

# 大数据征信在平台授信中的信息价值研究——基于互联网平台的消费行为信息视角

宋科 虞思燕 杨雅鑫

**【摘要】**为厘清大数据征信在互联网消费信贷平台授信决策中发挥的作用，从国内某偏重旅游的互联网消费信贷平台随机抽取约 50 万活跃用户数据进行实证分析。研究发现：① 大数据征信在互联网消费信贷平台的授信决策中提供了信息价值，消费信贷申请人的消费能力、消费习惯以及用户黏性均被纳入授信考量，累计消费金额越高、消费券使用越频繁、平台使用越久的消费者，在消费信贷授信过程中更具有优势。② 传统征信信息和消费行为所含的信息不同，在授信决策过程中存在替代关系，平台使用时长、信用卡支付比重等信息对传统征信信息产生替代。③ 从异质性特征看，平台授信过程中不存在地域歧视，并且更注重考察年轻人的传统征信情况及平台依赖度等信息，对年龄较大的群体则更偏重于对其消费能力等信息的考察，说明互联网消费信贷具有普惠性。④ 新冠疫情冲击下，金融科技平台风控加强，降低用户平台依赖程度在授信决策中的权重，这在一定程度上表现出顺周期性。

**【关键词】**大数据征信；授信决策；消费行为；传统征信；新冠疫情

**【文章编号】** IMI Working Papers No.2602



微博·Weibo



微信·WeChat

更多精彩内容请登陆 国际货币网

<http://www.imi.org.cn/>

# 大数据征信在平台授信中的信息价值研究

## ——基于互联网平台的消费行为信息视角

宋科<sup>1</sup>，虞思燕<sup>2</sup>，杨雅鑫<sup>3</sup>

**【摘要】**为厘清大数据征信在互联网消费信贷平台授信决策中发挥的作用，从国内某偏重旅游的互联网消费信贷平台随机抽取约 50 万活跃用户数据进行实证分析。研究发现：① 大数据征信在互联网消费信贷平台的授信决策中提供了信息价值，消费信贷申请人的消费能力、消费习惯以及用户黏性均被纳入授信考量，累计消费金额越高、消费券使用越频繁、平台使用越久的消费者，在消费信贷授信过程中更具有优势。② 传统征信信息和消费行为所含的信息不同，在授信决策过程中存在替代关系，平台使用时长、信用卡支付比重等信息对传统征信信息产生替代。③ 从异质性特征看，平台授信过程中不存在地域歧视，并且更注重考察年轻人的传统征信情况及平台依赖度等信息，对年龄较大的群体则更偏重于对其消费能力等信息的考察，说明互联网消费信贷具有普惠性。④ 新冠疫情冲击下，金融科技平台风控加强，降低用户平台依赖程度在授信决策中的权重，这在一定程度上表现出顺周期性。

**【关键词】**大数据征信；授信决策；消费行为；传统征信；新冠疫情

近年来，随着大数据、区块链、人工智能以及云计算等技术不断发展，中国数字经济步入高速增长轨道，新产业、新业态、新模式不断涌现，为高质量发展注入新动能。作为数字经济的核心生产要素，数据正成为科技创新的突破口。党的十九届四中全会明确提出，“健全劳动、资本、土地、知识、技术、管理、数据等生产要素由市场评价贡献、按贡献决定报酬的机制”。2022 年，中国人民银行发布的《金融科技发展规划（2022—2025 年）》明确指出，要发挥数据要素倍增作用，将数字元素注入金融服务全流程，将数字思维贯穿于业务运营全链条。2023 年，习近平总书记在中央金融工作会议上首次提出，要“做好科技金融、绿色金融、普惠金融、养老金融、数字金融‘五篇大文章’”，为推进金融高质量发展指明了方向。在新发展阶段，进一步激活数据要素潜能，充分发挥大数据征信在金融机构授信决策中的作用，对于统筹金融发展与安全，促进经济高质量发展具有重大战略意义。

在传统征信模式中，征信机构主动收集信贷申请主体的征信信息（包括基本信息、借贷信息、财务信息、公共信息等），进而对其开展信用评估。然而，由于数据来源有限、覆盖人群范围小、关键信息缺失、数源标准不统一、更新不及时等问题，传统征信模式无法构建起有效的征信系统，尤其缺

---

<sup>1</sup> 宋科，中国人民大学国际货币研究所副所长，中国人民大学财政金融学院教授

<sup>2</sup> 虞思燕，厦门市思明区财政局

<sup>3</sup> 杨雅鑫，新加坡国立大学房地产研究院

乏对下沉群体的征信覆盖，影响金融服务普惠性<sup>[1-2]</sup>。近年来，金融科技信贷公司的兴起，不仅为市场带来了规模巨大的信贷资源，也创造出大数据征信模式，极大程度上增强了传统征信的信息覆盖广度与挖掘深度<sup>[3]</sup>。Hau 等<sup>[4]</sup>发现蚂蚁集团信用评估的关键元素是将历史违约数据与来自电子商务交易平台的销售和财务数据结合起来，所用信息不仅包括借款人财务信息，还包括与电子商务平台和支付网络的其他参与者之间关系。Philippon<sup>[5]</sup>研究发现，信贷公司通过机器学习算法，将消费者手机账单、购物和订阅记录、网页浏览历史等行为数据用于贷款决策。Lin 等<sup>[6]</sup>利用美国 Prosper 借贷平台数据的实证分析发现，借款人的社会网络关系信息可以用于信贷审核，缓解信贷双方的信息不对称，从而优化市场信贷配给。Frost 等<sup>[7]</sup>研究发现，BigTech 公司能够通过其核心业务（如电子商务等）提供的交易额、平台信誉等信息来弥补个人消费者信贷记录缺失的局限。王正位等<sup>[2]</sup>将借款人是否得到花呗、京东白条授信作为消费者行为信息的替代变量、信用卡为代表的传统征信信息，评价消费行为信息在个人信用风险评估中的信息含量。黄益平等<sup>[8]</sup>提出一个新的信用风险管理框架，综合利用包括用户的收入、年龄、性别、不同时间段下银行信用卡信息（违约情况、使用频率、额度、使用历史）等 36 个传统征信信息，以及用户在网络平台上的消费数据、社交数据及其他行为信息等 99 个非传统信息变量。邱涌钦等<sup>[9]</sup>使用某大型消费金融平台数据，利用高维主动学习方法，从 50 个变量中筛选出影响用户违约的 15 个重要变量，分别为：学历、是否有车、公司行业、公司规模、工作时间、信用额度、利率、借款类型、还款期限。Óskarsdóttir 等<sup>[10]</sup>使用包括 9 000 万个电话号码、200 万银行客户及其他一系列社会人口统计和银行信息在内的匿名数据，对持卡人社会关系进行建模分析的结果表明，与仅依赖传统社会人口或银行数据的预测相比，考虑到持卡人电话联系网络的特征，可以提高信用和偿付能力分析预测的准确性。Wei 等<sup>[11]</sup>通过对比网络数据对于客户评分准确性的影响，发现利用网络数据能够提高客户评分准确性，而消费者可以策略性地经营其社交网络进而获得更高信用评分。Berg 等<sup>[12]</sup>基于德国电子商务公司用户足迹的实证分析也得到了相似结论，用户通过访问或注册网站而留在网上的信息成为信用机构信息的补充，能够提高信贷获取率、降低违约率。美国 Lending Club 将基于金融科技的非传统信息应用在信用分配中，这些基于金融科技的附加信息超出了传统信用评级或传统信用审批标准中通常包含的信息范畴，在平台内部信用评级过程中作用日渐凸显，有助于推动互联网信贷服务渗透到传统银行服务不足的领域<sup>[4, 13]</sup>。Jagtiani 等<sup>[13]</sup>研究表明，相比于传统银行，互联网信贷平台在信用评分较低、抵押贷款拒绝率较高的消费者群体中占据更大市场份额。一方面，大数据征信凭借云计算、机器学习等现代化技术手段，使得大量的、有效的数据采集成为可能，极大程度缓解信贷市场的信息不对称，有助于金融服务向广大下沉群体渗透，惠及更多长尾客户<sup>[13-14]</sup>；另一方面，大数据征信将用户被互联网记录下的行为数据，作为传统征信信息之外的补充信息，用于刻画互联网用户的信用状况，不仅优化了信用评估精度，提高了信用识别效率，也有效防范系统性金融风险<sup>[5, 15-17]</sup>。

从既有文献看，个人征信评估所使用的互联网用户行为数据主要来源于互联网借贷平台的用户数据以及沉淀在第三方金融科技平台的数据足迹。互联网借贷平台将传统借贷业务迁移到线上，信用评

估所用数据多提炼自互联网用户向平台提交的材料。其中一部分信息与传统征信信息类同，包括用户性别、人种、学历等基本信息，收入、房产等财务信息，借款期限、利率等借贷信息<sup>[18-20]</sup>；另一部分信息包括用户提交的照片、借款陈述、借款人的社交网络数量与质量等，借助图像分析、文本分析、网格分析等互联网信息技术，对传统征信信息进行补充<sup>[10,21-22]</sup>。随着大量数据在金融科技平台的不断集聚，数字足迹逐渐进入理论研究视野。数字足迹（数字影子、电子足迹）是指用户在使用互联网时留下的数据痕迹，包括网站浏览情况、在社交媒体上发帖、在线评论或在线购物信息等。国外部分研究已经将社交媒体帖子、电话账单支付、移动支付记录、付费广告观看情况等数字足迹用于个人信贷风险相关研究<sup>[12,23]</sup>。然而，囿于数据可得性，国内研究多直接引用大数据征信结果，使用用户的数字足迹的相关可信研究较少。例如，王正位等<sup>[2]</sup>将借款人是否得到花呗、京东白条授信结果作为消费者行为信息的替代变量，是否获得信用卡作为传统征信信息的替代变量，用于解释违约情况。王奇等<sup>[24]</sup>使用芝麻信用分度量数字信用，研究其对消费信贷的影响。可以看到，既有研究尚有较大拓展空间：①研究视角存在局限。既有研究主要聚焦于征信信息对违约情况的解释，较少关注授信决策层面。事实上，贷前授信决策作为信用风险管理的前置环节，与贷后违约管控共同构筑信用风险管理的稳固框架，均属于信用风险管理的重要部分，同样具有研究价值。若能在贷前授信决策环节排查出贷款申请人的信用风险，更有利于控制信贷风险。②定量分析较少。既有研究对消费者在金融科技平台上留下的消费行为大数据在金融科技平台授信过程中发挥的作用大多停留在定性讨论上，鲜有对其蕴藏的信息价值做定量分析。③对传统征信与大数据征信之间的关系讨论较少。作为集金融与消费于一体的金融科技平台，用户在该平台上的消费能力、用户黏性、消费支付习惯以及信贷依赖程度等消费行为蕴藏着用户信用风险信息，这些信息在平台授信过程中，对传统征信信息起到互补还是替代作用？这一研究问题尚未得到充分讨论。

