

# 数据要素对商业银行信贷效率的影响

## ——基于文本挖掘的实证研究

姜秋实<sup>1</sup>, 窦文章<sup>2,3</sup>

<sup>1</sup>苏银理财有限责任公司, 江苏 南京

<sup>2</sup>北京大学软件与微电子学院, 北京

<sup>3</sup>北京大学战略研究所, 北京

收稿日期: 2024年2月27日; 录用日期: 2024年3月19日; 发布日期: 2024年4月30日

### 摘要

数字经济时代, 数据成为与土地、劳动力、资本、技术并列的生产要素, 正深刻改变着各行业的生产方式和社会治理方式。商业银行是数据密集型行业, 数据要素作为推动自身转型发展的关键抓手受到了极大的重视。本文通过文本挖掘的方法, 选取2013~2021年全国34家上市商业银行年报中的原始词频数据, 构建了数据要素指数并提出了研究假设, 然后对商业银行要素指数与信贷效率之间的关系进行了实证检验。在此基础上, 以中国工商银行为案例, 分析了数据要素运用程度对信贷效率的影响。

### 关键词

数据要素, 商业银行, 信贷效率, 文本挖掘

# The Influence of Data Factors on the Credit Efficiency of Commercial Banks

## —An Empirical Study Based on Text Mining

Qiushi Jiang<sup>1</sup>, Wenzhang Dou<sup>2,3</sup>

<sup>1</sup>Suyin Wealth Management Co., Ltd., Nanjing Jiangsu

<sup>2</sup>School of Software and Microelectronics, Peking University, Beijing

<sup>3</sup>Institute of Strategic Studies, Peking University, Beijing

Received: Feb. 27<sup>th</sup>, 2024; accepted: Mar. 19<sup>th</sup>, 2024; published: Apr. 30<sup>th</sup>, 2024

文章引用: 姜秋实, 窦文章. 数据要素对商业银行信贷效率的影响[J]. 现代管理, 2024, 14(4): 753-763.

DOI: 10.12677/mm.2024.144089

## Abstract

In the context of the digital economy, data has emerged as a key factor of production alongside traditional inputs such as land, labor, capital, and technology. Its influence is profoundly changing the mode of production and social governance of various industries. Commercial banks, in particular, are data-intensive industries, underscoring the importance of data elements as a key driver for their transformation and growth. This paper constructs a data factor index and puts forward research hypotheses by selecting raw word frequency data in the annual reports of 34 listed commercial banks in China from 2013 to 2021 through text mining. The empirical testing of the relationship between the factor index and credit efficiency of commercial banks is then conducted. On this basis, the impact of the degree of data factor utilization on credit efficiency is analyzed with Industrial and Commercial Bank of China as a case study.

## Keywords

Data Factors, Commercial Banks, Credit Efficiency, Text Mining

Copyright © 2024 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

随着技术的不断变革与发展, 当今世界已然进入数字经济的时代。在数字经济时代背景下, 数据要素作为核心所在, 发挥着创新引擎的作用, 为经济发展注入源源不断的动力。伴随数据要素市场的发展, 数据要素对银行业乃至整个金融市场的影响日益深化, 特别是重塑传统金融的业务模式与服务, 从风险控制、支付结算、产品设计、业务管理、技术水平几个方面对商业银行的信贷效率带来了深远影响。

本文通过文本挖掘的方法, 选取了2013~2021年间34家上市商业银行年报中的原始词频数据, 并经过一系列数据整理得到并构建了具有各家银行个体因素的数据要素指数, 并对商业银行数据要素指数与信贷效率的关系进行了研究, 检验了不同银行类型和不同资产规模的商业银行数据要素水平对其信贷效率的影响差异, 随后在检验稳健性后进一步探究了两者之间的影响机制。

## 2. 理论分析与研究假设

随着大数据、区块链、云计算等新兴科学技术不断取得突破和发展, 作为生产资料“数据”已经成为重要的经济资源。数据的生产和合理利用、数据技术和有关产业创新已经逐步成为全世界经济发展的重要推动力。于立等以生产要素理论作为切入点, 重新定义了包括数据要素在内的生产要素基本概念, 其次探讨了数据要素的共性和特性[1]。李政等从政治经济学的角度分析了数据作为生产要素如何参与分配, 认为数据作为生产要素参与分配, 要遵循马克思主义政治经济学原理和“以按劳分配为主体, 多种分配方式并存”的分配原则[2]。内在属性方面, 戚聿东等指出, 数据要素是数字经济时代最为重要的生产要素, 并且进一步指出数据独具非竞争性、非排他性、可重复性和网络外部性等特点[3]。

