WORKING COLLECTIVE INTELIGENCE: THE CASE OF THE P2P LENDING

Seongmin Jeon, College of Business Administration, Seoul National University, Seoul, Korea, R.O.K., smjeon@snu.ac.kr

Byungjoon Yoo, College of Business Administration, Seoul National University, Seoul, Korea, R.O.K., byoo@snu.ac.kr

Abstract

Peer-to-peer (P2P) lending is an open marketplace for loans provided by individuals through the Internet rather than via traditional financial institutions. On a P2P lending Web site, potential borrowers create and post a listing with an overview describing their need for a loan while potential lenders place bids on listings they would be interested in funding. A borrower would be provided a loan only in the case that his or her listing garnered enough bids to exceed a predefined amount or to fulfill a loan request by a number of lenders. In regard to P2P lending, while it is the belief of some that collective intelligence can play a role in filtering out unreliable borrowers who are unlikely to repay the loan they received, others point out the existence of irrational herding with informational cascades. In light of the easy availability of ranking chart information on the bid participation rate for P2P lending, this unique financial tool offers an ideal environment for information cascades. In this study, I empirically examine the impact of ranking chart information on following bids. The relationship among the likelihood to get funded, timely repayment and ranking chart information is also examined to identify the impact of information regarding how many other borrowers participate in the bids.

Keywords: P2P lending, collective intelligence, herding, information cascades.

1 INTRODUCTION

The wisdom of crowds is said to enable businesses to profit when social networks try to establish the concept of a community into their decision making. For example, the underwriting decisions assessing the risk of each loan in microlending sites are made by individuals, while the value of a loan is established through lender bidding. Considering the context of the borrower, such lending decisions are expected to be superior to the same decisions currently made by loan officials at banks (Libert and Spector 2007).

Peer-to-peer (P2P) lending is an open marketplace for loans provided not by a bank but from individuals online. Financial transactions are facilitated directly between individuals ("peers") without any intermediation of a traditional financial institution. A market study by the Gartner Group forecasts that the scope of P2P lending will soar at least 66% to US$5 billion in outstanding loans by 2013 (Gartner 2010).

This rising presence of P2P lending has resulted in this type of funding becoming a popular research topic. A number of research papers have conducted studies based on data provided by Prosper.com, a U.S.-based P2P lending Web site. Social networks with Web 2.0 features found in P2P lending sites have been analyzed in the research of Lin, Prabhala, and Viswanathan (2009) which explained the effects and patterns of social networks on the fundability and appropriateness of a repayment. The intervention and coordination of groups and group leaders also play a key role in the full funding of an idea and subsequently timely repayment.

Puro, Teich, Wallenius and Wallenius (2010, 2011) found that there is a trade-off between having a low final rate and getting the loan funded and classifying bidding strategies in Prosper.com empirically, concluding that bidding behavior is not homogeneous among bidders. Shen, Krumme, and Lippman (2010) found that herding takes place when lenders make investments on loan listings rather than on more rational investments based on risk and returns.

The objective of this study is to empirically examine the impact of information cascades on bid participation in P2P lending and identify the relationship between the ranking chart information provided by the intermediary and the subsequent bids of lenders. The expectation is that a herding effect would take place in regard to the likelihood of funding and timely repayment.

The remainder of the study proceeds as follows. In Section 2, literature related to the ranking information and non-rational herding is briefly reviewed. In Section 3, the data used in this study is introduced and the numbers of bids in comparison to the rankings are charted. Section 4 describes the construction of the model and introduces the underlying methods. The current status of research and plans for completion are included in Section 5.

