Search and herding effects in peer-to-peer lending:

Evidence from Prosper.com

Efraim Berkovich

The University of Pennsylvania

November 9, 2009

Abstract

I examine loan data from Prosper.com—a website which allows borrowers to post loans and for lenders to bid on those loans. The Prosper market somewhat resembles the theoretical model of search, herding, and crowding in a large market described in Berkovich and Tayon (2009). That model predicts that regions of the market with high and low prices have high variance in the difference between price and true value. These regions are where private information provides excess returns. I find that low value loans do show higher variance of the difference between price and value. I also find that the Prosper market is not in equilibrium since areas of high prices in the Prosper loans market provide excess returns even after accounting for risk-aversion due to variance of returns.

Efraim Berkovich can be reached at efraim@econ.upenn.edu. I thank Petra Todd and Jeremy Tobacman for their advice on this work.

1 Introduction

The number of people eating in a restaurant is often a good indicator of the quality of the food. This intuitive approach to determining the value of a good or asset makes sense when some information about the good or asset is private—the crowd may know something. Even though eating at a less crowded restaurant may be more enjoyable, the crowd provides a signal of higher quality.

Berkovich and Tayon (2009) develop a model of a large market with public and private information. Payoffs to buyers in the model increase in the underlying value of the asset and decrease in the number of other buyers buying the asset. Asset values are initially unknown but can become known either by acquiring private information at a cost or by observing other buyers' behavior. With or without private information, buyers make a decision to purchase a fixed amount. Under this set-up, full information discovery through prices may not occur at high and low asset values except in the limit of time. A key testable prediction of the model is that high and low value assets will have higher variance of the difference of price and true value.

In this paper, I use data from an online market for peer-to-peer lending, Prosper.com, to examine the evidence for the theoretical model. The Prosper market has been described as an eBay for loans: borrowers post loan requests and lenders bid on the loans, with a bid consisting of an amount to lend and an interest rate Lenders can learn from the bidding history on a loan and use this information to form their opinion of the value of the loan. The Proper dataset contains information on loan listings and bidding history, as well as information on loan defaults.

Using the bidding history and loan listings, I calculate the estimated value of the loan. The residual difference between the actual price of a loan and the estimated value is the "adjusted price." In the loan market, the "price" of a loan to a lender is a combination of the interest rate and the default rate. The "adjusted price" of a loan is the difference between the interest rate and

the expected interest rate conditional on the characteristics of the loan, which takes into account the probability of default. To find the true value of a loan, I use the loan outcome information, which was obviously not available to the lenders at the time of bidding. I compare the adjusted price with the true value. According to the model, on average, this difference should be the same across all loans but with higher variance for low and high quality loans.

I find some evidence in the Prosper market that lenders respond to bidding history as loans with a lot of bids tend to receive more bids. Looking at variance versus loan returns, I find that the prediction of the model holds for low value loans–the adjusted prices exhibit higher variance for low return loans. I do not find higher variance for higher value loans.

The Prosper market does not appear to be in equilibrium. Lenders excessively avoid the high adjusted priced loans and may also excessively avoid the lower adjusted price loans (the evidence is unclear). These loans offer higher expected returns and lower variance of returns. If, as in the model, uninformed lenders avoid low and high adjusted price loans on the basis of perceived informational disadvantage, then they do it at rates in excess of equilibrium. Presumeably, this effect disappears as market learning evolves. An alternative explanation is that lenders behave in a contrarian way, over-bidding low adjusted price loans and under-bidding high adjusted price ones.

1.1 Related literature

The newness of the Prosper platform along with the wealth of publicly available data provided by Prosper has lead to a large number of papers on different aspects of the market. Pope and Sydnor (2008) and Theseira (2008) study discrimination effects on Prosper. Friedman and Jin (2008) look at the impact social networks have on alleviating information problems on Prosper loans. Iyer *et al* (2009) find that lenders effectively use the richer data set available on Prosper listings, and while

lenders mostly rely on standard banking variables to draw inferences on creditworthiness, they also use the listing-specific sources of information in their screening process, especially in the lower credit categories.