### 2.1. 数据要素对商业银行信贷效率的影响

商业银行作为传统金融体系的中坚力量, 近年来受到科学技术发展的影响程度较深, 关于数据要素

与商业银行的相关研究也纷至沓来。目前已经有研究的切入点包括了经营绩效、风险管理、盈利能力等。数据的发展运用对商业银行收入盈利[4]、管理能力[5]、服务质量[6]、防范风险[7]带来提升;数据要素的发展从挤占市场份额[8] [9]、挤出效应[10]等方面给商业银行的发展带来负面的影响。现有研究就数据要素对于银行信贷效率的影响集中在以下两种观点:一是数据要素可以通过大数据分析[11] [12]提升银行风控能力[13],对商业银行信贷效率有正向影响。二是数据要素对商业银行信贷效率有负向影响[14] [15] [16]。

从长尾理论角度来看,商业银行可以通过数据挖掘和分析,更加精准地了解长尾小微客户的需求和行为,进而为他们提供更加个性化、贴心的金融服务。唐也然认为,银行对数字技术的运用可以提升商业银行信贷普惠性和配置效率[17]。根据信息不对称理论,数据要素不断发展有助于降低银行信贷业务中的道德风险,改善了逆向选择现象,对促进降低商业银行的不良贷款率有积极作用。李权洋认为,商业银行可以通过移动支付、云计算、大数据等技术手段数据有效地解决由于信息不对称所带来的风险损失,进而降低银行营运成本[18]。从金融中介理论来看,商业银行目前是我国最重要也是最常见的一种金融机构。Muganyi 等认为,金融科技的发展可以促进金融创新、提高金融效率和降低金融成本,从而推动金融业向着更加普惠、高效和可持续的方向发展[19]。由此,提出第一个假设:

**假设 1:** 商业银行对数据要素的使用提高了其信贷效率。

商业银行的信贷配置效率与商业银行的类型有关。国有银行、股份制银行、城商行、农商行等不同类型商业银行的数据战略不同,目标客户人群存在显著差异。顾海峰指出,相较于其他类型的商业银行,大型国有商业银行拥有较高的信贷配置效率[20]。股份制银行有着最具活力与创造力,往往具有相对更高的信贷配置效率。城商行往往在其所在城市内占有举足轻重的地位,同时信贷业务的区域也相对集中,会导致其客户集中度可能较高,会影响其信贷配置的效率。农商行聚焦农村与普惠金融,触及长尾客户,相对来说,对数据要素的使用程度不高,较难通过技术手段提升其信贷效率。由此,提出第二个假设:

**假设 2:** 不同类型的商业银行,其运用数据要素对信贷配置效率的影响效果存在差异。

商业银行信贷配置效率与商业银行规模有关。资产规模较大的银行因其规模大、服务范围广最先受到数据要素的冲击,与资产规模较小的银行相比,大银行实力强、技术水平高,具备发展数据要素的先发优势,其面对数据要素浪潮时投入大量资金成立数据要素中心,在发展初期可能会对其盈利造成负面影响。规模较小的商业因其自身特殊的发展路径,服务客群主要是当地广大群众,客户黏性较大,受到数据要素冲击的幅度较小。面对数据要素的冲击,不同规模的银行反应有差异,其中数据要素加重了大型商业银行风险承担,而中小型商业银行表现则较为迟缓。基于此,提出第三个假设:

