

标点与字数会影响网络借贷吗

——来自人人贷的经验证据*

叶德珠 陈 霄

内容提要:就信息不对称程度而言,P2P网贷市场可能比传统金融市场更为严重。该市场的长期发展,取决于市场参与者如何有效利用有限的手段降低信息不对称的能力。本文以中国具有代表性的网贷平台“人人贷”2012年3月到2014年12月间的170817笔订单为研究对象,考察了网络借贷市场中借款描述的标点和字数所起到的产生增量信息的作用。实证结果发现:(1)在控制了其他因素的条件下,借款订单中借款描述的标点数每增加10个,借款成功率就会相对降低8%,而借款成本则会降低0.51%。(2)字数每增加10个,那么借款成功率就会相对提高2.3%,但是借款成本会提高0.28%。(3)标点数和字数对还款表现并未能产生显著的增量信息。上述结论在经过稳健性检验之后依然能够成立。

关键词:网络借贷 信息不对称 标点符号 借款成本

作者简介:叶德珠,暨南大学经济学院金融系教授、博士生导师,510632;

陈 霄,暨南大学经济学院金融系博士生,510632。

中图分类号:F832.4 **文献标识码:**A **文章编号:**1002-8102(2017)05-0065-15

一、引言

互联网信息技术的发展为建立新型商业模式包括金融活动创造了广泛的可能性。P2P网络借贷是一种新型的借贷模式,它允许借款人绕过金融中介机构,直接从个人投资者那里获取无抵押的借款。与传统金融相比,网络借贷具有直接性、分散性等特点(Mild等,2015)。这种模式不仅为投资者提供了一个新型理财途径,同时还能提高社会资金的利用效率(Duarte等,2012)。世界上第一家网贷平台Zopa于2005年成立于英国,短短数年,全球网贷市场就得到了长足的发展。但信息不对称问题依然是这个新兴市场所面临的基本问题,这个问题甚至在很大程度上要比传统金融市场中的更突出(Luo和Lin,2013)。因此,P2P网贷市场的长期发展,在很大程度上取决于

* 基金项目:国家自然科学基金“儒家文化与过度储蓄:微观机理、实证检验与宏观政策研究”(71473102);教育部人文社会科学基金“文化、偏好与消费”(13YJA790139)。作者感谢第十五届经济学会、互联网金融高层论坛和广东外语外贸大学金融双周论坛参会学者提供的宝贵建议,感谢审稿专家的建设性意见。文责自负。

市场参与者如何利用有效信息来减少这种信息不对称的能力。本研究主要针对中国的 P2P 网贷市场,在中国个人征信系统建设较为滞后的背景下,网贷市场中信息不对称问题更加严重,而未经认证的借款描述信息是否能够通过标点数和字数的差异来传递并产生增量信息?这一问题的答案,或许有助于全面深入了解中国的 P2P 网络借贷市场,为如何有效缓释该市场中信息不对称问题提供可能途径;同时,也能为中文文字和标点网贷市场中产生的经济后果分析提供重要的经验证据。

随着网络借贷市场规模的不断增长,近年来有关网络借贷的实证研究呈现激增的现象,关于网贷市场的研究内容也非常广泛。在这种模式中,市场主体只能使用借款人提供的信息决定是否参与经济交易,而借款人的客观定量数据在通常情况下又是不完整或者不可信的,因此,目前许多有关 P2P 网络借贷的研究,大多集中在市场参与主体的“软”“硬”信息在克服信息不对称问题的作用上(Burtch 等,2014)。在网贷市场中,“硬信息”通常是指一些能够被精确计算和量化的信用信息,如借款人的生理特征、信用等级等;而“软信息”通常是指借款人难以被量化的信用信息,如借款人的社会关系网络。借款人的“软”“硬”信息在前期的研究中已经被证实了能够对借款结果和还款表现产生影响(Barasinska 和 Schafer,2014;Liu 等,2015;Michels,2012)。

本文所关注的是“软”“硬”信息之外,借款人未经认证的借款描述信息。在借款人的订单中,借款人被提供了一个开放式的文本区域,在该区域内,借款人可以写下任何他们想要去表达的信息(Herzenstein 等,2011),如为什么需要贷款?为什么对他们的借款进行投资是可信并且值得的?本质上,这是借款人用于吸引投资者并提高借款成功率的一种方式。^①

借款人在 P2P 网络借贷中提供的借款描述具有重要作用(Herzenstein 等,2011)。在 P2P 网络借贷市场中,投资者和借款人之间的信息不对称程度更加严重,因此,投资者会关注任何有助于他们区别优质借款人的信息。前期有关 P2P 网络借贷中借款描述的研究大多集中在借款描述的内容和文字的数量方面(Dorfleitner 等,2016;Herzenstein 等,2011;Larrimore 等,2011;Lin 等,2013)。这些研究均证实投资者在借贷过程中会参照借款人书写的借款描述进行决策。但是这些研究主要使用的方法为文本分析法,如语气、语态对借款的影响。一方面,由于限定语态用词范畴有限,并且同样的单词在不同语境下的含义存在非常大的差异,因此,容易造成判断偏误(Loughran 和 McDonald,2011);另一方面,在语气和语态之外,学者们也关注过 P2P 网络借贷中借款描述的字数(Herzenstein 等,2011;Lin 等,2013),但均将其作为控制变量,同时尚未有文献关注标点数在网络借贷中的作用,这也为我们的研究提供了可拓展的空间。

标点符号在书面语言的表达中被用于平衡语句的韵律轮廓和句法结构(Bruthiaux,1993),它必须通过结合其与语言维度的关系进行理解(Ferreiro 和 Pontecorvo,1999)。研究表明,标点符号的使用能使读者对书面语言的理解产生作用(Liu, Wang 和 Jin,2010)。^②而书面语言的变化随时都在进行(Baron,2001),进入信息时代后,“网络语言”也随之出现,在“网络语言”中,标点符号无处不在,并且能够起到很大的作用(Steffes 和 Burgee,2009)。借款描述是借款人被指定在一个特定的区域内,写下任何他们想要去表达的信息(Herzenstein 等,2011)。在借款描述的表述形式中,借款人也难免会受到“网络语言”的影响,可能会忽略标点符号在书面语言中的作用,在互联网环

