Pacific-Basin Finance Journal 77 (2023) 101922

Contents lists available at ScienceDirect

Pacific-Basin Finance Journal

journal homepage: www.elsevier.com/locate/pacfin

The value of family social capital in informal financial markets: Evidence from China

Wanning Li [a], Xiuping Hua [b,*]

[a] Nottingham University Business School China, University of Nottingham Ningbo China, Ningbo 315100, China
[b] Ningbo Centre for New Structural Economics, UNNC-NFTZ Blockchain Laboratory, University of Nottingham Ningbo China, Ningbo 315100, China

ARTICLE INFO

JEL codes:
G23
G28

Keywords:
Family social capital
Loan performance
Credit markets
Risk management
Information asymmetry

ABSTRACT

This study investigates whether and how family social capital influences borrowers' loan performance in informal financial markets. Using a unique dataset collected from an informal financial institution in Eastern China from 2014 to 2018, we show that a higher level of family social capital is associated with better loan performance, captured by lower default probability, lower loss ratio, and lower loss amount. The effect is more pronounced for borrowers with less access to formal financial markets and those with poorer financial conditions, while it is less pronounced for borrowers with higher risk preferences and those born in regions with high social trust. These results continue to hold after addressing endogeneity concerns and conducting a series of robustness tests. Overall, our study provides evidence that borrowers' family social capital creates value in the informal financial markets.

1. Introduction

The main functions of financial institutions are risk assessment and control (Garmaise, 2015). Identifying factors or cues that can alleviate information asymmetry between counterparties has received special attention from academia (e.g., Petersen and Rajan, 2002; Norden and Weber, 2010; Garmaise, 2015; Fisman et al., 2017; Acheampong and Elshandidy, 2021). According to previous studies, social capital has been documented to have significant effects on corporate financing activities, such as reducing the cost of bank debt (Engelberg et al., 2012; Hasan et al., 2017), facilitating access to trade credit (Liu et al., 2016), reducing the cost of equity (Ferris et al., 2017a), influencing capital structure (Huang and Shang, 2019; Dudley, 2021), and increasing corporate financing efficiency (Yin et al., 2022). However, the role of social capital in individual financing, especially in informal financial markets, is not yet fully understood. This study fills this gap by examining whether individual social capital can influence loan performance in informal financial markets.

Garmaise (2015) finds that borrowers in formal financial markets usually have high credit scores, while informal financial markets usually serve the lower end of the market, such as small business owners, households, and individuals in rural areas (Ayyagari et al., 2010).[1] Owing to more opaque customers and weaker legal protections, informal financial markets often suffer from higher credit risk and severe transaction frictions. Given that financing activity relies greatly on the trust level between counterparties (Guiso et al., 2004), we expect that social capital, defined as "networks, norms, and trust that enable participants to act together more effectively to

* Corresponding author at: Nottingham University Business School China, UNNC-NFTZ Blockchain Laboratory, University of Nottingham Ningbo China, Ningbo, 315100, China.
E-mail address: Xiuping.hua@nottingham.edu.cn (X. Hua).

[1] Informal financing in China takes various forms, such as moneylenders, credit cooperatives, pawnshops, borrowing from families or friends, and trade credit (Allen et al., 2019).

https://doi.org/10.1016/j.pacfin.2022.101922
Received 5 March 2022; Received in revised form 3 November 2022; Accepted 13 December 2022
Available online 15 December 2022

pursue shared objectives" (Putnam, 1993), should play an important role in informal financial markets.

This study focuses on a specific type of social capital, family social capital, captured by the number of family members of an individual. We propose that family social capital is associated with better loan performance for two reasons. First, family social capital can improve borrowers' loan performance by increasing their repayment willingness. Ambrus et al. (2014) demonstrate that pre-existing relationships, such as family ties and friendships, can serve as social collateral to enforce people to fulfill their obligations. Karaivanov and Kessler (2018) also show that default on loans pledged with social collateral is costlier than default on loans pledged with physical collateral. Thus, borrowers with larger family sizes are perceived to pledge higher values of collateral and face greater penalties if a default occurs. Second, family social capital can improve loan performance by increasing borrowers' repayment abilities. More family members can provide an individual with greater access to social connections and resources. In addition, borrowing money from family and friends is believed to be an important and inexpensive form of informal financing (Lee and Persson, 2016). Therefore, borrowers with high family social capital are expected to have more financing channels to ensure liquidity, and thus, have higher repayment ability.

To empirically test our conjecture, we collect data from an informal financial institution for the period 2014 to 2018. This informal financial institution was first established in 2014 in Ningbo, an eastern city in China. It mainly provides small and short-term loans to individuals. Most of their customers are small business owners. All loans issued by the institution are credit loans without any physical collateral. To compensate for the high credit risk, relatively higher interest rates are charged than in formal financial markets. The city of Ningbo, where the financial institution is located, is one of China's oldest cities, with a long history of informal financial markets (Hua et al., 2016). Following Bedendo et al. (2020), our intra-regional study setting has great advantages in distinguishing the effects of individual social capital from other macro-specific factors.

We measure our key explanatory variable, family social capital, as the number of an individual's family members. We collect the information from the borrowers' residence booklets. In the Chinese context, family members in the same residence booklet usually live in the same house. Thus, our proxy of family social capital captures the closest and most valuable ties a person possesses. Sanders and Nee (1996) illustrate that a family is an important form of social capital. Family size has also been used as a direct measure of social capital in prior literature (e.g., Folland, 2006). Compared with proxies commonly used in previous studies, such as the community-level index compiled by Rupasingha et al. (2006), the Putnam Index (Putnam, 1993), and measures from experiments or surveys (Karlan, 2005), our proxy has several advantages. First, because community-level social capital reflects how groups behave rather than what individuals possess, our proxy responds to Glaeser et al. (2002) who suggest that the formation of social capital should be analyzed at the individual level. Second, our proxy satisfies the requirement proposed by Granovetter (2005), who indicates that social capital should be generated outside the studied economic setting. Third, our proxy captures pre-existing relationships, making the study setting less affected by the reverse causality problem.

The empirical results show that higher family social capital is associated with better loan performance, manifesting as lower default probability, the proportion of default amount, and loss amount. In particular, an increase of one unit in family social capital reduces the probability of default by 10.5% and the total loss amount by 29.39%. These results are in line with our expectations and demonstrate the value of family social capital in informal financial markets. By examining cross-sectional variation, we further find that the effect of family social capital on loan performance is more pronounced for borrowers with less access to formal financial markets for borrowers with poorer financial conditions, while it is less pronounced for borrowers with higher risk preferences and those born in regions with high social trust.

We adopt three strategies to address endogeneity problems. First, we use observable variables to estimate the bias from unobservable variables. Specifically, we calculate the selection ratios to see how much greater the effects of unobservable variables would need to be relative to observable variables to completely explain away the effect of family social capital. We also adopt an instrumental variable approach to further mitigate concerns about omitted variable bias. Since we cannot observe the loan performance of customers who are not successfully funded, our results might suffer from the sample selection bias. We further use the Heckman two-step model to deal with sample selection bias. Our results still hold after addressing the potential endogeneity problems. Moreover, by conducting a series of robustness tests, we show that our results remain robust when we include different dimensions of fixed effects and use observations excluding repeated borrowers, local borrowers, and extreme loans.

This study contributes to the literature in several ways. First, it contributes to the literature on personal financing activities in informal financial markets. Previous studies on financing activities in informal financial markets usually focus on the firm level, such as investigating trade credit financing (e.g., El Ghoul and Zheng, 2016; Liu et al., 2016; Kong et al., 2020). Owing to data constraints, studies on personal financing in informal financial markets are scarce. With the development of information technology, one strand of literature has emerged to show personal financing activities on online peer-to-peer (P2P) lending platforms. As described by Chen et al. (2020), P2P platform participants are relatively young and well-educated. For individuals who are unable to use new technology or do not even have a smartphone, traditional offline informal financial markets are important resources for meeting their financing needs. Our study provides evidence of personal financing activities in such markets with a unique dataset and thus adds to the literature.