Learning from other agents'actions in markets has extensive precedent in the literature. Banerjee (1992), (Bikhchandani et al. 1992) and (Welch 1992) describe models of herding where agents obtain a private signal about a choice of options and also observe the prior actions of a number of other agents who faced the same decision. They find that agents begin to ignore their private signal, forming an informational cascade. In an elaboration, Debo, Parlour, and Rajan (2008) describe a model where agents obtain a private signal about a firm's quality and also observe the length of a queue of prior randomly arriving agents. The agents infer the firm's quality from the length of the queue. Vives (1996) notes that informational cascades depend on discrete action spaces, though Huck and Oechssler (1998) find a counter-example. Related to the crowding aspect of the model, there exists a large literature on matching with capacity constraints. For example, Peters (1984) presents a directed search model with rationing of buyers by sellers. Burdett, Shi, and Wright (2001) discuss equilibrium pricing and matching in a similar model. Lester (2008) describes a search model where the choice to obtain costly private information is endogenized.

Herding and learning has been well-studied in relation to asset markets. For instance, Lang and Nakamura (1990) offer a dynamic model of learning in credit markets where a decrease in the number of traders decreases public information and increases risk thereby further driving out traders. They find that market prices have greater volatility than the underlying shocks because of this feedback. The effect of traders being driven out of market segments with low public information parallels effects in the model under examination in this paper.

Recent work has studied herding effects on online and experimental platforms. Simonsohn and

Ariely (2008) find an interesting equilibrium on eBay: bidders choose auctions with more bids, so sellers start the auction at lower prices. Cipriani and Guarino (2005) find contrarian behavior by agents in a laboratory financial market. Drehmann *et al.* (2007) perform experimental work on various herding models where they look at behavior under a variety of payoff externalities.

2 Prosper market

The Prosper website marketplace has been in operation since late 2005 and operates a peer-to-peer credit market providing fixed 36-month term unsecured loans from $1,000 to $25,000. Some have argued that the peer-to-peer lending business model creates economic gains through financial disintermediation by linking individual borrowers and lenders. Peer-to-peer credit markets have grown rapidly in the U.S. and U.K. since the creation of the first market, Zopa.com, in the U.K. in 2005.

All borrowers and lenders are U.S. residents. Borrowers must pass identity and credit verification checks and link a bank account electronically to Prosper.com for loan origination and repayments. Lenders pass identity verification for tax reasons and link a bank account to transfer funds to Prosper.com for the purpose of bidding.

Borrowers create a loan request listing on Prosper by choosing a specific amount to borrow and a maximum reservation interest rate and adding a written description, possibly with images. Listings display information from the borrower's credit report and other verified information by Prosper. In addition, Prosper assigns and displays a credit grade of AA (the highest) to HR (High Risk) based on 40-point bands of the borrower's Experian credit score. Prosper provides guidance on interest rates and historical default data to borrowers and lenders based on these grades. On the high end, loan interest rates are limited by applicable state usury laws.

Lenders search the listings for loans on which they wish to bid. Funding takes place through a descending uniform price auction, where lenders' price-quantity bids form a supply curve of funding for a loan. The minimum quantity bid is $50, and the median bid is $50. Since demand (loan value) is fixed, the auction clears at the interest rate charged by the marginal losing bidder. Bidding is partially open, as lenders always observe the number, and dollar quantities, of all bids. However, interest rates (prices) of winning bids are hidden, and as the auction clears, interest rates on the losing bids are revealed.

Borrowers select the length of time a listing is posted, pick a reservation highest rate they are willing to pay, write a text description, and optionally add pictures to the listing. Borrowers also have the option to choose whether the loan funds immediately once the requested amount has been bid.

3 Model of herding and crowding in a large market

Work by Berkovich and Tayon (2009) presents a simple overlapping generations model of a large market with incomplete information and publicly visible actions. Payoffs to agents in the model increase in the underlying value of the asset and decrease in the crowding of other agents buying the same asset. The market contains assets whose value can only be known directly by costly information acquisition or indirectly by observing the behavior of agents in the prior period. Initially, each buyer has no private information about the value of any asset, but he can choose to acquire costly private information. With or without private information, buyers make their purchase decision and acquire the asset. Because, as in many herding models, the action space of buyers is discrete, full value discovery does not immediately occur even when the number of buyers is infinite.

