

地理距离、信息不对称与借款人违约风险

张 笑 胡金焱

摘要: P2P网络借贷主要借助互联网技术进行线上借贷,一方面降低了获取借款人信息的成本,另一方面也加重了借贷双方的信息不对称,部分平台通过建立线下网点来解决这一问题。通过使用人人贷2010年10月至2016年12月的交易数据,运用Probit模型,将借款人所在地到线下网点的距离作为平台信息获取成本的代理变量,实证分析了地理距离对违约风险的影响,得出以下结论:借款人距离网点越远,违约概率越大,地理距离增加了平台监测和缓解道德风险的成本;借款金额不同,地理距离对借款人的违约风险的影响不同;平台改版后借款利率可以反映地理距离产生的违约风险。

关键词: P2P借贷; 违约风险; 地理距离; 信息不对称

DOI: 10.19836/j.cnki.37-1100/c.2020.01.014

一、引言

近年来,P2P网络借贷平台发展迅速,网络借贷在给人们带来更多投资机会的同时,风险问题也日益凸显。一方面平台借助互联网技术能够以较低的成本获取信息,另一方面网络中存在大量的虚假信息,使得信息可靠性大打折扣。目前,国家要求P2P平台作为信息中介,而不是信用中介,如果平台不能获取更多有效的信息或者不能很好地识别信息的真伪,就无法实现信息中介的职能。为了解决信息不对称的问题,部分平台在线下设立网点^①方便获取借款人的有效信息以及贷后追踪^②。理论上贷款人与借款人之间的地理距离越远,其信息的收集和监测成本就越高,进而会对违约概率产生影响。在网络借贷市场中,越来越多的平台采取线上线下相结合的方式进行的借贷,地理距离是否也会对借款人的违约风险产生影响,目前尚没有文献进行分析。

已有大量文献研究地理距离在借贷市场中的作用,主要从两个角度展开。一是研究地理距离对借款成功率的影响,Agarwal和Hauswald^③利用小公司向大型银行申请贷款的数据,研究了地理距离对银行获取和使用私人信息的影响,发现借款人离银行越远,越不利于收集软信息,距离削弱了银行收集私人信息和开拓本地市场的能力,距离近的公司更容易获得贷款。Flögel研究了地理距离对区域银行和大型银行贷款的影响,发现区域银行在本地的网点分布更广,能以较短的地理距离放贷,增加了

收稿日期: 2018-10-03

基金项目: 国家自然科学基金面上项目“网络借贷市场风险识别、传染与防控研究:基于复杂网络理论的视角”(71873079),国家自然科学基金重点项目“民间金融风险:变迁、区域差异与治理研究”(71333009),国家自然科学基金青年项目“民间金融及其风险研究:基于儒家文化与地方政府的视角”(71703080)

作者简介: 张笑,青岛大学经济学院,讲师(青岛266071;sdlwzhangxiao@163.com);胡金焱,青岛大学,教授、博士生导师(青岛266071;hwx@sdu.edu.cn)。

^① 目前,一些P2P平台如人人贷、翼龙贷等都纷纷在线下铺设网点。其中人人贷截至2018年已在近100个城市建立了300个线下网点,而翼龙贷已在全国200多个城市设立运营中心,各级营业网点1000余家。

^② 《网络借贷信息中介机构业务活动管理暂行办法》第十六条规定:网络借贷信息中介机构在互联网、固定电话、移动电话等电子渠道以外的物理场所只能进行信用信息采集、核实、贷后跟踪、抵质押管理等风险管理及网络借贷有关监管规定明确的部分必要经营环节。

^③ Agarwal S., Hauswald R., “Distance and Private Information in Lending”, *Review of Financial Studies*, 2010, 23(7), pp. 2757-2788.

获取软信息的机会,从而更好的对借款人进行筛选^①。Hollander 和 Verriest^② 研究了距离限制是如何影响银行贷款合同的设计,发现当企业向偏远贷款人寻求贷款时,贷款合同往往更具限制性。Degryse 和 Ongena^③ 研究了地理距离对贷款条件的影响,发现贷款利率随着企业与贷款银行之间距离的减小而减小。Pedrosa 和 Do^④ 分析了家庭住址与小额信贷机构办事处之间的距离对信贷市场的影响,研究发现距离越远,贷款条件越严格,贷款利率越高。而 Carling 和 Lundberg^⑤ 利用瑞典银行在 1994 到 2000 年间发放的企业贷款数据研究发现,地理距离不会对借款成功率产生影响,Rauterkus 和 Munchus^⑥ 也得出了相同的结论。二是研究地理距离对借款人违约率的影响,Deyoung et al.^⑦ 认为距离会干扰信息收集,从而对企业贷款违约的概率产生影响。Chen^⑧ 利用台湾地区 2005 至 2011 年间的数据库,探讨企业总部所在地的地理特征是否会影响企业的信用风险,结果显示地理距离越远违约概率越高,农村企业比城市企业具有更高的信用风险。Presbitero 和 Rabellotti^⑨ 将借款人与小额贷款公司之间的实际距离作为代理成本的代理变量,研究了其对违约风险的影响,发现小额信贷市场存在道德风险,代理成本随着地理距离的增加而增加。Milani^⑩ 研究了借贷双方地理距离对意大利公司贷款违约率的影响,发现距离阻碍了信息的收集和监测,并与违约率成正比。林丽琼^⑪ 利用 240 个民间借贷纠纷案件的数据,研究了地理距离与违约率之间的关系,发现地理距离越远,借款人违约风险越高。

