https://doi.org/10.1016/j.pacfin.2020.101338

Does Social Capital Matter for Peer-to-Peer-Lending? Empirical Evidence

Haitian Lu, Faculty of Business, The Hong Kong Polytechnic University

Bo Wang, Stuart School of Business, Illinois Institute of Technology

Haizhi Wang, Stuart School of Business, Illinois Institute of Technology

Tianyu Zhao, China Construction Bank

Abstract

This paper examines the relation between regional social capital and online peer-to-peer loans. The results indicate that borrowers from states with higher levels of social capital are less likely to be rejected during loan application, have a lower probability of default, and experience lower borrowing cost. In addition, loans granted to borrowers in states with higher levels of social capital yield higher rates of return after controlling for the loan defaults and loan prepayment. The effects of social capital on peer-to-peer loans are stronger in regions with more bank competition and for loans with higher risks.

JEL Classification: B55, G14, G21
Keywords: Social capital; Peer-to-peer lending; Bank competition

1. Introduction

With the development of internet technology, online peer-to-peer (P2P) lending has become increasingly popular as a new source of credit supply in the past few years. One of the largest P2P lending platforms, Lending Club, has served more than 4 million customers with more than $50 billion loans since 2007[1]. Another P2P lending platform, Prosper, has helped roughly 1 million customers obtain more than $15 billion loans since 2005[2]. Note that the vast majority of P2P loans are unsecured microloans with short maturity. Online P2P lending platforms skip the traditional financial intermediaries and bridge borrowers directly with lenders. As a result, the borrowers tend to experience lower interest rates, whereas the lenders are able to earn higher rates of return (Emekter et al., 2015).

The fundamental issue in P2P lending is the information asymmetry and the lack of financial intermediation as delegated monitors (Diamond, 1984). Both adverse selection problem (Akerlof, 1970) and moral hazard problem (Stiglitz and Weiss, 1981) are present and severe in P2P lending. P2P lending platforms employ multiple methods to address the issue in order to bridge the borrowers and the lenders and facilitate the transactions. The P2P lending platform requires the borrowers to provide “hard” information such as FICO credit score, annual income, debt-to-income ratio, and past credit history to certify their creditworthiness. In addition, researchers have identified other factors that are important determinants in the P2P lending. For example, Duarte et al. (2012) explore appearance-based impression and report that borrowers who appear more trustworthy tend to experience higher probabilities of loan issuing and lower

[1] https://www.lendingclub.com/info/statistics.action, last viewed on 6/30/2019.
[2] https://www.prosper.com/about, last viewed on 8/1/2019.

probabilities to default. Lin et al. (2013) further document that online friendships reduces the probabilities of loan rejection, the loan interest rates and ex post default rates.

In this paper, we focus on the regional social capital, and investigate whether and to what extent social capital may affect P2P lending. Social capital is embedded in local areas and has been shown to play an important role in the economy and society (Coleman, 1988, Sobel, 2002, Buonanno et al., 2009, Putnam, 1995, Knack and Keefer, 1997). Social capital consists of cooperative norms and social networks that together facilitate collective actions (Woolcock, 2001). Strong cooperative norms constrain the opportunistic behaviors, and dense social networks help to promote the trustworthiness (Hasan et al., 2017, Guiso et al., 2004, Coleman, 1988). Therefore, in this study, we posit that borrowers in states with higher level of social capital tend to be treated more favorably in P2P lending and are less likely to default after loan issuance. Moreover, loans granted to borrowers in states with higher levels of social capital may yield high rates of return for lenders.

Using an extensive loan-level dataset from Lending Club for the time span from 2008 to 2018, we document that regional social capital significantly alleviates the information asymmetry between borrowers and lenders and facilitates the P2P lending. Specifically, our empirical results show that loan applicants in states with higher levels of social capital are less likely to be rejected for loan issuance. Borrowers in states with higher levels of social capital have lower ex post loan default rates. Furthermore, we report that social capital helps borrowers to obtain higher credit grades and, in turn, experience more favorable interest rates. We also confirm one channel through which social capital affects the P2P lending process is to encourage online communication and voluntary disclosure. We examine the effect of local banking market structure on the relation between social capital and P2P lending. We find that such relation is stronger in markets with more intensive bank competition. Our results are robust to the instrumental variable (IV) approach and alternative measure of social capital.

We further investigate the effect of social capital on the rates of return from the lenders' perspective. The lender's rates of return are somewhat different from loan interest rates in several aspects. First, the loan interest rates are a nominal rate of return and equal to the actual rate of return only when borrowers make the contractual monthly payments in time. Second, borrowers may repay a portion of the principal before they default. Third, Lending Club allows prepayment without penalty. Our measure of rates of return takes into account the three abovementioned distinctions. Specifically, we calculate the internal rate of return (IRR) for P2P loans from the actual cash payments including both interests and principals received by the lenders. Our results reveal that the IRR is significantly greater for loans granted to borrowers in the states with high levels of social capital. In addition, we investigate the social capital effects for loans with different level of risks. Our findings suggest that the effect of social capital on IRR is stronger for riskier loans.

The remainder of the paper proceeds as follows. In section 2, we review related literature. Section 3 details the data, sample and measures. We present our empirical results in section 4, and summarize and conclude in section 5.

2. Related literature

Although there is no standard definition of social capital in the literature, social capital has been studied extensively since the 1990s (Putnam et al., 1993, Coleman, 1988, Sobel, 2002, Putnam, 1995). Previous studies identify overall positive effects of social capital on the economy, society, and individuals (Putnam, 2001, Buonanno et al., 2009, Knack and Keefer, 1997). An emerging line of research finds that social capital also plays a vital role in the capital market. Among the few studies on the relation between social capital and capital market, Guiso et al. (2004) find that households in higher social capital areas in Italy have broader access to institutional credit. Hasan et al. (2017) show that firms headquartered in top social capital counties in the US incur favorable bank loan spreads. In this study, we focus on the online P2P market and attempt to shed further light on the effect of social capital in this market.

The literature on P2P lending mainly focuses on the factors that are directly associated with the default risk of P2P loans. For example, Emekter et al. (2015) identify four critical determinants of P2P loan default risk, including credit grade, debt-to-income ratio, FICO score, and revolving line utilization. Existing studies also explore the role of "soft" information in the P2P lending process. Duarte et al. (2012) examine the role of borrowers' photograph and find that those borrowers who have a trustworthy appearance have a better chance to fund their loan, higher credit, and lower default probability. Gonzalez and Loureiro (2014) examine the "beauty premium" theory based on the photograph of borrowers and find that age and attractiveness are critical roles to the loan success by signaling the competence. Larrimore et al. (2011) emphasize the importance of extended narratives, concrete descriptions, and quantitative words on facilitating trust and funding success.

Note that we are not the first to investigate the effects of social capital on debt markets. For example, Hasan et al. (2017) report that regional social capital is negatively correlated with cost of bank loans as well as cost of public bonds by listed companies. Our study instead focuses on another fast-growth segment of capital market, namely P2P lending market. P2P loans are generally small, and uncollateralized personal loans (Galak et al., 2011, Paravisini et al., 2017). The lenders in this market lack of the sophisticaiton and expertise in assessing the creditworthiness and riskiness of the borrowers (Larrimore et al., 2011), and do not engage in personal interaction with borrowers (Herzenstein, Sonenshein, and Dholakia 2011, Larrimore et al. 2011). Traditiongal financial institutions (e.g., banks) are less likely to extend credit in this market (De Roure et al., 2016). Therefore, focusing on P2P market allows us to shed further light on the effect of social capital in facilitating economic transaction.

For another instance, other studies (Lin et al., 2013, Greiner and Wang, 2009) highlight the importance of social capital at indvidual level such as group affiliations and personal networks and examine the relation between social capital and P2P lending. Nonetheless, these measures are likely to be subject to endogeneity issue because there are other factors driving both the loan contractual terms and individual social capital. Moreover, the online P2P platforms may not be able to catpure such individual social capital and convey the message to potential lenders. In other words, it is unclear about the underlying mechanisms when the information at individual level is unavaiable to lenders (Freedman and Jin, 2017). Our measures of regional social capital has the advantages to capture social capital embedded in local markets which is stable over time and less subject to endogeneity issue.

