

大数据一定促进普惠金融实现吗?

——基于契约设计理论视角

洪洁瑛 武琦璠 李三希*

摘要: 本文在传统的信贷理论框架中引入大数据,探讨其对信贷契约设计、借贷双方效用以及社会福利的影响。研究发现,大数据的引入并不必然促进普惠金融,也可能强化信贷机构寻租并损害借款方利益。进一步地,数据精确度提升对借款方的影响是非单调的,数据的过度收集利用会损害借款方收益。基于此,本文提出并证明:赋予借款方对自身数据的控制权、防止信贷机构过度收集数据,是缓解上述负面影响、有效推动普惠金融发展的可行方案。

关键词: 大数据;信贷契约设计;普惠金融

DOI: 10.13821/j.cnki.ceq.2026.01.06

一、引言

近年来,中国政府高度重视普惠金融发展,持续推动完善面向小微企业的长效信贷机制。然而,从政策推动到服务落地仍存在一定差距,实践中普惠金融服务效能仍有待提升,大量中小微企业仍未摆脱融资困境。调研数据显示,近七成小微企业存在融资需求,其中三分之一未能获得贷款,“贷款难”仍是制约其发展的关键瓶颈。^① 这一困境的核心根源在于信息不对称。经典信贷理论早已指出,信息不对称会阻碍有效信贷契约的形成与供给(Jensen and Meckling, 1976; Stiglitz and Weiss, 1981)。在传统信贷市场中,金融机构的风险评估与信贷决策高度依赖结构化财务信息,但中小微企业普遍存在财务信息缺失等问题,导致金融机构难以准确评估其信用风险。由此,有效信贷契约供给不足,普惠金融的覆盖面与服务质量也随之受到限制。

围绕这一问题,过往研究提出了多种缓解信息不对称的路径(DeMarzo and Duffie, 1999; Casamatta, 2003; Yang and Zeng, 2019; Yang, 2020; Hong, 2020),包括由借款方提供抵押品或增加自有资金等信号机制(Freixas and Rochet, 2008; Tirole, 2010)、信

* 洪洁瑛,北京航空航天大学经济管理学院、北京航空航天大学低碳治理与政策智能实验室;武琦璠,首都经济贸易大学经济学院;李三希,中国人民大学经济学院、中国人民大学数字经济研究中心。通信作者及地址:武琦璠,北京市丰台区花乡张家路口 121 号首都经济贸易大学经济学院,100070;电话:18811796133;E-mail: wuyufan0103@126.com。本研究感谢国家自然科学基金项目(72203017、72573016、72192801)、教育部人文社会科学青年基金项目(21YJC790044)、首都经济贸易大学新入职青年教师科研启动基金项目(XRZ2026001)的支持。感谢匿名审稿人和期刊编辑的宝贵建议,当然文责自负。

① 资料来源:腾讯研究院、企鹅有调课题组,“中小微企业经营状况与数字化转型调研报告(2025 年一季度)”,腾讯网,https://news.qq.com/rain/a/20250424A06HNR00,访问时间:2025 年 8 月 22 日。

贷机构事后监督(Holmström and Tirole, 1997; Diamond, 1984; Gale and Hellwig, 1985),以及设计控制权分配机制(Aghion and Bolton, 1992)等。但这些传统手段往往实施成本高昂,难以适配中小微企业的实际情况,因而难以从根本上疏通信息不对称这一核心梗阻。在此背景下,数字技术的崛起为破解该难题提供了全新可能:数字金融通过降低信用信息的生成、收集与核验成本,把更多可用信息纳入授信与风控流程,深刻改变了传统金融的运行方式(黄益平和黄卓,2018),被视为突破普惠金融发展瓶颈的关键力量。例如,区块链技术的应用有助于提升信息记录与追溯的可靠性(龚强等,2021),平台承诺担保机制的引入(龚强和王璐颖,2018)在一定程度上缓释了风险分担问题,这些探索已在一定程度上解决了“能贷、会贷”问题。

在上述数字化实践的基础上,大数据的兴起与应用进一步拓展了信用信息来源,为破解信息瓶颈提供了新思路,却也催生了新的挑战。当前,社交媒体动态、网络搜索记录、消费行为踪迹等海量非结构化数据被广泛纳入信贷决策流程,用于精准评估借款方信用状况。依托数据优势,蚂蚁集团、字节跳动、度小满等大型科技公司通过大数据分析有效缓解了传统信贷市场的信息不对称,显著提升了信贷决策效率。但与此同时,大数据的引入也伴生了新的风险:部分金融科技信贷公司凭借数据优势实施寻租行为,通过抬高贷款利率损害社会福利(Ru and Schoar, 2016)，“普而不惠”的问题日益凸显,典型表现为部分小贷产品的年化利率远高于银行个人贷款水平^①。因此,大数据的广泛应用究竟能否推动普惠金融目标实现,抑或在特定条件下适得其反,成为亟待深入探讨的重要课题。

回答上述问题,需要将研究视角从大数据缓解信息不对称,推进到大数据如何影响信贷契约设计的微观机制分析。在这一研究脉络下,数据经济学与信贷领域的既有文献为理解数据要素的经济作用提供了重要分析基础。近些年,围绕数据要素信息价值的分析逐渐展开,研究者从不同视角探讨了数据要素的经济作用(Jones and Tonetti, 2020; Cong et al., 2021; Acemoglu and Restrepo, 2018; 陈彦斌等, 2019; 蔡跃洲和陈楠, 2019)。其中一类重要观点认为,数据可视为用于提高对随机变量的预估精确程度的信息资源,从而降低决策不确定性等(Begenau et al., 2018; Farboodi and Veldkamp, 2020)。在信贷领域,也有研究开始关注大数据在缓解信息不对称方面的作用,例如,有学者探究非结构化数据在信贷风险识别中的应用(Berg et al., 2020),或分析数据配置规则对信贷市场竞争的影响(谢丹夏等, 2022; He et al., 2023)。然而,这些研究主要关注放贷机构间的数据竞争,尚未从契约设计视角系统分析大数据对信贷契约条款及借贷双方福利的影响。

基于此,本文从契约设计视角出发,在经典不对称信息信贷理论的基础框架上(Holmström and Tirole, 1997; Tirole, 2010),构建一个纳入大数据的信贷契约模型。具体而言,我们借鉴 Begenau et al.(2018)及 Farboodi and Veldkamp(2020)的工作,将数

^① 某集团向上海证券交易所科创板递交上市招股说明书(申报稿)中披露,截至2020年6月30日12个月期间内,大部分贷款日利率为0.04%左右(年化利率14.6%),远高于银行个人贷款年化利率水平(3.85%)。

据要素视为一种帮助信贷机构甄别借款方类型的预测技术。^①在此框架下,首先,本文分别求解了当信贷机构没有数据与拥有数据两种情况下的最优信贷契约。对比分析发现:第一,大数据的引入具有正反双重效应。数据既可能提升借款方获得融资的可能性及降低其融资成本,也可能强化信贷机构的寻租行为,反而降低借款方融资概率并推高成本。第二,数据信息准确度对借款方福利的影响呈现非单调性。在一定范围内,信贷机构的数据收集与利用能提高借款方福利,但过度收集会损害借款方的利益。其次,鉴于信贷机构数据垄断的潜在负面影响,本文将分析视角拓展至数据所有权的配置问题。在现实制度层面,各国已陆续出台强化个人金融数据控制权的相关法规,对金融数据的权属与使用边界作出明确规范。借鉴这一制度背景,本文进一步探讨了数据所有权的配置问题,发现赋予借款方数据控制权,可有效缓解数据垄断与寻租行为,显著提高借款方的福利。这一发现为加强数据保护与规范大数据使用提供了理论支撑。最后,本文进一步拓展到社会福利分析,发现将数据所有权配置给借款方,不仅提高了借款方福利,还能够通过抑制数据垄断带来的效率损失,实现社会总福利的帕累托改进。