关于网络借贷平台中信息不对称问题的研究大都集中在理论层面,实证研究较少。部分学者认为与银行等传统金融机构相比,互联网金融在信息获取上存在优势,能较好地解决信息不对称问题。Collier 和 Hampshire^⑫ 使用 P2P 平台“Prosper”的数据研究发现,将个人信誉嵌入到社区信誉中后,能够激励投资者选择高质量的借款人,可以很好地解决逆向选择和道德风险问题。赵岳、谭之博通过建立理论模型,研究了银行信贷中的逆向选择问题,在引入电子商务平台后,借款人的违约成本增大,

-
- ① Flögel F., “Distance and Modern Banks’ Lending to SMEs: Ethnographic Insights from a Comparison of Regional and Large Banks in Germany”, *Journal of Economic Geography*, 2018, 18(1), pp. 3-23.
- ② Hollander S., Verriest A., “Bridging the Gap: The Design of Bank Loan Contracts and Distance”, *Journal of Financial Economics*, 2016, 119(2), pp. 399-419.
- ③ Degryse H., Ongena S., “Distance, Lending Relationships, and Competition”, *Journal of Finance*, 2005, 60(1), pp. 231-266.
- ④ Pedrosa J., Do Q T., *How Does Geographic Distance Affect Credit Market Access in Niger ?*, Policy Research Working Paper, 2008.
- ⑤ Carling K., Lundberg S., “Asymmetric Information and Distance: An Empirical Assessment of Geographical Credit Rationing”, *Journal of Economics and Business*, 2005, 57(1), pp. 39-59.
- ⑥ Munchus G., Rauterkus A., “Geographical Location: Does Distance Matter or What is the Value Status of Soft Information?”, *Journal of Small Business and Enterprise Development*, 2014, 21(1), pp. 87-99.
- ⑦ Deyoung R., Glennon D., Nigro P., “Borrower-lender Distance, Credit Scoring, and Loan Performance: Evidence from Informational-opaque small Business Borrowers”, *Journal of Financial Intermediation*, 2008, 17(1), pp. 113-143.
- ⑧ Chen T K., “Does Geography Matter in a Geographically Small and Culturally Homogeneous Country? Firm Location and Corporate Credit Risk”, *International Review of Economics and Finance*, 2016, 44, pp. 323-348.
- ⑨ Presbitero A F., Rabellotti R., “Geographical Distance and Moral Hazard in Microcredit: Evidence from Colombia”, *Journal of International Development*, 2014, 26(1), pp. 91-108.
- ⑩ Milani C., “Borrower-lender Distance and Loan Default Rates: Macro Evidence from the Italian Local Markets”, *Journal of Economics and Business*, 2014, 71, pp. 1-21.
- ⑪ 林丽琼:《地理距离、关系与民间借贷违约风险——基于 240 个法院纠纷案件调查数据的分析》,《亚太经济》2017 年第 2 期。
- ⑫ Collier B C., Hampshire R., “Sending Mixed Signals: Multilevel Reputation Effects in Peer-to-peer Lending Markets”, *ACM Conference on Computer Supported Cooperative Work*. ACM, 2010, pp. 197-206.

使得低风险的借款人即使没有抵押品也可以获得贷款,电子商务平台可以解决中小企业融资难的问题^①。刘征驰、赖明勇通过引入“软信息约束”和“虚拟抵押品”后,发现P2P平台可以满足没有抵押品的中小借款人的融资需求^②。但是,也有学者认为P2P平台中依然存在着信息不对称问题。Freedman和Jin^③利用P2P平台“Prosper”的数据进行了研究,他们发现虽然P2P平台公开了部分借款人的信用历史,但贷款人面临严重的信息不对称问题,会受到逆向选择的影响,但是贷款人可以通过干中学(Learning By Doing)来减轻市场中的信息不对称问题。谈超、王冀宁和孙本芝^④通过理论分析则认为,P2P平台并不能很好的解决信息不对称问题,平台中存在着逆向选择和道德风险问题,但是可以通过信号传递和引入信用担保来解决。

