Unobserved Risk Type and Sorting:
Signaling Game in Online Credit Markets[*]

Kei Kawai[†]
New York University

Ken Onishi[‡]
Northwestern University

Kosuke Uetake[§]
Northwestern University

March 12, 2012
VERY PRELIMINARY and INCOMPLETE — PLEASE DO NOT CIRCULATE

Abstract

This paper studies how signaling can facilitate the functioning of a market with classical adverse selection problems. Using data from Prosper.com, an online credit market where loans are funded through auctions, we provide evidence that reserve interest rates that potential borrowers post work as a signaling device. We then develop and estimate a structural model of borrowers and lenders where low reserve interest rate can credibly signal low default risk. Announcing high reserve interest rate increases the probability of trade at the cost of higher expected interest payment conditional on trade. Borrowers regard this trade-off differentially, which results in a separating equilibrium. In our counterfactual, we estimate the credit supply curve when signaling is not available.

1   Introduction

Inefficiencies arising from adverse selection figure importantly in many markets. Examples of such markets range from used car markets (Akerlof, 1970) to health insurance markets (Rothchild and Stiglitz, 1976). A key source of inefficiency in these markets is the distortion created by the uninformed party to reduce information rents accrued to the party with more information. To mitigate such information asymmetry and restore market inefficiency, there exist several mechanisms that allow people who have less information to distinguish "good" type from "bad" type. A key insight of Spence (1973), for example, is that markets with appropriate signaling devices make it possible to overcome adverse selection problems.

In this paper, we study how signaling can help the functioning of a credit market with classical adverse selection problems (e.g. Stiglitz and Weiss, 1981) using the data from an online peer-to-peer lending market, Prosper.com. In contrast to institutions in the traditional credit market such as commercial banks or credit unions, Prosper.com is a

---

[†]New York University, Stern School of Business: kkawai@stern.nyu.edu
[‡]Northwesern University: kenonishi2013@u.northwestern.edu
[§]Northwesern University: uetake@u.northwestern.edu

marketplace where potential borrowers are directly matched to potential lenders through auctions as described below. Although it is possible that the market suffers adverse selection problems, we find evidence that the reservation interest that borrowers can post work as a signaling device.

In the paper, we present descriptive results which show that (1) posting a higher reservation interest makes it more likely that the loan will be funded, (2) posting a higher reserve interest tends to raise the contract interest rate and (3) the probability that the borrower defaults is higher among those who submit higher reserve interest rates *conditional on the same contract interest rate* and other observable characteristics. The third finding suggests that the reserve interest rates are correlated with the unobserved "credit worthiness" of the borrower. The second finding suggests that lenders factor this into their lending decision. Finally, our first finding and our second finding together suggest that there is a trade-off that borrowers face when posting the reserve interest rate, i.e. the trade-off between the probabilities of being funded versus the actual interest rate they face. This is consistent with the idea that low risk types value more the likelihood of being funded relative to a higher interest on the loan, perhaps because they have access to credit through local banks, while high risk types do not have other sources of credit and hence view the trade-off differently. To the extent that borrowers differentially value the trade-off, there will be incentives for different types of borrowers to post different reservation interest rates – and hence signal their type.

These descriptive findings motivate us to develop and estimate a structural model of the online credit market with informational asymmetry between the lenders and borrowers. Our structural model involves three parts. First, potential borrowers post a listing on the website of Prosper.com. In each listing, a borrower specifies an amount he wants to borrow and a reserve interest rate, which is the maximum interest rate he is willing to borrow. A key feature of our model is that borrowers are heterogeneous with respect to their credit worthiness. As borrowers with different credit worthiness value the trade-off from increased likelihood of being funded relative to a higher interest rate on the loan, separation of the types can occur in equilibrium.

Second, after observing the listing characteristics including the reserve interest rate, requested amount, credit grade etc, lenders who are also heterogeneous in terms of their attitude toward risks decide their bidding strategy. A bid consists of the pair of interest rate and the amount she is willing to lend. The contract interest rate is then determined through a uniform price ascending auction, and we provide some characterizations of bidding strategy on which our estimation is based.

Finally, the third part of the model concerns borrowers' repayment behavior conditional on being funded. We model the repayment decision as a finite horizon single agent dynamic discrete choice problem. At each period, the borrower chooses whether to pay back the loan or not, depending on whether the disutility from paying back outweighs the disutility from default. This disutility tends to be lower for "good" borrowers because they are more capable of obtaining a fund from outside financial institutions.

Exploiting rich variation of listing characteristics, lenders' bidding behavior, and borrowers' repayment decisions in the data, we can identify the distribution of borrowers' unobservable heterogeneity and that of lenders' risk attitude on top of the structural parameters of borrowers' preference over the loans they can obtain. We find evidence of heterogeneity

across lenders and borrowers respectively. The distribution of lenders' risk parameter varies significantly by the credit grade. We also find that the distribution of borrowers' type is heterogeneous with regard to the credit grade, requested amount, and the contract interest rate.

Our final goal is implementing a counterfactual policy experiment to estimate the effect of signaling device on the market credit supply function. As Stiglitz and Weiss (1981) pointed out, the credit supply curve in the market with adverse selection problems may not be monotonically increasing in interest rates, and sometimes it becomes backward bending. We examine this hypothesis by prohibiting the use of reserve interest rates. Using the estimates we obtain, we re-compute lenders' and borrowers' behavior, and simulate a credit supply curve in the case where no reserve interest rate is allowed. The result supports Stiglitz and Weiss's prediction: Credit supply curve becomes more backward bending if posting the reserve interest rate is not allowed.

The plan of the paper is as follows. In the next subsection, we review several related literature. Section 2 describes market background and data we use in the estimation. In Section 3, we show descriptive evidence of signaling in our data. We then develop our structual model of the borrower and lenders in Section 4. Section 5 present estimation results and Section 6 demonstrates the result of counterfactual policy experiment. Section 9 concludes.

1.1   Related Literature

Our paper is related to several strands of literature. First, our study is related to papers structurally estimating a model with adverse selection and/or moral hazard problems. Recent papers in this literature also concern the identification of people's attitude toward risks on top of the distribution of private information. For example, Cohen and Einav (2007) examine the identification of loan applicants' joint distribution of risk type and risk attitude using the data of individual-level deductible choices in the car insurance market in Israel. Einav *et al.* (2011) consider how the supply side pricing and contract design decisions affect the consumers' behavior who are heterogeneous regarding their ability of repayment and risk attitude in a used car sale and subprime loan market. Jenkins (2009) studies moral hazard problems of repayment behavior using the same data of Einav *et al.* (2011). We study both demand and supply side of an online credit market with potential adverse selection and moral hazard problems, but attempt to shed new light on this literature by analyzing how the existence of signaling device mitigates the market inefficiencies.

Second, to the best of our knowledge, the empirical literature on signaling and screening is still scarce. Some exceptions include papers by Kim (2010), Gayle and Golan (2011) and Aryal *et al.* (2009). Kim (2010) considers a signaling game with two players and provides an identification strategy of type distribution under some equilibrium selection mechanism through an equilibrium refinement by Cho and Kreps (1987). Gayle and Golan (2011) formulate a dynamic general equilibrium labor supply model with endogenous gender wage discrimination. In their model, employers with different participation cost use their labor supply decisions as a signaling device to potential employees. Aryal *et al.* (2009) study the identification of a screening model in which potential insurees possess multidimensional private information in their risk attitude and risk type. Our paper complements these

3

papers by providing a step to extending the literature of empirical signaling games.

Third, there are few papers empirically investigating the effect of adverse selection on the credit supply function, although the theoretical literature on the adverse selection in the financial credit market is vast. A few exceptions include Berger and Udell (1992) and Adams *et al.* (2009). Using the micro level contract-term data in the commercial banking industry, Berger and Udell (1992) test a implication of the credit rationing model, which is that the commercial bank loan rate is sticky, or it does not fully respond to changes in Treasury rate. Recently, Adams *et al.* (2009) investigate the effect of asymmetric information in the subprice market. They find that consumers in their sample face liquidity constraints, and lenders face a informational problem.

There are also a few theoretical papers on signaling in auctions. Cai *et al.* (2007) consider the model of a second-price auction in which the seller who has private information about the quality of a good can transmit her private information by announcing a reserve price. They provide a characterization of the unique separating equilibrium and show that the lowest quality seller cannot earn informational rent. Jullien and Mariotti (2006) also consider a similar second-price auction model and compare the unique separating equilibrium outcome with the optimal mechanism for a monopoly broker who buys from the seller and sells to the buyers. Our paper adds to this literature by empirically studying the effect of signaling in auctions.

Finally, there is a growing body of papers that study online peer-to-peer lending markets such as Prosper.com and Zopa.com. Freedman and Jin (2010) use the data of credit scores which are unobservable to potential lenders to access the existence of adverse selection. They find potential borrowers with lower credit scores are more likely to stay in the market even conditional on a credit grade that Prosper.com determines. Rigbi (2008) identifies and estimates the effect of usury laws on the market outcome using the fact that Prosper.com had different state-level legislated caps on interest rates prior to Apr. 2008, and then set 36% cap uniformly after that. He finds that funding probability was increased by the increase of interest caps while the default probability did not increase. We construct and estimate an equilibrium model of an online P2P lending market, which provides deeper understanding of this market.

2 Institutional Background and Data

2.1 Institutional Background

Propser.com is an online peer-to-peer lending website that matches borrowers with lenders in addition to providing loan administrative service for the lenders. Established in 2006, it has become America's largest peer to peer lending marketplace with more than a million members and over $280 million in loans. In this section, we describe how Prosper works with particular emphasis on the auction mechanism which was in effect until December 19, 2010.[1] For details on other aspects of Prosper, see Freedman and Jin (2010).

---

[1]Prosper no longer uses auctions: Instead, each listing has a "posted price", or Prosper determined pre-set rates, which are based on the borrowers' credit risk. During our sample period, the terms of the loan and the match between borrowers and lenders were determined through auctions.

The allocation (which of the potential lenders end up lending and how much money they lend), the interest on the loan, and the subsequent repayment process occur according to the following timeline:

1. A borrower posts a listing

2. Lenders bid

3. Funding decision is made

4. If the borrower receives a loan in step 3, the borrower makes loan repayments

The first three steps of the timeline relate to how the borrowers and the lenders are matched as well as how the terms of the loan are determined. The last step concerns the repayment process conditional on the loan being funded. We explain each step in turn.

Borrower posts a listing A potential borrower creates an account on Prosper by providing the social security number, the driver's license number and the home address. Prosper then pulls the applicant's credit history from Experian, a third party credit scoring agency. If the applicant's credit score exceeds the minimum threshold, then a listing is created which contains information regarding the borrower's characteristics, the funding option (either "close when funded" or "open for duration"), the amount of loan requested, and the maximum interest rate (hereafter, *reserve interest rate*) he is willing to pay.[2] There is no fee for posting a listing.[3] The characteristics of the borrower that appear in the listing include the credit grade, home ownership status, debt to income ratio, purpose of the loan, as well as any other additional information that the borrower wishes to post. The credit grade (AA, A, B, C, D, E, and HR) corresponds to 7 distinct credit score bins and this information is verfied by Prosper and included in the listing automatically.[4] Home ownership status is another characteristic that is acertained by Prosper. Other information such as debt to income ratio and purpose of the loan are provided by the borrower without verification by Prosper. Finally, a key part of the listing that is important for our analysis is the reserve interest rate. The reserve interest rate is the maximum interest rate the borrower is willing to pay on the loan, and it plays a similar role to the reserve price in regular auctions. The requested loan amount and the reserve interest rate are both variables that the borrower chooses, subject to Prosper's conditions and State usury laws.[5]

---

[2]The listing remains active for 14 days. For "closed when funded (hereafter, *closed*)" listings, the listing becomes inactive after 14 days or after the listing attracts enough lenders to fund the whole loan, whichever occurs sooner. For "open for duration (hereafter *open*)" listings, the listing remains active after the requested amount is fully funded. Since less than 1/4th of the listings are closed listings and we throw away closed listings in our sample, and we will explain how the contract interest rate of open listings is determined in the following section.

[3]Prosper charges fees to both borrowers and lenders only if the loan originates. See Freedman and Jin (2010) for details.

[4]A credit grade of AA corresponds to a credit score of 760+, a grade of A corresponds to 720-759, B to 680-719, C to 640-679, D to 600-639, E to 540-599, and HR to 540-. The actual credit score is not listed.

[5]The minimum loan amount was $1,000 and the maximum amount was $25,000. As regards the interest rate, it was capped by the usury law of the State in which the borrower resided, before April 15, 2008. After April 15, 2008, the interest rate cap was uniformly set at 36% across all States. See Rigbi (2011) for more information.

Lenders Bid Prosper maintains a list of active listings on its website for potential lenders. Each listing contains information we described above, such as the credit grade of the borrower, etc. as well as the "active interest rate" and the "fraction funded", which we will explain later. If a lender finds a listing to which she wishes to lend money, she can then submit a bid on the listing, similar to a proxy bid in online auctions. The bid consists of an amount that the lender is willing to lend (typically a small fraction of the loan amount that the borrower requests), and the minimum interest rate that the lender is willing to accept. The lender can submit a bid with an amount anywhere between \$50 and the borrower's requested amount but the modal bid amount is \$50. The bidding is similar to other online auctions such as eBay auctions in the sense that the lender can bid on any active listing at any time. At the time the lender bids, the lender observes the "active interest rate" of the listings for which the total amount of outstanding submitted bids exceeds the requested loan amount. The "active interest rate" is similar to the lowest/highest bid in usual auctions. We will discuss how the "active interest rate" is determined below. Otherwise, the lender observes the "fraction funded", which is the ratio of the total amount of submitted bids to the requested loan amount.

Funding Decision The auction used in Prosper is a combination of an ascending auction and an uniform price auction. We explain how the allocation and the terms of the loan are determined using an example. Suppose a borrower creates a listing with a requested amount of \$10,000 and a reserve interest rate of 25%. For the purpose of this example, let us fix the the bid amount to be \$50. Potential lenders who are interested in lending money will bid on the listing. At the the time the lender submits her bid, she observes the bid amount for all of the submitted bids (i.e., \$50 for each bid). However, for listings that are not fully funded (i.e. less than 200 bids in this example, see left panel of Figure 1), she does not observe the interest rate of each bid. As for listings that have already been fully funded, (i.e. more than 200 bids, see right panel of Figure 1) the lender observes the active interest rate, i.e. the interest rate of the marginal bid that brings the supply of money over the requested amount. In our example, this is the interest of the 200th bid if we ordered the submitted bids according to its interest rate from the lowest to the highest. Moreover, for fully funded listings, the bidder also observes the interest rate of the losing bids, i.e. the interest rate of bids of the 201st bid, 202nd bid and so on. The bidder does not observe the interest rate of the bids below the marginal bid, however.

At the end of the bid submission period, the loan is made only to listings that are fully funded. There are no partial loans for listings that have failed to attract enough lenders to fund the total requested amount. In the first panel of Figure 1, a loan would not be made even though \$8,000 out of \$10,000 have been funded. For fully funded listings, the loan is made at an interest equal to the interest rate of the marginal bid, and the same interest rate applies to all the lenders. In the second panel of Figure 1, the loan is made at 24.8% and the same rate applies to all the lenders. In this sense, the auction is similar to uniform price auctions.

