2 on the number of bids submitted within the last 5 minutes and the time difference between the last bid and the end of the auction.[51] Contrary to expectation, the first variable is negatively correlated with the returns and the coefficient of the second variable is close to zero. Both suggest that bidder sniping contributes little to the heterogeneity of returns.

8 Conclusion

We examine how online lenders in a peer-to-peer lending market cope with information asymmetry that is likely to be exaggerated on the Internet. Evidence suggests that individual lenders on Prosper.com do face increased adverse selection because they do not observe borrowers' actual credit scores, and many lenders have made mistakes in their loan selection due to lack of expertise in risk evaluation. The latter is gradually improved by lender learning, and the former is partially offset by lenders charging higher interest rates on minus grade loans since Prosper began posting more detailed credit information in February 2007. In October of 2008 Prosper's lending market temporarily closed while it registered with the SEC. Since reopening in July 2009, many changes were made to the site's institutional structure, including a refinement of the definition of credit grades to be based on predicted rates of return instead of coarse credit score intervals. This change may further alleviate the adverse selection within grades.

Online P2P lending also differs from traditional banks through its use of social networks. Our results suggest that, depending on the incentive structure, some social network variables may convey soft information about borrower risk and therefore has a potential to compensate the lack of hard information on Prosper.com. That being said, the data reveal higher returns

[50]By definition, auto-funded loans are ended automatically when the loan is fully funded so they are closed before the specified auction time.

[51]We also control for a full set of year-week fixed effects and a dummy indicating autofunded loans.

for loans with the positive signals and lower returns for loans with the negative signals, implying that the market has not fully understood the signaling effect of social networks to the point where returns are equalized.

Overall, the past two years have seen substantial changes in the nature of P2P lending. Although the estimated rate of return to Prosper loans is still on average lower than alternative investment vehicles such as CDs and T-bills, both the degree of lender learning and the distinctive effects of social networks point towards an optimistic future. However, as lenders realize the actual risk on the Internet, the P2P market has excluded more and more subprime borrowers and evolved towards the population served by traditional credit markets. This suggests that, unless P2P platforms can improve the power of social networks for borrowers with low credit scores, P2P lending is likely to compete head-to-head with traditional banks in the future and would not provide a viable alternative for those excluded from traditional credit markets.

9 Appendix: IRR algorithm and potential bias

For each loan funded in our sample period, we use the following algorithm to compute the expected internal rate of return (IRR):

- Given a loan amount and the interest rate of the loan, calculate the amortized monthly payment ($MonthlyPay$) and the proportion that goes into the payment of principal ($MonthlyPrincipalPay$). Specifically, we define $MonthlyPay = [(InterestRate/12) * LoanSize{*}(1{+}InterestRate/12)^{36}]/[(1{+}InsterestRate/12)^{36}{-}1]$. $MonthlyPrincipalPay_t = (MonthlyPay - InterestRate * LoanSize/12) * (1 + InterestRate/12)^{(t-1)}$.

- Regress the observed monthly performance of each loan (which includes four dummies that indicates default, default or late, miss payment and paid off)[52] on an exhaustive set of loan age dummies plus all the listing variables available since June 1, 2006. To be comprehensive, we also interact all of the listing variables with each credit grade. These four prediction regressions yield the predicted probability of the four outcomes by loan-month for each loan that was originated on Prosper between June 1, 2006 and July 31, 2008. Because we do not observe any loans older than 24 months we have to make assumptions about the performance in months 25 through 36. In one version we assume that the cumulative mis-performance remains constant after month 24 (referred to as "flat IRR"). The other version assumes that the mis-performance rate will follow a linear projection after month 24 ("linear IRR"). In this version the predicted mis-performance rate at month x (where $x \geq 25$) is equal to [predicted risk at month 24 + (x-24)*(predicted

[52]If a loan is defaulted at month t, it is counted as default in all months after t.

risk at month 24-predicted risk at month 23)]. In both versions, we assume no new early payoff occurs after month 24.

- For loan i at month t, define $PrincipalRemain_{it}$ as the remaining principal after considering the possibility of mis-performance or early payoff in that month. This is calculated iteratively. In the first month of the loan, $PrincipalRemain_{i1} = LoanSize$. For the other months, $PrincipalRemain_{it} = PrincipalRemain_{it-1} - MonthlyPay * (1 - Prob_{it}(nopay) - Prob_{it}(paidoff))$.

- For loan i and month t, define $NetMonthlyReturn_{it}$ as the difference between the expected payment and lender service fees, where the expected monthly payment is set equal to $PrincipalRemain * (1 + InterestRate/12) * Prob(paidoffinthismonth) + MonthlyPay * (1 - Prob_{it}(nopay) - Prob_{it}(paidoff))$ and the monthly lender service fee is set equal to $LenderFee/12 * PrincipalRemain_{it}$.

- Solve for the annual discount factor R that equalizes LoanSize to the sum of the present value of $NetMonthlyReturn$ from month 1 to month 36, while using R as the annual discount factor.

- Compute IRR as the annual percentage yield after compounding R monthly. That is $IRR = (1 + R/12)^1 2 - 1$. By definition, R is bounded between -1200% and the contract rate, and IRR is bounded between -100% and $(1 + \text{contract rate}/12)^1 2 - 1$.

Our method is more comprehensive than the ones used by Prosper and LendingStats.[53] Specifically, when we predict mis-performance in a specific month, we regress observed loan-month performance on all listing attributes posted online, their interactions with each credit grade, and a set of loan age dummies (in months). This method utilizes more listing information than Prosper's grade-specific predictions.[54] Unlike LendingStats, we consider the fact that every loan has a positive risk at the time of origination even if ex post it is paid in full.[55] In this sense,

[53] LendingStats is a popular independent website that tracks Prosper activities in real time.

[54] For a given portfolio, Prosper assumes that roll rates from one loan status to another are revealed in the historical performance of each grade. For example, a 1-month-late AA loan has a 42% probability of becoming 2-months late, and a 2-months-late AA loan has a 71% probability of becoming 3-months-late. Prosper also makes assumptions on the probability of early payment-in-full (3.5% for AA, 0.5% for HR) and the probability of loss recovery if default. These assumptions enter the calculation of monthly returns. Annualizing this figure and averaging it across the whole life of the loan results in an overall rate of return.

[55] Compared to Prosper, LendingStats uses more pessimistic roll rate assumptions but puts more emphasis on actual performance than predicted performance. In particular, it takes the current status of a loan as given, and does not project its future risk until it is late. In that case, it is assumed that a 1-month-late loan will default with a 50% probability and loans that are more than 1 month late will default for sure.

we capture the return *expected* at the time of origination, not the return that is realized ex post.

The algorithm described above is subject to potential bias in both directions. On the one hand, our IRR estimates may be downward biased because we try to be conservative in the calculation of cash flows. Specifically, we assume away any loss recovery from default loans, and we do not account for the late fees that a lender may receive from a late-but-non-defaulting borrower. When we count early payoff as a bulk cash flow that arrives in the paid-off month, it effectively assumes that the paid off amount is reinvested in a loan that is identical to the loan under study. This assumption may be conservative because lenders may learn to fund better loans over time.

On the other hand, our IRR estimates may have overestimated the return of investment because we do not consider any cost that lenders may incur in processing Prosper information. The time that lenders spend on screening listings and digesting Prosper history could be long and stressful. Lastly, our IRR estimates are based on the average loan performance observed from June 1, 2006 to July 31, 2008, a period that stretches from the end of a boom to the beginning of a recession. If the recession prolongs and worsens over time, the reported IRR will overestimate the actual rate of return.

10 References

Adam, William; Liran Einav and Jonathan Levin (2009): "Liquidity Constraints and Imperfect Information in Subprime Lending" *American Economic Review* 99(1): 49-84.

Akerlof, George A. (1970): "The Market for 'Lemons': Quality Uncertainty and the Market Mechanism" *Quarterly Journal of Economics* 84(3): 488-500.

Berger, Allen N. and Gregory F. Udell (1992): "Some Evidence on the Empirical Significance of Credit Rationing" *The Journal of Political Economy* 100(5): 1047-1077.

Besley, Timothy; Stephen Coate and Glenn Loury (1993) "The Economics of Rotating Savings and Credit Associations" *The American Economic Review* September 1993, 792-810.

Besley, Timothy and Stephen Coate (1995) "Group Lending, Repayment Incentives and Social Collateral" *Journal of Development Economics* Vol. 46, 1-18.

Caskey, John P. (2005): "Fringe Banking and the Rise of Payday Lending", Chapter 2 of *Credit Markets for the Poor*, Russell Sage Foundation, New York.

Cawley, John, and Tomas Philipson (1999), "An Empirical Examination of Information Barriers to Trade in Insurance," *American Economic Review*, Vol. 89, pp. 827-846.

Chiappori, Pierre-Andre, and Bernard Salanie (2000), "Testing for Asymmetric Information in Insurance Markets," *Journal of Political Economy*, Vol. 108, pp. 56-78.

Cohen, Alma and Peter Siegelman (2009), "Testing for Adverse Selection in Insurance Markets"

NBER working paper #15586.

Cox, Donald and Tullio Japelli (1990) "Credit Rationing and Private Transfers: Evidence from Survey Data" *The Review of Economics and Statistics* 72(3): 445-454.

