

# 软信息能否缓解 P2P 网贷市场中的羊群行为?

都红雯,周文婷,胡文彬

(杭州电子科技大学 经济学院,浙江 杭州 310018)

**摘要:** 文章以“人人贷”网络借贷平台2011年8月15日至2014年8月20日交易数据为样本,通过 Python 编写网络爬虫程序抓取14936个有效借款订单,采用线性回归模型,按投标记录进行等时间段划分,分析借款人提供的借款陈述等软信息能否缓解 P2P 网贷市场投资者的羊群行为。实证表明,借款陈述的文本长度与羊群行为存在负相关关系,即借款描述等软信息的增加可适度缓解投资者的羊群行为。对此,建议国内网贷平台应借鉴国外 P2P 网贷平台的运行经验,提供更多的软信息以引导投资者合理决策。

**关键词:** P2P 网贷;软信息;借款陈述;羊群行为

**中图分类号:**F830 **文献标志码:**A **文章编号:**1000-2154(2018)01-0087-10

**DOI:**10.14134/j.cnki.cn33-1336/f.2018.01.008

## Could the Soft Information Alleviate the Herding Behavior in P2P Loan Market?

DU Hong-wen, ZHOU Wen-ting, HU Wen-bin

(School of Economics, Hangzhou Dianzi University, Hangzhou 310018, China)

**Abstract:** Using Python to compose a web crawler program to extract 14936 valid loans from “Renrendai” website from August 15, 2011 to August 20, 2014, this study tries to expose whether the soft information has effects on the herding behavior in P2P loan market, by using linear regression model according to the tender record for equal time division. The results show the length of the text of the loan statement is positively correlated with the herding behavior, that is, the increase in the soft information such as the description of the borrower may moderately alleviate the investor’s herding behavior. Therefore, domestic network loan platform should learn from foreign P2P network loan platforms’ operating experience, and provide more soft information to guide investors to make reasonable decisions.

**Key words:** P2P network loan; soft information; loan statement; herd behavior

## 一、问题的提出

金融市场中的羊群行为是投资者在信息环境不确定情况下所采取的一种非理性行为,指投资者的决策行为容易受到其他投资者行为的影响,出现趋同现象。通常表现为,某个投资者根据现有信息集已经决定投资,但这时观察到其他投资者并没有投资,于是放弃之前的投资决定,选择不投资;或者这个投资者本来决定不投资的,但观察到市场上其他投资者都选择投资,于是也改变自己的决定选择投资(Duan 等,2009)<sup>[1]</sup>。

目前,对金融领域羊群行为的研究多集中在股票、证券等传统金融市场,认为羊群行为的发生会加剧市场的风险,导致市场价格波动过度,引发金融风险(Kaminsky 等,1999;蔡庆丰和宋友勇,2010;

收稿日期:2017-06-03

基金项目:教育部人文社会科学基金项目“网络借贷信用风险评估的结构化方法及应用研究”(16YJCZH031)

作者简介:都红雯,女,教授,主要从事金融工程研究;周文婷,女,硕士研究生,主要从事互联网金融研究;胡文彬,男,讲师,金融学博士,主要从事金融工程、互联网金融研究。

Ouarda 等,2013)<sup>[2-4]</sup>。对此,Bikhchandani 和 Sharma(1990)<sup>[5]</sup>认为,在分析羊群行为的过程中应注意区分“伪羊群行为”与羊群行为。“伪羊群行为”是指不同的投资者在面对相同的投资信息时,根据信息做出相同的决策判断从而产生相同的投资行为。这种行为并不是一种有意的模仿,而是主动的自主选择。但是在现实生活中,影响到投资者投资行为的因素太多,导致这两种行为往往难以区分。如股票市场中,从经济的宏观信息到行业层面的微观信息,以及个股的历史信息等,这些信息都会影响到投资者的最终决策。相比之下,P2P网贷市场中,投资者可获得的信息主要是平台上披露的借款者的个人信息以及借款标的相关信息,信息相对可控,这就为我们对羊群行为的分析提供了很好的实证样本(廖理等,2015)<sup>[6]</sup>。

作为传统信贷和互联网相结合的一种新型模式,P2P网络借贷市场为低收入者和个体经营人群提供了良好的融资渠道和融资平台。但由于投资者通过平台能够获得的信息有限,不能完全辨别借款人的真实情况,因此容易发生羊群行为;而且,相比传统金融市场,羊群行为在网络环境中更容易发生,因为在互联网平台,人们能频繁观察到其他人的行为并进行模仿。