Second, this study contributes to the literature on individual loan outcomes. Existing studies document several determinants that have significant effects on individual loan outcomes, such as gender (Chen et al., 2020; Shahriar et al., 2020), image (Duarte et al., 2012), reporting behavior (Garmaise, 2015; Dorfleitner et al., 2016) and borrower-lender relationships (Fisman et al., 2017; Beck et al., 2018). Our study enriches this strand of literature by providing evidence that family social capital has a significant influence on loan performance.

Third, our study contributes to the literature on the role of social capital in informal financial markets. Previous studies have provided evidence for the role of social capital in online P2P platforms. For example, Lin et al. (2013) find that borrowers who build more online "friendships" are more likely to be funded and have lower default rates. Hasan et al. (2022) construct a provincial social

capital index and find that borrowers located in regions with high social capital have higher funding success rates and lower default rates. However, evidence from online platforms does not apply to the setting of offline financial institutions, since offline financial institutions usually serve local customers who have the same level of regional social capital. Therefore, we focus on a particular type of individual social capital, namely, family social capital, to complement the literature. Our study also differs from the literature on group lending, which requires joint liability to enforce repayment, in that our unique research setting helps us better distinguish the effects of social capital from other confounders such as joint liability.

Finally, our study contributes to the literature by highlighting a potential channel through which family social capital promotes self-employment. Sanders and Nee (1996) document that Asian and Hispanic immigrants rely heavily on their family social capital, captured by family composition, to start their businesses. Since credit constraints are an important factor impeding Chinese households from starting businesses (Ma et al., 2019), our study provides implications for promoting self-employment by showing the value of family social capital in facilitating households' access to credit.

The remainder of this paper is organized as follows. Section 2 reviews related literature and develops our hypotheses. Section 3 describes the data and reports descriptive statistics. Section 4 presents the empirical results and analysis. Section 5 addresses endogeneity concerns and provides the results of additional robustness tests. Section 6 concludes the paper.

2. Literature review and hypothesis development

2.1. The role of social capital in financial markets

Bourdieu (1986) states that "social capital is an attribute of an individual in a social context". Granovetter (2005) describes three main mechanisms through which social capital can affect economic outcomes: influencing information flow and quality, contributing sources of reward and punishment, and nourishing trust. As financial contracts are trust-intensive contracts that depend highly on how the two counterparties trust each other (Guiso et al., 2004), social capital, as an important determinant of trust, is believed to play an important role in financial markets.

Studies have documented that social capital can significantly facilitate corporate financing. By measuring social capital at the managerial level, existing studies show that higher managerial social capital can significantly reduce the cost of bank debt (Engelberg et al., 2012), reduce the cost of equity (Ferris et al., 2017a), lower the loan spreads of newly issued loans (Fogel et al., 2018), and improve corporate financing efficiency (Yin et al., 2022). In addition, Lins et al. (2017) find that during a financial crisis, firms with high social capital, as captured by corporate social responsibility intensity, have higher stock returns and raise more debt than firms with low social capital. Community-level social capital has also been found to promote corporate financing. For example, in the United States, firms located in counties with high social capital have lower bank loan spreads (Hasan et al., 2017). Dudley (2021) shows that firms with entrepreneurs from counties with high social capital have more access to external financing. Moreover, Álvarez-Botas and González (2021) document that trust can help firms reduce the cost of bank loans when the country's legal enforcement and economic development are weak.

Complementing formal financial markets (e.g., banks and stock markets), informal financial markets also play a key role in nurturing millions of small businesses (Allen et al., 2019). These small businesses often face severe financing constraints from formal financial markets, especially in emerging markets with underdeveloped financial systems (Yuan et al., 2021). Loans granted by informal financial institutions are usually small, short term, and unsecured (Ayyagari et al., 2010). In many cases, informal lenders cannot enforce loan repayments through legal mechanisms and must rely on trust, reputation, or even violence to collect payments (Allen et al., 2019). Therefore, informal financial markets are considered to have a high degree of information opacity. Given its ability to facilitate information flow and foster trust (Granovetter, 2005), social capital as an informal institution should be valued more in informal financial markets.

Nevertheless, studies investigating the role of social capital in informal financial markets are scarce. The main reason for this may be that transactions in these markets are less transparent. Some exceptions are studies that use data from peer-to-peer lending platforms, a new form of informal online financial markets. For example, Lin et al. (2013) measure a borrower's social capital by the number of online friendships the borrower builds on the platform and find that such friendships are associated with higher funding success rates, lower interest rates, and lower default rates. Using data from the Chinese market, Hasan et al. (2022) construct provincial social capital index and find that borrowers located in provinces with high social capital have higher funding success rates and lower default rates. However, evidence from P2P platforms does not apply well to offline financial institutions. Offline financial institutions do not have platforms on which borrowers can build connections. In addition, offline informal financial institutions, such as moneylenders or credit cooperatives, often deal with local borrowers who have the same level of regional social capital. Therefore, the impact of individual social capital in informal financial markets is still not fully understood. Our study fills this gap by focusing on a specific type of individual social capital, namely, family social capital, and its role in offline informal financial markets.

2.2. Hypothesis development

To provide complementary evidence on the role of individual-level social capital in informal financial markets, our study directly examines the relationship between borrowers' social capital and their loan performance in informal financial markets. Specifically, we measure borrowers' social capital by family size. The family consists of a social network that can be utilized to pursue collective goals (Coleman, 1988). Social interactions among family members create interdependence and expectations based on the daily performance of tasks and responsibilities (Sanders and Nee, 1996). Sanders and Nee (1996) demonstrate that families constitute an important form

of social capital. Thus, it is plausible that families are the closest and most valuable capital for groups with extreme financing constraints, such as poor rural households.

Since loan performance is affected by both repayment ability and willingness, we propose two mechanisms through which social capital can affect a borrower's loan performance. On one hand, family social capital can improve borrowers' loan performance by increasing their repayment willingness. Coleman (1988) points out that when insurance devices are absent, family ties can provide insurance to facilitate transactions. The theoretical work of Ambrus et al. (2014) also states that pre-existing relationships such as family ties and friendships can serve as social collateral to enforce individuals to fulfill their obligations. The collateral value depends on the number of relationships valued by the borrower. From this perspective, a borrower with high social capital (a large family size) is regarded as pledging more valuable collateral. When a default occurs, the penalty for losing reputation or the social stigma borne by the whole family should be more severe for borrowers with larger families. Therefore, a higher level of family social capital can increase opportunistic costs and reinforce contract enforcement, thereby increasing borrowers' willingness to repay.

On the other hand, family social capital can increase borrowers' repayment ability by providing them with more financing sources. Lee and Persson (2016) demonstrate that the most common form of informal finance is money from family and friends. Although borrowing from family and friends has been documented to have high shadow costs (Lee and Persson, 2016), loans from family and friends often have below-market interest rates. In addition, larger families can provide individuals with more connections to the outside as well as access to alternative financing sources. Therefore, borrowers with high family social capital are expected to have more financing channels to ensure liquidity, and thus, have a higher repayment ability.

In light of the above discussion, we formulate the hypothesis as follows:

H1. Borrowers with high family social capital are associated with better loan performance.

If borrowers' family social capital can improve their loan performance in informal financial markets, a reasonable extension of this hypothesis is that the role of family social capital should be more important when the degree of information asymmetry is more severe between borrowers and lenders and when borrowers have greater demand for credit. Formal financial markets, such as commercial banks, usually prefer to serve more transparent customers with physical collateral and standard financial statements (Mertzanis, 2019; Kong et al., 2020). As a result, borrowers with less access to formal financial markets are perceived to be more information-opaque and more dependent on informal financial markets. Moreover, if family social capital can improve loan performance by increasing borrowers' repayment willingness, such an effect should be more pronounced when borrowers are in financial distress. Thus, we hypothesize as follows:

H2a. The relationship between family social capital and loan performance is more pronounced for borrowers with less access to formal financial markets.

H2b. The relationship between family social capital and loan performance is more pronounced for borrowers in financial distress.

In addition, we investigate potential factors that can mitigate the influence of family social capital on loan performance. First, there is a strand of literature documenting that high social capital is associated with a high degree of risk-taking. For example, Bloch et al. (2008) demonstrate that social connections provide an informal channel that can change individuals' risk tolerance. More social connections can magnify an individual's sense of power (Rowley, 1997), thereby allowing them to ignore potential risks when making decisions (Anderson and Galinsky, 2006). Ferris et al. (2017b) also find that CEOs' social connections are positively associated with corporate risk-taking. According to these findings, it is reasonable to conjecture that borrowers with high family social capital are more likely to invest in risky projects with funds granted, and such risk-taking tendencies increase borrowers' default probabilities.