鉴于此，本文基于互联网消费信贷申请人的消费行为视角，随机抽取国内某旅游在线平台约 50 万活跃用户数据，实证分析了大数据征信在互联网消费信贷平台授信中的信息价值，并进一步厘清了传统征信信息与消费行为信息在平台授信决策过程中的互补关系。

相较于已有研究，本文的边际贡献在于：

（1）拓宽并丰富研究视角。与传统征信依赖基本信息、财务信息、公共信息等静态、低频数据不同，消费行为信息具有发生频率高、及时性强等特征。本文从授信决策这一视角切入，基于国内某旅游消费平台随机抽取的约 50 万活跃用户数据，提炼其消费行为信息，验证了大数据征信在互联网消费信贷平台授信决策中的信息价值。

（2）验证了消费行为信息在平台授信决策中对传统征信信息的互补和替代效应。借鉴已有文献将是否得到花呗、京东白条等粗颗粒度信息作为消费者行为信息的替代变量，本文以微观个体的消费特征（如累计消费金额、平台黏性等），作为消费行为信息的衡量变量，研究发现消费行为信息为平台授信决策提供了额外的价值信息，增加对消费行为信息的使用，能够提高对用户的风险识别精度，降

低平台潜在的信贷风险。其中，用户的消费金额与传统征信有互补性，但用户黏性以及支付习惯对传统征信有替代性。

（3）从征信角度验证了金融科技平台在提供信贷服务中的普惠性。一方面，传统征信信息和大数  
据征信信息在平台授信决策过程中的权重并不存在地区差异，说明平台授信没有地域歧视。另一方面，  
由于年龄较大的人对互联网依赖程度较轻，且其在传统信贷市场更容易受到排挤，因此平台授信过程  
中更注重其在平台上的交易信息，这也表明了金融科技平台在授信过程中不存在年龄歧视。

## 一、理论分析与研究假设

信息不对称是信贷配给产生的主要原因。在完全竞争且信息完全的信贷市场中，商业银行会针对  
不同信用风险的借款人设计出不同的最优信贷合约<sup>[25-26]</sup>。然而，在信息不对称的信贷市场中，贷款人  
无法准确识别借款人的还款概率，会基于风险中值给出平均利率。这导致高风险借款人模仿低风险借  
款人，获得低利率借款。此举不仅会对低风险借款人造成挤出，也将导致贷款人难以获得最优期望收  
益<sup>[27-28]</sup>。征信通过依法采集、评价、利用、维护和管理信用信息，可有效缓解信贷市场信息不对称问  
题，显著提升信贷效率。

伴随着通信技术、大数据、区块链、人工智能等科技的进步，金融科技企业异军突起，凭借社交、  
购物等新型场景的渗透以及大数据、人工智能技术，金融科技公司可以获得更多借款人的软信息及其  
在网络上的数字足迹。相比于结构化数据，这些大数据具有海量、高频、动态、非结构化等特征，能  
够弥补传统征信信息单一、低频、覆盖人群少等缺陷，大数据中蕴藏的信用信息价值可在贷款人授信  
决策中替代传统抵押品<sup>[4-8]</sup>。如 Lending Club、Prosper、蚂蚁集团、京东科技等国内外典型金融科技公  
司均将交易数据、物流数据、个人资料等平台数据用于信贷决策<sup>[6,13,24,29]</sup>；用户在社交平台上留下的社  
交网络信息也可用于小额信贷信用评级<sup>[22]</sup>。基于此，本文提出以下假设：

**假设 1** 大数据征信能够通过基于金融科技平台形成的海量消费行为信息，有效识别消费信贷申请  
人信用水平，提升授信通过概率。

由商业银行主导的传统征信，通常以个体特征信息、资产状况、收入稳定性、历史财务交易数据  
等作为信用评估依据。个人信用历史记录是用于消费信贷信用评估的主要依据。数字金融出现以前，  
我国金融服务的可及性存在明显的地区差异，农村地区和欠发达地区的居民因没有银行账户、缺乏历  
史财务交易数据，不仅无法被纳入银行征信系统，甚至被排除在正规金融市场之外，无法申请消费信  
贷。

相较于传统征信，金融科技平台依赖的大数据征信，凭借其信息技术以及海量数据的优势，很大  
程度上弥补了传统征信的缺失。一是提供替代、补充信息。大数据征信将用户在金融科技平台留下的  
行为数据用于信用评估，如消费行为信息能够反映用户购买能力、信贷依赖程度等信息，社交网络信  
息隐含用户的社会阶层、兴趣偏好、信誉水平等多维度信息。这些信息能够与传统征信信息进行交叉

验证，进一步细化对借款人还款能力、还款意愿层面的刻画<sup>[30]</sup>。二是提高群体覆盖。大数据征信面向全部互联网用户进行信息收集并用于信用评估，能够对传统征信系统无法覆盖、没有信用记录人群的信用状况进行高效刻画，降低金融服务的门槛<sup>[12, 31]</sup>。三是优化征信方式。与传统征信需要大量人力和物力进行信息的收集、录入、管理、维护和处理不同，大数据征信的数据产生于金融科技平台日常运营过程，利用专门的软件对网页信息进行批量抓取，并通过大数据分析和机器学习技术提高信息处理的效率，不仅能提高征信信息的更新速度，还能大幅降低征信成本<sup>[32]</sup>。基于此，本文提出以下假设：

**假设 2** 在授信决策中，大数据征信因与传统征信提供的不同的信息价值而呈现出替代效应。

## 二、数据、变量与实证策略

### （一）数据来源与变量描述

#### 1. 数据来源

本文的数据来自国内某头部在线旅游平台，该平台提供含场景消费信贷服务。消费信贷申请人向平台提交本人身份证号、手机号、储蓄卡号进行实名验证，平台基于个人用户画像，利用内部风险评估模型对借款人的信用风险进行判断，确定是否通过信贷消费申请以及审批可获得的授信额度。

本文在该平台上随机抽取了约 50 万用户数据的消费行为和信贷行为，样本区间为 2018 年 3 月～2021 年 12 月。由于旅游消费具有极强的季节性因素，因此，本文在处理时，将用户数据处理为季度数据。具体做法为：若数据为时期指标，则按季度计算累计规模，如季度的累计消费金额，为对应月度消费金额的累加值；若数据为时点指标，则按季度末的状态取值，如用户是否绑定信用卡，若季度末仍未绑定信用卡的状态，则该季度未绑定信用卡；若该季度末用户绑定信用卡，则计该用户本季度绑定信用卡。

#### 2. 变量描述

授信是否通过。参考 Cornelli 等<sup>[33]</sup>、Kowalewski 等<sup>[34]</sup>，本文采用平台消费信贷申请人授信是否通过作为授信决策结果的代理变量。若平台通过用户申请，则将该信贷申请成功，记为 1，否则信贷申请失败，记为 0。此外，用户授信额度会随着用户信用评分的增加而提高，因此本文将授信额度作为平台对用户信用评估的替代变量，用于稳健性检验。

传统征信信息。信用卡兼具支付和信贷功能，是缓解消费者流动性约束，帮助其实现收入与消费跨期配置的金融工具。根据 Telyukova 等<sup>[35]</sup>、廖理等<sup>[36]</sup>的研究，传统商业银行会根据申请人的年龄、婚姻状况、工作性质、收入水平、资产状况、历史财务信息等传统征信信息综合考虑决定是否对申请人发放信用卡。如果消费者获得信用卡，则说明商业银行根据财务状况和个人特征信息判定其信用风险相对较低、还款能力较强。在我国，高学历、高收入和工作稳定人群往往是商业银行信用卡产品的潜在客户<sup>[37]</sup>，而同样存在消费信贷需求的低收入和弱势群体难以享受到传统商业银行的信贷服务<sup>[38-79]</sup>。参考宋科等<sup>[39]</sup>，本文使用是否绑定信用卡作为衡量传统征信对消费信贷申请人信用风险评估结果的代

理变量。当用户在平台绑定至少一张信用卡时，取值为 1，否则取值为 0。

消费行为信息。本文使用用户在消费信贷平台留存的历史信息构建消费行为信息变量：①消费能力。本文使用累计消费金额对用户消费能力进行度量，即截至提交信贷消费申请时点的前一期，客户在平台的累计消费金额；②用户黏性。本文使用消费券使用频次和平台使用时长对用户黏性进行度量。前者为截至申请时点的前一期，用户在平台累计使用消费券的频次，后者为用户在平台活跃的时间长度，由于存在用户注册平台账户，但长期未使用的情况，因此本文将用户首次出现交易行为记为第一期，计算用户在平台的有效活跃时长；③消费支付习惯以及信贷依赖程度。本文使用信用卡支付比重对消费支付习惯以及信贷依赖程度进行度量，即截至申请时点的前一期，用户在平台消费过程中，使用信用卡支付的金额占总消费金额的比重。信用卡支付比重不仅说明了用户是否得到传统商业银行授信支持，更重要的是该变量反映了消费者的支付习惯。如果信用卡支付比重较大，则说明用户倾向于使用信贷支付手段平滑收入和旅游消费，更容易接受信贷产品。

疫情冲击变量。本文在拓展讨论中探究了疫情冲击对互联网消费信贷平台授信决策的影响，分别使用疫情时间虚拟变量和疫情百度搜索指数两个变量模拟疫情冲击。参考王建红等<sup>[40]</sup>，本文构建疫情冲击的虚拟变量，若时间在 2020 年第一季度至 2021 年第四季度，疫情冲击的虚拟变量取值为 1，反之为 0。参考章晟等<sup>[41]</sup>，构建疫情百度搜索指数，用以衡量疫情冲击程度，具体构建方法为：结合百度资讯指数的现有分词，确定将“交叉感染”“突发公共卫生事件”“新冠病毒”“新冠肺炎”“新型冠状病毒”“疑似病例”“疫情”等作为关键词，分别爬取各地级市百度资讯指数中各关键词的每日搜索热度，将所有关键词按季度加总取对数，并在此基础上根据互联网消费信贷使用者所在地级市进行逐一匹配。百度资讯指数以百度智能分发和推荐内容数据为基础，将网民的阅读、评论、转发、点赞、不喜欢等行为的数量加权求和、指数化处理后得出，全面衡量网民对智能分发和推荐内容的被动关注程度，能够很好地拟合现实中突发危机对居民行为的影响<sup>[42-43]</sup>。

