

信贷市场中的学历价值及其区域异质性研究

刘琼^{1,2}, 池仁勇³

(1. 浙江工业大学 管理学院, 浙江 杭州 310023; 2. 浙江工业大学 经济学院, 浙江 杭州 310023;
3. 浙江工业大学 中国中小企业研究院, 浙江 杭州 310023)

摘要: 网络借贷平台的兴起, 为研究借款者学历在信贷市场的实际价值及影响机制提供了新的机会窗口。依托 A 网贷平台的公开业务数据发现: 信贷市场的信用配给存在学历价值区分效应, 拥有高学历者在获取贷款及更低借款成本上具有优势; 学历价值的具体效用存在区域异质性, 相同学历背景的借款人在东部地区较中西部能够获得更多融资优势。研究结果表明: 学历作为易被观察的显性标签, 可以在信贷市场风险识别中发挥更大作用; 教育价值评估需充分考虑区域经济社会发展水平, 关注区域异质性。

关键词: 信贷市场; 学历价值; 区域异质性

中图分类号: F830 **文献标志码:** A **文章编号:** 1000-2154(2021)06-0038-14

DOI: 10.14134/j.cnki.cn33-1336/f.2021.06.004

Value of Academic Qualifications and Regional Heterogeneity in the Credit Market

LIU Qiong^{1,2}, CHI Renyong³

(1. School of Management, Zhejiang University of Technology, Hangzhou 310023, China;

2. School of Economics, Zhejiang University of Technology, Hangzhou 310023, China;

3. China Institute for Small and Medium Enterprises, Zhejiang University of Technology, Hangzhou 310023, China)

Abstract: The rise of online lending platforms has provided a new window of opportunity for studying the actual value of borrowers' academic qualifications in the credit market and the mechanism of influence. Relying on the public business data of the On-line Loaning Platform A, which is widely used in China, it is found that the credit rationing in the credit market has the effect of distinguishing the value of academic qualifications, and those with a high degree of education have an advantage in obtaining loans and lower borrowing costs. The specific utility of the value of academic qualifications exists with regional differences. Borrowers with the same educational background can obtain more financing advantages in the eastern regions than in the central and western regions. The research results show that: academic qualifications can play a greater role in the identification of risks in the credit market as a visible label that is easy to be observed. In the evaluation of educational value, we should fully consider the level of regional economic and social development and pay attention to regional heterogeneity.

Key words: credit market; education; regional heterogeneity

一、引言

学历作为标识人才类别的显性标签, 一直被作为衡量人才综合素质的重要尺度。特别是在信贷领域,

收稿日期: 2021-02-07

基金项目: 国家社会科学重大项目“大数据背景下中小微企业征信体系的理论与方法”(17ZDA088); 浙江省软科学项目“产业变革背景下人才陷阱的形成机理及高校角色优化对策研究——浙江的理论与实践”(16JJDGH046)

作者简介: 刘琼, 女, 讲师, 博士研究生, 主要从事高等教育、创新政策研究; 池仁勇, 男, 教授, 博士生导师, 管理学博士, 主要从事中小企业发展及创新管理研究。

受信方的学历水平一直是授信方评估额度的核心要素^[1-2]。但目前学历价值在信贷额度的映射大多基于常识和经验判断,其背后科学机理的长期缺位是部分高学历信贷坏账产生的重要原因^[3]。因此,学历对信贷市场的影响机制正成为学界和业界广泛关注的热点问题之一。同时,由于中国不同地区经济发展不平衡、人才政策导向不同、资源禀赋存在地区差异等问题,不同地区信贷市场中的学历价值效应是否发挥作用?学历价值如何影响信贷市场?这种影响背后的内在作用机理与理论逻辑是什么?这些问题都有待进一步深入分析。

之前学者对学历价值的研究大多集中在求职、议价等劳动力市场回报^[4],以及个人收入、幸福指数^[5]等方面,较少从市场信贷视角分析学历价值的重要性及其影响效应。特别是由于银行系统中个人信息的保密性原则,无法为分析信贷市场中学历价值重要性和作用导向的实证研究提供有效数据支撑,客观上使得学历背景提高信贷市场中信用配给效率的作用路径研究遭遇了瓶颈。近年来,互联网技术的快速发展带动了网络信贷平台的崛起,信贷平台下大量包含学历信息在内的借款信息能够被有效识别,为信贷市场中学历价值研究提供绝好的“机会窗口”。已有研究利用网络信贷数据从不同角度分析学历在借贷市场中的价值体现,认为高学历背景的借款者能够获得投资者的青睐^[3,6-7]。但是现有研究仍然存在不足,一方面较少有文献从整合视角将借贷成本、借贷成功率以及违约风险这三方面的价值表征纳入统一研究框架中,从而无法为学历价值在信贷市场中的作用研究提供一个整体画像;另一方面更为重要的是,现有研究并未就学历信息所展现的风险逻辑和价值逻辑进一步展开区域异质性分析。然而现实中,中国地缘辽阔、省份众多,各地区之间教育水平投入的差异、经济发展状况存在的差异性,往往会导致学历价值在信贷市场中的作用受到“扭曲”^[8-9]。其一,不同地区资源禀赋存在差异性,东部发达地区无论是金融基础设施还是金融配套能力都远远超过西部欠发达地区,能够为高学历价值借款人提供更多融资额度;其二,人才政策的招致效应不同,发达地区的人才政策力度更大,高学历价值人才能够获得更多政策上的金融倾斜;其三,人才心理建设导向不同,发达地区的高学历人才市场竞争激烈、社会与家庭期望值高,在信贷市场具有更强烈的“价值诉求”;其四,人才集聚呈现规模效应,相对于西部欠发达地区,东部发达地区人才集聚效应显著,相同学历下个人价值总体上优于西部地区人才,在信贷市场上获得融资资源也会更多。因此,基于全国东、中、西部不同地区层面探讨学历价值对信贷市场影响的区域异质性具有很强的现实与理论意义。

基于此,本文将在教育回报与信贷配给等相关理论基础,研究信贷市场中学历信息背后的低信用风险逻辑与价值主张逻辑对提高信贷配给效率的内在机理,并重点分析不同地区金融发展水平对学历价值的信贷效率提升效应的扭曲程度。此外,通过使用国内应用较为广泛的A网络信贷平台上的个人借贷数据对此作用机制进行实证分析,以期规避一直以来学历价值在信贷市场中重要性研究的“学而不术”问题。

