

文章编号:1003-207(2018)10-0030-11

DOI:10.16381/j.cnki.issn1003-207x.2018.10.004

P2P 借贷投资者的群体智慧

廖理, 向佳, 王正位

(清华大学五道口金融学院, 北京 100083)

摘要:互联网金融领域“去专家化”的投资趋势正在凸显群体意见的重要性,而群体是否有智慧就是一个非常重要的话题。本文使用了中国一家 P2P 借贷平台的交易数据,借助构建的一个羊群行为指标研究了群体智慧对于违约率的预测。我们发现,在控制借款各项特征的情况下,随着羊群效应程度的增强,借款的违约率在下降。进一步的研究发现,对于两笔特征相似的借款,投资者羊群效应程度高的借款违约率显著低于未发生羊群效应的借款,因此投资者羊群效应程度可以为判别借款违约概率增加新的信息。久期分析的结果指出,在借款已经正常还款若干个月的条件下,投资者羊群效应程度高的借款在下个月发生违约的风险也会显著降低。总而言之,本文研究结果表明:在 P2P 网贷的投资中,群体是有智慧的,并且群体智慧能够提供新的信息。

关键词:群体智慧;羊群效应;P2P 借贷

中图分类号:F832 文献标识码:A

1 引言

随着互联网经济的发展,基于互联网的投资机会日益增多。在 2008 年,至少有四分之一的美国人依赖社交媒体上的投资建议进行投资。相对于传统金融投资机会,互联网金融投资的一个重要特征是其失去了专家的指导意见。这种“去专家化(De-professionalized)”的投资凸显了群体智慧在金融市场中的重要性。

群体是否有智慧的话题早在 1852 年就被 Mackay 讨论过,他警告认为群体是疯狂的。这可能是由于群体总体判断被群体系统性的认识偏差所扭曲;也有可能是个体无法独立思考而导致他们的判断被群体观点所影响。

但是近年来的更多的研究却表明群体是有智慧的,群体决策要比个体决策更准确。Larrick 和 Soll^[13]认为群体决策更准确原因在于将不同个体的意见整合,能够抵消不同个体的判断误差。苏寒和胡笑旋^[20]也认为互联网上涌现出的群体智慧成为解决复杂决策问题的宝贵资源。为了进一步优化意

见整合的方法,Budescu 和 Chen^[3]提出了贡献度加权模型,根据每个个体的不同专业程度分配以不同的权重。当然,这样的意见整合机制有效的前提是个体判断产生满足独立性。Davis—Stober 等^[9]进一步指出如果个体进行判断的独立性增强,则群体判断的差异系统性地增大,群体的智慧将表现地更为明显。还有一种对于群体智慧产生的解释是 Zhang Juanjuan 和 Liu Peng^[19]提出的,他们认为群体行为是理性的,因此每个个体都能够使用公共信息来理性修正自己的判断而不是简单地模仿。

群体智慧的重要性体现在其对于市场的预测能力。李存金和王俊鹏^[21]发现了群体智慧在重大航天工程设计方案形成中的重要价值。在体育比赛中,群体的智慧通常被用来预测球队的竞争力。Herzog 和 Hertwig^[11]利用 2006 年世界杯和 2008 年欧洲杯的数据分析发现群体预测的准确性和官方权威预测相当,且比赌球者的预测结果更准确。在经济金融领域,利用群体智慧进行预测准确度依旧较高:Ray^[16]发现群体的智慧在对于利率、汇率、通胀率、股价、商品价格以及许多其他的宏观经济金融变量的预测上都有惊人的准确性。从微观层面看,Mollick 和 Nanda^[15]利用美国众筹平台 Kickstarter 的数据研究发现群体和专家对于大多数项目的意见是比较一致的,群体智慧可以预测专家意见。Chen Hailiang 等^[5]则研究了美国一家股票交流社交平台

收稿日期:2017-09-11; 修订日期:2018-03-09

基金项目:国家自然科学基金资助项目(71472100); 国家社科基金重大项目(15ZDA029)

通讯作者简介:向佳(1990-),男(汉族),福建人,清华大学五道口金融学院博士生,研究方向:行为金融, E-mail:xiangj.14@pbesf.tsinghua.edu.cn.

Seeking Alpha 的文字性投资意见的信息含量,结果发现这些投资建议在一定程度上能够预测股票的未來收益,由于这些投资建议是由“人民群众”而非“专家”提供的,因此这一定程度上体现了群体的智慧。

群体的智慧不仅有预测作用,更重要的是它还有修正专家意见的作用。Baruch 等^[2]的研究表明,如果存在部分具有独立观点的市场参与者时,群体甚至能够修正专家发布的错误信息,使得市场价格信号趋于正常。

但是上述研究还存在着一些不足。虽然之前的文章都提到群体的智慧确实存在,且能够预测乃至修正专家意见,也给出了群体智慧产生的机制解释,但是鲜有文献提出市场参与者该如何通过观测群体行为来发现这样的群体智慧所传达的新信息,并将这种新的信息作为我们决策的评判标准。