In this study, we posit that social capital may also play an important role in the P2P lending through several underlying channels. Particularly, we follow the definition from Woolcock (2001), and define social capital as the norms and networks that facilitate collective actions. The social capital theory generally identifies three interrelated dimensions: relational, cognitive, and structural (Nahapiet and Ghoshal, 1998). The relational dimension describes the relations developed through the social interactions and civic engagements of individuals. The cognitive dimension refers to the resources providing shared goals and cultures. The structural dimension emphasizes the properties of the social networks and the overall pattern of connections.

The relational dimension of social capital emphasizes the importance of private social links. Freedman and Jin (2017) demonstrate that borrowers' social ties within the online platform are associated with higher probability to get loans funded and lower interest rates, but high social ties are linked with higher delinquency or default rates. Lin et al. (2013) provide evidence that online friendships of borrowers serve as a signal of the credit quality. Galak et al. (2011) find that lenders are more likely to pick borrowers who are socially proximate, which are measured by gender, occupation, and initial of the first name.

The cognitive dimension of social capital stressess the shared goals and culture in social norms. Coleman (1988) argues that reciprocity and cooperative norms help to limit opportunistic behaviors in high social capital regions. In high social capital areas, strong cooperative norms help individuals to view self-serving behaviors as contradictory to common values (Coleman, 1988, Elster, 1989, Uhlaner, 1989). Moreover, Hasan et al. (2017) show that social capital provides environmental pressure constraining opportunistic behaviors. Thus, borrowers from high social capital regions are less likely to default.

The structural dimension of social capital highlights the properties of social networks. Social capital helps to promote trust and collaborations from civic engagement and activities. People tend to trust others who belong to the same group (Rupasingha et al., 2006) through repeated interactions and collaborations. The information asymmetry between borrowers and lenders as well as the monitoring cost in the P2P market can be greatly lessened in a closely connected network (Fukuyama, 2001). Therefore, we hypothesize that, all else being equal, borrowers from regions with higher level social capital tend to experience lower likelihood of loan rejection, lower cost of loan and lower likelihood of default in the P2P lending market.

3. Data, sample, and measures

We rely on several data sources to conduct our analysis. We obtain the P2P lending data from the Lending Club website. Lending Club reports its loan information from 2007 onward. In particular, we construct two samples from 2008 to 2018 according to the loan application status. The first sample includes all the issued loans ("issued sample"). Lending Club provides detailed information for each issued loan including the loan amount, loan term, interest rate, borrower's annual income, borrower's home-owning status, borrower's first 3-dight zip code, borrower's state, borrower's debt-to-income ratio, borrower's fico score range, current loan status, and loan purpose. In addition, Lending Club assigns a credit grade ranging from A to G with A and G being the grades from the lowest and highest risk levels, respectively. For each credit grade, there are five sub-credit grades from 1 to 5. We exclude loans granted in 2007 since Lending Club used a pilot credit model in that year (Serrano-Cinca and Gutiérrez-Nieto, 2016). We further require that each borrower have a debt-to-income ratio between 0 and 100 percent, and that the application type be "individual." We further exclude the loans with the "current" status so that we can identify the default status.

Lending Club also provides certain information for declined loans. We thereby include the declined loans after 2007 in our first sample to form the second sample, for which we label as "rejected sample." To ensure that the rejected loans and issued loans are comparable, we require that the loan applicant should have a FICO score of 640 or higher and a debt-to-income ratio between 0 and 100 percent.

In our regression analysis, we construct several dependent measures. For example, *Reject* is a dummy in the "rejected sample" which equals to 1 for rejected loans and 0 for issued loans. We also construct three variables to evaluate loan performance. The first measure is a dummy variable (*Default*) which equals 1 if the loan is default and zero, otherwise. Particularly, we categorize a loan as a default loan if the loan status is "Charged Off," "Default," "Late (16-30 days)", or "Late (31-120) days". Second, we count the number of months between the issue date and the date of last payment received (*Loan life*). Third, we calculate the internal rate of return (*IRR*) using information on the loan principal amount, actual payment amount, and the actual time of each payment. In other words, the IRR captures the lender's effective interest rate (Serrano-Cinca and Gutiérrez-Nieto, 2016), which takes into account the different repayment schedules and helps us to compare across different loans.

Our main explanatory variable is social capital. Following Rupasingha et al. (2006), we construct an index to capture regional social capital (*social capital*) from four state-level measures. The first measure is the aggregate number of all social associations in a state divided by the state's population (*assn*). Social associations are defined by the County Business Patterns of United States Census Bureau and include religious organizations, civic and social associations, business associations, political organization, professional organizations, labor organizations, bowling centers, fitness, and recreational sports centers, golf courses and county clubs, and sports teams and clubs. The second measure is the number of tax-exempt non-profit organizations from the National Center for Charitable Statistics scaled by population (*nccs*). *Assn* and *nccs* together proxy the social network density. The third measure is the percentage of voters (*pvote*) who voted in presidential elections (Alesina and La Ferrara, 2000). The last measure is the state-level response rate (*respn*) to the Census Bureau's decennial census (Knack, 2002). *Pvote* and *respn* together describe the civic norms. We further correct the *assn* and *nccs* to get time-consistent coverage. As suggested by Hasan et al. (2017a), we adjust the *nccs* in the year 1990 by estimating the historical growth rate from 1997 to 2009 and recalculate the *assn* in 1990 and 1997 using the social association category in the year 2005, 2009, and 2014. Following Rupasingha et al. (2006), we perform a principle components analysis (PCA) using the abovementioned four factors, and the first principal component is interpreted as the index of social capital. Our adjusted index has an overall correlation of 0.98 with the original index provided by the Northeast Regional Center for Rural Development.

We also include an alternative measure of social capital – social trust. Social trust information is obtained from the US section of World Values Survey data for the period from 2010 to 2014. In the survey, Questions 102 to 107 ask about interviewees' trust level about their family, neighborhood, and various types of people. We use a 1-4 scale to recode their answers. Specifically, 4, 3, 2 and 1 represent "trust completely," "trust somewhat," "do not trust very much," and "do not trust at all," respectively. Then we calculate the mean of the scores across the six questions as the social trust score for a particular respondent. We then calculate the average score of all respondents in a particular state as our measure of social trust.

We construct a set of variables capturing various loan characteristics. *Interest rate* is the borrower's contractual interest rate expressed in percentage. The *loan amount* measures the size of the loan in thousand dollars. The loan term in Lending Club is either 36 months or 60 months. We thereby use a dummy (*Loan term*) to control the loan term, which equals to 1 if a loan has a 60-month length and 0 if it is a 36-month loan. In addition, we also include a set of variables controlling for the borrower characteristics. We define an applicant to be a homeowner if the person owns a home or in mortgage (*Own home*). We use a 1-6 scale to reflect the FICO scores with 6 being the highest FICO score group, and 1 being the lowest FICO score group. We use a 1-7 scale to recode the A-G grading assigned by Lending Club. In particular, *Credit grade* has a score of 7 for A-grading and a score of 1 for G-grading. *Debt-to-income ratio* is the percentage of borrowers' debt divided by annual income at the time of application. *Employment length* is the number of years the applicant being employed at the time of loan application. We also enter several other variables including an indicator of whether the revolving line is utilized (*Revolving utilization*), the number of delinquency in past two years (*Delinquency*), the number of credit inquiries in past 6 months (*Credit inquiries*), the number of open accounts (*Accounts opened*), and the number of derogatory public records (*Public records*).