**假设 3:** 不同资产规模的商业银行,其运用数据要素对信贷配置效率的影响效果存在差异。

## 2.2. 数据要素对商业银行的影响机制

影响机制方面,风险承担是数据要素对商业信贷效率的作用渠道之一。郭品等发现,对于数据要素的应用可以促进商业银行提高借款人信息的透明度和准确性,从而帮助商业银行更好地评估借款人的信用风险[21]。例如,通过分析经济趋势、市场状况和客户行为的数据,商业银行可以评估风险并制定策略来减轻风险。通过分析客户信用记录、交易行为等数据,银行可以更好地评估客户的信誉度,更有效地分配信贷。通过使用数据分析工具来监控交易行为,商业银行可以发现潜在的欺诈行为,并采取积极的措施来预防。通过分析客户数据、交易数据及其他数据来源,银行可以识别客户需求、发现新市场,使得服务产品的开发与市场需求变化保持紧密联系,以减轻产品服务创新失败的风险。商业银行运用数据分析工具和技术,可以显著提高其防范风险的能力,进一步提高信贷业务的配置效率。通过利用数据,银行可以做出更明智的决策,更有效地分配信贷,并为客户提供更有针对性和个性化的服务。由此提出

第四个假设:

**假设 4:** 商业银行对数据要素的使用可以提高银行风险承担能力, 从而提高其信贷效率。

信贷结构调整也是数据要素作用于商业银行信贷效率的渠道之一。宋献中等发现, 信贷集中度对商业银行绩效和信贷效率具有显著的正向影响[22]。商业银行可以利用数据要素评估贷款项目的可行性。例如, 通过分析贷款项目的市场需求、行业前景、财务状况等数据要素, 银行可以评估项目的风险水平, 确定项目的还款期限、利率等信贷条件, 并最终优化其信贷结构, 提高信贷效率。在投资组合优化方面, 通过分析贷款组合的数据, 商业银行可以确定潜在客户高度集中的领域, 并采取措施降低这种集中度。例如, 银行可以利用数据来确定贷款高度集中的行业或地区, 并制定策略来降低这种集中度。这有助于降低贷款违约风险, 提高信贷业务的配置效率。其次, 通过使用数据分析工具监测贷款组合, 银行可以发现潜在问题, 并采取积极措施解决这些问题。例如, 若某商业银行发现它在某个特定行业或地区的贷款高度集中, 它可以采取相应措施降低集中度。通过持续监测贷款组合并报告潜在风险, 银行可以降低贷款违约风险, 提高信贷业务的配置效率。总之, 商业银行运用数据分析工具和技术, 可以显著降低客户贷款集中度, 提高信贷业务配置效率。通过利用数据, 银行可以做出更明智的决策, 更有效地分配信贷, 并为客户提供更有针对性和个性化的服务。这可以帮助银行在不断变化的市场中保持竞争力, 并与客户建立长期关系。由此提出五个假设:

**假设 5:** 商业银行对数据要素的使用可以优化银行信贷结构, 从而提高其信贷效率。

### 3. 研究设计

#### 3.1. 构建数据要素指数

近年来, 国内不少机构为了更好地衡量和量化数字经济的发展水平, 尝试构建出诸如互联网金融指数、金融科技指数等科技类指数。2015 年, 北京大学互联网金融研究中心联合蚂蚁金服、中国人民银行等机构共同编制并发布了中国首个北京大学互联网金融发展指数。2017 年, 深圳交易所运用 71 家金融科技公司的数据测得了香蜜湖金融科技指数这一我国首个金融科技指数。目前学界对于数据要素的定量研究较少, 而基于金融科技的相关定量研究较多, 考虑到数据要素指数与金融科技指数两者在内容上紧密联系, 从金融功能理论入手, 通过划分依据互联网金融模式建立原始词库, 利用百度引擎计算词频、量化原始词库, 借助因子分析法的方式构建互联网金融指数的方式, 构建数据要素指数。

本研究采用规范度更高、可靠性更强、可信度更好的商业银行年度报告信息作为数据来源。首先本研究采用 Python 中的 PyPDF2 库, 对 PDF 文件进行拆分合并解密等操作, 以便对商业银行年报进行文本数据的分析。其次, 借助文本挖掘中的 Word2vec 对银行年报进行分词处理, 生成可用于执行自然语言处理任务的词的向量表示。最后, 本研究将数据要素原始词库分为风险控制、支付结算、产品设计和业务管理、技术水平五个维度, 梳理出 20 个原始关键词, 构建数据要素关键词(见表 1)。