2 LITERATURE REVIEW

Even though research on the P2P lending is just emerging, there have been earlier works in the area of information systems, economics and marketing related to direct financing on the Internet.

a. Collective intelligence

Generally collective intelligence is defined as the synergistic and cumulative channelling of the efforts of many minds towards selecting actions in response to some challenge (Walton and Krabbe 1995). Libert et al. (2007) point out that collective intelligence itself is widely used for systems of crowd reasoning with the adoption of the Internet and social networks. According to Tapscott and Williams (2008), collective intelligence means mass collaboration. In order for this concept to happen, four principles – openness, peering, sharing and acting globally – are necessary.

b. Herd behavior and information cascades

Herd behavior is particularly prominent in the IS area. Computer users frequently adopt popular software products, resulting in making them even more popular (Brynjolfsson and Kemerer 1996). Simonsohn and Ariely (2004) found that bidders are repeatedly drawn toward herd favoring auctions with more existing bids. Informational cascades offer an information-based explanation for herd behavior (Bikhchandani, Hirshleifer and Welch, 1992; Bikhchandani and Sharma, 2001), implying that individuals lacking certain information may ignore their own information and choose to follow the crowd.

c. Ranking information

Duan, Gu, and Whinston (2009) empirically examined information cascades in the context of software adoption. Ghose and Yang (2009) found that the monetary value of a click is not uniform across all positions. Clemons, Gao and Hitt (2006) empirically demonstrated that the variance of ratings and the strength of the top quartile of reviews play a significant role in determining which new products grow in the marketplace.

3 RESEARCH HYPOTHESES AND DATA

a. Research hypotheses

This empirical study is conducted in the context of P2P lending at Popfunding.com. Popfunding.com presents an ideal environment for this research in that the site's P2P lending market follows the rule of Dutch auctions for borrowers' requests in the same format as found on the site Prosper.com. A potential borrower's requested funding amount is limited to KRW25 million (roughly equivalent to US$25,000) while each lender's bid amount is limited to less than KRW100,000 (approximately US$100). Like other P2P lending sites, Popfunding.com's user profiles show that most borrowers have low credit scores, with 96% of potential borrowers ranking 7 or lower on Korea's credit ranking system (based on a 1-10 scale where 1 represents outstanding credit). Borrowers who succeed in getting financing through the Popfunding.com site were found to be highly likely to repay the loan, as seen in the 95% repayment rate. Such a rate far exceeds expectations of the public who are unlikely to assume such a high repayment percentage.

To test whether collective intelligence works in the process of reviewing the loan requests of potential borrowers, I established the following hypotheses based on previous research. The informational cascade theory suggests that decision making changes with the relative ranking of products, that is, the number of bids in this study. Therefore, I propose:

H1. Herding is expected to take place in the case that requests are displayed on the ranking charts.

It can be inferred that a request could get funded with a higher probability if the request appears on the ranking chart.

H2. Whether a request has appeared on the ranking chart or not increases the likelihood of getting funded.

I would test if the ranking information could influence repayment additionally.

H3. The higher rank is, the timelier the repayment is.

If the collective intelligence can identify the likelihood to repay in timely manner, the more bid amount the requests raise, the timelier the repayment is.

b. Data

Transactional data from Popfunding.com was collected from December 1-31, 2009 for registration dates. During that period, the total number of requests was 307, while 6,282 bids were generated for those requests. On average, the bidding period for a request lasted from 7 to 10 days. For the period, I obtained 2,261 day-request combinations. Overall, 14% of total requests received a sufficient number of bids to get funded.

Variable	Mean	Std. Dev.	Min	Max
Rank	21.78838	24.09899	1	114
Rankyn (on the ranking chart =1, else =0)	0.3695976	0.4830559	0	1
Max_rank	3.549925	2.321244	1	8
Bid_num	7.870343	14.21182	0	99
Daily_bid_amount (KRW)	106,451.60	336,870	0	3,850,000
Total_money (KRW)	2,070,790	989,035	500,000	5,000,000
Funded (yes =1, no =0)	0.4262295	0.4948969	0	1
Timely_repayment (not funded= -1, no = 0, yes = 1)	-0.2146051	0.9424459	-1	1

Table 1. Sample descriptive statistics, N = 2,261

On Popfunding.com, a ranking snapshot is recorded at the beginning of a day. The number of bids was counted for any particular day. In this light, ranking data preceded the action of bidding data in the perspective of time. Figure 1 shows the relationship between the number of bids and the ranking information. The majority of requests attracted only a small number of bids while highly ranked requests were found to have significantly more bids.