Loan Repayments The loans originated by Prosper are unsecured and the length of the loan is 36 months: The borrower pays both the principal and the interest in equal
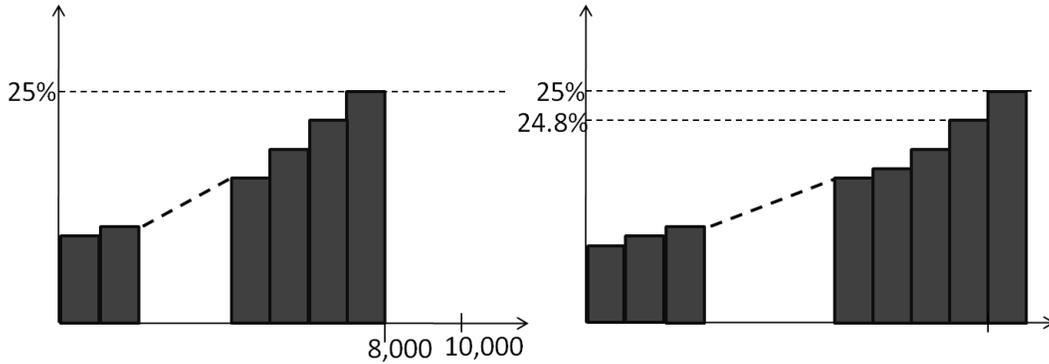
Figure 1: Allocation of Loans

installments over the 36 month period.[6] If a borrower's monthly payment is more than 15 days late, a late fee is charged in addition to the principal and the interest. If a borrower defaults, the defualt is reported to the credit beareau, and a third party collection agency is hired by Prosper to retrieve any money from the borrower. From the perspective of the borrower, defaulting on a loan originated by Prosper is just like defaulting on any other loan, resulting in a damaged credit history.

2.2 Data

The data for our analysis comes directly from Prosper.com, which makes the data available for researchers through their website. The dataset is unique in the sense that almost all the information that is relevant to potential lenders are available to the researcher: We have data on the borrower's credit grade, debt to income ratio, home ownership etc. as well as additional text information that borrowers provide to lenders and conversation that took place between borrowers and lenders through the Prosper website.[7]

We retrieved the data from the website of Prosper.com on January, 2011. Our data include all listings from May 2008 to December of 2008.[8] Note that all loans in our sample have been either matured or defaulted. We then drop observations that were either

---

[6]There is no penalty for early payment: early repayments go directly into paying off the principal.

[7]The one piece of information which may be relevant that lenders observe that we do not, are pictures that borrowers post.

[8]We use the data of this period because of the following reasons. First, the borrowers were subject to the state-level usury laws before April 2008. After April 2008, Prosper removed the state level restrictions, and set 36% maximum interest rate for all states (except a few states). We drop the observations before April 2008 to avoid the effect of state-level regulation on the borrowers and bidders' behavior. Second, Prosper entered into a settlement with state securities regulators over sales of unregistered securities at December 1st, 2008. Due to this, Prosper was shut down until July 2009. Hence, we have no observations during December 2008 and June 2009. Third, Prosper set the minimum bid amount as $25 after its relaunch, and changed the definition of the credit grade. We drop the observations after July 2009 to avoid the effect of such changes.

Grade	Amount Requested		Reserve Rate		Debt/ Income		Home Owner		Bid Count		Fund Pr.	Obs.
	mean	sd	mean	sd	mean	sd	mean	sd	mean	sd		
AA	12,845.2	8,311.6	0.132	0.057	0.349	0.923	0.813	0.39	86.7	91.5	0.546	1,461
A	12,126.3	7,816.7	0.167	0.066	0.366	0.671	0.618	0.486	62.5	74.0	0.418	1,851
B	10,460.6	6,032.6	0.215	0.073	0.374	0.632	0.593	0.491	50.4	61.5	0.349	2,951
C	7,293.2	4,866.9	0.250	0.076	0.344	0.528	0.554	0.497	27.8	38.6	0.269	4,817
D	6,032.0	4,309.7	0.292	0.072	0.349	0.568	0.379	0.485	15.0	25.7	0.175	5,852
E	4,397.8	4,398.5	0.315	0.071	0.314	0.505	0.325	0.487	4.7	7.5	0.103	3,795
HR	4,131.6	4,359.3	0.318	0.071	0.273	0.608	0.218	0.413	2.8	4.5	0.048	7,160
All	6,769.6	5,917.2	0.269	0.090	0.328	0.603	0.422	0.494	23.3	46.6	0.203	27,877

Table 1: Descriptive Statistics – Listings

withdrawn by the borrower, cancelled by Prosper, or any observations for which parts of the data were missing. We also dropped any observations that the borrowers chose closed options. We are left with a total of 27,877 listings of which 5,648 were funded. Below, we report some summary statistics.

Listings   Table 1 reports sample statistics of the listings by credit grade. The mean of the requested amount is reported in the first colum of the Table, and it ranges from a high of more than $12,000 dollars for AA listings to a low of less than $5,000 for HR listings. The average among the whole sample is $6,769. Not surprisingly, there is also a monotonic relationship between the reserve interest rate and the credit grade, reported in the second column, with higher credit grade borrowers posting lower reserve rates. The third and fourth column report the debt to income ratio and home ownership of the borrower. Note that information regarding home ownership is verified by Prosper.com, while the debt to income ratio is self-reported by each potential borrower. The debt to income ratio, reported in colum 3 is the only variable in the Table that does not seem to be related to the credit grade in an obvious way. The nonmonotnic relationship bewteen this variable and the credit grade may be partly due to the fact that people with a very low credit grade have trouble borrowing money. On the other hand, there is a monotonic relationship between home ownership and the credit grade: Borrowers with a higher credit grade are more likely to own their house. The bid count is the number of average bids submitted to a listing and this is reported in column 5. We note that the bid count and the funding probability, reported in column 6 are also related to the credit grade in a monotonic way.

Bid   In Figure 2, we report the distributions of the bid amount, again by credit grade.[9] The fraction of bidders who post $50 to a listing is more than 70% in all credit grades, and that of bidders who post $100 is more than 10% in all credit grades. Hence, more than 80% of bidders choose either $50 or $100 of bidding. We also find that a small fraction of

---

[9] The sample of bids used to create table 2 only consists of the final bids for each bidder, i.e. if a potential lender bids more than once in a listing, only the last bid is used.
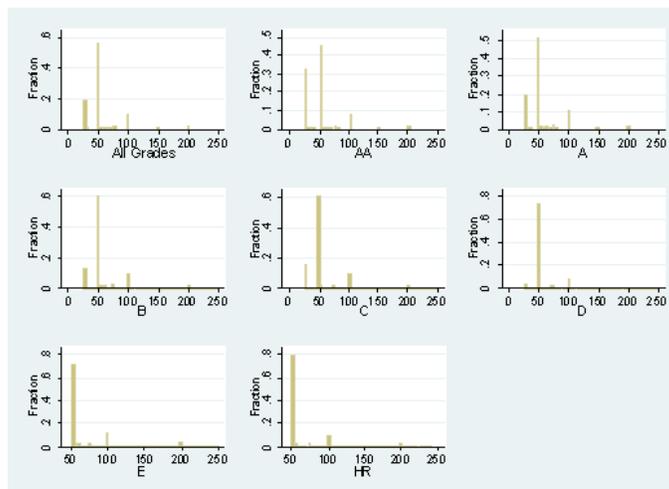
Figure 2: Distribution of Bids

bidders make \$150 or \$200, and that bidders rarely bid more than \$250. These facts help us model potential lenders' optimal bidding strategy in that most of them choose the bid amount from { \$50, \$100 and \$200} rather than from continuous set.

Identity of the bidder is observed in the data. This fact Not yet written.

Loans (funded listings)   Table 2 reports sample statistics of listings that were funded. The second column reports the average loan amount by credit grade. The loan amount is highest for AA listings with an average of \$9,537 and lowest for HR listings with an average of \$1,689. The mean for the total sample is \$5,821. Note that the mean loan amount reported in this Table is smaller than the mean requested amount shown in Table 1, indicating that loan amount for funded loans tend to be smaller than those for unfunded ones. In the third column, we report the contract interest rate determined by the interest rate of the marginal bid, as we described earlier. As expected, the contract interest rate is lowest for AA loans and highest for HR loans. The debt to income ratio and home ownership status are reported in the fourth and fifth column respectively. There is no clear relationship between the credit grade and debt to income ratio, while there is clearly a monotonic relationship between the credit grade and home ownership status. Column 5 reports the bid count for the funded listings. The average bid count for AA loans is about 130, and it becomes smaller as the credit grade is worse. The total sample average is 80. Comparing with the average bid count for all listings in Table 1, funded listings obviously attract more potential lenders. Column 6 summarizes the default rate of all loans. Note that all loans in our sample are either matured or defaulted. The default rate is defined as *the number of defaulted loans divided by the number of loans originated during May 2008 to December 2008.* The average default rate is lowest for AA loans at 14.7%, while it is highest for HR loans at 43.7%.

9

Grade	Loan Amount		Contract Rate		Debt/ Income		Home Owner		Bid Count		Default	Obs
	mean	sd	mean	sd	mean	sd	mean	sd	mean	sd		
AA	9,536.7	7,324.3	0.095	0.033	0.211	0.378	0.798	0.401	129.7	98.8	0.147	794
A	8,678.7	6,597.2	0.126	0.045	0.232	0.143	0.556	0.497	113.7	84.2	0.213	772
B	7,321.7	4,853.1	0.164	0.046	0.264	0.341	0.554	0.497	100.5	67.5	0.295	1,030
C	4,688.1	2,996.1	0.181	0.062	0.248	0.211	0.480	0.500	68.2	45.7	0.309	1,294
D	3,595.0	2,406.8	0.21	0.066	0.242	0.166	0.256	0.437	53.4	38.3	0.322	1,025
E	1,890.6	1,187.9	0.292	0.057	0.223	0.221	0.263	0.441	21.6	11.7	0.375	392
HR	1,689.1	1,285.2	0.299	0.058	0.196	0.440	0.173	0.379	17.6	10.4	0.437	341
All	5,821.8	5,276.0	0.178	0.079	0.238	0.273	0.475	0.499	80.0	71.7	0.286	5,648

Table 2: Descriptive Statistics – Loans

Grade	mean	sd	10%	25%	50%	75%	90%	Obs
AA	18.1	8	9	12	17	23	30	104
A	18.4	8.7	8	11	16	25	32	149
B	17.4	7.5	8	11	16	23	28	279
C	17.4	8.3	8	11	16	23	30	368
D	18.1	8.7	7	10	17	25	31	311
E	18.9	9.4	7	10	17	27	34	141
HR	15	8	7	9	13	20	27	145
All	17.6	8.4	7	11	16	24	31	1,497

Table 3: Descriptive Statistics – Default Timing

Repayment   For each loan originated by Prosper, we have monthly data regarding repayment decisions of the borrower, i.e., we observe whether the borrower repaid the loan or not every month, and whether the borrower defaulted.[10] In Table 3, we report sample statistics regarding default timing conditional on being default. The average default period among all loans is 17.6, the lowest is 15 for grade HR listings, and the greatest is 18.9 for grade E listings. Hence, we do not find any clear relationship between the default timing and the credit grade conditional on being default.

---

[10]Prosper records loans that are more than 4 months late as "charge off". There are exceptions where the loans are kept on Prosper's books even after being late for 4 months. **Find prospers' policy on charge off. Maybe fraction of loans that are not charged off after 4 months.** Our definition is the same as Freedman and Jin (2010).

	mean	sd	10%	25%	50%	75%	90%	Obs
All	-.044	.400	-.856	-.090	.121	.185	.281	5,648
AA	-.008	.278	-.431	.061	.081	.106	.132	794
A	-.027	.331	-.767	.072	.094	.135	.180	772
B	-.072	.403	-.870	-.196	.135	.168	.210	1,030
C	-.060	.413	-.870	-.211	.134	.196	.256	1,294
D	-.035	.423	-.865	-.192	.153	.230	.316	1,025
E	.000	.475	-.861	-.315	.249	.345	.394	392
HR	-.110	.530	-.886	-.800	.198	.345	.398	341

Table 4: Descriptive Statistics – Internal Rate of Return

Internal Rate of Return   Finally, we report the internal rate of return (IRR) for the loans originated by Prosper in Table 4.[11]   The average IRR for all listings is -4.4% and it is all negative in any credit grade except grade E whose average IRR is 0%. The standard error of IRR for grade AA is 0.278, and it monotonically increases as the credit grade becomes worse. Table 4 also reports the distributions of IRR for each credit grade. The median of IRR is lowest for the grade AA listings at 0.08, and is highest for the grade E listings at 0.25.

2.3   Preliminary Analysis

In this section, we provide some evidence that the borrower's reserve rate works as a signaling device. To that end, we first examine the effect of the reserve interest rate on the funding probability and the realized contract interest rate. We show below that borrowers face a trade-off between the funding probability and the contract rate in setting the reserve rate: posting a lower reserve rate leads to a lower funding probability but a higher probability of receiving a favorable contract interest rate conditional on being funded. This suggests that borrowers who post a high reserve rate weighs this trade-off differently from those who post a low reserve rate, i.e., borrowers who post high reserve rates care more about their loan being funded than what interest they will get on the loan and vice versa. We then examine if there are any systematic difference between those who post high reserve rate and low reserve rate. We find that those who post a high reserve rate are "high risk" in the sense that they are more likely to default than those who post low reseve rates.

To sum, we find evidence that (1) there is a trade-off in setting the reserve rate, i.e., a trade-off between bigger funding probability and higher contract interest rate and (2) "low risk" types and "high risk" types see this trade-off differentially, that "low risk" types are more willing to sacrifice favorable interest rate for a bigger probability of being funded, (3)

---

[11]If we denote the (monthly) IRR by R, then R is the interest rate which equalizes the loan amount to the discounted sum of the stream of actual monthly repayments, i.e.,

$$\text{Loan Amount} = \sum_{t=1}^{T} \frac{t\text{-th Monthly Payment}}{(1+R)^t}.$$

In Table 4, we report the annualized IRR.

high reserve rate signals "high risk" and low reserve rate signals "low risk".

Funding Probability and Contract Interest Rate   In order to analyze the effect of the reserve rate on the funding probability, we run a probit model as follows:

$$\text{Funded}_j = 1\{\alpha_1 s_j + \mathbf{x}'_j \boldsymbol{\alpha} + \upsilon_j \geq 0\},$$

where Funded_j is a dummy variable for whether listing j is funded or not, s_j is the reserve rate, $\mathbf{x}_j$ are controls such as requested amount, debt to income ratio, dummy variable for home ownership, and dummy variable for credit grade, and υ_j is an error term following a standard Normal distribution. The first column of Table 5 reports the results of this regression. The coefficient on the reserve rate is 1.72 and is statistically significant. This implies that a listing is more likely to be funded as higher reserve interest rate is posted even after controlling for observed listing characteristics. Other coefficients also seem to have natural sign. Coefficients on the requested amount, debt-to-income ratio, and the home ownership dummy are all negative and significant, indicating that these characteritics have negative impact on the funding probability. We report the estimated coefficients for each credit grade dummy in the bottom half of the column. These coefficients show that the listing with higher credit grade is more likely to be funded.