控制变量。消费信贷申请人在个人特征层面的控制变量有：年龄、性别、是否东部地区。年龄变量分为 7 组，分别为 25 岁以下、25~35 岁、35~45 岁、45~55 岁、55~65 岁、65~75 岁及 75 岁以上，其中，25 岁以下记为 1，以此类推，75 岁及以上则记为 7。年龄变量属于有序多项变量，只要年龄与消费之间不存在非线性关系，直接将有序多项变量引入回归模型也是可以的。但为了避免年龄可能与消费之间存在非线性关系从而对回归结果造成影响，本文参考尹志超等<sup>[44]</sup>和李响等<sup>[45]</sup>将年龄的平方项也加入回归模型中。性别为虚拟变量，女性为 1，男性为 0。是否东部地区为虚拟变量，根据信贷申请人所在省份进行分组，若在东部地区，则取值为 1，反之为 0。

表 1 和表 2 分别汇报了本文的主要变量定义及描述性统计。

表1 主要变量说明

Tab.1 Explanation of main variables

变量名	变量说明
-----	------

授信是否通过	平台是否通过用户消费信贷申请，若通过记为1，否则为0
授信额度	平台授予用户的信用额度
累计消费金额	截至申请时点的前一期，用户在平台累计消费金额
消费券使用频次	截至申请时点的前一期，用户在平台累计使用消费券次数
平台使用时长	用户在平台活跃的时间长度
信用卡支付比重	截至申请时点的前一期，用户在平台使用信用卡支付的金额占总消费额的比重
是否绑定信用卡	是为1，否为0
年龄	用户年龄段：1为25岁以下，2为25-34岁，3为35-44岁，4为45-54岁，5为55-64岁，6为65-74岁，7为75岁以上
年龄的平方项	年龄的平方
性别	0男性，1女性
是否东部地区	居住地是否为东部地区，是为1，否为0
疫情冲击	若申请时点在2020年第一季度至2021年第四季度，疫情冲击的虚拟变量均设定为1，反之为0
疫情冲击（百度搜索指数）	关键词“交叉感染”“突发公共卫生事件”“新冠病毒”“新冠肺炎”“新型冠状病毒”“疑似病例”“疫情”词条的百度资讯指数按季度加总之后取对数

表2 主要变量描述性统计

Tab.2 Summary statistics

变量名	样本量	均值	标准差	最小值	最大值
授信是否通过	422,598	0.2944	0.4558	0	1
授信额度	422,598	2.4757	3.8577	0	11.1050
累计消费金额	422,598	5.1168	3.9611	0	17.2036
消费券使用频次	422,598	0.1147	0.1632	0	12
平台使用时长	422,598	5.080	5.3799	0	23
信用卡支付比重	422,598	0.0019	0.0243	0	1
是否绑定信用卡	422,598	0.0863	0.2807	0	1
年龄	422,598	2.4745	0.8807	1	7
年龄的平方项	422,598	6.8987	5.3848	1	49
性别	422,598	0.3142	0.4642	0	1
是否东部地区	422,598	0.3354	0.4721	0	1
疫情冲击（虚拟变量）	422,598	0.3149	0.4645	0	1
疫情冲击（百度搜索指数）	121270	7.3207	0.9637	0	9.6771

## （二）模型设定

在信用风险评价方法中，Logit 回归模型凭借模型解释性强、预测精准度高、稳定性强等优势，在



信用风险评价中被广泛采用<sup>[7,46-47]</sup>。鉴于此，本文基于 Logit 回归构建以下基准模型：

$$\Pr(supplysucc_{i,t}=1) = \alpha + \beta_1 information_{i,t} + \gamma X_{i,t} + \delta_t + \varepsilon_{i,t} \quad (1)$$

其中， $supplysucc_{i,t}$  是平台授信是否通过，若平台第  $t$  期通过第  $i$  个消费信贷申请人的授信申请，则记为 1，否则记为 0。 $information_{i,t}$  既包括传统征信信息，也包括用户  $i$  的消费行为信息，具体为：① 传统征信信息：是否绑定信用卡，即截至申请时点的前一期，信贷申请人  $i$  是否绑定信用卡 0。② 累计消费金额，即截至申请时点的前一期，消费信贷申请人  $i$  在平台累计消费金额；③ 消费券使用频次，即截至申请时点的前一期，消费信贷申请人  $i$  在平台累计使用消费券次数；④ 平台使用时长，即用户在平台活跃的时间长度；⑤ 信用卡支付比重，即截至申请时点的前一期，用户在平台使用信用卡支付的金额占总消费额的比重。 $X_{i,t}$  是关于消费信贷申请人  $i$  在  $t$  期的个体控制变量，包括年龄、性别、居住地区等。 $\alpha$  是常数项， $\delta_t$  是时间虚拟变量，用于控制不同时间点对申请通过率的影响， $\varepsilon_{i,t}$  是随机误差项。

为研究消费行为大数据征信与传统征信之间的关系，本文在模型（1）的基础上加入传统征信  $ccardinf_{i,t}$  与消费行为大数据  $consuminf_{i,t}$  之间的交互项，构建模型（2）：

$$\Pr(supplysucc_{i,t}=1) = \alpha + \beta_1 ccardinf_{i,t} + \beta_2 consuminf_{i,t} + \beta_3 ccardinf_{i,t} * consuminf_{i,t} + \gamma X_{i,t} + \delta_t + \varepsilon_{i,t} \quad (2)$$

由于式（2）涉及交互项，因此对相关交互变量进行去中心化处理。若交互项系数为正，则说明大数据信息并不会削弱传统征信信息在平台授信决策中的权重，两者互为补充，大数据征信与传统征信存在互补性；若交互项系数为负，则说明大数据征信信息会削弱传统征信信息在平台授信决策中的权重，大数据征信与传统征信存在互为替代关系。

### 三、实证结果分析

#### （一）基准回归结果

模型（1）回归结果如表 3 所示。可以看到，消费信贷申请人是否绑定信用卡、累计消费金额、消费券使用频次、平台使用时长和信用卡支付比重等 5 个变量的回归系数均显著为正。信用卡是商业银行根据用户的基本信息、财务状况等传统征信信息对用户做出的授信决定。申请人绑定信用卡，则说明其获得了传统银行的征信，具有一定的还款能力，还款风险相对较低，因此更容易获得平台授信。累计消费金额直接衡量申请人的消费能力，根据收入效应和财富效应，预期未来收入稳定、具有一定财富积累的消费者往往表现出较高的消费能力和消费水平，这也从侧面反映出申请人具备较高的支付能力。消费券是平台为用户提供的专属优惠，仅限于在用户线上预定消费券使用范围内的度假产品使用，并不是覆盖全平台的商家或产品。因此，是否频繁使用消费券有助于识别用户是不是平台核心客户群、能否为平台提供高价值，进而平台更倾向于向消费券使用频次越高的消费者提供消费信贷。平台使用时长可用于衡量消费信贷申请人的平台粘性，当申请人在平台活跃时间越长，在平台留存的有价值信

息就越多，更能帮助平台去了解用户消费、支付习惯及验证用户真实性（以防存在借用他人信息注册、盗号等情况出现），进而获得授信的可能性越高。

表3第（6）列将是否绑定信用卡、累计消费金额、消费券使用频次、平台使用时长和信用卡支付比重这5个变量均放入模型（1），可以看到，消费行为信息影响平台对消费者的信用评估，消费信贷申请人的消费能力、消费习惯以及用户黏性均被纳入授信考量，累计消费金额越高、消费券使用越频繁、平台使用越久的消费者，在消费信贷授信过程中更具有优势，假设1成立。值得注意的是，本文将所有征信信息放入模型中的模型拟合度为0.206，与第（2）列所示仅将累计消费额作为因变量的模型拟合度（为0.202）接近，可见消费信贷申请人的消费能力是平台授信过程中的重要考量因素。值得一提的是，列（6）中信用卡支付比重的回归系数与列（5）相反，这可能是因为信用卡支付比重也代表了用户是否得到银行授信，因此是否绑定信用卡一定程度上替代了信用卡支付比重在平台授信决策过程中的权重。但信用卡支付比重不仅说明了用户是否得到传统商业银行授信支持，更重要的是该变量反映了消费者的支付习惯。在传统征信思路下，信用卡支付比重较高的申请人可能存在过度借贷、共债风险，因此平台会减少信用卡支付比重较高的申请人授信通过的概率。

从消费者特征看，女性授信通过率更高；对于年龄越大的用户，平台更倾向通过其授信申请。地域虚拟变量的回归系数显著，东部地区的消费者更容易通过消费信贷申请，表明生活在不同发展水平城市的消费信贷申请人在平台授信成功率上出现明显差异，即平台的信用评级对申请人的信用风险有一定的识别能力。

表3 大数据征信对平台授信通过概率的影响

Tab.3 The impact of big-data credit reporting on the probability of platform credit approval

变量	授信是否通过					
	(1)	(2)	(3)	(4)	(5)	(6)
是否绑定信用卡	0.971*** (0.012)					0.325*** (0.013)
累计消费金额		0.261*** (0.001)				0.224*** (0.002)
消费券使用频次			0.551*** (0.027)			0.289*** (0.028)
平台使用时长				0.145*** (0.001)		0.029*** (0.001)
信用卡支付比重					3.163*** (0.146)	-0.274** (0.138)
年龄	1.797*** (0.021)	1.842*** (0.023)	1.908*** (0.021)	1.671*** (0.022)	1.911*** (0.021)	1.743*** (0.023)
年龄平方项	-0.226*** (0.003)	-0.232*** (0.004)	-0.238*** (0.003)	-0.205*** (0.004)	-0.239*** (0.003)	-0.219*** (0.004)

性别	0.732*** (0.007)	0.645*** (0.008)	0.736*** (0.007)	0.676*** (0.008)	0.737*** (0.007)	0.643*** (0.008)
是否东部地区	0.249*** (0.692)	0.125*** (0.008)	0.270*** (0.007)	0.134*** (0.008)	0.267*** (0.007)	0.111*** (0.008)
常数项	5.119*** (7.396)	6.421*** (0.691)	5.051*** (0.677)	5.215*** (0.676)	5.052*** (0.676)	6.251*** (0.697)
时间固定效应	控制	控制	控制	控制	控制	控制
样本量	422,246	422,246	422,246	422,246	422,246	422,246
R2	0.094	0.202	0.082	0.164	0.082	0.206