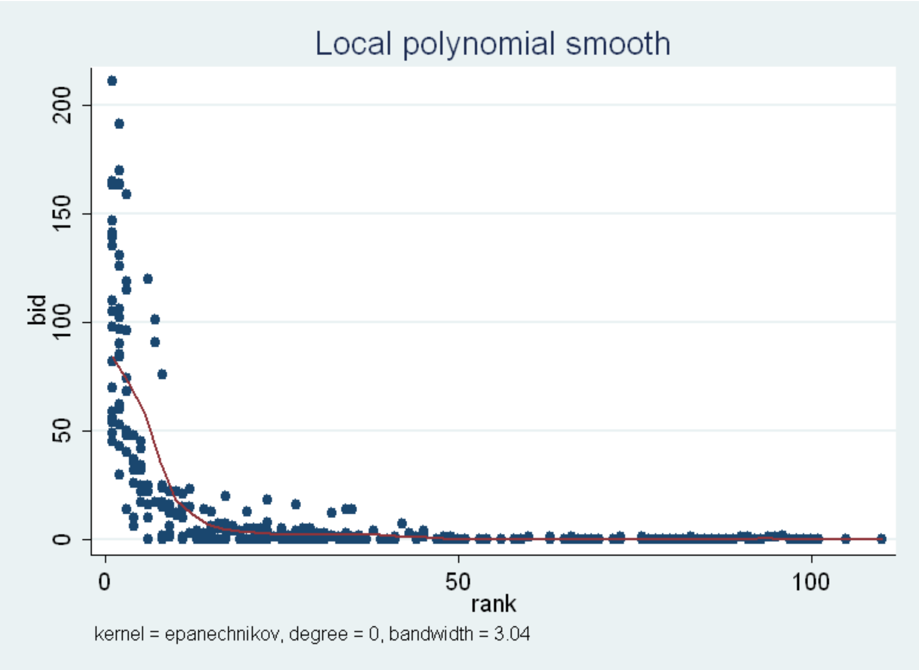

Figure 1. The relationship between the number of bids and the ranking information

4 MODEL

To test the hypothesis that the requests on the ranking chart attract more bids, a dummy variable model was used to analyze the seperate groups of the "on-ranking chart" bids as well as those not on the ranking chart. Then in order to verify the difference between the requests on the day when a request is listed on the ranking chart, a T-test was undertaken to determine the difference in bid numbers between two groups. Finally, in order to examine whether highly ranked requests would be repaid in a timely manner, the discrete dependent model was applied.

Specifically, the differences between the two groups – on-the-chart bids and others – were what were sought. The dummy variable model was able to clarify such a distinction.

$$Y_i = \beta_0 + \beta_1 (\text{Daily Average Bid Amount}_i) + (\text{RankN}_i) + D_1(\text{RankY}_i) + \varepsilon_i \qquad (1)$$

In the model utilized in this study, the daily average bid amount was used as an independent variable to compare the differences between the groups of on-chart bids and those not on the chart. The coefficient for the not-on-chart dummy variable was omitted as RankN_i, with the binary to be 0 if the request is not on the ranking chart. Thus, the author's interest concerns the effect of the ranking chart.

Variable	Definition
Dependent Variable	
bid_num (Yi)	Number of bids per day i
Independent Variables	
Daily Bid Amount_i	The amount of participating bids per day i
Daily AverageBid Amount_i	Daily Bid Amount / bid_num per day i
RankN_i	Dummy variable for a request not on the ranking chart per day i
RankY_i	Dummy variable for a request on the ranking chart per day i

Table 2. Variable Definitions

The author verifies the existence of a ranking information effect between the groups of days when the requests are on the chart and not on the chart. The mean difference between the group of bids on the days on the chart and the other group of bids on the days not on the chart were tested after extracting the requests that have been on the chart more than once. A T-test was used to determine the difference between the two groups.

The requests with higher ranking were found to be more likely to get funded as expected because the ranking corresponds with the order of the amount the request raises. Furthermore, the requests with a higher ranking were more likely to be repaid in a timelier manner. Such findings were tested using a probit and logit model to see the relationship among the likelihood to get funded, the timely repayment and the best ranks the requests had.