现有文献对于P2P平台能否解决信息不对称问题存在争议。由于网络借贷大多是纯信用借贷,没有抵押品,部分学者认为P2P平台在获取信息的成本上有优势,且借款人存在声誉约束,但是本文认为这些因素的作用有限。尽管P2P平台网上审核信息的成本很低,但借款人可能存在提供虚假信息和隐瞒重要信息的行为,信息的可靠性很难保证,平台在有效信息获取上并不一定就有优势。所以越来越多的平台建立线下网点,用于信息收集和贷后追踪,以解决信息不对称问题。

本文有以下创新之处:一是目前的研究较少从实证的角度去检验P2P平台中是否存在信息不对称问题,主要是因为信息不对称程度难以进行度量,本文将借款人到平台线下网点的距离作为平台信息收集和监测成本的代理变量来进行检验;二是本文研究发现借款人距离网点越远,违约概率越高,地理距离增加了平台监测和缓解道德风险的成本,为平台防范信用风险、建立合理的贷后审查机制提供支持;三是本文发现当借款利率由借款人设定时,利率不能充分反映借款人的违约风险,而当借款利率改为平台设定后,利率可以反映地理距离带来的违约风险。

本文的结构如下:第二部分为理论分析和研究假设;第三部分为数据来源和变量描述;第四部分为实证分析;第五部分是结论。

二、理论分析和研究假设

在银行借贷市场中,地理距离对借款人违约率的作用主要是通过影响平台获取信息的成本。在信息不透明的信贷市场中,地理距离会对银行获取和使用私人信息产生影响,距离削弱了贷款人收集借款人软信息的能力,而地理距离越远,贷款人在放贷之前收集信息的成本就越高,信息不对称程度就越高,这是事前的信息不对称。贷款人与借款人之间的地理距离越远,会增加贷后监测的成本,越无法监测借款人的贷后行为,这是事后的信息不对称。在传统的民间借贷市场中,地缘和人缘是贷款人考虑的重要因素,这是因为在一定的地域、人际关系范围内,民间借贷有其信息优势,这种优势使得贷款人在放贷之前能够以低成本获取借款人的信息,同时在贷后也以较低的成本进行管理。

P2P网络借贷具有一定的特殊性,以人人贷为例,信用认证标是在线上进行信息审核,线下会通过网点进行贷后管理。这种事前线上审核、事后线下管理的方式使得地理距离会影响平台的贷后信息搜集成本。在市场交易中,当事人签约之后,由于存在信息不对称问题导致委托人不能直接观测代理人行为或因观测成本较高时,代理人行为的变化使委托人利益受到损害。代理人(借款者)的事后行为将直接影响委托人(出借人)的利益,但是在网络借贷市场中,一个借款人往往对应着多个出借

① 赵岳、谭之博:《电子商务、银行信贷与中小企业融资——一个基于信息经济学的理论模型》,《经济研究》2012年第6期。

② 刘征驰、赖明勇:《虚拟抵押品、软信息约束与P2P互联网金融》,《中国软科学》2015年第1期。

③ Freedman S., Jin G Z., *Learning by Doing with Asymmetric Information: Evidence from Prosper. Com.*, NBER Working Papers, 2011.

④ 谈超、王冀宁、孙本芝:《P2P网络借贷平台中的逆向选择和道德风险研究》,《金融经济学研究》2014年第5期。

人,出借人很难对借款人的行为进行有效管理,而平台作为信息中介,贷后追踪是其应尽的责任。如果不能做好贷后审查工作,借款人就可能利用这种不对称信息做出对出借人不利的行为选择,道德风险就此产生。

防范事后信息不对称的有效方法就是贷后追踪,但这需要极高的成本,P2P网络借贷主要借助互联网进行线上借贷,大部分平台都没有线下网点,且网络借贷跨区域、广覆盖的特征使得借款人分布在全国各地,贷后追踪很难实现。即使平台拥有线下网点,也只是在少数城市,不可能像银行等金融机构那样广泛。因此距离网点城市越远,平台的信息获取成本就会越高。例如,两个完全相同的借款人A和B,其中A所在地有网点,B距离网点城市较远,平台对B贷后跟踪时就需要支付更高的交通费和人工费,当这个成本高于平台的收益时(手续费和管理费),平台就不会对B贷后追踪,由于“天高皇帝远”,借款人B更容易发生道德风险。此外,即使两个借款人都在同一城市,借款金额不同也会影响平台是否贷后追踪,当借款金额较低时,即使借款人距离网点很近,由于成本较高平台选择监测则无法获得收益,此时平台不会进行贷后管理;而当借款金额较高时,即使借款人距离网点较远,平台监测依然可以获得收益,此时平台会进行贷后管理。目前,国家要求P2P平台遵守“小额分散”的原则,而小额借贷可能导致平台获取信息和贷后追踪的相对成本较高,信息不对称问题就越严重。