① 网贷平台不会对这些内容进行验证,借款人可以自由撰写任何内容以提高他们的借款成功率。

② 如:弟弟找不到爸爸妈妈,很着急;弟弟找不到爸爸,妈妈很着急。

境下随意使用脱离实际意义的标点符号^①，而这能够在一定程度上反映出借款人的个人特征，这些特征将会起到一个信号的作用传递给投资者。

在网络环境中，嵌入描述中的文字长短可以影响产品的销售和股票市场的表现(Antweiler 和 Frank, 2004; Archak 等, 2011)。产品市场中，消费者对产品的评论性文字越长，则越容易提高其他消费者对该产品的感知和认知程度，进而促成消费者对产品的购买。网络借贷市场中的借款描述也能够起到类似的帮助投资者筛选借款人的作用，借款人通过进行借款描述，给投资者提供了一个深入了解借款人的机会，对投资者在信息处理过程中的认知资源产生影响，并且成为网贷中可以被利用的定量数据。借款描述的长度可能会刺激投资者仔细浏览，提高投资者对借款人的认知度进而促使投资者有足够的动力对其进行投资。

网络借贷中的借款描述中的标点数和字数是否能够产生增量信息呢？我们将借款描述中的标点符号数和字数作为衡量借款人的一个重要指标进行考察，结果发现，中国网贷市场中借款描述的标点数和字数能够产生增量信息。在控制了其他因素的条件下，标点数每增加 10 个，借款成功率就会相对降低 8%，而借款成本则会降低 0.51%。借款订单中借款描述的字数每增加 10 个，借款成功率就会相对提高 2.3%，借款成本则会提高 0.28%。但标点数和字数对借款人的还款表现并未能产生显著的增量信息。这些结果在经过稳健性检验之后依然能够成立。

本文可能的边际贡献主要体现在：(1) 深化并丰富了 P2P 网络借贷“借款描述”作用的分析框架，P2P 网贷平台的发展积累了大量个体之间的交易数据，但是较少有文献关注借款描述中标点符号数量的作用，而本文是对其首次进行探讨的文献；同时，我们还对文字数量的作用进行了更全面的考察。(2) 本文从标点和文字长度两个维度，探讨了借款人语言特征的作用，从而形成了对中国 P2P 网贷市场中借款描述作用的立体解读。这不仅有助于探索中国 P2P 网贷市场的发展规律，同时也是对现有研究中该领域国际经验的有益补充。

二、理论分析与研究假设

心理学和语言学上有关描述内容的信息传递作用有很广泛的研究。描述中的信息是无处不在的，从传统的媒介如报纸、财报，到网络媒介如博客、微博等。近年来，越来越多的语言学者们通过挖掘描述内容中的情绪来判别描述者的态度。这种研究多应用于产品评价和金融市场中，研究表明，嵌入在描述中的文字长短可以影响和预测产品的销售和股票市场的表现(Antweiler 和 Frank, 2004; Archak 等, 2011; Tetlock 等, 2008)。而在借贷市场中，这种借款描述能够帮助投资者筛选借款人。我们想要关注的问题是，当一个借款描述被提供时，会在多大程度上影响到经济交易的发生？

作为书面语言组成部分之一的标点符号，在书面表达中具有不可或缺的作用(Hirotani 等, 2006)。一般形式下，它不具有词的形式和内涵；与文字相比，标点符号也有独特性，它只能表达出语态、停顿等现象(Baldwin, 1978)。这就决定了标点符号对文字形成活动所产生的作用，是在特定条件下传达出来的一种模糊的语义信息(Schriefers 等, 1990)。但随着语言发展以及言语表达方式的日趋丰富，标点符号的功能也在不断增强，尤其是在 20 世纪进入信息时代后，伴随着互联网的崛起，“网络语言”也随之出现，在“网络语言”中，标点符号无处不在，并且能够起到很大的作

^① 如：本人有实体店一家，只为提高额度，请大家多多支持。

用(Steffes 和 Burgee, 2009)。网络借贷同样是以互联网为载体,因此在借款描述的表述形式中,也难免会受到“网络语言”的影响。

在借款描述中以“网络语言”的方式被使用,往往会让标点符号脱离其实际的意义(例如:本人有实体店一家,只为提高额度,请大家多多支持)。借款人使用借款描述主要的作用是促进投资者对该笔借款的了解,但是这种标点符号的大量使用会提高借款的非正式性。在一种正式的借贷市场中使用非正式的表达方式,会大大降低借款描述的可读性,而文本阅读难度的提高会导致投资者因为无法理解借款人的这种表达,进而放弃阅读借款描述,从而限制甚至阻止投资者从借款描述中获得有关借款人的额外信息,最终降低投资者对借款人的信任。据此我们提出假设 1:

H1: 借款描述中的标点数越多,借款成功率越低。

非正式表达中的标点符号,能够在一定程度上反映借款人的个人特征。Li(2010)对上市公司年报的研究发现,管理层的个人特征决定会计文本披露的内容与特征。在 P2P 网络借贷中,借款描述本质上也是借款人用于印象管理的手段(Herzenstein 等, 2011)。我们认为,借款人如果在正式的借贷市场中使用非正式的表达,比如随意使用大量标点符号,那么,这种借款人可能会在自我控制能力方面存在不足,而自我控制也是认知偏差的一种表现形式(Hirshleifer, 2001; Thaler 和 Shefrin, 1977)。借款描述是促进投资者理解借款人的一种方式,借款人在借款描述中的非正式表达,会影响投资者对他们的形象判断。借款人在明知使用非正式的表达会对自身形象造成不好影响的前提下,依然去使用这种非正式表达,因此可以认为有强烈的迹象表明这种借款人存在自我控制能力不足,导致这种不足的原因很可能是认知偏差而带来的过度自信。研究发现,过度自信的 CEO 会存在乐观偏差,并高估企业未来收益(Hribar 和 Yang, 2016)。因此,这种过度自信的借款人有可能会高估自己的还款能力。同时,Chuang 和 Lee(2006)的研究发现,过度自信会导致对私人信息反应过度,而对公共信息反应不足,并且产生过激的行为。在 P2P 网络借贷中这会导致借款人忽视市场的整体借款利率,高估自身的质量水平,设置一个错误的利率。根据上述分析,我们提出假设 2 和 3:

H2: 借款描述中的标点数越多,借款成本越低。

H3: 借款描述中的标点数越多,还款表现越差。

前期的研究在很大程度上集中于产品的描述(Levy, 1981),描述的长度是最容易被量化的指标,而描述内容的长短被用来量化产品评论对消费者购买意愿的影响(Ghose 等, 2012)。在网贷领域中,Larrimore 等(2011)的研究发现,借款人的借款描述中提到的过去的借款经历大都是可以信任的,会减少借款人的不确定性并提高借款成功率,而且个人对于贷款的描述内容越多则可能的借款成功率越高。我们认为,借款描述的长度可能会对交易达成产生促进作用,借款描述为投资者提供了一个了解借款人的方式。通常情况下,借款描述的内容越长,说明投资者所能够了解到借款人的信息越多,那么投资者也能够更好地控制潜在的风险,因此会使得借款人的借款成功率提高;如果借款人提供的借款描述较少,则可能会限制投资者对这笔借款的信息了解程度,最终导致借款人融资失败。据此我们提出假设 4:

H4: 借款描述中的字数越多,借款成功率越高。

虽然借款描述能够给投资者提供信息,但是这个描述的内容既有可能是真实的,也有可能是具有误导性的。如上所述,通常情况下,网贷平台不会对这些叙述的内容进行验证,因此,借款人可以自由撰写任何内容以提高他们的借款成功率。但是,描述内容的长短,能够在一定程度上反映出描述者未来的行为(Butler, 1991)。借款描述是借款人定量信息的补充,在高风险领域,投资

者具有很高的认知加工能力(Petty和Wegener,1998),潜在的经济损失会强烈地导致认知加工,并且会对一个人在未来发生某种事件的概率做出判断(Osborne和Gilber,1992),因此借款描述为投资者理解借款人的过去和现在提供了一种积极的方式。在网络借贷中,借款人能够通过借款描述,来设计并塑造他们的自身形象,用以得到他们期望的结果。在一定程度上,描述内容更多的人,是由于他们认为他们有理由获取更多的支持,因此在交易有可能无法达成的情况下,他们依然会选择书写更多借款描述,用以说明自己是一名合格借款人。Herzenstein等(2011)也发现文字描述在一定程度上能够提高借款成功率,同时描述内容越多的借款人违约风险越小。据此我们提出假设5:

H5:借款描述中的字数越多,借款人的还款表现越好。

在P2P网络借贷中,借款描述中的内容是属于自愿性披露的。已有的文献证实,借款人会通过借款描述,来设计并塑造他们的自身形象,用以得到他们期望的结果(Herzenstein等,2011)。例如,使用更多的积极词汇或者在借款描述中表露出强烈的还款意愿以达到获取借款的目的(Ciuchta和O'Toole,2016)。借款描述中字数越多,越有可能说明借款人对于自身的“硬信息”(信用等级、收入水平和还款能力等)并不是非常有信心,因此需要大量的借款描述来进行印象管理。在一定程度上,借款人的借款描述内容越多,那么,更加有可能说明借款人可能是在采取策略性的方式,通过精心制作一个借款描述来达到提高借款成功率的目的,而这种“策略性”书写借款描述的借款人更有可能为迎合投资者对于收益的需求设置一个较高的借款利率,最终达到从投资者处获取借款的目的。据此我们提出假设6:

H6:借款描述中的字数越多,借款成本越高。

三、研究设计

(一)数据选取与网站特征描述

“人人贷”网贷平台成立于2010年5月,目前的服务已经覆盖了全国30余个省的2000多个地区,服务了几十万名客户,是行业内最具影响力的品牌之一。在人人贷网贷平台中,投资者和借款人不产生物理上的联系,因此,使用人人贷的数据可以有效评估借款描述在匿名的金融市场交易中所起到的作用。我们采用2012年3月1日到2014年12月1日人人贷网站上发布的全部借款订单作为初始样本,同时在借鉴相关文献后(廖理等,2014),对样本做了如下处理:(1)剔除信息不全的订单;(2)剔除借款人年龄小于22岁的订单;(3)剔除机构担保和实地认证的订单。最终样本为170817个,其中159681笔失败的借款,11136笔成功借款,成功的借款中包含了809笔违约的订单。需要说明的是,在成功借款中除了809笔违约的订单,其余成功的借款订单均为按时还款订单。

(二)变量设定

我们根据相关文献的研究设计(陈霄、叶德珠,2016),将平台公布的借款人的各项基本信息分为四大因素进行探讨。一是订单因素,包括借款利率、借款金额和借款期限;二是用户的信用因素,包括信用等级、房贷和车贷;三是用户个人因素,包括借款人性别、年龄、婚姻状况、收入、学历、有无房产、车产及工作时间;四是宏观因素,包括平台改版因素以及年度虚拟变量。各项指标具体说明如表1所示。^①

^① 限于篇幅,我们并没有列示详细的变量描述性统计,感兴趣的读者可向作者索取。

表 1

变量定义表

变量类型	变量	变量名	变量含义
主要变量	成功借款	<i>Success</i>	借款人发布在平台中的借款订单,当借款成功时为 1,否则为 0
	借款违约	<i>Default</i>	借款人在借款成功后,未能按时偿还借款为 1,否则为 0
	文字	<i>Word</i>	借款人发布的订单中借款描述的汉字数量
	标点	<i>Punctuation</i>	借款人发布的订单中借款描述的标点符号的数量
订单因素	借款利率	<i>Rate</i>	借款人自主决定发布订单的借款年利率,按规定,借款人利率不得超过央行同期基准利率的四倍,单位%
	借款金额	<i>Amount</i>	借款人在平台内发布该笔借款的借款金额,单位元
	借款期限	<i>Time</i>	借款人发布订单的借款期限,单位月
信用因素	信用等级	<i>Credit</i>	平台根据用户不同风险情况授予不同的等级,AA 为最高,HR 为最低,共七级,我们将 HR~AA 分别赋值 1~7
	房贷	<i>House_l</i>	借款人基本信息中有无房贷的情况,有为 1,否则为 0
	车贷	<i>Car_l</i>	借款人基本信息中有无车贷的情况,有为 1,否则为 0
个人因素	性别	<i>Gender</i>	借款人性别,男性为 1,女性为 0
	年龄	<i>Age</i>	借款人基本信息中的生理年龄
	婚姻	<i>Marry</i>	借款人若已婚则取值 1,否则为 0
	收入	<i>Income</i>	借款人月收入,1 为 1000 以下,2 为 1001~2000,3 为 2001~5000,4 为 5001~10000,5 为 10001~20000,6 为 20001~50000,7 为 50000 以上,单位元
	学历	<i>Education</i>	借款人学历,1 为高中或以下,2 为大专,3 为本科,4 为研究生及以上
	房产	<i>House</i>	若借款人有购房为 1,无购房为 0
	车产	<i>Car</i>	若借款人有购车为 1,无购车为 0
	工作时间	<i>Worktime</i>	借款人参加工作的时间,1 为 1 年或以下,2 为 1~3 年,3 为 3~5 年,4 为 5 年以上
宏观因素	改版	<i>Redesign</i>	人人贷平台自 2013 年 10 月之后进行改版,我们对改版之后的订单取值 1,否则为 0
	年份	<i>Year</i>	样本区间涉及到时间上共计 3 个年度,为控制经济波动对于借贷的影响,我们在实证分析中,按照订单所处的时间段产生 3 个虚拟变量对其进行控制