Additionally, we obtain data from American Community Survey from U.S. Census Bureau and U.S. Bureau of Economic Analysis (BEA) to construct measures to control for regional level characteristics (Hasan et al., 2017; Guiso et al., 2004). For example, *Population* measures the region size and is the natural logarithm of number of persons in millions. *Age* is the median age for all state residents. *Education* measures the regional education attainment, and is calculated as the percentage of population with a bachelor degree or higher degrees for individual age at 25 years or older. *Unemployment* is the unemployment rate. *Income per capita* is the natural logarithm of average personal income per capita.

Table 1 presents the summary statistics of the two samples. There are 5,667,774 loan applications from 2008 to 2018 in our "rejected sample". The average rejection rate is 74.6%. Potential borrowers choosing P2P platform generally apply for an average loan amount of $15,150. Among 1,438,799 loans in our "issued sample", 20.8% of loans default. According to our calculation, the mean and median IRR is -2.2% and 11.3%, respectively. Borrowers on average pay interest rate at about 13.22%. The mean and median of loan amount for issued loans are

$14,400 and $12,000, respectively. In Table 2, we show some key statistics of issued loans by their credit grade. Borrowers with credit grades of B or C received more than half of the loans in our sample, in terms of the number of granted loans and loan amounts. Consistent with the credit grades, the average default rate and interest rate are monotonically decreasing with the riskiness of the borrowers. For lenders, providing funds to low-risk borrowers yields a positive rate of return after considering the expected defaults.

[Insert Table 1 about here]

[Insert Table 2 about here]

4. Empirical results

4.1 Social capital and the probability of loan rejection

We present our empirical results for the "rejected sample." To address the potential endogeneity problem, we adopt an IV approach. Following the literature, we employ the distance of state centroid to the Canadian border as the instrument. According to Putnam (2001), "the best single predictor of the level of social capital in American states is the distance to the Canadian border. Being closer to the Canadian border means more social capital." In all model specifications, standard errors are clustered by the first 3-digit zip code to control for residual dependence created by the geographic effect (Petersen, 2009).

We posit that regional social capital help loan applicants to get the loans. Particularly, we model an indicator whether the loan is rejected as a function of social capital along with a set of controls including the applicants' FICO score, debt-to-income ratio, loan amount, employment length, regional controls, and year fixed effects. Table 3 presents the regression results. Column 1 reports the second-stage regression result using the fitted social capital as the main explanatory variable. We document a significantly negative coefficient for the fitted social capital. Our finding indicates that borrowers from regions with higher levels of social capital experience a lower likelihood of being turned down for loan applications. The Wald test of exogeneity rejects the null hypothesis, which ensures the appropriateness of the IV approach. Column 2 of Table 3 demonstrates the result of using social trust as an alternative measure of social capital and result is consistent with our main finding reported in column 1.

[Insert Table 3 about here]

4.2 Social capital and loan characteristics

In this section, we examine whether and to what extent regional social capital may affect loan characteristics and report our findings in Table 4. In column 1, we model the likelihood of loan default as a function of social capital along with other controls. We report the second stage result of IV-Probit regression. In addition, we include year fixed effects and loan purpose fixed effects in our model specification. We find a significantly negative relation between our measure of social capital and loan default probability. Regarding other control variables, the findings are generally consistent with existing literature. For example, borrowers have a lower likelihood of default when they have better credit grade, more annual income, lower debt-to-income ratio, longer employment length, smaller loan amount, shorter loan terms, better credit history, owning a home, and paying a low-interest rate.

Next, we conduct a Cox proportional hazard model which relates the time passed before an event occurs to some explanatory variables (Cox, 1972). This survival model has been widely adopted in analyzing P2P loan performance (e.g., Butler et al. 2017; Emekter et al. 2015; Lin et al. 2013; Serrano-Cinca et al. 2015; Wei and Lin 2017; Xu and Chau 2018; Zhang and Liu 2012). In our study, the events are the defaults of P2P loans. We gauge the number of months between the loan issuance and the date of loan default as our dependent variable. For fully paid loans, the dependent variable is right censored at the number of months between loan origination and the final payment date. We include the same set of control variables as used in previous tests and report the results in column 2 of Table 4. According to Lin et al. (2013), a hazard ratio greater (less) than 1 indicates a positive (negative) relation between the explanatory variable and the probability of default (Serrano-Cinca et al., 2015). We report a significant and less-than-one hazard ratio (i.e., 0.973) for social capital, which indicates that higher regional social capital significantly reduces the likelihood of loan defaults.

In column 3 of Table 4, we investigate the effect of regional social capital on the loan performance from the lenders' point of view. As detailed in the data and measure section, we calculate IRR to reflect the facts that borrowers may have made a few payments before the loan default and that borrowers can pay off the loans earlier without penalty. It is possible that two similar loans have different rate of return if they have different payment schedules. We report a positive and significant coefficient for social capital, which indicates that loans granted to borrowers in regions with higher levels of social capital yield higher rate of return for the lenders. In addition, we gauge the economic significance of our finding. It appears that a one standard deviation increase of social capital measure can be translated into a 218bps increase in the rate of return, which is obviously economically significant.

[Insert Table 4 about here]

To ensure the robustness of our findings, we re-estimate our model specifications in Table 4 using an alternative measure of social capital and report our results in Table 5. As can be evidenced in Table 5, adopting social trust as the alternative measure of social capital do not change our findings in a material way. For the sake of brevity, we do not report the coefficients for control variables in Table 5.

[Insert Table 5 about here]

4.3 Social capital and credit grades

Table 6 presents the regression results relating the loan interest rate to regional social capital. In the column 1 of Table 6, we use the fitted social capital as our main explanatory variable along with a set of controls for borrower and regional characteristics. In column 2 of Table 6, we include additional control variables for loan characteristics. In columns 3 and 4, we repeat our analysis in columns 1 and 2 by adding credit grades as control variables. Overall, our results reveal a strong and negative relation between regional social capital and loan interest rate. Using information in column 1 as an example, we report that a one-unit increase of social capital reduces the loan interest rate by 62bps. It is not surprising that, in columns 3 and 4, adding credit grades take away the explanatory power of social capital.

[Insert Table 6 about here]

In Lending Club P2P platform, all loans with the same credit grade issued in the same month have the same loan interest rate. A borrower with a higher credit grade pays lower interest rate than one with a lower credit grade. Moreover, the credit grades are assigned by Lending Club using its own algorithm. In this section, we investigate whether, all else being equal, borrowers from regions with higher levels of social capital tend to be assigned higher credit grades. We adopt an ordered Probit model in the analysis because the dependent measure, credit grade, has a natural order. In addition, we control for the FICO scores, borrower characteristics, loan characteristics, and regional characteristics. We use social capital and social trust in column 1 and column 2, respectively. Our empirical results reveal that borrowers in regions with higher levels of social capital tend to be associated with higher credit grades from Lending Club. Regarding other variables, we find that FICO score, credit history, and annual income are the three important determinants of the credit grades.

[Insert Table 7 about here]

4.4 Social capital and voluntary disclosure

In this section, we explore a potential channel through which social capital may affect lender decisions to participate in P2P loans. Online communications and voluntary disclosure are critical in the P2P lending process because borrowers and lenders are connecting exclusively through the online platform in most scenarios and lenders' decisions are based on information provided by the borrowers. Prior studies identify that the funding probability and loan performance in the online marketplace are influenced by the identity claims embedded in the borrower narratives (Herzenstein et al., 2011), language used (Larrimore et al., 2011), voluntary information disclosure (Michels, 2012), writing style (Gao et al., 2018), and the coverage of certain topics (Netzer et al., 2019). In particular, Larrimore et al. (2011) show that word usage and language choices help to facilitate the trust-building in P2P platforms. Even unverifiable disclosures may have a certain influence on lenders' decisions (Michels, 2012). In addition, the study by Valenzuela et al. (2009) reveals a positive relation between online communication and social trust, civic engagement, and political participation. Therefore, we expect that borrowers in regions with higher levels of social capital may voluntarily provide more information in the loan listings.

