

RESEARCH

Open Access



Borrower's default and self-disclosure of social media information in P2P lending

Ruyi Ge¹, Juan Feng^{2*} and Bin Gu³

* Correspondence:

juafeng@cityu.edu.hk

²Department of Information Systems, College of Business, City University of Hong Kong, Kowloon Tong, Hong Kong

Full list of author information is available at the end of the article

Abstract

Background: We examine the signaling effect of borrowers' social media behavior, especially self-disclosure behavior, on the default probability of money borrowers on a peer-to-peer (P2P) lending site.

Method: We use a unique dataset that combines loan data from a large P2P lending site with the borrower's social media presence data from a popular social media site.

Results: Through a natural experiment enabled by an instrument variable, we identify two forms of social media information that act as signals of borrowers' creditworthiness: (1) borrowers' choice to self-disclose their social media account to the P2P lending site, and (2) borrowers' social media behavior, such as their social network scope and social media engagement.

Conclusion: This study offers new insights for screening borrowers in P2P lending and a novel usage of social media information.

Keyword: P2P lending, Social media, Self-disclosure, Default, Difference-in-difference

Background

Peer-to-peer lending, also known as P2P lending, refers to the practice of lenders loaning money to unrelated individuals through a P2P lending platform, an online venue. One unique feature of such lending is that the loan amount is usually small. Traditional financial institutions tend to do very little screening for small borrowers and rely excessively on collateral (Stiglitz and Weiss 1981; Ang et al. 1995; Avery et al. 1998; Manove et al. 2001). However, in P2P lending markets, borrowers do not provide collateral as protection to lenders against default. This practice makes P2P lending particularly attractive for small borrowers who might otherwise turn to pay day lenders or credit card debt (Adams et al. 2009).

To facilitate lenders to identify credible borrowers, P2P lending platforms encourage borrowers to submit as much relevant information as possible. Lenders make use of "hard" information, such as credit score, debt-to-income ratio, and annual income, and "soft" information, such as a borrower's picture or a textual description of the future plan (Iyer et al. 2009), to infer borrowers' creditworthiness. Soft information is a useful supplement to hard information in the loan underwriting process, especially for borrowers with poor credit, whose hard information is usually unattractive (Khwaja et al. 2013). Prior studies have examined the signaling effect of different types of soft information, such as borrowers' pictures, textual descriptions of the usage of loans (Khwaja et al. 2013), facial features (Theseira 2009; Pope and Sydnor 2011; Ravina 2012), and social

network characteristics on P2P platforms (Freedman and Jin 2011; Lin et al. 2013). Unlike these studies, the present study focuses on a new and promising category of soft information—borrowers' self-disclosed social media information. In P2P lending platforms, some borrowers voluntarily disclose their social media account, which makes the related information accessible to lenders or P2P lending platforms. Are borrowers who choose to disclose their social media account more creditworthy than non-disclosing borrowers? Is the social media information they choose to disclose useful in assessing their default probability?

Our study intends to answer these two questions by examining a combined data set obtained from a P2P lending platform and a social media site. With this data, we model borrowers' default probability as a function of their choice to disclose their social media account, controlling for relevant factors such as borrowers' demographic characteristics and identity verification.

The result shows that borrowers who disclosed their social media information have a significantly lower default probability compared to those who did not. To rule out the effect of self-selection, we leverage a natural experiment introduced by the P2P lending site that enabled borrowers to link to their social media sites. We employ the propensity score matching (PSM) technique to assess the relationship and find that the results are consistent. Furthermore, we examine the relationship between borrowers' social media engagement and their default probability. We find that social media engagement, such as the scope of the borrowers' social network and their activity level on a social media site, act as predictors of their default probability.

Our study makes three contributions to the P2P lending literature. First, we discover a predictive relationship between borrowers' choice to disclose their social media information and their default probability. Second, we find borrowers' social media engagement also predicts their default probability. These findings identify a new category of soft information that is useful for screening borrowers on P2P lending platforms. Finally, by examining a unique data set that uniquely combines data from both a P2P lending site and a social media site, this study is the first in the literature to integrate borrowers' financial behavior with their social media characteristics.