In this dynamic model with an infinite number of buyers and assets, buyers see the purchase

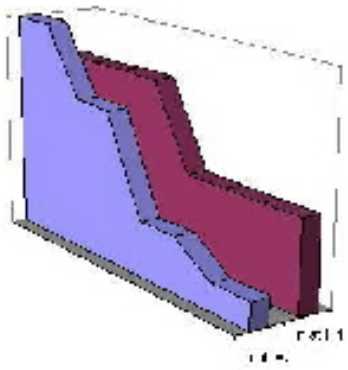

Figure 1: Distribution of buyers through time in the model. The x-axis the the quality of the asset and the y-axis is the number of buyers (price).

decisions of the previous period buyers as a public signal of asset quality. In equilibrium, higher number of buyers generally corresponds to higher quality assets, and based on this public signal, the market endogenously segments into areas of known value and unknown value. Figure 1 shows the distribution of buyers in time in the model. The number of buyers corresponds to the price of the asset as more buyers means lower payoff. In regions of unknown value, informed buyers drive out uninformed traders, reducing trade and forcing uninformed buyers into regions of known quality—a pattern that superficially resembles risk-aversion, though buyers are risk-neutral. As time goes on, public information grows, in contrast to many herding models where information is lost. This market becomes more efficient over time as fewer buyers incur search costs each period; the public signal for each asset shows the true value of the asset.

The Prosper market and the model are both large markets where assets have variable quality which is not visible to the buyer. In Prosper, finding valuable loans involves a costly time and or effort investment by lenders. The alternative (uninformed) Prosper investment strategy is to buy

lend across all loans. The table below summarizes some of the differences and similarities between the Prosper market and the model.

	Model	*Prosper*
Assets	Value distribution known.	Value distribution approximated from historical averages.
Timing	Discrete periods	Continuous time. loans listed over a number of days.
Size	Infinite buyers, assets	Finite lenders, loan interest rates limited by law.
Crowding	Lowers payoffs. Defined by number of buyers.	More bidders lower the listing's interest rate (in general).
Public information	Number of buyers.	Multi-variate public listing and bidding history.
Private information	Costly	No actual private information (in general), reading text description or analyzing information is costly private information acquisition.
Agent types	Informed and uninformed	Unknown distribution of types.

In the loan market, the price of a loan is a combination of the interest rate and the default rate. Since listings identical in their public information should have the same expected default rate and therefore price, the adjusted price is defined by the currently bid interest rate on the listing subtracted from the expected interest rate. Therefore, when y_i is the interest rate for listing i, the adjusted price is given by the residual

$$\varepsilon_i = E\left[y_i | x_i\right] - y_i$$

where x_i is the public information on listing i. Since low interest rates mean a loan is expensive, a positive residual implies low adjusted prices and a negative residual implies high adjusted prices.

The model predicts that assets with high and low true values are priced correctly only on average; the prices of these assets show variance from the true value. On the other hand, the price of medium value assets reflects their true value with little variance.

I find evidence for high variance in adjusted prices for low quality loans, though not for high quality loans. Analysis finds the Prosper market is not in equilibrium as it turns out that high adjusted price loans have excess returns even after accounting for risk-aversion due to variance in returns. In equilibrium for many models, all publicly identifiable groups of loans should provide the same (risk-adjusted) payoffs. In the model, uninformed lenders avoid the extreme value regions because of information asymmetry. It may be that this effect is pushed past the equilibrium avoidance rate when lenders see that prices are relatively high.

4 Estimating the loan rate

As a first step to finding the price of a loan, I estimate a linear regression model to predict the loan interest rate from publicly observable variables. The borrower-specific factors include standard measures of credit-worthiness such as the credit rating of the borrower, whether the borrower is a homeowner, borrower's debt to income ratio, borrower's location, and so on. These variables tend to be colinear. Prosper-specific factors include whether the borrower belongs to any Prosper groups (these are associations of users created by users on Prosper) and the number of prior loans generated on Prosper. Credit grades are assigned by Prosper (from AA to High Risk) based on 40 point bands of the Experian credit score. The debt to income ratio is the borrower's credit bureau reported debt divided by verifiable income. Income verification is usually done through paystubs,

so there exists potential for abuse and misreporting. Prosper caps the value of the debt to income ratio at 10.01, so I do not have a full range of data there. The borrower's location was grouped by U.S. state.