5 CURRENT STAGE OF RESEARCH AND DIRECTION

In this study, the author attempts to analyze the respective impacts of collective intelligence, informational cascades and ranking chart information by empirically investigating the impact of changes in ranking on lenders' bids. While the P2P lending market represents an extreme case of information overload in which only a limited amount of information regarding borrowers can be known, information about the participation of others in the bidding process could influence subsequent lenders' decisions. The author is currently working on analyzing the panel data using a GLS model with dummy variables to test that the lenders' choice of bids is influenced by a funding request's entry on the ranking chart with data covering a period of one month. It is expected that the findings will be consistent with informational cascade theories which show that individuals are very much influenced by the information inferred from others' behavior in the online market. In the case of P2P lending, the bids of lenders are expected to rely on the information provided by the intermediary. This study's basic analysis focuses mainly on lenders' participation in bids for the funding requests of potential borrowers. The analysis, however, will draw implications for e-commerce intermediaries as well, especially in regard to management of the ranking chart information.

The author plans to pursue future research in the following directions. First, the author would expand the amount of data to be sufficient for generalizability, as he is currently testing only a fraction of the data provided by the intermediary for the sake of easier handling of data. Second, the proposed model is not able to distinguish totally rational and irrational herding. Herding resulting from informational cascades is rational in that decision makers integrate antecendents' actions into their own decisions (Duan et al., 2008). The degree of rationality in herding information will be pursued in a further analysis. Third, characteristics of borrowers will be examined to determine differences between donating lenders and investing lenders. Lenders who bid a small amount are likely not to care much about a return on their investement in that they would not expect a return which exceeded the search and bidding cost.

References

Libert, B. and Spector, J. (2007). *We Are Smarter (than Me)*, Wharton School Publishing.

Gartner reports (2010). "Predicts 2010: Executive Decisions in Banking and Investment Services Demand a Longer View" and “Predicts 2010: Operational Technologies Present Threats and Opportunities in Banking and Investment Services”

Lin, M, Prabhala, N.R. and Viswanathan, S. (2009). Social Networks as Signaling Mechanisms: Evidence from Online Peer-to-Peer Lending, *WISE 2009.*

Puro, L., Teich, J.E., Wallenius, H. and Wallenius, J. (2010). Borrower Decision Aid for People-to-People Lending, *Decision Support Systems*, 49, 52–60.

Puro, L., Teich, J.E., Wallenius, H. and Wallenius, J. (2011). Bidding strategies for real-life small loan auctions, *Decision Support Systems*, 51, 31–41.

Shen, D., Krumme, C and Lippman, A. (2010). Follow the Profit or the Herd? Exploring Social Effects in Peer-to-Peer Lending, *IEEE International Conference on Social Computing / IEEE International Conference on Privacy, Security, Risk and Trust*

Walton, D.N. and Krabbe, E.C.W. (1995). *Commitment in dialogue: Basic concepts of interpersonal reasoning*, Albany, NY: State University of New York Press.

Tapscott, D. and Williams, A. D. (2008). *Wikinomics: How Mass Collaboration Changes Everything*, USA: Penguin Group.

Brynjolfsson, E. and Kemerer, C. F. (1996). Network Externalities in Microcomputer Software: An Econometric Analysis of the Spreadsheet Market, *Management Science* (42:12), 1627-1647.

Simonsohn, U., and Ariely, D. (2004). “e-Bay's Happy Hour: Non-rational herding in on-line auctions,” Working Paper, Wharton School, University of Pennsylvania.

Bikhchandani, S., Hirshleifer, D., and Welch, I. (1992). A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades, Journal of Political Economy (100:5), 992-1026.

Bikhchandani, S., and Sharma, S. (2001). Herd Behavior in Financial Markets,” IMF Staff Papers (47:3), 279-310.

Duan, W., Gu, B. and Whinston, A.B. (2008). Informational Cascades and Software Adoption on the Internet: An Empirical Investigation. *MIS Quarterly*, Vol. 33 No. 1, 23-48.

Ghose, A. and Yang, S. (2009). An empirical analysis of search engine advertising: Sponsored search in electronic markets, *Management Science.*

Clemons, E. K., Gao, G. and Hitt, L.M. (2006). When Online Reviews Meet Hyperdifferentiation: A Study of the Craft Beer Industry, *Journal of Management Information Systems*, V 23. No 2. 149 – 171.