**Table 1.** Thesaurus of data index keywords

**表 1.** 数据要素指数关键词词库

维度	风险控制	支付结算	产品设计	业务管理	技术水平
原始词库	用户画像	移动支付	互联网理财	电子交易	数字化
	人脸识别	在线支付	网络贷款	线上交易	大数据
	数据挖掘	刷脸支付	智能投顾	手机银行	区块链
	指纹解锁	数字货币	互联网保险	网上银行	云计算

资料来源: 本研究整理, 下同。



最后, 合成数据要素指数。本文采取姜世超等合成银行数据要素指数时所采用的熵值法[23]。在文本挖掘获取了词库中各个关键词词频后, 进一步采取熵值法合成数据要素指数:

第一, 标准化处理数据要素指数关键词。用  $X_{\alpha ij}$  表示第  $\alpha$  个样本年份, 第  $i$  家银行 ( $i=1, 2, 3, \dots, 34$ ), 第  $j$  个关键词 ( $j=1, 2, 3, \dots, 20$ ) 出现的次数。由于每个数据要素关键词的出现次数是非负整数, 因此进行标准化处理的公式是:

$$Y_{\alpha ij} = (X_{\alpha ij} - X_{\min}) / (X_{\max} - X_{\min}) \quad (1)$$

其中, 第  $j$  个数据要素关键词在 2013~2021 年 34 家银行年度报告中出现的最大次数为  $X_{\max}$ , 最少次数为  $X_{\min}$ 。  $Y_{\alpha ij}$  表述标准化之后数据, 其数值分布在  $[0, 1]$  闭区间上。经过第一步处理之后, 部分  $Y_{\alpha ij}$  的数值为 0, 本研究对所有标准后的数据进行了平移处理, 有  $Z_{\alpha ij} = Y_{\alpha ij} + A$ ,  $A = 0.01$ 。

第二, 计算 34 家商业银行各自对于数据要素指数关键词的贡献度:

$$P_{\alpha ij} = Z_{\alpha ij} / \sum_{\alpha=1}^h \sum_{i=1}^m Z_{\alpha ij} \quad (2)$$

其中,  $h$  指样本期间年度数,  $m$  指商业银行个数, 分别为 9 和 34。

第三, 得出数据要素指数关键词的熵值:

$$e_j = -1 / \ln(h * m) \sum_{\alpha=1}^h \sum_{i=1}^m P_{\alpha ij} \ln(P_{\alpha ij}) \quad (3)$$

第四, 得出数据要素指数关键词的差异性系数:

$$d_j = 1 - e_j \quad (4)$$

第五, 得出各个数据要素指数关键词权重:

$$W_j = d_j / \sum_{j=1}^n d_j \quad (5)$$

第六, 测得各年份各商业银行数据要素指数, 公式为:

$$S_{\alpha i} = W_j * P_{\alpha ij} \quad (6)$$

将文本挖掘分析的结果进行上述处理后, 得到了 34 家银行 2013~2021 年的数据要素指数。

## 3.2. 样本选取与数据来源

在充分考虑了数据的可读性、可获取性之后, 本文在上市商业银行中选取了 34 家商业银行 2013 年~2021 年的年报数据作为研究样本, 其中包括 6 家国有商业银行、9 家股份制商业银行、11 家城商行和 8 家农商行。银行相关的财务数据来源于 Wind 金融终端, 缺失数据通过查找相应商业银行年报和同花顺数据库进行补充。宏观经济数据来源于国家统计局, 数据要素指数来源于作者整理计算。

## 3.3. 变量定义

本文参照王秀丽等[24]的方法, 将不良贷款率作为衡量信贷效率的指标。高不良贷款率会使银行面临更多的风险和损失, 因此银行可能会减少信贷额度或者提高贷款利率, 从而限制贷款规模和范围, 影响信贷效率。文本核心解释变量为自行构建的数据要素指数。本文选取银行存贷比(LDR)、银行资产收益率(ROA)、银行收入成本比(CIR)、银行资本充足率(CAR)、银行总资产规模(Lnsize)、宏观经济指标(GDP)和货币政策(M2)作为七个控制变量。上述变量汇总在下表 2 中。

## 3.4. 实证模型

为了验证研究假设 1, 本文结合前文理论分析和相关参考文献, 加入了对时间、个体因素的考量后, 设定如下的多元回归模型:

**Table 2.** Variables sorting  
**表 2.** 变量梳理

类型	名称	符号	变量内涵
被解释变量	不良贷款率	<i>NPL</i>	银行不良贷款率
核心解释变量	数据要素指数	<i>DataF</i>	数据要素指数
调节变量	风险承担能力	<i>RISK</i>	加权风险资产取对数
	银行信贷结构	<i>LNC</i>	最大十家客户贷款占资本净额比例
控制变量	银行存贷比	<i>LDR</i>	银行贷款总额/存款总额
	银行资产收益率	<i>ROA</i>	银行总负债/总资产
	银行收入成本比	<i>CIR</i>	银行经营成本/营业收入
	银行资本充足率	<i>CAR</i>	银行资本总额/风险加权资产
	银行总资产规模	<i>Lnsiz</i>	银行总资产取对数
	宏观经济增长	<i>GDP</i>	GDP 同比增长率
	货币政策变化	<i>M2</i>	M2 同比增长率

$$NPL_{it} = \alpha + \beta_1 DataF_{i,t} + \beta_2 LDR_{i,t} + \beta_3 ROA_{i,t} + \beta_4 CIR_{i,t} + \beta_5 CAR_{i,t} + \beta_6 Lnsiz_{i,t} + \beta_7 GDP_{i,t} + \beta_8 M2_{i,t} + \delta_t + \mu_i + \varepsilon_{i,t} \quad (7)$$

其中,  $\alpha$  为截距项,  $\beta_1$  为核心解释变量数据要素指数的系数,  $\beta_2 \sim \beta_8$  为各个控制变量的系数,  $\mu_i$  代表商业银行个体固定效应,  $\delta_t$  代表时间固定效应,  $\varepsilon_{i,t}$  代表随机扰动项,  $i$  代表银行数量,  $t$  代表时间。

假设 2 和假设 3 分别根据商业银行类型和商业银行规模对样本进行异质性检验, 并通过对比核心解释变量 *DataF* 前系数的组间差异以检验假设。

在主回归模型的基础之上, 本文进一步研究数据要素通过何种渠道影响作用于商业银行的信贷配置效率, 在回归方程中加入调节变量、数据要素指数与调节变量的交叉相乘项。

为检验商业银行风险承担能力的调节效应(假设 4), 本文构建如下模型:

$$NPL_{it} = \alpha + \beta_1 DataF_{i,t} + \gamma_1 RISK_{i,t} + \gamma_2 DataF_{i,t} * RISK_{i,t} + \beta_2 LDR_{i,t} + \beta_3 ROA_{i,t} + \beta_4 CIR_{i,t} + \beta_5 CAR_{i,t} + \beta_6 Lnsiz_{i,t} + \beta_7 GDP_{i,t} + \beta_8 M2_{i,t} + \delta_t + \mu_i + \varepsilon_{i,t} \quad (8)$$

其中,  $RISK_{i,t}$  表示银行  $i$  在第  $t$  年的风险承担能力,  $DataF_{i,t} * RISK_{i,t}$  表示商业银行风险承担能力在数据要素影响银行信贷效率的调节效应。

为检验商业银行信贷结构的调节效应(假设 5), 本文构建如下模型:

$$NPL_{it} = \alpha + \beta_1 DataF_{i,t} + \gamma_1 LNC_{i,t} + \gamma_2 DataF_{i,t} * LNC_{i,t} + \beta_2 LDR_{i,t} + \beta_3 ROA_{i,t} + \beta_4 CIR_{i,t} + \beta_5 CAR_{i,t} + \beta_6 Lnsiz_{i,t} + \beta_7 GDP_{i,t} + \beta_8 M2_{i,t} + \delta_t + \mu_i + \varepsilon_{i,t} \quad (9)$$

其中,  $LNC_{i,t}$  表示银行  $i$  在第  $t$  年的信贷集中度情况,  $DataF_{i,t} * LNC_{i,t}$  表示商业银行信贷结构在数据要素影响银行信贷效率的调节效应。 $\gamma_1$  衡量商业银行信贷结构对其信贷效率的影响,  $\gamma_2$  衡量商业银行信贷结构与数据要素水平的交乘项影响的大小。