The listing-specific factors include the size of the loan requested, the highest rate the borrower is willing to pay, the number of days to display the listing on Prosper, whether the loan get issued immediately upon being funded, and the listing text and images. It is likely the specific wording and images in the listing has an impact on the lenders' assessment of loan quality. Work by Pope and Sydnor (2008), Theseira (2008), and, to a smaller extent, Iyer *et al.* (2009) indicates that, in fact, images and text have a significant impact on loan rate. But Prosper does not verify the images and wording of the listing, so that information is effectively cheap talk. I do not measure the effect of these listing-specific factors other than to check whether the listing contains images.

Environment related factors include the prime rate, date, and variables describing the Prosper market's loan supply and demand. The date is described by binary variables for each quarter from 2006 to the fourth quarter of 2008 and by the day of the week the listing ends. The supply and demand variables are the average listing loan request, the average bid size, and the bid to listing request ratio.

Given the newness of the Prosper lending platform, one expects that there would be some learning by borrowers and lenders in this market. Certain borrower traits may become more or less valuable to lenders. In order to see this effect, I add a binary variable denoting whether the listing is from before November 1, 2007. This division breaks the sample roughly in half and also somewhat coincides with the onset of the credit crisis.

The data sample contains 153,162 listings from between May 2006 and October 2008 resulting in 23,427 loans originated. Since many of the listings did not result in loans and therefore loan interest

rates, it makes sense to predict loan rate based on a tobit censored regression where the censored observations are those listings that did not lead to loans. Selected coefficients are displayed in the table below.

Variable	*Coefficient (Std.error)*
log(AMOUNTREQUESTED)	0.031742 (0.000379)
BORROWERMAXIMUMRATE	0.694990 (0.005630)
INGROUP	-0.001349 (0.001024)**
INGROUP*PRIORNOV2007	-0.019265 (0.001188)
HASIMAGES	-0.012736 (0.000712)
HASIMAGES*PRIORNOV2007	-0.003266 (0.000910)
FUNDIMMEDIATELY	0.005015 (0.000468)
CREDITE	-0.023774 (0.000781)
CREDITD	-0.058848 (0.000916)
CREDITC	-0.086966 (0.001101)
CREDITB	-0.112418 (0.001357)
CREDITA	-0.125168 (0.001527)
CREDITAA	-0.138453 (0.001675)

** Not significant at 95%

The results are generally as expected. The size of the requested loan increases the loan rate. Borrowers have a good sense of the loan rate they can get, as seen by the size of the coefficient on BORROWERMAXIMUMRATE, the highest rate loan the borrower accepts. Higher credit grades result in lower loan rates. Those listings with the fund immediately option had approximately 0.5% higher loan rate than comparable listings.

Looking at Prosper specific factors such as group membership and images on the listing before

and after November 1, 2007, I find that having images in a listing reduced loan interest rates by about 1.3% after the cutoff and by 1.6% prior to the cutoff. Clearly, lenders value the information contained in images. Because images are unverifiable cheap talk, one might expect that borrowers would provide whatever images lenders seem to value. So, either borrowers have not yet learned to take advantage of this fact, or many borrowers have an aversion to lying.

The effect of group membership decreased substantially. Before the cutoff, group membership reduced the loan rate by about 1.9%. Afterwards, lenders ignored group membership entirely. About 60.4% of loans prior to the cutoff had group membership while only 15.5% of loans after the cutoff did.

I report other results: Listings that closed on Thursday or Friday had a bit less than 1% lower interest rate than comparable listings that closed on Sunday. Loans that originated prior to November 1, 2007 had rates approximately 1.5% lower, though this variable is colinear.with the quarterly variables.

5 Herding

As evidence of the model in the Prosper market, I look at herding effects on loan rate. The model states that lenders look at other lenders' bids as a signal of loan quality. I use the censored regression and add variables on bidding history. Looking at the number of bids one day prior to the day the listing closed ($NumBids_{t-1}$) and the number of bids added on the last day ($\Delta NumBids_{t,t-1}$), I find that both variables have negative coefficients: $-0.000464(6.32 \times 10^{-6})$ for $\Delta NumBids_{t,t-1}$ and $-0.000280(4.40 \times 10^{-6})$ for $NumBids_{t-1}$, both significant at 99%. These numbers seem to indicate that lenders respond to the number of other bids by raising their estimate of loan quality and hence bidding down the interest rate. An additional 10 bids on the last day lowers the interest rate by

about 0.5% whereas an additional 10 bids on the penultimate day only lowers it by about 0.3%, perhaps implying that later bidders tended to value the loan more.