本文的贡献主要体现在对信贷契约理论的拓展、对数据作用机制的刻画,以及对金融数据治理的解释三个方面。第一,在传统信贷理论框架下引入数据要素,刻画了数字经济背景下数据对信贷契约设计的作用机制,并构建了适用于中国小微信贷市场的分析框架。第二,从契约设计视角揭示了大数据对普惠金融既有促进也有制约的非单调影响,为现实中金融科技助贷中“普而不惠”现象提供了理论解释。第三,从契约视角讨论金融数据治理问题,表明过度收集数据可能损害借款方利益,并分析赋予借款方对自身数据控制权以抑制数据滥用的机制,丰富了数据确权与数据保护的相关研究(张丰羽和汤珂,2023;张剑虎和林平,2023;李三希等,2023),为监管部门防止信贷机构过度收集数据以及加强借款方对金融数据控制权提供理论支撑和政策参考。

二、模型设定

本部分给出借贷双方的基本设定与博弈时序,明确本文使用的均衡概念,并求解模型的最优基准,为后文推导最优信贷契约奠定基础。

(一) 参与者

本模型包含借贷双方:一方为需要外部融资以实施项目的借款方,另一方为提供资金并通过契约安排回收收益的信贷机构。以下分别给出项目回报、外部效用、不对称信息等基础假设。

假设借款方面临一个需投资 I 的项目,但其自有资金不足,因而需要向信贷机构融资。该项目的未来收入为随机变量 \tilde{R} ,以概率 θ 取值为 R ,概率 $1-\theta$ 取值为 0 。其中 θ 为项目成功的概率,由借款方的能力决定,能力越高,项目成功的概率就越大。若借款方

^① 不同于 Farboodi and Veldkamp(2020)的研究,数据要素在本文模型中用于预测的变量不再是生产函数的某个参数,而是借款方的类型。

未能获得融资而无法开展项目,则可转向其他工作并获得外部效用 W 。该外部效用设定在契约文献中较为常见(Millon and Thakor, 1985; Pitchford, 1998; Li et al., 2013),既可理解为借款方的讨价还价能力,也可视为借款企业的运营机会成本。

借款方存在高质量 H 和低质量 L 两种类型:高质量借款方的项目成功概率为 θ_H , 外部效用为 W_H ;低质量借款方的项目成功概率为 θ_L , 外部效用为 W_L , 并满足 $\theta_H > \theta_L$, $W_H > W_L$ 。可以理解为高质量借款方因其管理能力或技术优势,项目成功的概率更高,将资源用于其他工作时也能产生更高回报,故其外部效用也更高。借款方类型为私人信息,但其先验分布为公共知识:以 $1/2$ 的概率为高质量 H , 以 $1/2$ 的概率为低质量 L 。此外,假设 $\theta_i R - I \geq W_i, i = H, L$, 即项目期望回报不低于外部效用。因此,在资金充足时投资项目是借款方的最优选择。

信贷机构为项目提供融资,并在获得与借款方类型相关的信息后,设计信贷契约菜单以最大化自身期望效用。假设信贷机构和借款方均为风险中性,且借款方受有限责任保护,即其还款的最高额度不会超过自身的收入(相关设定见 Holmström and Tirole, 1997; Laffont and Martimort, 2009; Tirole, 2010)。

(二) 博弈时序

本小节说明借贷双方在信息形成与契约交易中的行动顺序。博弈按如下时序展开:第0期,自然决定借款方的类型;第1期,信贷机构基于大数据获得与借款方类型相关的数据信号并据此设计契约;第2期,借款方在契约菜单中作出选择或退出;第3期,项目实现并按契约还款。

第1期的关键在于,信贷机构通过大数据分析获得与借款方类型相关的信息,可将其概括为一个可用于决策的数据信号 η 。本文将数据理解为借款方在社交媒体、网络搜索、消费及经营行为等场景中形成的非结构化的数据要素。以蚂蚁集团为例,当其向淘宝商家等小微企业提供贷款服务时,可利用商家在平台上留下的大量经营足迹数据进行分析,并据此形成与借款方类型相关的信号。该数据信号 η 对借款方类型具有预测能力,以概率 p 产生与真实类型一致的信号,以概率 $1-p$ 产生不一致的信号,即:

$$\Pr(\eta_H | \theta_H, W_H) = \Pr(\eta_L | \theta_L, W_L) = p;$$

$$\Pr(\eta_L | \theta_H, W_H) = \Pr(\eta_H | \theta_L, W_L) = 1 - p.$$

其中, $1/2 < p < 1$ 。信贷机构可基于 η 对借款方类型进行贝叶斯更新。^①

在得到后验信念后,信贷机构向借款方提供一个不容讨价还价的信贷契约菜单 (L, R) (矩阵)。其中, L (向量)代表第1期信贷机构发放给借款方的贷款额度, R (向量)代表第3期贷款到期时借款方的还款额。借款方只能在信贷机构提供的契约菜单中选择或拒绝合约。这一假设也与平台主导信贷模式下中小微企业融资渠道有限、议价能力较弱的现实相符。

^① 大数据提高了信贷机构对借款方类型的识别精度,使其能够据此提供差异化的信贷产品,体现了大数据在现实中的信贷契约中发挥的具体作用。

（三）均衡类型

本小节界定本文的均衡概念。在上述不完全信息博弈框架中，借款方依据契约菜单选择使其期望效用最大化的合约；而信贷机构则利用数据信号，通过贝叶斯更新修正对借款方类型概率分布的信念，并在满足借款方激励相容约束与参与约束的前提下，设计能够最大化自身期望效用的最优契约。该契约在设计时已预判借款方后续的最优反应。因而，本模型的均衡为精炼贝叶斯纳什均衡。

（四）最优基准

为便于后文比较，先考虑不存在信息不对称的最优基准。在该情形下，信贷机构能够准确识别借款方类型，并据此提供契约菜单。由于项目的回报高于外部效用，最优基准中，信贷机构为借款方提供资金，使其进行项目投资和开发。^① 通过求解，得到

$$R_H = R - \frac{W_H}{\theta_H}, R_L = R - \frac{W_L}{\theta_L}, L_H = L_L = I.$$

因此，在不存在信息不对称的理想情形下，两类借款方都将获得融资 I 来启动项目，信贷资金实现最优配置，数据也并不会产生额外的效用。此时，交易价值将全部由信贷机构获得。信贷机构的期望效用可以表示为：

$$V = \frac{1}{2}(\theta_H R - W_H) + \frac{1}{2}(\theta_L R - W_L) - I.$$

此外，本文假设 $W_H \theta_L > W_L \theta_H$ 。在相反情形下（即 $W_H \theta_L \leq W_L \theta_H$ ），信息不对称并不会妨碍社会达到最优基准的信贷分配结果，因此该情形下数据的引入不产生实质性作用。基于此，本文在 $W_H \theta_L > W_L \theta_H$ 这一假设条件下展开分析，以刻画信息不对称对信贷契约的扭曲以及数据要素可能发挥的缓解作用。该假设条件背后的经济学含义在于，风险更高的投资需要更高的回报。通过数学变换发现， $W_H \theta_L > W_L \theta_H$ 等同于 $R - \frac{W_H}{\theta_H} < R - \frac{W_L}{\theta_L}$ ，其中， $R - \frac{W_H}{\theta_H}$ 是最优基准中高质量借款方的还款额，而 $R - \frac{W_L}{\theta_L}$ 为低质量借款方的还款额。该不等式表明，为了补偿更高的失败风险，信贷机构需从风险高的投资中获得更高的回报。

三、最优契约设计分析

本部分在模型设定的基础上推导信贷机构的最优契约，并通过引入数据前后两种信息结构的对比识别大数据对契约条款的影响。具体而言，下文先求解信贷机构无法获得

^① 从建模角度看，模型引入借款方外部效用可以反映借款方的讨价还价能力。具体来说，当借款方外部效用较高时，代表着其讨价还价能力也增强。根据我们理论的结果，外部效用 W_i 影响信贷契约中的还款额度 $R - \frac{W_i}{\theta_i}$ 。当外部效用提升时，借款方向信贷机构支付的还款额降低。意味着借款方讨价还价能力增强时，能够争取到更优惠的合约，降低贷款成本。