Finkelstein, Amy, and Kathleen McGarry (2006), "Multiple Dimensions of Private Information: Evidence from the Long-Term Care Insurance Market," *American Economic Review*, Vol. 96, pp. 938-958.

Finkelstein, Amy, and James Poterba (2006) "Testing for Adverse Selection with 'Unused Observables"', NBER working paper #12112.

Finkelstein, Amy, and James Poterba (2004), "Adverse Selection in Insurance Markets: Policyholder Evidence from the U.K. Annuity Market," *Journal of Political Economy,* Vol. 112, pp. 183-208.

Freedman, Seth and Ginger Z. Jin (2008) "Do Social Networks Solve Information Problems for Peer-to-Peer Lending? Evidence from Prosper.com" Net Institute Working Paper 08-43.

Greenlaw, David; Jan Hatzius; Anil Kashyap, Hyun Song Shin (2008) : "Leveraged Losses: Lessons from the Mortgage Market Meltdown", mimeo, accessed at http://www.chicagogsb.edu/usmpf/docs/usmpf2008confdraft.pdf.

Gross, David B. and Nicholas S. Souleles (2002) "Do Liquidity Constraints and Interest Rates Matter for Consumer Behavior: Evidence from Credit Card Data" *Quarterly Journal of Economics* 117(1): 149-185.

Grossman, S. (1981) "The informational Role of Warranties and Private Disclosure about Product Quality," *Journal of Law and Economics* Vol. 24 (1981), pp. 461-489.

Hampshire, Robert (2008) "Group Reputation Effects in Peer-to-Peer Lending Markets: An Empirical Analysis from a Principle-Agent Perspective" *mimeo.*

Hendel, Igal; Aviv Nevo and Francois Ortalo-Magne (2008) "The Relative Performance of Real Estate Marketing Platforms: MLS versus FSBOMadison.com" *mimeo.*

Hoff, Karla and Joseph E. Stiglitz "Introduction: Imperfect Information and Rural Credit Markets – Puzzles and Policy Perspectives" *the World Bank Economic Review* 4(3): 235-250.

Iyer, Rajkamal; Asim Ijaz Khwaja; Erzo F. P. Luttmer; Kelly Shue (2009) "Screening in Alternative Credit Markets: Can Individual Lenders Infer Borrower Credit-worthiness in Peer-to-Peer Lending" mimeo.

Jaffee, Dwight M. (1971) *Credit Rationing and the Commercial Loan Market* New York: Wiley.

Jin, Ginger Z. and Andrew Kato (2007) "Dividing Online and Offline: A Case Study" *Review of Economic Studies* July 2007.

La Ferrara, Eliana (2003) "Kin Groups and Reciprocity: A Model of Credit Transactions in

Ghaha" *American Economic Review* 93(5): 1730-1751.

Mendelson, Haim (2006) "Prosper.com: A People-to-People Lending Marketplace" mimeo.

Milgrom, Paul (1981) "Good News and Bad News: Representation Theorems and Applications," *The Bell Journal of Economics* Vol. 12 (1981), pp. 380-391.

Parker, Jonathan A. (1999) "The Reaction of Household Consumption to Predictable Changes in Social Security Taxes" *American Economic Review* 89(4): 959-973.

Pope, Devin and Justin R. Sydnor (2008) "What's in a Picture? Evidence of Discrimination from Prosper.com" *mimeo.*

Portes, Alejandro (1998) "Social Capital: Its Origins and Applications in Modern Sociology" *Annual Reviews of Sociology* 24: 1-24.

Ravina, Enrichetta "Love & Loans: The Effect of Beauty and Personal Characteristics in Credit Markets," Available at SSRN: http://ssrn.com/abstract=972801.

Roth, Alvin and Axel Ockenfels (2002) "Last-Minute Bidding and the Rules for Ending Second-Price Auctions: Evidence from eBay and Amazon Auctions on the Internet" *American Economic Review* 92(4): 1093-1103.

Rothschild, Michael, and Joseph Stiglitz (1976), "Equilibrium in Competitive Insurance Markets," *Quarterly Journal of Economics*, 90(4): 629-649.

Salmon, Timothy C. (2001) "An Evaluation of Econometric Models of Adaptive Learning" *Econometrica* 69(6): 1597-1628.

Skiba, Paige and Jeremy Tobacman (2007) "Profitability of Payday Loans" Working paper available at http://www.law.vanderbilt.edu/faculty/faculty-personal-sites/paige-skiba/index.aspx.

Souleles, Nicholas S. (1999) "The Response of Household Consumption to Income Tax Refunds" *American Economic Review* 89(4): 947-958.

Stiglitz, Joseph E. (1990) "Peer Monitoring and Credit Markets" *The World Bank Economic Review* 4:3 351-366.

Stiglitz, Joseph E. and Andrew Weiss (1981): "Credit Rationing in Markets with Imperfect Information" *American Economic Review* 71(3): 393-410.

Udry, Cristopher (1994): "Risk and Insurance in a Rural Credit Market: An Empirical Investigation in Northern Nigeria" *Review of Economic Studies* July 1994, 61(3): 495-526.

Voridis, Hercules (1993): "Ceilings on Interest Rates and Investment: The Example of Greece" *The Review of Economics and Statistics* 75(2): 276-283.

Table 1: Evolution of Prosper Policies

Date initiated	Policy
Always	History updated online every day, Allow groups
May 30, 2006	Reveal more credit info, home ownership status, and bank account status
Oct. 19, 2006	Start group rating based on past loan performance
Feb. 12, 2007	Disallow borrowers with score<520 Reveal more credit info (e.g. amount delinquent)
Feb. 12, 2007	Allow friend endorsements
Sept. 12, 2007	Eliminate group leader rewards ($12/new borrower)
Oct. 30, 2007	Add bidder guidance
Feb. 23, 2008	Allow search by friend bids and endorsements
Apr. 15, 2008	Raise interest rate cap to 36% except for TX (10%) and SD (N/A)
Oct. 14, 2008	Switch to a quiet period, subject to SEC review
Jul. 13, 2009	Reopen with refined credit grades, prosper-defined credit score, secondary market, disallow borrowers with Experian credit score<640

Note: Gray rows indicate social network policies.

Table 2: Summary of Listings and Loans by Quarter

Listings:

Quarter	Number	Total Amount Requested ($100,000)	Mean Amount Requested ($)	Mean Borrower Max Interest Rate	% in a Group	% endorse no bid by group leader	% endorse with bid by group leader	% endorse no bid by friends	% endorse with bid by friends	Funding rate
20062	5375	26.65	4,957.22	16.86%	58.59%	0.00%	0.00%	0.00%	0.00%	10.01%
20063	19771	107.25	5,424.63	18.15%	61.84%	0.42%	0.71%	0.00%	0.00%	9.94%
20064	31629	196.57	6,214.85	17.45%	53.57%	1.33%	2.04%	0.00%	0.00%	7.98%
20071	31373	263.22	8,389.94	16.72%	48.24%	1.42%	3.46%	11.04%	0.58%	10.14%
20072	37505	331.62	8,841.98	17.51%	34.09%	1.07%	5.68%	20.86%	0.97%	8.07%
20073	39353	328.79	8,355.00	18.06%	23.64%	1.01%	4.10%	19.93%	1.14%	6.71%
20074	41585	334.23	8,037.29	18.41%	16.08%	1.42%	0.84%	16.48%	1.33%	6.32%
20081	33485	250.14	7,470.30	19.24%	12.77%	0.70%	0.75%	12.91%	1.86%	9.46%
20082	43371	318.53	7,344.20	24.50%	7.83%	0.54%	0.64%	9.36%	1.58%	10.08%
20083	10361	73.48	7,092.42	26.40%	7.53%	0.53%	0.61%	8.98%	1.89%	9.31%
Total	293808	2230.48	7,591.62	19.19%	28.82%	0.98%	2.23%	12.01%	1.04%	8.51%

Loans:

Quarter	Number	Total Amount Funded ($100,000)	Mean Amount Funded ($)	Mean Contract Interest Rate	% in a Group	% endorse no bid by group leader	% endorse with bid by group leader	% endorse no bid by friends	% endorse with bid by friends	% default
20062	385	1.47	3,822.17	19.03%	67.01%	0.00%	0.00%	0.00%	0.00%	30.39%
20063	1934	9.37	4,844.63	19.41%	71.30%	1.14%	3.10%	0.00%	0.00%	28.54%
20064	2403	11.54	4,804.05	18.97%	70.20%	4.04%	12.82%	0.00%	0.00%	29.09%
20071	3079	19.93	6,472.60	17.37%	67.49%	4.38%	17.93%	10.91%	2.24%	23.74%
20072	3118	23.47	7,527.98	17.42%	63.28%	4.36%	29.76%	27.77%	4.62%	17.54%
20073	2671	18.43	6,900.12	17.31%	44.85%	4.64%	23.40%	26.21%	5.13%	9.21%
20074	2593	18.98	7,320.17	17.11%	23.95%	2.70%	6.44%	22.33%	6.56%	4.09%
20081	3074	20.47	6,658.94	17.37%	19.00%	0.81%	3.81%	17.99%	5.50%	0.46%
20082	4344	26.33	6,061.10	17.98%	13.54%	1.31%	3.06%	14.11%	5.62%	0.00%
20083	1407	8.27	5,877.70	19.39%	10.80%	0.78%	2.70%	12.30%	6.54%	0.00%
Total	25008	158.27	6,328.65	17.90%	42.06%	2.71%	11.71%	15.28%	4.10%	12.04%