本文可能的贡献主要在以下两个方面:一方面本文丰富了教育价值回报研究,特别是扩展了教育价值在信贷市场中的研究。以往对教育价值的实证研究多停留在教育对个人收入^[10-13]、缓解歧视等方面^[3,7],较少从信贷市场的角度去探究借款者教育价值的意义。本文借助网络信贷市场中大量借款者学历的微观信息,肯定了信贷市场中教育价值的存在。另一方面本文拓展了地区间教育价值异质性研究的相关层面^[8-9]。地区间教育投入水平的差异会导致地区间教育回报率不同^[14],但较少考虑到教育价值对地区金融发展水平的影响,本文从地区金融发展这一视角,为地区教育价值异质性的研究提供了一定的有益补充。

二、影响机制与研究假设

教育的本质是培养并提升受教育者个人劳动技能、价值观、思想品德等综合素质^[15],学历是受教育程度的主要标识,具有权威性、可观性、级差性等特征,因此过去经验研究广泛认为受信方学历的提升能够让其在信贷市场具有更多优势。

1. 学历价值对借款成功率及成本的影响机制。一方面,受教育者学历价值越大,其违约信用风险越小,借款成功率越高。具体表现在:第一,受教育者的学历程度越高,其掌握的劳动技能知识也会随之增多^[16],

从而能够在就业市场中就任更高的职位,获得更多的劳动收入,进而具有更强的偿还能力,其违约的可能性也会较小。第二,良好的教育不仅能让受教育者具有追求更多资产回报的能力,更为重要的是在教育熏陶下培养了良好的思想品德、正确的是非价值观以及信用操守等^[17],因此具有高学历价值的借款者一般都会信守诺言,按时履约,违约风险较小。第三,学历价值高的借款者一般从事重要的工作岗位,可能具有广泛的社会影响力,会更加注重自身的声誉价值^[18],而违约受到的声誉惩罚机制会让其更加积极完成履约义务^[19]。第四,银行等金融机构资金借贷的信用配给大部分是基于市场主导逻辑展开^[20],当借款人具有更低违约风险、更高偿还能力以及更低违约倾向时,其获得借款的成功率以及借款额度均会得到提升^[21]。因此,高学历价值背后的低信用风险价值逻辑能够帮助借款人提高借款成功率。

另一方面,借款人学历价值越高,其议价能力越高,借款成本越低。具体表现在:第一,根据教育回报相关理论研究,借款人要想拥有更高学历水平,就需要在财力、脑力等方面付出更多^[18],继而对未来的教育回报预期也就会越高,从而对降低借贷成本价值的诉求意愿也会更强。第二,高学历价值借款人具有更高的偿还能力,在借贷市场具有一定优势^[22],因此在与银行等金融机构协商时拥有更强的议价能力,能够获得更低的借贷利率^[23]。因此,高学历背后的价值主张逻辑使得借款人能够获得更低的借款成本。基于此,本文提出以下2个假设:

假设1a:借款者学历价值越高,借款成功率越高。

假设1b:借款者学历价值越高,违约风险越小,借款成本越小。

2. 学历价值对信贷市场的地区异质性影响。中国省份较多、分布较广,各地区之间教育水平的投入以及经济发展状况等存在很大差异^[24],地区发展的不均衡导致地区金融发展水平以及人才学历价值也出现了不对称现象^[25]。这种不对称现象在一定程度上会扭曲学历价值在借贷市场中的促进作用,具体表现为:第一,基于“资本—技能互补”假说,发达地区的金融发展程度会提高教育回报率^[9],也就意味着地区金融市场越发达,受信方能够获得更高的收入回报,因此,相同学历背景的信方在金融市场发达地区具有更高的偿债能力,违约风险更低。第二,发达地区的资源禀赋更强,拥有更加完善的金融管理体系,受信方能够获得更为广泛的融资渠道^[26],从而具有更为宽裕的资金操作空间来提升自身的资金借贷冗余,继而会提升其履约能力^[27],在信贷市场上获得更多借贷资金的信用配给。第三,发达地区具有更强的人才集聚效应,人才竞争力度更大,通过“人才漏斗机制”使得人才进入门槛提高,因此相同学历背景的东部人才价值会优于西部地区,从而能够获得更好的信用评级和议价能力,使其能够拥有更多的借款优势。第四,发达地区政策引导力度更大,配套设施更丰富,能够为高学历借款人背书,这种积极的“信号效应”使其拥有超额信用配给,因此在借贷市场中可以更大程度上发挥其学历价值。第五,由于发达地区存在人才竞争效应以及经济压力的倒逼机制,一方面,相同学历背景的借款人在发达地区想要获得更高的职位来实现自身价值,从而在借贷市场中对自身价值“主动诉求”的意愿更加强烈。另一方面,在发达地区获得机会的难度大于欠发达地区,因此相同学历背景的借款人在东部地区所付出的劳动要远大于欠发达地区,致使借款人内心为了弥补这种付出差异,便会提高其心理预期,从而在信贷市场中“被动倒逼”地提高了其信用借贷需求,在发达地区供给充足的情况下,借款人能够获得更多融资。基于此,本文提出以下假设:

假设2:地区发达程度越高,信贷市场上学历价值越能发挥作用。

三、研究设计

(一) 变量选择与数据来源

本文选取A网络借贷平台上发布的有效借款订单作为实证样本。由于该平台的借贷条件不受地区限制,客户分布涵盖了我国34个省级行政区域,而且对借款者不设门槛,只要通过实名认证均可申请贷款业务,因此其数据具有相对较高的代表性和随机性。此外,考虑到新疆、西藏、香港、台湾以及澳门地区的样本相对不足,为了确保实证结果的准确性,本文剔除了来自以上地区的数据以及一些信息不全的订单,最终共获得65534笔有效借款订单,能满足本研究的基本数据需求。其中各变量的构成与说明如下:

本文选择的主要被解释变量包括三个:该笔借款申请最终是否成功(success),借款成功后借款者是否如约还款(default)以及借款者设定的借款利率水平(rate)。核心的解释变量为借款者的学历水平(edu)。该平台上公布的信息除本文主要研究的学历状况以外,还有借款金额、借款期限以及受信方的收入状况、年龄、婚姻状况等信息,而这些信息都会对借贷者行为产生影响。出于实证模型的稳健性考虑,将这些相关变量作为控制变量放入实证模型。其中最主要的是受信方的收入状况,当控制了与借款者收入相关的变量后,借款者的学历则更多地表示为与借款者诚信相关的特征。具体变量定义与赋值处理详见表1。