为了有效地观测群体行为,我们首先需要找到没有专家意见,只能观察到群体行为的交易市场。目前快速发展的互联网金融给我们提供了这样研究的基本条件。在很多互联网金融投资中,投资人看到的关于投资借款的信息都是一致的,没有专家的投资意见可供参考,唯一可以参考的外部信息就是群体中其他投资者的行为。由于中国的征信体系不完善,信息造假成本低,而信息验证成本较高,这种投资环境下,更需要投资者发挥智慧寻找优质借款,因此,基于中国市场研究群体的智慧更具现实意义。

此外,用互联网金融研究群体智慧还存在以下优点:首先,多数互联网金融中的投资没有价格杠杆的调节,投资者的投资行为和交易价格不会像股票市场那样交互影响,因此更容易甄别群体的智慧。以 P2P 网络借贷为例,交易价格——借款利率多数是平台在交易前即给定的。其次,相对于股票市场来说,P2P 网络借贷中的投资行为属于单向投资行为,不需要交易对手方,因此更容易清晰地界定目标群体。相反,二级股票市场上任何一笔股票买入都对应一笔股票卖出,因此一方群体的智慧往往会对应另一方面群体的非智慧,但是互联网金融则可以弥补这一缺陷。第三,对于股票交易来说,要验证群体的智慧,很难准确界定观测窗口——因为交易缺少自然的交易终止时点,但是 P2P 市场对于一笔借款的投资以该借款的满标为标志自然终结,这就克服了交易终止时点不清晰给研究带来的困扰。

在股票市场上,羊群行为通常通过对正常状态的偏离程度来度量。Lakonishok 等^[12]提出利用对某股票净买入的基金数与总体均值的偏离程度,来

衡量羊群行为的程度;Christie 和 Huang^[7]则利用横截面收益标准差(CSSD)计算股票收益率和市场收益率的偏离程度,来衡量投资者决策的一致性;Chang 和 Dong^[4]使用股票的机构持股情况变化来度量羊群行为。上述度量方法的局限性在于时效性和预测力不足,对投资的指导价值有限。此外,传统度量方法需要引入正常状态,即均值的估计值,不仅容易产生误差,也不便于快速计算,实用性不足。

为了构建领先性指标,我们首先需要寻找可即时观测的数据。在人人贷上,每笔借款的即时动态投资记录都能被观测得到,且所有投资者在投资时都能观测到别人对该借款的投资状况。基于这些信息,我们构建了一个独特的指标来反应每笔借款被投资的羊群程度。根据每笔借款的独立数据可以形成针对该笔借款的羊群效应估计值。如果累计投资额在投资期限内总体呈加速增长,投资活动愈发积极,羊群效应的度量值为正。如果累计投资额在投资期限内总体呈减速增长,投资活动愈发消极,羊群效应的度量值为负。如果累计投资额在投资期限内匀速增长,投资活动强度不变,羊群效应的度量值为零。

为了验证借款羊群行为所反应的群体的智慧,我们首先将样本中的借款分为违约借款组和未违约借款组,发现未违约借款组的投资者羊群效应确实要比违约借款组更明显。随后,为了验证借款的羊群效应的预测作用,我们以是否违约作为被解释变量,以羊群效应以及一系列控制变量作为被解释变量进行 Probit 回归,结果显示羊群效应程度越高的借款,其违约率越低。

我们接下来验证的是,对于两笔基本情况非常相似的借款,借款投资过程中的羊群效应程度能否作为判别其是否会违约的一个标准。我们通过倾向性得分匹配得到两笔倾向性得分非常接近的借款,如果一笔借款的投资者羊群效应程度较高,一笔借款的投资者羊群效应程度较低,则具有羊群效应的借款违约率更低。这表明群体的智慧传达了关于违约率的新信息。

此外,我们还引入久期分析模型,考察借款者已经按约定正常还款一段时间之后,投资者的羊群效应对其下一期发生违约风险的影响。久期分析模型的结果同样显示:随着羊群效应的增强,借款在其还款期间的违约风险会逐步降低。

在稳健性检验部分,我们考虑了若干种可能会影响本文结论成立的原因,发现本文结论依然成立。

2 数据描述及变量定义

2.1 数据样本

本研究中,我们使用中国一个 P2P 网络借贷平台“人人贷”信用借款的交易数据研究群体的智慧。人人贷成立于 2010 年 5 月,是中国影响较大的 P2P 网络借贷平台之一。

在人人贷上,借款者需要提交包含个人基本信息在内的借款,投资人可以对他们所感兴趣的借款进行投资具体来讲,潜在借款人通过在人人贷上发布借款来说明他们需要的借款金额、借款利率以及借款期限。每笔投借款投资金额和准确的投资时间也可以从人人贷网站上获得,投资人能够随时了解当前的投资进程。借款人还可以提供一系列信息来提升自身的可信度,而人人贷则会对借款人提供的信息进行认证,并根据自己掌握的信息对借款人信息进行信用打分和信用评级,评级包括 AA、A、B、C、D、E、HR 七个等级,其中 AA 表示信用情况最佳,HR 表示信用情况最差。

本研究的样本包括人人贷上自 2010 年 10 月 16 日至 2014 年 12 月 31 日的所有成功募集资金的信用标,在此期间共有 17,334 个借款成功募集资金,剔除 1,167 个尚在还款中的借款,研究共包含了 16,167 个样本。