注:关于年龄变量,在数据筛选中我们已经剔除了年龄小于 22 岁的借款人。在实证分析中,为控制离群值对回归结果的影响,我们对借款人年龄进行单边缩尾处理($p=0.001$)。

(三)研究方法

二元选择模型(Binary Choice Model)是模型中被解释变量只有两种选择的模型,在我们的研究中,借款人是否借款成功以及成功借款的借款人是否违约的被解释变量均属于这种类型的变量;同时该模型能够在一定程度上克服传统模型事后预测事前事件的缺陷,在经济学中的应用非常普遍。因此我们在下文的实证分析中选择二元 Logit 模型对网贷借款成功率以及还款表现进行实证分析。按照定义,Logit 模型的标准公式如公式(1)所示,其中 X_i 为解释变量, ϵ 为随

机扰动项：

$$L: \ln\left(\frac{Y_i}{1-Y_i}\right) = \Pr(Y_i = 1) = \beta_0 + \sum_{i=1}^{i=n} \beta_i X_i + \epsilon \quad (1)$$

四、实证研究

(一)文字、标点与借款成功率

我们首先对假设 1 和 4 进行检验,探讨标点和字数是否能够在网贷市场中产生增量信息。对此,我们构建了三个模型,回归结果呈现于表 2。

表 2 文字、标点对借款成功率的影响(Logit 模型)

自变量	(1)	(2)	(3)	(4)	(5)
<i>Rate</i>	-0.1763*** (-40.02)	-0.1787*** (-40.48)	-0.00841*** (-39.64)	-0.1803*** (-40.71)	-0.00847*** (-39.88)
<i>lnAmount</i>	-0.7468*** (-63.93)	-0.7582*** (-64.56)	-0.0357*** (-61.99)	-0.7653*** (-64.93)	-0.0360*** (-62.33)
<i>Time</i>	0.0241*** (17.19)	0.0245*** (17.46)	0.00115*** (17.40)	0.0246*** (17.50)	0.00115*** (17.43)
<i>Credit</i>	1.2470*** (57.57)	1.2385*** (57.39)	0.0583*** (58.35)	1.2436*** (57.55)	0.0584*** (58.52)
<i>House_l</i>	0.1794*** (5.59)	0.1853*** (5.77)	0.00872*** (5.770)	0.1849*** (5.75)	0.00869*** (5.751)
<i>Car_l</i>	-0.0138 (-0.30)	-0.0091 (-0.20)	-0.000426 (-0.195)	-0.0113 (-0.24)	-0.000529 (-0.242)
<i>Gender</i>	0.0220 (0.67)	0.0241 (0.73)	0.00113 (0.727)	0.0250 (0.75)	0.00117 (0.754)
<i>Age_h</i>	0.0331*** (17.89)	0.0326*** (17.58)	0.00153*** (17.53)	0.0319*** (17.14)	0.00150*** (17.08)
<i>Marry</i>	0.2247*** (8.57)	0.2290*** (8.72)	0.0108*** (8.715)	0.2269*** (8.64)	0.0107*** (8.628)
<i>Income</i>	0.4090*** (40.15)	0.4011*** (39.24)	0.0189*** (38.75)	0.3993*** (38.96)	0.0188*** (38.49)
<i>Education</i>	0.2491*** (18.08)	0.2478*** (17.96)	0.0117*** (17.91)	0.2483*** (17.97)	0.0117*** (17.92)
<i>House</i>	-0.1120*** (-4.09)	-0.1142*** (-4.17)	-0.00537*** (-4.169)	-0.1101*** (-4.02)	-0.00517*** (-4.017)
<i>Car</i>	0.3119*** (11.06)	0.3100*** (10.98)	0.0146*** (10.97)	0.3122*** (11.05)	0.0147*** (11.03)

续表 2

自变量	(1)	(2)	(3)	(4)	(5)
<i>Worktime</i>	0.3588*** (29.20)	0.3624*** (29.44)	0.0171*** (29.15)	0.3622*** (29.38)	0.0170*** (29.09)
<i>Word</i>		0.0032*** (14.18)	0.00015*** (14.16)	0.0055*** (18.08)	0.000258*** (18.03)
<i>Punctuation</i>				-0.0116*** (-10.18)	-0.00055*** (-10.17)
<i>Cons</i>	1.6219*** (13.30)	1.5540*** (12.74)		1.6562*** (13.51)	
<i>Redesign</i>	Yes	Yes	Yes	Yes	Yes
<i>Year</i>	Yes	Yes	Yes	Yes	Yes
N	170817	170817	170817	170817	170817
Chi2	21131.3	21312		21442.1	
Clas	93.81	93.83		93.84	
Pseudo R ²	0.2566	0.2588		0.2604	

注:***、**和* 分别表示在 1%、5%和 10%的水平上显著,下同。括号内为 Z 统计值。

在表 2 第(1)列的估计中,我们重点考察了前期文献中提及的主要变量对借款结果的影响,第(3)列和第(5)列分别是第(2)列和第(4)列各项指标对借款成功率的边际效应。^① 结果显示,第(1)列中大部分指标与前期文献所考察的保持一致。在第(2)列中,我们加入了反映借款人特征的借款描述字数(*Word*),此时 *Word* 项系数显著为正,同时该列模型预测的正确率达到了 93.83%,比第(1)列 93.81% 提高了 0.02%。在第(4)列中,我们进一步加入借款描述的标点项(*Punctuation*),结果发现标点项显著为负,并且该列模型预测的正确率达到了 93.84%,比第(1)列提高了 0.03%,比第(2)列提高了 0.01%。

由第(2)列的估计结果来看,借款描述字数项(*Word*)的系数为 0.0032,并且在 1%的置信水平上显著。而字数项的边际效应为 0.00015,这意味着在控制了其他因素的条件下,借款订单中借款描述的字数每增加 10 个,那么借款成功率就会相对提高大约 2.3%($0.00015 \times 10 / 0.065$,其中 0.065 为借款成功率均值)。第(4)列中,我们进一步将标点数(*Punctuation*)加入,发现标点数的系数为 -0.0116,并且同样在 1%的置信水平上显著,而标点变量的边际效应为 -0.00055,这意味着在控制了其他因素的条件下,借款订单中借款描述的标点数每增加 10 个,那么借款成功率就会相对降低大约 8%($0.00055 \times 10 / 0.065$)。