Next, we run the following Tobit regression to examine the effect of the reserve rate on the contract interest rate:

$$
\begin{aligned}
r_j^* &= \gamma_1 s_j + \mathbf{x}'_j \boldsymbol{\gamma} + e_j, \\
r_j &= \left\{ \begin{array}{cc} r_j^* & \text{if } r_j^* \leq s_j \\ \text{missing} & \text{otherwise} \end{array} \right.
\end{aligned}
\tag{1}
$$

where r_j denotes the observed contract interest rate, r_j^* is the latent contract interest rate, $\mathbf{x}_j$ is the same vector of controls as before and e_j is a Normally distributed error term. The first equation relates the latent contract interest rate to the reserve rate and other characteristics. Note that we require the selection equation in order to account for the fact that the contract interest rate r_j is always less than the reserve rate, s_j. The interpretation of this Tobit specification is that r_j^* is the (latent) interest rate at which the loan will be funded in the absence of any reserve rate. But since the observed contract rate is truncated above by s_j, we have our second equation. Note also that our Tobit specification is different from a simple regression of the contract interest rate, r_j, on the reserve rate, s_j, which would capture the mechanical truncation effect rather than the causal effect.[12] We report the result from this regression in the second column of Table 5. The coefficient on the reserve interest rate is significantly positive. Hence, posting a lower reserve interest rate leads to a higher contract interest rate conditional on the observable characteristics and censoring effect, which is consistent with out hypothesis. As in the previous regression, other coefficients seem to be natural.

---

[12] Even if s_j had no causal effect on the contract rate, r_j, then r_j and s_j will have positive correlation because the contract interest rate is only observed if $r_j \leq s_j$. In this case the conditional distribution of r_j, $F_{r_j}(\cdot|s_j)$, and $F_{r_j}(\cdot|s'_j)$ for s_j and s'_j ($s_j < s'_j$) will be the same for any point below s_j, but $F_{r_j}(\cdot|s'_j)$ will have positive support above s_j. This will induce mechanical positive correlation.

	Funded	Contract Rate
reserve rate	1.7235***	0.7104***
	(0.0268)	(0.0144)
amount	-0.0001***	7.35E-06***
	(5.87E-12)	(2.03E-07)
debt / income	-0.6471***	0.0623***
	(0.0015)	(0.0036)
homeownership	-0.1654***	0.0142***
	(0.0004)	(0.0017)
Grade		
AA	3.3395***	-0.2735***
	(0.0044)	(0.0059)
A	2.7679***	-0.2400***
	(0.0033)	(0.0053)
B	2.3198***	-0.2118***
	(0.0022)	(0.0045)
C	1.6639***	-0.1670***
	(0.0015)	(0.0038)
D	1.0961***	-0.1169***
	(0.0012)	(0.0033)
E	0.5011***	-0.0492***
	(0.0015)	(0.0037)
Observation	27,887	27,887
R^2	0.2358	0.7959

Table 5: Reduced Form Analysis - Funding Probability and Contract Interest Rate

In addition to the Tobit model above, we also estimate a censored quantile regression model (see *e.g.* Powell (1984)) using the same specification as equation (1) in order to confirm that the relationship between the contract interest rate, r_j, and the reserve rate, s_j, holds not only for the mean, but also for other quantiles. We do not report the result in the Table 5 for saving the space. We find that the coefficient on the reserve interest rate is positive and statistically significant in all quantiles we estimate. Hence, the results imply that $F(r|s)$ first order stochastically dominates $F(r|s')$ for any $s \geq s'$.

Repayment Behavior and Reserve Interest Rate   Next we examine how the repayment behavior of the borrower is related to the reserve interest rate. In particular, we provide evidence that high reserve rate is associated with high default rate, and vice versa, which suggests that the reserve rate signals the type of the borrower. In our first specification, we run a panel Probit of an indicator variable for default on observable characteristics of the loan as well as the reserve rate,

$$\text{Default}_{jt} = 1\{\beta_0 + \beta_1 s_j + \beta_2 r_j + \mathbf{x}'_j \boldsymbol{\beta}_3 + \mu_t + \alpha_j + \omega_{it} \geq 0\}$$

where Default_{jt} is a dummy variable that takes one if the borrower j defaults on the loan at period t, s_j is the reserve rate, r_j is the contract interest rate, $\mathbf{x}_j$ is a vector of control variables, μ_t is a period dummy, α_j is a borrower random-effects and ω_{it} is a random error following a Normal distribution. Note that because we control for the contract interest rate in the regression as well as other observable loan characteristics, the effect of the reserve rate is purely due to selection. That is, conditional on the contract rate, the reserve rate does not directly affect the borrower once the loan is made, since the interest rate is determined by r_j. The coefficient on s_j thus captures the difference in the risk among borrowers who posted different reserve rates, but ended up with the same contract rate.

The parameter estimates obtained from this regression is shown in the left panel of Table 6. The coefficient associated to the reserve interest rate and the contract interest rate are both positive and significant. It implies that posting higher reserve interest rate while keeping the contract interest rate fixed results in higher default rate.

In order to examine how the reserve rate relates to the borrower's repayment behavior from the perspective of the lender, we now analyze how the IRR is related to the reserve interest rate by estimating the following model:

$$IRR_j = \beta_0 + \beta_1 s_j + \beta_2 r_j + \mathbf{x}'_j \boldsymbol{\beta}_3 + u_j$$

where IRR_j is the internal rate of return of loan j. As we mentioned before, we have controlled for observable characteristics of the loan, so the coefficient on s_j captures the selection effect.

The parameter estimates obtained from this regression is shown in the center panel of Table 6. The reserve interest rate has negative and significant effect on IRR, which indicates that lenders are less willing to bid to the listing with higher reserve rate. The effect of the contract interest rate on IRR is positive. It is not statistically significant, though.

Finally, we run the following hazard model to see the effect of the reserve rate on the default timing. Specifically, we estimate a Cox's proportional hazard model as follows:

$$\lambda(t_j|\,(s_j, r_j, \mathbf{x}_j)\,, \boldsymbol{\beta}) = \lambda_0(t) \exp(\beta_0 + \beta_1 s_j + \beta_2 r_j + \mathbf{x}'_j \boldsymbol{\beta}_3),$$

	Default	Rate of Return	Default Time
reserve rate	2.2590***	-0.6069***	3.6325***
	(0.4608)	(-0.1288)	(0.8019)
contract rate	1.6141***	0.0660	4.1304***
	(0.4413)	(0.1354)	(0.7472)
amount	1.92E-05***	-4.51E-06***	3.30E-05***
	(4.38E-06))	(1.24E-06)	(7.64E-06)
debt / income	0.0305	-0.0311	0.0777
	(0.0618)	(0.0194)	(0.1048)
homeownership	0.0604	-0.0469***	0.1399***
	(0.0384)	(0.0115)	(0.0714)
Grade			
AA	-0.2015	0.0583	-0.2242
	(0.1298)	(0.0393)	(0.2124)
A	-0.1511	0.0431	-0.2809
	(0.1135)	(0.0358)	(0.2442)
B	-0.0898	0.0217	-0.0999
	(0.0940)	(0.0314)	(0.1700)
C	-0.0779	0.0371	-0.1396
	(0.0819)	(0.0283)	(0.1436)
D	-0.1186	0.0643**	-0.2093
	(0.0773)	(0.0268)	(0.1331)
E	-0.2814***	0.1149***	-0.4862***
	(0.0878)	(0.0293)	(0.1486)
Observation	87,572	5,648	87,572
R^2		0.0249	

Table 6: Reduced Form Analysis - Repayment Behavior and Reserve Interest Rate

where t_j indicates borrower j defaults at period t_j, and $\lambda_0(t)$ is the baseline hazard function. As in the regressions above, we controll for observable characteristics of the loan. The third panel of the Table 6 reports the parameter estimates of the regression. The result implies that the default time becomes greater as the reserve interest rate becomes higher.

All results clearly show that the coefficients on the reserve interest rate are significant, and have the expected sign.[13] The result suggests that borrowers announcing different reserve rate are different in their pay back ability.

3 Model

Our model has roughly three parts. The first part of the model concerns how the borrower posts a listing. The key element of this part is the borrower's decision regarding the reserve interest rate. The borrowers have an unobservable type which affects both the ease at which they can borrow money from alternative sources and also their credit risk. As we discussed in the previous section, the reserve interest rate is an observable characteristic of the listing that can potentially be used as a signal: It affects the probability that the loan is funded, and also affects the interest rate that the borrower faces conditional on being funded.

The second part of our model concerns the lender's bidding behavior. The allocation (*i.e.* which potential lenders become the actual lenders) and the contract interest rate is determined through an auction which is a combination of an ascending auction and an uniform price auction.The lenders are heterogeneous with regard to their attitude toward risk. The lenders decide whether to bid or not and what to bid. A bid consists of how much to lend and at what interest rate.

The third part of our model is on the borrowers' repayment behavior. We model the repayment decision as a finite horizon single agent dynamic programming problem. At each period, the borrower chooses whether to pay back the loan or not, depending on whether the disutility from paying back outweighs the disutility from default.

3.1 Borrowers

We first describe the repayment stage and work our way backwards. We model the re-payment behavior of the borrower as a seqential decision of 36 $(= T)$ months, which is the length of the loans that Prosper originates.[14] We write the terminal decision of the borrower at period T as follows:

$$\begin{cases} \text{full repayment: if } u_T(r) + \varepsilon_T \geq D(\varphi) \\ \text{default: otherwise} \end{cases}$$

where $u_T(r) + \varepsilon_T$ denotes the utility of the borrower if he repays the loan in full and $D(\varphi)$ denotes the cost of defaulting. r denotes the interest rate on the loan, φ captures the (unobservable) type of the borrower which shifts the cost of defaulting and ε_T is an idiosyncratic shock, with $\varepsilon_T \perp \varphi$. The independence of (ε_T, φ) is a strong assumption, but we come back to this point below. We suppress the dependence of u_T on other characteristics

---

[13]We also run separate regression for each credit grade. The results are basically the same.

[14]There were a total of 365,201 repayments in total, of which 28,076 were "early repayments". We abstract away from our model.

of the loan such as the loan amount, debt to income ratio, home-ownership, etc. We assume without loss of generality that $D(\varphi)$ is monotonicly decreasing in φ, *i.e.*, the disutility of defaulting is larger for borrowers with higher φ. Hence borrowers with high φ are "good" types who value avoiding default and maintaing their credit history.

Now let V_T denote the expected utitlity of the borrower at the final period T, defined as $V_T(r, \varphi) = E[\max\{u_T(r) + \varepsilon_T, u(\varphi)\}]$. The decision of the borrower in period $t < T$ is as follows:
$$\begin{cases} \text{repayment: if } u_t(r) + \varepsilon_t + \beta V_{t+1}(r, \varphi) \geq D(\varphi) \\ \text{default: otherwise} \end{cases}$$

where $u_t(r) + \varepsilon_t$ is period t utility of repaying the loan and $V_{t+1}(r, \varphi)$ is the continuation utility which can be defined recursively. We assume that $\{\varepsilon_t\}$ are independent (but not necessarily identical) across t but allow u_t to be time-dependent to capture any deterministic time-dependence. We discuss below the implications of assuming that $\{\varepsilon_t\}$ are independent below.

We make a few remarks concerning our specification. Our first remark is related to the interpretation of φ. In our specification, the unobservable type of the borrower is modeled as default cost. However, we could write down an alternative model which is isomorphic to the current set up where φ is modeled as unobserved income of the borrower. Consider the following alternative specification:

$$\text{full repayment} \Leftrightarrow \tilde{u}_T(\tilde{\varphi} - \text{repayment}) + \tilde{\varepsilon}_T \geq 0,$$

where $\tilde{\varphi}$ is now the (unobserved) income/asset of the borrower. The problem of the borrower for $t < T$ is defined analogously. Now rearranging terms in the previous expression and using the fact that (repayment) = (interest multipled by amount), $(r * amt)$, we obtain

$$\text{full repayment} \Leftrightarrow -(r * amt) - \tilde{u}_T^{-1}(-\tilde{\varepsilon}_T) \geq -\tilde{\varphi}.$$

Note that if we redifine u_T and D as $u_T(r) = -(r * amt)$, $\varepsilon_t = -\tilde{u}_T^{-1}(-\tilde{\varepsilon}_T)$, and $D(\tilde{\varphi}) = -\tilde{\varphi}$, then the two specifications are equivalent. As long as there is an unobservable type of the borrower that determines the propensity to default – whether it be default cost, income or some comibnation of both – the resulting specification will be similar and the difference will be in the interpretation of φ only. For our purposes, the source of heterogeneity among the borrowers is also unimportant, as borrower heterogeneity will be structural to our counterfactual policy. This is not to say, however, that the distinction may very be important in other contexts.

Our second remark concerns the independence assumption on ε_t, $\varepsilon_t \perp \varphi$. While this is a restrictive assumption, we note that mean independence of ε_t and φ, $E[\varepsilon_t|\varphi] = 0$, is without loss of generality. Since we have not made any assumption on the form of $D(\varphi)$, we can always redefine $D(\varphi)$ as $D(\varphi) - E[\varepsilon_T|\varphi]$, which will result in $E[\varepsilon_T|\varphi] = 0$. The independence assumption makes the model more tractable. Note that there is little hope of recovering any serial correlation in $\{\varepsilon_t\}$. What we observe in the data are binary decisions (repay or default) where defaulting is a terminal state: If a borrower defaults, the data ends there and we do not observe later repayment decisions anymore. Unlike in a situation where there are distinct decisions for each of the T periods, our particular data structure precludes us from identifying the joint distribution of $\{\varepsilon_t\}$. We can only hope to recover the marginals of $\{\varepsilon_t\}$.

Next we describe the borrower's decision regarding the reserve interest, s. When the borrower determines s, he has to trade-off the effect of s on the probability that the loan is funded, and the effect of s on the contract interest rate, r. Recall from the previous section that increasing s tends to increase the funding probability but it simultaneously tends to increase contract interest rate. The borrower's problem is as follows:

$$\max_s \Pr(s) \int V_1(r,\varphi)f(r|s)dr + (1 - \Pr(s))\lambda(\varphi), \tag{2}$$

where $\Pr(s)$ is the probability that the loan is funded, $f(r|s)$ is the conditional distribution of realized interest rate given s with its CDF denoted by $F(r|s)$, and $\lambda(\varphi)$ is the borrower's utility from the outside option, *i.e.* the borrower's utility in the event that the borrower cannot obtain a loan from Prosper. We suppress the dependence of $P(s)$, the probability that the loan is funded, and $f(r|s)$, the conditional distribution of the interest rate on the characteristics of the borrower, e.g., credit grade and home ownership. $P(s)$ and $f(r|s)$ are known and taken as exogenous by the borrower, although thye are equilibrium objects.

Note that the first term in equation (2) captures the borrower's expected utility in the event that the loan is funded: The value function of the borrower (at period $t = 1$) when the contract interest rate is r, $V_1(r,\varphi)$, is integrated against the distribution of the contract interet rate $f(r|s)$. The second term in equation (2) captures the utility of the borrower in the event the loan is not funded: $(1 - \Pr(s))$ is the probability that this event occurs, which is multiplied by the utility of the outside option, $\lambda(\varphi)$. We make the assumption that $\lambda(\varphi)$ is increaing in φ, where φ is the private type of the borrower we defined earlier that shifts the disutility of default. This assumption just reflects the idea that "good" types (high φ), who value their credit history, for example, have an easier time obtaining a loan from other sourses, such as relatives, friends, and local banks, etc., while "bad" types, with low cost of default, e.g., who have relatively damaged credit history or expecting to default in the future anyway, are likely to be more desperate, with limited alternative sources of funding. This implies that $\lambda(\varphi)$ would be increasing in φ.