2.2 群体智慧的度量

为了研究的进行,我们需要首先衡量 P2P 投资中群体的智慧。衡量群体智慧的一种方式是把群体的行为和行为的结果结合起来,在没有专家意见指导的情况下,如果群体的行为仍有“趋利避害”的特征,那么我们就可以推断群体是有智慧的——根据这种思想,Zhang Juanjuan 和 Liu Peng^[19]关于 P2P 中羊群行为的研究能带给我们一些启发。Zhang Juanjuan 和 Liu Peng^[19]认为 P2P 网络借款的投资者存在理性的羊群行为,具体表现为借款下一期的投资量会受到之前累计募集资金量的影响,而从结果来看,这种反馈机制是理性的。基于 P2P 市场的特殊性,本文提出了一个新的方法来度量每个借款的羊群程度:既然投资额在投资期限内的分布能在一定程度上衡量羊群行为,那么我们就可以通过捕捉投资额在投资期限内分布的不均匀程度来衡量一个具体借款的羊群程度——该方法和统计中广泛应用的“基尼系数”思想类似。具体地,我们通过每一笔投借款信息来刻画借款随时间的完成速度,图 1

展示了两种典型的借款完成速度。

如图 1 所示,纵轴表示的是累计完成的投资额,横轴表示投资随时间的进程。原点所对应的的时间是对于该借款的第一个投资时间。每一个矩形的高度代表其对应的时点的累计投资金额。每一个矩形的宽度代表两个投资时点之间的时间间隔。所有矩形的面积之和被表示为 $\sum_{n=1}^N S_n$ 。三角形 OTC 的面积表示为 S_T 。仿照基尼系数的定义,本文将借款 i 的羊群程度定义为:

$$Herding_i = \frac{S_{T,i} - \sum_{n=1}^N S_{n,i}}{S_{T,i}}$$

$Herding$ 的取值范围是 -1 到 1。当 $Herding$ 取值在 0 在 1 之间时,表示总体而言,投资者的投资活动在该时间范围内时逐渐增强的,投资者的羊群效应较强。当 $Herding$ 取值在 -1 在 0 之间时,表示总体而言,投资者的投资活动在该时间范围内时逐渐减弱的,投资者的羊群效应较弱。

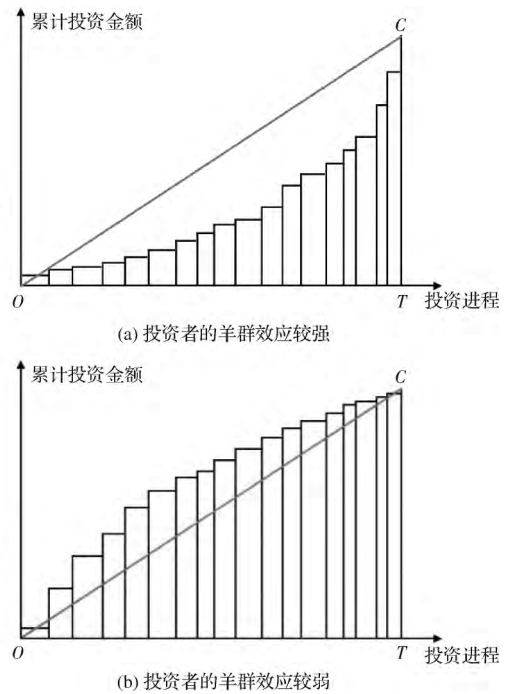


图 1 投资者羊群效应模式说明

本文通过 t 检验发现违约借款的 $Herding$ 值显著低于未违约借款,即投资者对于违约借款会呈现出更弱的羊群效应,这意味着羊群效应是有信息量的,符合预期。

2.3 主要变量定义

本文后续分析中涉及的变量定义如下:

表 1 主要变量描述性统计

符号	变量名	定义
<i>Default</i>	借款是否违约	被解释变量为借款的违约情况 <i>Default</i> , 如果借款最终违约, <i>Default</i> 取值为 1, 否则取值为 0。
<i>IRR</i> (%)	借款的内部收益率	人人贷披露了所有借款的还款记录, 我们可以通过借款的还款现金流以及如下公式 $Amount_i = \sum_{t=1}^T \frac{RepaymentCashFlow_{i,t}}{(1+IRR_i)^t}$ 计算出每笔借款的内部收益率, 单位: %。
<i>Interest</i> (%)	借款利率	定义为借款贷款的借款利率, 单位: %。
<i>Term</i> (月)	借款期限	定义为借款贷款的借款期限, 包括 3、6、9、12、18、24、36 个月
$Ln(Amount)$ (¥)	借款金额对数值	定义为每笔借款要求的借款金额的自然对数
<i>HR</i>	借款人信用级别	如果借款人的信用评级为 HR(High Risk) 级, 该变量取值为 1, 否则为 0
<i>Male</i>	借款人性别	借款人为男性时取值为 1, 否则为 0
<i>Age</i>	借款人年龄	借款人在人人贷平台提交的年龄信息
<i>Bachelor</i>	借款人的最高学历为本科	如果借款人的最高学历为本科, 该变量取值为 1, 否则为 0
<i>MasterOrAbove</i>	借款人的最高学历为硕士及以上	如果借款人的最高学历为硕士及以上, 该变量取值为 1, 否则为 0;
<i>Employ(3-5yrs)</i>	借款人的工作年限在 3-5 年之间	如果借款人的工作年限在 3-5 年之间, 该变量取值为 1, 否则为 0
<i>Employ(5yrs+)</i>	借款人的工作年限在 5 年以上	如果借款人的工作年限大于 5 年, 该变量取值为 1, 否则为 0
<i>Income</i> (¥5,000 - 10,000)	借款人月收入处于 0-10,000 元	如果借款人的月收入处于 0-10,000 元之间, 该变量取值为 1, 否则为 0
<i>Income</i> (¥10,000 - 20,000)	借款人月收入处于 10,000-20,000 元	如果借款人的月收入处于 10,000-20,000 元人民币之间, 该变量取, 否则为 0
<i>Income</i> (¥20,000 - 50,000)	借款人月收入处于 20,000-50,000 元	如果借款人的月收入处于 20,000-50,000 元人民币之间, 该变量取值为 1, 否则为 0
<i>Income</i> (¥50,000+)	借款人月收入超过 50,000 元	如果借款人的月收入超过 50,000 元人民币之间, 该变量取值为 1, 否则为 0
$Ln(Income/Debt)$	收入债务比对数值	我们将借款者月收入取区间中的最低值, 并将其除以其每月需偿还的本息金额, 并将该比值加 1 后取对数
<i>House</i>	借款人是否有房产	如果借款人有房产, 该变量取值为 1, 否则为 0
<i>Mortgage</i>	借款人是否有房贷	如果借款人有正在偿还的房贷, 该变量取值为 1, 否则为 0
<i>Car</i>	借款人是否有车产	如果借款人拥有私家车, 该变量取值为 1, 否则为 0
<i>CarLoan</i>	借款人是否有车贷	如果借款人有正在偿还的车贷, 该变量取值为 1, 否则为 0
<i>AnalysisTime</i> (%)	分析时间	每笔借款从开始还款到还清或者违约的时间间隔占其借款期限的百分比

3 实证结果

3.1 回归模型

虽然相关系数的结果一定程度上说明了投资者羊群行为程度与借款是否违约之间的相关关系, 但是这些结果尚未控制其它因素的影响。为此, 我们以投资者羊群行为程度 (*Herding*) 作为核心解释变量, 以借款是否发生违约 (*Default*) 作为被解释变

量, 进行回归分析。由于 *Default* 是 0-1 变量, 本文选择用 Probit 模型, 构建回归方程如下:

$$Pr(Default_i = 1) = F(\alpha + \beta * Herding_i + \gamma * Characteristic_i + \epsilon_i)$$

其中 i 表示的是借款 i , $Characteristic_i$ 表示借款 i 的其他特征变量。为了控制借款发布的时间 (例如上班时间与下班时间、工作日与周末、春节期间等) 可能会对投资者的羊群效应程度造成一定的

影响,我们在表 2 的回归中逐一加上 Hour-of-Week 固定效应、Day-of-Week 固定效应和 Month

表 2 羊群效应对借款是否违约的回归结果(控制时间固定效应)

	Default			
	(1)	(2)	(3)	
	系数	系数	系数	均值处边际效应
<i>Herding</i>	-0.108*** (-2.662)	-0.102** (-2.498)	-0.109*** (-2.642)	-0.017
<i>Interest</i>	0.028*** (4.655)	0.030*** (4.868)	0.040*** (6.198)	0.003
<i>Term</i>	0.156*** (4.305)	0.154*** (4.229)	0.144*** (3.932)	0.021
<i>Ln(Amount)</i>	0.050*** (19.115)	0.050*** (19.108)	0.050*** (18.882)	0.006
<i>HR</i>	1.326*** (26.378)	1.327*** (26.299)	1.317*** (25.911)	0.163
<i>Male</i>	0.144*** (3.212)	0.145*** (3.231)	0.148*** (3.290)	0.018
<i>Age</i>	0.015*** (5.889)	0.015*** (5.982)	0.015*** (6.012)	0.002
<i>Bachelor</i>	-0.367*** (-10.304)	-0.368*** (-10.321)	-0.367*** (-10.245)	-0.041
<i>MasterOrAbove</i>	-0.829*** (-6.313)	-0.831*** (-6.334)	-0.829*** (-6.288)	-0.067
<i>Employ(3 - 5yrs)</i>	-0.097** (-2.488)	-0.097** (-2.486)	-0.090** (-2.289)	-0.012
<i>Employ(5yrs+)</i>	-0.077** (-2.026)	-0.077** (-2.030)	-0.068* (-1.794)	-0.009
<i>Income(¥5,000 - 10,000)</i>	-0.078 (-0.787)	-0.074 (-0.747)	-0.054 (-0.539)	-0.013
<i>Income(¥10,000 - 20,000)</i>	-0.088 (-0.703)	-0.083 (-0.662)	-0.051 (-0.406)	-0.015
<i>Income(¥20,000 - 50,000)</i>	0.057 (0.386)	0.068 (0.457)	0.108 (0.719)	0.001
<i>Income(¥50,000+)</i>	0.086 (0.471)	0.101 (0.550)	0.163 (0.884)	0.003
<i>Ln(Income/Debt)</i>	0.071 (1.446)	0.069 (1.400)	0.058 (1.164)	0.010
<i>House</i>	0.017 (0.502)	0.018 (0.529)	0.019 (0.538)	0.002