Table 3: Summary of Listings and Loans by Credit Grade

	Prosper Listings			Prosper Loans						Experian Accounts opened in Sept. 2003
						observed on 8/1/2008				observed on 9/2005
Credit Grade	Number of listings	Mean Borrower Maximum Interest Rate	Funding Rate	Number of loans	Mean Contract Interest Rate	% Late	% Default	% 3m late or worse	% 3m late or worse (if loan life = 2 years)	% 3m late or worse (loan life = 2 years)
2A	9321	11.57%	32.08%	2990	9.70%	1.07%	1.64%	1.91%	2.27%	0.89%
A	11099	14.05%	25.45%	2825	12.29%	2.90%	3.58%	4.04%	4.08%	3.33%
B	17211	16.47%	21.88%	3766	15.00%	3.90%	5.50%	6.27%	12.24%	6.04%
C	30843	18.57%	15.76%	4862	17.49%	5.29%	8.97%	10.24%	22.73%	9.44%
D	43282	20.08%	10.35%	4479	20.66%	5.20%	11.23%	12.30%	23.08%	15.29%
E	52000	20.65%	5.56%	2891	24.82%	7.06%	21.27%	22.79%	43.64%	24.25%
HR	128633	19.83%	2.39%	3077	24.52%	7.12%	33.67%	35.36%	48.18%	34.40%
NC	1419	17.66%	8.32%	118	22.06%	4.24%	55.08%	56.78%		

Table 4: Summary Statistics of Listing Attributes (June 1, 2006 – July 31, 2008)

	Listings			Loans		
	Mean	STD	N	Mean	STD	N
Information available before Feb. 12, 2007						
Grade=AA	0.032	0.175	293808	0.120	0.324	25008
Grade=A	0.038	0.191	293808	0.113	0.317	25008
Grade=B	0.059	0.235	293808	0.151	0.358	25008
Grade=C	0.105	0.307	293808	0.194	0.396	25008
Grade=D	0.147	0.354	293808	0.179	0.383	25008
Grade=E	0.177	0.382	293808	0.116	0.320	25008
Grade=HR	0.438	0.496	293808	0.123	0.328	25008
Grade=NC	0.005	0.069	293808	0.005	0.069	25008
amountrequested	7592	6388	293808	6329	5679	25008
autofunded	0.311	0.463	293808	0.263	0.441	25008
borrowermaximumrate	0.192	0.084	293808	0.209	0.074	25008
yeshomeowner	0.327	0.469	293808	0.441	0.497	25008
debt-to-income (DTI) ratio	0.505	1.359	293808	0.330	0.978	25008
missing DTI	0.068	0.251	293808	0.035	0.183	25008
DTI topcoded if DTI>=10	0.083	0.275	293808	0.044	0.205	25008
have image	0.515	0.500	293808	0.659	0.474	25008
length of listing desc (in chars)	1058	772	293808	1295	866	25008
mention debt consolidation	0.358	0.480	293808	0.375	0.484	25008
mention business loan	0.231	0.421	293808	0.271	0.444	25008
mention car	0.689	0.463	293808	0.626	0.484	25008
mention mortgage	0.139	0.346	293808	0.187	0.390	25008
mention health	0.721	0.449	293808	0.790	0.407	25008
mention education	0.211	0.408	293808	0.248	0.432	25008
mention family	0.179	0.383	293808	0.189	0.392	25008
mention retirement	0.030	0.171	293808	0.041	0.199	25008
mention pay-day loan	0.057	0.233	293808	0.057	0.231	25008
concede relisting	0.008	0.089	293808	0.021	0.144	25008
# of listings (incld current one)	2.811	3.361	293808	2.912	2.863	25008
interest rate cap	0.243	0.093	293808	0.273	0.082	25008
borrower fee	1.800	0.794	293808	1.548	0.781	25008
lender fee	0.852	0.231	293808	0.790	0.258	25008
amountdelinquent ($)	3516	12374	221618	1176	6257	18618
missing amountdelinquent	0.004	0.066	221618	0.001	0.037	18618
currentdelinquency	3.833	5.303	293808	1.454	3.400	25008
delinquency in 7yrs	11.022	16.450	293808	5.800	12.356	25008
lengthcredithistory (in days)	152.208	84.472	293808	158.049	87.107	25008
totalcreditlines	24.354	14.393	293808	23.964	14.424	25008
in public records in past 10 years	0.657	1.395	293808	0.405	0.936	25008
# of inquiries in past 6 months	4.153	4.959	293808	2.927	3.979	25008

Table 4 Continued: Summary Statistics of Listing Attributes (June 1, 2006 – July 31, 2008)

	Listings			Loans		
	Mean	STD	N	Mean	STD	N
Credit info added after Feb. 12, 2007						
currentcreditlines	8.230	6.001	221618	9.566	5.931	18618
opencreditlines	7.224	5.303	221618	8.165	5.223	18618
band card utilization rate	0.629	0.431	221618	0.547	0.373	18618
revolving balance ($)	12087	31802	221618	16326	39388	18618
in publice records in past 1 year	0.075	0.346	221618	0.040	0.237	18618
working full time	0.821	0.383	221618	0.859	0.348	18618
working part time	0.040	0.196	221618	0.038	0.192	18618
income 25-75 K	0.670	0.470	202271	0.651	0.477	17782
income > 75K	0.144	0.351	202271	0.220	0.415	17782
missing income	0.297	0.457	293808	0.284	0.451	25008
no employment or income reported	0.014	0.119	293808	0.005	0.072	25008
missing new credit info posted after 2/07	0.000	0.013	293808	0.000	0.018	25008
missing credit info posted bef 2/07	0.008	0.087	293808	0.004	0.062	25008
Social network variables						
borrower in a group	0.288	0.453	293808	0.421	0.494	25008
borrower having any friend	0.191	0.393	293808	0.249	0.432	25008
listing with endorsement+nobid by group leader	0.010	0.098	293808	0.027	0.162	25008
listing with endorsement+nobid by friend	0.120	0.325	293808	0.153	0.360	25008
listing with endorsement+bid by group leader	0.022	0.148	293808	0.117	0.322	25008
listing with endorsement+bid by friend	0.010	0.101	293808	0.041	0.198	25008

Table 5: fund rate, interest rate and default or late (June 1, 2006 – July 31, 2008)

	Funded?	Contract interest rate	Default or late as of 8/1/2008
	Probit (marginal effects)	OLS	Probit (marginal effects)
Listing attributes available before Feb 2007			
Grade=AA	0.696*	-0.032*	-0.134*
	(21.041)	(-8.258)	(-17.371)
Grade=A	0.409*	-0.026*	-0.118*
	(14.144)	(-8.326)	(-14.290)
Grade=B	0.252*	-0.021*	-0.122*
	(11.146)	(-6.855)	(-11.878)
Grade=C	0.095*	-0.016*	-0.121*
	(8.135)	(-5.175)	(-9.000)
Grade=D	0.033*	-0.008*	-0.115*
	(6.020)	(-2.597)	(-9.295)
Grade=E	0.001	-0.003	-0.089*
	(0.578)	(-0.889)	(-7.667)
Grade=HR	-0.005*	-0.003	-0.048**
	(-2.711)	(-1.068)	(-2.370)
amountrequested	-0.000*	0.000*	0.000*
	(-32.188)	(17.345)	(11.306)
autofunded	0.011*	0.036*	0.047*
	(20.910)	(90.511)	(9.190)
borrowermaximumrate	0.702*	0.458*	1.732*
	(38.743)	(17.009)	(6.271)
borrowermaximumrate2	-1.161*	0.484*	-2.308*
	(-34.968)	(7.978)	(-3.871)
yeshomeowner	0.001***	0.001	-0.060*
	(1.729)	(0.608)	(-5.085)
Debt to income ratio	-0.017*	0.004*	0.015*
	(-15.799)	(5.433)	(2.822)
debt-to-income * homeowner	0.001**	-0.000	0.004
	(2.156)	(-0.711)	(1.205)
having an image	0.005*	-0.001*	-0.003
	(16.152)	(-3.449)	(-0.582)
length of description	0.000*	-0.000*	-0.000*
	(18.177)	(-5.380)	(-2.987)
mention debtconsolidation	0.001**	0.000	-0.010**
	(2.372)	(0.174)	(-2.493)
mention business	-0.001*	0.000	0.021*
	(-3.199)	(1.062)	(4.359)
mention car	-0.000	0.001**	0.011**
	(-1.207)	(2.524)	(2.244)
mention mortgage	0.000	0.000	0.008
	(0.315)	(0.026)	(1.514)
mention health	0.001*	0.001**	0.008
	(2.647)	(1.997)	(1.563)