Then first order condition associated with this problem is

$$\frac{\partial}{\partial s}\Pr(s)\left(\int V_1(r,\varphi)f(r|s)dr - \lambda(\varphi)\right) + \Pr(s)\int V_1(r,\varphi)\frac{\partial}{\partial s}f(r|s)dr = 0. \tag{3}$$

The first order condition captures the two trade-offs that the borrower faces in determining the reserve interest, s. The first term is the incremental utility gain that results from an increase in the funding probability and the second term is the incremental utility loss resulting from an increase in the contract interest rate. Recall from the previous section that both $\Pr(s)$ and the expected mean of the contract interest rate is increasing in s, and $F(r|s)$ first order stochastically dominates $F(r|s')$ for any $s \geq s'$. We note that under these conditions, the single crossing property (SCP) (see e.g. Mailath (1987)) is satisfied. Note also that from the perspective of the borrower, SCP is neccesary and sufficient to induce separation, *i.e.* no pooling among types. We state this as a proposition below.

Proposition 1 *If $\frac{\partial}{\partial s}\Pr(s) > 0$ and $F(r|s)$ FOSD $F(r|s')$ for $s' > s$, then we have SCP, i.e.*

$$\frac{\partial}{\partial s\partial\varphi}\left[\Pr(s)\int V_1(r,\varphi)f(r|s)dr + (1 - \Pr(s))\lambda(\varphi)\right] < 0$$

3.2 Lenders

In this subsection, we describe the model of lenders' behavior. Let N be the (random) number of potential lenders who view a particular listing on Prosper's website. We let F_N denote its cumulative distribution function with support $\{0, 1, \cdots, \bar{N}\}$. A potential lender who observes a listing on Prosper may then decide whether to submit a bid or not, which is a interest-amount pair. At the time of bidding, a potential lender observes various characteristics of the listing which include the credit grade, the reserve rate, etc., as well as the active interest rate which we explained in Section 2. We let the potential lenders to be heterogenous with regard to their attitude toward risk.

When the lender determines whether to bid on a listing or not, the lender must first form an expectation over the return she will make if she funds part of the loan. Following the standard specification used in the asset pricing literature (see e.g. Paravisini *et al.* (2011)), we assume that lender's utility from owning an asset depends on the mean and variance of the asset. Thus we specify the utility of lender j who holds an asset Z with mean return $E[Z] = \mu(Z)$ and variance $Var[Z] = \sigma^2(Z)$ as follows:

$$U_j^L(Z) = \mu(Z) - A_j \sigma^2(Z)$$

where A_j is a lender specific random variable known only to lender j that determines her attitude toward risk. [15] Note that this specification easily extends to the case with amount choice. If the lender holds q units of asset Z, then $E[qZ] = q\mu(Z)$ and $E[(qZ)^2] = q^2\sigma^2(Z)$. Hence, lender j's utility of having q amount of asset Z can be expressed as follows,

$$U_j^L(qZ) = q\mu(Z) - A_j(q\sigma(Z))^2.$$

Now, we are ready to describe the lenders' problem: Lender j chooses amount q_j and interest rate r_j to maximize her expected utility,

$$\max\left\{\max_{\substack{q_j \in M, \\ r_j \leq \bar{r}}} E\left[U_j^L(q_j Z(r^*))|r_j\right] - c(q_j), \varepsilon_{0j}\right\} \tag{4}$$

where M denotes the set of feasible amount choices, $Z(r^*)$ denotes the loan when the contract interest is r^*, $\bar{r}$ is the active interest rate, $c(q)$ is the cost of lending amount q, and

---

[15]That can be seen as a reduced form of more general utility functions. Suppose the "true" utility function is

$$EU = E[U(z + w)]$$

for some wealth level w_j. Assuming that 2nd order Taylor expansion gives sufficiently well approximation,

$$EU = E[U(Z)] = E[U(w_j) + U'(w_j)z + \frac{1}{2}U''(w_j)z^2 + R]$$

$$= U(w_j) + U'(w_j)E[z] + \frac{1}{2}U''(w_j)E[z^2] + R'.$$

If R and R' is negligible, the reduced form expression can be justified.
Incorporating asset specific error terms ends up with the same expression. Setting the primitive utility function as $EU_{ij} = E[U_i(Z + \epsilon_{ij})]$ and doing the same expansion gives $U^L(A_j, Z) = \mu - (A_j + \epsilon_{jz})\sigma^2$. We are interested in the distribution of the coefficient on variance and both expression are equivalent for this purpose.

lastly, ε_{0j} is the lender specific opportunity cost of investing in this listing.[16] There are a few things to note about this expression. First, we let "asset" Z to be a function of r^*, since the return from making a loan depends on the actual contract interest rate. Second, the expectation in the previous expression is taken over the realization of r^*, or

$$E\left[U_j^L(q_j Z(r^*))|r_j\right] = \int_{r_j}^{\bar{r}} U_j^L(q_j Z(r^*))dG(r^*|r_j) = \int_{r_j}^{\bar{r}} \{q\mu(Z(r^*)) - A_j(q\sigma(Z(r^*)))^2\}dG(r^*|r_j).$$

where $G(r^*|r_j)$ denotes the distribution of the final contract interest rate, r^*, conditional on submitting a bid with interest rate equal to r_j. We assume that the lenders know how the contract interest rate, r^*, maps to the mean $\mu(Z(r^*))$ and variance $\sigma^2(Z(r^*))$. Our third point concers M, the set of feasible amout choices. In principle, the lender is free to bid any amount between $50 and the full requested amount by the borrower. But as we showed in Section 2, the vast majority of the bid amount is either $50, $100, or $200. We therefore assume $M \equiv \{\$0, \$50, \$100, \$200\}$ in what follows. Finally, note that ε_{0j} is meant to capture the outside option of the lender: For example, ε_{0j} can be the opportunity cost of taking money away from an existing asset in the portfolio and putting it in this listing. To the extent that lenders must decide on which listings to bid from a large pool of Prosper listings, ε_{0j} can be interpreted as a reduced form way of capturing the value of investing in these other listings.[17]

In order to characterize the lender's strategy, it is useful to illustrate the problem of the borrower graphically. Figure 3 is a graphical representation of the lender's problem in a simplified setting without any amount choice. In the left panel of this figure, we have taken the horizontal axis as σ^2 and the vertical axis as μ: This lets us plot any loan whose mean is μ and variance is σ^2 as a point on this $\mu - \sigma^2$ plane. Now take some listing and suppose that the listing is funded at a contract interest rate equal to the reserve rate, so that $r^* = s$. At $r^* = s$, this listing has mean return $\mu(Z(s))$ and variance $\sigma^2(Z(s))$, which corresponds to the top end-point of the curve (Curve X) depicted in the left panel of the figure. Since the mean return, $\mu(Z(r^*))$ and the variance, $\sigma^2(Z(r^*))$ are functions of the contract interest rate, r^*, we can plot the point on the plane that corresponds to the listing as a function of r^*. In the figure, the trajectory of $(\mu(Z(r^*)), \sigma^2(Z(r^*)))$ as r^* changes is shown as a movement along Curve X in the direction of the arrows. Thus, as the contract rate is bid down from s, the utility of making the loan also changes. This is shown in the right panel of the Figure: We plot the utility of the lender as the contract interest rate falls from s.

Note that we have also drawn in a dashed line in the left panel of Figure 3. This is the lender's indifference curve, $U_j^L(\mu, \sigma^2) = \varepsilon_0$, *i.e.*, the set of points (μ, σ^2) that makes the lender just indifferent between lending and not lending. As the lender's utility function is linear with respect to μ and σ^2, the indifference curve is a straight line, *i.e.* $\mu - A_j\sigma^2 + c(50) =$

---

[16]The borrower's problem described in the text corresponds to the case when the requested amount has been exceeded by the outstanding bid amount. The problem of the borrower when the full requested amount has not been filled is simply obtained by replacing $\bar{r}$ in equation (4) by s.

[17]We assume that ε_{0j} is i.i.d. across each lender. This may be a cause of concern if the portfolio of the lenders are correlated. However, we only require ε_{0j} to be i.i.d. conditional on a set of time dummies, i.e. we only require independence net of common macro shocks. Also, the lenders of Prosper are relatively dispersed: Lenders are typically individuals with no dominant player. Because of these reasons, we think that the i.i.d. assumption is not unreasonable.
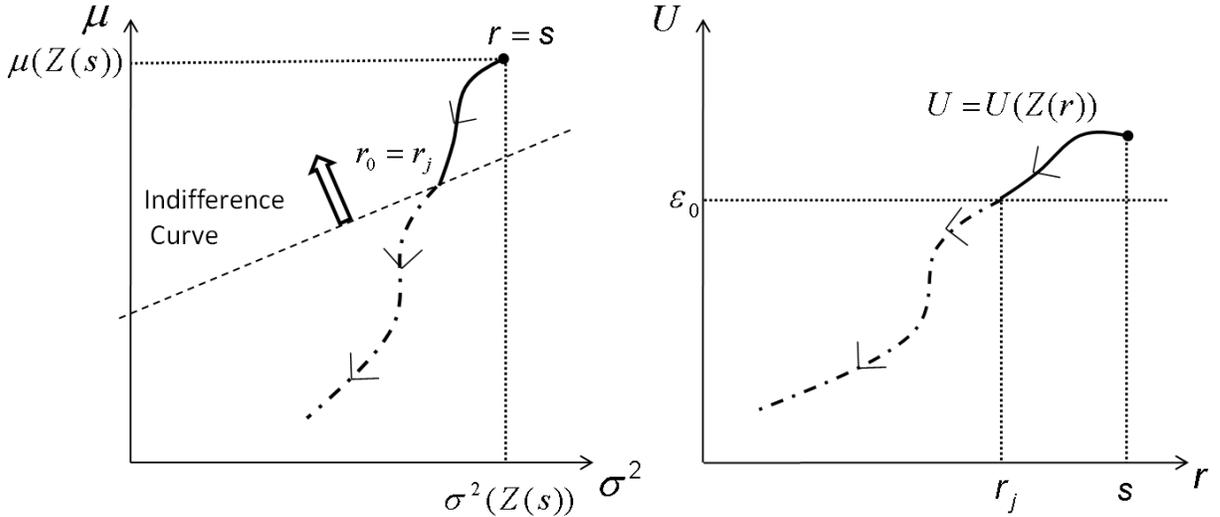
Figure 3: Determinants of Optimal Bidding Strategy: Case of No Amount Choice.

ε_0. Any points above this line gives the lender a strictly positive utility and vice versa. Now suppose that $(\mu(Z(r^0)), \sigma^2(Z(r^0)))$ is the intersection of curve X and the lender's indifference curve, that is, at contract interest $r^* = r^0$, the utility from lending money to the listing is exactly equal to ε_0. This can be seen in the right panel of the Figure as well, since the curve $U(r^*)$ crosses $U = \varepsilon_0$ at r_j. We show below that under the assumption that the lender is not pivotal, there is a dominant strategy for the lender.

Recall from Section 2 that the auction of Prosper has a flavor of the second price auction, *i.e.*, as long as the lender is not "pivotal" or marginal, (*i.e.*, the lender's bid is exactly the bid that brings the cumulative bid amount over the requested amount when we order the outstanding bids by their interest rate) the contract interest rate is determined by the bid of someone else. Hence, under this assumption, it is a dominant strategy for lenders to set $r_j = r^0$ in this auction. In order to see this, suppose that the lender bids an interest rate, r_j, which is higher than r^0, so that $r_j > r^0$. If the final contract interest r^* turns out to be above r_j, then the lender makes a loan at r^* regardless of whether she bids r^0 or r_j ($> r^0$). However, if the final contract interest rate is between r^0 and r_j, then the lender will be able to lend at a rate equal to r^* if she bids r^0, while she will not be able to lend if she bids r_j. Since lending at $r^* \in [r^0, r_j]$ gives the lender higher utility than the utility from the outside option, setting the rate equal to r^0 weakly dominates setting it to r_j.[18] Likewise, it is also easily shown that bidding $r_j(< r^0)$ is weakly dominated by bidding r^0 as well.

The preceding argument hinges on the assumption that bidders are never pivotal (or believe as if they are never pivotal). Although bidders sometimes do end up being pivotal, and hence this assumption is violated, we think that this is a reasonable approximation of the lender's behavior. Note that the average requested amount is $\$6,769$ for all listings ($\$5821$ for funded listings) while the average bid amount is a little over $\$70$. Hence, there

---

[18]If the contract interest r^* turns out to be less than r^0, then the the lender does not make the loan regardless of whether she bids r^0 or r_j.

is an average of about 80 winning bids conditional on the loan being funded. Moreover, the ability of the lender to strategically submit the interest rate is limited by the the lowest interest rate among the losing bids. In practice, the difference between the lowerst interest rate among the losing bids and the interest rate of the marginal bid is very small. Note also that a bidder who bids $50 is never pivotal if all other bidders bid ib multiples of $50 and the requested amount is in multiples of $50 as well.

As drawn in Figure 3, Curve X only intersects with the lender's indifference curve once. We have also abstracted from amount choice in our analysis of the lender's bidding strategy thus far. Below, we illustrate the lender's optimal strategy when the lender has an amount choice. We discuss the lender's optimal strategy when Curve X intersects with the indifference curve more than once in Appendix 2.

When the lender faces an amount choice, she needs to keep track of three curves, Curve 50, Curve 100 and Curve 200, where each of the curves are defined as $(\mu(qZ(r^*)),$ $\sigma^2(qZ(r^*)))$ with $q = 50$, 100, and 200. This is depicted in Figure 4: In this figure, we have plot the utitliy of the lender from lending different amounts as functions of r: $U_{50}(r)$, for example, corresponds to the utility of the lender if she lends $50 as a function of the contract interest rate, r. Now, just as before, there is a dominant strategy for the lender under the assumption that the bidder is not pivotal. The lender's dominant strategy is to bid an amount and interest rate so as to gurantee herself the maximum of $\{0, U_{50}, U_{100}, U_{200}\}$ at any contract interest rate r^*. As an example, for the case shown in Figure 4, this can be done by specifying the following bid strategy:

$$\text{bid amount \$200 and interest } r' \text{ if active interest rate} \in [r', s]$$
$$\text{bid amount \$100 and interest } r'' \text{ if active interest rate } \in [r'', r')$$
$$\text{bid amount \$50 and interest } r''' \text{ if active interest rate } \in [r''', r'')$$
$$\text{do not bid if active interest rate } \in [0, r''').$$

It is easy to check that this is the lender's dominant strategy by following the same logic we used to explain that bidding r_j is the lender's dominant strategy in Figure 4.

We now state the previous analysis in the form of a lemma.