Following pervious studies on textual analysis (Freedman and Jin, 2017; Lin et al., 2013) we construct two measures on the description sections in the loan listings. *Number of words* is the count of words a borrower used in the voluntary loan description section. *Number of characters* is the count of total characters in the loan description. We regress these two measures on social capital using an IV approach, and report the results in Table 8. We find that our regional measures of social capital has a significantly postive correlation with the number of words and the number of characters in the description of loan listings, which is in line with our expectation.

[Insert Table 8 about here]

4.5 Social capital, online P2P lending, and local bank competitions

Previous studies show that bank market structures have mixed effects on access to credit and cost of debt financing. On the one hand, a banking market characterized by intensive competition may benefit the would-be borrowers with more funds and lower interest rates (Guzman, 2000, Beck et al., 2004). On the other hand, bank competition makes the borrower-specific information more dispersed and increases the cost of information production on the quality of potential borrowers (Petersen and Rajan, 1994, Marquez, 2002). Furthermore, lenders may find it difficult to lock in their loan customers in highly competitive market (Petersen and Rajan, 1994). Therefore, in this section, we investigate the effect of local bank market structure on the relation between social capital and P2P lending.

We rely on FDIC Summary of Deposits (SOD) dataset to construct our measure of local bank market structure. For each 3-digit zip code area, we aggregate the total amounts of deposits for each bank in a particular year. We calculate the deposit-based Herfindahl index (HHI) to capture local bank market structure. On a yearly basis, we define an area as a highly competitive banking market if the HHI is within the 25^{th} percentile. A medium competitive region has the HHI index between the 25^{th} percentile and the 75^{th} percentile of HHI distribution. A low competitive region has the HHI index greater than the 75^{th} percentile. We then partition our sample into three subsamples, and re-run our regressions for each subsample.

Table 9 presents the regression results relating loan default likelihood, loan maturity, IRR, and credit grade to regional social capital. We include the same set of controls variables as detailed in previous sections, and we only present the coefficients of social capital for an easy comparison. Our findings suggest that the effects of social capital on P2P lending activities are stronger in local markets with high levels of bank competition. It is plausible that intensive bank competition results in significant information dispersion about borrowers and the difficulty to screen customers due to increased cost of information production (Marquez, 2002). Borrowers find themselves having difficulty to get access to traditional bank lending and resort to P2P lending market because the severity of information asymmetry problem. As an informal institution, social capital facilitates the transaction by promoting a mutual trust between borrowers and lenders when traditional banks are unwilling to expend resources in costly information production for this group of borrowers.

[Insert Table 9 about here]

4.6 Social capital and P2P lending for small business

Lending Club collects information on the purposes of the loans in the applications. We are particularly interested in the category for small business lending. In our sample, loans in this category amount to about $260 million. Given the volume of P2P lending to small business, we further explore the role of social capital to gain more insights. Small businesses usually lack the access to public debt or equity markets and depend heavily on banks to secure external financing (Mills and Mccarthy, 2014). Small firms are subject to high riskiness, information opacity, and limited tangible assets for collateral (Berger and Udell, 2002). In the online P2P lending platform, it is difficult for the two parties (i.e., the lenders and the borrowers) to meet physically. As a result, lenders may resort to other reference points. We argue that regional social capital could promote a trustworthy environment and facilitate the transaction, especially when the small businesses do not have sufficient track records and assets for collateral (Stuart et al., 1999, Uzzi, 1999).

We therefore form a sample with only loans cauterized as small business lending and we group the rest of all issued loans in another subsample. We re-run our analysis for these two subsamples and report our findings in Table 10. We document that social capital has a stronger effect for loan default likelihood, IRR, and credit grade in the subsample of small business purpose.

[Insert Table 10 about here]

4.7 Social capital and riskiness of P2P loan risks

In this section, we explore whether the effects of social capital on P2P loans are contingent on loan riskiness. We categorize P2P loans according to loan amount, loan maturity and borrower credit grades. Specifically, in Panel A of Table 1, we classify a loan as a high risk (low risk) loan if the loan amount is within the top (bottom) 30 percentile of all the loans granted in a particular year. In Panel B, we classify loans with 60-month (36-month) maturity as high-risk (low-risk) loans. In Panel C, following Polena and Regner (2018),we compare the loans granted to top grade (A) borrowers with loans granted to lower (E, F, G) borrowers. Consistently, we find that social capital has a stronger effect on loan performance (i.e., IRR) for high-risk loans. However, the effects of social capital on loan defaults and loan life are mixed.

[Insert Table 11 about here]

4.8 The effects of civic norm and social network on contractual terms of P2P loans

In this section, we further decompose our measure of social capital (Rupasingha et al., 2006) into two measures, namely civic norm and social network. Following Hasan et al. (2017), we conduct PCA analysis on *pvote* and *respn* and take the first principal component as the measure of *civic norm*. Similarly, the *social network* is the first principal component of the PCA on *assn* and *nccs*. We include civic norm and social network along with the same set of control variables used in our baseline regression, and report the results relating various loan features to both measures in Table 12. For the sake of brevity, we only report the coefficients of civic norm and social network. We find that civic norm has significant coefficients across all model specifications. Note that social network has significant coefficients when it is included in regression along, but it is only significant when loan life is the dependent measure and civic norm is entered. In other words, civic norm appears to play a more important role, and it takes away the significance of social network in some cases. These results are consistent with notion that social capital provides environmental pressures to limit borrowers' opportunistic behaviors.

[Insert Table 12 about here]

5. Summary and Conclusion

In this study, focusing on P2P lending, we attempt to shed further lights on the importance of regional social capital in the capital market. Using a comprehensive dataset from Lending Club from 2008 to 2018, we find that social capital affects both the borrowers and the lenders. Specifically, higher state-level social capital of the borrowers is associated with higher probabilities to secure P2P loans, higher credit grades assigned by Lending Club platform, and lower default likelihood conditional on loan issuance. Lenders realize better rates of return from loans issued to borrowers in states with higher levels of social capital. We find supportive evidence that borrowers in regions with higher levels of social capital are likely to engage voluntary disclosure of more information in loan applications. In addition, we report that social capital has a stronger effect on P2P lending in more competitive banking markets, for small business loans, and for riskier loans.

We recognize that our study is not without limitations. For example, we could implement a test to relate the changes of borrowers' location to P2P lending. Exploring such an exogenous shock would allow us to draw strong causal inferences. Nonetheless, the Lending Club data do not disclose the unique ID of individual borrower ostensibly for privacy concern. Moreover, Michels (2012) analyzes the Prosper data between February 2007 and October 2008 and finds that only 7 percent of borrowers obtained more than 1 loan within their sample and less than 1 percent of borrowers obtained more than 2 loans. We thereby may make a reasonable assumption that the proportion of P2P loans taken by the same borrower in different regions would be even less. For another instance, we could conduct the empirical analysis of the economic effects of social capital focusing on county as the geographic unit. However, the Lending Club data only provide 3-digit Zip code to identify the location of the borrowers.

Despite the above-mentioned possible limitations, we believe that our research contributes to the literature in several important ways. First, to our best knowledge, we are the first to examine the effects of regional social capital on P2P loans. We are able to add novel evidence and extend the existing literature on social capital. Second, we explore several important aspects of P2P lending and perform our analysis from both the borrowers' and the lenders' perspectives to gain a rather comprehensive understanding of the effects of social capital on P2P lending as an increasingly important source of financing. Third, our work investigates the interaction between local bank market structure and P2P lending.

Acknowledgements

The authors would like to thank Iftekhar Hasan, Xiaping Cao, Xian Gu, Yiwei Fang, Priyanka Sharma, and Kai Lisa Lo for valuable discussions and insightful comments.

References

Akerlof, G. A. 1970. The Market for Lemons: Quality Uncertainty and the Market Mechanism. *The Quarterly Journal of Economics,* 84**,** 488-500.

Alesina, A. & La Ferrara, E. 2000. Participation in heterogeneous communities. *The Quarterly Journal of Economics,* 115**,** 847-904.