利率作为资金的价格,在一定程度上反映了借款人的信用风险水平,借款人的违约风险越高,出借人要求补偿的风险溢价就越高,利率水平就越高。在市场化利率下,当利率达到均衡时可以反映借款人的所有风险,即借款人所在地距离越远,其违约风险越高,则借款利率越高。此外,由于存在监督成本,距离远的借款人可能需要支付一个更高的利率。本文基于上述分析,本文提出以下假设:

假设1:借款人距离线下网点越远,其违约概率越高。

假设2:借款人距离线下网点越远,其借款利率越高。

三、数据来源和变量描述

本文数据来自P2P网络借贷平台人人贷,从2010年10月19日至2016年12月31日^①所有已还清、已违约的数据。由于平台中的标的包括了实地认证、机构担保等形式,而这些标的相当于为借款人提供了担保,因此本文对样本进行处理,剔除实地认证标、机构担保标和智能理财标,仅使用信用认证标作为研究样本。在剔除缺失值和异常值之后,本文最终的样本为25744个标的,其中已还清的笔数为21627,违约的笔数为4117个。

(一)被解释变量

本文的被解释变量为借款人是否违约,人人贷中显示的借款状态为借款失败、已还清、还款中、严重逾期和已垫付。其中当借款人逾期天数大于30天时,借款人的状态为严重逾期,已垫付是指对于一些逾期时间过久的借款,平台可能会垫付。本文违约的数据是指借款状态为严重逾期或者已垫付的标的,赋值为1;没有违约的数据是指已还清的标的,赋值为0。

(二)关键解释变量

人人贷采取线上和线下相结合的方式,线下与人人友信合作设立网点。本文以借款人工作所在地到本省网点城市的距离作为平台信息获取成本的代理变量,部分省份在多个城市有网点^②,我们使用最近的距离,同时,新疆、西藏和宁夏均没有网点。我们认为省会城市交通更为便利,其信息获取成本也应小于本省其他城市,所以这三个自治区使用到省会城市的距离。Chen使用铁路距离和机场距

^① 由于人人贷在2017年1月以后就没有信用认证标,所以本文的样本只能到2016年12月31日。

^② 以下省份拥有多个网点:辽宁(沈阳和大连),山东(济南、烟台和青岛),江苏(南京和苏州),福建(福州和厦门),广东(广州和深圳),贵州(贵阳和遵义)。这些省份均选取借款人所在城市到最近的网点城市的距离。

离作为两城市间的实际距离,由于部分城市没有机场,所以本文使用两城市间的铁路距离^①(千公里)。若借款人距离网点越远,平台信息获取成本越高,就越难观测到借款人的贷后行为,信息不对称问题就越严重。

(三)控制变量

本文从4个角度去考虑控制变量:借款信息、借款人特征、借款人资产以及借款人信用。借款信息包括借款利率、借款金额和借款期限等3个变量,借款人特征包括年龄、婚姻状况和受教育程度等3个变量。具体来说,借款人未婚、离异和丧偶赋值为1,已婚赋值为0;按照借款人受教育程度的高低依次赋值,高中或以下赋值1、大专赋值2、本科赋值3、研究生或以上赋值4。借款人资产包括收入、有无房产和有无车产等3个变量,根据人人贷的数据将收入分为7个等级,收入为1000元以下赋值1,1001-2000元赋值2,2000-5000元赋值3,5000-10000元赋值4,10000-20000元赋值5,20000-50000元赋值6,50000元以上赋值7。借款人有房产赋值为1、没有赋值为0、有车产赋值为1、没有赋值为0。借款人的信用状况为借款人的信用等级,平台根据借款人提交的材料全面审核,并根据审核结果给出对应的信用评级,信用评级从高到低依次为AA、A、B、C、D、E、HR。本文将信用评级为HR的赋值为1、E赋值2、D赋值3、C赋值4、B赋值5、A赋值6、AA赋值7。此外,人人贷在2014年12月进行了较大幅度改版,为避免平台改版带来的影响,本文设置时间虚拟变量,若借款人在2015年1月1日以前贷款的赋值为0,2015年1月1日以后贷款的赋值为1。表1为各变量的变量说明。