数据时的最优契约,再求解其能够基于数据信号进行贝叶斯更新时的最优契约,最后在两种契约结果的对照下讨论大数据的作用机制。

(一) 引入数据前信贷机构的契约设计

在没有数据信号的情形下,借款方类型为私人信息,信贷机构只能基于先验信念设计信贷契约菜单 (L_i, R_i) ,其中 $i = H, L$ 。根据显示原理,信贷机构可提供包含两类合约的菜单,以满足借款方的自选择约束。^① 信贷机构的最优决策是在满足借款方激励相容和参与约束等条件下,设计契约以最大化自身期望效用。

为确保所得均衡解为全局最优,在求解过程中将同时考察以下四种可能出现的契约实施情形(见表1),包括借款额度大于($L_i > I$)或小于投资需求($L_i < I$)等均衡路径外的情形^②,以确保最优契约是在所有可能情况下均优于其他任何潜在的策略。

表1 信贷机构提供贷款额度的四种情形

信贷机构决策	
情形1	$L_i \geq I$, 其中 $i = H, L$, 信贷契约给两类借款方都提供足够的贷款额度。
情形2	$L_H \geq I$ 和 $L_L < I$ 。信贷机构给高质量借款方提供足够的贷款额度,而给低质量借款方的额度小于其投资需求。
情形3	$L_L \geq I$ 和 $L_H < I$ 。信贷机构向低质量借款方提供足够的贷款额度,而给高质量借款方的额度小于其投资需求。
情形4	$L_i < I$, 其中 $i = H, L$, 信贷机构提供给两种类型借款方的贷款额度都小于其投资需求。

首先,分析情形1中的最优均衡解,即信贷契约给两类借款方都提供足够的贷款额度。这种情况下,借款方的激励相容条件为:

$$\theta_i(R - R_i) + (L_i - I) \geq \theta_i(R - R_{-i}) + (L_{-i} - I), \forall i = H, L, \quad (1)$$

其中, $\theta_i(R - R_i) + (L_i - I)$ 代表类型 i 的借款方选择与其真实类型相对应的信贷契约 (L_i, R_i) 所获得的期望效用。 $\theta_i(R - R_{-i}) + (L_{-i} - I)$ 代表类型 i 的借款方选择了与类型 $-i$ 对应的信贷契约 (L_{-i}, R_{-i}) 所得到的期望效用。式(1)确保借款方选择真实类型对应的契约优于选择其他类型对应的契约。

借款方选择接受信贷契约需满足如下参与约束条件:

$$\theta_i(R - R_i) + (L_i - I) - W_i \geq 0, \forall i = H, L. \quad (2)$$

参与条件保证借款方选择自己真实类型相对应的信贷契约所获得的期望效用会大于等于外部效用。同时需满足以下约束: $0 \leq R_i \leq R$,即借款方享有有限责任保护; $L_i \geq I$,即项目获得足够融资。

^① 本文模型在形式上虽提供差异化契约菜单(两种选择),但最优契约为混同还是分离均衡是内生的结果,不排除混同均衡作为最优解的可能性。

^② 虽然部分情形在直观上看似乎与实际情况有所偏离,但在最优契约的求解中,比较均衡路径内外的可行配置以排除潜在的更优偏离属于常规处理(Casamatta, 2003; Fudenberg and Tirole, 1991; Holmström and Tirole, 1998, 2011; Tirole, 2010)。

在满足式(1)、式(2)以及其他两个条件下,信贷机构设计契约 (L_i, R_i) , $i = H, L$, 最大化自身期望效用。其中,信贷机构的效用函数如下:

$$V^0 = V(L_i, R_i) = \frac{1}{2}(\theta_i R_i - L_i) + \frac{1}{2}(\theta_{-i} R_{-i} - L_{-i}). \quad (3)$$

类型 i 的借款方的期望效用为:

$$U_i^0 = \theta_i(R - R_i) + (L_i - I). \quad (4)$$

通过求解该最优化问题,可得均衡结果为:

$$U_H^0 = W_H, R_H = R - \frac{W_H}{\theta_H}, L_H = I; \quad U_L^0 = \frac{\theta_L}{\theta_H} W_H, R_L = R - \frac{W_H}{\theta_H}, L_L = I.$$

相应地,信贷机构的期望效用为:

$$V^0 = \frac{1}{2}(\theta_H R - I - W_H) + \frac{1}{2}\left(\theta_L R - I - \frac{\theta_L}{\theta_H} W_H\right).$$

具体证明见附录 I^①。采用类似分析方法,可以求解其他三种情形下信贷机构的最优契约设计,详细分析过程见附录 II。

为便于刻画信息租金成本在最优契约中的相对重要性,定义

$$X = \frac{\theta_H R - I - W_H}{\frac{\theta_L}{\theta_H} W_H - W_L},$$

其中, $\theta_H R - I - W_H$ 代表信贷机构向高质量借款方提供贷款所获得的净收益, $\frac{\theta_L}{\theta_H} W_H - W_L$ 代表为防止低质量借款方模仿高质量借款方所支付的信息租金成本。因此, X 可以理解为信贷机构的收益相对于信息租金成本的大小。

综合上述分析,无数据情形下的最优契约呈现出如下分段结构(整理如表 2 所示):

① 当 $X \geq 1$, 即信息租金成本相对较低时,信贷机构提供 $(I, R - \frac{W_H}{\theta_H})$ 的信贷契约, 两个类型的借款方均接受契约。

② 当 $X < 1$, 即信息租金成本相对较高时,信贷机构提供 $(I, R - \frac{W_L}{\theta_L})$ 的信贷契约, 低质量的借款方会接受该契约,而高质量的借款方选择退出。

表 2 信贷机构无数据时的均衡信贷契约 (L, R)

参数取值区间	高质量借贷方的信贷契约	低质量借贷方的信贷契约
$X \geq 1$	$(I, R - \frac{W_H}{\theta_H})$	$(I, R - \frac{W_H}{\theta_H})$
$X < 1$	$(0, 0)$	$(I, R - \frac{W_L}{\theta_L})$

① 篇幅所限,附录未在正文列示,感兴趣的读者可在《经济学》(季刊)官网(<https://ceq.ccer.pku.edu.cn>)下载。

这一结果表明,在信息不对称的环境中,低质量借款方存在模仿高质量借款方以获得更优惠贷款条件的动机,从而迫使信贷机构在契约中向低质量借款方支付信息租金以维持激励相容。当信息租金成本较低($X \geq 1$)时,信贷机构能为两类借款方同时提供融资,使得项目均得以开展。而当信息租金成本较高($X < 1$)时,信贷机构倾向于牺牲高质量借款方的融资机会,以减少激励相容约束带来的成本,结果是高质量借款方因贷款条件不具有吸引力而退出信贷市场。这一机制与 Stiglitz and Weiss (1981)指出的信息不对称导致融资排斥等经典理论发现一致。