Beck, T., Demirgüç-Kunt, A. & Maksimovic, V. 2004. Bank competition and access to finance: International evidence. *Journal of Money, Credit and Banking***,** 627-648.

Berger, A. N. & Udell, G. F. 1995. Relationship lending and lines of credit in small firm finance. *Journal of Business***,** 351-381.

Berger, A. N. & Udell, G. F. 2002. Small business credit availability and relationship lending: The importance of bank organisational structure. *The Economic Journal,* 112**,** F32-F53.

Buonanno, P., Montolio, D. & Vanin, P. 2009. Does social capital reduce crime? *The Journal of Law and Economics,* 52**,** 145-170.

Butler, A. W., Cornaggia, J. & Gurun, U. G. 2017. Do local capital market conditions affect consumers' borrowing decisions? *Management Science,* 63**,** 4175-4187.

Coleman, J. S. 1988. Social Capital in the Creation of Human Capital. *American Journal of Sociology,* 94**,** S95-S120.

Cox, D. R. 1972. Regression models and life-tables. *Journal of the Royal Statistical Society: Series B (Methodological),* 34**,** 187-202.

De Roure, C., Pelizzon, L. & Tasca, P. 2016. How does P2P lending fit into the consumer credit market? *Bundesbank Discussion Paper No. 30/2016*

Diamond, D. W. 1984. Financial intermediation and delegated monitoring. *The Review of Economic Studies,* 51**,** 393-414.

Duarte, J., Siegel, S. & Young, L. 2012. Trust and credit: The role of appearance in peer-to-peer lending. *The Review of Financial Studies,* 25**,** 2455-2484.

Elster, J. 1989. Social norms and economic theory. *Journal of Economic Perspectives,* 3**,** 99-117.

Emekter, R., Tu, Y., Jirasakuldech, B. & Lu, M. 2015. Evaluating credit risk and loan performance in online Peer-to-Peer (P2P) lending. *Applied Economics,* 47**,** 54-70.

Freedman, S. & Jin, G. Z. 2017. The information value of online social networks: lessons from peer-to-peer lending. *International Journal of Industrial Organization,* 51**,** 185-222.

Fukuyama, F. 2001. Social capital, civil society and development. *Third World Quarterly,* 22**,** 7-20.

Galak, J., Small, D. & Stephen, A. T. 2011. Microfinance decision making: A field study of prosocial lending. *Journal of Marketing Research,* 48**,** S130-S137.

Gao, Q., Lin, M. & Sias, R. W. 2018. Words matter: The role of texts in online credit markets. *Available at SSRN: https://ssrn.com/abstract=2446114*.

Gonzalez, L. & Loureiro, Y. K. 2014. When can a photo increase credit? The impact of lender and borrower profiles on online peer-to-peer loans. *Journal of Behavioral and Experimental Finance,* 2**,** 44-58.

Greiner, M. E. & Wang, H. 2009. The role of social capital in people-to-people lending marketplaces. *ICIS 2009 proceedings***,** 29.

Guiso, L., Sapienza, P. & Zingales, L. 2004. The Role of Social Capital in Financial Development. *American Economic Review,* 94**,** 526-556.

Guzman, M. G. 2000. Bank structure, capital accumulation and growth: a simple macroeconomic model. *Economic Theory,* 16**,** 421-455.

Herzenstein, M., Sonenshein, S. & Dholakia, U. M. 2011. Tell me a good story and I may lend you money: The role of narratives in peer-to-peer lending decisions. *Journal of Marketing Research,* 48**,** S138-S149.

Hasan, I., Hoi, C. K., Wu, Q. & Zhang, H. 2017. Social capital and debt contracting: Evidence from bank loans and public bonds. *Journal of Financial and Quantitative Analysis,* 52**,** 1017-1047.

Iyer, R., Khwaja, A. I., Luttmer, E. F. & Shue, K. 2015. Screening peers softly: Inferring the quality of small borrowers. *Management Science,* 62**,** 1554-1577.

Knack, S. 2002. Social capital and the quality of government: Evidence from the states. *American Journal of Political Science***,** 772-785.

Knack, S. & Keefer, P. 1997. Does social capital have an economic payoff? A cross-country investigation. *The Quarterly Journal of Economics,* 112**,** 1251-1288.

Larrimore, L., Jiang, L., Larrimore, J., Markowitz, D. & Gorski, S. 2011. Peer to peer lending: The relationship between language features, trustworthiness, and persuasion success. *Journal of Applied Communication Research,* 39**,** 19-37.

Lin, M., Prabhala, N. R. & Viswanathan, S. 2013. Judging borrowers by the company they keep: Friendship networks and information asymmetry in online peer-to-peer lending. *Management Science,* 59**,** 17-35.

Marquez, R. 2002. Competition, adverse selection, and information dispersion in the banking industry. *The Review of Financial Studies,* 15**,** 901-926.

Michels, J. 2012. Do unverifiable disclosures matter? Evidence from peer-to-peer lending. *The Accounting Review,* 87**,** 1385-1413.

Mills, K. & Mccarthy, B. 2014. The state of small business lending: Credit access during the recovery and how technology may change the game. *Harvard Business School General Management Unit Working Paper*.

Nahapiet, J. & Ghoshal, S. 1998. Social capital, intellectual capital, and the organizational advantage. *Academy of Management Review,* 23**,** 242-266.

Netzer, O., Lemaire, A. & Herzenstein, M. 2019. When words sweat: Identifying signals for loan default in the text of loan applications. *Journal of Marketing Research,* 56**,** 960-980.

Paravisini, D., Rappoport, V. & Ravina, E. 2017. Risk Aversion and Wealth: Evidence from Person-to-Person Lending Portfolios. *Management Science,* 63**,** 279-297.

Petersen, M. A. 2009. Estimating standard errors in finance panel data sets: Comparing approaches. *The Review of Financial Studies,* 22**,** 435-480.

Petersen, M. A. & Rajan, R. G. 1994. The benefits of lending relationships: Evidence from small business data. *The Journal of Finance,* 49**,** 3-37.

Petersen, M. A. & Rajan, R. G. 1995. The effect of credit market competition on lending relationships. *The Quarterly Journal of Economics,* 110**,** 407-443.

Polena, M. & Regner, T. 2018. Determinants of borrowers' default in P2P lending under consideration of the loan risk class. *Games,* 9**,** 82.

Putnam, R. 2001. Social capital: Measurement and consequences. *Canadian Journal of Policy Research,* 2**,** 41-51.

Putnam, R. D. 1995. Tuning in, tuning out: The strange disappearance of social capital in America. *PS: Political science & politics,* 28**,** 664-684.

Putnam, R. D., Leonardi, R. & Nanetti, R. 1993. *Making Democracy Work: Civic Traditions in Modern Italy*. Princeton university press.

Rupasingha, A., Goetz, S. J. & Freshwater, D. 2006. The production of social capital in US counties. *The Journal of Socio-economics,* 35**,** 83-101.
Serrano-Cinca, C. & Gutiérrez-Nieto, B. 2016. The use of profit scoring as an alternative to credit scoring systems in peer-to-peer (P2P) lending. *Decision Support Systems*, 89**,** 113-122.
Sobel, J. 2002. Can we trust social capital? *Journal of Economic Literature,* 40**,** 139-154.
Stiglitz, J. E. & Weiss, A. 1981. Credit Rationing in Markets with Imperfect Information. *The American Economic Review,* 71**,** 393-410.
Stuart, T. E., Hoang, H. & Hybels, R. C. 1999. Interorganizational endorsements and the performance of entrepreneurial ventures. *Administrative Science Quarterly,* 44**,** 315-349.
Uhlaner, C. J. 1989. Rational turnout: The neglected role of groups. *American Journal of Political Science***,** 390-422.
Uzzi, B. 1999. Embeddedness in the making of financial capital: How social relations and networks benefit firms seeking financing. *American Sociological Review***,** 481-505.
Valenzuela, S., Park, N. & Kee, K. F. 2009. Is there social capital in a social network site?: Facebook use and college students' life satisfaction, trust, and participation. *Journal of Computer-mediated Communication,* 14**,** 875-901.
Wei, Z. & Lin, M. 2017. Market mechanisms in online peer-to-peer lending. *Management Science,* 63**,** 4236-4257.
Woolcock, M. 2001. The place of social capital in understanding social and economic outcomes. *Canadian Journal of Policy Research,* 2**,** 11-17.
Xu, J. J. & Chau, M. 2018. Cheap talk? The impact of lender-borrower communication on peer-to-peer lending outcomes. *Journal of Management Information Systems,* 35**,** 53-85.