表1 变量的定义及赋值处理

变量类型	变量名称	赋值和处理方法	
被解释变量	是否成功借款(success)	成功借款取1,反之取0	
	是否违约(default)	违约取1,无取0	
	借款利率(rate)	取实际值	
解释变量	学历(edu)	高中及以下取1,大专取2 本科取3,研究生及以上取4	
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
		收入状况(income)	1000元以下取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
		是否有房贷(house_loan)	有取1,无取0
		是否有车(car)	有取1,无取0
		是否有车贷(car_loan)	有取1,无取0
		工作时间(work)	1年以下取值为1,1-3年取值为2,3-5年取值为3,5年以上取值为4
	标的特征	借款金额(amount)	取自然对数,缩尾1%
		借款利率(rate)	取实际值,年利率
		借款期限(time)	取实际值,以月为单位
信用特征	信用等级(level)	信用等级从高到低七个等级依次取1~7	

(二) 描述性统计分析

在建模分析前,本文先对数据进行初步的描述性统计分析。通过从平台上获得的65534笔有效借款订单数据(如表2),研究发现随着借款者学历的提高,其借款成功率不断增大,这说明授信方在选择借款订单时更青睐学历高的借款者。从违约率来看,随着借款者学历的提高,其违约率也随之降低,说明学历这一指标的确能够在一定程度上反映受信方的违约风险,学历越高的借款者违约风险越低。从借款利率来看,整体表现为学历越高、利率水平越低,说明放贷者在设定借款利率时会充分考虑到借款人的学历价值,从而设定一个较低水平的借款利率。表3为主要变量的描述性统计结果。

表2 样本数据的学历差异与借贷行为关系分析

学历		高中或以下	大专	本科	硕士及以上	总计
成功	已还清	5573	6210	3682	277	15742
	违约	161	116	50	1	328
失败		22395	17878	8566	625	49464
总计		28129	24204	12298	903	65534
借款成功率		20.38%	26.14%	30.35%	30.79%	24.52%
违约率		2.81%	1.83%	1.34%	0.36%	2.04%
利率		15.88%	15.70%	15.11%	14.58%	15.65%

表3 主要变量的描述性统计

Variable	Obs	Mean	Std.	Min	Max
success	65534	0.2452	0.4302	0	1
default	16070	0.0204	0.1414	0	1
rate	65534	15.6539	3.5530	6	24
edu	65534	1.7860	0.7913	1	4
level	65534	2.0215	1.9515	1	7
time	65534	12.9827	9.2146	1	36
amount	65534	9.8250	1.4479	8.0064	13.1224
wage	65534	34.1487	7.5588	23	58
marriage	65534	0.5332	0.4989	0	1
income	65534	3.9527	1.3270	1	7
house	65534	0.3786	0.4850	0	1
house_loan	65534	0.1050	0.3066	0	1
car	65534	0.2823	0.4501	0	1
car_loan	65534	0.0436	0.2043	0	1
work	65515	2.4222	1.0068	1	4

(三) 实证模型

本文选择 Logistcs 回归模型和 OLS 回归模型分别检验区域差异和学历因素对借款成功率、违约率和借款利率的影响。具体模型如下：

$$Pr(Success)_i = \beta_0 + \beta_1 Edu_i + \sum \beta Control_i + \lambda_i + I_i + \xi_i \quad (1)$$

$$Pr(Default)_i = \beta_0 + \beta_1 Edu_i + \sum \beta Control_i + \lambda_i + I_i + \delta_i \quad (2)$$

$$Rate_i = \beta_0 + \beta_1 Edu_i + \sum \beta Control_i + \lambda_i + I_i + \mu_i \quad (3)$$

其中,模型设计的变量含义以及赋值情况如表 1 描述,control 表示的是一系列控制变量, β_0 表示截距项, ξ 、 δ 和 μ 为误差项。由于个体在所属行业和地区层面可能存在很大的异质性,因此本文在模型中加入地区固定效应 λ_i 和行业固定效应 I_i 。

如果式(1)中系数 β_1 显著为正,则说明学历越高的借款者更易获得借款。如果式(2)中系数 β_1 显著为负,则说明学历能够很好地反映受信方的违约风险,学历越高的受信方其违约风险越低。如果式(3)中系数 β_1 显著为负,则说明授信方在设定利率时会考虑借款人的学历价值,从而为学历越高的借款者设定一个较低水平的借款利率。

四、实证结果

(一) 基准回归

本文先将获得的样本进行回归分析,主要研究借款人的学历差异是否会显著影响借款成功率、违约率以及借款利率。具体回归结果如表4所示。

表4第(1)、(2)列是对借款成功率的分析结果,当不加入控制变量时,学历(edu)的系数为0.45且显著为正,这说明本科学历对借款成功率的影响较为显著。进一步计算其边际效应,发现借款者学历每增加一个单位,借款成功率增加了56.36%。加入一系列控制变量之后,学历(edu)的系数降为0.12但依然显著为正,说明在考虑了借款者收入状况、婚姻情况等控制变量后,借款者学历每增加一个单位,借款成功率增加12.84%,即在其他条件相同的情况下,投资者更青睐拥有较高学历的借款者,高学历的借款者相比低学历借款者更易筹得借款,显示出授信方对借款者学历价值的“认可”。

表4第(5)、(6)列是对借款利率的分析结果,学历(edu)的系数显著为负,借款者学历每增加一个单位,借款利率降低88.16% ($e^{-0.1260}$)。说明放贷者在设定利率时会考虑到借款人的学历价值,表现为对学历价值的“认可”。综合表4(1)、(2)、(5)、(6)列数据可得,假设1a成立。

表4第(1)、(2)列和第(3)、(4)列有一个共同现象,加入控制变量后,学历系数的绝对值变小但检验均显著不为零,说明在模型中相较于收入,学历才是借款成功的重要因素。第(3)(4)列是对违约率的分析结果,尽管加入控制变量,学历(edu)的系数均显著为负,说明在很大程度上学历能够反映借款人的违约风险。与其他条件相比,学历越高的借款者,其违约的可能性越低,借款者学历每增加一个单位,借款违约率降低67.61% ($e^{-0.3914}$),假设1b成立。