The dataset provides the status of the loan (*e.g.* delinquent, late, paid, *etc.*), allowing me to sum up all the loan payments and estimate the return on the loan and hence its realized value. Although Prosper temporarily halted listing activity at the end of 2008, the dataset provides loan status through March 2009. One issue with my estimation of loan return is that I do not have recovery rates for delinquent loans, so I assume a zero recovery rate on all delinquent loans. For the analysis that follows, the actual recovery rate is not crucial so long as there is not a correlation between the residual and the recovery rate.

As another check for herding effects, I use the estimated loan return to see if it is correlated with the number of bids. A linear regression to predict the loan return based on the public variables and the number of bids on the listing estimates the coefficient for the number of bids to be -6.67×10^{-5} (2.68×10^{-5}) (significant at 95%). Although the effect is small, the negative sign is consistent with an "over-bidding" effect—that bidders believed loans with more bids were more valuable.

Lenders may also be looking at the average bid size as an indication of the value other bidders have for the listing. Looking at the average bid size for listings with at least 10 bids, I find that, while it is significant at 99% and negative, the coefficient is extremely small and has little impact on predicting the interest rate. Lenders tend to bid fixed amounts, typically $50. This behavior (while not an institutional feature of the Prosper market) jibes with the discrete action setup of the model.

6 Adjusted prices

The adjusted price of the loan is the difference between the expected interest rate on a particular loan and the actual rate. If lenders expect a particular interest rate and find that the actual interest rate is higher, it is an indication that not enough lenders have bid on the loan to drive down the rate. This residual is given by $\varepsilon_i = E[y_i|x_i] - y_i$, as explained above, and indicates low adjusted prices when positive and high adjusted prices when negative.

The model predicts that for a new market, after one period, the areas of low and high value contain those assets whose values are not well known from the public information and where private information about those assets is most valuable. The region of medium value contains assets whose value can be determined from the price level (interest rate in our case). For those lenders who have no private information, buying loans in the extreme adjusted price regions is dangerous since they are at an informational disadvantage and do not know why the loans are priced so differently from their expected price. Of course, in the theoretic model, uninformed traders buy just enough in the extreme regions so that expected payoffs are equalized across all loans for uninformed traders.

In the model, lower quality assets have lower prices. In the Prosper market, each group of assets which have the same public information (other than bidding history) corresponds to a model market. The whole Prosper market consists of a large number of these market models. Therefore, lower adjusted price regions of the Prosper market correspond to, in the eyes of lenders, lower quality loans *ceteris paribus*. One would not necessarily expect that lower adjusted price loans be of lower credit grade, however, merely lower quality within any credit grade.

The distribution of residuals for the listings which resulted in loans has a mean of 0.056353, a median of 0.054255, and a standard deviation of 0.045262. I look at the adjusted prices versus the actual returns. I sort the loans by return and find the standard deviation of the adjusted prices of

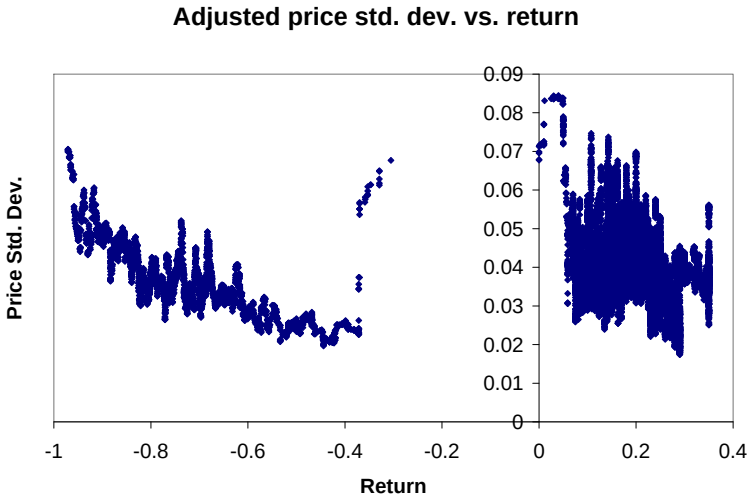

Figure 2: Standard deviation of adjusted prices plotted vs. return.

the adjacent 60 other loans. Results are graphed in Figure 2.