Table 1. Summary statistics

Variable	N	Mean	SD	Minimum	P25	Median	P75	Maximum
Panel A: Reject Sample								
Social capital	5,667,774	-0.66	0.81	-2.54	-1.22	-0.76	-0.17	7.18
Reject	5,667,774	0.75	0.44	0	0	1	1	1
FICO	5,667,774	2.62	1.26	1	2	3	3	6
Loan amount	5,667,774	15.15	10.51	0.5	6.5	12	20	40
Debt-to-income ratio	5,667,774	24.62	18.82	0	11.67	20.75	33.08	100
Employment length	5,667,774	2.05	3.50	0	0	0	3	10
Population (log)	5,667,774	2.40	0.87	-0.08	1.80	2.34	3.04	3.67
Age	5,667,774	38.05	2.04	33.90	36.50	38.20	39.40	42.20
Education	5,667,774	30.35	4.54	21.00	27.10	29.86	33.25	42.09
Unemployment	5,667,774	6.55	1.77	3.50	5.40	6.00	7.30	12.30
Income per capita (log)	5,667,774	10.80	0.16	10.48	10.69	10.78	10.93	11.14
Panel B: Issued Sample								
Social capital	1,438,799	-0.66	0.83	-2.54	-1.22	-0.76	-0.17	7.18
Default	1,438,799	0.21	0.41	0	0	0	0	1
Loan life	1,436,529	21.46	12.70	0	11	20	33	70
IRR	1,432,887	-2.19	34.30	-99.99	6.30	11.26	15.26	29.48
Interest rate	1,438,799	13.22	4.78	5.31	9.75	12.74	15.99	30.99
Loan amount	1,438,799	14.40	8.72	0.5	7.925	12	20	40
Loan term	1,438,799	41.84	10.30	36	36	36	36	60
Credit grade	1,438,799	5.26	1.29	1	5	5	6	7
Annual income	1,438,799	4.19	0.52	2.93	3.83	4.17	4.51	5.56
Debt-to-income ratio	1,438,799	18.04	8.39	0	11.75	17.55	23.94	49.96
Employment length	1,438,799	5.63	3.84	0	2	6	10	10
Revolving utilization	1,437,892	51.55	24.44	1.1	33.1	51.9	70.5	98.2
Own home	1,438,799	0.60	0.49	0	0	1	1	1
Delinquency	1,438,799	0.30	0.73	0	0	0	0	4
Credit inquiries	1,438,798	0.64	0.91	0	0	0	1	4
Accounts opened	1,438,799	11.58	5.29	3	8	11	14	29
Public records	1,438,799	0.20	0.46	0	0	0	0	2
Population (log)	1,438,799	2.41	0.87	-0.12	1.80	2.33	3.03	3.67
Age	1,438,799	37.98	2.02	33.90	36.40	38.10	39.30	42.20
Education	1,438,799	31.35	4.59	21.78	28.20	30.94	34.50	42.74
Unemployment	1,438,799	6.63	1.63	3.60	5.50	6.30	7.30	11.90
Income per capita (log)	1,438,799	10.81	0.16	10.48	10.69	10.79	10.93	11.14

Notes: This table reports the summary statistics for all variables in the sample of rejected loans and the sample of issued loans.

Table 2. The distribution of credit grade in the sample of issued loans

Credit grade	N	Total amount	Default rate	IRR	Loan maturity	Interest rate
G = Highest risk	9.39	191.87	50.19	-18.93	17.03	27.76
F	33.06	627.71	45.53	-13.51	19.43	24.99
E	98.75	1,730.05	39.10	-9.63	20.30	21.23
D	214.68	3,276.69	31.46	-6.39	20.42	17.78
C	409.32	5,806.41	23.57	-3.06	21.10	14.03
B	420.11	5,558.98	14.29	0.95	22.12	10.66
A = Lowest risk	253.48	3,523.38	6.49	2.50	22.69	7.10
Total	1,438.80	20,715.07	20.77	-2.19	21.45	13.22

Notes: This table presents the distribution of each credit grades in the issued sample. The number of loans is in thousands. The total amount is in thousand dollars. The default rate, IRR and Interest rate are in percentage.

Table 3. Social capital and P2P loan rejection

Independent variables	Dependent variable Reject			
	IV-Probit		Probit	
	(1)		(2)	
Social capital (fitted)	-0.047***	(0.005)		
Social trust			-0.086***	(0.010)
FICO = 2	-0.455***	(0.004)	-0.455***	(0.005)
FICO = 3	-0.463***	(0.004)	-0.463***	(0.005)
FICO = 4	-0.380***	(0.005)	-0.380***	(0.005)
FICO = 5	-0.342***	(0.005)	-0.343***	(0.006)
Loan amount	0.014***	(0.000)	0.014***	(0.000)
Debt-to-income ratio	0.018***	(0.000)	0.018***	(0.000)
Employment length	-0.206***	(0.000)	-0.206***	(0.000)
Population (log)	-0.015***	(0.002)	0.005***	(0.001)
Age	0.015***	(0.001)	0.012***	(0.000)
Education	-0.008***	(0.000)	-0.008***	(0.000)
Unemployment	-0.005***	(0.001)	0.004***	(0.001)
Income per capita (log)	0.137***	(0.011)	0.125***	(0.011)
Constant	0.016	(0.111)	0.402***	(0.109)
Year FE	Yes		Yes	
Observations	4,440,237		4,440,237	
Wald chi2 test of exogeneity	139.3***			

Notes: First column present the second stage of IV-Probit regression results. The second column presents the Probit regression results. Variables are introduced in section 3. Values between parentheses denote standard errors clustered by the first three-digit zip code. *,**,***Significant at the 10, 5 and 1 percent levels, respectively.

Table 4. Social capital and loan characteristics

Independent variables	Dependent variables					
	Default likelihood		Loan life		IRR	
	IV-Probit		Cox proportional hazard model		2SLS	
	(1)		(2)		(3)	
Social capital (fitted)	-0.147**	(0.058)			2.635**	(1.097)
Social capital			0.973***	(0.009)		
Credit grade = 2	0.049***	(0.016)	1.077***	(0.019)	1.709***	(0.622)
Credit grade = 3	0.076***	(0.015)	1.150***	(0.020)	2.351***	(0.602)
Credit grade = 4	0.081***	(0.017)	1.211***	(0.024)	1.876***	(0.621)
Credit grade = 5	0.041**	(0.019)	1.164***	(0.027)	1.623**	(0.659)
Credit grade = 6	-0.078***	(0.021)	0.944**	(0.025)	1.158	(0.708)
Credit grade = 7	-0.296***	(0.024)	0.558***	(0.017)	-1.079	(0.767)
Annual income	-0.225***	(0.005)	0.811***	(0.006)	3.707***	(0.108)
Loan amount	0.009***	(0.000)	1.010***	(0.000)	-0.215***	(0.006)
Loan term = 60	0.305***	(0.004)	1.309***	(0.008)	-6.911***	(0.116)
Debt-to-income ratio	0.006***	(0.000)	1.006***	(0.000)	-0.167***	(0.006)
Employment length	-0.008***	(0.001)	0.990***	(0.001)	0.216***	(0.012)
Revolving utilization	0.002***	(0.000)	0.999***	(0.000)	-0.020***	(0.001)
Own home	-0.162***	(0.007)	0.829***	(0.006)	3.428***	(0.131)
Delinquency	0.037***	(0.002)	1.015***	(0.003)	-0.613***	(0.041)
Credit inquiries	0.061***	(0.002)	1.100***	(0.003)	-1.295***	(0.042)
Accounts opened	0.006***	(0.000)	1.006***	(0.000)	-0.071***	(0.008)
Public records	0.059***	(0.004)	1.053***	(0.005)	-0.783***	(0.085)
Interest rate	0.036***	(0.001)	1.070***	(0.001)	-0.584***	(0.027)
Population (log)	-0.051	(0.044)	1.030***	(0.011)	0.813	(0.814)
Age	0.015***	(0.004)	1.004	(0.003)	-0.276***	(0.075)
Education	-0.011***	(0.003)	1.001	(0.002)	0.197***	(0.055)
Unemployment	-0.022**	(0.009)	1.012*	(0.007)	0.338**	(0.158)
Income per capita (log)	0.354***	(0.135)	0.952	(0.036)	-6.173**	(2.449)
Constant	-4.634***	(1.327)			60.050**	(24.096)
Year FE	Yes		Yes		Yes	
Loan Purpose FE	Yes		Yes		Yes	
Observations	1,437,890		1,427,281		1,431,994	
Wald chi2 test of exogeneity	3.76*					
Endogeneity test					18.91***	