mention education	0.000	-0.000	-0.008***
	(0.016)	(-0.132)	(-1.782)
mention family	0.001*	0.001**	0.014*
	(2.887)	(2.054)	(2.846)
mention retirement	-0.001**	-0.001	-0.004
	(-2.053)	(-1.343)	(-0.455)
mention pay-day loan	0.003*	0.003*	0.025*
	(5.157)	(3.113)	(2.884)
saidrelisting	0.008*	0.002	0.009
	(4.951)	(1.570)	(0.695)
count of relisting	-0.000*	0.001*	0.002*
	(-5.666)	(6.828)	(3.225)
currentdelinquencies	-0.001*	0.000*	0.008*
	(-18.347)	(4.454)	(10.886)
delinquencies in passt 7 yrs	-0.000*	0.000*	-0.000**
	(-11.790)	(6.091)	(-2.317)
length of credit history	-0.000*	0.000*	-0.000***
	(-7.724)	(3.973)	(-1.660)
totalcreditlines	0.000	0.000**	-0.001*
	(0.438)	(2.171)	(-4.257)
In public records in past 10 yrs	-0.001*	0.000	0.002
	(-8.563)	(0.126)	(1.350)
# of inquiries in past 6m	-0.001*	0.000*	0.007*
	(-13.047)	(6.197)	(12.927)
missing credit info	0.002	-0.005	0.032
	(0.957)	(-1.359)	(0.843)
Social network variables			
in_a_grp_borrower	0.004*	-0.004*	0.005
	(10.464)	(-9.299)	(0.988)
have endorsement + nobid by group leader	0.017*	-0.003*	-0.000
	(8.118)	(-2.798)	(-0.011)
have endorsement + bid by group leader	0.096*	-0.005*	0.005
	(21.708)	(-7.294)	(0.772)
have endorsement + nobid by friend	0.002*	0.001**	0.009
	(5.304)	(2.367)	(1.435)
have endorsement + bid by friend	0.050*	-0.007*	-0.041*
	(12.454)	(-6.150)	(-4.830)
Year-week FE	Yes	Yes	Yes
N	293,802	25,008	23,344
Adjusted R2	0.375	0.855	0.269

The sample includes all the listings and loans between June 1, 2006 and July 31, 2008. T-statistics are in parenthesis. * p<0.01, ** p<0.05, *** p<0.1. All regressions control for state dummies, year-week FE, macro variables, variables for prosper environment, duration of auction, and indicators for missing debt-to-income ratio and other credit attributes. All regressions do not include the new credit variables added after Feb. 12, 2007. In an unreported table, we show that regressions including these variables and condition on the post-Feb-2007 sample generate similar results.

Table 6-1: Half Grade Regressions: include month counts since June 2006

Unit of observation = census division by month by half-grade interval
Sample: the half-grade intervals that have credit scores at or above 520

	# of listings		**# of loans**		**Funding rate**		**Average contract interest rate**		**% late in 6m**		**% late in 12m**	
Dummy of minus grade	11.381*	4.351*	1.908*	-0.006	0.006	0.022	0.004*	-0.000	0.014**	0.012	0.023*	0.024
	(4.962)	(3.791)	(4.124)	(-0.024)	(0.671)	(0.877)	(2.989)	(-0.216)	(2.213)	(1.073)	(2.825)	(1.008)
Monthcount * nearprime		1.614*		0.090		-0.004*		0.000		0.001		0.001
		(5.996)		(1.525)		(-2.589)		(0.750)		(0.524)		(0.277)
Monthcount * subprime		1.416*		-0.227*		-0.004*		0.001*		0.001		-0.002
		(6.100)		(-3.121)		(-2.643)		(3.995)		(0.425)		(-0.622)
Monthcount * minus grade		0.547*		0.183*		-0.001		0.000		0.000		0.000
		(5.384)		(4.143)		(-0.926)		(0.941)		(0.233)		(0.116)
N	3,978	3,978	3,978	3,978	3,779	3,779	3,357	3,357	2,776	2,776	2,006	2,006
R2	0.679	0.700	0.552	0.581	0.269	0.271	0.795	0.801	0.145	0.144	0.208	0.207

T-statistics are in parenthesis. * p<0.01, ** p<0.05, *** p<0.1. All regressions control for credit grade FE, year-month FE, and census-division FE, standard errors clustered by census division.

Table 6-2: see next page.

Table 6-3: Placebo test: shift up the definition of half grade by 20 points in specification (4) without month count

Unit of observation = census division by month by half-grade interval
Sample: the half-grade intervals that have credit scores at or above 520

	# of listings	# of loans	Funding rate	Average contract interest rate
Dummy of minus grade	11.381*	1.908*	0.006	0.004*
(true data)	(4.962)	(4.124)	(0.671)	(2.989)
Dummy of minus grade	5.775*	-0.752*	-0.065*	0.019*
(placebo)	(5.730)	(2.897)	(4.769)	(13.091)
N	3,978	3,978	3,978	3,978

T-statistics are in parenthesis. * p<0.01, ** p<0.05, *** p<0.1. All regressions control for credit grade FE, year-month FE, and census-division FE, standard errors clustered by census division.

Table 6-2: Half Grade Regressions: use macro and Prosper policy variables instead of month counts

		# of listings		# of loans	Funding rate	Average Contract Interest Rate	% late in 6m	% late in 12m
Dummy of minus grade	11.381*	4.087*	3.876*	0.262	0.008	0.002	0.015	0.043
	(5.012)	(4.202)	(3.353)	(1.115)	(0.324)	(0.871)	(1.001)	(0.680)
AftNewInfo(Feb07)*minus grade		1.589*	1.841*	3.023*	-0.009	0.006*	0.015	0.018
		(2.724)	(2.768)	(3.569)	(-0.433)	(5.048)	(0.928)	(0.789)
AftBidguide(Oct07)*minus grade		-0.205	0.337	-0.411	-0.010	-0.003	-0.028	NA
		(-0.727)	(0.442)	(-1.617)	(-0.346)	(-0.658)	(-0.920)	.
AftRateCap(Apr08)*minus grade		3.876*	4.246**	1.554***	0.053	-0.004	NA	NA
		(5.370)	(2.511)	(1.783)	(0.925)	(-0.943)	.	.
AftNewInfo(Feb07)*HR*minus grade		61.338*	60.991*	-2.199*	0.001	-0.008***	0.019	0.069***
		(4.944)	(4.890)	(-3.200)	(0.091)	(-1.904)	(0.726)	(1.675)
AftNewInfo(Feb07)*nearprime		19.909*	20.443*	3.079**	-0.089**	-0.004	0.012	0.041
		(5.372)	(5.256)	(2.431)	(-2.355)	(-1.043)	(1.058)	(1.466)
AftBidguide(Oct07)*nearprime		0.550	11.532*	-3.966*	-0.049	0.003	-0.023	NA
		(0.400)	(4.231)	(-3.705)	(-0.924)	(0.562)	(-0.560)	.
AftRateCap(Apr08)*nearprime		14.167*	25.559*	2.459	0.041	0.015	NA	NA
		(4.933)	(3.891)	(1.225)	(0.736)	(1.432)	.	.
AftNewInfo(Feb07)*subprime		29.864*	32.004*	-0.783	-0.093*	-0.005	0.023	-0.051
		(5.492)	(5.862)	(-0.771)	(-2.659)	(-0.893)	(1.041)	(-1.589)
AftBidguide(Oct07)*subprime		-15.964*	-8.817*	-2.707*	-0.010	0.002	-0.066*	NA
		(-5.123)	(-3.220)	(-4.490)	(-0.248)	(0.207)	(-3.249)	.
AftRateCap(Apr08)*subprime		10.511*	14.600***	0.042	-0.016	0.023	NA	NA
		(5.195)	(1.851)	(0.064)	(-0.286)	(1.259)	.	.
N	3,978	3,978	3,978	3,978	3,779	3,357	2,776	2,006
r2_a	0.695	0.771	0.775	0.640	0.275	0.812	0.146	0.215

T-statistics are in parenthesis. * $p<0.01$, ** $p<0.05$, *** $p<0.1$. All regressions control for credit grade FE, year-month FE, census-division FE, all macro variables, and nearprime, subprime and minus-grade interacting with bank prime rate, unemployment rate and foreclosure rate. Standard errors clustered by census division.

Table 7: Summary Statistics of Social Network Variables

	Listings			Loans		
	Mean	SD	N	Mean	SD	N
% In a Group	0.288	0.453	293,808	0.421	0.494	25,008
% with Friends	0.191	0.393	293,808	0.249	0.432	25,008
% w/ Group Leader Endorsement no Bid	0.010	0.098	293,808	0.027	0.162	25,008
% w/ Group Leader Endorsement + Bid	0.022	0.148	293,808	0.117	0.322	25,008
% w/ Friend Endorsement no Bid	0.120	0.325	293,808	0.153	0.360	25,008
% w/ Friend Endorsement + Bid	0.010	0.101	293,808	0.041	0.198	25,008
Conditional on a borrower in a group:						
Number of Members	1799.214	2346.502	84,377	1176.963	1872.194	10,512
Number of Borrowers	1082.372	1311.981	84,377	724.992	1070.800	10,512
Number of Lenders	198.860	248.414	84,377	159.373	217.842	10,512
1-100 Borrowers	0.232	0.422	84,680	0.308	0.462	10,518
101-500 Borrowers	0.225	0.418	84,680	0.296	0.457	10,518
501-1000 Borrowers	0.251	0.434	84,680	0.209	0.406	10,518
> 1001 Borrowers	0.288	0.453	84,680	0.186	0.389	10,518
% of Members that are Borrowers	0.627	0.153	84,377	0.651	0.166	10,512
% of Members that are Lenders	0.138	0.116	84,377	0.202	0.169	10,512
Alumni Group	0.023	0.148	84,680	0.029	0.168	10,518
Military Group	0.019	0.137	84,680	0.014	0.119	10,518
Other Connection	0.017	0.128	84,680	0.022	0.145	10,518
Loose Connection	0.025	0.156	84,680	0.016	0.125	10,518
Listing Review Required	0.341	0.474	84,680	0.519	0.500	10,518
% Funded by Group Members				0.017	0.062	10,518
$ Funded by Group Members				95.818	553.280	10,518
% Funded by Group Leader				0.032	0.124	10,518
$ Funded by Group Leader				131.042	605.746	10,518
Conditional on a borrower in a group & after 10/19/06:						
Low Rated Group	0.414	0.493	66,062	0.275	0.447	8,416
High Rated Group	0.323	0.468	66,062	0.421	0.494	8,416
Nonrated Group	0.261	0.439	66,062	0.301	0.459	8,416
Conditional on a borrower that has friends:						
% Funded by Friends				0.033	0.143	6,205
$ Funded by Friends				190.498	1109.441	6,205
Conditional on a borrower that has endorsement(s):						
% Funded by Endorsing Friends				0.127	0.247	1,022
$ Funded by Endorsing Friends				775.496	2040.419	1,022
% Funded by Endorsing Group Leader				0.068	0.165	2,903
$ Funded by Endorsing Group Leader				322.991	907.019	2,903