Methods

Hypotheses

In our study, borrowers disclosing their social media account on a P2P lending platform raise the possibility that their default behavior would be revealed to their friends. A moral failure damages one's social image, and consequently damages social bonds to others (Baumeister and Leary 1995; de Waal 1996; Ahmed 2001) and can lead to social punishments such as being marginalized, ostracized, or excluded (de Waal 1996; Braithwaite 1989). Moreover, the literature on social capital finds it to be a valuable resource (Dore 1983; Adler and Kwon 2002). Social capital originates not only from the structure and content of our social relations but also from trust (Putnam 1995; Knoke 1999; Leana and Van Buren 1999). Finally, economic theories of social stigma point out that a default imposes a social stigma cost on borrowers if their friends know about the default (Crocker et al. 1998; Thorne and Anderson 2006; Cohen-Cole and Duygan-Bump 2008). We therefore propose:

HYPOTHESIS 1. Borrowers who voluntarily disclose their social media accounts on a P2P lending platform are less likely to default.

HYPOTHESIS 2a. Borrowers who have a larger social network on the social media site are less likely to default on the P2P lending platform.

HYPOTHESIS 2b. Borrowers who have more engagement on the social media site are less likely to default on the P2P lending platform.

Data

A key, notable contribution of our study is that we combine data related to borrowers' financial behavior from a P2P lending platform with their information on a social media site. The P2P lending data are from one of the largest online P2P lending platforms in China. Our data sample covers all peer-to-peer lending listings on this company between January 2011 and August 2013. It consists of 35,457 loan records and 11,047 borrower records in total. Variables related to listed loans include loan amount, interest rate, opening and closing dates, credit grade ranging from A (high quality) to HR (low quality), and the status of loan repayment. Variables related to borrowers contain their demographic characteristics, including age, gender, education level, and marital status, and verification items, including verification of identity card, education certificate, phone number, and image.

Over 40% of borrowers on the P2P lending platform have disclosed their Sina microblog account to the platform. Sina microblog, which was launched in September 2009, is the largest microblog site in China and had nearly 300 million users by the end of 2013. The dataset includes a variable that indicates whether a borrower disclosed their Sina microblog account. For those borrowers who did so (5239 borrowers in total), we accessed their microblog page and collected relevant data. The data we thus obtained include social network scope metrics and engagement metrics.

Results and discussions

We begin by analyzing the relationship between the default outcome and borrowers' choice to disclose their microblog account. We use a logit regression model first, and then utilize the PSM technique and instrument variable regressions to address endogeneity concerns. We estimate:

$$\text{logit}(\text{Default}) = \alpha + \beta_1 \text{Microblog_disclosed} + \beta_2 \text{Controls} + \varepsilon \quad (1)$$

The results show that the microblog disclosure is negatively related to the default probability (coefficient is -0.748), and is significant at the 0.01 level. However, the data are unbalanced in covariates between the group who discloses their microblog and those who do not. The unbalanced data weakens the reliability of the results of the regression model (Imbens and Wooldridge 2009). Therefore, we utilize PSM to adjust for the differences in covariates. The results after PSM support Hypothesis 1. We find significant differences in the default rate between treated and control groups.

Difference-in-difference (DID) model

Although the result of the logistic model shows that disclosure of a microblog account is a predictor of default probability, it does not identify the underlying cause: is it because the borrowers are afraid of social stigma costs? We use a DID model to identify the cause. In April of 2013, the P2P platform launched a marketing campaign to encourage borrowers to disclose their microblog accounts. We estimate the effect

of the campaign on the default probability of a loan whose borrower disclosed their social media account. The estimated model is:

$$\ln\left(\frac{P(\text{Default}_{it} = 1)}{1 - P(\text{Default}_{it} = 1)}\right) = \alpha + \beta_1 \text{Mb_disclosed}_i + \beta_2 \text{Cmp}_{it} + \beta_3 \text{Mb_disclosed}_i \times \text{Cmp}_{it} + \beta_4 \text{Controls}_i + \varepsilon_{it} \quad (2)$$

The dummy variable Mb_disclosed equals 1 if the borrower of a loan has disclosed his microblog, otherwise it equals 0. The dummy variable Cmp is a time variable, which takes the value of 0 or 1 for periods prior to or after the disclosure campaign. Controls represent a vector of loan characteristics, such as loan amount, interest rate, and lending period. The main parameter of interest is β_3 . The result of β_3 is negative (-0.652) and significant ($p < 0.01$), suggesting that this campaign negatively influenced the default probability of the loans whose borrowers have disclosed their social media account. One possible reason is that these borrowers care about social stigma costs. Borrowers may worry the P2P lending company could use their microblog account as an outlet to spread the word if a default occurs, which would increase their social stigma costs. With this worry in mind, they are less likely to default after the disclosure campaign.