表1 变量说明

变量名称	变量说明
是否违约	违约赋值1,没有违约赋值0
到网点的距离	借款人工作所在城市距本省网点城市的铁路距离(千公里)
借款利率	借款人的借贷利率(%)
借款金额	借款人的借款金额的对数
借款期限	还款的期限(月)
年龄	借款人的年龄(岁)
婚姻状况	借款人为单身(未婚、离异和丧偶)赋值1,已婚赋值0
受教育程度	借款人受教育程度为高中或以下赋值1,大专赋值2,本科赋值3,研究生或以上赋值4
收入	借款人收入为1000元以下赋值1,1001-2000元赋值2,2000-5000元赋值3,5000-10000元赋值4,10000-20000元赋值5,20000-50000元赋值6,50000元以上赋值7
有无房产	借款人有房产的赋值1,没有房产赋值0
有无车产	借款人有车产的赋值1,没有车产赋值0
信用等级	借款人的信用评级为HR赋值1,E赋值2,D赋值3,C赋值4,B赋值5,A赋值6,AA赋值7
时间	借款人在2015年1月1日以前贷款的赋值为0,2015年1月1日以后贷款的赋值为1

四、实证分析

首先需要检验借款人到网点的距离是否会影响违约概率,本文中的被解释变量为0,1变量,且数据类型为横截面数据,本文使用Probit模型进行研究,模型的回归形式为:

^① 各城市间距离的数据来自于 www. huochepiao. net.

$$\ln\left(\frac{p}{1-p}\right) = \alpha_0 + \alpha_1 Distance_i + \sum \alpha_n Control_i + \varepsilon_i \quad (1)$$

$$p \equiv P(y = 1 | X)$$

$\equiv P(\text{是否违约} | \text{到网点距离、借款信息、借款人特征、借款人资产和借款人信用})$

(一) 基准回归结果及解释

表 2 为地理距离对借款人违约概率的影响,回归(1)为单独回归的实证结果,显示借款人到网点的距离与借款人违约概率显著正相关,说明借款人距网点越远,其违约概率越高。回归(2)为控制其他变量的情况下,地理距离对违约率的影响,与原结果一致。根据理论分析,借款人到网点的距离与平台信息获取成本正相关,距离越远,平台监测时需要支付的费用越高。而平台的信息获取成本越高,平台越不会对借款人进行贷后追踪,由于“天高皇帝远”,借款人与平台之间的信息不对称问题就越严重,其违约率也就越高。因此,假设 1 是正确的。

表 2 地理距离对借款人违约风险的影响(基于 Probit 模型的回归)

变量名称	被解释变量:是否违约	
	回归(1)	回归(2)
到网点的距离	0.2511*** (5.1144)	0.1698*** (2.8320)
借款利率		0.0105* (1.8041)
借款金额		0.0995*** (5.4923)
借款期限		0.0584*** (37.0443)
受教育程度		-0.2583*** (-17.3765)
年龄		0.0183*** (9.7661)
婚姻状况		0.0492* (1.8749)
收入		0.0758*** (6.8522)
有无房产		-0.0276 (-1.1056)
有无车产		-0.1524*** (-5.6003)
信用等级		-0.9074*** (-41.9151)
时间		控制
观测值	25744	25744

注:***、**和*分别表示变量在1%、5%和10%的水平上显著,括号内表示 z-Statistic。以下各表含义相同。

(二) 不同借款金额下地理距离对违约率的影响

根据前文的理论分析,地理距离越远,平台的信息获取成本就越高。但平台是否选择监测不仅取决于绝对成本,还取决于相对成本。比如有两个借款人 a 和 b 在同一座城市, a 的借款金额为 L_a , b 的借款金额为 L_b ,且 $L_a > L_b$,其他条件均相同。平台网点到该城市的信息获取成本为 C ,同时平台收取费率为 f 的手续费。只有当 $L_b f - C > 0$ 时,平台才会对两个借款人进行监测,如果 $L_b f - C < 0$

且 $L_a f - C > 0$, 平台只会对借款人 a 监测, 如果 $L_a f - C < 0$, 平台均不会监测。不同的借款金额其相对信息获取成本也不同, 借款金额越小, 相对信息获取成本越大, 平台越不会进行贷后追踪, 信息不对称问题就越严重。

本文按照借款金额在 20000 元以下、20000-50000 元以及 50000 元以上分组回归。回归(3)、(4)、(5)为不同金额下, 借款人到网点的距离对违约概率的影响。表 3 结果显示借款金额在 20000 元以下以及 20000-50000 元时, 借款人距离网点越远, 违约概率越高, 而借款金额在 50000 元以上的不显著。这与理论分析是一致的, 借款金额越小, 其相对信息获取成本越大, 平台越不会选择监测, 越容易造成信息不对称。当借款金额较大时, 平台的收益较大, 且若借款人违约会给借款人带来较大的损失, 所以平台才有动力进行监测。