Table 8: Credit Grade Composition of Social Network Loans

Groups:

Grade	Group Listings	All Listings	Group Loans	All Loans
A	2.84%	3.78%	8.10%	11.30%
AA	2.18%	3.17%	7.70%	11.96%
B	4.76%	5.86%	11.62%	15.06%
C	9.16%	10.50%	17.83%	19.44%
D	13.10%	14.73%	18.42%	17.91%
E	19.20%	17.70%	15.61%	11.56%
HR	47.96%	43.78%	19.95%	12.30%
NC	0.81%	0.48%	0.78%	0.47%

Friend Endorsements (Listings & Loans after Feb 12, 2007):

Grade	Endorsed Listings	Endorsed + Bid Listings	All Listings	Endorsed Loans	Endorsed + Bid Loans	All Loans
A	3.96%	9.27%	4.35%	10.27%	12.82%	12.61%
AA	2.87%	8.84%	3.49%	9.50%	14.97%	13.23%
B	6.13%	11.96%	6.80%	14.89%	16.24%	16.82%
C	12.03%	16.20%	12.14%	20.98%	16.54%	21.30%
D	16.58%	19.49%	16.91%	19.93%	17.91%	18.55%
E	17.43%	11.20%	17.22%	10.79%	8.41%	8.64%
HR	41.00%	22.91%	39.08%	13.65%	13.11%	8.86%
NC	0.00%	0.03%	0.00%	0.00%	0.00%	0.00%

Table 9-1: fund rate, interest rate and default or late for social variables interacting with credit grade (June 1, 2006 – July 31, 2008)

	Funded?	Contract interest rate	Default or late as of 8/1/2008
	Probit (marginal effects)	OLS	Probit (marginal effects)
Social variables by credit grade			
In a group * Grade= AA or A	0.002**	-0.003*	0.002
	(2.289)	(-3.859)	(0.118)
In a group * Grade = B, C or D	0.004*	-0.005*	0.009
	(6.877)	(-7.768)	(1.245)
In a group * Grade = E or HR	0.007*	-0.006*	0.018**
	(10.242)	(-6.998)	(1.985)
Have group leader endorse+nobid * AA-A	0.009**	0.001	-0.036
	(2.361)	(0.694)	(-1.479)
Have group leader endorse+nobid * B-D	0.020*	-0.006*	0.008
	(6.432)	(-4.085)	(0.528)
Have group leader endorse+nobid* E-HR	0.016*	-0.001	0.004
	(5.004)	(-0.338)	(0.184)
Have group leader endorse+bid * AA-A	0.052*	-0.004*	0.026
	(6.428)	(-3.076)	(1.042)
Have group leader endorse+bid * B-D	0.081*	-0.004*	-0.003
	(15.277)	(-4.840)	(-0.363)
Have group leader endorse+bid * E-HR	0.124*	-0.005*	0.011
	(17.421)	(-4.890)	(1.068)
Have friend endorse+nobid * AA-A	0.002***	0.001	0.014
	(1.811)	(0.930)	(0.778)
Have friend endorse+nobid * B-D	0.001***	0.001	0.008
	(1.889)	(1.164)	(1.059)
Have friend endorse+nobid* E-HR	0.005*	0.002**	0.010
	(6.137)	(2.085)	(0.946)
Have friend endorse+nobid * AA-A	0.027*	-0.004**	-0.073*
	(5.105)	(-2.278)	(-4.896)
Have friend endorse+nobid * B-D	0.037*	-0.007*	-0.044*
	(8.876)	(-4.593)	(-3.837)
Have friend endorse+nobid* E-HR	0.110*	-0.011*	-0.019
	(9.366)	(-3.788)	(-1.096)
Year-week FE	Yes	Yes	Yes
N	293,802	25,008	23,344
Adjusted R2	0.376	0.855	0.270

The sample includes all the listings and loans between June 1, 2006 and July 31, 2008. T-statistics are in parenthesis. * p<0.01, ** p<0.05, *** p<0.1. All regressions control for state dummies, year-week FE, macro variables, variables for prosper environment, duration of auction, and indicators for missing debt-to-income ratio and other credit attributes as in Table 5. All regressions do not include the new credit variables added after Feb. 12, 2007.

Table 9-2: fund rate, interest rate and default or late, including mutual endorsement dummies (June 1, 2006 – July 31, 2008)

	Funded?	Contract interest rate	Default or late as of 8/1/2008
	Probit (marginal effects)	OLS	Probit (marginal effects)
in_a_grp_borrower	0.005*	-0.005*	0.011**
	(11.127)	(-10.448)	(1.981)
have endorsement + nobid by group leader	0.017*	-0.003*	0.001
	(8.180)	(-2.814)	(0.056)
have endorsement + bid by group leader	0.096*	-0.005*	0.004
	(21.607)	(-7.232)	(0.689)
have endorsement + nobid by friend	0.002*	0.001**	0.009
	(4.150)	(2.454)	(1.350)
have endorsement + bid by friend	0.048*	-0.007*	-0.039*
	(11.727)	(-5.884)	(-4.332)
mutual friend endorsement + nobid	0.004*	-0.001	0.004
	(3.453)	(-0.724)	(0.284)
mutual friend endorsement + bid	0.002	-0.001	-0.059**
	(0.551)	(-0.286)	(-2.080)
Year-week FE	Yes	Yes	Yes
N	293,802	25,008	23,344
Adjusted R2	0.375	0.855	0.270

The sample includes all the listings and loans between June 1, 2006 and July 31, 2008. T-statistics are in parenthesis. * p<0.01, ** p<0.05, *** p<0.1. All regressions control for state dummies, year-week FE, macro variables, variables for prosper environment, duration of auction, and indicators for missing debt-to-income ratio and other credit attributes as in Table 5. All regressions do not include the new credit variables added after Feb. 12, 2007. "Mutual friend endorsement + no bid" includes the cases where (1) neither bids on each other or (2) one bids on the other but not vice versa.

Table 10: Summary of IRR

	Flat projection: assume cumulative risks remain constant after month 24			Linear projection: assume cumulative risks grow linearly after month 24		
	IRR1	**IRR2**	**IRR3**	**IRR4**	**IRR5**	**IRR6**
performance measure	default	miss payment	default or late	default	miss payment	default or late
% loans with converged IRR	99.70%	99.65%	99.61%	99.68%	99.64%	99.63%
Per loan						
min	-99.98%	-96.88%	-98.05%	-99.98%	-95.42%	-99.23%
max	39.83%	39.83%	38.87%	39.83%	39.83%	38.84%
mean	3.23%	-0.48%	-1.22%	-0.22%	-1.24%	-1.53%
std. dev.	12.00%	12.43%	12.57%	13.72%	12.81%	12.69%
Per lender						
mean	3.36%	-0.17%	-0.64%	0.24%	-0.89%	-0.94%
std. dev.	6.96%	6.92%	7.08%	7.85%	7.16%	7.12%
Per dollar						
mean	3.19%	-0.62%	-1.24%	-0.06%	-1.38%	-1.56%
std. dev.	12.54%	12.35%	12.30%	14.00%	12.75%	12.43%

Note: The reported IRR is compounded monthly into the annual percentage rate.

Table 11: Lender Responses to Ever Late Loans

			Conditional on Funding a Loan in Week t					
			Amount Funded	% of Investment in:				
	Funded a Loan			AA to A	B to D	E to HR	IRR2	
	coef/t	coef/t	coef/t	coef/t	coef/t	coef/t	coef/t	coef/t
% of Portfolio Ever Late	-0.078*		-795.256*				0.217*	
	(-16.401)		(-10.999)				(44.288)	
% of NC Loans Ever Late		-0.116*		0.062*	0.046*	-0.078*		0.020*
		(-13.269)		(5.655)	(3.487)	(-8.662)		(8.107)
% of E to HR Loans Ever Late		-0.080*		0.159*	0.052*	-0.211*		0.041*
		(-21.487)		(17.567)	(5.380)	(-30.943)		(20.345)
% of B to D Loans Ever Late		-0.105*		0.475*	-0.416*	-0.061*		0.094*
		(-24.743)		(24.223)	(-21.191)	(-6.273)		(20.777)
% of A to AA Loans Ever Late		-0.126*		-0.146*	0.087*	0.060*		0.023*
		(-22.661)		(-10.693)	(5.541)	(5.436)		(5.911)
N	2,564,481	2,564,481	553,117	553,117	553,117	553,117	549,837	549,837

T-statistics are in parenthesis. * $p<0.01$, ** $p<0.05$, *** $p<0.1$. Column 1 is a linear probability model and all other columns are OLS regressions. Standard errors are clustered at the lender level.