注：表内数字均为变量的回归系数，对应括号内均为稳健标准误。\*、\*\*和\*\*\*分别表示在 10%、5%和 1%水平上显著（下同）。

## （二）进一步分析：大数据征信与传统征信的替代效应

为了进一步探究传统征信信息和消费行为信息之间的关系，本文构建“是否绑定信用卡”与消费行为信息做交互项，进行（2）式回归，结果如表 4 所示。列（1）表明，“是否绑定信用卡”与累计消费金额的交互项回归系数在 1%水平下显著为正，意味着消费金额并不会削弱是否绑定信用卡这一传统征信信息在平台授信决策中的权重，反而会增加其在授信过程中的重要性，两个信息互为补充。这可能是因为作为交易数据，消费金额本质上同样属于与财务相关的信息。列（2）～（3）分别为将用户“消费券使用频次”“平台使用时长”替换“累计消费金额”放入交互模型的回归结果，发现消费券使用频次和平台使用时长均不会影响传统征信（即是否绑定信用卡）对授信通过概率的影响。第（4）列为将累计消费金额中使用信用卡支付的比例加入交互模型的实证结果，发现交互项系数为-1.034，且在 1%的水平下显著，说明信用卡支付比重会减少传统征信在平台评估用户授信过程中的权重，形成替代作用。列（5）汇报了将“是否绑定信用卡”与累计消费金额、消费券使用频次、平台使用时长以及信用卡支付比例等消费行为信息的交互项共同纳入模型中的回归结果，发现平台使用时长会减少传统征信在平台授信评估中的权重，即对于平台使用时间越长的用户，传统征信信息对平台是否对用户授信的影响减弱，两者形成替代关系。

表4 大数据征信和传统征信之间关系

Tab.4 The relationship between big-data credit reporting and traditional credit reporting

变量	授信是否通过				
	(1)	(2)	(3)	(4)	(5)
是否绑定信用卡	0.142*** (0.024)	0.324*** (0.013)	0.322*** (0.016)	0.329*** (0.013)	0.140*** (0.024)
是否绑定信用卡*累计消费金额	0.057*** (0.006)				0.076*** (0.007)
是否绑定信用卡*消费券使用频次		0.089			0.080

		(0.065)		(0.065)	
是否绑定信用卡*平台使用时长			0.001	-0.016***	
			(0.003)	(0.003)	
是否绑定信用卡*信用卡支付比重				-1.034***	-1.252***
				(0.317)	(0.318)
常数项	5.077***	6.220***	6.076***	6.222***	4.925***
	(0.695)	(0.697)	(0.697)	(0.697)	(0.696)
消费行为信息	控制	控制	控制	控制	控制
控制变量	控制	控制	控制	控制	控制
时间固定效应	控制	控制	控制	控制	控制
样本量	422,246	422,246	422,246	422,246	422,246
R <sup>2</sup>	0.206	0.206	0.206	0.206	0.206

注：消费行为信息变量包括累计消费金额、消费券使用频次、平台使用时长、信用卡支付比重，下表同。

### （三）稳健性检验

#### 1. 变量替换

由于授信额度为连续型变量，因此本文采取多元线性回归模型进行稳定性检验。具体模型如下：

$$supply\_limit_{i,t} = \alpha + \beta information_{i,t} + \gamma X_{i,t} + \delta_t + \varepsilon_{i,t} \quad (3)$$

其中， $supply\_limit_{i,t}$  是平台授信额度，其他变量与基准模型一致。

首先，本文将平台授信额度作为被解释变量，将是否绑定信用卡、累计消费金额、消费券使用频次、平台使用时长和信用卡支付比重作为核心解释变量依次加入式（3），结果如表5第（1）～（6）列所示，在平台授信额度审核上，以“是否绑定信用卡”衡量的传统征信和消费行为信息均能够帮助平台识别低风险用户。平台更青睐于获得传统征信授信、在平台累计消费金额高、消费券使用频繁、在平台活跃时间长的用户，并给予更高的授信额度。

表5 稳健性检验：大数据征信对平台授信额度的影响

Tab.5 Robustness: The impact of big-data credit reporting on the probability of platform credit limits

变量	授信额度					
	(1)	(2)	(3)	(4)	(5)	(6)
是否绑定信用卡	1.936*** (0.023)					0.819*** (0.022)
累计消费金额		0.355*** (0.001)				0.263*** (0.002)
消费券使用频次			0.952*** (0.030)			0.497*** (0.024)
平台使用时长				0.244*** (0.001)		0.078*** (0.002)

信用卡支付比重					6.323***	0.354
					(0.299)	(0.261)
常数项	-3.341***	-4.244***	-3.350***	-3.045***	-3.350***	-3.880***
	(1.040)	(0.981)	(0.961)	(0.966)	(0.961)	(1.013)
控制变量	控制	控制	控制	控制	控制	控制
时间固定效应	控制	控制	控制	控制	控制	控制
样本量	422,246	422,246	422,246	422,246	422,246	422,246
R2	0.118	0.227	0.100	0.201	0.100	0.235

进一步，本文将平台授信额度作为被解释变量，将“是否绑定信用卡”与消费行为信息做交互项，进行（4）式回归，结果如表6第（1）～（5）列所示。本文发现，除列（5）外大部分回归结果与表4回归结果基本保持一致，表明前述“传统征信信息与消费行为信息相互补充”结果稳健。

$$supply\_limit_{i,t} = \alpha + \beta_1 ccardin f_{i,t} + \beta_2 consuminf_{i,t} + \beta_3 ccardin f_{i,t} * consuminf_{i,t} + \gamma X_{i,t} + \delta_t + \varepsilon_{i,t} \quad (4)$$

表6 稳健性检验：大数据征信和传统征信之间关系

Tab.6 Robustness: The relationship between big-data credit reporting and traditional credit reporting

变量	授信额度				
	(1)	(2)	(3)	(4)	(5)
是否绑定信用卡	0.289*** (0.027)	0.819*** (0.020)	0.680*** (0.023)	0.827*** (0.020)	0.294*** (0.027)
是否绑定信用卡*累计消费金额	0.187*** (0.777)				0.224*** (0.008)
是否绑定信用卡*消费券使用频次		0.049 (0.067)			(0.019) (0.067)
是否绑定信用卡*平台使用时长			0.042*** (0.004)		-0.030*** (0.005)
是否绑定信用卡*信用卡支付比重				-2.535*** (0.501)	-3.499*** (0.502)
常数项	-2.504*** (0.871)	-3.804*** (0.872)	-3.449*** (0.872)	-3.805*** (0.872)	-2.099** (0.871)
消费行为信息	控制	控制	控制	控制	控制
控制变量	控制	控制	控制	控制	控制
时间固定效应	控制	控制	控制	控制	控制
样本量	422,246	422,246	422,246	422,246	422,246
R <sup>2</sup>	0.237	0.235	0.236	0.235	0.237

## 2. 模型替换

本文选择 Logit 模型作为基准信用评分模型。王正位等<sup>[2]</sup>在利用某现金贷平台的借款人数据，验证

不同信息的信用评估方法是否能有效预测现金贷借款人的信用风险时，选择使用 probit 模型。鉴于此，利用 probit 模型对基准研究结果进行稳健性检验，回归结果如表 7 所示，与 Logit 模型回归结果保持一致，说明文章结论具有稳健性。

表7 稳健性检验：大数据征信对平台授信通过概率的影响（probit模型）

Tab.7 Robustness: The impact of big-data credit reporting on the probability of platform credit approval (probit)

变量	授信是否通过					
	(1)	(2)	(3)	(4)	(5)	(6)
是否绑定信用卡	0.589*** (0.007)					0.205*** (0.008)
累计消费金额		0.146*** (0.001)				0.123*** (0.001)
消费券使用频次			0.336*** (0.016)			0.176*** (0.016)
平台使用时长				0.086*** (0.000)		0.019*** (0.001)
信用卡支付比重					1.865*** (0.083)	-0.119 (0.084)
常数项	2.945*** (0.368)	3.620*** (0.377)	2.947*** (0.371)	2.979*** (0.371)	2.947*** (0.371)	3.501*** (0.375)
控制变量	控制	控制	控制	控制	控制	控制
时间固定效应	控制	控制	控制	控制	控制	控制
样本量	422,246	422,246	422,246	422,246	422,246	422,246
R2	0.093	0.199	0.082	0.164	0.082	0.203

同样的，利用 probit 模型对大数据征信与传统征信的关系进行稳健性检验，表 8 展示了回归结果，与表 4 结果保持一致，说明大数据征信信息与传统征信信息之间的关系具有稳健性。

表8 稳健性检验：大数据征信和传统征信之间关系（probit模型）

Tab.8 Robustness: The relationship between big-data credit reporting and traditional credit reporting (probit)

变量	授信是否通过				
	(1)	(2)	(3)	(4)	(5)
是否绑定信用卡	0.094*** (0.013)	0.205*** (0.008)	0.322*** (0.016)	0.207*** (0.008)	0.094*** (0.013)
是否绑定信用卡*累计消费金额	0.036*** (0.003)				0.048*** (0.004)
是否绑定信用卡*消费券使用频次		0.050 (0.037)			0.042 (0.037)
是否绑定信用卡*平台使用时长			0.001		-0.010***

			(0.003)		(0.002)
是否绑定信用卡*信用卡支付比重				-0.662***	-0.819***
				(0.193)	(0.194)
常数项	2.859***	3.482***	6.076***	3.482***	2.758***
	(0.375)	(0.375)	(0.697)	(0.375)	(0.375)
消费行为信息	控制	控制	控制	控制	控制
控制变量	控制	控制	控制	控制	控制
时间固定效应	控制	控制	控制	控制	控制
样本量	422,246	422,246	422,246	422,246	422,246
R <sup>2</sup>	0.203	0.203	0.206	0.203	0.203

### 3. 排除疫情期间样本

在基准模型中，本文使用样本时间覆盖 2018 年 3 月~2021 年 12 月，包含受新冠肺炎疫情影响的时间段。新冠肺炎疫情是全球性重大突发公共卫生事件，疫情暴发初期，我国果断采取严密防控措施，企业停工停产、居民居家管制，信贷供需端均受到不同程度的影响。鉴于此，筛去疫情期间样本，对基准研究结果进行稳健性检验，结果如表 9 所示，分别与表 3 的回归结果保持一致，说明基准结果是稳健的。