表4 基准回归结果

变量	success		default		rate	
	(1)	(2)	(3)	(4)	(5)	(6)
edu	0.4470 *** (0.0139)	0.1208 *** (0.0269)	-0.4398 *** (0.0879)	-0.3914 *** (0.0894)	-0.3686 *** (0.0182)	-0.1260 *** (0.0174)
rate		-0.1149 *** (0.0063)		0.2786 *** (0.0199)		
time		-0.0094 *** (0.0027)		0.0190 (0.0139)		0.0436 *** (0.0016)
wamount		-0.0000 *** (0.0000)		0.0000 ** (0.0000)		-0.0000 *** (0.0000)
wage		0.0575 *** (0.0030)		-0.0040 (0.0083)		0.0051 *** (0.0019)
marriage		0.3070 *** (0.0505)		-0.1244 (0.1745)		-0.2042 *** (0.0314)
income		0.5129 *** (0.0184)		0.1780 *** (0.0512)		0.0795 *** (0.0115)
house		-0.3499 *** (0.0523)		0.5591 *** (0.1810)		-0.0118 (0.0353)
house_loan		-0.0847 (0.0730)		-0.5056 ** (0.2135)		-0.3336 *** (0.0484)
car		0.2441 *** (0.0526)		-0.3820 ** (0.1665)		-0.0735 ** (0.0341)
car_loan		-0.1578 * (0.0934)		0.1123 (0.2469)		0.0666 (0.0686)
work		0.2374 *** (0.0224)		0.0290 (0.0800)		-0.0396 *** (0.0146)
level		1.3569 *** (0.0185)				-0.7815 *** (0.0071)
_cons	-3.3117 *** (0.0906)	-7.5352 *** (0.1972)	-4.7978 *** (1.0280)	-10.030 *** (1.2063)	16.1247 *** (0.1020)	15.9690 *** (0.1170)
行业固定效应	YES	YES	YES	YES	YES	YES
省份固定效应	YES	YES	YES	YES	YES	YES
N	64552	64536	15573	15564	64552	64536

注:(1)括号内为估计标准误,为了防止异方差问题,所报告的均为稳健标准误;(2)*表示显著性水平为10%,**表示显著性水平为5%,***表示显著性水平为1%

(二) 学历价值的地区异质性分析

中国省份较多、分布较广,各地区之间经济发展状况及金融发展水平存在很大差异,因此学历价值对

信贷市场的地区异质性影响分析就很有必要。基于此,本文整理出借款人工作所在地,并按照国家统计局划分标准,将数据分成东、中、西三组,从区域差异视角探究借款者学历价值的表现,具体实证结果如表5所示。

从第(1)、(2)、(3)这3列可以看出,贷款者对借款者学历认可程度呈现区域性差异。第(4)、(5)、(6)列是对借款者违约行为的分析结果,东部地区高学历的借款者相比西部地区高学历的借款者表现出更低的借款违约率。第(7)、(8)、(9)列是对借款人利率水平的分析结果显示,无论是在东、中部地区还是西部地区,学历的系数都显著为负,说明无论在哪个地区,在设定借款利率时,借贷双方都会考虑到借款人的学历价值。

综合以上分析可知,在我国信贷市场中借款者的学历能够发挥一定的价值,学历较高的借款者违约风险较低,且这一价值在东部地区体现得更为明显。此外,借贷双方都会考虑到学历的价值,从而设定一个较低水平的借款利率。而从授信方的角度来看,贷款者能够很好地捕捉到隐含借款者风险信息——“学历”这一指标,更倾向于向高学历借款者放贷,且对东部地区的高学历者更为青睐。^①

表5 分组回归结果

变量	success			default			rate		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	东部	中部	西部	东部	中部	西部	东部	中部	西部
edu	0.2997*** (0.0967)	0.1057*** (0.0280)	0.2313** (0.0907)	-1.4908*** (0.5277)	-0.3950*** (0.0906)	-0.3253 (0.5338)	-0.1701*** (0.0597)	-0.1156*** (0.0187)	-0.2586*** (0.0581)
edu(dy/dx)	0.0100*** (0.0033)	0.0041*** (0.0011)	0.0075** (0.0030)	-0.0146*** (0.0054)	-0.0076*** (0.0017)	-0.0018 (0.0031)			
_cons	-7.7860*** (0.5943)	-7.1110*** (0.2377)	-6.4166*** (0.5296)	-7.4094*** (2.5210)	-8.5272*** (0.8360)	-9.194*** (2.7886)	14.8677*** (0.3389)	16.5367*** (0.1171)	15.3502*** (0.3262)
控制变量	YES	YES	YES	YES	YES	YES	YES	YES	YES
行业固定效应	YES	YES	YES	YES	YES	YES	YES	YES	YES
省份固定效应	YES	YES	YES	YES	YES	YES	YES	YES	YES
N	5915	55857	6189	1467	14049	834	5922	55857	6189

注:(1)括号内为估计标准误,为了防止异方差问题,所报告的均为稳健标准误;(2)*表示显著性水平为10%,**表示显著性水平为5%,***表示显著性水平为1%

为更好地考察学历价值的地区异质性,本文进一步引入地区金融发展程度(finance)和借款人学历状况的交叉变量来探究学历价值地区差异的深层机制。其中,地区金融发展水平用地区金融机构年末存贷款余额占该地区GDP的比重表示。实证结果如表6所示。

第(7)、(8)、(9)列分别是东、中、西部地区借款利率的回归结果,交互项均显著为负,说明地区金融发展水平在借款者学历对借款利率的影响方面起到负向调节作用,但交互项的系数大小和显著性水平在区域之间并不存在明显的规律。

第(1)、(2)、(3)列分别是东、中、西部地区借款成功率的回归结果,交互项的回归系数均显著为正,说明地区金融发展水平在借款者学历对借款违约率的影响方面的确起到了正向调节作用;从边际效应和系数的显著性水平来看,该调节效果在东部地区更为明显,而中西部地区均有所降低,说明学历价值表现出区域的差异性是由于地区金融发展水平所导致的。