(二) 引入数据后信贷机构的契约设计

数据的引入使得信贷机构能够基于数据信号对借款方类型进行判断,从而在契约设计中使用后验信念进行优化。具体而言,当信贷机构获得数据信号 η_H 时,其对借款方类型的后验概率由贝叶斯法则给出:

$$\Pr(\theta_H, W_H | \eta_H) = \frac{\Pr(\eta_H | \theta_H, W_H) \Pr(\theta_H, W_H)}{\Pr(\eta_H | \theta_H, W_H) \Pr(\theta_H, W_H) + \Pr(\eta_H | \theta_L, W_L) \Pr(\theta_L, W_L)} = p,$$

$$\Pr(\theta_L, W_L | \eta_H) = 1 - \Pr(\theta_H, W_H | \eta_H) = 1 - p.$$

类似地,当信贷机构获得数据信号 η_L 时,更新后的后验分布为:

$$\Pr(\theta_L, W_L | \eta_L) = p, \quad \Pr(\theta_H, W_H | \eta_L) = 1 - p.$$

由于 $p > 1/2$,表明数据信号具有有效的针对借款人类型的识别能力。因此, p 可以视为数据信息的有效程度。 p 值越大,数据信号的预测精度越高。

在获得后验信念后,信贷机构设计信贷契约菜单 (L_i^j, R_i^j) , 其中 $i, j = H, L$ 。信贷契约设计的目标是在满足借款方激励相容约束和参与约束的条件下,最大化信贷机构的期望效用。与引入数据前的求解思路一致,同样分四种情形进行最优化求解(求解及证明见附录 III),整理可得表 3:

表 3 信贷机构拥有数据时的均衡信贷契约 (L, R)

参数取值区间	高质量借贷方的信贷契约	低质量借贷方的信贷契约
$X > \frac{p}{1-p}$	$(I, R - \frac{W_H}{\theta_H})$	$(I, R - \frac{W_H}{\theta_H})$
$\frac{1-p}{p} \leq X \leq \frac{p}{1-p}$	$(0, 0)$, 信贷机构获得 η_L $(I, R - \frac{W_H}{\theta_H})$, 信贷机构获得 η_H	$(I, R - \frac{W_L}{\theta_L})$, 信贷机构获得 η_L $(I, R - \frac{W_H}{\theta_H})$, 信贷机构获得 η_H
$X < \frac{1-p}{p}$	$(0, 0)$	$(I, R - \frac{W_L}{\theta_L})$

上述结论表明,数据会影响信贷机构的契约设计。当信息租金处于中间水平,即 $\frac{1-p}{p} \leq X \leq \frac{p}{1-p}$ 时,数据会带来有效的边际影响:信贷机构在获得不同信号时将提供不同的最优契约。在该区间内,当信贷机构观察到信号 η_H 时,两类借款方均可获得融资。而当信贷机构观察到信号 η_L 时,仅低质量的借款方获得融资,高质量借款方因合约不具有

吸引力退出市场。相比之下,当信息租金成本较高($X > \frac{p}{1-p}$)或较低($X < \frac{1-p}{p}$)时,最优契约对数据信号不敏感,契约条款与无数据情形下的最优结构一致。

这一理论发现揭示了在适中信息租金水平下大数据引入对信贷市场产生的复杂筛选效应。结合现实来看,这一微观机制在 Berg et al.(2020)基于德国电商平台的研究中得到了有力的经验支撑。尽管该实证研究未直接界定信息租金区间,但其关于中等征信评分群体的发现与本文适中信息租金的理论界定存在精确的映射关系。该研究表明在传统评分中等的信贷模糊区,数据的引入导致了显著的挤出效应,即拥有中等传统评分但在数字足迹上表现不佳的申请人失去了信贷机会。这在逻辑上印证了本文模型结论,即当数据产生负面信号时信贷机构倾向于提供严苛的契约。面对不再具有吸引力的合约条款,被信号误伤的高质量借款人选择主动退出市场,而只有真实的低质量借款人会选择接受。这种由定价机制引发的良币驱逐效应导致留存的信贷池中低质量样本高度富集,在宏观上表现为 Berg et al.(2020)所观察到的负面数字足迹群体的高违约率与部分潜在合格借款人的信贷缺失并存的现象。

(三) 大数据的影响

在前两小节求解结果的基础上,我们进一步刻画大数据引入对均衡契约的边际影响。具体而言,首先对比无数据与有数据两种情形下的均衡契约条款,识别数据对融资可得性与融资成本的影响方向;随后在此基础上,通过比较静态分析,进一步考察数据信息精度 p 的变化如何影响均衡结果。

对比无数据和有数据情形下的均衡差异发现,数据信息对不同类型借款方的作用路径存在明显差异:对高质量借款方而言,数据的引入主要影响其能否获得融资,一旦获得融资后信贷契约保持不变;而对低质量借款方而言,数据的引入并不改变其融资可得性,却会通过还款额的调整影响其融资成本。为了更直观地刻画这一差异化效应,本文将参数取值范畴变换为数据信息精度 p 的参数范畴。图 1 中的点划线和虚线分别代表信贷机构通过分析数据得到信号 η_L 和 η_H 的信贷情况,实线代表信贷机构没有数据时的信贷情况。

基于图 1 所刻画的区间变化特征,可以得到如下命题。

命题 1 当 $X \geq 1$ 时,若 $\frac{1}{2} < p < \frac{X}{1+X}$,数据的引入不会改变均衡信贷契约;若 $p \geq \frac{X}{1+X}$,与无数据情形相比,基于数据的信贷契约将降低高质量借款方获得融资的概率,并提高低质量借款方的融资成本,从而降低借款方整体的期望效用。

当 $X < 1$ 时,若 $\frac{1}{2} < p < \frac{1}{1+X}$,数据的引入同样不会改变均衡信贷契约;若 $p \geq \frac{1}{1+X}$,与无数据情形相比,基于数据的信贷契约将提高高质量借款方获得融资的概率,同时降低低质量借款方的融资成本,从而提高借款方整体的期望效用。

证明:见附录 IV。

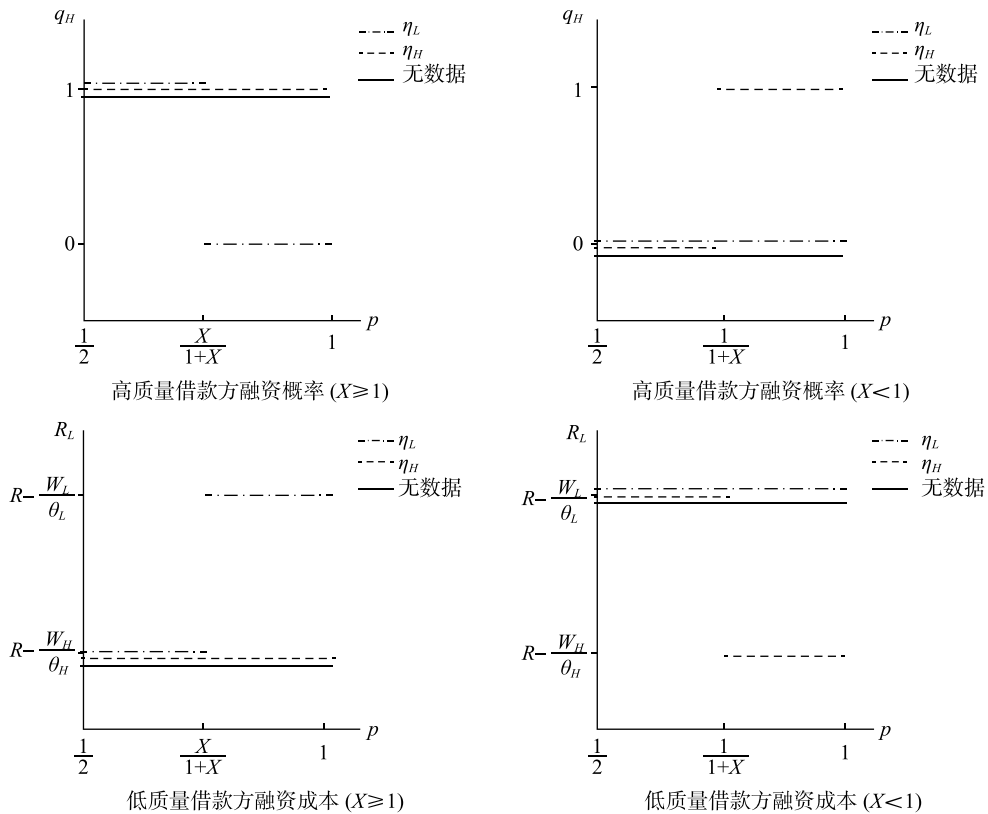


图1 数据引入前后的信贷契约对比情况

命题1的理论结果表明,数据的作用不仅体现在能够更准确识别借款人类型,更体现在信贷机构据此扩展了差异化定价与筛选的策略空间。当信息租金相对较高($X < 1$)时,传统不对称信息造成的扭曲较强,数据引入带来的识别精度提升会更更多地被用于缓解逆向选择,从而提升高质量借款方获得融资的机会,或降低低质量借款方的融资成本,因而更接近“既普又惠”。相反,当信息租金相对较低($X \geq 1$)时,原本由激励相容约束带来的让利压力较小,数据优势更可能被转化为定价权与剩余攫取:信贷机构在维持放贷收益的同时,通过更严的筛选或更高的还款要求,将部分风险与成本转嫁给借款方,表现为融资可得性降低或融资成本上升。这也意味着,数据红利并不自动转化为借款方红利,其福利效应取决于数据在契约中被用来“纠偏”还是被用来“攫取”。换言之,数据驱动的信贷扩展在某些制度与市场结构下会导致“普而不惠”甚至“既不普又不惠”的现实结果。