Table 12: Lender response to ever late social loans

A: By Endorsement Status (Conditional on Listing After Feb. 12, 2007)

	% Endorsed No Bid	% Endorsed w/ Bid	% Not Endorsed
% of Endorsed No Bid Loans Ever Late	-0.053*	0.013**	0.039*
	(-6.242)	(2.553)	(4.272)
% of Endorsed w/ Bid Loans Ever Late	0.011	-0.029*	0.018**
	(1.478)	(-6.488)	(2.199)
% Not Endorsed Loans Ever Late	-0.078*	0.048*	0.030**
	(-5.708)	(5.971)	(2.004)
N	471,470	471,470	471,470

B: By Group Status

	% In Group	% Not in Group
% of Group Loans Ever Late	-0.136*	0.136*
	(-11.397)	(11.397)
% of Non-Group Loans Ever Late	0.245*	-0.245*
	(18.391)	(-18.391)
N	553,117	553,117

C: By Own Group Status (Conditional on Lender Being a Group Member)

	% In Own Group	% In Other Group	% In No Group
% of Own Group Loans Ever Late	-0.119*	0.041	0.079*
	(-8.373)	(1.609)	(3.369)
% of Other Group Loans Ever Late	0.036**	-0.105*	0.069**
	(2.325)	(-3.532)	(2.371)
% of No Group Loans Ever Late	0.069*	0.190*	-0.258*
	(4.872)	(6.349)	(-8.582)
N	82,387	82,387	82,387

T-statistics are in parenthesis. * p<0.01, ** p<0.05, *** p<0.1. Standard errors are clustered at the lender level.

Table 13: Heterogeneous learning

A: By Time

	Funded A Loan coef/t	Amount Funded coef/t	IRR2 coef/t
% of Portfolio Ever Late	-0.054*	-750.121*	0.228*
	(-10.060)	(-6.978)	(29.921)
X 2007 Half 1	-0.109*	-231.772**	-0.041*
	(-15.510)	(-2.076)	(-3.464)
X 2007 Half 2	-0.043*	-64.349	-0.002
	(-9.260)	(-0.739)	(-0.254)
X 2008 Half 1	-0.153*	-9.113	-0.011***
	(-4.930)	(-0.148)	(-1.690)
N	2,564,481	553,117	549,837

B: By Lender Cohort

	Funded A Loan coef/t	Amount Funded coef/t	IRR2 coef/t
% of Portfolio Ever Late	-0.167*	-1,161.596*	0.243*
	(-9.340)	(-4.102)	(18.997)
X 2006 Half 2	0.073*	384.006	-0.018
	(3.890)	(1.403)	(-1.366)
X 2007 Half 1	0.108*	453.434***	-0.034**
	(5.740)	(1.659)	(-2.411)
X 2007 Half 2	0.142*	572.299**	-0.056*
	(7.070)	(1.987)	(-2.935)
X 2008 Half 1	0.136*	331.377	-0.218*
	(4.750)	(0.969)	(-5.915)
	2,564,481	553,117	549,837

T-statistics are in parenthesis. * p<0.01, ** p<0.05, *** p<0.1. The Funded a Loan columns are linear probability model and all other columns are OLS regressions. The Amount Funded and IRR2 regression samples are conditional on funding a loan in week t. Standard errors are clustered at the lender level.

Table 14: Detection of lender heterogeneity in initial sophistication

	Avg IRR2
Dummy of Above-mean IRR in first-month portfolio	0.073*
	(15.741)
Dummy of above-mean * lender cohort	-0.004*
	(-13.036)
Dummy of above-mean * lender age	-0.002*
	(-7.851)
Dummy of above-mean * lender cohort * lender age	0.0001*
	(3.789)
N	844
Adjusted R2	0.736

Unit of observation is defined by the combination of lender cohort (by month), lender age (by month), and whether a lender's first month portfolio has an average IRR2 above the market-wide average of the corresponding month. The sample includes all the lenders that started on Prosper since June 1, 2006. Cohort is defined as the count of quarters from June 2006. Lender age is measured in month since the first day of investment on Prosper. Regression controls for cohort fixed effects and lender age fixed effects. T-stat in parentheses. *p<0.01, **p<0.05, ***p<0.1.

Table 15: Regressions of IRR2 on Social Variables

	IRR2 (1)	IRR2 (2)
Borrower in a group	-0.042*	-0.039*
	(-19.145)	(-7.515)
Borrower in a group * no group leader reward		0.013*
		(3.289)
Borrower in a group * group rating adopted		-0.010
		(-1.629)
Have endorsement no bid from group leader	0.041*	0.042*
	(8.041)	(8.040)
Have endorsement + bid from group leader	0.003	0.005
	(1.106)	(1.452)
Have endorsement no bid from friends	-0.032*	-0.032*
	(-14.304)	(-14.395)
Have endorsement + bid from friends	0.045*	0.045*
	(12.531)	(12.404)
N	24,815	24,815
Adjusted R2	0.0670	0.0789

The sample includes all the funded loans between June 1, 2006 and July 31, 2008 that yield a valid calculation of IRR2. All regressions have year-week fixed effects but do not control for other listing attributes. The calculation of IRR2 has already incorporated these attributes. T-stat in parentheses. *p<0.01, **p<0.05, ***p<0.1.

Figure 1: Macro Economic Indicators (2005 – June 2008)