Moreover, we conjecture that the effect of family social capital can be weakened by community-level social trust. The trust levels of different regions in China are highly heterogeneous (Ang et al., 2015). An area with high social trust emphasizes the value of honesty, thus people born there are more likely to behave by social norms. Based on this argument, a rational conjecture is that the effect of individual-level social capital is less pronounced for borrowers from regions with high social trust. In other words, regional social trust acts as a substitute for individual social capital to improve borrowers' loan performance. We consider social trust in borrowers' birthplaces because the cultural environments there leave inherited imprints on individuals (Guiso et al., 2004). Thus, we hypothesize as follows:

H3a. The relationship between family social capital and loan performance is less pronounced for borrowers with high-risk preferences.

H3b. The relationship between family social capital and loan performance is less pronounced when borrowers come from regions with high social trust.

3. Data, variables, and descriptive statistics

3.1. Data

We collect our data from an offline informal financial institution located in Ningbo, China. The financial institution was founded in 2014 and mainly serves to provide customers with small credit loans. Their target clients are self-employed small business owners, who often have a high degree of information opacity and limited access to formal financial sources. Unlike formal financial institutions,

Table 1
Variable definitions.

Variable Name	Variable Definition
Default	Takes the value of 1 if the term loan is in default, 0 otherwise
Lossratio	The percentage of total default amount to the loan size
Totalloss	Natural logarithm of total default amount
Success	Takes the value of 1 if the loan has been granted successfully, 0 otherwise
FSC	Family social capital, measured by the number of family members of a borrower
AppSize	Natural logarithm of the size of loan applied by the borrower
LoanSize	Natural logarithm of the size of the loan granted to a borrower
Loanterm	Loan term, measured in months
InterestRate	The annual interest rate of a loan
Age	Natural logarithm of the borrower's age
Gender	Takes the value of 1 if the borrower is male, 0 otherwise
College	Takes the value of 1 if the borrower has a college degree or above, 0 otherwise
Marital	Takes the value of 1 if the borrower has ever been married, 0 otherwise
EMC	The number of emergency contacts that a borrower has provided
Creditline	Natural logarithm of the amount of credit line a borrower has received from formal banks
CardUsage	Natural logarithm of the amount of credit line that has been used at the time of applying for the loan
Aveuse	Natural logarithm of the average amount of credit line that has been used within 6 months before applying for the loan
Inquiries	The number of credit inquiries shown on a borrower's credit report within 1 month before applying for the loan

which require physical collateral and standard financial statements, loans issued by this informal financial institution are credit loans with no physical collateral pledged. Because of the high default risk, the financial institution charges customers relatively high-interest rates. When applying for loans, all customers are required to submit a series of documents related to their basic information, including application forms, identification cards, household registers, education certificates, marriage certificates, and credit reports from the People's Bank of China. The financial institution then verifies the authenticity of the information and keeps copies of the documents.

We originally obtained 11,058 observations of applications, of which 5100 were successfully funded by the financial institution. All observations we received are anonymous, that is, we cannot know the identity of any customer. Despite our key variable of interest, we also control for other factors that may influence loan performance, including borrowers' demographic characteristics, loan-specific characteristics, and credit histories. After deleting observations with missing variables, our dataset finally contains 4795 observations without controlling for credit report information and 3912 observations when controlling for it.

3.2. Variables

3.2.1. Family social capital

Our key independent variable, family social capital (*FSC*), is measured by the number of family members shown in a borrower's residence booklet. In the Chinese context, family members in the same residence booklet usually live in the same house. According to China's specific regulations, residence booklets have the legal effect of proving the relationship between family members. Although we are unable to obtain copies of these documents due to privacy protection, the financial institution helps provide us with the family-size data of each borrower. In contrast to social connections established through education, employment, and other entertainment activities, our proxy of social capital captures the closest and most valuable ties a person possesses.

3.2.2. Loan performance

We measure a borrower's loan performance using three proxies that capture different levels of delinquent behavior. The first dependent variable is *Default*, defined as whether a borrower defaults. Following Norden and Weber (2010), we define a borrower as default if the overdue period is more than 90 days. The second dependent variable is *Lossratio*, which is defined as the ratio of the amount owed by the borrower to the total amount borrowed. The third dependent variable is *Totalloss*, defined as the total amount of money owned by the borrower.

3.2.3. Control variables

In addition to the main explanatory variable, we control for a series of variables that can influence loan performance. Specifically, we control for borrowers' demographic characteristics, including age, gender, education level, and marital status. We also control for loan-specific characteristics, including loan size, maturity, and interest rates. Since authentic financial statements are absent, it is difficult to identify borrowers' real income status; therefore, we try to capture borrowers' financial condition using information from their credit reports. Specifically, we control for the total credit line that a borrower has been granted from formal banks and the monthly usage of the credit line. Following Lin et al. (2013), we control for the number of inquiries in credit reports within a month before the loan application date, since more inquiries signal a higher frequency of borrowing. We further control the number of emergency contacts provided by borrowers on the application forms as a proxy for willingness to pay. The detailed definitions of the variables are provided in Table 1.

Table 2
Descriptive statistics.

Variable	N	Mean	Std Dev	Min	Max
Panel A			Descriptive Statistics of Full Observations		
Success	11,058	0.461	0.499	0	1
FSC	10,692	1.339	0.834	0	7
AppSize	11,058	11.97	1.127	1.253	16.12
Age	11,058	3.719	0.235	2.996	4.745
Gender	11,058	0.809	0.393	0	1
College	11,058	0.0945	0.293	0	1
Marital	11,058	0.888	0.315	0	1
EMC	11,058	2.078	0.855	0	6
Creditline	8910	10.05	3.209	0	16.220
CardUsage	8910	0.376	3.658	0	14.680
Aveuse	8910	10.080	3.534	0	15.860
Inquiries	8910	1.395	2.075	0	58
Panel B			Descriptive Statistics of Funded Observations		
Default	5100	0.179	0.383	0	1
Lossratio	5100	0.083	0.228	0	1.600
Totalloss	5100	1.877	3.849	0	14.866
FSC	4795	1.518	0.879	0	5
LoanSize	5100	11.21	1.090	4.615	14.910
Loanterm	5100	9.399	3.767	1	36
InterestRate	5100	20.55	5.190	0	72
Age	5100	3.720	0.228	2.996	4.205
Gender	5100	0.811	0.392	0	1
College	5100	0.110	0.313	0	1
Marital	5100	0.909	0.287	0	1
EMC	5100	2.170	0.955	0	6
Creditline	3912	11.220	2.922	0	16.220
CardUsage	3912	10.430	3.402	0	14.340
Aveuse	3912	10.480	3.266	0	15.200
Inquiries	3912	1.181	1.836	0	37

The table reports the descriptive statistics of the main variables. Panel A presents the descriptive statistics of full sample, Panel B presents the descriptive statistics of funded sample. The definitions of variables are shown in Table 1. The descriptive statistics are the mean, standard deviation, minimum, and maximum of the variables.

3.3. Descriptive statistics

The descriptive statistics are reported in Table 2. Panel A reports the descriptive statistics of all loan applications that we obtained. Panel B reports the descriptive statistics of the observations that were successfully funded. Panel A shows that, on average, 46.1% of loan applications were successfully funded. The average amount of loans requested is approximately 298,170 Chinese yuan. In terms of demographic characteristics, it shows that the financial institution's borrowers are mostly middle-aged, and only about 10% have a college degree. These patterns differ from those of borrowers on online lending platforms. According to Chen et al. (2020), borrowers on online lending platforms are relatively young and well-educated. These differences can be understood intuitively because young and well-educated individuals tend to adapt more quickly to new technologies such as Fintech platforms, while middle-aged individuals with less education rely more on traditional financial markets. Of the total applicants, 88.8% were married. Panel A also shows that males account for more than 80% of all applicants. The dominant role of male borrowers in informal financial markets is consistent with previous studies (Duarte et al., 2012; Chen et al., 2020).