第(4)(5)(6)列分别是东、中、西部地区借款违约率的回归结果,交互项的回归系数均为负,且存在显著差异,东部地区的显著性水平最高,其次是中部地区,西部地区并不显著,说明在借款者学历对借款违约率影响的过程中,地区金融发展水平的负向调节作用存在地区差异。这一结果进一步证明了借款者学历价值呈现地区差异是由地区金融水平的不同所造成的。综上所述,假设2成立。

^①进一步为了证明两组之间的回归系数存在显著性差异,在分组回归的基础上,进行suest检验,验证了显著性差异的存在,后文的分组回归均进行了组间系数差异检验。

表6 金融差异的地区分组回归结果

变量	success			default			rate		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	东部	中部	西部	东部	中部	西部	东部	中部	西部
fina_edu	5.7481*** (2.1259)	2.1848** (1.0448)	4.2059** (1.9392)	-25.5205*** (9.5228)	-17.3890** (8.2805)	-15.3221 (11.4520)	-5.2650*** (1.3145)	-4.5040*** (0.6793)	-6.0583*** (1.2790)
dy/dx	0.1925*** (0.0720)	0.8532** (0.0409)	0.1361** (0.0633)	-0.1270*** (0.0475)	-0.3422** (0.1629)	-0.0856 (0.0719)			
_cons	-7.961*** (0.6440)	-7.016*** (0.2384)	-6.5137*** (0.5663)	-9.532*** (3.5304)	-7.631*** (0.5513)	-7.8557* (4.3249)	15.225*** (0.3651)	16.5623*** (0.1176)	15.6453*** (0.3526)
控制变量	YES	YES	YES	YES	YES	YES	YES	YES	YES
行业固定效应	YES	YES	YES	YES	YES	YES	YES	YES	YES
省份固定效应	YES	YES	YES	YES	YES	YES	YES	YES	YES
N	5915	55857	6189	1155	14162	834	5915	55857	6189

注:(1)括号内为估计标准误,为了防止异方差问题,所报告的均为稳健标准误;(2)*表示显著性水平为10%,**表示显著性水平为5%,***表示显著性水平为1%

(三) 内生性检验

本文因变量是否成功借款(success)和借款利率(rate)的观测样本均为65534,而因变量是否违约(default)的观测样本仅为15564,大量观测样本缺失可能导致该模型由于样本选择偏差而存在内生性问题。借鉴 Yu 等的研究方法,本文将采用 Heckman 二阶段模型来缓解潜在的内生性问题^[28]。同时,借鉴 Chrisman 和 Patel 等研究,本文将分别选择以下三个工具变量:(1)个体*i*所在省份中披露是否违约的人数占比(report ratio1);(2)个体*i*所在城市中披露是否违约的人数占比(report ratio2);(3)个体*i*所在行业中披露是否违约的人数占比(report ratio3)^[29]。选择上述三个工具变量主要有两点原因:一方面,个体*i*的同省份、同城市、同行业相似者披露是否违约这一数据的比例越高,就越容易产生同群效应,能够促进个体*i*也倾向于披露是否违约数据;另一方面,同省份、同城市、同行业相似者披露是否违约数据的占比并不会对个体*i*是否违约产生显著影响,因此上述变量符合工具变量的基本要求。

表7列示了本文 Heckman 二阶段模型处理内生性问题的检验结果。其中,模型(1)和模型(2)为本文利用个体*i*所在省份中披露是否违约的人数占比(report ratio1)为工具变量的内生性检验模型。结果显示:工具变量回归系数显著为正(coef. = 1.998, $p < 0.01$),表明同省份内其他个体披露因变量否成功借款的比率越高,个体*i*也更倾向于披露这一变量。同时,本文将通过模型(1)估计获得的尼米尔斯比率(Inverse mills ratio1)放入到模型(2)中作为控制变量,自变量学历(edu)的回归系数依旧显著为负(coef. = -0.742, $p < 0.01$),表明在控制内生性问题后,学历对个体违约的负向作用依旧显著。

模型(3)和模型(4)为本文利用个体*i*所在城市中披露是否违约的人数占比(report ratio2)为工具变量的内生性检验模型。结果显示:工具变量回归系数显著为正(coef. = 2.209, $p < 0.01$),表明同城市内其他个体披露因变量否成功借款的比率越高,个体*i*也更倾向于披露这一变量。同时,本文将通过模型(3)估计获得的尼米尔斯比率(Inverse mills ratio2)放入到模型(4)中作为控制变量,自变量学历(edu)的回归系数依旧显著为负(coef. = -0.648, $p < 0.01$),表明在控制内生性问题后,学历对个体违约的负向作用依旧显著。

模型(5)和模型(6)为本文利用个体*i*所在行业中披露是否违约的人数占比(report ratio3)为工具变量的内生性检验模型。结果显示:工具变量回归系数显著为正(coef. = 2.016, $p < 0.01$),表明同行业内其他个体披露因变量是否成功借款的比率越高,个体*i*也更倾向于披露这一变量。同时,本文将通过模型(5)估计获得的尼米尔斯比率(Inverse mills ratio3)放入到模型(6)中作为控制变量,自变量学历(edu)的回归系数依旧显著为负(coef. = -0.769, $p < 0.01$),表明在控制内生性问题后,学历对个体违约的负向作用依旧显著。