一方面,借款描述中的标点符号数量是能够在一定程度上反映借款人自我控制力的指标,在正式借贷市场中使用非正式的表达,如大量使用标点符号,会降低投资者的信任感,从而降低借款成功率。另一方面,借款描述的字数越多,越能够提高投资者的信任感,因为字数越多说明其借款描述的内容越详细。这也有利于缓解借贷双方的信息不对称问题,投资者也能够更好地控制潜在

^① 在二元 Logit 模型中,由于被解释变量是离散的,因此极大似然估计让我们无法直接在平均值的意义下解释回归结果的系数。为直观地考察变量系数的经济意义,我们使用 Stata 中的 *margins* 命令来计算变量的边际效应。

风险,进而促进了借款成功率的提高。根据实证结果,假设 1 和 4 是能够成立的。我们另外还发现,在所控制的两个年度变量中,系数均显著为负值,这或许与中国宏观经济的走势存在一定关联。

(二)文字、标点与还款表现

假设 3 和 5 提出,借款人如果在正式的借贷市场中使用非正式的表达,比如大量使用标点符号,那么这种借款人的还款表现可能会更好。同时,使用字数越多的借款人还款表现也更好。对此,我们构建了三个模型,回归结果呈现于表 3。

表 3 文字、标点对借款违约率的影响(Logit 模型)

自变量	(1)	(2)	(3)	(4)	(5)
<i>Rate</i>	0.1669*** (10.67)	0.1679*** (10.70)	0.00974*** (10.74)	0.1675*** (10.67)	0.00971*** (10.71)
<i>lnAmount</i>	0.4030*** (7.09)	0.4071*** (7.14)	0.0236*** (7.166)	0.4042*** (7.06)	0.0234*** (7.090)
<i>Time</i>	-0.0180*** (-3.26)	-0.0181*** (-3.29)	-0.00105*** (-3.286)	-0.0180*** (-3.27)	-0.00105*** (-3.267)
<i>Credit</i>	-4.7347*** (-10.49)	-4.7311*** (-10.48)	-0.274*** (-10.28)	-4.7326*** (-10.48)	-0.274*** (-10.28)
<i>House_l</i>	-0.6454*** (-5.13)	-0.6461*** (-5.13)	-0.0375*** (-5.128)	-0.6470*** (-5.14)	-0.0375*** (-5.134)
<i>Car_l</i>	0.2147 (1.33)	0.2107 (1.30)	0.0122 (1.304)	0.2121 (1.31)	0.0123 (1.313)
<i>Gender</i>	0.4913*** (3.64)	0.4910*** (3.63)	0.0285*** (3.631)	0.4929*** (3.64)	0.0286*** (3.644)
<i>Age_h</i>	0.0159** (2.38)	0.0161** (2.40)	0.000934** (2.406)	0.0160** (2.39)	0.000929** (2.395)
<i>Marry</i>	-0.0935 (-0.98)	-0.0922 (-0.97)	-0.00535 (-0.967)	-0.0928 (-0.97)	-0.00538 (-0.973)
<i>Income</i>	0.1123*** (2.99)	0.1146*** (3.05)	0.00664*** (3.048)	0.1146*** (3.05)	0.00665*** (3.050)
<i>Education</i>	-0.4678*** (-8.83)	-0.4675*** (-8.83)	-0.0271*** (-8.820)	-0.4671*** (-8.82)	-0.0271*** (-8.812)
<i>House</i>	0.2434*** (2.65)	0.2429*** (2.64)	0.0141*** (2.642)	0.2430*** (2.64)	0.0141*** (2.644)
<i>Car</i>	-0.3643*** (-3.57)	-0.3632*** (-3.56)	-0.0211*** (-3.557)	-0.3649*** (-3.57)	-0.0212*** (-3.571)
<i>Worktime</i>	-0.0666 (-1.45)	-0.0678 (-1.48)	-0.00393 (-1.476)	-0.0674 (-1.47)	-0.00391 (-1.468)
<i>Word</i>		-0.0008 (-1.03)	-0.000046 (-1.031)	-0.0003 (-0.29)	-0.000019 (-0.294)

续表 3

自变量	(1)	(2)	(3)	(4)	(5)
<i>Punctuation</i>				-0.0027 (-0.58)	-0.00015 (-0.583)
<i>Cons</i>	-3.3184*** (-4.56)	-3.3175*** (-4.56)		-3.2858*** (-4.50)	
<i>Redesign</i>	Yes	Yes	Yes	Yes	Yes
<i>Year</i>	Yes	Yes	Yes	Yes	Yes
N	11136	11136	11136	11136	11136
Chi2	1276.2	1277.3		1277.6	
Clas	92.74	92.66		92.67	
Pseudo R ²	0.2200	0.2202		0.2203	

注:括号内为 Z 统计值。

在表 3 第(1)列的估计中,我们重点考察了前期文献中提及的主要变量对借款结果的影响,第(3)列和第(5)列分别是第(2)列和第(4)列各项指标对违约率的边际效应。从第(2)列和第(3)列的估计结果可知,字数的系数为-0.0008,边际效应为-0.000046,这意味着在控制了其他因素的情况下,借款订单中借款描述的字数每增加 10 个,那么借款违约率就会相对降低大约 0.63% ($0.000046 \times 10 / 0.072$, 其中 0.072 为借款违约率的平均值)。在第(4)列和第(5)列中,我们将标点数这一指标加入,发现标点的系数为-0.0027,边际效应为-0.00015,这意味着在控制了其他因素的情况下,借款订单中借款描述的标点每增加 10 个,那么借款违约率就会相对降低大约 0.21% ($0.00015 \times 10 / 0.072$)。

我们发现,字数和标点项这两个指标均未能在 10% 的置信水平上通过显著性检验。同时,在第(1)列中,模型预测的正确率为 92.74%,而第(2)列中的模型预测正确率为 92.66%,第(4)列为 92.67%,在加入了字数和标点项之后模型的预测率反而会降低。另外,第(4)列中的 Pseudo R² 较第(1)列而言,仅仅提高了 0.03%,这反映出字数和标点并没能能够在还款表现中产生显著的增量信息。标点与字数对借款成功率存在显著的影响,但对违约并没有产生显著影响,因此我们认为,事实上这些信息只起到心理上的增量信息作用,根据前景理论(Kahneman 和 Tversky, 1979),这是由于投资者在决策中的信息编辑阶段赋予了借款人信用较高的权重。对此,我们认为假设 3 和 5 并未能得到实证结果的强烈支持。