Panel B shows that the average default rate in the sample is 17.9%. The average loss ratio for the funded sample is 8.3%. The maximum loss ratio is larger than one owning to interest and pecuniary penalties. The average loss amount is 10,071 Chinese yuan. The average loan size granted is 140,915 Chinese yuan, and the average maturity is above nine months. The mean value of the annual interest rate is 20.55%, which is higher than the market level. The maximum annual interest rate is unusually high at 72%, underscoring the lack of effective regulations in informal financial markets. Our main independent variable, *FSC*, takes the average values of 1.339 and 1.518 in the full sample and the funded sample, respectively, and ranges from 0 to 7 in the full sample and 0 to 5 in the funded sample.

4. Methodology and empirical results

4.1. Baseline results

In this section, we empirically test the relationship between borrowers' family social capital and their loan performance. Our main specification is as follows:

Table 3
Family social capital and loan performance.

Dependent Variable	Default	Lossratio	Totalloss	Success
Variable	(1)	(2)	(3)	(4)
FSC	−0.135**	−0.017***	−0.348***	0.507***
	(0.055)	(0.004)	(0.067)	(0.031)
Gender	0.205*	0.016*	0.280*	0.058
	(0.116)	(0.009)	(0.148)	(0.060)
College	−0.060	0.015	0.264	0.166**
	(0.148)	(0.012)	(0.193)	(0.084)
Age	0.188	0.043**	0.521*	−0.086
	(0.218)	(0.019)	(0.300)	(0.122)
Marital	0.099	−0.011	−0.060	0.074
	(0.169)	(0.018)	(0.258)	(0.095)
LoanSize	−0.157***	−0.018***	−0.189**	
	(0.059)	(0.005)	(0.075)	
Loanterm	0.112***	0.001	0.054**	
	(0.016)	(0.001)	(0.021)	
InterestRate	0.077***	0.003***	0.036*	
	(0.015)	(0.001)	(0.019)	
EMC	−0.208***	−0.020***	−0.386***	0.227***
	(0.053)	(0.003)	(0.064)	(0.030)
Creditline	−0.039	−0.007**	−0.178***	0.111***
	(0.032)	(0.003)	(0.047)	(0.019)
CardUsage	−0.105**	−0.008*	−0.094	−0.009
	(0.044)	(0.004)	(0.072)	(0.025)
Aveuse	0.126**	0.010**	0.136*	0.013
	(0.053)	(0.005)	(0.082)	(0.028)
Inquiries	0.050**	0.007***	0.105***	−0.063***
	(0.020)	(0.002)	(0.035)	(0.012)
AppSize				0.046*
				(0.025)
_cons	−2.042*	0.202**	3.379**	−3.353***
	(1.101)	(0.085)	(1.437)	(0.466)
Year FEs	Yes	Yes	Yes	Yes
Observations	3912	3912	3912	8668
Pseudo/Adj. R^2	0.073	0.072	0.065	0.103

This table reports the regression results on loan performance and funding success. The sample covers the period from 2014 to 2018. Columns (1) and (4) are estimated by Logit regression. Columns (2) and (3) are estimated by OLS regression. The dependent variables are *Default*, taking a value of 1 if the loan has been defaulted and 0 otherwise; *Lossratio*, the ratio of the default amount to the loan size; *Totalloss*, the natural logarithm of the total default amount; *Success*, taking a value of 1 if the loan is successfully funded and 0 otherwise. The detailed definitions of variables are shown in Table 1. Robust standard errors are used and reported in parentheses. Pseudo R-squares are reported in columns (1) and (4). Adjusted R-squares are reported in columns (2) and (3). *, ** and *** refer to statistical significance levels at the 10%, 5%, and 1%, respectively.

$$Loan\ Performance_i = \alpha + \beta\ FSC_i + \beta' Controls_i + \text{Year Dummies} + \varepsilon_i \quad (1)$$

where the dependent variable (*Loan Performance*) is either *Default*, *Lossratio*, or *Totalloss*, capturing different levels of delinquent behavior. When *Default* is used as the dependent variable, a logit regression model is adopted, because the dependent variable is binary. The OLS estimation method is used when *Lossratio* and *Totalloss* are used as dependent variables. The independent variable (*FSC*) is defined as a borrower's family size. *Controls* is a set of control variables defined in Table 1, including demographic characteristics, loan-specific characteristics, and credit history information. Year fixed effects are included in all regressions to control for the influence of time trends.$\varepsilon$is the error term. In heterogeneity tests, we add additional factors and their corresponding interaction terms with the key explanatory variable to explore whether the relationship is moderated by other factors. Robust standard errors are used in all specifications, and are reported in parentheses.

Table 3 reports the baseline regression results of the relationship between family social capital and loan performance. Column (1) presents the Logit estimation results with *Default* as the dependent variable. Columns (2) and (3) report the OLS estimation results with *Lossratio* and *Totalloss* as the dependent variables. The results show that borrowers' family social capital significantly improves their loan performance, as the coefficients on *FSC* remain negative and statistically significant at least at the 5% level in all regressions. The marginal effect of *FSC* in column (1) of Table 3 is −0.0188, indicating that at the mean, an increase of one unit in social capital reduces the probability of default by 10.5% (0.0188/0.179).[2] In terms of the other two dependent variables, when social capital increases by one unit, the loss ratio (*Lossratio*) decreases by 1.7% and the total loss amount (*Totalloss*) decreases by 29.39% (equivalent to about 2960 RMB). Overall, the results support our first hypothesis.

[2] The marginal effects of variables are not reported and are available upon request.

Table 4
Effect of access to external finance on the relationship between family social capital and loan performance.

Dependent Variable	Default	Lossratio	Totalloss
Variable	(1)	(2)	(3)
Panel A: Effect of bank relationship			
FSC	−0.284***	−0.030***	−0.686***
	(0.102)	(0.008)	(0.133)
Bankrelationship	−0.066	−0.008**	−0.117*
	(0.048)	(0.004)	(0.066)
*FSC*Bankrelationship*	0.037	0.004*	0.092***
	(0.027)	(0.002)	(0.034)
Controls	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes
Observations	3521	3521	3521
Pseudo/Adj. R^2	0.075	0.073	0.068
Panel B: Effect of credit line			
FSC	−0.160***	−0.022***	−0.426***
	(0.058)	(0.004)	(0.071)
Creditline	−0.047	−0.009***	−0.204***
	(0.033)	(0.003)	(0.049)
*FSC*Creditline*	0.104*	0.021***	0.307***
	(0.057)	(0.008)	(0.092)
Controls	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes
Observations	3912	3912	3912
Pseudo/Adj. R^2	0.074	0.075	0.068

This table reports the regression results on loan performance. The sample covers the period from 2014 to 2018. Panel A report results of regressions which include interactions with *Bankrelationship*. Panel B report results of regressions which include interactions with *Creditline*. Column (1) is estimated by Logit regression. Columns (2) and (3) are estimated by OLS regression. The dependent variables are *Default*, taking a value of 1 if the loan has been defaulted and 0 otherwise; *Lossratio*, the ratio of the default amount to the loan size; *Totalloss*, the natural logarithm of the total default amount. The detailed definitions of variables are shown in Table 1. Robust standard errors are used and reported in parentheses. Pseudo R-square is reported in column (1). Adjusted R-squares are reported in columns (2) and (3). *, ** and *** refer to statistical significance levels at the 10%, 5%, and 1%, respectively.

Coefficients on control variables suggest that male borrowers, older borrowers and loans with longer terms are associated with worse repayment performance, consistent with previous evidence in other informal financial markets (e.g., Chen et al., 2020). In addition, we find that borrowers who provide more emergency contacts have better loan performance. We also find strong evidence that borrowers who require their credit reports frequently have poorer loan performance, which is consistent with our expectation.

We further test whether family social capital is associated with funding success rates. Column (4) of Table 3 reports the logit regression results. The dependent variable *Success* is a binary variable that takes the value of one if the borrower is funded by the informal financial institution, and zero otherwise. The results show that family social capital is positively and significantly related to funding success rates. However, we should be aware that the lending decisions in our setting are made within the financial institution with limited loan officers, where the investors are not as diverse as those in online lending platforms. Therefore, we should be more cautious when making inferences about the external validity of the results. We also rerun the regressions without controlling for credit history information to keep larger sample size. The estimated results, which are consistent with the baseline findings, are not reported for the sake of brevity and are available upon request.