现有的研究表明,借款人在提供性别、年龄、收入等硬信息基础上,更多地提供社交网络、借款用途以及个人描述等软信息,会影响借款成本、借款成功率和违约率。如 Lin 等(2009)<sup>[7]</sup>认为借款人的社交网络等软信息能够使最终获得的借款利率更低、违约率也更低,且成功借款的可能性更高。Larrimore 等(2011)<sup>[8]</sup>从 Prosper 网站上摘录了2005-2008年的213510个数据进行文本分析,发现文本说明中每增加一个词能够增加0.017个百分点的借款成功率。Michels(2011)<sup>[9]</sup>认为投资者会试图从借款者的描述信息中提炼出其他信息,帮助其做出最终的投资决策。Herzenstein 等(2011)<sup>[10]</sup>也得出了类似的结论,他们使用 Prosper 网站上的数据,研究发现不可验证的信息(借款人的描述说明)对借款成功率的影响甚至大于可验证的客观信息,那些被投资者认为是值得信赖的借款者更易获得投资。李焰等(2014)<sup>[11]</sup>以手工收集的“拍拍贷”网贷平台数据为样本,从借款人品质、能力和生活状况三个角度对借款人的描述性信息进行分析,发现提供描述性信息的借款人更易获得借款。王会娟和何琳(2015)<sup>[12]</sup>根据“人人贷”平台数据进行实证分析,发现借款描述中展现的信息量越多,越能降低借贷双方的信息不对称,使借款者更易成功获得借款,且满标时间越短,违约率越低。刘征驰等(2015)<sup>[13]</sup>、庄雷等(2015)<sup>[14]</sup>也认为软信息能够降低信息不对称,使借款者更易筹资成功。彭红枫等(2016)<sup>[15]</sup>根据 Prosper 网贷平台数据进行实证研究,结果表明借款陈述能降低借款成本,但是不一定能增加借款成功率。王琴英等(2017)<sup>[16]</sup>发现,软信息能够降低借贷双方的信息不对称,预判违约行为发生的可能性。于军(2017)<sup>[17]</sup>从语言信息含量和语言内容两个角度进行实证分析,结果显示借款描述中包含的信息量越大,借款人越不易违约。

但上述这些软信息能否缓解 P2P 网贷市场借贷双方的信息不对称、从而减少投资者的逆向选择和羊群行为?本文借鉴 Zhang 和 Liu(2012)<sup>[18]</sup>的实证方法,将历史投标记录按等时间划分,采用线性回归模型,以“人人贷”为例,运用 Python 编写网络爬虫程序,抓取其2011年8月15日至2014年8月20日期间的数据进行实证分析,并根据实证结果提出相应建议以完善我国 P2P 网贷市场的发展。

## 二、P2P 网贷市场羊群行为形成的机理分析

国内外学者通过构建各种理论模型来分析羊群行为。Banerjee(1992)<sup>[19]</sup>于1992年最早提出了基于信息不对称的序列性羊群行为模型(Sequential Herding);同年,Bikhchandani 等(1992)<sup>[20]</sup>也对基于信息不对称的羊群行为模型加以阐述,发展成为经典的羊群行为模型——BHW 模型。此外,较经典的羊群行为模型还包括 Scharfstein 和 Stein(1990)<sup>[21]</sup>在1990年提出的基于声誉的羊群行为模型(Reputation-based Herding)、Maug 和 Naik 于1996年提出的基于报酬机制的羊群行为模型(Compensation-based Herding)。后两类理论模型主要适用于委托代理关系中的羊群行为,与本文研究的主体——P2P 网贷市场关系不大,故本文主要运用 BHW 模型,根据 P2P 网贷市场运作规律分析 P2P 网贷市场中羊群行为形成的理论机制。

在P2P网络借贷中,假设市场上有两类可投资的借款订单 $v \in V = \{-1, 1\}$ ,其中 $v = 1$ 表示是优质的借款订单,在其他条件相同的情况下违约风险相对较低; $v = -1$ 表示借款订单是劣质的借款订单,相比之前违约风险更高。由于信息不对称,每个投资者在作出投资决策之前都可以获得一定的私人信息,这些私人信息可能是投资者凭借自身的专业知识根据市场上的公共信息进行价值判断所形成的私人信息,也可能是投资者从其他渠道所获得的私人信息。投资者只能根据市场上总的信息 $S$ (包括公共信息以及投资者获得的私人信息)做出判断: $s = 1$ 表示借款订单的相关指标传递出偏积极的信息; $s = -1$ 表示借款订单的相关指标传递出偏消极的信息。信息的精确度 $p$ 是指信息判断正确的概率,且 $p > 1/2$ , ( $Prob(s = 1/v = 1) = Prob(s = -1/v = -1) = p$ )。

市场上投资者的行为决策可以被观测,每个投资者可以根据其他投资者的行为用贝叶斯法则更新其决策依据。若第一个投资者得到积极的信息( $s = 1$ )后决定投资,得到消极的信息( $s = -1$ )后决定不投资。第二个投资者可以根据第一个投资者的行为决策推断出其掌握的信息。

假设第一个投资者不投资,说明他掌握的信息是消极的,此时如果第二个投资者得到的信息也是消极的,根据贝叶斯法则,此时他判断该借款订单为优质订单的概率为 $(1 - p^2) / [p^2 + (1 - p)^2] < 1/2$ ,决定不投资;如果第二个投资者得到的信息是积极的,此时他判断该借款订单为优质订单的概率为 $1/2$ ,由于风险中性,他以 $1/2$ 的概率投资, $1/2$ 的概率不投资。

假设第一个投资者投资,说明他掌握的信息是积极的,此时如果第二个投资者得到的信息也是积极的,根据贝叶斯法则,此时他判断该借款订单为优质订单的概率为 $p^2 / [p^2 + (1 - p)^2] > 1/2$ ,决定投资;如果第二个投资者得到的信息是消极的,此时他判断该借款订单为优质订单的概率为 $1/2$ ,由于风险中性,他以 $1/2$ 的概率投资, $1/2$ 的概率不投资。