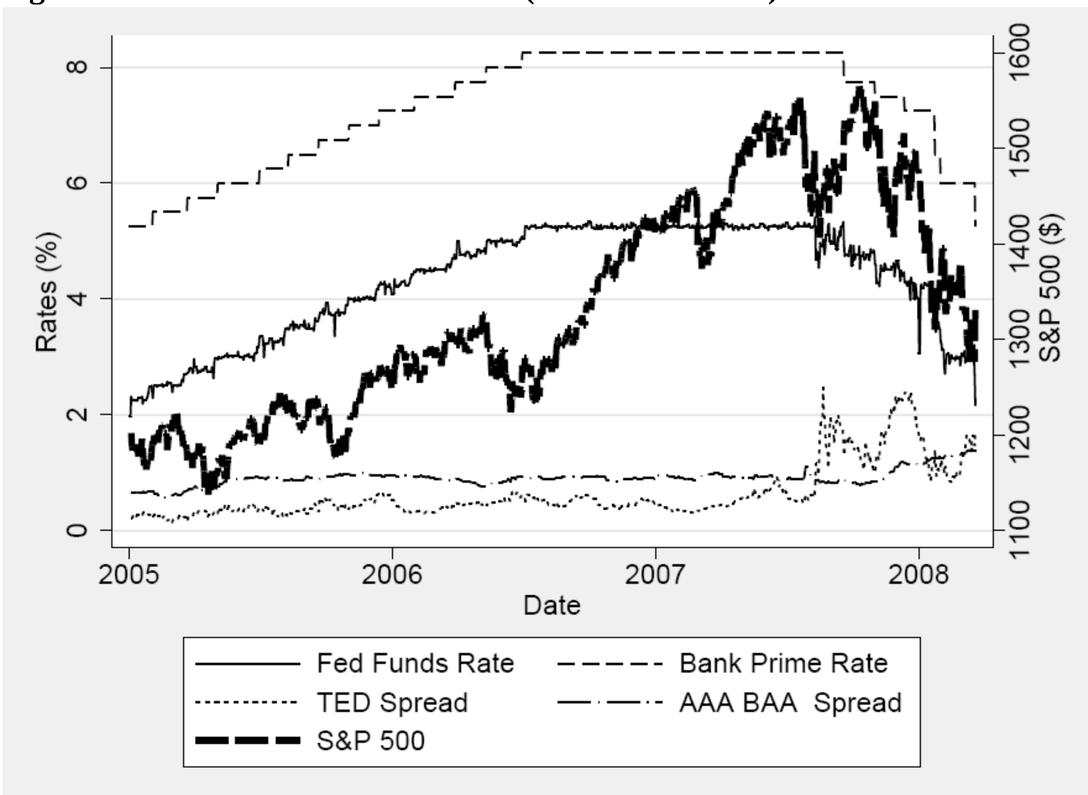


Figure 2: CDF of Prosper and Experian Listings

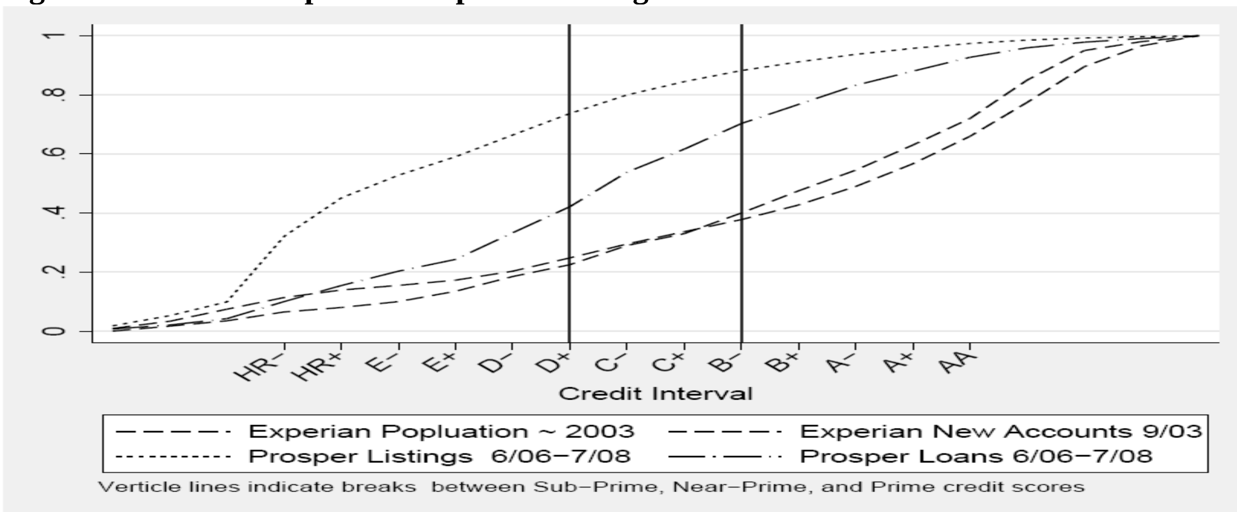


Figure 3: PDF of Prosper Listings by Time

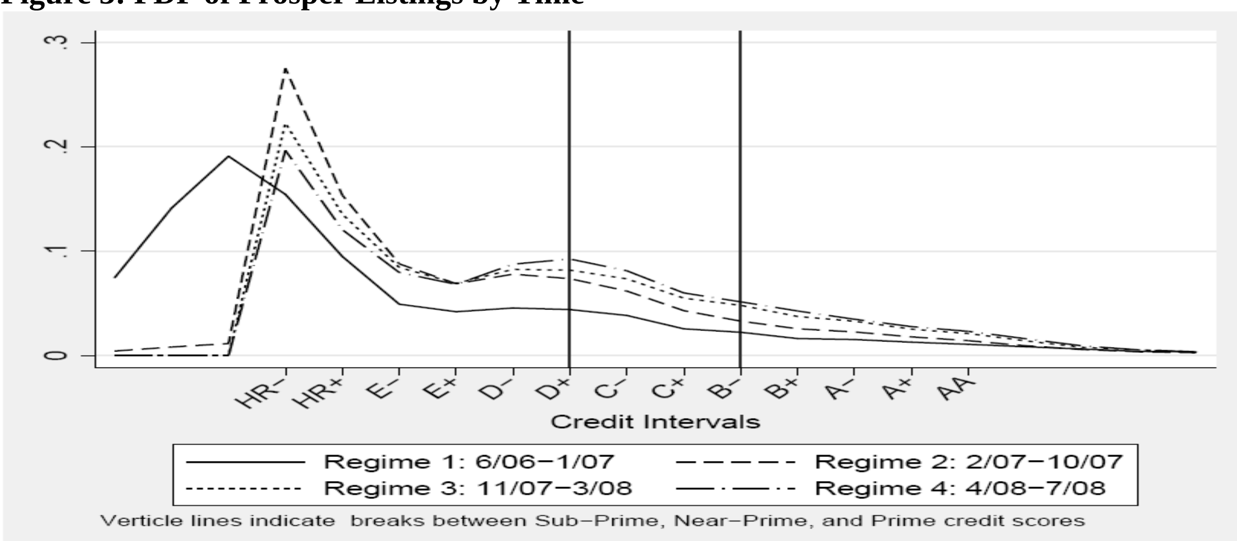


Figure 4: PDF of Prosper and Experian Loans by Time

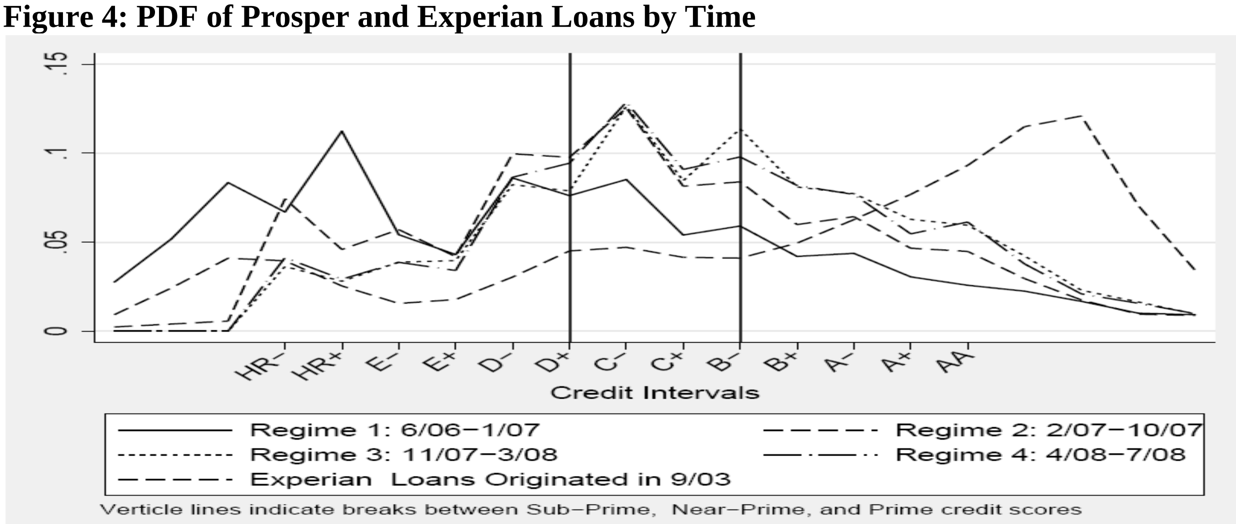


Figure 5: Observed Cumulative Performance by Loan Age

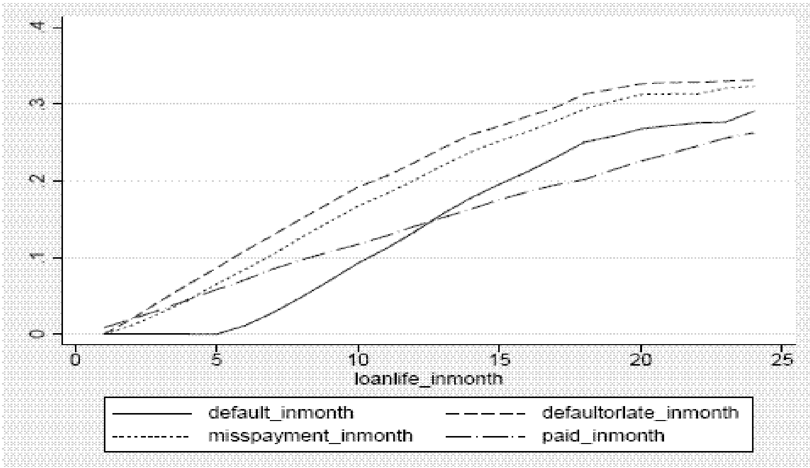


Figure 6: Actual, predicted and unpredicted default or late

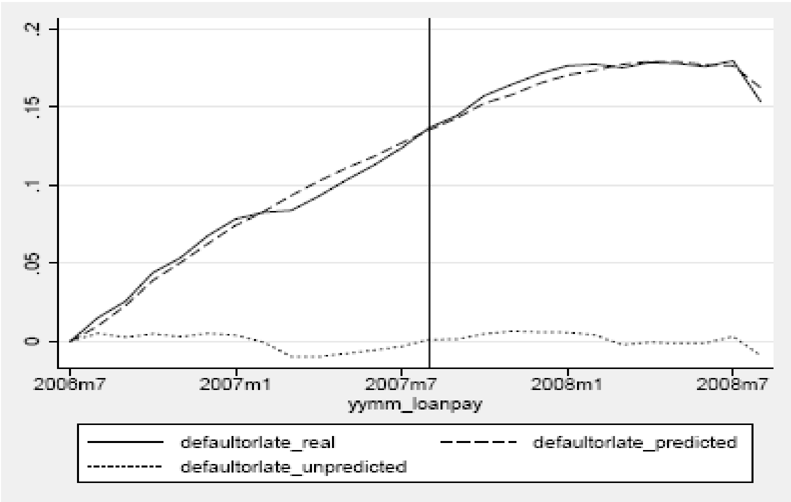


Figure 7: Density of IRR2 by Credit Grade

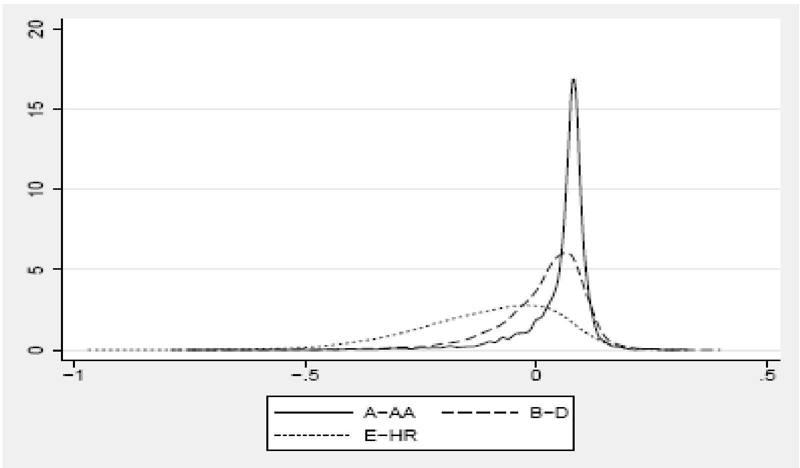
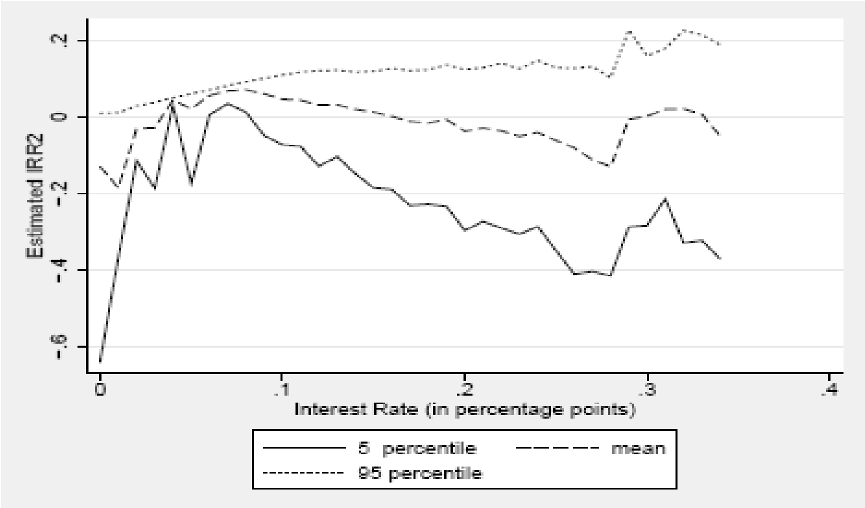


Figure 8: IRR2 by Contract Interest Rate



Note: Vertical line corresponds to the start of the subprime crisis (8/07).

Figure 9: Density of IRR2 by borrower's group affilation