Lemma 1 *Take $I_0 = [0, r_1]$, $I_1 = [r_1, r_2], \cdots I_M = [r_M, s]$ to be intervals so that $U_{q(k)}(r)$ is largest for all $r \in I_k$, where $q(k) \in \{\$0, \$50, \$100, \$200\}$ and $U_0(r) = 0$ for all r. Under the assumption that the lender behaves as if she is not pivotal, it is a dominant strategy to bid $q(k)$ and interest rate r_k when the active interest rate is in I_k.*

We make a few remarks at this point. First, note that this optimal strategy *does* require the lenders to submit new bids as the active interest rate changes. In the example above, the lender would submit new bids as the active interest rate drops below r', r'', and r'''. This implicitly takes as given that lenders have low cost of revising their bid.[19] (Actually, with our specification, we can pick listings where it is optimal for everybody to bid smaller amount as r goes down. Then, if we limit our estimating sample to such listings, we

---

[19]Some bidding strategies can be replicated with proxy bidding: For example, one can submit four $50 bids, two of which with corresponding interest rate at r' and two others with r' and r''. This would be equivalent to the bidding strategy described in the main text.
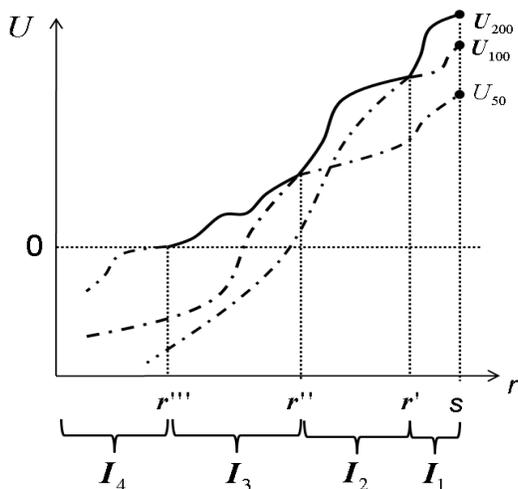
Figure 4: Determinants of Optimal Bidding Strategy: Case with Amount Choice

can potentially get around this problem: One can always replicate a complicated bidding strategy with multiple $50 bids) Second, there are other bidding strategies that ensure the same payoffs as the one described in the Lemma above: The strategy we described requires the least number of bid revisions necescsary, but it is only one of potentially many dominant strategies. In our estimation and identification, we only assume that the lenders are playing one of the dominant strategies, not necessarily the one described above.

Our final remark previews how the model of the lenders tie together with our identification and estimation. For any given listing, we can identify the mean-variance that corresponds to the listing if it were funded at the reserve interest rate, i.e., we can identify the "starting end point" of the mean-variance curve. This means that for a fixed distribution for A and N (the risk aversion parameter of the lender and the number of potential bidders) the model will induce a probability distribution over (i) whether a listing is funded and (ii) the number of lenders who bid $0, $50, $100, and $200. In the next section, we show that there is a one-to-one mapping from the primitives to the induced probability distribution. Correspondingly, our estimation is based on matching predicted moments with sample moments.

4   Identification

4.1   Identification of the Borrower's Primitives

The primitives of the borrower that we would like to identify are the period utility function, u_t, the distribution of borrower types $F_{\varphi|X}$, the cost of default, D, the utility from the outside option, λ, and the distribution of the shocks ε_t, F_{ε_t}. We begin with a few remarks. First, note that we allow the distribution of φ, $F_{\varphi|X}$, to be dependent on borrower characteristics, X, i.e. the type distribution of the borrower is identified for each X. Second, note that we can normalize D without loss of generality: A specification with $\tilde{D} = D \circ F_\varphi^{-1}$,

$\tilde{F}_\varphi = U[0,1]$, and $\tilde{c} = c \circ F_\varphi^{-1}$ is going to be observationally equiavalent to one with D, F_φ, and c.[20] Therefore, we normalize $D(\varphi) = -\varphi$. It is also easy to see that we can normalize the intercept of u_T and : Hence, we set $u_T(0) = 0$.[21]

Our identification result relies on the observation that there is a one-to-one (monotonic) mapping of s to φ conditional on X. This means that conditioning on a quantile of s (given X), is equivalent to conditioning on a quantile of $\varphi|X$. If we take observations (loans) that are conditional on an α - quantile of s, $F_{s|X}^{-1}(\alpha)$, the borrowers all have φ equal to $F_{\varphi|X}^{-1}(\alpha)$.

We start with our discussion on how to identify primitives from the last period, $t = T$ and work backward. Consider the repayment decision of the borrower with $\varphi_\alpha = F_{\varphi|X}^{-1}(\alpha)$ at period $t = T$. The borrower's problem is as follows:

$$\begin{cases} \text{repay: if } u_T(r) + \varepsilon_T \geq -F_{\varphi|X}^{-1}(\alpha) \\ \text{default: otherwise} \end{cases}$$

This is a binary threshold crossing model the identification of which has been studied extensively (See Matzkin (1992, 1994) for example). Hence by the results of (1992) the distribution of shocks ε_T, and the utiltiy u_T are nonparametrically identified as well as the α - quantile of φ, $F_{\varphi|X}^{-1}(\alpha)$, under certain conditions on u_T and F_{ε_T}. We will show in the Appendix that the specification that we use for estimation satisfies these conditions.

Now consider the $t = T - 1$ period problem:

$$\begin{cases} \text{repay: if } u_{T-1}(r) + \beta V_T(r, F_{\varphi|X}^{-1}(\alpha)) + \varepsilon_{T-1} \geq -F_{\varphi|X}^{-1}(\alpha) \\ \text{default: otherwise} \end{cases}$$

Again, $u_{T-1}(r) + \beta V_T(r, F_\varphi^{-1}(\alpha))$ is nonparametrically identified as well as the distribution of ε_{T-1}, using Matzkin's results. To see that u_{T-1} and β are separately identified, take $F_\varphi^{-1}(\alpha)$ and $F_\varphi^{-1}(\alpha')$. Then we can identify $u_{T-1}(r) + \beta V_T(r, F_\varphi^{-1}(\alpha))$ and $u_{T-1}(r) + \beta V_T(r, F_\varphi^{-1}(\alpha'))$, as well as the difference, $\beta V_T(r, F_\varphi^{-1}(\alpha)) - \beta V_T(r, F_\varphi^{-1}(\alpha'))$. Since $V_T(r, F_\varphi^{-1}(\alpha))$ and $V_T(r, F_\varphi^{-1}(\alpha'))$ are known functions of u_T, F_{ε_T}, and $F_{\varphi|X}^{-1}(\alpha)$, they are already identifed at period T, which implies that β is also identified. Lastly, because $u_{T-1}(r) + \beta V_T(r, F_\varphi^{-1}(\alpha))$ and $\beta V_T(r, F_\varphi^{-1}(\alpha))$ are identified, we can identify $u_{T-1}(r)$. Applying the same logic to earlier periods identifies $\{u_t, F_{\varepsilon_t|X}\}$, for all t.

Finally, we argue how to identify $\lambda(\varphi)$. Recall the borrower's FOC in equation (3):

$$\frac{\partial}{\partial s} \Pr(s) \left(\int V_1(r, \varphi) f(r|s) dr - \lambda(\varphi) \right) + \Pr(s) \int V_1(r, \varphi) \frac{\partial}{\partial s} f(r|s) dr = 0.$$

Solving for $c(\varphi)$, we obtain

$$\lambda(\varphi) = \int V_1(r, \varphi) f(r|s) dr + \frac{\Pr(s)}{\frac{\partial}{\partial s} \Pr(s)} \int V_1(r, \varphi) \frac{\partial}{\partial s} f(r|s) dr.$$

Note that all the terms in the right hand side are identified. V_1 is a known function of $\{u_t, F_{\varepsilon_t|X}\}$, F_φ and β. $\Pr(s)$ is the probability that the loan is funded at reserve interest rate

---

[20] The important component of the model is the distribution of $D(\varphi)$, not the distribution of φ or the shape of $D()$ per se. Hence we can either normalize $D()$ or F_φ.

[21] If we set $\tilde{u}_t = u_t + \kappa$ $(\forall t)$, $\tilde{F}_\varphi(t) = F_\varphi(t + \kappa)$, $\tilde{c}(\varphi) = c(\varphi + \kappa)$, it will be observationally equivalent to u_t, F_φ, and c.

equal to s, which is directly observed in the data. $f(r|s)$ is the conditional distribution of contract interest rate r given s, which is also directly observed. Hence the previous equation identifies $\lambda(\varphi)$.

4.2 Identification of the Lender's Primitives

There are N lenders, where each lender j chooses amount q_j and interest rate r_j to maximize her expected utility:

$$\max\left\{\max_{\substack{q_j \in M, \\ r_j \leq \bar{r}}} E\left[U_j^L(q_j Z(r^*))|r_j\right], \varepsilon_0\right\},$$

where

$$U_j^L(qZ) = q\mu(Z) - A_j(q\sigma(Z))^2 - c(q).$$

The primitives of the lender that we need to identify are the distribution of coefficient of risk, F_{A_j}, the distribution of the outside option, F_{ε_0}, the cost of lending, $c(q)$, and the distribution of the number of potential bidders, F_N. We first show how to identify F_{A_j}, F_{ε_0}, and $c(q)$ under the assumption that $P_q(\mu, \sigma)$, which we will define below, is known for all values of (μ, σ) and $q \in M \cup \{\$0\} \equiv \{\$0, \$50, \$100, \$200\}$. We will then show how $P_q(\mu, \sigma)$ is identified. Now let us define $P_q(\mu, \sigma)$ by

$$P_q(\mu, \sigma) = \Pr(\{q\mu - A_j(q\sigma)^2 - c(q) \geq \max\{\max_{q' \in M}\{q'\mu - A_j(q'\sigma)^2 - c(q')\}, \varepsilon_0\}) \text{ for } q \in M$$

$$P_0(\mu, \sigma) = \Pr(\varepsilon_0 \geq \max_{q' \in M}\{q'\mu - A_j(q'\sigma)^2 - c(q')).$$

$P_q(\mu, \sigma)$ is just the probability that holding q units of an asset whose mean return is μ and standard error is σ, gives higher utility than holding q' units of the asset for all $q' \neq q$. Note that $\sum_{q \in M \cup \{0\}} P_q(\mu, \sigma) = 1$. Note also that $P_q(\mu, \sigma)$ corresponds to the probability that (A_j, ε_0) falls into particular regions. Figure 5 shows this graphically. In Figure 5, we plot U_j^L for $q = \$50$, $\$100$, and $\$200$, denoted as U_{50}, U_{100}, U_{200}, as a function of A_j for the case of $\frac{100\mu - c(200) + c(100)}{30000\sigma^2} < \frac{c(50) - c(100) + 50\mu}{7500\sigma^2}$, i.e. $U_{50} = 50\mu - A_j(50\sigma)^2 - c(50)$, $U_{100} = 100\mu - A_j(100\sigma)^2 - c(100)$ and so on. Figure 5 also shows the regions of A_j that correspond to $P_q(\mu, \sigma)$, when $\varepsilon_0 = 0$.

We first discuss how to identify the cost of bidding, $c(q)$. We begin by noting that it is possible to assume $c(50) = 0$ without loss of generliaty.[22] We also assume that $c(200) < 3c(100)$ for our proof below. This second restriction is just for exposition: Identification for $c(200) > 3c(100)$ can be shown analogously. Now, consider $P_{200}(\mu, \sigma)$. Note that $P_{200}(\mu, \sigma)$ can be expressed as follows[23],

$$P_{200}(\mu, \sigma) = \Pr(\{U_{200}(\mu, \sigma) > \max\{U_{50}(\mu, \sigma), U_{100}(\mu, \sigma)\} \cap \{\varepsilon_0 < U_{200}(\mu, \sigma)\})$$

$$= \Pr(A_j < \frac{c(100) - c(200) + 100\mu}{30000\sigma^2} \wedge \varepsilon_0 < 200\mu - A_j(200\sigma)^2 - c(200)).$$

---

[22] We can add a constant to $c(50)$, $c(100)$, $c(200)$ and shift the distribution of ε to the right without changing the distribution of outcomes.

[23] This is true as long as μ is "big" enough, i.e. $\frac{-c(100) + 50\mu}{7500\sigma^2} > \frac{100\mu - c(200) + c(100)}{30000\sigma^2}$ $(\Leftrightarrow \mu > \frac{-c(200) + 5c(100)}{100})$
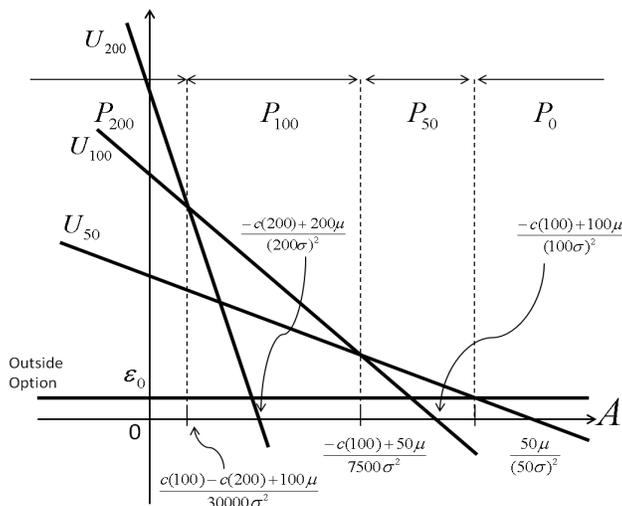
Figure 5: Determinants of $P_q(\mu, \sigma)$. In the figure, three lines represent $U_{50} = 50\mu + 2500\sigma A - c(50)$, $U_{100} = 100\mu + 10000\sigma A - c(100)$, and $U_{200} = 200\mu + 40000\sigma A - c(200)$, respectively.

Observe that if $\frac{c(100) - c(200) + 100\mu}{30000\sigma^2} < 0$, then as we let $\sigma \to 0$ (while keeping μ fixed), $P_{200}(\mu, \sigma)$ would tend to 0. However, if $\frac{c(100) - c(200) + 100\mu}{30000\sigma^2} = 0$, then if we let $\sigma \to 0$, $P_{200}(\mu, \sigma)$ would converge to a positive number,

$$\lim_{\sigma \to 0} P_{200}(\mu, \sigma) \to \Pr(A_j > 0 \wedge \varepsilon_0 < c(200) - 2c(200))$$

where we have substituted out 200μ using the fact that $c(100) - c(200) + 100\mu = 0$.[24] Hence define μ^* as $\mu^* = \sup_\mu \{\lim_{\sigma \to 0} P_{200}(\mu, \sigma) = 0\}$. μ^* is identified because everything in the right hand side of this expression is identified. Note that μ^* solves $c(100) - c(200) + 100\mu^* = 0$. Hence $c(100) - c(200)$ is identified. Similarly, working with the intersection between U_{50} and U_{100}, we can identify $c(100)$.[25]

Now we consider identification of F_{A_j}. Again note that

$$P_{200}(\mu, \sigma) = \Pr(A_j < \frac{c(100) - c(200) + 100\mu}{30000\sigma^2} \wedge \varepsilon_0 < 200\mu - A_j(200\sigma)^2 - c(200)).$$

Now take μ and σ so that $\frac{c(100) - c(200) + 100\mu}{30000\sigma^2} = \delta^+$, or equivalently, $\mu = \frac{c(200) - c(100) + 30000\sigma^2\delta^+}{100}$, where δ^+ is some negative number.[26] Now think about keeping $\frac{c(100) - c(200) + 100\mu}{30000\sigma^2}$ constant

---

[24]Recall that the expression for P_{200} takes the form in the text only if $\mu > \frac{-c(200) + 5c(100)}{100}$. Hence implicitly, we are assuming that the value of μ which solves $\frac{c(100) - c(200) + 100\mu}{30000\sigma^2} = 0$ satisfies this restriction, i.e. $\frac{c(200) - c(100)}{100} > \frac{-c(200) + 5c(100)}{100} \Leftrightarrow c(200) > 3c(100)$.