第三个投资者可能面临的情况有两种。一种情况是,如果前两个投资者都选择投资(或不投资),根据贝叶斯法则,无论第三个投资者得到的信息如何,他判断该借款订单为优质订单的概率都大于 $1/2$ 、选择投资(或小于 $1/2$ 、选择不投资)。此时,第三个投资者的决策行为没有给后续投资者提供任何信息,第四个投资者面临的情况和第三个投资者一致。以此类推,后面所有的投资者都会选择和第三个投资者一致的投资决策,投资者的羊群行为产生。

另一种情况是,如果前两个投资者的行为决策相反,一个选择投资而一个不投资,两者信息相互抵消。于是第三个投资者面临的情况和第一个投资者相似,第三个投资者会依据自己的信息做出决策。依次类推,当决定投资的人数比决定不投资的人数多时,后续的投资者都会选择投资,即投资的羊群行为由此产生。

### 三、软信息对P2P网贷市场羊群行为影响的研究设计

国内外P2P平台交易机制不同,软信息所包含的内容也不尽相同。在国外学者的研究中,软信息主要是指P2P网贷平台中的社会资本(即借款人所在组群、社交网络中的朋友数量等)以及借款人提供的描述信息。但国内P2P平台目前尚未开启社交板块,因此所反映的软信息主要是借款人在借款陈述中提供的描述性信息,如借款用途、借款人的个人性格特征等。鉴于此,本文拟以“人人贷”平台为例,着重分析借款人提供的借款陈述等软信息能否缓解市场中的信息不对称,从而减弱羊群行为。

#### (一) 研究模型选择

对羊群行为的检验模型很多,基于P2P网贷市场的特殊性,目前学者们主要采用Logistic回归模型、需求加总的市场份额模型、线性回归模型三种。

Logistic回归模型以后续投标概率为因变量,以当前投标人数为羊群行为的代理变量,检验当前投标人数与后续投标概率之间的关系来判断羊群行为是否存在(Herzenstein等,2011)<sup>[10]</sup>;需求加总的市场份

额模型将某一时刻  $t$  平台上所有正处于筹标状态的借款订单作为一个市场,分析  $t-1$  时刻订单的筹集进度或投标人数对  $t$  时刻借款订单的市场份额的影响来检验羊群行为(Lee 等,2012<sup>[22]</sup>;廖理等,2015<sup>[6]</sup>);线性回归模型将每一借款订单中的投标记录按照时间排序划分成若干份,以  $t-1$  期累计投标人数作为自变量, $\Delta t$  期新增的投标人数为因变量,通过分析新增投标人数是否会受到上一期累积投标人数的影响来验证羊群行为(Zhang 和 Liu,2012)<sup>[18]</sup>。

三种模型中,Logistic 回归模型需要对借款订单进行后续跟踪,由此其样本只限于处于借款申请状态的订单,能获得的样本数据量相对较少;需求加总的市场份额模型在模型构建过程中存在一个假设前提,即 P2P 网贷市场中每个投资者在面对可投资的借款申请时最终只会选择投资一个借款标的,但从实际情况来看,大多数投资者在进行投资时并不会只选择投资一个借款标的,出于风险分散考虑会投资于多个借款申请。相比之下,线性回归模型则没有太多的限制,新增投标人数可以根据借款投标记录获取。因此,出于数据获取的可行性以及模型的准确性考虑,本文选择等时间段线性回归模型对羊群行为的存在性进行实证分析,具体如式(1):

$$\Delta Bid = C + \beta_1 Bid + \beta_2 Controls + \varepsilon \quad (1)$$

其中, $\Delta Bid$  代表后续投标者投标行为,用  $\Delta t$  时间段中新增的投标人数表示; $Bid$  表示其他投资者的投资行为,用上时刻累计的投标人数表示; $Controls$  代表 P2P 市场上可获得的其他影响到后续投标者投标行为的因素。

自变量  $Bid$  和因变量  $\Delta Bid$  根据每个借款订单中投标记录整理算出。根据投标时间先后排序,取小时为  $t$  的观测单位,计算每等份  $\Delta t$  内新增的投标人数,并记为  $\Delta Bid$ ,将上一时间段累计投标人数记作  $Bid$ ,以小时面板数据为样本进行研究。

如果变量  $Bid$  前的系数显著为正,说明借款标的当前投标人数会影响到后续投资者的行为决策,即在我国 P2P 网络借贷市场中羊群效应是存在的;反之,则说明羊群行为并不存在。

由于本文研究的重点在于 P2P 网贷市场中软信息的增加对羊群行为是否具有减弱的调节作用,因此,本文的实证模型将在羊群行为的存在性模型的基础上加以修改。通常,计算机字节信息与表达信息之间存在正向关系,在不考虑其他因素的情况下,人在语言表达方面与之有一定的相似性。Loughran 和 McDonald(2014)<sup>[23]</sup> 的研究表明,文本长度能够作为文件可读性度量的有效指标;廖理等(2015)<sup>[24]</sup> 也认为,从信息传递的角度来看,文字描述中包含的字数越多,传达给投资人的信息也可能更多,更多的信息能够减少借贷双方的信息不对称程度。基于此,本文假定在不对称信息环境下,行为人符合该规律,用借款人在借款陈述中文本长度即文本字符数来表示软信息的信息量大小,具体实证模型如式(2)所示:

$$\Delta Bid = C + \beta_1 Bid + \beta_2 Length + \beta_3 Length * Bid + \beta_4 Controls + \varepsilon \quad (2)$$