表7 内生性检验

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
	report1	default	report2	default	report3	default
report ratio1	1.998 *** (0.058)					
report ratio2			2.029 *** (0.058)			
report ratio3					2.016 *** (0.043)	
edu	0.172 *** (0.009)	-0.742 *** (0.124)	0.171 *** (0.009)	-0.648 *** (0.110)	0.202 *** (0.009)	-0.769 *** (0.126)
rate	-0.198 *** (0.003)	0.727 *** (0.104)	-0.198 *** (0.003)	0.610 *** (0.084)	-0.190 *** (0.003)	0.679 *** (0.093)
time	0.048 *** (0.001)	-0.082 *** (0.026)	0.048 *** (0.001)	-0.055 ** (0.022)	0.037 *** (0.001)	-0.052 *** (0.020)
wamount	-0.000 *** (0.000)	0.000 *** (0.000)	-0.000 *** (0.000)	0.000 *** (0.000)	-0.000 *** (0.000)	0.000 *** (0.000)
wage	0.055 *** (0.001)	-0.105 *** (0.026)	0.055 *** (0.001)	-0.078 *** (0.020)	0.055 *** (0.001)	-0.100 *** (0.024)
marriage	0.424 *** (0.017)	-1.021 *** (0.251)	0.425 *** (0.017)	-0.783 *** (0.221)	0.440 *** (0.018)	-0.993 *** (0.243)
income	0.306 *** (0.006)	-0.449 *** (0.151)	0.305 *** (0.006)	-0.280 ** (0.121)	0.291 *** (0.006)	-0.374 *** (0.133)
house	-0.726 *** (0.019)	1.985 *** (0.383)	-0.730 *** (0.019)	1.612 *** (0.321)	-0.669 *** (0.019)	1.777 *** (0.339)
house_loan	0.125 *** (0.025)	-0.767 *** (0.224)	0.127 *** (0.025)	-0.709 *** (0.220)	0.155 *** (0.025)	-0.794 *** (0.227)
car	0.269 *** (0.018)	-0.979 *** (0.215)	0.269 *** (0.019)	-0.826 *** (0.197)	0.231 *** (0.019)	-0.854 *** (0.198)
car_loan	-0.143 *** (0.034)	0.396 (0.257)	-0.145 *** (0.034)	0.339 (0.255)	-0.091 *** (0.034)	0.269 (0.251)
work	0.130 *** (0.008)	-0.257 ** (0.101)	0.129 *** (0.008)	-0.183 * (0.093)	0.134 *** (0.008)	-0.240 ** (0.099)
Inverse mills ratio1		-2.934 *** (0.662)				
Inverse mills ratio 2				-2.164 *** (0.524)		
Inverse mills ratio 3						-2.693 *** (0.601)
Constant	-2.663 *** (0.060)	-3.562 * (1.892)	-2.670 *** (0.060)	-5.302 *** (1.625)	-2.680 *** (0.060)	-4.057 ** (1.797)
Observations	64536	15564	64536	15564	64536	15564
行业固定效应		YES		YES		YES
省份固定效应		YES		YES		YES

注:括号里是估计标准误,*** $p < 0.01$,** $p < 0.05$,* $p < 0.1$

(四) 倾向匹配得分回归结果和分析

为了避免样本的“自选择”而引起的“选择性偏差”,本文进一步采用倾向匹配得分法(PSM)进行分

析。由于PSM需要确定实验组和对照组,所以本文将东部地区设定为实验组,中西部地区设定为对照组,采用1:1无放回近邻匹配原则为对照组中每一个借款订单在实验组中找到一个与之匹配的另一借款订单。^①最后,基于匹配后的样本重新进行实证分析。

在实证分析前,先对匹配的平衡性假设进行检验,如表8所示,匹配后偏差均小于5%,对照组和实验组之间不存在显著性差异,表明匹配的效果较好,满足匹配的平衡性假设。图1是匹配前后实验组和对照组的分布密度函数图,从图中可以看出匹配后实验组和对照组之间的分布密度几乎一致,同样说明两组样本不存在显著性差异。

在满足平衡性假设的前提下,对匹配后的样本重新进行分组回归。就借款利率水平而言,无论是在东部还是中西部地区,学历的系数都显著为负,该结果和上文表5的分析结果表现一致。表9是匹配后学历价值的地区异质性结果,可以看出东部地区高学历的借款者相比中西部地区高学历的借款者表现出更低更显著的借款违约率,学历的低风险价值在东部地区表现得更为显著。表10是匹配后加入地区金融发展水平和借款者学历交互项的回归结果,回归结果与上文表6的分析结果一致,说明借款者学历价值呈现地区的差异化的确是因为地区金融水平的不同导致的。

表8 PSM的平衡性假设检验

变量	Unmatched	Mean		% bias	t-test	
	Matched	Treated	Control		t	P > t
edu	U	1.8839	1.775	13.2	10.59	0
	M	1.8839	1.8653	2.3	1.26	0.208
time	U	10.715	13.236	-30.1	-21.09	0
	M	10.715	10.603	1.3	0.88	0.381
wamount	U	55963	50165	6.2	5.02	0
	M	55963	56096	-0.1	-0.07	0.94
rate	U	15.461	15.674	-5.9	-4.61	0
	M	15.461	15.605	-4	-2.26	0.024
work	U	2.4247	2.4219	0.3	0.21	0.83
	M	2.4247	2.4074	1.7	1	0.318
income	U	3.9436	3.9532	-0.7	-0.55	0.579
	M	3.9436	3.9008	3.2	1.85	0.064
wage	U	34.261	34.133	1.7	1.3	0.195
	M	34.261	34.027	3.1	1.82	0.068
Car-loan	U	0.0422	0.04361	-0.7	-0.53	0.594
	M	0.0422	0.03976	1.2	0.7	0.481
House-loan	U	0.12401	0.10271	6.7	5.34	0
	M	0.12401	0.11807	1.9	1.04	0.297
house	U	0.4025	0.37582	5.5	4.23	0
	M	0.4025	0.39732	1.1	0.61	0.545

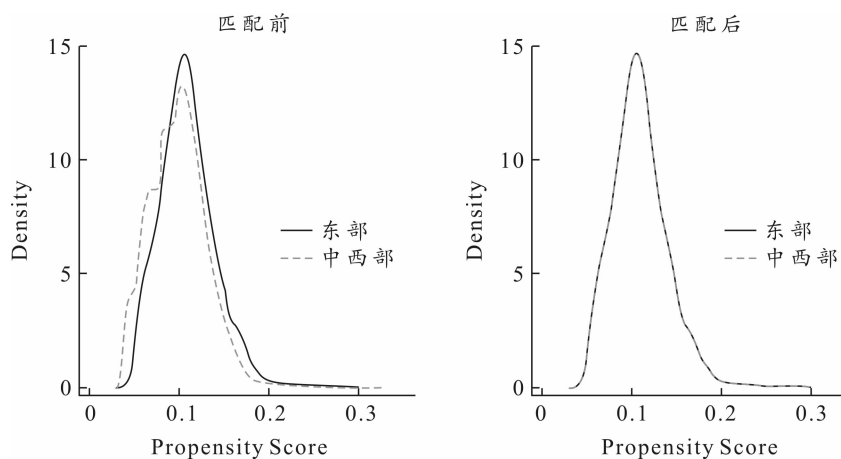


图1 匹配前后密度函数图

^①用 pselect 法对协变量的一阶形式进行筛选,确定匹配变量 edu、income、wage、work、car-loan、house、marriage、level、wamount、car、time、house-loan、rate。