The effect of microblog behavior on default probability

We select borrowers who disclose their microblog accounts on the P2P lending site, and collect microblog metrics (e.g. #Followers, #Friends, #Fans and #Microblogs) from their profile pages on sina.com. This combined data sample includes 5239 listings.

We use a logit model to estimate the default probability of the effect of the microblog metrics on default likelihood for borrowers who have disclosed his microblog.

$$\text{logit}(\text{Default}) = \alpha + \beta_1 \text{Microblog_Metrics} + \beta_2 \text{Controls} + \varepsilon \quad (3)$$

Because of the large variance and scale of the microblog metric variables, we use their natural logs in the model.

We first analyze the effect of #Followers and #Microblogs, respectively. The independent variable in both models is negatively related to the default probability at the 0.01 significance level. The results demonstrate that the larger the scope of the social network a borrower has on a social media site, the less likely they are to default on a loan; the more engagement a borrower has with his social media site, the less likely he is to default. Both Hypotheses 2a and 2b are supported.

We next examine the effect of two different types of social network, that is, friends and fans. For a borrower, both friends and fans on the microblog site are sources of social capital. Either friends or fans knowing about a borrower's default can damage his social image and cause a social stigma cost; therefore, both #Friends and #Fans influence the borrower's default likelihood. However, as previous studies have demonstrated, close friends have a stronger behavioral effect on each other than strangers do (Bond et al. 2012; Christakis and Fowler 2013). We therefore expect the effect of #Friends on borrowers' default likelihood to be more intensive than that of #Fans. Our result shows that #Friends and #Fans are both negatively related to the default probability with $p < 0.01$, but the coefficient of #Friends (-0.153) is almost double to that of #Fans (-0.079). The results indicate that although both variables are predictors of default likelihood, #Friends is a stronger signal than #Fans.

We also consider that a borrower having a large #Followers is more likely to have a healthy financial situation as an influential person. Therefore, their low default probability may be due to their financial well-being instead of avoiding costs in social capital. We include an additional term to represent a borrower's influence, which can also be regarded as a proxy for financial position. From the Sina microblog site, we received not only data showing how many followers a borrower has (e.g., #Followers), but also data showing how many people the borrower is following (e.g., #Following). It is reasonable to assume that #Followers of influential borrowers is always greater than #Following. Therefore, we created a dummy variable "Influential," whose value equals 1 when #Followers is greater than #Following, and otherwise, equals 0. The result shows that #Followers remain significant while Influential is not significant.

Conclusion

In this study, we investigate the signaling effect of social media information on borrowers' credit worthiness in P2P lending. The results suggest that social media information can be a signal of creditworthiness on two levels. On the first level, for all the borrowers in the market, their decision on whether to disclose their social media accounts is a predictor of their default probability. On the second level, for the borrowers who choose to disclose their social media accounts, their social media metrics, such as their social network scope and their inputs on the social media site, are predictors of their default probability.

Our study contributes to the literature across the Information Systems and Finance disciplines. To our knowledge, it is the first study that examines the usage of social media in personal finance. While most prior literature regards social media as a marketing tool, we provide a new perspective by regarding social media as an information source for individual creditworthiness. Moreover, our results provide a new insight to improve risk control in P2P lending in China. On the one hand, individuals in China do not have a well-verified credit score, such as a FICO score, which enlarges the information asymmetry in Chinese P2P lending markets. On the other, about 80% of Internet users in China have a social media account (Report on Internet Development Status in China 2016). Their social media activity provides a rich set of information that could be used by P2P lending markets for credit assessment. Our study demonstrates the validity of this approach and highlights the importance of leveraging social media information in P2P lending markets in China.

This paper is not without limitation. For example, due to data limitation, we are not able to provide the money borrowers' detailed social media behavior information, especially their behavior before and after the default takes place. In the future we will work with the P2P lending company to design better experiment, to more clearly identify the impact of self-disclosure, as well as to better address the endogeneity problem.

Acknowledgement

The authors would like to thank professor Pengzhu Zhang for his insightful comments and support of this paper.

Funding

Juan Feng would like to acknowledge GRF (General Research Fund) 9042133 and City U SRG grant 7004566. Bin Gu would like to acknowledge National Natural Science Foundation of China [Grant 71328102].

Authors' contributions

All three authors contribute equally. All authors read and approved the final manuscript.

Competing interests

The authors declare that they have no competing interests.

Author details

¹College of Information and Computer Science, Shanghai Business School, Shanghai 200235, China. ²Department of Information Systems, College of Business, City University of Hong Kong, Kowloon Tong, Hong Kong. ³Department of Information Systems, WP Carey School of Business, Arizona State University, Tempe, USA.