上述分析表明,数据对借款方可能利害并存;但对信贷机构而言,数据带来的选择空间扩大意味着其最优收益不会下降。下面给出信贷机构侧的单调性结论(命题2),并以此作为后文讨论数据治理与权利配置的出发点。

命题2 与引入数据前的信贷契约相比,基于数据要素的信贷契约提高了信贷机构的期望效用。

证明:见附录V。

命题2的含义在于,数据要素通过提升借款人类型识别与合同差异化设计的可行性,扩展了信贷机构的最优决策集合,使其在均衡中的期望效用不低于无数据情形。直

观地说,数据要素给信贷机构带来了额外的信息,而信贷机构在合约设计中可以选择是否采用以及如何采用该信息。在最优选择下,新增信息不会使其效用下降。近年来,部分依托大数据开展助贷服务的平台型机构已经通过卖数据及风控服务等模式获得了可观的超额收益,这与命题 2 的发现相契合。

在分析了数据的引入对借贷双方各自的影响后,下面进一步通过比较静态分析探究数据信息精度 p 对均衡结果的影响。具体而言,考察在信贷机构拥有数据情况下,数据信息精度 p 对借款方融资概率、期望融资成本、借贷双方期望效用的影响。

为刻画融资概率与融资成本,本文将高质量借款方的期望融资概率定义为在两类信号下融资结果的加权平均,权重为信贷机构获得两类信号的概率。例如,当 $X \geq 1$ 且 $p > \frac{X}{1+X}$ 时,高质量借款方的融资概率为 $p \times 1 + (1-p) \times 0 = p$ 。这是因为,面对高质量借款方,信贷机构会以概率 p 获得信号 η_H ,此时高质量借款方获得贷款;信贷机构还会以概率 $1-p$ 获得信号 η_L ,此时高质量借款方未获得融资。其他参数情况下的融资概率也同样求解。类似地,低质量借款方的期望融资成本定义为两类信号下还款额度的加权平均,权重同样由信贷机构获得两类信号的概率决定。图 2 进一步展示了借款方融资概率与融资成本随数据信息精度 p 变化的规律。

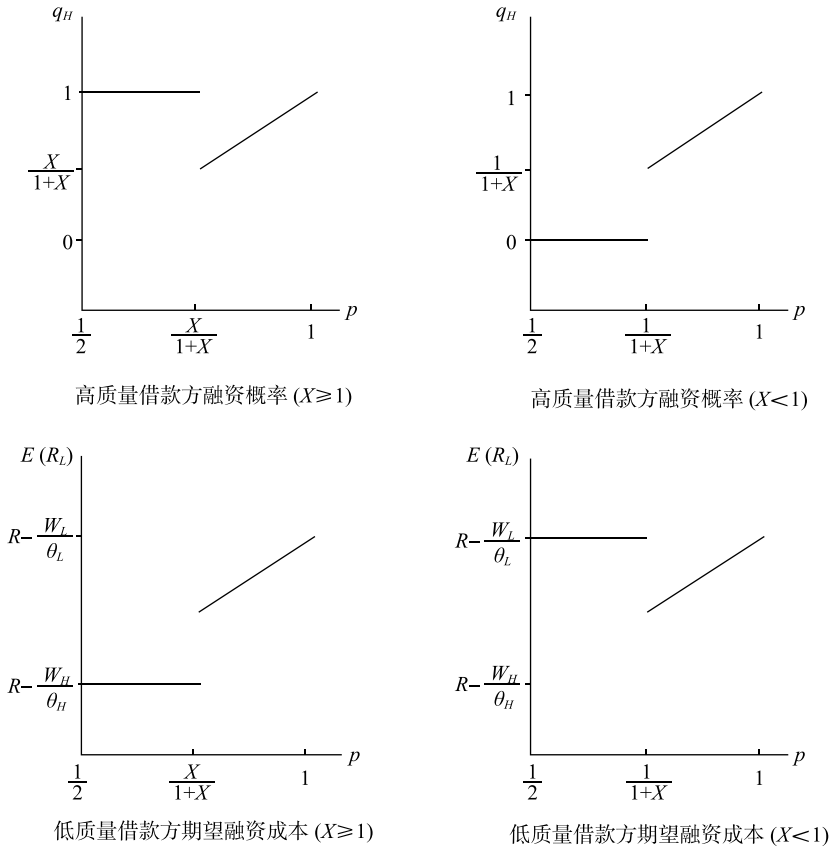


图 2 数据信息精度对借款方融资概率及融资成本的影响

由此,我们得到命题3。

命题3 对于高质量借款方而言,当 $X \geq 1$ 时,数据信息精度 p 对其融资概率的影响是非单调的,融资概率在 $p = \frac{X}{1+X}$ 时取值最小;当 $X < 1$ 时,数据信息精度 p 的增加会提高其融资概率。对于低质量借款方而言,当 $X \geq 1$ 时,数据信息精度 p 的增加会提高其期望融资成本;当 $X < 1$ 时,数据信息精度 p 对其期望融资成本的影响呈非单调关系,该成本在 $p = \frac{1}{1+X}$ 处取最小值。

证明:见附录VI。

命题3刻画了一个关键机制:数据信息精度 p 的提升并不必然改善借款方的融资结果,其作用方向取决于参数 X 所反映的信息租金成本。随着数据信息精度 p 的提高,信贷机构确实能够更精准地区分高质量和低质量借款方,但这种区分能力的增强会同时改变契约筛选与定价的均衡边界。在 $X \geq 1$ 的情形下,均衡调整可能表现为对部分借款方更严格的筛选或更不利的合同条款,从而使高质量借款方的融资概率在某一精度水平后反而下降,并使低质量借款方面临更高的期望融资成本。而在 $X < 1$ 的情形下,数据则更主要发挥缓解信息不对称的作用,因此更可能提升高质量借款方的融资概率,并在一定区间内降低低质量借款方的融资成本。

为了更加清楚地分析数据信息精度 p 对借款方群体的总体福利影响,下一步继续探究借款方的平均期望效用随数据信息精度 p 的变化规律。其中,借款方的平均期望效用定义为高质量和低质量借款方的平均期望效用,用以刻画借款方群体福利。通过求解得到命题4。

命题4 当信息租金相对较低时($X \geq 1$),数据信息精度 p 的增加会损害借款方利益;然而,当信息租金相对较高时($X < 1$),数据信息精度 p 的增加对借款方利益的影响是非单调的。当 $p = \frac{1}{1+X}$ 时,借款方的期望效用达到最大。