The model predicts higher variance of prices for low and high value assets. From the graph, it appears that lower value loans do have higher variance in adjusted price. The situation for higher value loans does not show significant differences in variance across returns. Looking at average standard deviations by region backs up this assessment. However, as Prosper loans are limited in interest rate (highest rate is 35%) and because the full domain of all possible loans is not realized, it may be the case that there are not enough high value loans in the sample to be easily observable.

I plot the residual versus the estimated loan return and smooth the displayed loan returns by averaging over a region of 60 loans with adjacent residual values. This region is also used to find a variance to loan returns for loans with adjacent residual values.

The shape of the graph in Figure 3 appears to indicate that higher and lower priced loans tend to have higher returns. Moreover, the variance of returns appears to be lower or, at worst,

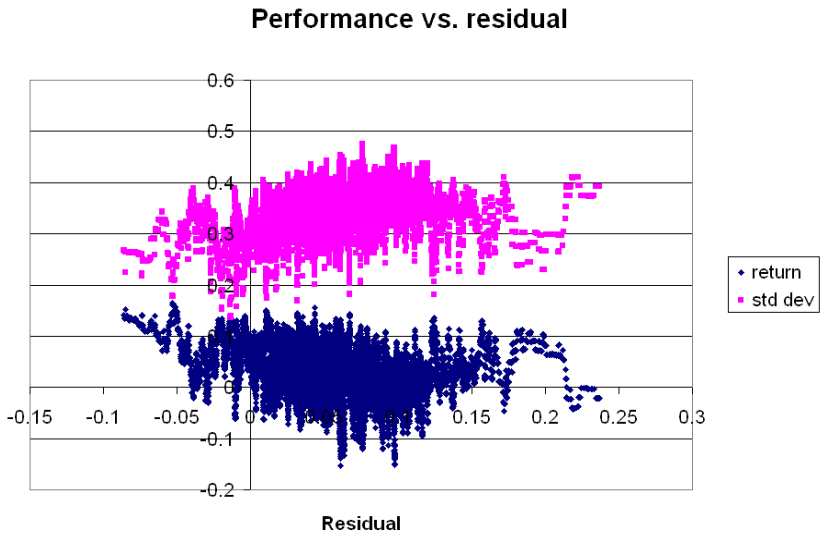

Figure 3: Smoothed estimated loan return and loan return standard deviations plotted vs. residuals. Increasing residual implies lower adjusted price.

flat for those regions so that a risk-averse lender receives even higher payoff to loans from those regions. The following table summarizes characteristics of loans with residuals in high and low adjusted price regions. I define the regions to each be somewhat less than 10% of the total sample and approximately one standard deviation away from the mean. While the divisions are arbitrary, moving the cutoffs slightly does not have a big effect on the results. I ignore the extremely low adjusted price loans (high residuals) as they represent a small number of loans with high variance.

	Low price	*Med. price*	*High price*
Residuals	0.12 to 0.25	0.0 to 0.12	< 0.0
Num. loans	1788	19583	2036
Mean return	0.043836	0.023489	0.072190
Return std. dev.	0.330648	0.351725	0.286297
Credit grades	AA: 4.7% A: 5.3%	AA: 11.1% A: 11.3%	AA: 27.9% A: 17.2%
	B: 10.5% C: 21.5%	B: 14.6% C: 19.4%	B: 18.4% C: 18.7%
	D: 26.1% E: 11.7%	D: 17.6% E: 12.4%	D: 11.1% E: 5.0%
	HR,NC: 20.2%	HR,NC: 13.6%	HR,NC: 1.7%

Even though lower adjusted price loans should not necessarily be lower credit quality, it seems that they are. The low priced area has lower credit quality while the high priced area has higher credit quality. It may be that lenders are willing to overpay (in the sense of paying more than the expected return) for higher credit quality because higher grades have lower variance of returns and lenders are risk-averse.