4.2. Factors influencing the relation between family social capital and loan performance

In this section, we conduct a series of cross-sectional tests to provide further evidence that supports our findings. Specifically, we examine the effects of individual-level (borrowers' access to formal financial markets, financial conditions and risk preferences) and community-level factors (regional social trust) on the relationship between family social capital and loan performance.

4.2.1. Effects of borrowers' access to formal financial markets

First, we investigate whether the relationship established in our baseline regressions varies across borrowers with different levels of access to formal financial markets. We argue that borrowers with greater access to formal financial markets have greater information transparency and rely less on “soft collateral” to signal their creditworthiness. We use two proxies to capture the degree of access to formal financial markets. The first proxy is *Bankrelationship*, which is defined as the number of banks that have issued loans to a borrower. The second proxy is *Creditline*, which is defined as the amount of credit line that a borrower has received from formal banks. Panel A of Table 4 reports the results using the first proxy, *Bankrelationship*, and Panel B of Table 4 reports the results using the second proxy, *Creditline*. The estimated coefficients on the first interaction term are positive and significant at least at the 10% level when *Lossratio* and *Totalloss* are used as the dependent variables. The estimated coefficients on the second interaction term are positive and

Table 5
Effect of financial conditions on the relationship between family social capital and loan performance.

Dependent Variable	Default	Lossratio	Totalloss
Variable	(1)	(2)	(3)
FSC	−0.129**	−0.014***	−0.328***
	(0.057)	(0.004)	(0.070)
FC	0.228	0.071**	0.742
	(0.300)	(0.032)	(0.463)
*FSC*FC*	−0.062	−0.030**	−0.211
	(0.186)	(0.014)	(0.227)
Controls	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes
Observations	3912	3912	3912
Pseudo/Adj. R^2	0.073	0.074	0.066

This table reports the regression results on loan performance. The sample covers the period from 2014 to 2018. Column (1) is estimated by Logit regression. Columns (2) and (3) are estimated by OLS regression. The dependent variables are *Default*, taking a value of 1 if the loan has been defaulted and 0 otherwise; *Lossratio*, the ratio of the default amount to the loan size; *Totalloss*, the natural logarithm of the total default amount. The detailed definitions of variables are shown in Table 1. Robust standard errors are used and reported in parentheses. Pseudo R-square is reported in column (1). Adjusted R-squares are reported in columns (2) and (3). *, ** and *** refer to statistical significance levels at the 10%, 5%, and 1%, respectively.

Table 6
Effect of risk preferences on the relationship between family social capital and loan performance.

Dependent Variable	Default	Lossratio	Totalloss
Variable	(1)	(2)	(3)
FSC	−0.160**	−0.022***	−0.453***
	(0.063)	(0.005)	(0.078)
Guarantee	−0.320	−0.056***	−1.129***
	(0.233)	(0.016)	(0.285)
*FSC*Guarantee*	0.132	0.025***	0.472***
	(0.121)	(0.008)	(0.142)
Controls	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes
Observations	3912	3912	3912
Pseudo/Adj. R^2	0.073	0.074	0.068

This table reports the regression results on loan performance. The sample covers the period from 2014 to 2018. Column (1) is estimated by Logit regression. Columns (2) and (3) are estimated by OLS regression. The dependent variables are *Default*, taking a value of 1 if the loan has been defaulted and 0 otherwise; *Lossratio*, the ratio of the default amount to the loan size; *Totalloss*, the natural logarithm of the total default amount. The detailed definitions of variables are shown in Table 1. Robust standard errors are used and reported in parentheses. Pseudo R-square is reported in column (1). Adjusted R-squares are reported in columns (2) and (3). *, ** and *** refer to statistical significance levels at the 10%, 5%, and 1%, respectively.

significant at least at the 10% level across all three columns. These results are consistent with our expectation that family social capital plays a more important role for borrowers who have less access to informal financial markets.

4.2.2. *Effects of borrowers' financial conditions*

We next examine the influence of borrowers' financial conditions on the relationship between family social capital and loan performance. Following Garcia-Appendini and Montoriol-Garriga (2013), who consider a firm's unused credit line to measure liquidity, we measure borrowers' financial conditions by how much they use their credit line. Specifically, we calculate the ratio of the average use of credit line within six months before loan applications to the total amount of credit line granted. We define a dummy variable *FC* that equals 1 if the ratio is above the value at the 90 percentile and 0 otherwise. We add the interaction term between *FC* and *FSC* to our baseline specifications and report the results in Table 5. The estimated coefficient on the interaction term is negatively significant at the 5% level when *Lossratio* is used as the dependent variable. As for economic significance, compared to borrowers in good financial condition, borrowers in poor financial condition experience an additional 3% decrease in the loss ratio with a one-unit increase in social capital. The coefficients on the interaction term in columns (1) and (3) are not significant, but all with negative signs as expected. In general, the results provide weak but suggestive evidence and support our hypothesis to some extent.

4.2.3. *Effects of borrowers' risk preferences*

We next investigate whether the effects of family social capital differ among borrowers with different risk preferences. Previous studies document that social capital can enhance the sense of power and prompt risk-taking behaviors (Rowley, 1997; Ferris et al., 2017b). We propose that borrowers with high social capital are more likely to invest in high-risk projects, thereby increasing their default risks. We measure borrowers' risk preferences based on their guaranteed behavior. The underlying rationale is that individuals

Table 7
Effect of regional social trust on the relationship between family social capital and loan performance.

Dependent Variable	Default	Lossratio	Totalloss
Variable	(1)	(2)	(3)
FSC	−0.097	−0.034***	−0.655***
	(0.107)	(0.008)	(0.134)
Trust	0.158	−0.041**	−0.810***
	(0.220)	(0.019)	(0.314)
*FSC*Trust*	−0.049	0.023***	0.419***
	(0.122)	(0.009)	(0.153)
Controls	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes
Observations	3912	3912	3912
Pseudo/Adj. R^2	0.073	0.073	0.067

This table reports the regression results on loan performance. The sample covers the period from 2014 to 2018. Column (1) is estimated by Logit regression. Columns (2) and (3) are estimated by OLS regression. The dependent variables are *Default*, taking a value of 1 if the loan has been defaulted and 0 otherwise; *Lossratio*, the ratio of the default amount to the loan size; *Totalloss*, the natural logarithm of the total default amount. The detailed definitions of variables are shown in Table 1. Robust standard errors are used and reported in parentheses. Pseudo R-square is reported in column (1). Adjusted R-squares are reported in columns (2) and (3). *, ** and *** refer to statistical significance levels at the 10%, 5%, and 1%, respectively.

who provide a guarantee to others indicate that they are willing to undertake additional risks beyond their control. The guarantee information is obtained from borrowers' credit reports. We define a dummy variable *Guarantee* that equals one if a borrower has ever guaranteed someone else and zero otherwise, and add the variable and its interaction with *FSC* to the baseline regressions. Table 6 presents the results. The estimated coefficients on the interaction terms are positive and significant at the 1% level in columns (2) and (3), supporting our argument that the relationship between family social capital and loan performance is less pronounced in borrowers with higher risk preferences.

4.2.4. Effects of regional social trust level

We next examine whether community-level informal institutions influence the effects of family social capital. Specifically, we focus on regional social trust and investigate its moderating effects on the relationship between family social capital and loan performance. Borrowers born in high-trust regions are nurtured by justice and honesty; thus, they are more likely to behave consistently with social norms. Therefore, we examine whether regional social trust and family social capital play substitute roles in informal financial markets. Our proxy for social trust is taken from a survey conducted by the "Chinese Enterprise Survey System" in 2001, which has been widely used in empirical studies (e.g., Wu et al., 2014). We define the variable *Trust* which equals one if the social trust level of a borrower's birthplace is above the national average and zero otherwise. The results are presented in Table 7. The estimated coefficients on the interaction term are positive and significant at the 1% level in columns (2) and (3), indicating that family social capital plays a more important role in regions with lower social trust.