[25]As before, we need that the value of μ which equates $\frac{-c(100) + 50\mu}{7500\sigma^2} = 0$ ($\Leftrightarrow \mu = \frac{c(100)}{50}$) satisfies $\mu > \frac{-c(200) + 5c(100)}{100}$, or equivalently, $\frac{c(100)}{50} > \frac{-c(200) + 5c(100)}{100}$ ($\Leftrightarrow c(200) > 3c(100)$)

[26]As before, we need μ to satisfy $\mu > \frac{-c(200) + 5c(100)}{100}$. This means that $\frac{c(200) - c(100) + 30000\sigma^2\delta^+}{100} > \frac{-c(200) + 5c(100)}{100} \Leftrightarrow c(200) > 3c(100) - 15000\sigma^2\delta^+$. If $c(200) > 3c(100)$, this restriction will be satisfied for all σ and δ^+.

26

at δ^+, but moving $200\mu - A_j(200\sigma)^2 - c(200)$ by changing both μ and σ. In particular, consider $\sigma \to 0$. Then $\mu \to \frac{c(200)-c(100)}{100}$ and $P_{200}(\mu,\sigma) \to \Pr(A_j > \delta^+ \wedge \varepsilon_0 < c(200) - 2c(100)) = \Pr(A_j > \delta^+)\Pr(\varepsilon_0 < c(200) - 2c(100))$. By moving $\delta^+ > 0$, we identify $\Pr(A_j > t)\Pr(\varepsilon_0 < c(200)-c(100))$ for all $t > 0$. Now think about keeping $\frac{c(100)-c(200)-100\mu}{30000\sigma^2}$ constant at some negative number δ^-, but moving $200\mu - A_j(200\sigma)^2 - c(200)$.[27] In particular, consider $\sigma \to 0$ i.e., $\mu = \frac{c(200)-c(100)+30000\sigma^2\delta^-}{100}$. Then $\mu \to \frac{c(200)-c(100)}{100}$. $P_{200}(\mu,\sigma) \to \Pr(A_j > \delta^- \wedge \varepsilon_0 < c(200) - 2c(100) = \Pr(A_j > \delta^-)\Pr(\varepsilon_0 < c(100) - 2c(200))$. Hence we identify $\Pr(A_j > t)\Pr(\varepsilon_0 < c(200) - 2c(100))$ for all $t < 0$. Hence we identify F_{A_j}. Once F_{A_j} is satisfied, it is straightforward to identify the distribution of F_{ε_0}. We provide a proof of identification in Appendix X.

Now we discuss how to identify $P_q(\mu,\sigma)$ for all values of μ and σ and $q \in M$ and the distribution of the number of lenders, F_N. For the purpose of exposition, we first start our discussion when $M = \{\$0, \$50\}$, *i.e.*, when the lenders do not have any amount choice. To begin, recall that we assumed that F_N has finite support, *i.e.* the support is $\{0, 1, \cdots, \bar{N}\}$ for some finite $\bar{N}$. The upper bound $\bar{N}$ is identified by the maximum amount requested by the borrower that has positive probability of being funded. If the borrower requests an amount that is larger than $\$50 \times \bar{N}$, then the loan is never funded. Conversely, for loan amount less than $\$50 \times \bar{N}$, there is a positive probability of being funded. Hence $\bar{N}$ is identified by the maximum loan amount for which the probability of being funded is nonzero.

Next we identify $\{P_0(\mu,\sigma), P_{50}(\mu,\sigma)\}$ and $f_N(0), \ldots, f_N(\bar{N})$, where $f_N(\cdot)$ is the pdf of F_N. In order to do so, consider listings which, if funded at an interest rate equal to the reserve interest rate, yields (ex-ante) mean return μ and variance σ^2. Among such listings, consider listings with requested amount just equal to $\$50 \times \bar{N}$. Then it follows that

$$\Pr(\text{fund} = 1|x_{amt} = \$50 \times \bar{N}) = f_N(\bar{N}) \times P_{50}(\mu,\sigma)^{\bar{N}},$$

where the left hand side is the observed funding probability, and the right hand side is the probability that $\bar{N}$ potential lenders visit the listing, multiplied by the probability that all lenders who visited the listing are willing to bid $\$50$ (i.e., the lenders obtain higher utility from lending than from not lending). This equation holds by Lemma 3: Lenders submit a bid if and only if the utility she receives at the active interest rate exceeds the utility of not funding the loan. This implies that a listing with loan amount equal to x_{amt} is funded if and only if there are more than $\lfloor \frac{x_{amt}}{50} \rfloor + 1$ (i.e. the smallest interger that exceeds the loan amount divided by $\$50$).

Similarly, consider a listing with requested amount equal to $\$50 \times (\bar{N} - 1)$. Then

$$\Pr(\text{fund} = 1|x_{amt} = \$50 \times (\bar{N} - 1))$$
$$= f_N(\bar{N}) \times \binom{\bar{N}}{1} P_0(\mu,\sigma)P_{50}(\mu,\sigma)^{\bar{N}} + f_N(\bar{N} - 1) \times P_{50}(\mu,\sigma)^{(\bar{N}-1)}.$$

The right hand side is equal to the sum of two probabilities: the first term is the probability that $\bar{N}$ potential lenders visit the listing and $\bar{N} - 1$ of them decide to bid, and the second

---

[27]We need μ to satisfy $\mu > \frac{-c(200)+5c(100)}{100}$. This means that $\frac{c(200)-c(100)+30000\sigma^2\delta^-}{100} > \frac{-c(200)+5c(100)}{100} \Leftrightarrow c(200) > 3c(100) - 15000\sigma^2\delta^-$. If $c(200) > 3c(100)$, for each δ^-, there will be some small interval $(0, \epsilon_{\delta^-})$ such that for any $\sigma \in (0, \epsilon_{\delta^-})$ this restriction is satisfied.

term corresponds to the probability that $\bar{N} - 1$ potential lenders visit the listing, and all of them decide to bid. We repeat this process for all amounts $\{50, 100, \cdots, 50 \times \bar{N}\}$. This yields $\bar{N}$ equations (for each loan amount) and $\bar{N}+1$ unknowns, $P_{50}(\mu, \sigma)$, $f_N(1), \ldots, f_N(\bar{N})$ (Note that $P_0(\mu, \sigma) = 1 - P_{50}(\mu, \sigma)$ and $f_N(0) = 1 - \sum_{n=1}^{\bar{N}} f_N(n)$). Now consider repeating the above excercise with a different μ and σ (say μ' and σ'). Then this yields $\bar{N}$ additional equations. If we assume that F_N is invariant to (μ, σ), we have a total of $2 \times \bar{N}$ equations and $\bar{N} + 2$ unknowns, $P_{50}(\mu, \sigma)$, $P_{50}(\mu', \sigma')$, $f_N(0), \ldots, f_N(\bar{N})$. Hence $P_0(\mu, \sigma)$, and $f_N(0), \ldots, f_N(\bar{N})$ are identified for all μ and σ.

The preceeding identification argument focused on the case when $M = \{\$0, \$50\}$, *i.e.*, when there is no amount choice. We now briefly discuss identification when $M = \{\$0, \$50, \$100, \$200\}$. As before, we start with identification of $\bar{N}$: $\bar{N}$ is again identified by the maximum loan amount for which the probability of being funded is nonzero: $\$200 \times \bar{N}$ is the threshold loan amount, below which the probability of being funded is positive, and above which the probability is zero.

In the presense of amount choice, the objects that we would like to identify are now $\{P_0(\mu, \sigma), P_{50}(\mu, \sigma), P_{100}(\mu, \sigma), P_{200}(\mu, \sigma)\}$ and $f_N(0), \ldots, f_N(\bar{N})$. As before, consider listings which, if funded at an interest rate equal to the reserve interest rate, yields mean return μ and variance σ^2. Moreover, if we consider listings with amount equal to $\$200 \times \bar{N}$, we see that

$$\Pr(\text{fund} = 1 | x_{amt} = \$200 \times \bar{N}) = f_N(\bar{N}) \times P_{200}(\mu, \sigma)^{\bar{N}}.$$

For listings with amount equal to $\$200 \times (\bar{N} - 1) + 100$,

$$\Pr(\text{fund} = 1 | x_{amt} = \$200 \times (\bar{N} - 1) + 100)$$
$$= \binom{\bar{N}}{1} f_N(\bar{N}) \times P_{200}(\mu, \sigma)^{\bar{N}-1} P_{100}(\mu, \sigma).$$

Similarly, we can express the probability that the loan is funded for different loan amounts as a function of $P_q(\mu, \sigma)$ and f_N. The number of equations we end up with is $\bar{N}$, and the number of unobservables is $\bar{N} + 3$. Assuming that F_N is invariant to (μ, σ), we can increase the number of equations at a faster rate than the number of observables. Hence $P_q(\mu, \sigma)$ and f_N are identified for all μ and σ.

Step2wo owattekara x ni izonsasetemo iitoka kaku.

5 Estimation

We estimate our model in two steps. First, we estimate the conditional distribution of the contract interest rate given the reserve rate $f(r|s)$ and the funding probability, $\Pr(s)$. We estimate these two functions nonparametrically: As $f(r|s)$ and $\Pr(s)$ are both equilibrium objects we do not want to place any parametric assumptions. The second step involves estimating the primitives of the model of the borrower and the lender. While our discussion of model identification argument was focused on achieving nonparametric identification, we estimate our model with a parametric assumption.

5.1 Estimation of $f(r|s)$ and $\Pr(s)$

Our estimation proceeds first by estimating the probability density function of the contract interest rate, r, conditinal on the reserve interest rate, s, and observable listing characteristics, x, including the requested amount, debt-to-income ratio, and home ownweship, *i.e.*, $f(r|s,x)$, and the funding probability as a function of s and x, $\Pr(s,x)$. Since we observe the empirical distribution of r given s and x, and the empirical probability given s in the data, we can nonparametrically estimate these objects. Our estimation is based on Gallant and Nychka (1987), which propose a maximum likelihood estimation with a Hermite series function approximation.

In order to do so, we first divide the observations into 14 subsamples by each borrower's credit grade (AA, A, B, C, D, E, and HR) and whether he has his own home or not. This step is necessary because the estimation strategy by Gallant and Nychka (1987) requires the continuous support for each covariate. Hence, we estimate the parameters by each subsample in what follows. In each partition, we nonparametrically estimate the joint distributions of $f(r,s,x)$ and $\Pr(s,x)$ as the function of r, s and also observable listing characteristics x except credit grade and home ownership. We then compute the conditional distribution, $\widehat{f}(r|s,x)$, by analytically integrating $\widehat{f}(r,s,x)$.

In addition to $\widehat{f}(r|s,x)$ and $\widehat{\Pr}(s,x)$, we need to obtain the derivatives of $\widehat{f}(r|s,x)$ and $\widehat{\Pr}(s,x)$ with respect to s, $\frac{\partial \widehat{\Pr}(s,x)}{\partial s}$ and $\frac{\partial \widehat{f}(r|s,x)}{\partial s}$, when we evaluate the first order condition of the borrower's problem in equation (3). We compute $\frac{\partial \widehat{\Pr}(s,x)}{\partial s}$ and $\frac{\partial \widehat{f}(r|s)}{\partial s}$ by analytically taking derivatives.

5.2 Estimation of Borrower Side

The second part of the estimation concerns the borrower's behavior. We parameterize the borrower's period t utility function and outside option with parameters θ_B, *i.e.* $u_t(r,x;\theta_B)$ and $\lambda(\varphi;\theta_B)$. Recall that the default cost $D(\varphi)$ is normalilzed so that $D(\varphi) = -\varphi$. We note that θ_B is estimated for each credit grade and for each home ownership dummy.

For estimating θ_B, we use maximum likelihood, *i.e.*, we maximize the likelihood given realized sequences of repayment behaviour of each borrower. Given θ_B, our model of the borrower's repayment behavior generates a probability distribution over various sequences of repayment and default decisions for each borrower type φ. But we cannot use this to form a likelihood directly, as we do not observe the type of the borrowers. However, recall that there is a monotone relationship between φ and s (conditional on x), where this relationship is implicitly defined by the borrower's first order condition. This means that we can use the first order condition and the observed reserve interest rate (s) to back out the type of the borrower. Once we can assign a φ for each borrower, we can then compute the likelihood. We explain each of these steps in more detail below.

First note that we can compute $V_1(r,\varphi,x;\theta_B)$ given θ_B. That is, for any value of $\{r,\varphi,x\}$, we can recursively solve the borrower's dynamic problem, and compute the value function, $V_1(r,\varphi,x;\theta_B)$, given θ_B. This implies that the borrower's first order condition,

$$\frac{\partial}{\partial s} \Pr(s, x) \left(\int V_1(r, \varphi, x; \theta_B) f(r|s, x) dr - \lambda(\varphi; \theta_B) \right) + \Pr(s, x) \int V_1(r, \varphi, x; \theta_B) \frac{\partial}{\partial s} f(r|s, x) dr = 0, \tag{5}$$

can be seen as an equation in φ. In other words, the first order condition reveals, for each choice of s, the type of borrower φ that found it optimal to choose s. Since we have estimated $\Pr(s, x)$, and $f(r|s, x)$ in the first step, we can replace these objects with our nonparametric estimate $\widehat{\Pr}(s, x)$, $\widehat{f}(r|s, x)$. This allows us to back out the borrower type, $\widehat{\varphi}(s, x; \theta_B)$, for each borrower. Note that Proposition 1 shows that the right hand of equation (5) is monotonic in φ, guaranteeing that a unique solution exists given s and x.

Now we can compute the likelihood for a given sequence of repayment decisions for each borrower using $\widehat{\varphi}(s, x; \theta_B)$ that we obtained from the first order condition. Let d_{it} be an indicator variable that is equal to 1 if borrower i defaults at period t, and 0, otherwise. Then, the borrower i's default probability at period t is

$$\Pr(d_{it} = 1; \theta_B) = \int \mathbb{1} \left\{ D(\widehat{\varphi}(s_i, x_i; \theta_B) \geq u_t(r_i, x_i; \theta_B) + \varepsilon_{it} + \beta V_{t+1}(r_i, \widehat{\varphi}(s_i, x_i; \theta_B)) \right\} dF_\varepsilon. \tag{6}$$

Similarly, the probability of paying back at period t is

$$\Pr(d_{it} = 0; \theta_B) = \int \mathbb{1} \left\{ D(\widehat{\varphi}(s_i, x_i; \theta_B) \leq u_t(r_i, x_i; \theta_B) + \varepsilon_{it} + \beta V_{t+1}(r_i, \widehat{\varphi}(s_i, x_i; \theta_B)) \right\} dF_\varepsilon.$$

The likelihood of a sequence of repayment decisions, $\{d_{it}\}$ is then

$$\prod_{t=1}^{T_i} \Pr(d_{it} = 1)^{d_{it}} \times \Pr(d_{it} = 0)^{(1-d_{it})},$$

where $T_i \equiv \max\{1 + \sum_{\tau=1}^{T} d_{i\backslash\tau}, 36\}$, *i.e.* the number of periods until default or 36 periods, whichever is smaller.

Finally, the likelihood function is written as

$$L(\theta_B) = \prod_{i=1}^{N_L} \prod_{t=1}^{T_i} \Pr(d_{it} = 1)^{d_{it}} \times \Pr(d_{it} = 0)^{(1-d_{it})}, \tag{7}$$

where N_L is the number of loans. We obtain our parameter estimates by maximizing the likelihood function.

5.3 Estimation of Lender Side

The last part of the estimation considers the lender's bidding behaviour. In this section, we discuss how to estimate the distribution of the number of potential bidders, F_N, the distribution of the lender's risk attitude, F_A, and the lender's cost of bidding, $c(q)$. We let θ_L denote the vector of parameters. We note that we estimate a (potentially) different θ for each credit grade.