(三)文字、标点与借款成本

假设 2 和 6 提出,标点符号数量越多,借款成本越低。另外,借款描述的字数越多,借款成本越高。同样,我们构建了三个模型,回归结果呈现于表 4。^①

由第(2)列结果来看,借款描述字数变量的估计系数为 0.0042,并且在 1% 的置信水平上显著。在本文的描述统计分析中,借款利率的平均值为 14.98,这说明在控制了其他因素的情况下,借款订单中借款描述的字数每增加 10 个,就会提高借款利率大约 0.28% ($= 0.0042 \times 10 / 14.98$)。第

^① 在实证分析的这部分内容中,被解释变量不属于离散型数值,因此,我们使用的估计方法是普通的线性回归模型 OLS。

表 4 文字、标点对借款成本的影响

自变量	(1)	(2)	(3)
<i>lnAmount</i>	-0.0158** (-2.00)	-0.0254*** (-3.22)	-0.0285*** (-3.62)
<i>Time</i>	0.0321*** (34.50)	0.0323*** (34.72)	0.0321*** (34.62)
<i>Credit</i>	-0.8597*** (-39.39)	-0.8685*** (-39.83)	-0.8669*** (-39.78)
<i>House_l</i>	-0.2736*** (-10.64)	-0.2662*** (-10.36)	-0.2652*** (-10.33)
<i>Car_l</i>	-0.0253 (-0.66)	-0.0184 (-0.48)	-0.0196 (-0.51)
<i>Gender</i>	0.2488*** (10.83)	0.2517*** (10.97)	0.2548*** (11.11)
<i>Age_h</i>	-0.0006 (-0.44)	-0.0012 (-0.82)	-0.0017 (-1.21)
<i>Marry</i>	-0.1766*** (-9.77)	-0.1714*** (-9.50)	-0.1724*** (-9.55)
<i>Income</i>	0.1218*** (15.59)	0.1115*** (14.26)	0.1093*** (13.97)
<i>Education</i>	-0.2612*** (-25.64)	-0.2646*** (-26.01)	-0.2653*** (-26.09)
<i>House</i>	-0.0628*** (-3.26)	-0.0679*** (-3.53)	-0.0662*** (-3.44)
<i>Car</i>	-0.4932*** (-22.26)	-0.4989*** (-22.55)	-0.4981*** (-22.53)
<i>Worktime</i>	0.0149* (1.69)	0.0178** (2.02)	0.0182** (2.07)
<i>Word</i>		0.0042*** (21.65)	0.0059*** (24.43)
<i>Punctuation</i>			-0.0077*** (-11.75)
<i>Cons</i>	16.9084*** (208.65)	16.7508*** (206.16)	16.7911*** (206.56)
<i>Redesign</i>	Yes	Yes	Yes
<i>Year</i>	Yes	Yes	Yes
N	170817	170817	170817
F	1573.8	1512.9	1437.7
Adj-R ²	0.1284	0.1308	0.1315

注:括号内为 T 值。

(3)列的估计中我们将标点数这一指标加入,结果发现该指标系数为-0.0077,并且同样在1%的置信水平上显著,这意味着在控制了其他因素的情况下,借款订单中借款描述的标点数每增加10个,那么借款利率就会降低大约0.51% $(=0.0077 \times 10/14.98)$ 。

这些结果说明,借款描述的字数越多,借款人对“硬信息”内容越不自信,需要使用更长的借款描述来进行说明,达到自己获取融资的目的。这对投资者而言,是一种高风险的暗示,因此投资者会对这类借款要求更高的风险补偿。在中国正式的文字表达中,标点数一般而言会随着字数的增加而增加,但是在网贷市场的借款描述中,标点符号的数量一般不会受到字数长度的约束(例如:本人有实体店一家,只为提高信用额度,请大家多多支持。。。。),但这不仅会提高借款描述的非正式性质,也同样会对投资者的信任产生影响。因此,借款描述中标点符号的数量越多,借款人越有可能表现为过度自信,过分相信自己能够以低于市场的成本获取融资。网贷市场中的借款利率是借款人自行设定的,从而使标点数量与借款成本之间呈现出显著的负相关,而实证结果也支持了这一论点。因此,我们认为假设2和6是能够成立的。

(四)稳健性检验

为使上述的实证结论更加可靠,我们还进行了如下的稳健性测试:^①

第一,改变被解释变量。本文实证分析的前提在于投资者会阅读这些借款描述信息,对此,我们将投资者人数视为被解释变量,通过考察标点和字数是否对投资者人数产生显著的影响,进而判断投资者是否会阅读这些信息。我们的结果依然证实标点和字数能够对投资者人数产生显著的影响,因此可以认为本文的逻辑前提是存在的,换句话说,投资者在进行投资的过程中会参考借款人提供的借款描述信息进行决策。

第二,控制更多的宏观变量。宏观因素也是影响投资者及借款人行为的重要因素之一,在正文部分的分析中,我们控制了年度效应和平台改版效应;而在样本期间内,监管部门发布的一些有关P2P网络借贷的管理办法和风险提示我们并没有考虑。对此,我们定义了12个月度虚拟变量,使用月度虚拟变量的方法定义样本期间内政策等宏观因素变化。在控制了月度虚拟变量之后,标点和字数对借款成功率、借款违约率和借款利率的影响依然与正文部分的内容一致。

第三,内生性问题。样本中存在重复借贷的借款人,而借款人有可能存在学习效应,一方面重复借贷的借款人可能会发现标点符号数量的价值,从而在后续的借贷中自我复制,另一方面其他的借款人通过观察借款人使用标点符号的这种行为能给他们带来什么影响,从而也会模仿。换句话说,本文的结果可能受到由遗漏变量引起的内生性问题的影响。对此,本文借鉴同伴效应(Peer Effect)原理,我们认为相同年龄段—受教育程度—收入水平的借款人在书写借款描述时,对标点符号的使用有可能会影响借款人对于标点符号的使用,而相同年龄段—受教育程度—收入水平的其他借款人使用标点符号的数量却无法对该借款人的借款成功率、借款违约率和借款利率产生影响。对此,我们构建了与使用借款人年龄相仿、受教育程度相同和收入水平相同的其他借款人平均使用标点符号的数量作为工具变量,使用IVPROBIT模型对借款成功率和借款违约率进行回归,使用两阶段最小二乘法(2SLS)对借款利率进行回归。结果表明,本文的基本结论依然没有发生改变,可以认为结论是稳健的。