目前, 国家要求 P2P 平台坚持“小额分散”的经营策略, 防范信贷集中风险, 但如果不能解决市场中的信息不对称问题, 小额借贷也会提高平台贷后监测的相对成本, 使得借款人违约风险上升, 不利于平台自身的发展。

表 3 不同借款金额下地理距离对违约率的影响(基于 Probit 模型的回归)

变量名称	被解释变量: 是否违约		
	<20000 回归(3)	20000-50000 回归(4)	>50000 回归(5)
到网点的距离	0.1616** (2.1584)	0.2171* (1.8164)	0.2042 (0.8269)
借款利率	0.0085 (1.2400)	0.0262 (1.4707)	0.4534*** (5.7986)
借款金额	0.2357*** (6.6500)	-0.2640*** (-2.7827)	0.0059 (0.2179)
借款期限	0.0530*** (24.8069)	0.0679*** (21.6353)	0.0770*** (12.6966)
受教育程度	-0.2565*** (-12.8702)	-0.3137*** (-11.0899)	-0.1405*** (-3.0615)
年龄	0.0206*** (7.9780)	0.0193*** (5.5426)	0.0102** (2.0574)
婚姻状况	0.0715** (2.1636)	0.0036 (0.0687)	0.1388 (1.3072)
收入	0.0455*** (2.9982)	0.1655*** (7.6584)	0.1601*** (4.1768)
有无房产	-0.0362 (-1.1087)	-0.0530 (-1.1106)	-0.0223 (-0.2784)
有无车产	-0.1912*** (-5.1910)	-0.1217** (-2.4455)	0.0604 (0.7020)
信用等级	-1.0487*** (-28.8608)	-0.8414*** (-22.4474)	-0.7735*** (-18.5686)
时间	控制	控制	控制
观测值	17601	6063	2080

(三)改版前后地理距离对违约率的影响

至此, 本文研究发现 P2P 平台中借款人距离网点越远, 违约概率越高。人人贷在 2014 年 12 月底对平台进行了大幅度改版, 平台重新调整了信用标借贷的规则, 使得信用借贷的审核标准更为严格, 大幅度地增加了 AA 和 A 信用等级的人数, 利率由借款人自己设定改为平台设定。本文考虑平台能

不能通过线上改版来解决信息不对称问题。

由于样本是截面数据,本文只能逐年回归。将样本按照借款人的借款时间划分为两个时间段:2015年以前(2010年10月19日到2014年12月31日)和2015年以后(2015年1月1日到2016年12月31日),从而研究在平台改版前后借款人到网点的距离与违约概率的关系。

表4为回归结果,2015年以前和2015年以后的样本均为借款人到网点距离越远违约概率越高。P2P平台作为信息中介,向出借人提供有效信息,解决借贷双方信息不对称的问题,撮合两者交易。信息不对称的存在使得出借人的利益受到损失,目前来看,平台仅通过线上改版无法彻底解决这一问题。

表4 改版前后地理距离对违约率的影响(基于Probit模型的回归)

变量名称	2015年以前	2015年以后
	回归(6)	回归(7)
到网点的距离	0.1134* (1.7008)	0.2235** (2.3702)
借款利率	0.0242*** (4.0446)	-0.4532*** (-10.3979)
借款金额	0.1264*** (5.8761)	0.0450 (1.2715)
借款期限	0.0457*** (24.7036)	0.1365*** (22.1236)
受教育程度	-0.2508*** (-12.7914)	-0.3025*** (-12.6008)
年龄	0.0121*** (4.9656)	0.0306*** (10.0252)
婚姻状况	0.0226 (0.6445)	0.0732* (1.7771)
收入	0.0755*** (5.4087)	0.0952*** (5.0234)
有无房产	-0.0217 (-0.6661)	-0.0602 (-1.4871)
有无车产	-0.1603*** (-4.4909)	-0.1262*** (-2.8850)
信用等级	-0.7921*** (-27.2148)	-1.1610*** (-34.5379)
观测值	16200	9544

(四)借款利率与地理距离的关系

在借贷市场中,利率表示信贷资金的价格,借款人的风险越高,就需要支付一个更高的风险溢价,即利率越高,那么信息不对称的风险是否反映在利率之中了呢?本文进一步研究借款人到网点的距离对借款利率的影响,回归模型如(2)所示,其中 R 表示借款利率。

$$R_i = \beta_0 + \beta_1 Distance_i + \sum \beta_n Control_i + \mu_i \quad (2)$$

由于借款利率不是二值选择变量,本文使用OLS进行回归,表5为回归结果。回归(8)的结果显示,借款人到网点的距离越远,其借款利率反而越低,说明利率并没有充分反映借款人潜在的违约风险,假设2错误。