续表 2 羊群效应对借款是否违约的回归结果(控制时间固定效应)

	<i>Default</i>			
	(1)	(2)	(3)	
	系数	系数	系数	均值处边际效应
<i>Mortgage</i>	-0.219*** (-5.017)	-0.221*** (-5.063)	-0.223*** (-5.102)	-0.027
<i>Car</i>	-0.205*** (-5.471)	-0.203*** (-5.391)	-0.200*** (-5.307)	-0.026
<i>CarLoan</i>	0.180*** (3.012)	0.179*** (2.993)	0.170*** (2.834)	0.023
常数项	-5.080*** (-12.564)	-5.142*** (-12.475)	-5.137*** (-12.299)	
Hour-of-Day 固定效应	控制	控制	控制	
Day-of-Week 固定效应	未控制	控制	控制	
Month-of-Year 固定效应	未控制	未控制	控制	
观测数	16,167	16,167	16,167	
Pseudo R-squared	0.250	0.251	0.254	

注:括号内为 *z* 值;***表示在 1%的水平下显著,**表示在 5%的水平下显著,*表示在 10%的水平下显著。

如果群体拥有智慧,那么我们猜测投资者的羊群效应程度本身就反映了投资者群体对借款质量的判断,投资者的羊群效应程度越高,表示该借款越受投资者群体的追捧,其违约可能性应该越低。因此我们预期 *Herding* 的系数符号将为负且显著。从表 2 的结果可以看到,*Herding* 的系数依然为负且在 1%水平下显著。在同时控制了三组的固定效应之后,我们计算了各解释变量在均值处的边际效应,*Herding* 值每增大 1,能使得违约概率下降 1.7%。

以上结果说明,本文所设计的借款的 *Herding* 的度量方式含有借款信息、借款人信息所不包含的信息,*Herding* 值越高的借款其违约率越低。由于借款的羊群效应程度所反映的是群体对借款的追捧程度,因此该结果说明群体能发现那些违约概率更低的借款,并对之表现出更强的羊群行为,这说明群体总体来说是有智慧的。

3.2 倾向性得分匹配

前文的结果表明,投资者的羊群效应程度对于借款是否发生违约确实有预测作用,在一定程度上肯定了群体是有智慧的。但是以上研究还不能回答一个问题:群体的智慧有没有产生新的信息?是否有可能是群体的智慧并不包含新的信息,仅仅是群体对相似的借款表现出了相似的投资风格。为此,我们使用倾向性得分匹配的方法,筛选出发生羊群效应可能性接近、但是一组是实际发生了羊群行为、

而另一组未实际发生羊群行为的借款。这种情况下,站在投资者的视角,两组借款的发生羊群行为的概率是相近的。如果实际发生羊群行为的借款仍然有较低的违约率,那么我们就可以判断投资者的羊群行为带来了新的信息。

我们将借款依照其 *Herding* 值从大到小排序,取 *Herding* 值排在前 25%的借款作为投资者羊群效应程度高的借款,即实验组;取 *Herding* 值排在后 25%的借款作为投资者羊群效应程度低的借款,即对照组。我们利用 Probit 模型估计出每笔借款发生投资者羊群行为的概率,即倾向性得分,将倾向性得分最相近的借款进行“实验组-对照组”的匹配,我们的匹配过程遵照无放回的匹配原则,保证每笔借款都不会被重复匹配。此外,我们还对倾向性得分的差值进行了限制,所有匹配的样本需保证倾向性得分的差值在 0.005 以内。

在表 3 中,我们对实验组借款和对照组借款的借款基本特征以及借款人基本特征进行了差异性比较。两组借款的 *t* 检验结果显示:匹配后的实验组与对照组借款在基本特征上已经不存在显著的差异。也就是说,站在投资者的视角,匹配后的实验组和对照组两组借款的风格是相似的,发生羊群行为的程度和概率也是相近的。但是,如前所述,从事后的角度来看,实验组实际上发生了羊群行为,而对照组则没有发生,因此接下来我们就可以比较两组样

本违约率的差异——如果实验组仍然有较低的违约率,那么我们就可以判断羊群行为带来了新的信息。

基于以上逻辑,我们对配对后的样本的违约概率进行进一步的 t 检验,检验结果见表 3 的最后一

行。从该结果可以看到:实验组的违约率小于发生对照组的违约率,且在 1% 的统计水平上显著。这说明羊群效应确实有识别违约率差异的功能,群体的智慧也传达了关于违约率判定的新信息。