Notes: This table reports the effects of social capital on P2P loan performance. Variables are introduced in section 3. Values between parentheses denote standard errors clustered by the first three-digit zip code. *, **, ***Significant at the 10, 5 and 1 percent levels, respectively.

Table 5. Alternative measure of social capital and loan characteristics

Independent variables	Dependent variables					
	Default likelihood		Loan life		IRR	
	Probit		Cox proportional hazard model		OLS	
	(1)		(2)		(3)	
Social trust	-0.151***	(0.048)	0.860**	(0.057)	2.486***	(0.937)
All control variables	Yes		Yes		Yes	
Year FE	Yes		Yes		Yes	
Loan Purpose FE	Yes		Yes		Yes	
Observations	1,437,890		1,427,281		1,431,994	

Notes: This table reports the results of the robustness check on the effects of social capital on P2P loan performance using social trust as alternative measure. Only the coefficients on social trust are reported. Variables are introduced in section 3. Values between parentheses denote standard errors clustered by the first three-digit zip code. *, **, ***Significant at the 10, 5 and 1 percent levels, respectively.

Table 6. Social capital, credit grade, and interest rate

Independent variables	Dependent variable - Interest rate							
	(1)		(2)		(3)		(4)	
Social capital (fitted)	-0.619**	(0.241)	-0.585**	(0.232)	-0.045**	(0.019)	-0.053**	(0.022)
Credit grade = 2					-2.726***	(0.028)	-2.700***	(0.029)
Credit grade = 3					-6.407***	(0.025)	-6.329***	(0.026)
Credit grade = 4					-9.914***	(0.025)	-9.765***	(0.026)
Credit grade = 5					-13.624***	(0.026)	-13.431***	(0.027)
Credit grade = 6					-16.950***	(0.026)	-16.693***	(0.027)
Credit grade = 7					-20.375***	(0.026)	-20.064***	(0.027)
Annual income	-1.057***	(0.017)	-2.093***	(0.017)	-0.098***	(0.003)	-0.191***	(0.003)
Debt-to-income ratio	0.072***	(0.001)	0.048***	(0.001)	0.006***	(0.000)	0.005***	(0.000)
Employment length	0.008***	(0.002)	-0.012***	(0.001)	0.001**	(0.000)	-0.000	(0.000)
Revolving utilization	0.051***	(0.000)	0.052***	(0.000)	0.004***	(0.000)	0.005***	(0.000)
Own home	-0.505***	(0.024)	-0.713***	(0.020)	-0.050***	(0.003)	-0.069***	(0.003)
Delinquency	0.451***	(0.006)	0.490***	(0.005)	0.027***	(0.001)	0.037***	(0.001)
Credit inquiries	1.240***	(0.006)	1.198***	(0.005)	0.074***	(0.001)	0.091***	(0.001)
Accounts opened	-0.004***	(0.001)	0.002**	(0.001)	-0.001***	(0.000)	-0.001**	(0.000)
Public records	0.771***	(0.014)	0.839***	(0.012)	0.051***	(0.003)	0.069***	(0.003)
Loan amount			0.064***	(0.001)			0.006***	(0.000)
Loan term = 60			4.300***	(0.010)			0.228***	(0.003)
Population (log)	-0.221	(0.193)	-0.204	(0.187)	-0.011	(0.014)	-0.014	(0.017)
Age	0.040***	(0.015)	0.027*	(0.014)	0.003**	(0.001)	0.003*	(0.002)
Education	0.001	(0.013)	-0.005	(0.011)	-0.001	(0.001)	-0.001	(0.001)
Unemployment	-0.061*	(0.031)	-0.045*	(0.027)	-0.008**	(0.003)	-0.008**	(0.003)
Income per capita (log)	0.408	(0.588)	0.627	(0.549)	0.067	(0.045)	0.085*	(0.052)
Constant	5.798	(5.738)	7.559	(5.378)	25.439***	(0.446)	25.279***	(0.512)
Year FE	Yes		Yes		Yes		Yes	
Loan Purpose FE	No		Yes		No		Yes	
Observations	1,437,890		1,437,890		1,437,890		1,437,890	
R-squared	0.176		0.393		0.926		0.926	
Endogeneity test	28.52***		26.04***		14.31***		15.61***	
Kleibergen-Paap rk Wald F statistic (Weak identification test)	6.2		6.2		6.2		6.2	
Kleibergen-Paap rk LM statistic (Underidentification test)	6.5**		6.4**		6.4**		6.4**	

Notes: This table reports the effect of social capital on interest rates. All four columns present the second-stage results of 2SLS regression. The first model uses borrower characteristics and state controls. The second model adds controls for loan characteristic. The last two models additionally control for the credit grades. Variables are introduced in section 3. Values between parentheses denote standard errors clustered by the first three-digit zip code. *, **,***Significant at the 10, 5 and 1 percent levels, respectively.

Table 7. Social capital and credit grades

Independent variables	Dependent variable: Credit grade			
	Ordered Probit		Ordered Probit	
	(1)		(2)	
Social capital	0.013***	(0.004)		
Social trust			0.063***	(0.024)
Annual income	0.394***	(0.004)	0.394***	(0.004)
Loan amount	-0.019***	(0.000)	-0.019***	(0.000)
Loan term = 60	-1.235***	(0.003)	-1.235***	(0.003)
Debt-to-income ratio	-0.021***	(0.000)	-0.021***	(0.000)
Employment length	-0.002***	(0.000)	-0.002***	(0.000)
Revolving utilization	-0.005***	(0.000)	-0.005***	(0.000)
Own home	0.089***	(0.004)	0.090***	(0.004)
Delinquency	-0.058***	(0.001)	-0.058***	(0.001)
Credit inquiries	-0.305***	(0.001)	-0.305***	(0.001)
Accounts opened	0.000	(0.000)	0.000	(0.000)
Public records	-0.056***	(0.003)	-0.056***	(0.003)
Total accounts	0.009***	(0.000)	0.009***	(0.000)
FICO	0.522***	(0.002)	0.523***	(0.002)
Credit history	0.242***	(0.003)	0.242***	(0.003)
Population (log)	-0.003	(0.003)	-0.007**	(0.003)
Age	0.003**	(0.002)	0.004***	(0.001)
Education	0.002*	(0.001)	0.002*	(0.001)
Unemployment	-0.015***	(0.003)	-0.018***	(0.003)
Income per capita (log)	-0.140***	(0.046)	-0.140***	(0.044)
Year FE	Yes		Yes	
Loan Purpose FE	Yes		Yes	
Observations	1,437,890		1,437,890	
(Pseudo)R2	0.179		0.179	

Notes: This table reports the Ordered Probit regression results. Variables are introduced in section 3. Values between parentheses denote standard errors clustered by the first three-digit zip code. *,**,***Significant at the 10, 5 and 1 percent levels, respectively.