5. Additional tests

5.1. Addressing endogeneity concerns

The relationships established in the baseline regressions indicate that higher family social capital is associated with better loan performance. Our key independent variable, family social capital, is measured by pre-existing relationships, while our dependent variable, loan performance, can only be observed ex post. Thus, the relationship is less affected by the reverse causality. However, our results may still suffer from other types of endogeneity problems, such as omitted variables and sample selection bias. In this section, we adopt three methods to examine whether the correlations documented to this point are causal. First, we use selection on observable variables to estimate the likelihood that our results are driven by unobservable variables. Second, we use instrumental variable strategy to alleviate omitted variable bias. Finally, we conduct Heckman two-step estimations to deal with sample selection bias.

5.1.1. Using observable variables to estimate the bias from unobservable variables

In this section, we address the endogeneity problems by assessing the possibility of the results being biased by unobserved factors. The strategy of selecting observable variables to estimate the degree of bias from unobservable variables was first developed by Altonji et al. (2005), and has been successfully exploited in empirical studies (e.g., Nunn and Wantchekon, 2011). Specifically, the strategy measures that concerning the selection of observable variables, how much stronger the selection on unobservable variables should be to fully explain away the estimated results. The method uses two regressions, one containing a full vector of control variables and the other containing a restricted set of control variables. We use $\widehat{\beta}^F$ (where F stands for full) to denote the estimated coefficient for the key explanatory variable from the full regression, and $\widehat{\beta}^R$ (where R stands for restricted) to denote the estimated coefficient for the key explanatory variable from the restricted regression. The selection ratio is calculated as $\widehat{\beta}^F$ divided by the difference between $\widehat{\beta}^R$ and $\widehat{\beta}^F$. The intuition underlying the selection ratio is illustrated as: the smaller the denominator (difference between $\widehat{\beta}^R$ and $\widehat{\beta}^F$), the less the

Table 8
Using selection on observable variables to assess the bias from unobservable variables.

		Default	Lossratio	Totalloss
Controls in the restricted set	Controls in the full set	(1)	(2)	(3)
None	Full set of controls	1.91	1.59	2.00
Demographic characteristics	Full set of controls	1.84	1.76	2.11

Each cell of the table presents selection ratios based on the coefficients on *FSC* from two regressions. The sample covers the period from 2014 to 2018. The dependent variables are *Default*, taking a value of 1 if the loan has been defaulted and 0 otherwise; *Lossratio*, the ratio of the default amount to the loan size; *Totalloss*, the natural logarithm of the total default amount. The control variables included in the first restricted set are only time fixed effects. The control variables included in the second restricted set are demographic characteristics with time fixed effects. Robust standard errors are used. The coefficient for the restricted regression is $\widehat{\beta}^R$. The coefficient for the full regression is $\widehat{\beta}^F$. The selection ratio is calculated as $\widehat{\beta}^F/(\widehat{\beta}^R\text{-}\widehat{\beta}^F)$.

Table 9
Instrumental variable approach.

	First Stage	Second Stage		
Dependent Variable	FSC	Default	Lossratio	Totalloss
Variable	(1)	(2)	(3)	(4)
Child	0.464***			
	(0.035)			
FSC		−0.093	−0.037**	−1.011***
		(0.126)	(0.015)	(0.276)
Constant	−0.280	−1.344**	0.194**	3.110**
	(0.280)	(0.581)	(0.085)	(1.448)
Controls	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes
Observations	3912	3912	3912	3912
F-statistic	58.16			

This table reports the IV Probit regression results and 2SLS regression results on loan performance. The sample covers the period from 2014 to 2018. Column (1) reports the first-stage regression results estimated by OLS regression, Column (2) reports the second-stage regression results estimated by IV Probit regression, and Columns (3) and (4) report the second-stage regression results estimated by 2SLS regression. The dependent variables are *FSC*, the number of family numbers of a borrower; *Default*, taking a value of 1 if the loan has been defaulted and 0 otherwise; *Lossratio*, the ratio of the default amount to the loan size; *Totalloss*, the natural logarithm of the total default amount. The detailed definitions of variables are shown in Table 1. Robust standard errors are used and reported in parentheses. *, ** and *** refer to statistical significance levels at the 10%, 5%, and 1%, respectively.

result is affected by selection on observable variables, and the stronger requirement of selection on unobservable variables to fully explain away the effect. Similarly, the larger the numerator ($\widehat{\beta}^F$), the stronger the effect that should be explained away by selection on unobservable variables. A higher selection ratio indicates that the baseline relationship is less affected by unobservable factors.

Specifically, we conduct two restricted regressions with two sets of restricted control variables. The first restricted regression includes no controls but only time fixed effects. The second restricted regression includes demographic control variables (including age, gender, education and marital status) as well as time fixed effects. We calculate the selection ratios using the estimated coefficients from the two restricted regressions and the estimated coefficients from the full specifications. The calculated selection ratios are presented in Table 8. The results show that all the six selection ratios calculated are larger than one, indicating that the selection on unobservable variables should be at least one more time greater than the selection on observable variables to fully explain away the effect. The results suggest that the relationship established in baseline regressions is less likely to be fully driven by unobservable variables.

5.1.2. Instrumental variable approach

We next exploit the instrumental variable approach to further address the concerns of omitted variables bias. The key point of conducting the estimation is to find an instrumental variable (IV) that is correlated with our measure of social capital but does not directly influence loan performance except through the channels brought by social capital. We use whether a borrower has children or not as an instrument to account for family size. Specifically, we construct a dummy variable *Child* which equals one if the borrower has at least one child and zero otherwise. It is reasonable to conjecture that having children is positively related to family size, thereby satisfying the correlation assumption. In addition, little evidence has documented that having children or not influences borrowers' loan performance through channels other than social capital, satisfying the requirement of exclusion restriction.

The results are presented in Table 9. Column (1) presents the results of the first-stage regression. The estimated coefficient on the instrumental variable (*Child*) is positive and statistically significant at the 1% level, indicating that a borrower who has at least one child is more likely to have a larger family size. The minimum eigenvalue statistic is much larger than 10, thus mitigating the concern about the weak instrument. Columns (2) to (4) report the results of the second-stage regressions with three dependent variables. We

Table 10
Heckman two-step selection model.

	First step	Second step		
Dependent Variable	Success	Default	Lossratio	Totalloss
Variable	(1)	(2)	(3)	(4)
Peerrate	1.763***			
	(0.509)			
FSC	0.385***	−0.072**	−0.055***	−0.764***
	(0.017)	(0.030)	(0.018)	(0.284)
Inverse Mills ratio		−0.220*	−0.160**	−1.726
		(0.120)	(0.071)	(1.133)
Constant	1.120	0.215	0.290***	4.321***
	(1.000)	(0.163)	(0.095)	(1.552)
Borrower variables	Yes			
Controls		Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes
Observations	10,692	9809	9809	9809

This table reports the Heckman two-step regression results on loan performance. The sample covers the period from 2014 to 2018. Column (1) reports the first-step results estimated by Probit regression. Columns (2) to (4) report the second-step results of the Heckman two-step model. The dependent variable in Column (1) is *Success*, which takes the value of 1 if the loan has been successfully funded and 0 otherwise. The dependent variables in Column (2) to (4) are *Default*, taking a value of 1 if the loan has been defaulted and 0 otherwise; *Lossratio*, the ratio of the default amount to the loan size; *Totalloss*, the natural logarithm of the total default amount. The exogenous predictive instrument in the first-step is *Peerrate*, which takes the value of the average success rates of borrowers in each education group. The detailed definitions of variables are shown in Table 1. Robust standard errors are used and reported in parentheses. *, ** and *** refer to statistical significance levels at the 10%, 5%, and 1%, respectively.

conduct IV Probit regression estimation in column (2) and two-stage least squares regression estimation in columns (3) and (4). The coefficients on the predicted *FSC* are both negative and statistically significant at least at the 5% level in columns (3) and (4), and the coefficient in column (2) is insignificant but has a sign consistent with expectation, meaning that the overall loan performance is better for borrowers with higher family social capital. These results are consistent with the baseline results.