证明:见附录VII。

命题3和命题4从理论层面证明了数据收集对借款方的非单调影响。现实中也存在与这一机制相呼应的现象:随着数据信息精度的提高,借款方群体的总体融资条件并非单向改善,而是呈现出先改善后弱化的非单调特征。例如,Jagtiani and Lemieux (2019)对互联网借贷平台的实证分析表明,在引入大数据驱动的信用评分模型之后,尽管部分借款人的利率有所改善,但在更精细的风险分层与定价机制作用下,中等风险区间的借款成本并未显著降低甚至有所上升,从而导致借款人的平均融资成本并未在总体水平上持续下降。这种非单调的融资成本变化印证了命题3和命题4揭示的非单调效应,说明对借款方数据的过度收集与精细使用可能强化信贷机构的定价优势,并降低普惠效果。这一发现不仅为现行个人信息保护法提供了理论支持,也为限制企业过度收集数据的相关立法提供了依据。

最后,分析信贷机构的期望效用如何随着数据信息精度 p 变化。与前文分析方法一致,通过求解可以得到命题5。

命题 5 数据信息精度 p 的增加会提高信贷机构期望效用。

证明：见附录 VIII。

命题 5 说明当信贷机构拥有数据时，数据信息精度的提升有助于其更精准地识别借款方类型，从而依据更准确的信号优化契约设计，提高自身效用。大量实践经验也表明，提高数据信息精度能显著增强信贷机构的风险识别与管控能力，并提升其盈利水平。例如，金融机构借助人工智能和大数据构建“全景式”借款方画像，风险判断更为精确，使得贷款决策更高效、坏账率降低，从而提高整体收益率。不少数字银行与互联网银行已经在实践中体现出数据越丰富、模型越精细、利润空间越大的特征。

因此，本文发现，引入数据要素以及数据信息精度的提升能提高信贷机构的利润，但却可能损害借款方的利益。这表明，有必要建立更加有效的机制，以规范信贷机构或企业对用户个人信息的收集行为，切实保护用户权益。基于此，下一部分进一步分析，若将数据控制权赋予借款方自身，增强个体对自身数据的控制和保护，是否能够化解数据要素带来的负面效应。

四、数据所有权的分配

前文分析表明，大数据能提升信贷机构识别借款人类型的精度，进而影响契约设计与借贷双方期望效用。然而，数据不仅是一种信息载体，其本质上还具有明确的所有权属性。在现实中，借款方在数字平台留下的行为数据被金融科技机构广泛收集与利用，却往往未明确其权利归属。这引发了一个关键问题：数据所有权应配置给作为收集方的信贷机构还是作为原发者的借款方？所有权的不同安排是否会改变信贷契约的均衡结果与社会福利？为回答这一问题，我们将系统考察所有权配置的经济影响。

本部分扩展基准模型，将数据控制权赋予借款方。基于前文的模型设定，并借鉴 Choi et al.(2019)和 Acemoglu et al.(2022)的工作，我们在信贷机构设计契约之前引入借款方的数据披露决策。通过这一拓展，我们将分析借款方在拥有数据控制权下的策略性披露行为，及其对信贷契约与社会福利的深层影响。具体如图 3 所示，考虑一个四期博弈模型：

第 0 期，自然决定借款方类型，高质量与低质量借款方的概率各为 $1/2$ 。

第 1 期，借款方决定自己的数据披露策略 $d_i \in \{0, 1\}$ ，其中 $i = H, L$ 。 $d_i = 1$ 表示类型 i 的借款方选择披露数据信息； $d_i = 0$ 表示不披露。

第 2 期，信贷机构根据借款方的披露决策及所披露的数据信息，对借款方类型的概率分布进行贝叶斯更新，并据此设计契约菜单 (L_i, R_i) ， $i = H, L$ 。其中 L_i 为贷款额度， R_i 为还款额度。

第 3 期，借款方从贷款契约菜单中选择合适的契约。

第 4 期，项目结束。借款方按照信贷契约的规定进行还款，信贷机构和借款方按照信贷契约获得相应收益。

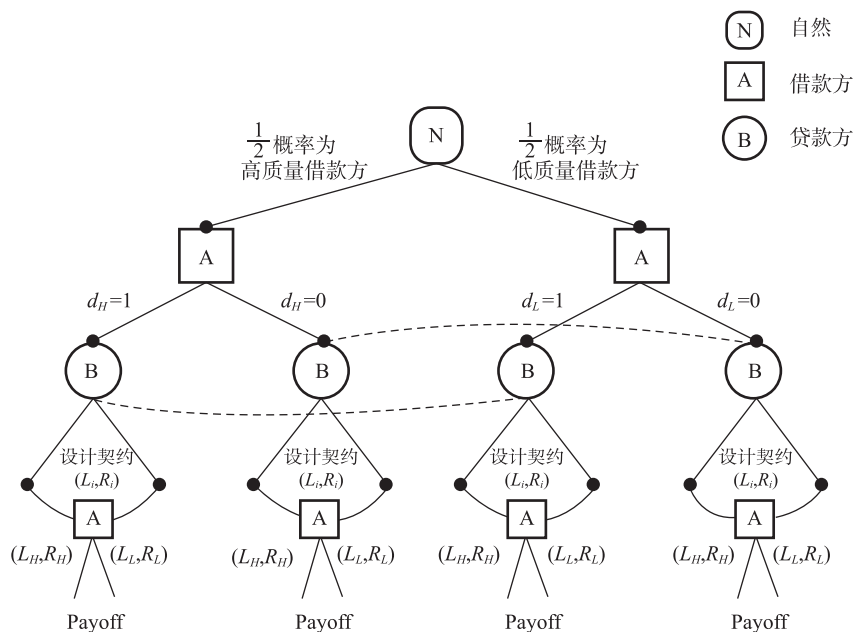


图3 博弈树

关于第1期借款方的信息披露策略,本文聚焦于纯策略均衡。在模型设定下,存在四种可能的纯策略组合:

- ① $(d_H=0, d_L=0)$, 两类借款方均不披露数据。
- ② $(d_H=1, d_L=0)$, 高质量借款方选择披露数据, 低质量借款方选择不披露。
- ③ $(d_H=0, d_L=1)$, 高质量借款方选择不披露数据, 低质量借款方选择披露。
- ④ $(d_H=1, d_L=1)$, 两类借款方均选择披露数据。

第2期, 信贷机构根据借款方在第1期的策略和披露的数据信息, 通过贝叶斯法则更新对其类型的概率判断, 并在此基础上设计满足激励相容和参与约束的最优信贷契约, 以实现自身期望效用最大化。均衡外的贝叶斯预期设定如下: 如果信贷机构观察到上述策略外的其他情况, 则将基于其对借款方类型的先验信念设计最优契约。这一设定意味着, 若借款方偏离上述策略, 信贷机构将忽略其披露的信息, 仍然按照先验概率分布预期对借款方进行评估和放贷。该假设也较为合理, 因为均衡策略外的路径, 信贷机构无法解读其背后信号。

经过分析, 我们得到如下命题:

命题6 $(d_H=0, d_L=0)$ 是均衡策略; 当 $X < 1$ 时, $(d_H=1, d_L=1)$ 也是均衡策略。

证明: 见附录IX。

基于命题6, 进一步分析借款方在第1期的最优均衡策略选择。通过比较不同均衡策略下的期望效用, 借款方选择最大化其效用的均衡策略。通过求解, 得到下面的结论:

命题7 当 $\frac{1-p}{p} \leq X < 1$ 时, 借款方选择均衡策略 $(d_H=1, d_L=1)$; 当 $1 \leq X \leq$

$\frac{\rho}{1-\rho}$ 时, 借款方选择均衡策略 ($d_H = 0, d_L = 0$)。

证明: 见附录 X。

根据命题 7, 可以得到如下推论:

推论 1 当借款方被赋予数据所有权时, 若信息租金相对较高, 借款方倾向于选择披露信息, 从而实现信贷机构和借款方双重受益; 若信息租金相对较低, 借款方则倾向于选择不披露信息, 以避免信贷机构的寻租行为。