表9 稳健性检验：大数据征信对平台授信通过概率的影响（排除疫情期间样本）

Tab.9 Robustness: The impact of big-data credit reporting on the probability of platform credit approval (excluding samples from the Covid-19 period)

变量	授信是否通过					
	(1)	(2)	(3)	(4)	(5)	(6)
是否绑定信用卡	0.974*** (0.013)					0.334*** (0.015)
累计消费金额		0.266*** (0.002)				0.217*** (0.002)
消费券使用频次			0.545*** (0.037)			0.244*** (0.040)
平台使用时长				0.183*** (0.001)		0.046*** (0.002)
信用卡支付比重					3.231*** (0.161)	-0.183 (0.151)
常数项	5.034*** (0.690)	6.424*** (0.692)	4.992*** (0.674)	5.351*** (0.680)	4.985*** (0.673)	6.240*** (0.698)
控制变量	控制	控制	控制	控制	控制	控制
时间固定效应	控制	控制	控制	控制	控制	控制
样本量	289,205	289,205	289,205	289,205	289,205	289,205

R2	0.087	0.196	0.072	0.163	0.073	0.200
----	-------	-------	-------	-------	-------	-------

同样的，本文筛去疫情期间样本后，对大数据征信与传统征信的替代效应进行稳健性检验，回归结果如表 10 所示，从第（5）列的结果可以看出，是否绑定信用卡与消费券使用频次的交互项系数在 10%的水平上为负，说明消费券使用频次所包含的信息量会降低传统征信对授信是否通过的解释力度，形成替代关系。

表10 稳健性检验：大数据征信和传统征信之间关系（排除疫情期间样本）

Tab.10 Robustness: The relationship between big-data credit reporting and traditional credit reporting (excluding samples from the Covid-19 period)

变量	授信是否通过				
	(1)	(2)	(3)	(4)	(5)
是否绑定信用卡	0.094*** (0.013)	0.205*** (0.008)	0.322*** (0.016)	0.207*** (0.008)	0.094*** (0.013)
是否绑定信用卡*累计消费金额	0.071*** (0.007)				0.084*** (0.008)
是否绑定信用卡*消费券使用频次		-0.144* (0.085)			-0.165* (0.085)
是否绑定信用卡*平台使用时长			0.014*** (0.004)		-0.012*** (0.005)
是否绑定信用卡*信用卡支付比重				-0.689** (0.342)	-0.954*** (0.344)
常数项	5.097*** (0.696)	6.208*** (0.698)	5.985*** (0.697)	6.210*** (0.698)	4.855*** (0.696)
消费行为信息	控制	控制	控制	控制	控制
控制变量	控制	控制	控制	控制	控制
时间固定效应	控制	控制	控制	控制	控制
样本量	289,205	289,205	289,205	289,205	289,205
R <sup>2</sup>	0.201	0.200	0.200	0.200	0.201

#### 4. 增加控制变量

为进一步验证文章结论的稳健性，本文在原有模型基础上控制了用户的婚姻状况、受教育水平以及所在城市的固定效应。其中，婚姻状况为虚拟变量，如果用户已婚，则为 1，如果用户为单身或者离异，则为 0。受教育水平为有序多项变量，初中及以下受教育水平为 1，中等职业教育（职高）为 2，普通高中为 3，大专为 4，本科为 5，研究生及以上为 6。实证结果如表 11 和 12 所示，表 11 结果与表 3 基准回归结果保持一致；表 12 的结果则进一步表明，在平台授信决策过程中，消费金额作为财务信息的替代，是对传统征信信息的补充。



表11 稳健性检验：大数据征信对平台授信通过概率的影响（增加控制变量）

Tab.11 Robustness: The impact of big-data credit reporting on the probability of platform credit approval (adding control variables)

变量	授信是否通过					
	(1)	(2)	(3)	(4)	(5)	(6)
是否绑定信用卡	0.450*** (0.029)					0.097*** (0.031)
累计消费金额		0.143*** (0.003)				0.129*** (0.004)
消费券使用频次			0.844*** (0.044)			0.685*** (0.045)
平台使用时长				0.081*** (0.002)		0.009*** (0.003)
信用卡支付比重					1.485*** (0.307)	0.086 (0.322)
常数项	4.740*** (1.349)	4.756*** (1.368)	4.667*** (1.343)	4.191*** (1.354)	4.764*** (1.344)	4.587*** (1.368)
控制变量	控制	控制	控制	控制	控制	控制
时间固定效应	控制	控制	控制	控制	控制	控制
城市固定效应	控制	控制	控制	控制	控制	控制
样本量	93,042	93,042	93,042	93,042	93,042	93,042
R2	0.031	0.067	0.034	0.053	0.029	0.071

表12 稳健性检验：大数据征信和传统征信之间关系（增加控制变量）

Tab.12 Robustness: The relationship between big-data credit reporting and traditional credit reporting (adding control variables)

变量	授信是否通过				
	(1)	(2)	(3)	(4)	(5)
是否绑定信用卡	-0.059 (0.051)	0.088*** (0.031)	0.031 (0.036)	0.096*** (0.031)	-0.061 (0.051)
是否绑定信用卡*累计消费金额	0.053*** (0.013)				0.132*** (0.004)
是否绑定信用卡*消费券使用频次		0.217* (0.111)			0.188* (0.112)
是否绑定信用卡*平台使用时长			0.023*** (0.006)		0.012 (0.007)
是否绑定信用卡*信用卡支付比重				0.177 (0.735)	0.027 (0.739)
常数项	2.457***	3.101***	3.082***	3.105***	2.410***

	(1.184)	(1.184)	(1.185)	(1.185)	(1.184)
消费行为信息	控制	控制	控制	控制	控制
控制变量	控制	控制	控制	控制	控制
时间固定效应	控制	控制	控制	控制	控制
城市固定效应	控制	控制	控制	控制	控制
样本量	93,042	93,042	93,042	93,042	93,042
R <sup>2</sup>	0.071	0.071	0.071	0.071	0.071

#### （四）异质性分析

为进一步探究消费行为信息对平台授信是否通过的异质性影响，本文基于地区差异与年龄差异进行实证分析。① 根据消费信贷申请人的所在地区不同，探究消费信贷申请人的地区差异（是否在东部地区）如何导致平台授信的差异；② 根据消费信贷申请人年龄，将样本分为 25 岁以下、25~35 岁、35~45 岁、45~55 岁、55~65 岁、65 以上等 6 组，探究不同年龄层的消费信贷申请人在平台授信中的特征。

由于不同地区和不同年龄段的消费信贷申请人拥有的个人财富、面临的流动性约束不同，平台在授信决策上也会存在差异。从地区差异来看，一直以来，我国区域发展不均衡，东中西部地区金融发展程度存在显著差异，中西部地区居民收入水平较低，消费金融覆盖度较低。表 13 第（1）列将居住地区与传统征信及大数据征信变量做交互项，这些交互项系数均不显著，说明传统征信与大数据征信并不存在地区差异，这也在一定程度上验证了金融科技平台在对用户授信时并不存在地域歧视。从年龄差异来看，如表 13 第（2）列所示，是否绑定信用卡、消费券使用频次和平台使用时长等 3 个变量与年龄的交互项均显著为负，表明消费信贷申请人年龄层次越小，这 3 个大数据征信信息对授信是否通过的正向影响越显著。一个可能的解释在于，年轻群体是信用卡、（电子）消费券等产品的主要使用群体，其对互联网依赖性更高，用网时间越长。因此，互联网消费平台更注重对年轻人的传统征信情况、平台依赖度等信息的考察。相反，平台对年龄较大用户授信时，会加大累计消费金额在决策中的权重，反而会减少其他消费行为在决策中的权重，这可能是因为年龄较大用户对互联网熟悉程度较低，因此在平台上留下除消费行为外的其他行为足迹较少。

表13 异质性分析：大数据征信对平台授信通过概率的影响

Tab.13 Heterogeneity: The impact of big-data credit reporting on the probability of platform credit approval		
变量	授信是否通过	授信是否通过
	(1)	(2)
是否东部地区*是否绑定信用卡	-0.012 (0.029)	
是否东部地区*累计消费金额	0.002 (0.004)	
是否东部地区*消费券使用频次	0.017 (0.082)	

是否东部地区*平台使用时长	0.000 (0.004)	
是否东部地区*信用卡支付占比	-0.156 (0.304)	
年龄*是否绑定信用卡		-0.029* (0.017)
年龄*累计消费金额		0.034*** (0.002)
年龄*消费券使用频次		-0.126*** (0.048)
年龄*平台使用时长		-0.007*** (0.002)
年龄*信用卡支付占比		0.248 (0.171)
常数项	4.834*** (0.697)	0.577 (0.698)
是否绑定信用卡	控制	控制
消费行为信息	控制	控制
控制变量	控制	控制
时间固定效应	控制	控制
样本量	289,205	289,205
R <sup>2</sup>	0.200	0.201

#### 四、拓展讨论：重大突发公共卫生事件冲击下的平台授信决策

2020 年伊始，新冠肺炎疫情不断蔓延并逐渐席卷全球，重大突发公共卫生事件往往会对区域乃至全球经济、社会、文化造成重大冲击。早期关于西欧“黑死病”、伦敦霍乱事件等重大疫情对经济社会与个人福利的影响已有丰富的研究<sup>[48-49]</sup>。21 世纪以来，重大突发公共卫生事件在全球各国时有发生。例如，2003 年“非典型肺炎”疫情、2011 年日本大地震及其引发的福岛核电站泄漏、2014 年埃博拉病毒以及至今仍在肆虐全球的新冠疫情等。与其他突发公共卫生事件相比，新冠疫情对经济的负向冲击具有诸多共同特征。如，率先影响服务行业，改变消费者心理预期和行为模式，进而对内需等方面产生外溢性和连锁影响。Worthington<sup>[50]</sup>指出，突发公共事件除对受影响地区造成直接损害外，供应链中断也会使得风险转移至原本未受影响的区域，直接或间接造成重大经济损失。杨子晖等<sup>[51]</sup>发现，重大突发公共事件所特有的紧迫性与不确定性为全球经济发展带来极大挑战。消费者信心遭受冲击，交通运输严重受阻，从而引发居民恐慌情绪与悲观预期，消费意愿大幅减弱，同时会在各部门之间产生明显的风险共振和风险溢出效应。

新冠疫情暴发以来，国内外学者已经围绕着疫情对社会经济与贸易<sup>[52-54]</sup>、劳动力市场<sup>[55-57]</sup>、金融市场<sup>[51,58-59]</sup>、居民预期<sup>[60-61]</sup>以及经济政策不确定性<sup>[62-63]</sup>等方面展开深入研究。作为平滑居民收入和消费支出的