Table 8. Social capital and voluntary disclosure

Independent variables	Dependent variables			
	Number of words		Number of characters	
	(1)		(2)	
Social capital (fitted)	2.694***	(0.946)	10.987***	(3.866)
Credit grade = 2	-5.706***	(1.854)	-25.640***	(7.998)
Credit grade = 3	-8.127***	(1.778)	-36.231***	(7.675)
Credit grade = 4	-13.901***	(1.819)	-61.448***	(7.829)
Credit grade = 5	-17.638***	(1.923)	-77.391***	(8.280)
Credit grade = 6	-20.844***	(2.135)	-91.339***	(9.142)
Credit grade = 7	-26.750***	(2.408)	-116.850***	(10.286)
Annual income	-1.819***	(0.315)	-2.640**	(1.345)
Loan amount	0.368***	(0.018)	1.600***	(0.077)
Loan term = 60	0.484	(0.337)	2.231	(1.449)
Debt-to-income ratio	-0.009	(0.018)	-0.036	(0.074)
Employment length	-0.525***	(0.031)	-2.394***	(0.132)
Revolving utilization	0.033***	(0.006)	0.134***	(0.026)
Own home	-3.521***	(0.303)	-14.465***	(1.291)
Delinquency	-0.305	(0.188)	-1.345*	(0.805)
Credit inquiries	-0.185	(0.126)	-1.001*	(0.535)
Accounts opened	0.092***	(0.030)	0.328**	(0.130)
Public records	-1.136***	(0.314)	-4.962***	(1.324)
Interest rate	-1.416***	(0.104)	-6.134***	(0.447)
Population (log)	0.714	(0.791)	2.834	(3.111)
Age	-0.775***	(0.154)	-3.123***	(0.664)
Education	0.108	(0.080)	0.537	(0.341)
Unemployment	0.845***	(0.181)	3.435***	(0.772)
Income per capita (log)	-5.024*	(2.909)	-25.355**	(11.949)
Constant	168.624***	(29.539)	733.199***	(120.674)
Year FE	Yes		Yes	
Loan Purpose FE	Yes		Yes	
Observations	106,915		106,915	
R squared	0.208		0.213	
Endogeneity test	8.64***		7.06***	
Kleibergen-Paap rk Wald F statistic (Weak identification test)	8.3		8.3	
Kleibergen-Paap rk LM statistic (Underidentification test)	8.5***		8.6***	

Notes: This table reports the effect of social capital on the length of loan description. All two columns present the second-stage results of 2SLS regression. The dependent variable in the first column is the number of words used in the description. The dependent variable in the second column is the number of characters in the description. Variables are introduced in section 3. Values between parentheses denote standard errors clustered by the first three-digit zip code. *, **,***Significant at the 10, 5 and 1 percent levels, respectively.

Table 9. The effect of local bank competitions

Dependent variable	Model	Local bank market competition					
		High competition		Medium competition		Low competition	
Default likelihood	IV-Probit	-0.107***	(0.029)	-0.242	(0.240)	-0.069	(0.069)
Loan maturity	Cox proportional hazard model	0.954**	(0.020)	0.982**	(0.009)	0.991	(0.014)
IRR	2SLS	2.122***	(0.591)	4.422	(4.639)	0.821	(1.243)
Credit grade	Ordered Probit	0.066***	(0.010)	0.011**	(0.005)	0.004	(0.007)

Notes: This table reports the coefficient of social capital among different level of bank competitions. The first row presents the second stage estimation of IV-Probit regressions. The second row presents the results of the Cox hazard model. The third row presents the second stage estimation results of the 2SLS model. The last row reports the coefficients of Ordered Probit regression. Variables are introduced in section 3. Values between parentheses denote standard errors clustered by the first three-digit zip code. *,**,***Significant at the 10, 5 and 1 percent levels, respectively.

Table 10. Social capital and P2P loans purpose

Dependent variable	Model	Loan Purpose			
		Small Business lending		Non-Small Business lending	
Default likelihood	IV-Probit	-0.179**	(0.073)	-0.148***	(0.009)
Loan maturity	Cox proportional hazard model	0.991	(0.018)	0.973***	(0.008)
IRR	2SLS	7.344*	(4.197)	2.573**	(1.071)
Credit grade	Ordered Probit	0.024*	(0.014)	0.013***	(0.004)

Notes: This table reports the coefficient of social capital among different loan purpose. The issued sample is split into two subsamples, one including all loans with "small business" as loan purpose and the other including the rest loans. The first row presents the second stage estimation of IV-Probit regressions. The second row presents the results of the Cox hazard model. The third row presents the second stage estimation results of the 2SLS model. The last row reports the coefficients of Ordered Probit regression. Variables are introduced in section 3. Values between parentheses denote standard errors clustered by the first three-digit zip code. *, **, ***Significant at the 10, 5 and 1 percent levels, respectively.

Table 11. Social capital and P2P loans risks

Panel A. Subsample by loan amount		High risk		Low risk	
Dependent variable	Model	Loan amount > P30		Loan amount < P30	
Default likelihood	IV-Probit	-0.199**	(0.092)	-0.115**	(0.048)
Loan life	Cox proportional hazard model	0.975***	(0.008)	0.970**	(0.012)
IRR	2SLS	3.970**	(1.974)	1.718**	(0.758)
Credit grade	Ordered Probit	0.013***	(0.004)	0.014***	(0.005)
Panel B. Subsample by loan length		**High risk**		**Low risk**	
Dependent variable	Model	Term = 60		Term = 36	
Default likelihood	IV-Probit	-0.149***	(0.053)	-0.153**	(0.064)
Loan life	Cox proportional hazard model	0.990	(0.007)	0.963***	(0.010)
IRR	2SLS	3.744***	(1.411)	2.312**	(1.019)
Credit grade	Ordered Probit	0.012***	(0.005)	0.013***	(0.004)
Panel C. Subsample by credit grade		**High risk**		**Low risk**	
Dependent variable	Model	Credit grade = E,F,G		Credit grade = A	
Default likelihood	IV-Probit	-0.132**	(0.062)	-0.182**	(0.088)
Loan life	Cox proportional hazard model	0.993	(0.009)	0.937***	(0.014)
IRR	2SLS	4.120**	(2.026)	1.524*	(0.792)

Notes: This table reports the coefficient of social capital among subsamples with different risk measures. We construct subsamples by loan amount, loan terms, and borrower credit grade, respectively. The first row presents the second stage estimation of IV-Probit regressions. The second row presents the results of the Cox hazard model. The third row presents the second stage estimation results of the 2SLS model. The last row reports the coefficients of Ordered Probit regression. Variables are introduced in section 3. Values between parentheses denote standard errors clustered by the first three-digit zip code. *, **, ***Significant at the 10, 5 and 1 percent levels, respectively.

Table 12. The effects of civic norm and social networks

Independent variables	Dependent variables							
	Default		Loan life		IRR		Credit grade	
	Probit		Cox proportional hazard model		OLS		Ordered Probit	
	(1)		(2)		(3)		(4)	
Civic norm	-0.028***	(0.004)	0.976***	(0.006)	0.483***	(0.092)	0.027***	(0.005)
Social network	-0.004	(0.007)	0.974***	(0.009)	0.082	(0.129)	0.000	(0.004)
All control variables	Yes		Yes		Yes		Yes	
Year FE	Yes		Yes		Yes		Yes	
Loan Purpose FE	Yes		Yes		Yes		Yes	
Observations	1,437,890		1,427,281		1,431,994		1,437,890	
(Pseudo)R2	0.0924		0.0251		0.0641		0.179	

Notes: This table reports the results of decomposing social capital into civic norm and social network. Only the coefficients of civic norm and social network are reported. Variables are introduced in section 3. Values between parentheses denote standard errors clustered by the first three-digit zip code. *, **, ***Significant at the 10, 5 and 1 percent levels, respectively.