## 4. 实证分析

### 4.1. 描述性统计

表 3 列出了相关变量的观测值、均值、标准差、最小值和最大值, 可见, 各变量标准差都比它们平均值小。表明原始数据波动范围较小, 数据质量良好。

**Table 3.** Descriptive statistics of variables**表 3.** 变量描述性统计

变量名称	观测值	平均值	标准差	最小值	最大值
<i>NPL</i>	306	1.37	0.41	0.51	2.90
<i>DataF</i>	306	0.24	0.12	0.10	0.71
<i>LDR</i>	306	0.74	0.14	0.29	1.16
<i>ROA</i>	306	0.95	0.22	0.50	1.76
<i>CIR</i>	306	0.31	0.06	0.19	0.67
<i>CAR</i>	306	13.37	1.59	8.84	18.02
<i>Lsize</i>	306	9.52	1.76	6.41	12.77
<i>GDP</i>	306	8.83	2.89	2.74	13.38
<i>M2</i>	306	10.49	2.06	8.10	13.60

## 4.2. 基准回归

在全样本回归中, 本文先根据静态面板的混合效应、固定效应和随机效应进行估计, 回归结果如表 4 前三列所示, 根据 Breusch 与 Pagan 拉格朗日乘子检验的结果, 在混合效应模型与随机效应模型之间选择随机效应模型。本文进一步根据 Hausman 检验的结果判断回归模型在固定效应与随机效应间如何选择, Hausman 检验的原假设为个体效应与回归变量无关, 检验结果显示 P 值为 0.0000, 故强烈拒绝原假设, 故选择固定效应模型。

**Table 4.** Baseline regression results**表 4.** 基准回归结果

变量名	混合回归	固定效应	随机效应	双向固定效应
<i>DataF</i>	-1.127** (0.427)	-1.677*** (0.388)	-1.603*** (0.379)	-1.094*** (0.538)
控制变量	Yes	Yes	Yes	Yes
<i>_cons</i>	0.309 (0.475)	5.625*** (1.070)	1.147** (0.537)	5.696*** (1.951)
时间固定	--	--	--	Yes
个体固定	--	--	--	Yes
样本量	306	306	306	306
调整后的 R <sup>2</sup>	0.197	0.257		0.426
豪斯曼检验		Prob > chi <sup>2</sup> = 0.0000		

注: \*, \*\*, \*\*\* 分别表示在 10%、5%、1% 显著性水平下显著; 括号内数值为回归标准差, 下同。

本研究最终选择加入个体效应与时间效应的双向固定效应模型进行回归, 并将回归结果记录在表 4 第四列中。根据回归结果, 核心解释变量数据要素指数的系数为-1.094, 并且在 1% 的水平上显著。说明解释变量数据要素指数与被解释变量商业银行信贷效率之间为显著的负相关关系, 因此假设 1: 商业银行对数据要素的使用提高了其信贷效率得到了验证。

### 4.3. 稳健性检验

为了验证上述估计结果的稳健性。本文将被解释变量不良贷款率(NPL)替换成不良贷款拨备覆盖率(NPC), 以此为基础进行稳健性检验。不良贷款拨备覆盖率用于衡量银行应对不良贷款风险的能力。较高的不良贷款拨备覆盖率意味着较高的信贷效率, 稳健性检验的具体结果如表 5 所示。

**Table 5.** Robustness test of the replacement of credit efficiency measures  
**表 5.** 更换信贷效率衡量方式稳健性检验

变量名	(1)	(2)	(3)
<i>DataF</i>	0.847*** (0.214)	1.878*** (0.281)	1.177*** (0.315)
控制变量	No	Yes	Yes
_cons	5.268*** (0.053)	1.040 (0.776)	3.187** (1.440)
样本量	306	306	306
调整后的 R <sup>2</sup>	0.064	0.319	0.455
个体固定	--	--	Yes
时间固定	--	--	Yes

核心解释变量 *DataF* 的估计系数是我们主要关心的系数。列(1)是单变量回归的结果; 列(2)是加入了一系列控制变量后的回归结果; 列(3)表示在双向固定效应模型下的回归结果。综上, *DataF* 的系数在 1% 的显著度水平下为正, 数据要素可以显著提升商业银行信贷效率, 这一结论与基准回归相同。