We use (simulated) method of moments and match conditional funding probability in order to estimate θ_L. For each listing j, we first take a draw from a log normal distribution

$F_N(\cdot; \theta_L)$, which corresponds to the number of potential bidders. Denote this draw as $\tilde{N}$. Then for each potential bidder, we draw the lender's risk attitude, $(A_j)_{j=1}^{\tilde{N}}$, for each potential bidder. We then compute the optimal bid choice, denoted by q_j^*, for each potential bidder, *i.e.*, whether to bid 0, 50, 100, or 200, at the reserve interest rate s by solving the following problem:

$$q_j^* = \arg\max_{q_j \in M} \left\{ E\left[U_j^L(q_j Z(s)) \right] - c(q_j; \theta_L) \right\},$$

where

$$U_j^L(qZ) = q\mu(Z) - A_j(q\sigma(Z))^2.$$

We then check whether there are enough bids to cover the requested amount or not. Define a dummy $\widehat{fd}$ which takes the value one if the loan is funded and zero otherwise,

$$\widehat{fd} = 1\left\{ \sum_{j=1}^{\tilde{N}} q_j^* \geq x_{amt} \right\}$$

where amt is the requested loan amount. Recall that a consequence of Lemma 3 is that a listing is funded if and only if there are enough people that want to fund the loan at the reserve rate. We repeat the preceeding steps NS times and compute the fraction of instances, denoted by $\hat{p}$, in which the loan is funded, i.e., $\hat{p} = \frac{1}{NS} \Sigma_{k=1}^{NS} \widehat{fd}_k$.

We match the funding probability by amount, and by debt-to-income ratio. The moment conditions we use is the following:

$$\sum_{A \in AMT} \sum_{i=1}^N \frac{1\{amt_i = A\}(fd_i - \hat{p}_i)}{\sum_{i'}^N 1\{amt_{i'} = A\}},$$

where AMT are discrete bins for loan amount, $x_{i,amt}$ is the loan amount for listing i, fd_i is a dummy variable that equals one if listing i is funded and zero otherwise. We similarly form moment conditions for debt-to-income ratio.

6   Result (preliminary)

6.1   Borrower Side

(Estimates are still preliminary. Standard errors will be updated soon.) The parameter estimates of the borrower side model is reported in Table (xxx).[28] The exact specification we estimate is

$$u_t(r_j, x_j; \theta_B) = \begin{bmatrix} [\theta_{const}, \theta_r, \theta_{r^2}, \theta_{r \times amt}, \theta_{amt}, \theta_{amt^2}, \theta_{debt}]'\mathbf{X}_j + \\ \theta_{4-6}, \theta_{7-9}, \theta_{10-12}, \theta_{13-15}, \theta_{16-18}, \theta_{19-21}, \\ \theta_{22-24}, \theta_{25-27}, \theta_{28-30}, \theta_{31-33}, \theta_{34-36} \end{bmatrix}' \mathbf{T},$$

$$D(\varphi) = -\varphi,$$

$$\lambda(\varphi; \theta) = \theta_\varphi \varphi,$$

---

[28] We estimate the coefficients using the observations with credit grade AA.

Parameter	Estimates
θ_r	-0.064
θ_{r^2}	0.105
θ_{amt}	0.040
θ_{amt^2}	-0.002
$\theta_{r \times amt}$	-5.24E-08
θ_{debt}	-0.008
θ_{ho}	-0.002
θ_{φ}	2.20E-08
θ_{const}	-0.188
Observation	794
Log likelihood	-672.8

Table 7: Parameter Estimates of the Borrower's Utility: time dummies are also included, but omitted from the table. We obtain the estimates using the observations from grade AA listings.

where $\mathbf{X}_j = [1, r_j, r_j^2, r_j \times amt_j, amt_j, amt_j^2, debt_j]$ and $\mathbf{T} = [1\{4 \leq t \leq 6\}, 1\{7 \leq t \leq 9\}, 1\{10 \leq t \leq 12\}, 1\{13 \leq t \leq 15\}, 1\{16 \leq t \leq 18\}, 1\{19 \leq t \leq 21\}, 1\{22 \leq t \leq 24\}, 1\{25 \leq t \leq 27\}, 1\{28 \leq t \leq 30\}, 1\{31 \leq t \leq 33\}, 1\{34 \leq t \leq 36\}]$. We also assume that the idiosyncratic preference shock ε_{it}s follow an *i.i.d.* Type I extremum distribution. The borrower i's default probability at period t in equation (6) has the analytical form as follows:

$$\Pr(d_{it} = 1; \theta_B) = \frac{\exp(D(\widehat{\varphi}(s_i, x_i; \theta_B); \theta_B))}{\exp(u_t(r_i, x_i; \theta_B) + \beta V_{t+1}(r_i, \widehat{\varphi}(s_i, x_i; \theta_B), x_i; \theta_B)) + \exp(D(\widehat{\varphi}(s_i, x_i; \theta_B); \theta_B))}.$$

Finally, we do not estimate the discount factor, β, and set at $0.95^{1/12}$.

The estimates for the contract interest rate and for the squared interest rate, θ_r and θ_{r^2}, are -0.064 and 0.105. That is, the higher interest rate would be, the higher the default probability would be, while the magnitude of such negative relationship diminishes as the interest rate increases. We estimate the coefficient for the requested amount and the squared amount, θ_{amt} and θ_{amt^2}, to be 0.04 and -0.002. These estimates imply that the borrower who borrows more is less likely to default, but the effect becomes smaller as the amount borrowed increases. The interaction term between the interest rate and the amount, $\theta_{r \times amt}$, is very small. To control for observable characteristics regarding the borrower's credit worthiness, we include the debt to income ratio and home ownership

Parameter	Estimates
μ_N	4.3822
σ_N^2	0.6361
c_{50}	3.4601
c_{100}	0.2154
C_{200}	-0.8480
μ_A	0.6856
σ_A^2	3.7757
Observation	1,188

Table 8: Parameter Estimates of the Lender Side Model: These estimates are still preliminary and subject to changes.

status. The coefficients for these variables are both negative, indicating that the borrower whose debt to income ratio is higher is more likely to default, and that the borrower who owns his house is more likely to default, respectively. Lastly, the borrower's outside option, $\lambda(\varphi)$, is close to 0.

6.2   Lender Side

(Parameter estimates of the lender side model is still incomplete. The specification is slightly different from one in Section 4. Standard errors are yet esimated.) Next we report the parameter esimates for the model of lenders' side in Table 8. The parameters for the lender's utility function are the costs of bidding for each amount choice, $\{c_{50}, c_{100}, c_{200}\}$. We assume that the distribution of potential bidders, F_N, follows a log normal distribution with parameters μ_N and σ_N^2, and that the distribution of risk attitude, F_{A_j}, is a normal distribution with mean μ_A and variance σ_A^2, $N(\mu_A, \sigma_A^2)$.

Table 8 indicates that parameters for the distribution of the number of potential bidders are 4.4 and 0.6. These imply that on average more than 80 potential bidders browse a listing. Cost of bidding is highest for bidding \$50, and is lowest for bidding \$200, which is consistent with the fact that more than 70% of bidders choose \$50 as in the descriptive statistics in Section 2. Finally, the distribution of lenders' risk parameter has its mean at 0.69 and variance at 3.78, meaning that more than half of lenders are risk averse, while as many as 42.7% of lenders are risk loving.
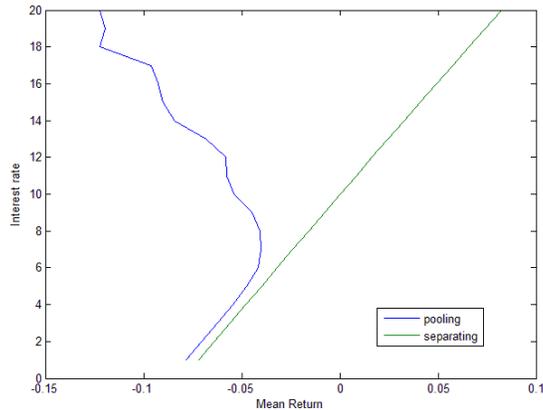
Figure 6: Mean Return v.s. Contract Interest Rate: The mean return-interest rate curve for the case with reserve interest rate (right) is computed by setting the reserve interest rate at 12%.

7 Counterfactual Experiment (preliminary)

7.1 Counterfactual Experiment 1: Credit Supply Curve without Signaling Possibility

In our first counterfactual experiment, we investigate what the credit supply curve would have been if the signaling though the reserve interest rate had been prohibited. Without any signaling possibility, all borrowers with observationally same characteristics who are intrinsically different in their credit worthiness are pooled. Hence, the market might suffer the traditional adverse selection problem as Stiglitz and Weiss (1982) point out. They show that the expected return curve under the adverse selection problem might become non-monotonic in the contract interest rate. This is bacause there is a trade-off between the interest earnings and adverse selection problem. As the interest rate rises, the adverse-selection effect predominates the positive effect of higher interest rate, and the expected return becomes non-monotonic in the interest rate. We provide the computational algorithm of credit supply curves with and without reserve interest rate in Appendix 4.

In Figure 6, we plot the mean return and the contract interest rate for the case where posting the reserve interest rate is allowed (right) and for the case where doing so is not allowed (left). Note that these curves are not credit supply curves because lenders take into account not only the mean return but also the variance when they make a bidding decision. Notably, the Figure indicates that the mean return is increasing in the contract interest rate when the reserve interest rate is allowed. On the other hand, the mean return is first increasing in the interest rate, but then decreasing when the reserve interest rate is prohibited. Moreover, the mean return is greater for the separating case than the pooling case. Hence, if the lenders are risk neutral, then the credit rationing may occur due to the adverse selection problem as is consistent with the prediction by Stiglitz and Weiss (1981).[29]

---

[29]Adams *et al.* (2009) found a similar non-monotonic relationship between the loan amount and the

We further investigate the effect of signaling device by estimating the credit supply curves. To do so, we re-compute the lender's bidding strategy when posting the reserve interest rate is prohibited, and obtain the funded amount. The computation algorithm for estimating the credit suppy curves is discussed in detail in Appendix 4. **(The figure of credit supply curves will be added here.)**

8   Conclusion

In this paper, we study how the signaling device can restore some of inefficiencies arising from the adverse selection problem using the data from an online peer-to-peer lending market, Prosper.com. We find some evidence showing that the reserve interest rate posted by the potential borrower works as a signaling device. Based on the evidence, we then develop and estimate a structural model of borrowers and lenders, where low reserve interest rate can credibly signal low default risk. Finally, we conduct two counterfactual experiments. First, we compute the credit supply curve when posting a reserve interest rate is prohibited. As Stiglitz and Weiss (1981) point out, the supply curve of the credit becomes non-monotonic and the credit rationing happens when the borrowers are pooled. Second, we compare the efficient market outcome to the current outcome with signaling possibility.

There are some issues left for future researches. One issue we could not address in this paper is the dynamics of the market. Some potential borrowers obtain a loan more than once. It is important to understand the evolution of the market resulting from entry/exit of borrowers. Another issue is the design of the platform. Other peer-to-peer lending markets, such as Lending Club, use different mechanisms to match potential borrowers with investors. Studying how the choice of the mechanism affect the market outcome is still an open and important question.

References

[1] Adams, William, Liran Einav, Jonathan Levin (2009) "Liquidity Constraints and Imperfect Information in Subprime Lending," American Economic Review, Vol 99(1), pp. 49–84.

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

[3] Aryal, Goyal, Isabelle Perrigne, and Quang Vuong (2009) "Identification of Insurance Models with Multidimensional Screening," mimeo.

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

[5] Cai, Hongbin, John Riley, and Lixin Ye (2007) "Reserve price signaling," *Journal of Economic Theory*, Vol. 135, pp. 253 – 268.

expected loan payments using the data of loans in the subprime market. They find the customers self-select the loan amount based off of their unobservable repayment ability.

[6] Cho, In-Koo and David M. Kreps "Signaling Games and Stable Equilibria," *Quarterly Journal of Economics*, Vol. 102, No. 2, pp. 179-222.

[7] Cohen, Alma and Liran Einav (2007) "Estimating Risk Preferences from Deductible Choice," *American Economic Review*, Vol. 97(3), pp. 745-788.

[8] Einav, Liran, Mark Jenkins, and Jonathan Levin (2011) "Contract Pricing in Consumer Credit Markets," forthcoming in *Econometrica.*

[9] Freedman, Seth and Ginger Jin (2010) "Learning by Doing with Asymmetric Information: Evidence from Prosper.com," mimeo.

[10] Gallant, A. Ronald, and Douglas W. Nychka (1987) "Semi-Nonparametric Maximum Likelihood Estimation," *Econometrica* , Vol. 55(2), pp. 363-390

[11] Gayle, George-Levi and Limor Golan (2011) "Estimating a Dynamic Adverse-Selection Model: Labour-Force Experience and the Changing Gender Earnings Gap 1968-97," forthcoming in *Review of Economic Studies.*

[12] Jenkins, Mark (2009) "Measuring the Costs of Ex Post Moral Hazard in Subprime Lending," mimeo.

[13] Matzkin, Rosa (1992) "Nonparametric and Distribution-Free Estimation of the Binary Threshold Crossing and the Binary Choice Models," *Econometrica*, Vol. 60(2), pp. 239-270.

[14] Jullien, Bruno and Thomas Mariotti (2006) "Auction and the Informed Seller Problem," *Games and Economic Behavior*, Vol. 56(2), pp. 225-58.

[15] Kim, Il Kyoo (2010) "Semiparametric Estimation of Signaling Games with Equilibrium Refinements," mimeo.

[16] Mailath, George J. (1987) "Incentive Compatibility in Signaling Games with a Continuum of Types," *Econometrica*, Vol. 55(6), pp. 1349-1365.

[17] Rigbi, Oren (2008) "The Effects of Usury Laws: Evidence from the Online Loan Market," mimeo.

[18] Paravisini, Daniel, Veronica Rappoport, and Enrichetta Ravina (2011) "Risk Aversion and Wealth: Evidence from Person-to-Person Lending Portfolios," mimeo.

[19] Spence, Michael (1973) "Job Market Signaling," *Quarterly Journal of Economics*, Vol. 87(3), pp. 355-374.

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

9 Appendix 1: Proof of Proposition 1

In this section we provide the proof of Proposition 1. We do so by proving the following two lemmas. The first lemma shows a property of the value function of the borrower with type φ when he obtains a loan with interest rate r, and the second lemma shows the proposition.

Lemma 2 $\frac{\partial}{\partial\varphi}V_1(r,\varphi)$ *is non-increasing in* r.