^① 由于篇幅,我们并没有列示详细的结果,感兴趣的读者可向作者索取。感谢匿名审稿专家在此处提出的建议。

五、结 语

P2P 网贷市场的长期发展,在很大程度上取决于市场参与者如何利用有效信息来减少这种信息不对称的能力。作为借款的唯一线索,借款者借款描述文字的内容与形式都非常重要。以往文献对文字内容进行了详尽的研究,本文则主要从文字表达形式方面进行补充,考察中国网贷市场中借款描述的标点数和字数是否能产生增量信息。

我们的研究表明,在本文数据样本区间内,标点符号数量、文字字数皆与借款成功率、借款利率呈显著相关关系,具有明显的增量信息作用,这种显著性在控制多个因素、经过多种稳健性检验之后依然能够成立。我们的实证结果表明:标点数每增加 10 个,借款成功率就会相对降低 8%,而借款成本则会降低 0.51%;借款订单中借款描述的字数增加 10 个,借款成功率就会相对提高 2.3%,但借款成本则会提高 0.28%。文字的增加会直接增加信息含量,因此有助于提高借款成功率;但标点符号没有直接的信息含量,却反而会因为网络标点符号的随意性,容易造成投资者认知超载(Sweller,1988),并使人们难以较好地进行决策制定,进而降低投资者的信任度,从而影响借款人借款订单的表现。

标点与字数对借款成功率存在显著的影响,但对违约并没有产生显著影响,因此我们认为,事实上这些信息只起到心理上的增量信息作用。根据前景理论(Kahneman 和 Tversky,1979),这是由于投资者在决策中的信息编辑阶段赋予了借款人信用较高的权重。另外,标点和字数对违约率不仅没有预测作用反而会降低模型的效果,但却能提高借款成功率的预测效果。根据行为金融学的理论,这在一定程度上说明投资者在决策过程中也可能受到了心理偏差中易记性偏差的影响,在投资中注意力往往集中在具体的借款描述的内容当中,因为具体的事情比抽象的概念容易记住,给人印象深刻。

本文的研究不仅为中国网贷市场借款描述的标点和字数能够产生的经济后果提供了经验证据,同时也证实了未经认证的借款描述能够提供增量信息。本文在实践层面的意义主要体现在两个方面:对借款人而言,在书写借款描述时应该使用正式的表达方式,既要注重借款描述的质量也要控制借款描述的长度;对于投资者而言,在 P2P 网络借贷中投资那些书写的借款描述很不规范的订单也并不会带来任何损失,但是投资那些借款描述质量更高的订单能够获得更高的投资回报。本文的政策意义在于两个方面:一是规范或者标准化借款描述中信息披露内容,让借款描述更加符合普通投资者的阅读习惯;二是保证借款人在借款描述中披露信息的真实性,避免借款描述中“空谈博弈”行为的发生。

参考文献:

1. 陈霄、叶德珠:《中国互联网金融中的性别歧视研究》,《金融评论》2016 年第 2 期。
2. 廖理、李梦然、王正位:《聪明的投资者:非完全市场化利率与风险识别——来自 P2P 网络借贷的证据》,《经济研究》2014 年第 7 期。
3. Antweiler, W. , & Frank, M. Z. , Is All That Talk Just Noise? The Information Content of Internet Stock Message Boards. *Journal of Finance* , Vol. 59, No. 3, 2004, pp. 1259—1294.
4. Archak, N. , Ghose, A. , & Ipeirotis, P. G. , Deriving the Pricing Power of Product Features by Mining Consumer Reviews. *Management Science* , Vol. 57, No. 8, 2011, pp. 1485—1509.
5. Baldwin, R. S. , Psycholinguistic Approaches to a Theory of Punctuation. *Journal of Literacy Research* , Vol. 10, No. 4,

1978, pp. 363–375.

6. Barasinska, N. , & Schafer, D. , Is Crowdfunding Different? Evidence on the Relation between Gender and Funding Success from a German Peer-to-Peer Lending Platform. *German Economic Review*, Vol. 15, No. 4, 2014, pp. 436–452.

7. Baron, N. S. , Commas and Canaries: The Role of Punctuation in Speech and Writing. *Language Sciences*, Vol. 23, No. 1, 2001, pp. 15–67.

8. Bruthiaux, P. , Knowing When to Stop: Investigating the Nature of Punctuation. *Language & Communication*, Vol. 13, No. 1, 1993, pp. 27–43.

9. Burtch, G. , Ghose, A. , & Wattal, S. , Cultural Differences and Geography as Determinants of Online Prosocial Lending. *Mis Quarterly*, Vol. 38, No. 3, 2014, pp. 773–794.

10. Butler, J. K. , Toward Understanding and Measuring Conditions of Trust: Evolution of a Conditions of Trust Inventory. *Journal of Management: Official Journal of the Southern Management Association*, Vol. 17, No. 3, 1991, pp. 643–663.

11. Chuang, W. I. , & Lee, B. S. , An Empirical Evaluation of the Overconfidence Hypothesis. *Journal of Banking & Finance*, Vol. 30, No. 9, 2006, pp. 2489–2515.

12. Ciuchta, M. P. , & O’Toole, J. , Looks and Linguistics: Impression Formation in Online Exchange Marketplaces. *Journal of Social Psychology*, Vol. 156, No. 6, 2016, pp. 648–663.

13. Dorfleitner, G. , Priberny, C. , Schuster, S. , Stoiber, J. , Weber, M. , de Castro, I. , & Kammler, J. , Description-text Related Soft Information in Peer-to-Peer Lending—Evidence from Two Leading European Platforms. *Journal of Banking & Finance*, Vol. 64, No. 3, 2016, pp. 169–187.

14. Duarte, J. , Siegel, S. , & Young, L. , Trust and Credit: The Role of Appearance in Peer-to-Peer Lending. *Review of Financial Studies*, Vol. 25, No. 8, 2012, pp. 2455–2483.

15. Ferreira, E. & Pontecorvo, C. , Managing the Written Text: The Beginning of Punctuation in Children’s Writing. *Learning & Instruction*, Vol. 9, No. 6, 1999, pp. 543–564.