消费信贷，在疫情暴发以来也引起学者广泛关注<sup>[64-65]</sup>。事实上，有关危机对于信贷冲击影响的研究主要围绕“债务-通缩”理论<sup>[66]</sup>，以及此后的金融加速器理论<sup>[67-69]</sup>。上述理论认为金融系统面对重大危机冲击时，往往存在明显的顺周期性，并会放大经济冲击。从贷款人角度来看，疫情冲击使得银行等信贷供给方资产负债表收缩，影响其提供信贷的意愿和能力，社会信贷投放总量最终有所下降<sup>[70]</sup>。金融科技平台的信贷资本主要来源往往是商业银行，因此金融科技平台的借贷资本会受到银行信贷收缩的影响，最终导致其在放贷过程中表现出一定的趋同性。从借款人角度来看，疫情会明显影响借款人的财富水平、收入水平及未来收入预期，对其还款能力造成严重损害。首先，疫情期间，社会经济活动停滞，大量企业停工停产，中小企业发展尤其艰难，就业岗位减少，失业率增加，就业形势严峻，消费者就业信心普遍降低<sup>[60]</sup>。这不仅会影响消费者当下的收入水平，也会显著影响居民未来收入预期<sup>[60]</sup>。消费信贷是将未来收入折现后用于当下消费的金融手段，因此收入预期的下降将损害借款人的还款能力和意愿，增加其信用风险。因此，贷款人会减少放贷行为。

但另一方面，互联网平台与传统商业银行在信贷市场中存在一定的竞争关系，这会导致金融科技平台出于获客动机，增加对借款人授信概率。疫情冲击下，消费者可能会产生新的信贷需求。根据反周期信贷需求理论，在经济下行，收入水平降低情况下，消费者往往会通过申请消费信贷来满足消费需求<sup>[72]</sup>。面对暂时性负向收入冲击，借款者倾向于使用无担保消费信贷来平滑消费<sup>[73]</sup>。疫情期间，互联网消费信贷凭借其低门槛、无担保、覆盖广、方便快捷等特征，特别是其独特的非接触式优势，在疫情冲击下其信贷需求和供给相较于传统消费信贷更具稳定性<sup>[74-75]</sup>。此外，相比于商业银行依赖的传统征信，金融科技平台的大数据征信能够弱化收入水平、财富水平、财务数据等传统征信信息在疫情冲击中产生的波动，提高信贷风险的识别精度。在此情况下，互联网消费平台可能会凭借自身在信用评估效率上的优势，适当增加信贷供给，提升信贷市场占有率，一定程度上缓解危机时期的顺周期性<sup>[3,76]</sup>。为进一步探索重大突发公共卫生事件是如何影响传统征信和大数据征信在互联网消费平台授信决策中的作用，本文通过模型（5）开展实证分析。

$$\Pr(\text{supplysucc}_{i,t}=1) = \alpha + \beta_1 \text{yiqing}_{i,t} + \beta_2 \text{information}_{i,t} + \beta_3 \text{yiqing}_{i,t} * \text{information}_{i,t} + \gamma X_{i,t} + \delta + \varepsilon_{i,t} \quad (5)$$

其中， $\text{yiqing}_{i,t}$  是疫情冲击的哑变量，若申请时间在 2020 年第一季度之前，则  $\text{yiqing}_{i,t} = 0$ ，反之则  $\text{yiqing}_{i,t} = 1$ 。其他变量设定与前文一致。 $\text{information}_{i,t}$  为消费者的征信信息，包括传统征信信息（ $\text{ccardinf}_{i,t}$ ）和消费行为信息（ $\text{consuminf}_{i,t}$ ）。我们着重关注  $\beta_2 + \beta_3 \text{yiqing}_t$ ，即征信信息对授信通过概率的影响，也会受到  $\beta_3 \text{yiqing}_t$  这一变量的影响。疫情冲击前，即  $\text{yiqing}_t = 0$  时，征信信息对授信通过概率的影响为  $\beta_2$ ，疫情冲击后，即  $\text{yiqing}_t = 1$  时，征信信息对授信通过概率的影响为  $\beta_2 + \beta_3$ ，如果  $\beta_3 < 0$ ，则说明疫情冲击会削弱传统征信和大数据征信在平台授信决策过程中的权重，即使消费者的传统征信信息和大数据征信信息与疫情前完全相同，平台授信通过的概率也会下降，这也意味着新冠疫情冲击下，平台会表现出顺周期性，收紧授信风险控制，降低授信通过概率。

回归结果如表 14 所示。在控制疫情变量的基础上，列（1）～（2）、（4）～（5）可以看出疫情冲击分别与“是否绑定信用卡”“累计消费金额”“平台使用时长”“信用卡支付比重”交互项系数显著为负。但当把这些交互项一并放入模型中，只有疫情冲击与“平台使用时长”的交互项系数在 1%的水平下显著为负，这说明疫情冲击并不会显著影响传统征信信息以及大数据征信信息中与交易相关信息在平台授信决策过程中的权重，但会显著影响用户黏性在授信决策中的权重，因此疫情冲击下在线旅游平台一定程度上表现为顺周期。

表14 疫情冲击对平台授信通过概率的影响

Tab.14 The impact of the COVID-19 shock on platform credit approval probability

变量	授信是否通过					
	(1)	(2)	(3)	(4)	(5)	(6)
疫情冲击*是否绑定信用卡	-0.125*** (0.030)					-0.044 (0.031)
疫情冲击*累计消费金额		-0.027*** (-0.003)				0.004 (0.004)
疫情冲击*消费券使用频次			0.066 (0.057)			0.096* (0.057)
疫情冲击*平台使用时长				-0.025*** (0.002)		-0.027*** (0.002)
疫情冲击*信用卡支付占比					-0.967*** (0.359)	-0.581 (0.364)
常数项	5.870*** (0.478)	4.761*** (0.479)	5.890*** (0.478)	5.689*** (0.479)	5.746*** (0.478)	4.541*** (0.479)
疫情冲击	控制	控制	控制	控制	控制	控制
是否绑定信用卡	控制	控制	控制	控制	控制	控制
消费行为信息	控制	控制	控制	控制	控制	控制
控制变量	控制	控制	控制	控制	控制	控制
时间固定效应	控制	控制	控制	控制	控制	控制
样本量	422,246	422,246	422,246	422,246	422,246	422,246
R <sup>2</sup>	0.206	0.206	0.206	0.206	0.206	0.206

虽然我国在 2020 年～2021 年一直处于疫情常态化阶段，但由于季节更替、防护措施落地、疫苗接种等原因，疫情在部分阶段得到良好控制，也在部分阶段出现爆发。为了更精确地模拟疫情冲击的严重程度，本文使用新冠疫情的百度搜索指数替换疫情冲击哑变量。互联网用户对疫情相关内容搜索量增加，表明对疫情的关注程度上升。一方面，居民要及时更新当地确诊人员的行程轨迹，及时防范时空交叠风险；另一方面，由于疫情导致的封控措施，会影响居民的出行、工作、就医等，因此，新冠疫情的百度搜索指数也在一定程度体现疫情的严重程度。在用疫情百度搜索指数替代疫情冲击后，回归结果如表 15 所示，与表

14 基本保持一致。

表15 疫情冲击对平台授信通过概率的影响：使用百度搜索指数

Tab.15 The impact of the COVID-19 shock on platform credit approval probability: using Baidu Index

变量	授信是否通过					
	(1)	(2)	(3)	(4)	(5)	(6)
疫情冲击（百度搜索指数）*是否绑定信用卡	-0.010* (0.005)					-0.003 (0.005)
疫情冲击（百度搜索指数）*累计消费金额		-0.003*** (0.000)				-0.001 (0.001)
疫情冲击（百度搜索指数）*消费券使用频次			0.005 (0.012)			0.008 (0.012)
疫情冲击（百度搜索指数）*平台使用时长				-0.003*** (0.000)		-0.002*** (0.000)
疫情冲击（百度搜索指数）*信用卡支付占比					-0.057 (0.058)	-0.023 (0.059)
常数项	6.084*** (86.919)	4.979*** (73.409)	6.104*** (87.575)	6.018*** (84.643)	6.109*** (87.647)	5.296*** (1.168)
疫情百度搜索	控制	控制	控制	控制	控制	控制
是否绑定信用卡	控制	控制	控制	控制	控制	控制
消费行为信息	控制	控制	控制	控制	控制	控制
控制变量	控制	控制	控制	控制	控制	控制
时间固定效应	控制	控制	控制	控制	控制	控制
样本量	198,416	198,416	198,416	198,416	198,416	198,416
R2	0.206	0.206	0.205	0.206	0.205	0.206

## 五、结论

本文从国内某旅游在线平台随机抽取了 50 万活跃用户的数据，实证分析了传统征信与大数据征信在互联网平台授信决策中的信息价值。研究发现：

（1）大数据征信在互联网平台授信决策中提供了信息价值，消费信贷申请人的消费能力、消费习惯以及用户黏性均被纳入授信考量，累计消费金额越高、消费券使用越频繁、平台使用越久的消费者，获得平台授信的可能性越大。

（2）传统征信信息和大数据征信信息在互联网平台的授信决策过程中以替代关系为主。平台使用时长、信用卡支付比重等消费行为信息对传统征信信息产生了替代，但一定程度上反映财务现状的交易信息，如累计消费金额，则与传统征信信息之间互为补充。

（3）从异质性分析结果来看，一方面，不同于传统商业银行，平台授信过程中并不存在地域歧视。

另一方面，在线旅游平台更注重对年轻人传统征信情况、平台依赖度等信息的考察。相反，平台对年龄较大用户授信时，会加大累计消费金额在决策中的权重，反而会减少其他消费行为在决策中的权重，这可能是年龄较大用户对互联网熟悉程度较低，因此在平台上留下除消费行为外的其他行为足迹较少所致。这很大程度上验证了互联网消费信贷的普惠特征。

（4）重大突发公共卫生事件会降低大数据征信信息在平台授信决策中的权重，风险控制加强，授信通过概率降低，表现出顺周期性。

本文的研究结论，一方面为揭示大数据征信与传统征信之间的相互作用机制提供了独特视角，丰富了关于多源信息融合与征信机制的理论讨论；另一方面则强调了互联网平台授信过程中针对不同人群的灵活性和包容性，凸显出平台授信机制在普惠性上的制度潜力。不仅为信贷风险管理提供了新的理论基础和微观证据，更为互联网消费信贷的规范有序发展提供了有益的经验证据和政策启示。