Proof. We show by induction that $\frac{\partial}{\partial r\partial\varphi}V_1(r,\varphi) \leq 0$. We first show that $\frac{\partial}{\partial r\partial\varphi}V_T(r,\varphi) \leq 0$ and $0 > \frac{\partial}{\partial\varphi}V_T(r,\varphi) \geq D'(\varphi)$. We then show that for any $\tau \leq T$, $\frac{\partial}{\partial r\partial\varphi}V_\tau(r,\varphi) \leq 0$, and $0 > \frac{\partial}{\partial\varphi}V_\tau(r,\varphi) \geq D'(\varphi)$ imply that the same condition holds for $t = \tau - 1$. First, for $t = T$,

$$\frac{\partial}{\partial\varphi}V_T(r,\varphi) = \frac{\partial}{\partial\varphi}\int \max\{u_T(r) + \varepsilon_T, D(\varphi)\}dF_{\varepsilon_T}(\varepsilon_T)$$

$$= \int \left(0 \times \mathcal{I}\{u_T(r) + \varepsilon_T \geq D(\varphi)\} + D'(\varphi)\mathcal{I}\{u_T(r) + \varepsilon_T < D(\varphi)\}\right) dF_{\varepsilon_T}(\varepsilon_T)$$

$$= D(\varphi)' \Pr\left(u_T(r) + \varepsilon_T < D(\varphi)\right) \geq D(\varphi)'$$

Also, it is easy to see $\frac{\partial}{\partial r\partial\varphi}V_T(r,\varphi) \leq 0$ since $\frac{\partial}{\partial r}Pr\left(u_T(r) + \varepsilon_T < D(\varphi)\right) > 0$.

Now, for $t < T$, assume $\frac{\partial}{\partial r\partial\varphi}V_{t+1}(r,\varphi) \leq 0$ and $0 > \frac{\partial}{\partial\varphi}V_{t+1}(r,\varphi) \geq D'(\varphi)$

$$\frac{\partial}{\partial\varphi}V_t(r,\varphi) = \frac{\partial}{\partial\varphi}\int \max\{u_t(r) + \varepsilon_t + \beta V_{t+1}(r,\varphi), D(\varphi)\}dF_{\varepsilon_t}(\varepsilon_t)$$

$$= \int \left(\begin{array}{c}\frac{\partial}{\partial\varphi}\beta V_{t+1}(r,\varphi)\mathcal{I}\{u_t(r) + \varepsilon_t + \beta V_{t+1}(r,\varphi) \geq D(\varphi)\}+ \\ D'(\varphi)\mathcal{I}\{u_t(r) + \varepsilon_t + \beta V_{t+1}(r,\varphi) < D(\varphi)\}\end{array}\right) dF_{\varepsilon_t}(\varepsilon_T)$$

$$= \begin{array}{c}\frac{\partial}{\partial\varphi}\beta V_{t+1}(r,\varphi)Pr\left(u_t(r) + \varepsilon_t + \beta V_{t+1}(r,\varphi) \geq D(\varphi)\right) + \\ D(\varphi)'Pr\left(u_t(r) + \varepsilon_t + \beta V_{t+1}(r,\varphi) < D(\varphi)\right)\end{array}$$

$$\geq D(\varphi)'$$

The last inequality holds since $\frac{\partial}{\partial\varphi}V_{t+1}(r,\varphi) \geq D'(\varphi)$. Again, it is easy to see $\frac{\partial}{\partial r\partial\varphi}V_t(r,\varphi) \leq 0$.

From induction we conclude $\frac{\partial}{\partial r\partial\varphi}V_1(r,\varphi) \leq 0$. Following Lemma XXX, we then prove the borrower's utility function at deciding the reserve interestrate s satisfies the single crossing property in the next lemma. ∎

Lemma 3 *Let* $G(s,\varphi) = \Pr(s)\int V_1(r,\varphi:\theta)f(r|s)dr + (1-\Pr(s))\lambda(\varphi)$. *Then,* $\frac{\partial}{\partial s\partial\varphi}G(s,\varphi) < 0$.

Proof. First, let us consider the second term. Note that $\frac{\partial}{\partial s\partial\varphi}(1-\Pr(s))c(\varphi) = -\Pr'(s)\lambda_\varphi(\varphi) < 0$. This is because $\frac{d\Pr(s)}{ds} > 0$ and $\frac{\partial\lambda(\varphi)}{\partial\varphi} > 0$ by assumption. Second, we consider first term. If $\frac{\partial}{\partial s\partial\varphi}\Pr(s)\int V_1(r,\varphi:\theta)dF(r|s)$ is also negative, the statement is proved. To show this, note that for $s_0 < s_1$, $F(r|s_1)$ first-order stochastically dominates $F(r|s_0)$. Hence if $\frac{\partial}{\partial\varphi}V_1(r,\varphi)$ is non-increasing in r, then $\int \frac{\partial}{\partial\varphi}V_1(r,\varphi)dF(r|s_0) > \int \frac{\partial}{\partial\varphi}V_1(r,\varphi)dF(r|s_1)$ for any s_0 and s_1 s.t. $s_0 < s_1$. This implies that $\frac{\partial}{\partial s\partial\varphi}\Pr(s)\int V_1(r,\varphi)dF(r|s) < 0$ if $\frac{\partial}{\partial\varphi}V_1(r,\varphi)$ is nonincreasing in r. Due to Lemma XXX, we have known it is non-decreasing in r. Thus, we complete the proof. ∎
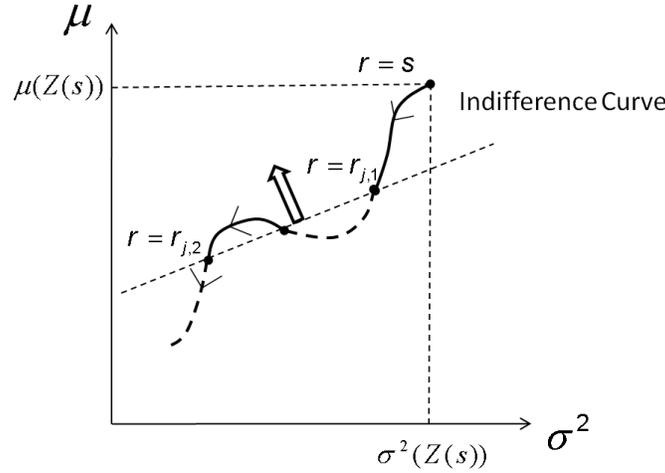
10 Appendix 2: Discussion on Lender's Behavior (incomplete)

We briefly discuss the lender's optimal bidding strategy when the mean-variance curve of a listing intersects with the lender's indifference curve more than once. Let $I_0 = [0, r_1]$, $I_1 = [r_1, r_2], \cdots I_M = [r_M, s]$ be intervals so that $U_0(r)$ is largest for all $r \in I_k$, where $q(k) \in \{\$0, \$50, \$100, \$200\}$ and $U_0(r) = \varepsilon_0$ for all r.

11 Appendix 3: Identification of Borrower-Side Model (incomplete)

12 Appendix 4: Identification of Lender-Side Model

In this Appendix, we consider identification of F_ε. Consider $P_{200}(\mu, \sigma)$,[30]

$$P_{200}(\mu, \sigma) = \Pr(A_j < \frac{c(100) - c(200) + 100\mu}{30000\sigma^2} \wedge \varepsilon_0 < 200\mu - A_j(200\sigma)^2 - c(200)).$$

Suppose we take a μ so that $c(100) - c(200) + 100\mu > 0$ ($\Leftrightarrow \mu > \frac{c(200) - c(100)}{100}$). Now consider holding μ constant and taking the limit as σ goes to zero. Then $P_{200}(\mu, \sigma) \to \Pr(\varepsilon_0 < 200\mu - c(200))$. Because we can move μ in the region $\mu > \frac{c(200) - c(100)}{100}$,[31] the distribution of ε_0 is identified for at all $t > c(200) - 2c(100)$.

Now consider $P_{100}(\mu, \alpha)$:[32]

$$P_{100}(\mu, \sigma) = \Pr(\frac{c(100) - c(200) + 100\mu}{30000\sigma^2} < A_j < \frac{-c(100) + 50\mu}{7500\sigma^2} \wedge \varepsilon_0 < 100\mu - A_j(100\sigma)^2 - c(100)).$$

---

[30] This is true as long as μ is "big" enough, i.e. $\frac{c(50) - c(100) + 50\mu}{7500\sigma^2} > \frac{100\mu - c(200) + c(100)}{30000\sigma^2}$ ($\Leftrightarrow \mu > \frac{-c(200) + 5c(100)}{100}$)

[31] $c(200) > 3c(100)$ implies $\frac{c(200) - c(100)}{100} > \frac{-c(200) + 5c(100)}{100}$.

[32] This is true as long as $\frac{c(50) - c(100) + 50\mu}{7500\sigma^2} > \frac{100\mu - c(200) + c(100)}{30000\sigma^2}$ ($\Leftrightarrow \mu > \frac{-c(200) + 5c(100)}{100}$)

Again, take a μ so that $-c(100) + 50\mu > 0$ and $c(100) - c(200) + 100\mu < 0$ ($\Leftrightarrow \frac{c(100)}{50} < \mu < \frac{c(200)-c(100)}{100}$). As before, we take $\sigma \to 0$, while holding μ constant. Then $P_{100}(\mu, \sigma) \to \Pr(\varepsilon_0 < 100\mu - c(100))$. Because we can move μ in the region $\frac{c(100)}{50}, < \mu < \frac{c(200)-c(100)}{100}$ [33] the distribution of ε_0 is identified for all $t \in [c(100), c(200) - 2c(100)]$.

Likewise, consider $P_{50}(\mu, \alpha)$,[34]

$$P_{50}(\mu, \sigma) = \Pr(A_j > \frac{-c(100) + 50\mu}{7500\sigma^2} \wedge \varepsilon_0 < 50\mu - A_j(50\sigma)^2).$$

As before, take a μ so that $-c(100) + 50\mu < 0$ ($\Leftrightarrow \mu < \frac{c(100)}{50}$). Then $P_{50}(\mu, \sigma) \to \Pr(\varepsilon_0 < 50\mu)$. Because we can move μ in the region $\frac{c(100)}{50} > \mu$ ($> \frac{-c(200)+5c(100)}{100}$), the distribution of ε_0 is identified for at all $\frac{-c(200)+5c(100)}{2} < t < c(100)$

Lastly, consider $P_{50}(\mu, \alpha)$,[35]

$$P_{50}(\mu, \sigma) = \Pr(A_j > \frac{150\mu - c(200)}{37500\sigma^2} \wedge \varepsilon_0 < 50\mu - A_j(50\sigma)^2).$$

If we take a μ so that $150\mu - c(200) < 0$ ($\Leftrightarrow \mu < \frac{c(200)}{150}$). Then $P_{50}(\mu, \sigma) \to \Pr(\varepsilon_0 < 50\mu)$. Because we can move μ in the region $\mu < \min\{\frac{-c(200)+5c(100)}{100}, \frac{c(100)}{150}\}$ ($\Leftrightarrow \mu < \frac{-c(200)+5c(100)}{100}$), the distribution of ε_0 is identified for at all $t < \frac{-c(200)+5c(100)}{2}$. This identifies the distribution of ε_0, F_{ε_0}.

13 Appendix 5: Computation of the Counterfactual

In this section, we describe the computation procedure of the counterfactual experiments. First, we show how to compute the credit supply curve for the case that posting a reserve interest rate is allowed. The procedure is as follows:

1. For each listing i with observable characteristics x_i and reserve interest rate s_i, compute her credit worthiness, φ_i, by solving FOC in equation (3).

2. Given θ_B, simulate each potential borrower i's repayment decision if he were funded at the contract interest rate of r for any $r \in [0, 0.36]$. Repeat it many times to obtain its mean return and variance, $(\mu(r, x_i, s_i), \sigma^2(r, x_i, s_i))$.

3. Given θ_L, simulate the number of potential bidders, $\widetilde{N}$, drawn from F_N, and each bidder's risk attitude, A_j, drawn from F_A. For each potential bidder j, obtain her optimal amount choice $q_j^*(r)$ by solving the following problem for any $r \in [0, 0.36]$,

$$\max_{q_j \in M} \left\{ \left[q\mu(r, x_i, s_i) - A_j(q\sigma(r, x_i, s_i))^2 \right] - c(q_j; \theta_L) \right\}.$$

We repeat this S times. Denote potential borrower j's optimal choice at interest rate r for s-th simulation draw by $q_{js}^*(r)$.

---

[33] $c(200) > 3c(100)$ implies $\frac{c(200)-c(100)}{100} > \frac{c(100)}{50} > \frac{-c(200)+5c(100)}{100}$.

[34] This is true as long as $\frac{c(50)-c(100)+50\mu}{7500\sigma^2} > \frac{100\mu-c(200)+c(100)}{30000\sigma^2}$ ($\Leftrightarrow \mu > \frac{-c(200)+5c(100)}{100}$)

[35] This is true as long as $\frac{100\mu-c(200)+c(100)}{30000\sigma^2} > \frac{c(50)-c(100)+50\mu}{7500\sigma^2}$ ($\Leftrightarrow \mu < \frac{-c(200)+5c(100)}{100}$)

4. For each simulation draw, compute the total supply of credits by $S_i(r) = \frac{1}{S} \sum_{s=1}^{S} \sum_{j=1}^{\tilde{N}} q_{js}^*(r)$ for any $r \in [0, 0.36]$.

Second, we show how to compute the credit supply curve for the case that posting a reserve interest rate is prohibited. The procedure is as follows:

1. For all observed listings, solve FOC in equation (3) to obtain each potential borrowers' credit worthiness, φ_i. This gives the empirical joint distribution of unobserved type φ and listing characteristics x. From this empirical distribution, estimate a conditional distribution of φ given x, or $F_{\varphi|X=x}$.

2. Fix any listing characteristics x,[36] and compute the threshold value of the borrower's type for being indifferent between obtaining a loan with interest rate $r \in [0, 0.36]$ and leaving the market. We obtain it by solving equation

$$V_1(r, \varphi, x; \theta_B) = \lambda(\varphi; \theta_B)$$

with respect to φ. Denote the solution by $\varphi^*(r, x; \theta_B)$. Borrowers whose unobservable credit worthiness φ is less than $\varphi^*(r, x; \theta_B)$ leave the market.

3. Compute the conditional distribution of type φ by $F_{\varphi|X=x}(\varphi|\varphi^*(r, x; \theta_B) > \varphi)$.

4. Simulate many borrowers with listing characteristics x and assign a value of φ, which is drawn from $F_{\varphi|X=x}(\varphi|\varphi^*(r, x; \theta_B) > \varphi)$, and simulate the borrowers' repayment behavior given φ, r and x.

5. Repeat step 4 many times to obtain the mean return and the variance, $(\mu(r, x), \sigma^2(r, x))$.

6. For each value of $(\mu(r, x), \sigma^2(r, x))$, simulate the number of potential bidders and their risk aversion parameter as we did in Step 3 of the procedure above. Given the simulated values, calculate the amount of money supplied to the $(\mu(r, x), \sigma^2(r, x))$ pair. That gives a credit supply curve.

Finally, the computation procedure of estimating the efficient market outcome is as follows:

1. Assume the lenders can see borrowers type.

2. Assign a (arbitrary or draw from) single value of φ.

3. For any $r \in [0, 0.36]$, φ, and x, simulate borrowers' repayment behavior many times to calculate $(\mu(r, x, \varphi), \sigma^2(r, x, \varphi))$.

4. Calculate the interest rate that equates $V_1(r, \varphi, x; \theta^B)$ and $\lambda(\varphi; \theta^B)$, denoted by $r^*(\varphi, x; \theta^B)$, and also calculate the signal value of $s(\varphi, x; \theta^B)$.

5. Simulate potential bidders and see if the listing will be funded, what the realized interest rate would be for case where the reserve rate equals to $r^*(\varphi, x; \theta^B)$

---

[36] We set the requested amount to be $\$10,000$ and the debt to income ratio to be 0.3.

6. integrate over φ

7. Do the same calculation for case where the reserve rate equals to $s(\varphi, x; \theta^B)$ and compare the result.