本文考虑到2014年底平台改版前后的利率定价机制有较大变化,改版前利率是由借款人自己设

定,改版以后利率是由平台设定,因此进一步将样本按照借款人的借款时间分组。回归(9)和回归(10)分别为改版前和改版后的回归结果,可以看出在平台改版以前,借款人到网点的距离越远,其借款利率越低,说明利率不能反映地理距离带来的违约风险。由于借款利率是由借款人自己设定的,在这种非完全市场化利率机制下,有相当高比例的违约风险未被反映在利率当中。在平台改版以后,借款利率是由平台设定的,此时借款人所在地距离网点越远,借款利率越高,借款人的违约风险能够反映在利率中,说明平台在利率设定上考虑了借款人所在地到网点的距离。

表 5 地理距离对借款利率的影响(基于 OLS 的回归)

变量名称	被解释变量:借款利率		
	回归(8)	2015 年以前 回归(9)	2015 年以后 回归(10)
到网点的距离	-0.4211*** (-5.7235)	-0.5391** (-4.5779)	0.0542** (2.2799)
借款金额	-0.1793*** (-10.0794)	-0.1897*** (-7.7612)	0.0492*** (5.7850)
借款期限	0.0762*** (37.7554)	0.0505*** (17.0882)	0.1303*** (171.0498)
受教育程度	-0.2321*** (-13.5514)	-0.2932*** (-11.1288)	-0.0460*** (-7.7817)
年龄	0.0384*** (16.8195)	0.0496*** (14.4313)	0.0016** (2.0215)
婚姻状况	0.1133*** (3.5104)	0.1904*** (3.7532)	-0.0031 (-0.2925)
收入	0.0434*** (3.4325)	0.0077 (0.4165)	-0.0314*** (-6.5765)
有无房产	-0.1389*** (-4.5793)	-0.1649*** (-3.5341)	-0.0361*** (-3.4852)
有无车产	-0.0025 (-0.0793)	-0.0233 (-0.4862)	-0.0096 (-0.8977)
信用等级	-0.1455*** (-15.3419)	-0.1566*** (-12.4124)	-0.2620*** (-48.2856)
时间	控制	控制	控制
观测值	25744	16200	9544

(五)稳健性检验

尽管本文的关键解释变量借款人到网点的距离(借款人工作所在地到本省网点城市的距离)为外生变量,但可能存在以下三种情况使得实证结果出现偏误。

(1)本文使用借款人到网点的距离作为衡量平台信息获取成本的代理变量,因为距离越远平台信息获取成本越高,但是距离越远平台上门催债的成本也越高,而催债成本可能会影响违约概率。比如 A 城市有网点,B 城市没有网点,若平台不上门催债,A 与 B 的违约概率一样,由于 A 城市有网点其上门催债的成本较低,在借款金额相同时,平台可能只对 A 城市的贷款进行催债,而 B 城市的成本较高平台不选择催债,最终使得 A 城市的违约率低于 B 城市。在这种情况下,距离网点越远违约概率越高,不是由于信息不对称引起的,而是催债成本不同导致的。考虑两种情况:一种是借款人先逾期(逾期小于 30 天的),借款人再违约(本文定义逾期超过 30 天或平台已垫付的是违约),最后平台上门催债;另一种是借款人先逾期,平台上门催债,借款人选择违约。如果是第一种情况催债成本不会影响

借款人的违约概率,因为借款人违约发生在催债之前,此时借款状态已经确定了,不管平台能否追回贷款,借款人都是违约。如果是第二种情况,上门催债发生在借款人是否违约之前,若平台追回贷款,则会对借款状态产生影响。所以平台是否上门催债可能会对违约概率产生影响。由于不知道人人贷是借款人逾期后立刻上门催债,还是借款人违约后上门催债,如果是前者则会对本文的实证结果产生影响。因此催债成本会影响平台是否上门催债,而是否催债对违约率可能有影响,但可以确定的是借款人是否逾期一定发生在平台催债之前,所以催债不会对逾期率产生影响。本文把被解释变量替换为借款人是否逾期,若借款人发生过逾期赋值为 1,没有发生过逾期赋值为 0。回归(11)为实证结果,借款人到网点的距离越远其逾期率越高,与原结果一致。

(2)本文使用的借款人与平台之间的距离为两个城市间的铁路距离,对于部分距离较短城市来说,汽车是最方便的出行方式,为了使回归结果更稳健,我们用两个城市间的行车距离作为平台信息获取成本的代理变量,所用数据来自百度地图。回归(12)为实证结果,借款人工作所在地到网点城市的公路距离越远,违约概率越高,与原回归结果一致。

(3)2016 年银监会出台的《网络借贷信息中介机构业务活动管理暂行办法》中明确规定同一自然人在同一网络借贷信息中介机构平台的借款余额上限不超过人民币 20 万元,为了确保回归结果的稳健性,本文剔除了借款金额在 20 万以上的样本。回归(13)为实证结果,借款人距离网点越远,违约概率越高,与原回归结果一致。稳健性检验的实证结果如表 6 所示。