表 3 匹配后实验组与对照组借款特征对比

	实验组		对照组		实验组减去对照组	
	均值	标准差	均值	标准差	t 值	p 值
借款基本特征						
<i>Interest</i>	13.335	3.051	13.309	2.793	0.325	0.372
<i>Term</i>	9.708	7.205	9.556	6.881	0.799	0.212
<i>Ln(Amount)</i>	9.357	1.089	9.354	0.945	0.097	0.461
借款人基本特征						
<i>HR</i>	0.625	0.484	0.608	0.488	1.277	0.101
<i>Male</i>	0.871	0.336	0.865	0.341	0.637	0.262
<i>Age</i>	32.982	6.474	33.041	6.785	-0.332	0.370
<i>Bachelor</i>	0.307	0.461	0.307	0.009	0.029	0.488
<i>MasterOrAbove</i>	0.032	0.177	0.030	0.003	0.623	0.267
<i>Employ(3 - 5yrs)</i>	0.224	0.417	0.222	0.416	0.194	0.423
<i>Employ(5yrs+)</i>	0.329	0.470	0.336	0.472	-0.544	0.293
<i>Income(¥5,000 - 10,000)</i>	0.270	0.444	0.268	0.443	0.152	0.440
<i>Income(¥10,000 - 20,000)</i>	0.148	0.355	0.144	0.351	0.497	0.310
<i>Income(¥20,000 - 50,000)</i>	0.131	0.337	0.137	0.344	-0.713	0.238
<i>Income(¥50,000+)</i>	0.106	0.308	0.110	0.313	-0.521	0.301
<i>Ln(Income/Debt)</i>	1.268	1.141	1.282	1.128	-0.478	0.317
<i>House</i>	0.561	0.496	0.559	0.497	0.190	0.425
<i>Mortgage</i>	0.224	0.417	0.218	0.413	0.585	0.279
<i>Car</i>	0.421	0.494	0.422	0.494	-0.082	0.467
<i>CarLoan</i>	0.081	0.273	0.084	0.277	-0.393	0.347
借款违约情况						
<i>Default</i>	0.115	0.320	0.138	0.343	-2.359***	0.009

注:***表示在 1% 的水平下显著,**表示在 5% 的水平下显著,*表示在 10% 的水平下显著

3.3 久期分析

之前的分析中,我们仅考虑了借款是否出现违约,但我们常常更关心给定借款已经按约定正常还款若干个月的条件下,其在下个月发生违约的概率。本节中,我们使用久期分析模型来考察在借款已经正常还款若干个月的条件下,投资者的羊群效应程度对于借款在下个月违约风险的影响。每笔借款正常还款的持续时间为借款从开始还款到还清或者违约之间的时间间隔,考虑到每笔借款的期限不同,我们定义 *AnalysisTime* 为其正常还款持续时长与借

款期限的比值,并将 *AnalysisTime* 作为久期分析的被解释变量。

在引入解释变量之前,我们先进行一些初步的描述性分析。我们还画出了两组借款的风险函数,即借款在该时刻的瞬间违约率的对比图。在图 2 中,我们发现, *Herding* 值最高的 25% 的借款的风险函数值始终高于 *Herding* 值最低的 25% 的借款,且它们的 95% 置信区间完全不重合,说明在分析时间内,投资者羊群行为程度高的借款的违约风险在任意时刻都要显著地比投资者羊群行为程度低的借

款高。

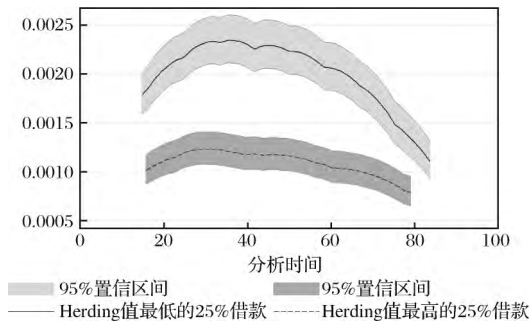


图 2 Herding 值最高 25% 与最低 25% 借款的风险函数对比

随后我们引入解释变量,并利用久期模型进行参数回归和半参数回归。在参数回归部分,由于指数回归为威布尔回归的特例,而我们的统计检验拒

绝了使用指数回归的原假设,因此本文中我们使用威布尔回归;由于参数回归的具体分布存在不确定性,因此我们也使用了半参数 Cox 模型作为补充。我们在表 4 中展示了两组回归的回归系数以及各解释变量的风险比。在控制了借款特征、借款人特征以及时间固定效应的情况下,我们发现两组回归都显示羊群程度高的借款的在下个月的违约风险会显著降低:威布尔回归的结果显示,Herding 值每增大 1,则借款在下个月的违约风险将降为原来的 83.5%左右;Cox 模型的实证结果显示,Herding 值每增大 1,则借款下个月的违约风险将降为原来的 83.4%左右,两组回归方法估计出来的违约风险降低程度非常接近。此外,在其他的控制变量中,我们发现借款人的信用等级对于借款的违约风险影响最大;如果借款人的信用等级为高风险等级,则借款在下个月的违约风险将为低风险等级的 10 倍以上。