16. Ghose, A. , Ipeirotis, P. G. , & Li, B. B. , Designing Ranking Systems for Hotels on Travel Search Engines by Mining User-Generated and Crowdsourced Content. *Marketing Science*, Vol. 31, No. 3, 2012, pp. 493–520.

17. 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. *Journal of Marketing Research*, Vol. 48, No. SPL, 2011, pp. 138–149.

18. Hirotnani, M. , Frazier, L. , & Rayner, K. , Punctuation and Intonation Effects on Clause and Sentence Wrap-up: Evidence from Eye Movements. *Journal of Memory and Language*, Vol. 54, No. 3, 2006, pp. 425–443.

19. Hirshleifer, D. , Investor Psychology and Asset Pricing. *Journal of Finance*, Vol. 56, No. 4, 2001, pp. 1533–1597.

20. Hribar, P. , & Yang, H. , CEO Overconfidence and Management Forecasting. *Contemporary Accounting Research*, Vol. 33, No. 1, 2016, pp. 204–227.

21. Kahneman, D. , & Tversky, A. , Prospect Theory: An Analysis of Decision under Risk. *Econometrica*, Vol. 47, No. 2, 1979, pp. 140–170.

22. Larrimore, L. , Jiang, L. , Larrimore, J. , Markowitz, D. , & Gorski, S. , Peer to Peer Lending: The Relationship between Language Features, Trustworthiness, and Persuasion Success. *Journal of Applied Communication Research*, Vol. 39, No. 1, 2011, pp. 19–37.

23. Levy, S. J. , Interpreting Consumer Mythology: A Structural Approach to Consumer Behavior. *Journal of Marketing Research*, Vol. 45, No. 2, 1981, pp. 49–62.

24. Li, F. , The Information Content of Forward-Looking Statements in Corporate Filings—A Naive Bayesian Machine Learning Approach. *Journal of Accounting Research*, Vol. 48, No. 5, 2010, pp. 1049–1102.

25. Lin, M. F. , Prabhala, N. R. , & Viswanathan, S. , Judging Borrowers by the Company They Keep: Friendship Networks and Information Asymmetry in Online Peer-to-Peer Lending. *Management Science*, Vol. 59, No. 1, 2013, pp. 17–35.

26. Liu, D. , Brass, D. J. , Lu, Y. , & Chen, D. Y. , Friendships in Online Peer-to-Peer Lending: Pipes, Prisms, and Relational Herding. *Mis Quarterly*, Vol. 39, No. 3, 2015, pp. 729–742.

27. Liu, B. L. , Wang, Z. N. , & Jin, Z. , The Effects of Punctuations in Chinese Sentence Comprehension: An ERP Study. *Journal of Neurolinguistics*, Vol. 23, No. 1, 2010, pp. 66–80.

28. Loughran, T., & McDonald, B., When Is a Liability Not a Liability? Textual Analysis, Dictionaries, and 10-Ks. *Journal of Finance*, Vol. 66, No. 1, 2011, pp. 35–65.
29. Luo, B. J., & Lin, Z. X., A Decision Tree Model for Herd Behavior and Empirical Evidence from the Online P2P Lending Market. *Information Systems and E-Business Management*, Vol. 11, No. 1, 2013, pp. 141–160.
30. Michels, J., Do Unverifiable Disclosures Matter? Evidence from Peer-to-Peer Lending. *Accounting Review*, Vol. 87, No. 4, 2012, pp. 1385–1413.
31. Mild, A., Waitz, M., & Wockl, J., How Low Can You Go? —Overcoming the Inability of Lenders to Set Proper Interest Rates on Unsecured Peer-to-Peer Lending Markets. *Journal of Business Research*, Vol. 68, No. 6, 2015, pp. 1291–1305.
32. Osborne, R. E., & Gilber, D. T., The Preoccupational Hazards of Social Life. *Journal of Personality & Social Psychology*, Vol. 62, No. 2, 1992, pp. 219–228.
33. Petty, R. E., & Wegener, D. T., Matching versus Mismatching Attitude Functions: Implications for Scrutiny of Persuasive Messages. *Personality and Social Psychology Bulletin*, Vol. 24, No. 3, 1998, pp. 227–240.
34. Schriefers, H., Meyer, A. S., & Levelt, W. J., Exploring the Time Course of Lexical Access in Language Production: Picture-Word Interference Studies. *Journal of Memory & Language*, Vol. 29, No. 1, 1990, pp. 86–102.
35. Steffes, E. M., & Burgee, L. E., Social Ties and Online Word of Mouth. *Internet Research*, Vol. 19, No. 1, 2009, pp. 42–59.
36. Sweller, J., Cognitive Load during Problem Solving: Effects on Learning. *Cognitive Science*, Vol. 12, No. 2, 1988, pp. 257–285.
37. Tetlock, P. C., Saar-Tsechansky, M., & Macskassy, S., More than Words: Quantifying Language to Measure Firms' Fundamentals. *Journal of Finance*, Vol. 63, No. 3, 2008, pp. 1437–1467.
38. Thaler, R. H., & Shefrin, H. M., An Economic Theory of Self-Control. *Journal of Political Economy*, Vol. 89, No. 2, 1977, pp. 392–406.
39. Yum, H., Lee, B., & Chae, M., From the Wisdom of Crowds to My Own Judgment in Microfinance Through Online Peer-to-Peer Lending Platforms. *Electronic Commerce Research and Applications*, Vol. 11, No. 5, 2012, pp. 469–483.

Is Punctuation and Word useful ——Evidence from P2P Lending

YE Dezhu (Jinan University, 510632)

CHEN Xiao (Jinan University, 510632)

Abstract: In terms of degree of information asymmetry, P2P network credit market may be more severe than traditional financial markets. With China's representative network credit platform "Renrendai" during March 2012 to December 2014, 170817 orders as the research object, this paper investigates the network in the lending market borrowing the punctuation and word of description on the role of incremental information. The results find that: (1) after controlling for other factors, to borrow loan described in the order number of punctuation, every ten increase borrowing success rate will be relatively reduced by 8%, while borrowing costs will be reduced by 0.51%; (2) the number of words, every ten increase the success rate of borrowing will relatively increase by 2.3%, but the cost of borrowing will increase by 0.28%; (3) number of punctuation and word for repayment performance and produced no significant incremental information. This conclusion after robustness test can still set up.

Keywords: P2P Lending, Information Asymmetry, Punctuation, Borrowing Costs

JEL: G10, G11, G14

责任编辑:康 邑