推论 1 表明, 当信息租金较高时, 借款方通过披露数据信息可以获得更优惠的贷款条件, 实现双方共赢。而当信息租金较低时, 借款方隐瞒数据信息可以避免信贷机构利用数据优势寻租, 从而保护自身利益。近年来, 开放银行、个人信息保护立法等实践正是通过增强个体对自身数据的控制权, 来约束机构的数据滥用行为, 推论 1 为这类政策与法规的出台提供了理论依据。

五、不同数据所有权下的社会福利分析

本部分将比较信贷机构拥有数据所有权和借款方拥有数据所有权两种情况下, 借贷双方的期望效用和社会总福利的变化。其中, 社会总福利定义为借款方效用与信贷机构效用之和。

图 4 详细展示了两种所有权配置下借贷双方的期望效用和社会福利对比(证明见附录 XI)。

通过图 4, 我们可以得到命题 8:

命题 8 当 $X < 1$ 时, 两种数据所有权配置下借贷双方的期望效用与社会总福利均相同; 当 $X \geq 1$ 时, 将数据所有权赋予借款方会降低信贷机构的期望效用, 但同时提高借款方的期望效用与社会总福利。

命题 8 的结论意味着, 当信息租金较低时, 赋予借款方数据所有权能有效抑制信贷机构的寻租行为, 保护借款方利益, 并提升社会整体福利。这一结论与近年来的政策实践相呼应: 通过增强借款方对其自身金融数据的控制权, 有助于打破数据垄断, 促进信贷市场竞争, 从整体上提升社会总福利。以英国推行的开放银行政策为例, 其监管框架的核心正是要求金融机构开放客户交易数据接口, 客户可自主授权第三方获取自身金融信息, 从而在制度上赋予借款方对自身数据的控制权。截至 2024 年年底, 英国开放银行生态已拥有超过 1 209 万活跃用户, 并形成了规模约 41 亿英镑的相关产业。^① 这些现实证据表明, 数据控制权的重新配置能够通过打破数据垄断将理论预期的福利提升转化为可观测的社会经济效益。

^① Open Banking UK, “Open Banking and Smart Data Expert Predictions 2025”, Open Banking UK Official Website, <https://www.openbanking.org.uk/insights/open-banking-and-smart-data-expert-predictions-2025>, Accessed on: August 22, 2025.

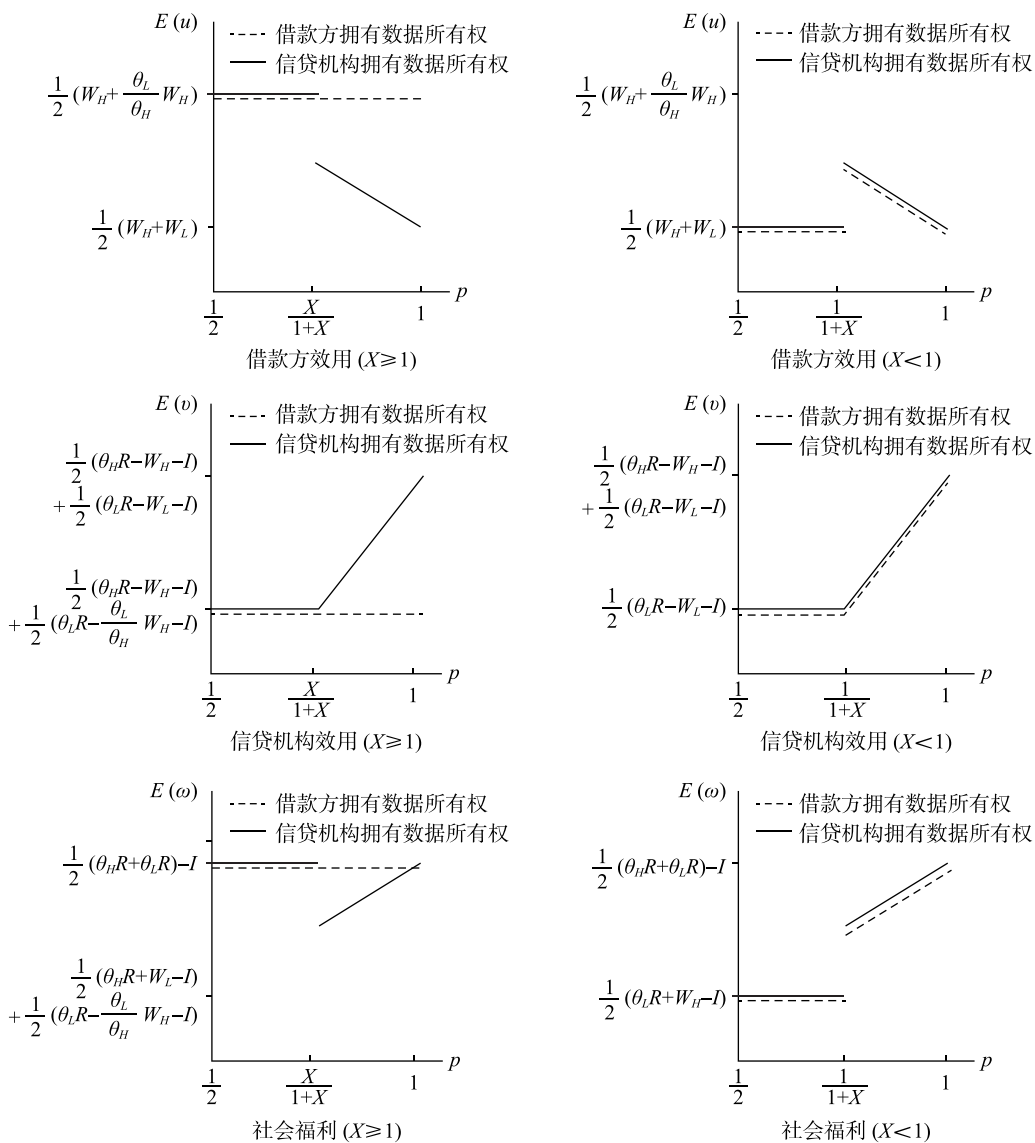


图 4 不同数据所有权下借贷双方期望效用和社会福利的比较

需要指出的是,在实际应用场景中,大数据的引入会增加信贷机构的信息搜集成本,借款者类型分布也可能是连续的。基准模型对大数据特征的刻画仍具有一定简化性。为此,本文在附录Ⅹ中讨论了大数据能够降低信贷机构信息收集成本的情形,在附录Ⅺ中考察了当借款方质量是连续分布的情形。分析结果表明,在上述更一般的设定下,本文的核心结论仍然稳健。

六、政策启示与研究局限

本研究的发现为数字金融监管与数据治理提供了重要的政策启示。政府部门应设计合理的机制,引导信贷机构将数据优势用于缓解信息不对称、提升金融服务效率,而非

进行寻租,从而避免社会福利的减损。具体而言,监管机构可通过加强借款方对个人数据的控制权,有效约束信贷机构的数据垄断与寻租行为,推动数据要素的合理利用,进而促进普惠金融的实现,缓解大数据可能带来的负面效应。

本研究仍存在若干局限,未来可以在以下方向进一步拓展。首先,可深入分析大数据在不同信贷场景下的具体应用效果,并研究市场上存在多家信贷机构竞争时大数据的应用对借贷双方的差异化影响,从而更加全面地评估普惠金融在小微企业和偏远地区产业中的作用机制;其次,可进一步研究数据控制权的实施机制,探索如何在技术架构、法律规范与市场实践中有效赋予并保障借款方对自身数据的控制权;最后,可探讨数据要素与其他金融创新工具的协同作用,并系统评估其对金融市场整体稳定性与长期发展的影响。这些方向的探索将有助于进一步完善数字金融的理论框架,并为相关政策的制定提供更全面的依据,从而更好地推动数字金融的健康发展,实现普惠金融的目标。