其中, $Length$  代表借款陈述中文本长度(文本字符数)。为考察软信息对羊群行为的调节作用,在模型中加入软信息和羊群行为的代理变量的交叉变量,即  $Length * Bid$ ,通过分析该交叉变量前的回归系数来对软信息的调节作用进行判断。当  $Length * Bid$  前的回归系数为负且通过了显著性检验,说明借款陈述的文本长度能够对市场中的羊群行为起到减弱调节作用,即软信息的文本长度越长,包含的信息量越大,越能够减弱市场的羊群行为。

## (二) 数据来源及处理

“人人贷”平台成立于2010年5月,是我国较早成立的 P2P 网络借贷平台之一,根据该平台公布的数据,截至2016年底平台累计成交额达243亿元,全年共完成133687笔交易,撮合交易累计达2724万人次,借款人分布在全国各地区372个城市,具有一定的代表性,因此本文的实证数据以“人人贷”为例。

考虑到平台运行之初存在大量实验标,所以选取的时间不宜过早。由于最长借款期限是3年,考虑到实证数据涉及到借款者最终的还款表现,如果选取的时间过近则存在许多正处于还款状态的借款订单。因此,本文选择该平台2011年8月15日到2014年8月20日交易数据为研究样本。通过运用 Python 编写网络爬虫

程序,抓取样本期借款标的相关信息,删除信息不完整的借款标,最终获得14936个有效借款订单,并根据投标时间先后排序,取小时为 $t$ 的观测单位,采用面板数据进行实证分析。

根据其他学者,如吴佳哲(2015)<sup>[25]</sup>的研究和“人人贷”平台的特点,从借款者的个人特征、标的特征以及历史借款记录三个角度选择以下控制变量放入模型中,具体变量定义赋值处理如表1所示。

表1 实证变量定义

变量类型	变量名称	变量说明	
核心变量	新增投标人数( $\Delta Bid$ )	取实际值	
	上一期累计投标人数( $Bid$ )	取实际值	
	借款陈述的文本长度( $Length$ )	文本描述的字符数	
控制变量	个人特征	年龄( $Age$ )	取实际值
		婚姻状况( $Marriage$ )	已婚取1,未婚、离异或丧偶取0
		学历( $Education$ )	高中或以下取1,专科取2 本科取3,研究生或以上取4
		信用等级( $Level$ )	“人人贷”列表信用等级从HR到AA由低到高七个等级依次取1~7
		收入( $Income$ )	月收入1000元以内取1,1000~2000取2,2000~5000取3,5000~10000取4,10000~20000取5,20000~50000取6,50000以上取7
		是否有房( $House$ )	有取1,无取0
		是否购车( $Car$ )	有取1,无取0
		是否车贷( $CarL$ )	有取1,无取0
	是否房贷( $HouseL$ )	有取1,无取0	
	标的特征	借款金额( $Amount$ )	取实际值,单位元
		借款利率( $Rate$ )	取实际值,以年为单位
		借款期限( $Time$ )	取实际值,以月为单位
		提前还款费率( $EarlyRR$ )	取实际值
		保障方式( $SaveW$ )	本金保障取1,用户利益保障机制取0
	历史借贷记录	成功借款次数( $BorrowN$ )	取实际值
		正常还清次数( $PayoffN$ )	取实际值
		申请借款次数( $AppliN$ )	取实际值
		逾期还清次数( $OverdueN1$ )	借款用户未按协议约定时间进行足额还款。取实际值
		严重逾期还清次数( $OverdueN2$ )	借款用户逾期时间超过30天时,从第31天开始该标的状态为严重逾期。取实际值

### (三) 数据的描述性统计

实证之前,先对所获得的数据进行简单的描述性统计分析,具体结果如表2所示。样本中,每个时间段中平均新增投标人数6人;借款陈述的文本长度最小为1个字符,最长为491个字符,每个借款申请的平均字符长度为143。通常,借款陈述的文本长度越长,包含的信息量越大,投资者能够从中捕捉到的关于借款人的信息也越多,就越能根据所掌握的现有信息做出理性判断,从而缓解市场上的羊群行为。

从借款的标的特征来看,大多采取的是本金保障方式,平均借款金额为5.56万元,期限1年左右,符合P2P网络借贷市场中小额短期借贷的特点。样本中平均借款利率为12.91%,说明P2P市场中的融资成本相对较高。