### 4.4. 异质性分析

为了进一步研究不同类型的商业银行在运用数据要素对其信贷效率影响的异质性, 本文选择分组样本回归的方法。本文将 34 家商业银行样本按照银行类型为国有商业银行、股份制商业银行、城市商业银行与农村商业银行四组, 分别带入回归模型, 回归结果如表 6 所示。

**Table 6.** Regression results of sample grouping by bank category  
**表 6.** 按银行类别分组样本回归结果

变量名	国有商业银行	股份制银行	城市商银行	农村商银行
<i>DataF</i>	-4.096*** (0.753)	-6.745*** (1.188)	-2.098*** (0.651)	-0.588 (1.399)
控制变量	Yes	Yes	Yes	Yes
_cons	-26.210*** (7.593)	4.826** (2.031)	1.542 (1.335)	12.802*** (1.898)
样本量	54	81	99	72
调整后的 R <sup>2</sup>	0.467	0.648	0.678	0.558



从分组样本实证结果中可以看出,不同类型的商业银行存在组间差异。国有商业银行、股份制商业银行和城商行均在 1%的水平上显著,而农村商业银行组数据要素指数的估计系数不显著,说明国有商业银行、股份制商业银行和城市商业银行运用数据要素对信贷效率提升作用显著,而农村商业银行运用数据要素对信贷效率的提高不显著。故验证本文研究的假设 2:不同类型的商业银行,其运用数据要素对信贷配置效率的影响效果存在差异。

参考喻微锋等以商业银行资产规模作为分类标准[25],本文将 34 家商业银行样本按照资产规模的大小分为两组。其中,资产规模在样本中排名前 50%的银行纳入大规模的银行组,剩下的银行纳入小规模的银行组。具体的回归结果记录在表 7 中。

**Table 7.** Regression results of sample grouping by bank size

**表 7.** 按银行规模分组样本回归结果

变量名	大规模银行组	小规模银行组
<i>DataF</i>	-5.439*** (0.459)	-2.749*** (1.209)
控制变量	Yes	Yes
<i>_cons</i>	-0.679 (2.037)	1.287 (1.402)
样本量	153	153
调整后的 R <sup>2</sup>	0.461	0.340

结果显示,当商业银行按照资产规模大小分为规模大和规模小两组时,数据要素指数的估计系数显著为负,说明资产规模较大的商业银行与资产规模较小的银行相比,数据要素指数对信贷效率的影响更大,即验证了本文的研究假设 3:不同资产规模的商业银行,其运用数据要素对信贷效率的影响效果存在差异。

#### 4.5. 影响机制分析

本文进一步研究数据要素影响商业银行信贷效率的潜在传导机制。

一是数据要素提高了商业银行的抗风险能力,进而提高了银行的信贷效率。本文借鉴江曙霞等[26]的研究,选取加权风险资产(RISK)表示商业银行的风险承担能力,RISK 越高说明商业银行风险承担能力越强。二是数据要素降低了银行的贷款集中度,进而提高了银行的信贷效率。本文选用商业银行最大十家客户贷款占资本净额(LNC)表示商业银行贷款集中度,该比率越高贷款集中度越高。

结合第四章提到的检验调节效应的模型进行回归,结果展示在表 8 中。

在第(1)列中,加权风险资产的估计系数为负,在 10%的水平上显著,表明商业银行风险承担能力越高,其信贷配置效率就越高,但这一影响作用有限。在第(2)列中,交互项系数的显著负值表明,对于风险承担能力较强的商业银行而言,数据要素对其信贷配置效率的促进作用越大。可以验证本研究假设 4:商业银行对数据要素的使用可以提高风险承担能力,从而提高其信贷效率。在第(3)列中,衡量商业银行贷款集中度的指标显著为正,说明商业银行信贷结构中贷款集中度越低,其信贷配置效率越高。在第(4)列中,数据要素指数与银行信贷结构指标的交互项系数显著为正,可以验证文本研究假设 5:商业银行对数据要素的使用可以优化信贷结构,从而提高其信贷效率。