表 6 稳健性检验的实证结果(基于 Probit 模型的回归)

变量名称	被解释变量:是否逾期	被解释变量:是否违约	
	回归(11)	回归(12)	回归(13)
到网点距离	0.2320*** (5.0007)	0.1753*** (2.7005)	0.1738*** (2.8942)
借款利率	0.0196*** (4.8174)	0.0104* (1.7945)	0.0093 (1.5910)
借款金额	0.1445*** (11.9712)	0.0994*** (5.4853)	0.0857*** (4.5866)
借款期限	0.0404*** (29.3075)	0.0584*** (37.0520)	0.0586*** (36.9472)
受教育程度	-0.1075*** (-9.9178)	-0.2587*** (-17.4128)	-0.2622*** (-17.5432)
年龄	0.0022 (1.4986)	0.0183*** (9.7889)	0.0181*** (9.6103)
婚姻状况	0.0250 (1.2303)	0.0490* (1.8648)	0.0462* (1.7556)
收入	0.0404*** (4.9376)	0.0755*** (6.8325)	0.0773*** (6.9669)
有无房产	-0.0404** (-2.1178)	-0.0271 (-1.0845)	-0.0291 (-1.1613)
有无车产	-0.0769*** (-3.8492)	-0.1523*** (-5.5950)	-0.1535*** (-5.6258)
信用等级	-0.4187*** (-53.3230)	-0.9074*** (-41.9111)	-0.9151*** (-41.4036)
时间	控制	控制	控制
观测值	25744	25744	25477

五、结论和建议

本文将借款人工作所在地到线下网点的距离作为平台信息获取成本的代理变量,使用人人贷2010年10月至2016年12月的交易数据,运用Probit模型,研究了地理距离对借款人违约率的影响,得出如下结论:(1)借款人到网点的距离越远,违约概率越高。借款人到网点的距离与平台的信息获取成本正相关,距离越远,平台监测时需要支付的费用越高。信息获取成本越高,平台越无法观测到借款人的行为,借款人更容易造成事后的信息不对称。(2)只有借款金额较小时,借款人离平台网点越远,其违约概率越高。不同的借款金额其相对信息获取成本也不同,借款金额越小,相对信息获取成本越大,信息不对称问题越严重。(3)尽管平台通过各种手段提高对借款人的线上审核力度,但目前平台还无法彻底解决信息不对称问题。(4)平台改版前借款利率不能充分反映借款人的违约风险,但在平台改版后利率可以反映借款人违约风险。最后本文考虑其他可能会对实证结果产生影响的因素,通过稳健性检验说明原结果是可靠的。根据本文所得结论,提出如下建议:

第一,发展线下网点,加强贷后跟踪。解决道德风险问题的有效方法就是贷后追踪,平台一方面可以扩大线下网点,另一方面可以与当地的金融机构(如小额贷款公司等)合作,降低贷后信息获取成本,定期对借款人的财务状况、资金使用情况等信息跟踪调查。

第二,对于平台中的信用标的,投资人一定要小心谨慎、仔细审查,尽量选择那些提供了较多信息,且进行了认证的借款人,同时也应注意借款人所在地是否有平台的线下网点,从而可以降低信息不对称所带来的风险。

第三,建立信息共享机制。由于各个平台的信息是相互独立的,形成了信息孤岛的问题。通过信息共享可以避免借款人在多个平台借贷的行为,同时对于违约客户建立黑名单制度,可以提高借款人的违约成本,有效防范借款人的信用违约风险。

Geographical Distance, Information Asymmetry, and Borrower's Default Risk

Zhang Xiao Hu Jinyan

(School of Economics, Qingdao University, Qingdao 266071, P. R. China)

Abstract: P2P network lending mainly uses Internet technology for online lending. On the one hand, it reduces the cost of information acquisition. On the other hand, it also causes information asymmetry. Some platforms solve this problem by setting up the offline network. By using peer-to-peer lending's trading data from October 2010 to December 2016 and Probit Model, taking the distance from the location of the borrower's location and the offline network as the proxy variable of the platform information monitoring cost, this paper empirically analyzes the impact of geographic distance on the default risk. The study found that the farther the borrower was from the network, the greater the probability of default, and the geographical distance increases the cost of platform monitoring and mitigating moral hazard; the impact of geographical distance on the borrower's default risk carries with the amount of loan; after the platform is revised, the loan interest rate can reflect the default risk caused by geographical distance.

Keywords: Peer-to-peer lending; Default risk; Geographical distance; Information asymmetry

[责任编辑:邵世友]