从借款者的信用特征来看,成功获得借款者信用等级较高,平均借款等级在B级以上,平均有2次以上的借款申请经历,逾期还清次数和严重逾期还清次数都较低,平均都低于1次。

表2 变量的描述性统计

变量类型	变量名称	极小值	极大值	均值	标准差
核心变量	新增投标人数( $\Delta Bid$ )	0.00	732.00	5.70	12.39
	上一期累计投标人数( $Bid$ )	0.00	1824.00	25.05	47.07
	借款陈述的文本长度( $Length$ )	1.00	491.00	142.98	92.89
个人特征	年龄( $Age$ )	22.00	74.00	39.00	8.54
	婚姻状况( $Marriage$ )	0.00	1.00	0.73	0.44
	学历( $Education$ )	1.00	4.00	1.95	0.75
	信用等级( $Level$ )	1.00	7.00	5.43	1.49
	收入( $Income$ )	1.00	7.00	4.62	1.35
	是否有房( $House$ )	0.00	1.00	0.38	0.49
	是否购车( $Car$ )	0.00	1.00	0.25	0.43
	是否车贷( $CarL$ )	0.00	1.00	0.16	0.37
标的特征	是否房贷( $HouseL$ )	0.00	1.00	0.16	0.37
	借款金额( $Amount$ )	3000.00	3000000.00	55052.52	75479.00
	借款利率( $Rate$ )	6.00	24.00	12.92	1.23
	借款期限( $Time$ )	1.00	36.00	22.95	10.41
	提前还款费率( $EarlyRR$ )	0.00	1.00	0.30	0.46
历史借贷记录	保障方式( $SaveW$ )	0.00	1.00	0.17	0.38
	成功借款次数( $BorrowN$ )	1.00	144.00	1.91	6.79
	正常还清次数( $PayoffN$ )	0.00	144.00	1.84	6.76
	申请借款次数( $AppliN$ )	1.00	148.00	2.26	7.33
	逾期还清次数( $OverdueN1$ )	0.00	54.00	0.75	3.15
严重逾期还清次数( $OverdueN2$ )	0.00	7.00	0.04	0.25	

## 四、软信息对 P2P 网贷市场羊群行为影响的实证分析

### (一) 面板回归验证

首先,对面板数据进行 Hausman 检验以确定采用固定效应还是随机效应。以新增投标人数  $\Delta Bid$  为因变量时, Hausman 检验的统计值为 204, 对应的  $p$  值为 0.00, 拒绝原假设, 故选用固定效应模型。实证结果如表 3 所示, 在控制了其他变量的情况下, 上一期累计投标人数  $Bid$  前的回归系数为 0.0137, 且通过显著性检验, 说明借款标的当前投标人数会影响到后续投资者的行为决策, 即在我国 P2P 网络借贷市场中羊群效应是存在的。借款陈述的文本长度 ( $Length$ ) 前的回归系数为 0.0160, 且通过显著性检验, 说明借款者在借款描述中的文本长度越长, 越能够帮助借款者获得后续投标。借款陈述的文本长度和羊群行为代理变量的交叉变量  $Length * Bid$  前的回归系数为 -0.0007, 与预期相一致, 但回归结果并不显著。通过逐一分析样本数据发现, 其原因在于借款保障方式中属于本金保障机制的借款陈述基本上是以一种固定表达方式给出, 所含信息多是借款人已披露过的一些硬信息的简单重复, 如“此次借款用于  $\times \times$ , 借款人在  $\times \times$  公司, 该公司位于  $\times \times$ , 主要从事  $\times \times$ , 每月有稳定的流水。上述信息已经  $\times \times$  实地认证。同时, 经审核借款人所提供资料真实有效, 符合人人贷的借款审核标准”等。该类信息内容经平台审核后规范化处理, 不属于本文对软信息的定义, 包含这部分借款陈述的借款订单在总样本中占比近 40%, 会造成对实证结果的干扰, 因此在实证中应该删除此类借款订单。

表3 面板回归结果1

变量类型	变量名称	标准系数	标准误差	t	P
主要自变量	上一期累计投标人数(Bid)	0.0137	0.0037	3.7152	0.0000
	借款陈述的文本长度(Length)	0.0160	0.0031	5.2069	0.0000
	Length * Bid	-0.0007	0.2914	-0.0024	0.2791
个人特征	年龄(Age)	0.0260	0.0067	3.8509	0.0001
	婚姻状况(Marriage)	-0.0294	0.1132	-0.2594	0.7954
	学历(Education)	-0.1431	0.0607	-2.3592	0.0183
	信用等级(Level)	0.1020	0.0432	2.3632	0.0181
	收入(Income)	0.2693	0.0387	6.9601	0.0000
	是否有房(House)	-0.1222	0.1189	-1.0284	0.3038
	是否购车(Car)	0.0147	0.1155	0.1269	0.8990
标的特征	是否车贷(CarL)	-0.5031	0.1818	-2.7680	0.0056
	是否房贷(HouseL)	-0.3097	0.1354	-2.2869	0.0222
	借款金额(Amount)	0.0001	0.0000	40.7271	0.0000
	借款利率(Rate)	-0.0362	0.0181	-1.9959	0.0459
	借款期限(Time)	-0.0137	0.0074	-1.8544	0.0637
	提前还款费率(EarlyRR)	-12.5362	0.7560	-16.5820	0.0000
历史借贷记录	保障方式(SaveW)	13.6734	0.7783	17.5683	0.0000
	成功借款次数(BorrowN)	-0.1058	0.1792	-0.5902	0.5551
	正常还清次数(PayoffN)	0.1483	0.1775	0.8354	0.4035
	申请借款次数(AppliN)	-0.0502	0.0231	-2.1694	0.0301
	逾期还清次数(OverdueN1)	0.0415	0.0163	2.5451	0.0109
严重逾期还清次数(OverdueN2)	-0.3016	0.2486	-1.2133	0.2250	
	常数	0.1772	0.6073	0.2918	0.7704