One may ask if excess returns are due to varying credit quality in the described regions, especially since the credit crisis which began in 2007 may have increased defaults amongst lower credit borrowers. An issue with that explanation is that it does not account for the higher returns to low adjusted price, low credit quality borrowers. Nevertheless, to check for any such effect, I look at returns by credit quality. The sample yields the following estimated returns by credit grade:

Credit grade	*Mean return (std. dev.)*	*Num. loans (%)*
AA	0.044420 (0.220841)	2821 (12.0%)
A	0.033316 (0.285435)	2663 (11.4%)
B	0.024248 (0.330540)	3425 (14.6%)
C	0.021391 (0.355854)	4569 (19.5%)
D	0.038197 (0.368976)	4159 (17.8%)
E	0.041492 (0.388010)	2747 (11.7%)
HR and NC	0.006270 (0.405719)	3043 (13.0%)

The expected return to lending with uniform probability to any funded loan is 2.93%. The expected return, based only on credit grade distribution, for the low priced area is 2.71%, for the medium priced area is 2.91%, and for the high priced area is 3.3%. Therefore, since returns in the high and low priced regions deviate significantly from returns based solely on credit quality, it seems the credit quality explanation is not the whole story.

I evaluate the returns to adjusted price by time-period, splitting the sample between loans originated prior to November 1, 2007 and those after.

	Return by price region (num loans)		
Time period	*Low*	*Medium*	*High*
Before Nov 2007	-6.05% (412)	-1.92% (11330)	3.55% (1081)
After Nov 2007	7.51% (1376)	8.22% (8253)	11.38% (955)

This data provides an explanation for the higher returns to the lower priced region—there are more lower priced loans with higher returns after November 2007.

I repeat the residual analysis with only the after November 2007 sub-sample. I estimate a new censored regression and plot residuals versus estimated return.

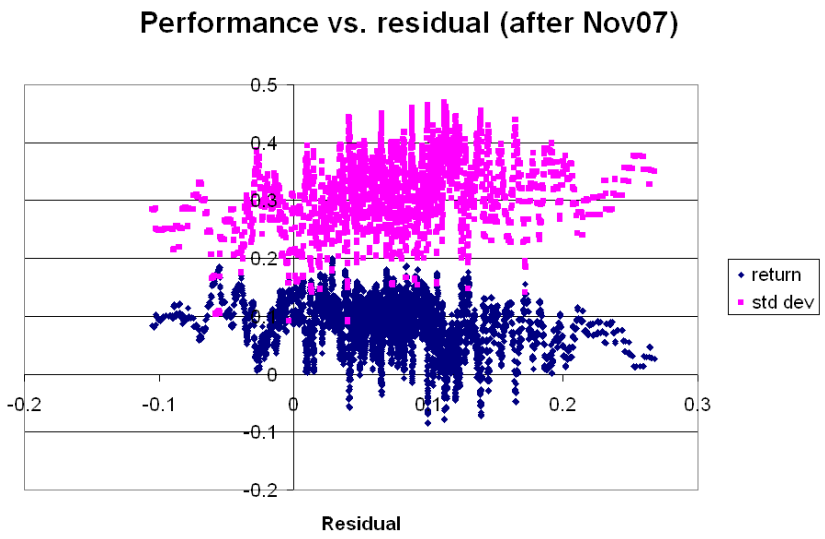

Figure 4: Smoothed estimated loan return and loan return standard deviations plotted vs. residuals for loan originated after Nov. 2007. Increasing residual implies lower adjusted price.

The shape of the graph in Figure 4 is similar to that for the overall sample. There are 10602 loans in the sub-sample and the residuals have a mean of 0.076106, a median of 0.076736, and a standard deviation of 0.060966. The statistics grouped by price regions are in the table below.

	Low price	*Med. price*	*High price*
Residuals	0.145 to 0.24	0.0 to 0.145	< 0.0
Num. loans	1120	8333	1073
Mean return	0.071310	0.084448	0.097991
Return std. dev.	0.319251	0.321390	0.272418

In this sub-sample, the low adjusted price region does not have higher returns than the medium price region while the higher price region does. These results echo the initial assessment above, that the time period is responsible for the higher returns to the low price region in the overall sample. However, the effect may still exist as the shape of the residuals graph roughly has a smile shape.

Given the relatively low number of sample points, a small number of loans in default significantly changes the mean return. In this sub-sample, the region returns are more sensitive to the cutoffs chosen. Nevertheless, the evidence for lenders avoiding low priced regions is not there.

An alternative explanation may be that lenders are using a contrarian strategy, over-bidding low priced loans and under-bidding highly priced ones. Evidence for this kind of contrarian behavior by actual people has been found by Cipriani and Guarino (2005) in a laboratory financial market.