表9 PSM 分组回归结果

变量	success		default		rate	
	(1)	(2)	(3)	(4)	(5)	(6)
	东部	中西部	东部	中西部	东部	中西部
edu	0.2997*** (0.0967)	0.2171*** (0.0751)	-1.4908*** (0.5277)	-0.0044 (0.0050)	-0.1701*** (0.0597)	-0.2013*** (0.0559)
edu(dy/dx)	0.0100*** (0.0033)	0.0089*** (0.0031)	-0.0146*** (0.0054)	-0.0044 (0.0050)		
_cons	0.2997*** (0.0967)	0.2171*** (0.0751)	-1.4908*** (0.5277)	-0.0044 (0.0050)	-0.1701*** (0.0597)	-0.2013*** (0.0559)
控制变量	YES	YES	YES	YES	YES	YES
行业固定效应	YES	YES	YES	YES	YES	YES
省份固定效应	YES	YES	YES	YES	YES	YES
N	5915	6524	1467	1315	5915	6533

注:(1)括号内为估计标准误,为了防止异方差问题,所报告的均为稳健标准误;(2)*表示显著性水平为10%,**表示显著性水平为5%,***表示显著性水平为1%

表10 PSM 金融差异的地区分组回归结果

变量	success		default		rate	
	(1)	(2)	(3)	(4)	(5)	(6)
	东部	中西部	东部	中西部	东部	中西部
Fina-edu	5.7481*** (2.1259)	5.6559** (2.7496)	-25.5205*** (9.5228)	-8.2666 (6.9579)	-5.2650*** (1.3145)	-8.3532*** (2.0295)
dy/dx	0.1925*** (0.0720)	0.2324** (0.1134)	-0.1270*** (0.0475)	-0.2092 (0.1784)		
_cons	-7.9606*** (0.6440)	-7.6278*** (0.6303)	-9.5320*** (3.5304)	-6.2074*** (1.5326)	15.2251*** (0.3651)	16.7954*** (0.4029)
控制变量	YES	YES	YES	YES	YES	YES
行业固定效应	YES	YES	YES	YES	YES	YES
省份固定效应	YES	YES	YES	YES	YES	YES
N	5915	6524	1155	1321	5915	6533

注:(1)括号内为估计标准误,为了防止异方差问题,所报告的均为稳健标准误;(2)*表示显著性水平为10%,**表示显著性水平为5%,***表示显著性水平为1%

(五) 稳健性检验

为了避免伪回归,进行以下稳健性检验。首先,替换金融发展水平这一变量。表10的回归结果中地区金融发展水平采用的是金融机构年末存贷款余额占GDP的比重,在这里本文用樊纲构造的地区金融市场化指数来替代并重新进行回归,结果如表11所示,可以看出,当替换解释变量之后原结论依旧成立。其次,替

换学历变量,采用学历的哑变量^①代替之前的有序变量,对东部和中西部地区进行分组回归,结果如表12所示,该分析结果与上文保持一致。第三,用地区(east)和借款者学历(edu)的交互项代替分组回归,结果也依然不变(如表13)。

表11 稳健性检验1——替换金融发展水平

变量	success		default		rate	
	(1)	(2)	(3)	(4)	(5)	(6)
	东部	中西部	东部	中西部	东部	中西部
finadex_edu	14.1350*** (4.6851)	7.6649** (3.3050)	-83.9616** (42.1823)	-13.4790 (9.3584)	-10.0369*** (2.9203)	-9.6549*** (2.5081)
dy/dx	0.4730*** (0.1585)	0.3149** (0.1362)	-0.4159** (0.1980)	-0.3412 (0.2415)		
_cons	-7.9361*** (0.6203)	-7.6977*** (0.6313)	-8.9950** (3.5007)	-6.0024*** (1.5537)	15.0475*** (0.3542)	16.8064*** (0.4053)
控制变量	YES	YES	YES	YES	YES	YES
行业固定效应	YES	YES	YES	YES	YES	YES
省份固定效应	YES	YES	YES	YES	YES	YES
N	5915	6524	1155	1321	5915	6533

注:(1)括号内为估计标准误,为了防止异方差问题,所报告的均为稳健标准误;(2)*表示显著性水平为10%,**表示显著性水平为5%,***表示显著性水平为1%

表12 稳健性检验2——替换学历变量

变量	success		default		rate	
	(1)	(2)	(3)	(4)	(1)	(2)
	东部	中西部	东部	中西部	东部	中西部
edu	0.3658** (0.1864)	0.1895 (0.1525)	-1.1002* (0.6120)	-0.0104 (0.0088)	-0.4193*** (0.1123)	-0.4313*** (0.1038)
_cons	-7.2048*** (0.5414)	-7.3347*** (0.6032)	-8.8739*** (2.6135)	0.1394** (0.0604)	14.6529*** (0.3077)	16.4615*** (0.3915)
控制变量	YES	YES	YES	YES	YES	YES
行业固定效应	YES	YES	YES	YES	YES	YES
省份固定效应	YES	YES	YES	YES	YES	YES
N	5915	6524	1467	1315	5915	6533

注:(1)括号内为估计标准误,为了防止异方差问题,所报告的均为稳健标准误;(2)*表示显著性水平为10%,**表示显著性水平为5%,***表示显著性水平为1%

①本科及其以上取值为1,本科以下取值为0。

表13 稳健性检验3——地区交互项代替分组回归

变量	(1)	(2)	(3)
	success	default	rate
east_edu	0.2676 *** (0.0942)	-0.6572 *** (0.1684)	-0.1434 ** (0.0572)
_cons	-7.5596 *** (0.4196)	-7.4450 *** (1.1862)	15.4393 *** (0.2493)
控制变量	YES	YES	YES
行业固定效应	YES	YES	YES
省份固定效应	YES	YES	YES
N	12448	2788	12448

注:(1)括号内为估计标准误,为了防止异方差问题,所报告的均为稳健标准误;(2)*表示显著性水平为10%,**表示显著性水平为5%,***表示显著性水平为1%

五、结论与启示

公民受教育水平对经济社会活动影响的研究成果丰硕,为我们理解学历价值的特征提供了重要参考,但现有的研究较少从信贷视角来考察学历的价值。网络信贷市场的迅速发展,不仅积累了大量包含受信者学历信息的借贷数据,也为这一研究提供了很好的契机。基于此,本文运用A网络信贷平台公开数据,探究借款者学历价值的存在性及地区异质性,得出如下结论与启示:

第一,在中国信贷市场中,借款者的受教育水平能够发挥其价值,表现为授信方能够很好地识别出借款者学历这一因素并表现出对高学历价值的“认可”。高学历借款者的借款成功率更高,并可获得较低借款利率。

第二,从地区异质性的角度来看,在东部地区,借款者的学历价值表现得更加显著,即高学历的借款者相比西部地区具有更低的借款违约率,且更易获得授信方的青睐。说明经济社会发展水平越高的地区,高学历借款者越有可能获得更高的教育回报率。

本文的研究结论对学历价值的存在性提供了有力支撑,证明了教育在信贷市场中的确发挥了其应有之义,且能够被授信方很好地识别和认可。教育一直是中国社会建设的重要组成部分,历年来在财政支出方面教育的投入巨大,但现有研究更多地还是关注教育对收入等层面的影响,较少地关注到信用这一层面,教育的目的不仅在于提高受教育者收入水平,其终极的目标在于对受教育者的道德约束以及诚信理念的提升。由于教育是一项较为复杂的系统工程,其最终效用不仅依赖于经费资源投入,而且与制度设计、人才环境、资源禀赋等因素也存在重要关系。此外,本文的结论也为信贷市场的风险防控提供了重要参考。借贷双方的信息不对称和识别困难是投资市场风险增加、资金错配的重要原因,如何在借款人的众多信息中识别出能够反映借款人真实违约风险的指标则对于贷款者而言具有重要意义。本文为贷款者对借款者学历价值评估提供了一定的借鉴思路。作为市场上容易观测的“学历”这一因素,可以帮助贷款者较为准确地识别出潜在的投资风险,合理地进行投资决策。

参考文献:

- [1]王先柱,吴义东,吴璟.住房按揭贷款逾期风险识别与管理研究——基于借款人学历视角的实证检验[J].浙江工商大学学报,2020(4):125-137.
- [2]张雪丽,朱天星,于立新.基于判别分析的商业银行个人信用风险评价模型研究[J].工业技术经济,2011(10):131-138.

- [3] 廖理,吉霖,张伟强. 借贷市场能准确识别学历的价值吗? ——来自 P2P 平台的经验证据[J]. 金融研究,2015(3):146-159.
- [4] 丁琳. 从学历参透人才价值[J]. 中国邮政,2012(8):7.
- [5] 胡海燕. 学历对个人收入的影响分析[J]. 企业导报,2010(6):248-249.
- [6] GATHERGOOD J. Self-control, financial literacy and consumer over-indebtedness[J]. Journal of Economic Psychology,2011,33(3):590-602.
- [7] 孙武军,樊小莹. 从业经历和教育背景是否能提高借贷成功率? ——来自 P2P 平台的经验证据[J]. 中央财经大学学报,2016(3):33-41.
- [8] 邢春冰,贾淑艳,李实. 教育回报率的地区差异及其对劳动力流动的影响[J]. 经济研究,2013(11):114-126.
- [9] 申广军,龚雅娴,姚洋. 金融发展与教育回报率的地区差异[J]. 金融研究,2015(3):131-145.
- [10] 赵西亮. 教育、户籍转换与城乡教育收益率差异[J]. 经济研究,2017(12):164-178.
- [11] 姚先国,张海峰. 教育、人力资本与地区经济差异[J]. 经济研究,2008(5):47-57.
- [12] 郭庆旺,贾俊雪. 公共教育政策、经济增长与人力资本溢价[J]. 经济研究,2009(10):22-35.
- [13] 程名望, JIN Y H, 盖庆恩, 等. 农村减贫:应该更关注教育还是健康? ——基于收入增长和差距缩小双重视角的实证[J]. 经济研究,2014(11):130-144.
- [14] 关会娟,李昕,谭莹. 教育投入、交易成本与区域收入差距[J]. 财经研究,2019(7):97-111.
- [15] 樊文皓,郝刚. 大学生综合素质存在的问题与提升对策[J]. 产业与科技论坛,2010(2):173-175.
- [16] 冯丹. 高学历者受教育程度与消费行为的相关性分析[D]. 四川:西南大学教育学部,2008.
- [17] 孙红艳. 社会主义核心价值观体系与高校思想政治教育方法创新研究[J]. 思想教育研究,2014(9):67-70.
- [18] 李滨源. 从业经历和学历水平对借贷行为的影响[D]. 天津:天津大学管理与经济学部,2018.
- [19] 李霖魁,张成虎. P2P 网络借贷中的借款人社会资本、风险甄别与市场均衡实现[J]. 当代财经,2017(10):46-57.
- [20] 黄亮,付伟,倪克勤. 逆向选择与信用配给:中小企业融资难根源分析[J]. 中央财经大学学报,2005(1):39-43.
- [21] 谭惠文. P2P 网络借贷:借款描述与借款成功率及违约率关系[D]. 湖南:湖南大学金融与统计学院,2019.
- [22] 程瑶. 学历水平在借贷市场中的作用——来自 P2P 市场的经验证据[J]. 上海金融,2018(2):48-55.
- [23] 蒋彧,周安琪. P2P 网络借贷中存在地域歧视吗? ——来自“人人贷”的经验数据[J]. 中央财经大学学报,2016(9):29-39.
- [24] 张男星,王纾,孙继红. 我国高等教育综合发展水平评价及区域差异研究[J]. 教育研究,2014(5):28-36.
- [25] 王奕莹. 金融发展、高端制造业集聚与地区经济发展差距——基于新经济地理学的分析框架以及来自中国 31 省市的证据[J]. 经济问题探索,2017(6):129-137.
- [26] 陈慧. 区域金融发展水平、债务资本成本与资本结构动态调整速度[D]. 江西:江西财经大学会计学院,2016.
- [27] 吴汉洪,徐国兴. 信用本质的经济学分析[J]. 中国人民大学学报,2004(4):56-62.
- [28] YU W, MINNITI M, NASON R. Underperformance duration and innovative search: evidence from the high-tech manufacturing industry[J]. Strategic Management Journal,2019,40(5):836-861.
- [29] CHRISMAN J J, PATEL P C. Variations in R&D investments of family and nonfamily firms: behavioral agency and myopic loss aversion perspectives[J]. Academy of Management Journal,2012,55(4):976-997.



(责任编辑 郭宝才)