5.1.3. Heckman two-step selection model

It has been noticed that the credit scoring model suffers from sample selection bias (Jacobson and Roszbach, 2003). Taking our study setting as an example, we can only observe the repayment behavior of borrowers who have been successfully funded but cannot observe those who fail to be funded. We, therefore, exploit Heckman's (1979) two-step selection model to address the concern of sample selection bias. In the first step, we estimate a Probit model to predict the probability of being successfully funded. The specification of the Probit model is as follows:

$$Pr(SUCCESS)_i = \alpha + \beta X_i + \beta^{'} Controls_i + \text{Year Dummies} + \varepsilon_i \quad (2)$$

where the dependent variable (*SUCCESS*) is a binary variable that equals one if the loan has been granted and zero otherwise, and *X* is an exogenous variable that can be used to predict a loan application's funding outcome. We use the average funding success rate in each education group as the exogenous predictive variable because previous studies document that the loan success rate of borrowers' peers influences their funding success rates, but not their loan performance (Chen et al., 2020). We first divide our observations into five education groups: below the middle school, middle school, high school, college, and above college. We then define a variable *Peerrate* that takes the value of the average funding success rate of each education group and use it as our exogenous predictive variable.

In addition to *Peerrate*, we also control for borrowers' demographic characteristics and time fixed effects in the first stage. We cannot control for loan-specific characteristics because such factors can only be obtained when borrowers are successfully funded. Because controlling credit history information leads to more losses of observations, we also remove such information to obtain a larger predictive sample in the first stage. The Inverse Mills ratio, calculated as the ratio of the probability density function to the cumulative distribution function of distribution, can be obtained from the first-step estimation. In the second step, the Inverse Mills ratio is added to our main regression models to correct for selection bias.

Table 10 presents the results of the Heckman two-step selection model estimation. Column (1) of Table 10 presents the first-step results. Our exogenous variable *Peerrate* is positively and significantly associated with the probability of funding success. Columns (2) to (4) present the second-step results. The Inverse Mills ratio is statistically significant at least at the 10% level in columns (2) and (3), and is insignificant when the dependent variable is *Totalloss*, indicating that sample selection bias exists to some extent in our study setting. The estimated results are robust and remain consistent with our baseline results after addressing sample selection bias.

5.2. Robustness tests

We conduct several robustness tests to corroborate our results. First, we add additional dimensions of fixed effects, including loan purpose and loan officer fixed effects, to our baseline specifications. In our setting, borrowers apply for loans for different purposes, such as business expansion, decoration, and the purchase of equipment and raw materials. Some borrowers also borrow for daily

Table 11
Additional dimensions of fixed effects.

Dependent Variable	Default	Lossratio	Totalloss
Variable	(1)	(2)	(3)
FSC	−0.122*	−0.012***	−0.257***
	(0.063)	(0.004)	(0.075)
Controls	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes
Loan Purpose FEs	Yes	Yes	Yes
Loan Officer FEs	Yes	Yes	Yes
Observations	3912	3912	3912
Pseudo/Adj. R^2	0.118	0.074	0.073

Table 12
Robustness tests.

Dependent Variable	Default	Lossratio	Totalloss
Variable	(1)	(2)	(3)
Panel A: First-granted borrowers			
FSC	−0.173***	−0.020***	−0.350***
	(0.063)	(0.005)	(0.083)
Controls	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes
Observations	2816	2816	2816
Pseudo/Adj. R^2	0.074	0.103	0.082
Panel B: Non-local borrowers			
FSC	−0.121*	−0.015***	−0.347***
	(0.064)	(0.004)	(0.077)
Controls	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes
Observations	2667	2667	2667
Pseudo/Adj. R^2	0.070	0.083	0.081
Panel C: Without extreme loans			
FSC	−0.170**	−0.016***	−0.271***
	(0.069)	(0.005)	(0.087)
Controls	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes
Observations	2412	2412	2412
Pseudo/Adj. R^2	0.053	0.051	0.046

This table reports the regression results on loan performance. The sample covers the period from 2014 to 2018. Panel A reports the robustness tests with each borrower's first granted loan. Panel B reports the robustness tests with non-local borrowers. Panel C reports the robustness tests without extreme loans. Column (1) is estimated by Logit regression. Columns (2) and (3) are estimated by OLS regression. The dependent variables are *Default*, taking a value of 1 if the loan has been defaulted and 0 otherwise; *Lossratio*, the ratio of the default amount to the loan size; *Totalloss*, the natural logarithm of the total default amount. The detailed definitions of variables are shown in Table 1. Robust standard errors are used and reported in parentheses. Pseudo R-square is reported in column (1). Adjusted R-squares are reported in columns (2) and (3). *, ** and *** refer to statistical significance levels at the 10%, 5%, and 1%, respectively.

consumption. Because different borrowing purposes might reflect different financial conditions and different fund flows, we include loan purpose fixed effects to control for such information. Moreover, since each loan officer has specific lending preferences, we also add loan officer fixed effects to absorb the effects of unobservable factors in lending preferences. The estimated results are reported in Table 11, which shows that our results remain robust after controlling for the loan purpose and loan officer fixed effects.

We next provide evidence to rule out the possibility that repeated borrowers dominate our results. Individual borrowers with long-term financing requirements tend to maintain good relationships with familiar financial institutions. In our sample, 2454 individuals have borrowed once, 1293 individuals have borrowed twice, and the remaining individuals have borrowed more than twice from this financial institution. Our sample also shows that the family sizes of borrowers who apply more than once did not change during the period. We re-estimate our baseline regressions using each borrower's first-granted loan from this financial institution. The results are reported in Panel A of Table 12 and are consistent with the baseline results.

Local borrowers live and conduct their business in their birthplace, with family members and friends all close to them. Intuitively, family social capital plays a more important role for local borrowers. Therefore, we examine whether our results are driven by local borrowers. We re-estimate the baseline regressions with only non-local borrowers and report the results in Panel B of Table 12. The

estimated coefficients are consistent with our baseline results, ruling out the possibility that the results are driven by local borrowers.

We then explore whether extreme loans dominate the results. Since the financial institution in our setting is informal and lacks effective regulation, interest rates are extremely high in some circumstances. In our sample, the maximum annual interest rate is 72%. The average loan size in our sample is around 140,000 RMB. To further corroborate our results, we delete observations with loan sizes larger than 150,000 RMB and interest rates higher than 24% and re-estimate the baseline regressions. The results in Panel C of Table 12 are consistent with our baseline results.

6. Conclusions

This study examines whether and how family social capital influences borrowers' loan performance in informal financial markets. We provide empirical evidence that family social capital, captured by family size, can reduce borrowers' opportunistic behavior, thus leading to better loan performance. Specifically, we find that an increase of one unit of family social capital reduces the default probability by 10.5%, the loss ratio by 1.7%, and the total loss amount by 29.39%. Our results remain robust after controlling for a vector of factors, time fixed effects, and after addressing endogeneity concerns. We further find that the relationship between family social capital and loan performance is more pronounced for borrowers with less access to formal financial markets and borrowers with poorer financial conditions, while it is less pronounced for borrowers with higher risk preferences and those born in regions with high social trust.

This study has important implications for both practitioners and policymakers. First, credit risk management plays a crucial role in financial market development. By revealing the important role of family social capital in improving loan performance, our study helps reduce information asymmetries between borrowers and lenders, thus improving credit risk management in financial markets. Second, our study also provides implications for small business owners to use their most valuable capital to overcome financing constraints. Meanwhile, because financing constraints significantly discourage households from establishing small businesses, our findings have implications for policymakers to take advantage of family social capital, a valuable informal institution, to support small business development.

CRediT authorship contribution statement

Wanning Li: Data curation, Formal analysis, Methodology, Writing – original draft. **Xiuping Hua:** Conceptualization, Writing – review & editing, Validation.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors are grateful to the Editor-in-Chief and the anonymous referee for their constructive comments and suggestions. This work was supported by the National Office for Philosophy and Social Science of China under the National Social Science Fund programmes 2019 (Grant No. 19BJY252), Ningbo Science and Technology Bureau for S&T Innovation 2025 Major and Special Program (Project No. 2019B10038, 2021Z017 and 2022Z173), and Ningbo Science and Technology Bureau for Soft Science Program (Project No. 2022R018). All errors remain the responsibility of the authors.