7 Conclusion

There appears to be some evidence that lenders on Prosper behave according to the model in Berkovich and Tayon (2009). I hypothesize that reading and analyzing the "soft" data contained in the images and text on a Prosper loan listing is a form of private information acquisition. Those lenders who can effectively obtain this data are at an advantage over uninformed lenders.

The model predicts high and low value regions of the assets in the market have higher variance in the difference between price and true value. I find evidence that the low value region has this property, but not the high value region. Defining adjusted prices based on the deviation from the expected loan interest rate, I find that high adjusted price loans are under-priced given their actual realized returns. While this effect means the market is not in the equilibrium described by the model, it does seem to indicate that lenders excessively avoid loans which are overpriced given the "hard" public variables. It may be that there are not enough informed lenders to bid down these loans and that the uninformed lenders avoid these loans.

Low adjusted price loans do not appear to be under-bid. If, in fact, low priced loans are slightly over-bid, then the explanation may be simply that Prosper lenders over-react to higher prices, a sort of contrarian strategy. They under-bid high adjusted price loans thinking the interest rate is

too low and over-bid low adjusted price loans thinking the interest rate is high and a good deal.

8 References

Banerjee, Abhijit. 1992. "A simple model of herd behavior," *Quarterly Journal of Economics*, v.107(3), pp.797-817.

Berkovich, Efraim and Robert Tayon. 2009. "Herding and crowding to efficiency," *Working paper.*

Bikhchandani, Sushil, David Hirshleifer, and Ivo Welch. 1992. "A theory of fads, fashion, custom, and cultural change as informational cascades," *The Journal of Political Economy*,v.100, pp.992–1026.

Burdett, Kenneth, Shouyong Shi, and Randall Wright. 2001. "Pricing and Matching with Frictions," *Journal of Political Economy*, v.109, pp. 1060–1085.

Cipriani, Marco, and Antonio Guarino. 2005. "Herd Behavior in a Laboratory Financial Market," *American Economic Review,* v.95(5) pp. 1427-1443.

Debo, Laurens. G., Christine A. Parlour, and Uday Rajan. 2008. "Inferring Quality from a Queue," *manuscript.*

Drehmann, Mathias, Jorg Oechssler, and Andreas Roider. 2007. "Herding with and without payoff externalities — an internet experiment," *International Journal of Industrial Organization* v. 25 pp. 391–415.

Freedman, Seth and Ginger Zhe Jin. 2008. "Do Social Networks Solve Information Problems for Peer-to-Peer Lending? Evidence from Prosper.com". *Working paper.*

Iyer, Rajkamal, Asim Ijaz Khwaja, Erzo F.P. Luttmer, and Kelly Shue. 2009. "Screening in New Credit Markets: Can Individual Lenders Infer Borrower Creditworthiness in

Peer-to-Peer Lending?" *NBER Working Paper 15242* http://www.nber.org/papers/w15242

Huck, Steffen and Jorg Oechssler. 1998. "Informational cascades with continuous action spaces," *Economics Letters*, v. 60(2), pp.163-166.

Lang, William W. and Leonard I. Nakamura.1990. "The dynamics of credit markets in a model with learning," *Journal of Monetary Economics*, v. 26(2), pp. 305-318.

Lester, Benjamin. 2007. "Differential information, limited availability, and prices," *manuscript.*

Peters, Michael. 1984. "Bertrand Equilibrium with Capacity Constraints and Restricted Mobility," *Econometrica*, v.52,(5) pp. 1117-1127

Pope, Devin and Justin Sydnor. 2008. "What's in a Picture? Evidence of Discrimination from Prosper.com," *Working paper.*

Simonsohn, Uri and Ariely Dan. 2008. "When Rational Sellers Face Non-Rational Consumers: Evidence from Herding on eBay," *Management Science,* v54(9) pp.1624-1637

Theseira, Walter. 2008. "Competition to Default? Racial Discrimination in the Market for Online Peer-to-Peer Lending," *Working paper.*

Vives, Xavier. 1996. "Social learning and rational expectations," *European Economic Review*, v.40, pp.589-601

Welch, Ivo. 1992. "Sequential sales, learning and cascades," *The Journal of Finance*, v.47, pp.695–732.