参 考 文 献

- [1] Acemoglu, D., A. Makhdoumi, A. Malekian, and A. Ozdaglar, "Too Much Data: Prices and Inefficiencies in Data Markets", *American Economic Journal: Microeconomics*, 2022, 14(4), 218-256.
- [2] Acemoglu, D., and P. Restrepo, "The Race Between Man and Machine: Implications of Technology for Growth, Factor Shares, and Employment", *American Economic Review*, 2018, 108(6), 488-1542.
- [3] Aghion, P., and P. Bolton, "An Incomplete Contracts Approach to Financial Contracting", *Review of Economic Studies*, 1992, 59(3), 473-494.
- [4] Begenau, J., M. Farboodi, and L. Veldkamp, "Big Data in Finance and The Growth of Large Firms", *Journal of Monetary Economics*, 2018, 97, 71-87.
- [5] Berg, T., V. Burg, A. Gombović, and M. Puri, "On The Rise of Fintechs: Credit Scoring Using Digital Footprints", *Review of Financial Studies*, 2020, 33(7), 2845-2897.
- [6] 蔡跃洲、陈楠, "新技术革命下人工智能与高质量增长、高质量就业", 《数量经济技术经济研究》, 2019 年第 5 期, 第 1—21 页。
- [7] Casamatta, C., "Financing and Advising: Optimal Financial Contracts with Venture Capitalists", *Journal of Finance*, 2003, 58(5), 2059-2085.
- [8] 陈彦斌、林晨、陈小亮, "人工智能、老龄化与经济增长", 《经济研究》, 2019 年第 7 期, 第 47—63 页。
- [9] Choi, J., D. Jeon, and B. Kim, "Privacy and Personal Data Collection with Information Externalities", *Journal of Public Economics*, 2019, 173, 113-124.
- [10] Cong, L., D. Xie, and L. Zhang, "Knowledge Accumulation, Privacy, and Growth in a Data Economy", *Management Science*, 2021, 67(10), 6480-6492.
- [11] DeMarzo, P., and D. Duffie, "A Liquidity-Based Model of Security Design", *Econometrica*, 1999, 67(1), 65-99.
- [12] Diamond, D. W., "Financial Intermediation and Delegated Monitoring", *Review of Economic Studies*, 1984, 51(3), 393-414.
- [13] Farboodi, M., and L. Veldkamp, "Long-run Growth of Financial Data Technology", *American Economic Review*, 2020, 110(8), 2485-2523.
- [14] Freixas, X., and J. Rochet, *Microeconomics of Banking*. Cambridge: MIT Press, 2008.
- [15] Fudenberg, D., and J. Tirole. *Game Theory*. Cambridge: MIT Press, 1991.
- [16] Gale, D., and M. Hellwig, "Incentive-Compatible Debt Contracts: The One-Period Problem", *Review of Eco-*

- nomics Studies*, 1985, 52(4), 647-663.
- [17] 龚强、班铭媛、张一林,“区块链、企业数字化与供应链金融创新”,《管理世界》,2021年第2期,第22—34页。
- [18] 龚强、王璐颖,“普惠金融、风险准备金与投资者保护——以平台承诺担保为例”,《经济学》(季刊),2018年第4期,第1581—1598页。
- [19] He, Z., J. Huang, and J. Zhou, “Open Banking: Credit Market Competition When Borrowers Own The Data”, *Journal of Financial Economics*, 2023, 147(2), 449-474.
- [20] Holmström, B., and J. Tirole, “Financial Intermediation, Loanable Funds, and the Real Sector”, *Quarterly Journal of Economics*, 1997, 112(3), 663-691.
- [21] Holmström, B., and J. Tirole, “Private and Public Supply of Liquidity”, *Journal of Political Economy*, 1998, 106(1), 1-40.
- [22] Holmström, B., and J. Tirole. *Inside and Outside Liquidity*. Cambridge: MIT Press, 2011.
- [23] Hong, J., “The Financing of Alliance Entrepreneurship”, *Journal of Business Venturing*, 2020, 35(1), 105916.
- [24] 黄益平、黄卓,“中国的数字金融发展:现在与未来”,《经济学》(季刊),2018年第4期,第1489—1502页。
- [25] Jagtiani, J., and C. Lemieux, “The Roles of Alternative Data and Machine Learning in Fintech Lending: Evidence from the Lending Club Consumer Platform”, *Financial Management*, 2019, 48(4), 1009-1029.
- [26] Jensen, M. C., and W. H. Meckling, “Theory of The Firm: Managerial Behaviour, Agency Costs and Ownership Structure”, *Journal of Financial Economics*, 1976, (3), 305-360.
- [27] Jones, C. I., and C. Tonetti, “Nonrivalry and the Economics of Data”, *American Economic Review*, 2020, 110(9), 2819-58.
- [28] Laffont, J., and D. Martimort, *The Theory of Incentives*, Princeton: Princeton University Press, 2009.
- [29] Li, S., H. Xiao, and D. Yao, “Contract Bargaining with a Risk-Averse Agent”, *The B. E. Journal of Theoretical Economics*, 2013, 13(1), 285-301.
- [30] 李三希、王泰茗、刘小鲁,“数据投资、数据共享与数据产权分配”,《经济研究》,2023年第7期,第139—155页。
- [31] Millon, M. H., and A. V. Thakor, “Moral Hazard and Information Sharing: A Model of Financial Information Gathering Agencies”, *Journal of Finance*, 1985, 40 (5), 1403-1422.
- [32] Pitchford, R., “Moral Hazard and Limited Liability: The Real Effects of Contract Bargaining”, *Economics Letters*, 1998, 61(2), 251-9.
- [33] Ru, H., and A. Schoar, “Do Credit Card Companies Screen for Behavioral Biases?”, *National Bureau of Economic Research*, 2016, No. w22360.
- [34] Stiglitz, J. E., and A. Weiss, “Credit Rationing in Markets with Imperfect Information”, *American Economic Review*, 1981, 71(3), 393-410.
- [35] Tirole, J., *The Theory of Corporate Finance*. Princeton: Princeton University Press, 2010.
- [36] Williams, B., “Search, Liquidity, and Retention: Screening Multidimensional Private Information”, *Journal of Political Economy*, 2021, 129(5), 1487-1507.
- [37] 谢丹夏、魏文石、李尧、朱晓武,“数据要素配置、信贷市场竞争与福利分析”,《中国工业经济》,2022年第8期,第25—43页。
- [38] Yang, M., “Optimality of Debt under Flexible Information Acquisition”, *Review of Economic Studies*, 2020, 87(1), 487-536.
- [39] Yang, M., and Y. Zeng, “Financing Entrepreneurial Production: Security Design with Flexible Information Acquisition”, *Review of Financial Studies*, 2019, 32(3), 819-863.
- [40] 张丰羽、汤珂,“数字时代的算法滥用及其规制研究”,《经济学动态》,2023年第2期,第71—87页。
- [41] 张剑虎、林平,“个人信息保护的市场机制与法律监管的福利效果”,《经济学》(季刊),2023年第4期,第1477—1494页。

Does Big Data Promote Inclusive Finance? An Optimal Contract Perspective

HONG Jieying

(Beihang University)

WU Yufan*

(Capital University of Economics and Business)

LI Sanxi

(Renmin University of China)

Abstract: This study incorporates big data into the traditional credit theory framework to explore its influence on credit contract design and the welfare of both lenders and borrowers. Our findings suggest that big data could intensify rent-seeking behavior among credit institutions, thereby harming borrowers' interests. Additionally, the impact of data precision on borrowers is non-monotonic—excessive data collection can erode their benefits. To mitigate these negative effects and promote inclusive finance, we propose that granting borrowers control over their data and restricting excessive data collection by credit institutions is a viable solution.

Keywords: big data; credit contract design; inclusive finance

JEL Classification: D53, D84, G14

* Corresponding Author: WU Yufan, School of Economics, Capital University of Economics and Business, No. 121 Zhangjialukou, Fengtai District, Beijing 100070, China; Tel: 86-18811796133; E-mail: wuyufan0103@126.com.