References

Acheampong, A., Elshandidy, T., 2021. Does soft information determine credit risk? Text-based evidence from European banks. J. Int. Financ. Mark. Inst. Money 75, 101303.
Allen, F., Qian, M., Xie, J., 2019. Understanding informal financing. J. Financ. Intermed. 39, 19–33.
Altonji, J.G., Elder, T.E., Taber, C.R., 2005. Selection on observed and unobserved variables: assessing the effectiveness of Catholic schools. J. Polit. Econ. 113 (1), 151–184.
Álvarez-Botas, C., González, V.M., 2021. Does trust matter for the cost of bank loans? J. Corp. Finan. 66, 101791.
Ambrus, A., Mobius, M., Szeidl, A., 2014. Consumption risk-sharing in social networks. Am. Econ. Rev. 104 (1), 149–182.
Anderson, C., Galinsky, A.D., 2006. Power, optimism, and risk-taking. Eur. J. Soc. Psychol. 36 (4), 511–536.
Ang, J.S., Cheng, Y., Wu, C., 2015. Trust, investment, and business contracting. J. Financ. Quant. Anal. 50 (3), 569–595.
Ayyagari, M., Demirgüç-Kunt, A., Maksimovic, V., 2010. Formal versus informal finance: evidence from China. Rev. Financ. Stud. 23 (8), 3048–3097.
Beck, T., Behr, P., Madestam, A., 2018. Sex and credit: do gender interactions matter for credit market outcomes? J. Bank. Financ. 87, 380–396.
Bedendo, M., Garcia, A.E., Siming, L., 2020. Cultural preferences and firm financing choices. J. Financ. Quant. Anal. 55 (3), 897–930.
Bloch, F., Genicot, G., Ray, D., 2008. Informal insurance in social networks. J. Econ. Theory 143 (1), 36–58.
Bourdieu, P., 1986. The forms of capital. In: Richardson, J.G. (Ed.), Handbook of Theory and Research for the Sociology of Education. Greenwood Press, New York, pp. 241–258.
Chen, X., Huang, B., Ye, D., 2020. Gender gap in peer-to-peer lending: evidence from China. J. Bank. Financ. 112, 105633.
Coleman, J.S., 1988. Social capital in the creation of human capital. Am. J. Sociol. 94, 95–120.
Dorfleitner, G., Priberny, C., Schuster, S., Stoiber, J., Weber, M., Castro, I.D., Kammler, J., 2016. Description-text related soft information in peer-to-peer lending-evidence from two leading European platforms. J. Bank. Financ. 64, 169–187.
Duarte, J., Siegel, S., Young, L., 2012. Trust and credit: the role of appearance in peer-to-peer lending. Rev. Financ. Stud. 25 (8), 2455–2484.
Dudley, E., 2021. Social capital and entrepreneurial financing choice. J. Corp. Finan. 70, 102068.

El Ghoul, S., Zheng, X., 2016. Trade credit provision and national culture. J. Corp. Finan. 41, 475–501.
Engelberg, J., Gao, P., Parsons, C.A., 2012. Friends with money. J. Financ. Econ. 103 (1), 169–188.
Ferris, S.P., Javakhadze, D., Rajkovic, T., 2017a. The international effect of managerial social capital on the cost of equity. J. Bank. Financ. 74, 69–84.
Ferris, S.P., Javakhadze, D., Rajkovic, T., 2017b. CEO social capital, risk-taking and corporate policies. J. Corp. Finan. 47, 46–71.
Fisman, R., Paravisini, D., Vig, V., 2017. Cultural proximity and loan outcomes. Am. Econ. Rev. 107 (2), 457–492.
Fogel, K., Jandik, T., McCumber, W.R., 2018. CFO social capital and private debt. J. Corp. Finan. 52, 28–52.
Folland, S., 2006. Value of life and behavior toward health risks: an interpretation of social capital. Health Econ. 15 (2), 159–171.
Garcia-Appendini, E., Montoriol-Garriga, J., 2013. Firms as liquidity providers: evidence from the 2007–2008 financial crisis. J. Financ. Econ. 109 (1), 272–291.
Garmaise, M.J., 2015. Borrower misreporting and loan performance. J. Financ. 70 (1), 449–484.
Glaeser, E.L., Laibson, D., Sacerdote, B., 2002. The economic approach to social capital. Econ. J. 112 (483), 437–458.
Granovetter, M., 2005. The impact of social structure on economic outcomes. J. Econ. Perspect. 19 (1), 33–50.
Guiso, L., Sapienza, P., Zingales, L., 2004. The role of social capital in financial development. Am. Econ. Rev. 94 (3), 526–556.
Hasan, I., Hoi, C.K., Wu, Q., Zhang, H., 2017. Social capital and debt contracting: evidence from bank loans and public bonds. J. Financ. Quant. Anal. 52 (3), 1017–1047.
Hasan, I., He, Q., Lu, H., 2022. Social capital, trusting, and trustworthiness: evidence from peer-to-peer lending. J. Financ. Quant. Anal. 57 (4), 1409–1453.
Heckman, J., 1979. Sample selection bias as a specification error. Econometrica 47 (1), 153–161.
Hua, X., Chen, Y., Prashantham, S., 2016. Institutional logic dynamics: private firm financing in Ningbo (1912-2008). Bus. Hist. 58 (3), 378–407.
Huang, K., Shang, C., 2019. Leverage, debt maturity, and social capital. J. Corp. Finan. 54, 26–46.
Jacobson, T., Roszbach, K., 2003. Bank lending policy, credit scoring and value-at-risk. J. Bank. Financ. 27 (4), 615–633.
Karaivanov, A., Kessler, A., 2018. (Dis)advantages of informal loans–theory and evidence. Eur. Econ. Rev. 102, 100–128.
Karlan, D.S., 2005. Using experimental economics to measure social capital and predict financial decisions. Am. Econ. Rev. 95 (5), 1688–1699.
Kong, D., Pan, Y., Tian, G.G., Zhang, P., 2020. CEO's hometown connections and access to trade credit: evidence from China. J. Corp. Finan. 62, 101574.
Lee, S., Persson, P., 2016. Financing from family and friends. Rev. Financ. Stud. 29 (9), 2341–2386.
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. Manag. Sci. 59 (1), 17–35.
Lins, K.V., Servaes, H., Tamayo, A., 2017. Social capital, trust, and firm performance: the value of corporate social responsibility during the financial crisis. J. Financ. 72 (4), 1785–1823.
Liu, Q., Luo, J., Tian, G.G., 2016. Managerial professional connections versus political connections: evidence from firms' access to informal financing resources. J. Corp. Finan. 41, 179–200.
Ma, S., Wu, X., Gan, L., 2019. Credit accessibility, institutional deficiency and entrepreneurship in China. China Econ. Rev. 54, 160–175.
Mertzanis, C., 2019. Family ties, institutions and financing constraints in developing countries. J. Bank. Financ. 108, 105650.
Norden, L., Weber, M., 2010. Credit line usage, checking account activity, and default risk of bank borrowers. Rev. Financ. Stud. 23 (10), 3665–3699.
Nunn, N., Wantchekon, L., 2011. The slave trade and the origins of mistrust in Africa. Am. Econ. Rev. 101 (7), 3221–3252.
Petersen, M.A., Rajan, R.G., 2002. Does distance still matter? The information revolution in small business lending. J. Financ. 57 (6), 2533–2570.
Putnam, R.D., 1993. Making Democracy Work: Civic Traditions in Modern Italy. Princeton University Press, Princeton, New Jersey.
Rowley, T.J., 1997. Moving beyond dyadic ties: a network theory of stakeholder influences. Acad. Manag. Rev. 22 (4), 887–910.
Rupasingha, A., Goetz, S.J., Freshwater, D., 2006. The production of social capital in US counties. J. Soc.-Econ. 35 (1), 83–101.
Sanders, J.M., Nee, V., 1996. Immigrant self-employment: the family as social capital and the value of human capital. Am. Sociol. Rev. 61 (2), 231–249.
Shahriar, A.Z.M., Unda, L.A., Alam, Q., 2020. Gender differences in the repayment of microcredit: the mediating role of trustworthiness. J. Bank. Financ. 110, 105685.
Wu, W., Firth, M., Rui, O.M., 2014. Trust and the provision of trade credit. J. Bank. Financ. 39, 146–159.
Yin, H., Jin, X., Quan, X., Yu, J., 2022. Does social network improve corporate financing efficiency? Evidence from China. Pac. Basin Financ. J. 74, 101802.
Yuan, Y., Rong, Z., Xu, N., Lu, Y., 2021. Credit cards and small business dynamics: evidence from China. Pac. Basin Financ. J. 67, 101570.