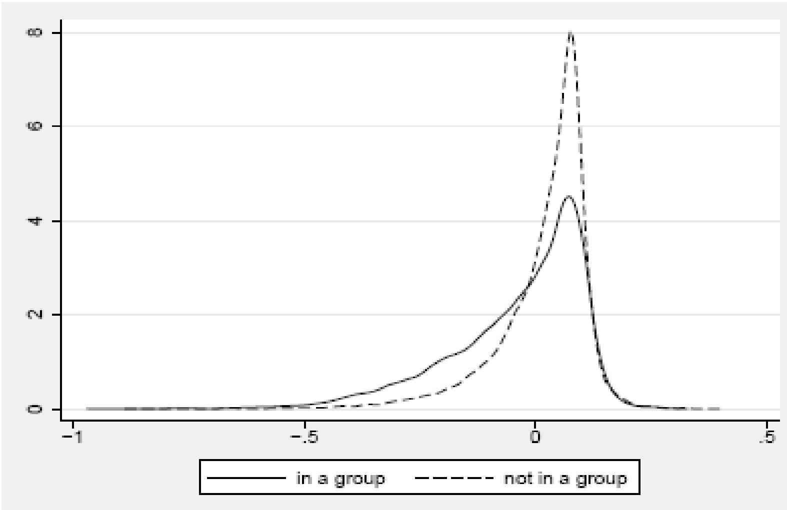


Figure 10: Density of IRR2 by friend endorsement

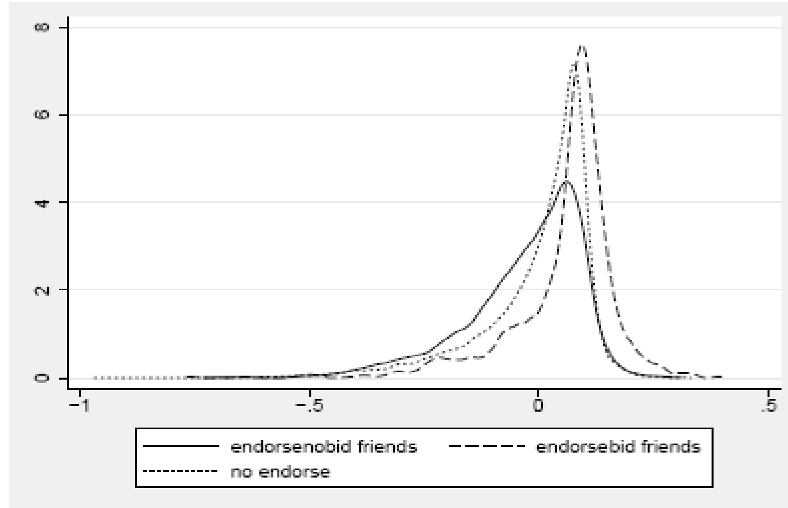
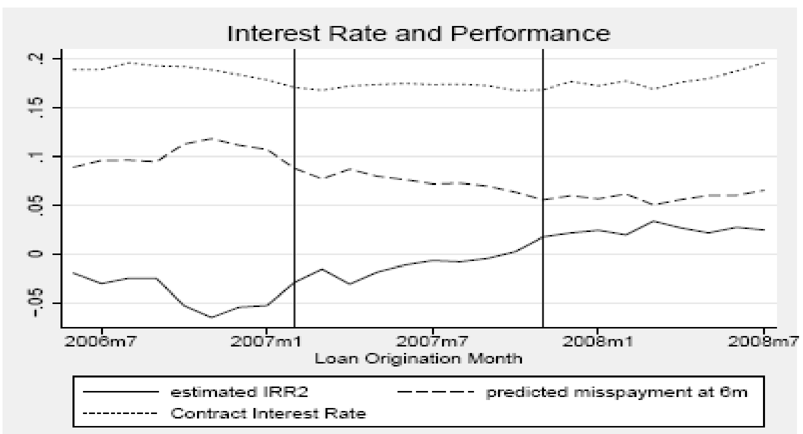


Figure 11: IRR2, contract rate and performance by time



Vertical lines indicate Prosper's Feb. 12, 2007 policy of redefining E and HR plus posting more credit information and Oct. 30, 2007 introduction of bidder guidance.

Figure 12: IRR2 distribution by loan origination time

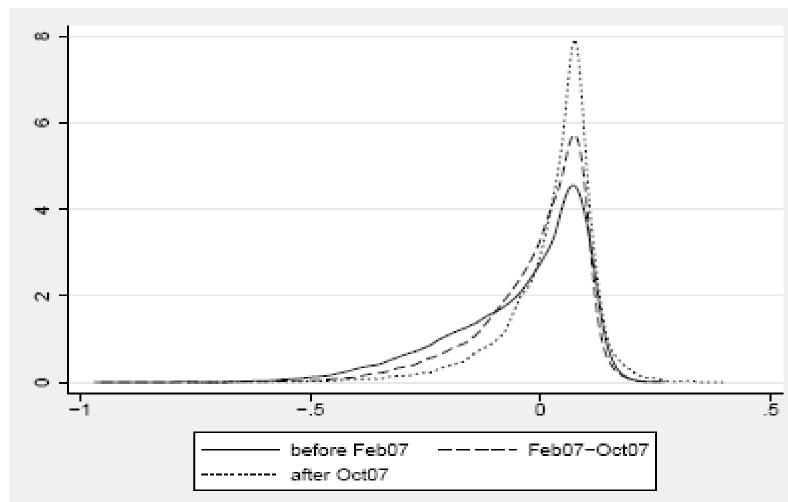


Figure 13: Average IRR2 of new investments by lender age (in week) and whether the average IRR2 of the lender's first month investment is above or below the market-wide average IRR2 for that month

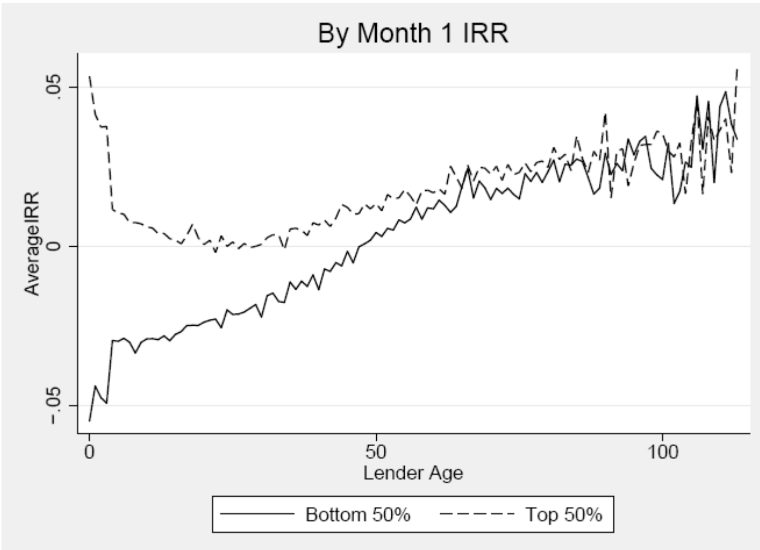


Figure 14: Average IRR2 of New Investments by Investment Date and Lender Cohort