Received: 16 November 2016 Accepted: 2 December 2016

Published online: 17 December 2016

References

- Adams W, Einav L, Levin J (2009) Liquidity constraints and imperfect information in subprime lending. *American Econ Rev* 99(1):49–84
- Adler PS, Kwon SW (2002) Social capital: Prospects for a new concept. *Acad Manag Rev* 27(1):17–40
- Ahmed E, Harris N, Braithwaite J, Braithwaite V (2001) Shame management through reintegration, Cambridge University Press, 978-0-521-00370-4
- Ang JS, Lin JW, Tyler F (1995) Evidence on the lack of separation between business and personal risks among small businesses. *J Entrepreneurial Finance* 4(2):197–210
- Avery RB, Bostic RW, Samolyk KA (1998) The role of personal wealth in small business finance. *J Bank Financ* 22(6):1019–1061
- Baumeister RF, Leary MR (1995) The need to belong: desire for interpersonal attachments as a fundamental human motivation. *Psychol Bull* 117(3):497
- Bond RM, Fariss CJ, Jones JJ et al (2012) A 61-million-person experiment in social influence and political mobilization. *Nature* 489(7415):295–298
- Braithwaite J (1989) *Crime, shame and reintegration*. Cambridge University Press, 1st edition.
- Christakis NA, Fowler JH (2013) Social contagion theory: examining dynamic social networks and human behavior. *Stat Med* 32(4):556–577
- Cohen-Cole E, Duygan-Bump B (2008) Household bankruptcy decision, the role of social stigma vs. information sharing. Working paper, Federal Reserve Bank of Boston, Boston
- Crocker J, Major B, Steele C (1998) Social stigma. In: Gilbert DT, Fiske ST, Lindzey G (eds) *The Handbook of Social Psychology*, Vols. 1 and 2. McGraw-Hill, New York, pp 504–553
- de Waal FBM (1996) *Good Natured: The Origins of Right and Wrong in Humans and Other Animals*. Harvard University Press, Cambridge
- Dore R (1983) Goodwill and the spirit of market capitalism. *The British Journal of Sociology* 71(3):459–482
- Freedman SM, Jin GZ (2011) Learning by doing with asymmetric information: evidence from Prosper.com (No. w16855). National Bureau of Economic Research. Cambridge, MA
- Imbens GW, Wooldridge JM (2009) Recent Developments in the Econometrics of Program Evaluation. *J Econ Lit* 47(1):5–86
- Iyer R, Khwaja AI, Luttmer EF, Shue K (2009) Screening in new credit markets: Can individual lenders infer borrower creditworthiness in peer-to-peer lending? In: AFA 2011 Denver Meetings Paper
- Iyer R, Khwaja AI, Luttmer EFP, Shue K (2016) Screening peers softly: Inferring the quality of small borrowers. *Management Science*, 62(6):1554–1577
- Knoke D (1999) Organizational networks and corporate social capital. In: Th R, Leendersand AJ, Gabbay SM (eds) *Corporate Social Capital and Liability*. Kluwer, Boston, pp 17–42
- Leana CR, Van Buren HJ (1999) Organizational social capital and employment practices. *Acad Manag Rev* 24(3):538–555
- Lin M, Prabhala NR, Viswanathan S (2013) Judging borrowers by the company they keep: friendship networks and information asymmetry in online peer-to-peer lending. *Manag Sci* 59(1):17–35
- Manove M, Padilla AJ, Pagano M (2001) Collateral versus project screening: A model of lazy banks. *RAND Journal of Economics* 71(3):726–744
- Pope DG, Sydnor JR (2011) What's in a picture? evidence of discrimination from Prosper.com. *J Hum Resour* 46(1):53–92
- Putnam RD (1995) Bowling alone: America's declining social capital. *J Democr* 6:68
- Ravina E (2012) Love and loans: the effect of beauty and personal characteristics in credit markets., Available at SSRN 1101647
- Report on Internet Development Status in China (2016) 7. <http://www.cnnic.net.cn/gywm/xwzx/rdxw/2016/201608/W020160803204144417902.pdf>
- Stiglitz JE, Weiss A (1981) Credit rationing in markets with imperfect information. *American Economic Review*, 71(3):393–410
- Theseira W (2009) *Competition to Default: Racial Discrimination in the Market for Online Peer-to-Peer Lending*. Dissertation #AAI3395722, University of Pennsylvania, Philadelphia, PA
- Thorne D, Anderson L (2006) Managing the stigma of personal bankruptcy. *Social Focus* 39(2):77–97