表 4 利用久期分析模型考察羊群效应对借款违约风险的影响

	威布尔回归		Cox 模型	
	系数	风险比	系数	风险比
<i>Herding</i>	-0.180*** (-2.749)	0.835	-0.182*** (-2.766)	0.834
<i>Interest</i>	0.077*** (7.477)	1.080	0.076*** (7.384)	1.079
<i>Term</i>	0.369*** (6.236)	1.446	0.377*** (6.377)	1.458
<i>Ln(Amount)</i>	0.055*** (15.507)	1.056	0.050*** (14.173)	1.052
<i>HR</i>	2.380*** (21.762)	10.810	2.357*** (21.581)	10.559
<i>Male</i>	0.238*** (3.367)	1.269	0.234*** (3.304)	1.264
<i>Age</i>	0.019*** (5.129)	1.019	0.019*** (5.106)	1.019
<i>Bachelor</i>	-0.607*** (-10.149)	0.545	-0.604*** (-10.116)	0.546
<i>MasterOrAbove</i>	-1.269*** (-5.332)	0.281	-1.258*** (-5.287)	0.284
<i>Employ(3-5yrs)</i>	-0.152** (-2.438)	0.859	-0.151** (-2.408)	0.860
<i>Employ(5yrs+)</i>	-0.158*** (-2.649)	0.854	-0.155*** (-2.598)	0.856
<i>Income(¥5,000-10,000)</i>	-0.489*** (-3.021)	0.613	-0.508*** (-3.138)	0.602

续表 4 利用久期分析模型考察羊群效应对借款违约风险的影响

	威布尔回归		Cox 模型	
	系数	风险比	系数	风险比
<i>Income</i> (¥10,000—20,000)	-0.533*** (-2.600)	0.587	-0.558*** (-2.721)	0.572
<i>Income</i> (¥20,000—50,000)	-0.319 (-1.316)	0.727	-0.359 (-1.482)	0.698
<i>Income</i> (¥50,000+)	-0.351 (-1.187)	0.704	-0.401 (-1.354)	0.670
<i>Ln(Income/Debt)</i>	0.265*** (3.403)	1.303	0.274*** (3.521)	1.315
<i>House</i>	0.027 (0.503)	1.028	0.029 (0.536)	1.030
<i>Mortgage</i>	-0.340*** (-4.803)	0.712	-0.340*** (-4.803)	0.712
<i>Car</i>	-0.282*** (-4.703)	0.754	-0.277*** (-4.611)	0.758
<i>CarLoan</i>	0.192** (2.030)	1.211	0.187** (1.977)	1.205
常数项	-15.012*** (-21.472)	3.02e-7		
Hour-of-Day 固定效应	控制		控制	
Day-of-Week 固定效应	控制		控制	
Month-of-Year 固定效应	控制		控制	
观测数	16,167		16,167	
Wald chi2	6,089.94		28,016.54	
Prob>chi2	0.000		0.000	

3.4 稳健性检验

本节中我们对三种可能会影响本文结论的干扰因素进行排除,作为稳健性检验。

首先,我们认为最后一笔投资可能会对 *Herding* 的度量造成一定的影响:当借款投资接近尾声时,剩余的可投资金额可能会较少,从而使得最后一笔投资金额也变小,投资者投资的强度被迫减慢,这样就会影响到借款的羊群效应度量的准确性。我们去掉每笔借款的最后一笔投资,在此基础上再次计算借款的 *Herding* 数值,发现微调后的羊群效应估计值仍然具有预测违约率的能力,这说明群体智慧是稳健存在的。

此外,如果一笔借款的被投资次数过少,*Herding* 度量值也可能存在偏差,因此,我们将投资次数小于总体样本的 25%分位值,即 9 次的借款样本

删去之后重新进行回归,羊群效应估计值仍具有预测违约率的能力。

最后,由于借款是否违约是借款还款表现的一个相对粗略的衡量,我们又使用了内部收益率作为衡量投资者投资业绩的另一种方式,我们发现投资者羊群效应越明显的借款,其内部收益率也越高,这和前文的研究结论一致,即投资者投资时对于借款的羊群效应越明显,则借款的收益表现越好。

5 结语

伴随着互联网经济的发展,越来越多的投资呈现出专家化的趋势,因此群体的智慧就显得越来越重要了。本文以一个 P2P 网贷平台的投资参与数据为例,研究了群体在投资中是否有智慧。

本文首先构建了一个独特的指标来反应每个投

资借款羊群程度。基于 Probit 回归分析的结果表明,在控制了其它因素之后,羊群效应变量的系数统计上显著为负,这一结果十分稳健。并且,这一效应在经济意义上也较为显著。这说明受投资者追捧的借款其违约率更低——这在一定程度上反映了群体的智慧。

为了验证群众的智慧是传达了新的信息还是仅仅反映了群众对某些类型借款的特殊偏好,本文通过倾向性得分匹配的方法,筛选出两组发生羊群行为概率相近的借款。其中实验组真实发生了羊群行为,而对照组实际未发生羊群行为。匹配后的结果表明:实验组的违约率会更低,这说明羊群效应程度可以作为识别借款违约率的有效特征,也就是说群体的羊群行为传达了关于借款是否会违约的新信息。

此外,我们还利用了久期分析模型,考察对于已经按约定正常还款一段时间的借款而言,投资者的羊群效应程度越高,其在下个月的违约风险就越低。

在稳健性检验部分,我们使用了去除最后一笔投借款羊群效应度量值、去除了投资次数较少的借款样本、利用内部收益率作为借款的收益表现的替代度量方式,发现考虑到这些可能影响文章结论的因素之后,投资者的羊群效应程度依然传达了关于借款是否会发生违约的新信息。

本文的结论对于广大投资者具有一定的启示,即投资者在投资决策时,除了一些确定性的信息之外,还可以适当参考其他投资者的投资行为,从中获取额外的信息。当其他投资者在投资过程中投资热情逐步高涨,说明该笔借款得到了投资者的认可,则该笔借款更值得投资。反之,当其他投资者在投资过程中投资热情逐步下降,则说明该笔借款未得到投资者的认可,该笔借款的投资价值存疑。