（1）鼓励和引导互联网平台充分利用自身优势构建核心能力，服务于我国不断增长的消费信贷需求。互联网平台依托于大数据优势，利用海量、高频的微观消费行为数据，对用户进行信用风险识别，可扩大消费信贷服务的覆盖广度，惠及下沉人群，有助于缩小不同人群间因信用识别不完全导致的信贷不平等，进一步释放居民消费潜力。尤其是在传统金融机构难以触及的农村和偏远地区，互联网平台通过网络与这部分用户群建立联系，为其合理的消费信贷需求提供有效支撑。

（2）要注重消费行为信息在消费信贷评估中运用。传统金融机构凭借个人信用历史和单一机构内部积累的客户数据，难以实现信贷服务群体的有效扩张。尤其面临外部冲击时，基于银行授信审核体系评估结果，反而会使得市场流动性收紧。消费行为信息作为传统征信信息的补充，能够对消费信贷申请人进行多角度精准刻画，在一定程度上缓解信用记录问题，有效弥补传统征信方法的不足。

## 参考文献

- [1] 张晶, 张喆, 方匡南, 等. 基于稀疏结构连续比率模型的消费金融风控研究[J]. 统计研究, 2020, 37(11): 57-67.  
ZHANG J, ZHANG Z, FANG K N, *et al.* Sparse structural continuation ratio model with its application in consumer finance risk control[J]. Statistical Research, 2020, 37(11): 57-67.
- [2] 王正位, 周从意, 廖理, 等. 消费行为在个人信用风险识别中的信息含量研究[J]. 经济研究, 2020, 55(1): 149-163.  
WANG Z W, ZHOU C Y, LIAO L, *et al.* Informational content of consumption behavior in consumer credit risk evaluation[J]. Economic Research Journal, 2020, 55(1): 149-163.
- [3] GIRARDONE C, NIERI L, PISERA S, SANTULLI R. Does fintech credit affect firms' cost of capital and capital structure? [J]. The European Journal of Finance, 2024, 30 (10): 1-21.
- [4] HAU H, HUANG Y, LIN C, SHAN H Z, SHENG Z X, WEI L. Fintech credit, financial inclusion and entrepreneurial growth[J]. The Journal of Finance, 2024, 79(5): 52.
- [5] PHILIPPON T. On fintech and financial inclusion[R]. Bank for International Settlements Working Paper, 2020: 841.
- [6] LIN M, PRABHALA N R, VISWANATHAN S. Judging borrowers by the company they keep: Friendship networks and information asymmetry in online peer-to-peer lending[J]. Management Science, 2013, 59(1): 17-35.
- [7] FROST J, GAMBACORTA L, HUANG Y, *et al.* Big tech and the changing structure of financial intermediation[R]. Bank for International Settlements Working Paper, 2019: 779.
- [8] 黄益平, 邱晗. 大科技信贷: 一个新的信用风险管理框架[J]. 管理世界, 2021, 37(2): 12-21.  
HUANG Y P, QIU H. Big tech lending: A new credit risk management framework[J]. Journal of Management World, 2021, 37(2): 12-21.
- [9] 邱涌钦, 方匡南, 张庆昭, 等. 高维主动 PU 学习及其在信用评分中的应用[J]. 中国管理科学, 2025, 33(9): 57-66.  
QIU Y Q, FANG K N, ZHANG Q Z, *et al.* High-dimensional active PU learning with its application to credit scoring[J]. Chinese Journal of Management Science, 2025, 33(9): 57-66.
- [10] ÓSKARSDÓTTIR M, BRAVO C, SARRAUTE C, Vanthienen J, Baesens B. The value of big data for credit scoring: Enhancing financial inclusion using mobile phone data and social network analytics[J]. Applied Soft Computing, 2019, 74: 26-39.
- [11] WEI Y, YILDIRIM P, BULTE C V, *et al.* Credit scoring with social network data[J]. Marketing Science, 2016, 35(2): 234-258.
- [12] BERG T, VALENTIN B, ANA G, *et al.* On the rise of fintechs: credit scoring using digital footprints[J]. Review of Financial Studies, 2020, 33(7): 2845-2897.
- [13] JAGTIANI J, CATHARINE L. Do fintech lenders penetrate areas that are underserved by traditional banks?[R]. FRB of Philadelphia Working Paper, 2018: 18-23.
- [14] THAKOR A V. Fintech and banking: What do we know?[J]. Journal of Financial Intermediation, 2020, 41(1): 1-13.
- [15] AMANOLLAHI G F. The influence of external factors on the credit risk in leasing industry[J]. Management Science Letters, 2016, 3(6): 251-258.
- [16] SKOGLUND J, CHEN W. The application of credit risk models to macroeconomic scenario analysis and stress testing[J]. Journal of Credit Risk, 2016, 12(2): 1-45.
- [17] VIVES X. Digital disruption in banking[J]. Annual Review of Financial Economics, 2019, 11(1): 243-272.
- [18] POPE D G, SYDNOR J R. What's in a picture? Evidence of discrimination from Prosper.com[J]. Journal of Human Resources, 2011, 46(1): 53-92.
- [19] 廖理, 向佳, 王正位. 网络借贷的角色转换与投资者学习效应[J]. 中国工业经济, 2018, 9: 60-78.  
LIAO L, XIANG J, WANG Z W. Role transition and investor learning in online lending[J]. China Industrial Economics, 2018, 9: 60-78.



- [20] IYER R, KHWAJA A I, LUTTMER E F, *et al.* Screening in new credit markets: Can individual lenders infer borrower credit worthiness in peer-to-peer lending?[R]. National Bureau of Economic Research Working Paper, 2009: 15242.
- [21] DUARTE J, SIEGEL L, YOUNG L. Trust and credit: The role of appearance in peer-to-peer lending[J]. Review of Financial Studies, 2012, 25: 2455-2484.
- [22] DE-CNUDDÉ S, MOEYERSOMS J, STANKOVA M, *et al.* What does your facebook profile reveal about your credit worthiness? Using alternative data for microfinance[J]. Journal of the Operational Research Society, 2018, 70: 1-11.
- [23] ANAMITRA D, ARJUNA C, MICHAEL K. Big data, small credit: The digital revolution and its impact on emerging market consumers[J]. Innovations, 2015, 10 (3-4): 49-80.
- [24] 王奇,晋晓妹,王子奇. 数字信用与中国家庭消费: 理论模型和经验证据[J]. 数量经济技术经济研究, 2025, 42 (3): 111-133.  
WANG Q, JIN X S, WANG Z Q. Digital credit and household consumption in china: theoretical models and empirical evidence[J]. Journal of Quantitative & Technological Economics, 2025, 42 (3): 111-133.
- [25] BESTER H. Screening vs. rationing in credit markets with imperfect information[J]. The American Economic Review, 1985, 75(4): 850-855.
- [26] BESANKO D, THAKOR A V. Collateral and rationing: Sorting equilibria in monopolistic and competitive credit markets[J]. International Economic Review, 1987, 28(3): 671-689.
- [27] JAFFEE D M, RUSSELL T. Imperfect information, uncertainty, and credit rationing[J]. Quarterly Journal of Economics, 1976, 90(4): 651-666.
- [28] STIGLITZ J E, WEISS A. Credit rationing in markets with imperfect information[J]. American Economic Review, 1981, 71(3): 393-410.
- [29] TSAI C, WU J. Using neural network ensembles for bankruptcy prediction and credit scoring[J]. Expert Systems with Applications: An International Journal Volume, 2008, 4(34): 2639-2649.
- [30] BABAEI G, GIUDICI P, RAFFINETTI E. Explainable finTech lending[J]. Journal of Economics and Business, 2023, 125-126 (5-6): 106126.
- [31] JAGTIANI J, CATHARINE L. The roles of alternative data and machine learning in fintech lending: Evidence from the lending club consumer platform[J]. Financial Management, 2019, 48(4): 1009-1029.
- [32] ZHU Y, WU D. P2P credit risk management with KG-GNN: A knowledge graph and graph neural network-based approach[J]. Journal of the Operational Research Society, 2024, 76(5): 1-15.
- [33] CORNELLI G, FROST J, GAMBACORTA L, RAU P R, WARDROP R, ZIEGLER T. Fintech and big tech credit: drivers of the growth of digital lending[J]. Journal of Banking & Finance, 2023, 148(3): 106742.
- [34] KOWALEWSKI O, PISANY P. Banks' consumer lending reaction to fintech and bigtech credit emergence in the context of soft versus hard credit information processing[J]. International Review of Financial Analysis, 2022, 81: 102-116.
- [35] TELYUKOVA I A, WRIGHT R. A model of money and credit, with application to the credit card debt puzzle[J]. Review of Economic Studies, 2008, 75(2): 629-647.
- [36] 廖理, 沈红波, 苏治. 如何推动中国居民的信用卡消费信贷——基于住房的研究视角[J]. 中国工业经济, 2013, 12: 117-129.  
LIAO L, SHEN H B, SU Z. How to promote credit card consumption credit for chinese residents: A research perspective based on housing[J]. China Industrial Economics, 2013, 12: 117-129.
- [37] 黄丹, 沈红波. 生命周期、消费者态度与信用卡使用频率[J]. 经济研究, 2010, 1: 108-117.  
HUANG H, SHEN H B. Life-cycle, consumer attitude and credit card use frequency[J]. Economic Research Journal, 2013, 12: 117-129.
- [38] 易行健, 周利. 数字普惠金融发展是否显著影响了居民消费——来自中国家庭的微观证据[J]. 金融研究, 2018, 11: 47-67.  
YI X J, ZHOU L. Does digital financial inclusion significantly influence household consumption? Evidence from household survey data in china[J]. Journal of Financial Research, 2018, 11: 47-67.