将借款数据中属于本金保障机制的借款订单删除后,剩下9522个有效借款订单。按照本文的实证思路对样本重新筛选后再采用面板回归的方式进行验证,结果如表4所示。在控制其他变量的情况下,Bid前的回归系数通过了0.01水平下的显著性检验且为正,说明市场中羊群行为的存在。借款陈述的文本长度(Length)前的回归系数为0.0042,且通过显著性检验,说明借款者在借款描述中的文本长度越长,越能够帮助借款者获得后续投标。借款陈述的文本长度和羊群行为代理变量的交叉变量Length \* Bid前的回归系数为-0.1046,且通过了显著性检验,说明借款陈述的文本长度能够对市场中的羊群行为起到减弱的调节作用,实证结果与本文假设一致,通过验证。也就是说P2P网贷市场中借款陈述等软信息能够影响到市场上的羊群行为,借款陈述越长的借款订单补充的信息量更多,从而有效地缓解了市场上借贷双方的信息不对称,并减弱了市场上的羊群行为。

从其他控制变量来看,借款人年龄、收入能够帮助借款人获得后续投标,借款人年龄越大,向借款人传递出成熟稳重的可靠信息,一般来说年龄较大的人群工作收入也相对比较稳定。收入越大展示出更强的还款能力,更易获得投资者的借款。借款期限显著为负,说明借款期限越短的借款订单越易获得投资者的青睐,因为借款期限在一定程度上反映了出借人所面临的风险,当借款期限越长时,借款人违约的风险也越大,因此当其他情况相同时,借款人更倾向于借款期限较短的订单。从历史借贷记录来看,成功借款次数、正常还清次数的回归系数显著为正,这两个指标反映出借款人如约还款的信息,展现出借款人信用品质,能够帮助借款人获得后续借款。严重逾期还清次数的回归系数显著为负,该数据反映出借款人的违约风险,对借款人获得后续借款产生消极影响。

表4 面板回归结果2

变量类型	变量名称	标准系数	标准误差	t	Sig.
主要自变量	上一期累计投标人数 (Bid)	0.0246	0.0027	9.1283	0.0000
	借款陈述的文本长度 (Length)	0.0042	0.0006	7.0867	0.0000
	Length * Bid	-0.1046	0.0330	-3.1695	0.0006
个人特征	年龄 (Age)	0.0172	0.0064	2.6866	0.0072
	婚姻状况 (Marriage)	0.1216	0.1014	1.1993	0.2304
	学历 (Education)	-0.0665	0.0502	-1.3249	0.1852
	信用等级 (Level)	-0.0456	0.0302	-1.5134	0.1302
	收入 (Income)	0.2041	0.0328	6.2155	0.0000
	是否有房 (House)	-0.3493	0.0988	-3.5347	0.0004
	是否购车 (Car)	0.1129	0.0987	1.1441	0.2526
	是否车贷 (CarL)	-0.3132	0.1529	-2.0482	0.0405
	是否房贷 (HouseL)	-0.0273	0.1169	-0.2332	0.8156
标的特征	借款金额 (Amount)	0.0001	0.0000	65.3488	0.0000
	借款利率 (Rate)	-0.0154	0.0126	-1.2231	0.2213
	借款期限 (Time)	-0.0358	0.0071	-5.0129	0.0000
	提前还款费率 (EarlyRR)	0.3313	1.1227	0.2951	0.7679
历史借贷记录	成功借款次数 (BorrowN)	0.9184	0.1899	4.8372	0.0000
	正常还清次数 (PayoffN)	0.8522	0.1895	4.4976	0.0000
	申请借款次数 (AppliN)	-0.0517	0.0157	-3.2826	0.0010
	逾期还清次数 (OverdueN1)	0.0248	0.0113	2.2047	0.0275
	严重逾期还清次数 (OverdueN2)	-1.1148	0.2156	-5.1713	0.0000
	常数	0.6526	1.1849	0.5508	0.5818

## (二) 稳健性检验

出于实证的稳健性考虑,本文用两种方法对实证结果进行稳健性检验:第一种方法是对软信息代理变量即借款陈述的文本长度进行1%水平下的缩尾处理,以减小极端值的影响;第二种方法是将借款陈述的文本长度即字符数用借款陈述的中文字符数来代替,去掉借款陈述中的非中文单词所占字符数。实证结果分别如表5、表6所示,其中表5是第一种稳健性检验的实证结果,表6是第二种稳健性检验的实证结果。