参考文献:

- [1] Agrawal A, Catalini C, Goldfarb A. Some simple economics of crowdfunding[J]. *Innovation Policy and the Economy*, 2014, 14(1): 63-97.
- [2] Baruch S, Saar G, Zhang Xiaodi. News, influence, and the evolution of prices in financial markets[J]. SSRN Working Paper, 2014.
- [3] Budescu D V, Chen E. Identifying expertise to extract the wisdom of crowds[J]. *Management Science*, 2014, 61(2): 267-280.
- [4] Chang E C, Dong S. Idiosyncratic volatility, fundamentals, and institutional herding: Evidence from the Japa-

nese stock market[J]. *Pacific-Basin Finance Journal*, 2006, 14(2): 135-154.

- [5] Chen Hailiang, De P, Hu Y J, et al. Wisdom of crowds: The value of stock opinions transmitted through social media[J]. *The Review of Financial Studies*, 2014, 27(5): 1367-1403.
- [6] Chen Yubo, Xie Jinhong. Online consumer review: Word-of-mouth as a new element of marketing communication mix[J]. *Management Science*, 2008, 54(3): 477-491.
- [7] Christie W G, Huang R D. Following the pied piper: Do individual returns herd around the market? [J]. *Financial Analysts Journal*, 1995, 51(4): 31-37.
- [8] Cogent Research. Social media's impact on personal finance & investing[R]. 2008, Available from: <http://www.cogentresearch.com>.
- [9] Davis-Stober C P, Budescu D V, Dana J, et al. When is a crowd wise? [J]. *Decision*, 2014, 1(2): 79-82.
- [10] Grinblatt M, Titman S, Wermers R. Momentum investment strategies, portfolio performance, and herding: A study of mutual fund behavior[J]. *The American Economic Review*, 1995, 85(5): 1088-1105.
- [11] Herzog S M, Hertwig R. The wisdom of ignorant crowds: Predicting sport outcomes by mere recognition [J]. *Judgment and Decision Making*, 2011, 6(1): 58-72.
- [12] Lakonishok J, Shleifer A, Vishny R W. The impact of institutional trading on stock price[J]. *Journal of Financial Economics*, 1992, 32(1): 367-403.
- [13] Larrick R P, Soll J B. Intuitions about combining opinions: Misappreciation of the averaging principle[J]. *Management Science*, 2006, 52(1): 111-127.
- [14] Mannes A E, Larrick R P, Soll J B. The social psychology of the wisdom of crowds[M]//Krueger J I. *Frontiers of social psychology. social judgment and decision making*. New York: Psychology Press, 2012: 227-242.
- [15] Mollick E, Nanda R. Wisdom or madness? Comparing crowds with expert evaluation in funding the arts[J]. *Management Science*, 2015, 62(6): 1533-1553.
- [16] Ray R. Prediction markets and the financial "wisdom of crowds" [J]. *The Journal of Behavioral Finance*, 2006, 7(1): 2-4.
- [17] Simmons J P, Nelson L D, Galak J, et al. Intuitive biases in choice versus estimation: Implications for the wisdom of crowds[J]. *Journal of Consumer Research*, 2010, 38(1): 1-15.
- [18] Surowiecki J. *The wisdom of crowds: Why the many*

- are smarter than the few[N]. The Spectator, 2004, 295(9177):295.
- [19] Zhang Juanjuan, Liu Peng. Rational herding in micro-loan markets[J]. Management Science, 2012, 58(5): 892-912.
- [20] 苏寒,胡笑旋.基于群体智慧的复杂问题决策模式[J].中国管理科学,2012,20(S2):783-789.
- [21] 李存金,王俊鹏.重大航天工程设计方案形成的群体智慧集成机理分析——以阿波罗登月计划为例[J].中国管理科学,2013,21(S1):103-109.

Wisdom of Crowds from P2P Lending Investors

LIAO Li, XIANG Jia, WANG Zheng-wei

(PBC School of Finance, Tsinghua University, Beijing 100083, China)

Abstract: The de-professionalized trend in Internet financial investment highlights the significance of crowd views, and the wisdom of crowd is increasingly attracting researchers' attention. The objective of this study to explore the predictive power of crowd views towards default rate by analyzing the evidence from P2P lending. The transaction data from Renrendai, a main P2P lending platform in China, is employed and the herding measure index of each loan is constructed for data analysis. It is found that the default rate of loan decreases as the herding effect of investors boosts under the control of all the other characteristics of loans and borrowers using Probit regression model. Further, regarding two loans with very similar characteristics, the loan with stronger herding effect is significantly less likely to default compared with the loan with weaker herding effect using the propensity score matching method. Herding, therefore, is informative in judging the probability to default. The duration analysis also indicates that loans with higher herding effect will significantly be less likely to default in the next month conditional on the loan repaid punctually for several months. With all the conclusion above, it is concluded that the crowd has wisdom in P2P investment and the wisdom of crowd offers new information to this investment. This conclusion is also helpful during investors' investment decisions process.

Key words: wisdom of crowd; herding; P2P lending