- [39] 宋科, 武沛璋, 李鸿翔, 等. 互联网消费信贷与传统消费信贷: 互补还是替代?[J]. 管理科学学报, 2023, 6(4): 41-61.  
SONG K, WU P Z, LI H X, *et al.* Internet consumer credit and traditional consumer credit: Complementation or substitution?[J]. Journal of Management Sciences in China, 2023, 26(4): 41-61.
- [40] 王红建, 曹晴, 曹瑜强. 金融科技赋能与实体经济修复——基于新冠疫情的冲击[J]. 南开管理评论, 2024, 27(6): 16-26.  
WANG H J, CAO Q, CAO Y Q. Fintech empowerment and real economy repair: Based on the covid-19 shock[J]. Nankai Business Review, 2024, 27(6): 16-26.
- [41] 章晟, 景辛辛. 新冠疫情冲击下投资者信心与宏观经济政策作用效果[J]. 中南财经政法大学学报, 2021, 6: 77-92.  
ZHANG S, JING X X. Covid-19 epidemic, investor confidence and the effect of macroeconomic policy[J]. Journal of Zhongnan University of Economics and Law, 2021(6), 77-92.
- [42] BAKER S R, BLOOM N, DAVIS S J. Measuring economic policy uncertainty[J]. Quarterly Journal of Economics, 2016, 131(4): 1593-1636.
- [43] MANELA A, MOREIRA A. News implied volatility and disaster concerns[J]. Journal of Financial Economics, 2017, 123(1): 137-162.
- [44] 尹志超, 吴雨, 甘犁. 金融可得性、金融市场参与和家庭资产选择[J]. 经济研究, 2015, 3: 87-99.  
YIN Z C, WU Y, GAN L. Financial availability, financial market participation and household portfolio Choice[J]. Economic Research Journal, 2015, 3: 87-99.
- [45] 李响, 杨雅鑫, 邹玥萌, 等. 金融知识、金融市场参与度与消费升级[J]. 金融论坛, 2023, 28 (12): 28-42.  
LI X, YANG Y X, ZOU Y M, *et al.* Financial literacy, household financial market participation and consumption upgrade[J]. Journal of Modern Finance, 2023, 28 (12): 28-42.
- [46] NIE G, WEI R, ZHANG L, *et al.* Credit ard churn forecasting by logistic regression and decision tree[J]. Expert Systems with Applications, 2011, 38(12): 15273-15285.
- [47] CHERRY S, JIANG E X, MATVOS G, *et al.* Government and private household debt relief during covid-19[J]. Brookings Papers on Economic Activity, 2021, 10: 141-199.
- [48] BRIDBURY A R. The black death[J]. Economic History Review, 1973, 26(4): 577-592.
- [49] AMBRUS A, FIELD E, GONZALEZ R. Loss in the time of cholera: Long-run impact of a disease epidemic on the urban landscape[J]. American Economic Review, 2020, 110: 475-525.
- [50] WORTHINGTON A C. The impact of natural events and disasters on the australian stock marker: A GARCH-M analysis of storms, floods, cyclones, earthquakes and bushfires[J]. Global Business and Economics Review, 2008, 10(1): 1-10.
- [51] 杨子晖, 陈雨恬, 张平森. 重大突发公共事件下的宏观经济冲击、金融风险传导与治理应对[J]. 管理世界, 2020, 36 (5): 13-35.  
YANG Z H, CHEN Y T, ZHANG P M. Macroeconomic shock, financial risk transmission and governance response to major public emergencies[J]. Journal of Management World, 2020, 36(5): 13-35.
- [52] 张友国, 孙博文, 谢锐. 新冠肺炎疫情的经济影响分解与对策研究[J]. 统计研究, 2021, 38(8): 68-82.  
ZHANG Y G, SUN B W, XIE R. Study on decomposition and counter measures of the economic impact of covid-19[J]. Statistical Research, 2021, 38(8): 68-82.
- [53] 刘洪铎, 张钊, 卢阳, 等. 新冠肺炎疫情对全球贸易的影响研究[J]. 统计研究, 2021, 38(12): 61-76.  
LIU H D, ZHANG N, LU Y, *et al.* A study on the impact of covid-19 on global trade[J]. Statistical Research, 2021, 38(12): 61-76.
- [54] 刘洪波, 邸建亮, 王冉. 新冠肺炎疫情对居民消费的影响研究[J]. 统计研究, 2022, 39(5): 38-48.  
LIU H B, DI J L, Wang R. The research of the impact of covid-19 on household consumption[J]. Statistical Research, 2022, 39 (5): 38-48.
- [55] FORSYTHE E, KAHN L B, LANGE F, WICZER D. Labor demand in the time of covid-19: Evidence from vacancy postings and UI claims[J]. Journal of Public Economics, 2020, 189 (9): 104238.

- [56] FARRELL D, GANONG P, GREIG F, LIEBESKING M, NOEL P, VAVRA J. Consumption effects of unemployment insurance during the covid-19 pandemic[J]. Social Science Electronic Publishing. Available at SSRN: <http://dx.doi.org/10.2139/ssrn.3654274>, 2020.
- [57] 蒋昉, 张丹丹, 刘雅玄. 新冠肺炎疫情对中国劳动力市场的影响——基于个体追踪调查的全面分析[J]. 经济研究, 2021, 56(2): 4-21.
- CAI F, ZHANG D D, LIU Y X. The impact of covid-19 on the chinese labor market——A comprehensive analysis based on the individual tracking survey[J]. Economic Research Journal, 2021, 56(2): 4-21.
- [58] 陈赞, 沈艳, 王靖一. 重大突发公共卫生事件下的金融市场反应[J]. 金融研究, 2020, 6: 20-39.
- CHEN Y SHEN Y, WANG J Y. Financial market reaction to dramatic public health shocks[J]. Journal of Financial Research, 2020, 6: 20-39.
- [59] 蒋涛. 疫情对企业融资的影响研究——来自银团贷款市场的经验证据[J]. 国际金融研究, 2020, 4: 65-75.
- JIANG T. The impact of epidemics on enterprises' funding——evidences from the syndicated loan market[J]. Studies of International Finance, 2020, 4: 65-75.
- [60] BINDER C. Coronavirus fears and macroeconomic expectations[J]. Review of Economics and Statistics, 2020, 102(4): 721-730.
- [61] COIBION O, GORODNICHENKO Y, WEBER M. The cost of the covid-19 crisis: lockdowns, macroeconomic expectations, and consumer spending[R]. NBER Working Paper, 2020: 27141.
- [62] PATA U K. Covid-19 induced economic uncertainty: A comparison between the united kingdom and the united states[J]. Available at SSRN: <https://ssrn.com/abstract=3637652>, 2020.
- [63] BAKER S R, BLOOM N, DAVIS S J, TERRY S J. Covid-induced economic uncertainty[R]. NBER Working Paper, 2020: 26983.
- [64] NORDEN L, MESQUITA D, WANG W. Covid-19, policy interventions and credit: The brazilian experience[J]. Journal of Financial Intermediation, 2021, 48(10): 100933.
- [65] HORVATH A, KAY B, WIX C. The covid-19 shock and consumer credit: evidence from credit card data[J]. Available at SSRN: <https://ssrn.com/abstract=3613408>, 2021.
- [66] FISHER I. The debt-deflation theory of great depressions[J]. Econometrica, 1933, 1: 337-357.
- [67] BERNANKE B, BLINDER A. Credit, money, and aggregate demand[J]. American Economic Review, 1988, 78: 435-439.
- [68] BERNANKE B, BLINDER A. The federal funds rate and the channels of monetary policy transmission[J]. American Economic Review, 1992, 82: 901-921.
- [69] BERNANKE B, GERTLER M. Inside the black box: The credit channel of monetary policy transmission[J]. Journal of Economic Perspectives, 1995, 9(4): 27-48.
- [70] BLOOM N, Bond S, REENEN J V. Uncertainty and investment dynamics[J]. Review of Economic Studies, 2007, 74(2): 391-415.
- [71] LEVINE R, LIN C, TAI M, XIE W. How did depositors respond to covid-19?[J]. Review of Financial Studies, 2021, 34(11): 5438-5473.
- [72] BROWNING M, CROSSLEY T F. Shocks, stocks, and socks: Smoothing consumption over a temporary income loss[J]. Journal of the European Economic Association, 2009, 7(6): 1169-1192.
- [73] Braxton J C, Herkenhoff K F, Phillips G M. Can the unemployed borrow? Implications for public insurance[R]. NBER Working Paper, 2020, No.27026.
- [74] Liu Y, Zhang Y, Zhang Y, Xiao H. Small business owners' fintech credit in crises: Theory and evidence from farmers under the covid-19[J]. Pacific-Basin Finance Journal, 2022, 71(1): 101692.
- [75] CUMMIN D J, MARTINEZ-SALGUEIRO A, REARDON R S, SEWAID A. Covid-19 bust, policy response, and rebound: Equity crowd funding and p2p versus banks[J]. The Journal of Technology Transfer, 2021, 47(6): 1825-1846.
- [76] XING S, LI T, CHENG L, Sha F. Fintech model performance analysis in credit loans: Evidence from china city commercial banks[J]. Applied Economics, 2024, 57(49): 8210-2

# **Research on Information Value of Big Data Credit Reporting in Credit Decision: From the perspective of consumption behavior information on the Internet platform**

**Song Ke<sup>1</sup>, Yu Siyan<sup>2</sup>, Yang Yaxin<sup>3</sup>**

(1. School of Finance, Renmin University of China, Beijing 100086, China; 2. Xiamen Siming District Finance Bureau, Xiamen 361005, Fujian, China; 3. Institute of Real Estate and Urban Studies, National University of Singapore, Singapore 117602)

**Abstract :** This study investigates the role of big data credit assessment in credit-granting decisions on an internet consumer credit platform. Utilizing a random sample of approximately 500,000 active users from a domestic internet consumer credit platform with a focus on tourism, we conduct an empirical analysis and obtain the following findings: ① Big data credit assessment provides informational value in credit-granting decisions. The applicant's consumption capacity, consumption habits, and platform engagement are incorporated into credit assessments. Consumers with higher cumulative spending, more frequent use of consumption vouchers, and longer platform engagement enjoy a relative advantage in the credit-granting process. ② Traditional credit information and consumption behavior capture distinct dimensions of information and exhibit a substitution relationship in credit decisions. Variables such as platform usage duration and the proportion of credit card payments serve as substitutes for traditional credit information. ③ In terms of heterogeneity, no evidence of discrimination is found in platform credit-granting decisions. The platform places greater emphasis on the traditional credit profile and platform dependency of younger users while prioritizing consumption capacity in assessing older users. This suggests that internet consumer credit exhibits inclusive characteristics. ④ Under the shock of the COVID-19 pandemic, enhanced risk management by fintech platforms reduces the weight of platform dependency in credit decisions, demonstrating a certain degree of procyclicality.

**Key words:** big data credit reporting; credit decision; consumption behavior; traditional credit reporting; COVID-19 pandemic



中国人民大学国际货币研究所

INTERNATIONAL MONETARY INSTITUTE OF RUC

地址：北京市海淀区中关村大街 59 号文化大厦 605 室，100872    电话：010-62516755    邮箱：imi@ruc.edu.cn