表5 稳健性检验1

变量类型	变量名称	标准系数	标准误差	t	Sig.
主要自变量	上一期累计投标人数 (Bid)	0.0306	0.0037	8.3514	0.0000
	借款陈述的文本长度 (Length)	0.0068	0.0023	2.9646	0.0495
	Length * Bid	-0.2070	0.0409	-5.0611	0.0000
个人特征	年龄 (Age)	0.0149	0.0065	2.2715	0.0231
	婚姻状况 (Marriage)	0.1306	0.1022	1.2780	0.2012
	学历 (Education)	-0.1061	0.0511	-2.0776	0.0377
	信用等级 (Level)	-0.0490	0.0307	-1.5973	0.1102
	收入 (Income)	0.1996	0.0334	5.9783	0.0000
	是否有房 (House)	-0.3448	0.1000	-3.4495	0.0006
	是否购车 (Car)	0.0870	0.1002	0.8687	0.3850
	是否车贷 (CarL)	-0.1896	0.1564	-1.2121	0.2255
	是否房贷 (HouseL)	0.0173	0.1184	0.1462	0.8838
标的特征	借款金额 (Amount)	0.0001	0.0000	61.3625	0.0000
	借款利率 (Rate)	-0.0265	0.0127	-2.0903	0.0366
	借款期限 (Time)	-0.0339	0.0075	-4.4937	0.0000
	提前还款费率 (EarlyRR)	0.5084	1.4070	0.3613	0.7178



(续表5)

变量类型	变量名称	标准系数	标准误差	t	Sig.
历史借贷记录	成功借款次数(BorrowN)	1.0273	0.1994	5.1515	0.0000
	正常还清次数(PayoffN)	0.9657	0.1989	4.8542	0.0000
	申请借款次数(AppliN)	0.0514	0.0156	3.3015	0.0010
	逾期还清次数(OverdueN1)	0.0193	0.0115	1.6713	0.0947
	严重逾期还清次数(OverdueN2)	-1.1753	0.2269	-5.1797	0.0000
	常数	0.4052	1.4676	0.2761	0.7825

表6 稳健性检验2

变量类型	变量名称	标准系数	标准误差	t	Sig.
主要自变量	上一期累计投标人数(Bid)	0.0257	0.0040	6.3455	0.0000
	借款陈述的文本长度(Length)	0.0053	0.0009	3.7351	0.0003
	Length * Bid	-0.2028	0.0501	-4.0476	0.0000
个人特征	年龄(Age)	0.0153	0.0065	2.3388	0.0194
	婚姻状况(Marriage)	0.1348	0.1022	1.3194	0.1870
	学历(Education)	-0.1069	0.0511	-2.0946	0.0362
	信用等级(Level)	-0.0501	0.0307	-1.6346	0.1021
	收入(Income)	0.2006	0.0334	6.0072	0.0000
	是否有房(House)	-0.3437	0.0999	-3.4396	0.0006
	是否购车(Car)	0.0844	0.1002	0.8424	0.3996
	是否车贷(CarL)	-0.1853	0.1564	-1.1846	0.2362
	是否房贷(HouseL)	0.0163	0.1184	0.1379	0.8903
标的特征	借款金额(Amount)	0.0001	0.0000	61.4249	0.0000
	借款利率(Rate)	-0.0269	0.0127	-2.1142	0.0345
	借款期限(Time)	-0.0341	0.0075	-4.5262	0.0000
	提前还款费率(EarlyRR)	0.4924	1.4069	0.3500	0.7264
历史借贷记录	成功借款次数(BorrowN)	1.0317	0.1993	5.1767	0.0000
	正常还清次数(PayoffN)	0.9696	0.1988	4.8769	0.0000
	申请借款次数(AppliN)	0.0518	0.0156	3.3271	0.0009
	逾期还清次数(OverdueN1)	0.0192	0.0115	1.6659	0.0957
	严重逾期还清次数(OverdueN2)	-1.1778	0.2268	-5.1926	0.0000
	常数	0.4352	1.4694	0.2961	0.7671

从表5和表6可以看出,上一期累计投标人数(Bid)和借款陈述的文本长度(Length)前的回归系数均通过了显著性检验且为正。借款陈述的文本长度和羊群行为代理变量的交叉变量Length \* Bid前的回归系数分别为-0.2070和-0.2028,且均通过了显著性检验,说明借款陈述的文本长度会对市场上的羊群行为起到减弱的调节作用。

以上两种方法得出的结论表明,本文的实证分析具有稳健性。

## 五、研究结论与建议

综上所述,在我国P2P网贷市场中存在羊群行为,但借款人提供的软信息能够在一定程度上减弱该羊群行为。借款人在借款陈述中提供的信息越丰富,P2P网贷市场上的信息不对称程度随之减弱,投资者将越倾向于依据自己的判断做出投资决策,从而减弱了市场上的羊群行为。对此,本文提出如下建议:

第一,P2P网贷平台在完善市场硬信息的同时,应该进一步丰富软信息的内容,如借鉴国外Prosper等P2P平台运行机制,构建平台社交网络,以缓解市场上的信息不对称,引导投资者合理做出投资决策,降低投资者的羊群行为。

第二,借款者在申请借款时,除提交借款人个人信息以及借款标的等必要的硬信息外,还应该提供更多展示借款人还款能力、信用状况以及借款用途等软信息,让投资人更多了解,以缓解市场上借贷双方的信息不对称。

第三,投资者在筛选优质借款申请、进行投资决策时,可以适当结合借款人的软信息特征,从中挖掘出更多关于借款人相关状况的信息,以此结合市场上的其他信息来做出更加合理、理性的投资决策。

相信通过借贷双方以及网络借贷平台的共同努力,可以引导投资者做出更加理性的投资决策,从而缓解 P2P 网贷市场上羊群行为,促进网贷平台健康良性发展。

#### 参考文献:

- [1] DUAN W, GU B, WHINSTON A B. Informational cascades and software adoption on the internet: an empirical investigation[J]. *Mis Quarterly*, 2009, 33(1): 23-48.
- [2] KAMINSKY G, SCHMUKLER T. What triggers market jitters: a chronicle of the Asian crisis[J]. *Journal of International Money and Finance*, 1999, 18(4): 537-560.
- [3] 蔡庆丰, 宋友勇. 超常规发展的机构投资者能稳定市场吗? ——对我国基金业跨越式发展的反思[J]. *经济研究*, 2010(1): 90-101.
- [4] OUARDA M, BOURI A, BERNARD O. Herding behavior under markets condition: empirical evidence on the European financial markets[J]. *Journal of Economics and Financial Issues*, 2013, 3(1): 214-228.
- [5] BIKHCHANDANI S, SHARMA S. Herd behavior in financial markets[M]. Oxford: Oxford University Press, 1990: 206-233.
- [6] 廖理, 李梦然, 王正位, 等. 观察中学习: P2P 网络投资中信息传递与羊群行为[J]. *清华大学学报(哲学社会科学版)*, 2015(1): 156-165.
- [7] LIN M. Peer-to-Peer lending: an empirical study[EB/OL]. (2009-07-31)[2017-03-20]. [http://aisel.aisnet.org/amcis2009\\_dc/17](http://aisel.aisnet.org/amcis2009_dc/17).
- [8] LARRIMORE L, JIANG L, LARRIMORE J. Peer to Peer lending: the relationship between language features, trustworthiness, and persuasion success[J]. *Journal of Applied Communication Research*, 2011, 39(1): 19-37.
- [9] MICHELS J. Do unverifiable disclosures matter? —Evidence from Peer-to-Peer lending[J]. *Social Science Electronic Publishing*, 2011, 87(4): 1385-1413.
- [10] HERZENSTEIN M, SONENSHEIN S, DHOLAKIA U M. Tell me a good story and I may lend you money: the role of narratives in Peer-to-Peer lending decisions[J]. *Journal of Marketing Research*, 2011, 48(SPL): 138-149.
- [11] 李焰, 高弋君, 李珍妮, 等. 借款人描述性信息对投资人决策的影响——基于 P2P 网络借贷平台的分析[J]. *经济研究*, 2014(S1): 143-155.
- [12] 王会娟, 何琳. 借款描述对 P2P 网络借贷行为影响的实证研究[J]. *金融经济研究*, 2015(1): 77-85.
- [13] 刘征驰, 赖明勇. 虚拟抵押品、软信息约束与 P2P 互联网金融[J]. *中国软科学*, 2015(1): 35-46.
- [14] 庄雷, 周勤, 伏玉林. P2P 网络借贷用途、投资偏向效应与软信息价值[J]. *国际商务(对外经济贸易大学学报)*, 2015(6): 77-85.
- [15] 彭红枫, 赵海燕, 周洋. 借款陈述会影响借款成本和借款成功率吗? ——基于网络借贷陈述的文本分析[J]. *金融研究*, 2016(4): 158-173.
- [16] 王琴英, 由林青, 王佳佳, 等. 互联网金融背景下“软信息”对于违约行为的影响——基于 Logistic 模型[J]. *金融理论与实践*, 2017(3): 60-65.
- [17] 于军. 描述性信息与违约行为的关系研究——基于 P2P 平台的分析[J]. *当代经济管理*, 2017(5): 86-92.
- [18] ZHANG J, LIU P. Rational herding in microloan markets[J]. *Social Science Electronic Publishing*, 2012, 58(5): 892-912.
- [19] BANNERJEE A. A simple model of herd behavior[J]. *Quarterly Journal of Economics*, 1992, 107(3): 797-817.
- [20] BIKHCHANDANI S, HIRSHLEIFER D, WELCH I. A theory of fads, fashion, custom, and cultural change as informational cascades[J]. *Social Science Electronic Publishing*, 1992, 100(5): 992-1026.
- [21] SCHARFTEIN D, STEIN J. Herd behavior and investment[J]. *The American Economic Review*, 1990, 80(3): 465-479.
- [22] LEE E, LEE B. Herding behavior in online P2P lending: an empirical investigation[J]. *Electronic Commerce Research & Applications*, 2012, 11(5): 495-503.
- [23] LOUGHRAN T, MCDONALD B. Measuring readability in financial disclosures[J]. *The Journal of Finance*, 2014, 69(4): 1643-1671.
- [24] 廖理, 吉霖, 张伟强. 语言可信吗? 借贷市场上语言的作用——来自 P2P 平台的证据[J]. *清华大学学报(自然科学版)*, 2015(4): 413-421.
- [25] 吴佳哲. 基于羊群效应的 P2P 网络借贷模式研究[J]. *国际金融研究*, 2015(11): 88-96.