**Table 8.** Mechanism test results  
**表 8.** 机制检验结果

变量名	(1)	(2)	(3)	(4)
<i>DataF</i>		-15.248*** (2.184)		-2.056*** (0.460)
<i>RISK</i>	-0.029* (0.017)	-0.130*** (0.025)		
<i>DataF * RISK</i>		-1.161*** (0.185)		
<i>LNC</i>			0.006** (0.002)	0.011*** (0.004)
<i>DataF*LNC</i>				0.036** (0.018)
<b>控制变量</b>	No	Yes	No	Yes
<i>_cons</i>	1.108*** (0.154)	5.421*** (1.006)	1.481*** (0.048)	6.020*** (1.065)
样本量	306	306	306	306
调整后的 R <sup>2</sup>	0.113	0.351	0.100	0.476

## 5. 结论与建议

### 5.1. 研究结论

本文得出以下结论：第一，不同商业银行其数据要素发展程度不同。国有商业银行与股份制商业银行由于体量规模较大且具备先发优势，故其数据要素发展程度较城市商业银行和农村商业银行高。第二，商业银行对于数据要素的使用有助于提高其信贷效率。在实证分析中，本文采用自行构建的数据要素指数作为解释变量，选用不良贷款率来衡量信贷效率，得到了这一结论。第三，不同类型、不同资产规模的商业银行，其运用数据要素对信贷配置效率的影响效果存在差异。第四，商业银行对数据要素的使用可以通过提高银行风险承担能力和优化信贷结构两方面提高信贷效率。

### 5.2. 对策建议

对于商业银行而言，发展数据要素可以显著提升信贷效率。银行可以加大建设数字化平台等软硬件基础设施的力度；具备一定实力的银行可以考虑成立金融科技子公司，集中力量提升数据要素水平；此外，银行可以注重金融科技类人才的引进与内部职工的培训与转型，以适应不断提升的数据要素水平。

对于整个银行业来说，应当承担一定的社会责任，为推进数据要素市场的建设探索出针对金融业的发展模式，鼓励同业银行之间就数据要素相互交流，不断推动数据资产的开放与共享，促进银行业不断向好发展。行业内监管机构应当积极探索出与数据要素发展水平相适配的监管体系，防范发展数据带来的信息技术风险，也要防止系统性金融风险。

## 参考文献

- [1] 于立, 王建林. 生产要素理论新论——兼论数据要素的共性和特性[J]. 经济与管理研究, 2020, 41(4): 62-73.

- <https://doi.org/10.13502/j.cnki.issn1000-7636.2020.04.006>
- [2] 李政, 周希祺. 数据作为生产要素参与分配的政治经济学分析[J]. 学习与探索, 2020(1): 109-115.
- [3] 戚聿东, 刘欢欢. 数字经济下数据的生产要素属性及其市场化配置机制研究[J]. 经济纵横, 2020(11): 63-76.  
<https://doi.org/10.16528/j.cnki.22-1054/f.202011063>
- [4] 刘忠璐. 互联网金融对商业银行风险承担的影响研究[J]. 财贸经济, 2016, 37(4): 71-85, 115.  
<https://doi.org/10.19795/j.cnki.cn11-1166/f.2016.04.007>
- [5] 徐晓萍, 李弘基, 戈盈凡. 金融科技应用能够促进银行信贷结构调整吗?——基于银行对外合作的准自然实验研究[J]. 财经研究, 2021, 47(6): 92-107.
- [6] 巴曙松, 慈庆琪, 郑焕卓. 金融科技浪潮下, 银行业如何转型[J]. 当代金融研究, 2018(2): 22-29.
- [7] 王依婷, 王世文, 周雷. 利率市场化与金融科技发展降低了商业银行风险吗? [J]. 金融理论探索, 2021(2): 29-40.
- [8] 王均山. 金融科技对商业银行零售业务盈利的影响研究[J]. 湖北社会科学, 2020(10): 81-88.  
<https://doi.org/10.13660/j.cnki.42-1112/c.015481>
- [9] 李向前, 贺卓昇. 金融科技发展对商业银行影响研究[J]. 现代经济探讨, 2021(2): 50-57.  
<https://doi.org/10.13891/j.cnki.mer.2021.02.008>
- [10] 熊健, 张晔, 董晓林. 金融科技对商业银行经营绩效的影响: 挤出效应还是技术溢出效应? [J]. 经济评论, 2021, 229(3): 89-104.
- [11] 姜增明, 陈剑锋, 张超. 金融科技赋能商业银行风险管理转型[J]. 当代经济管理, 2019, 41(1): 85-90.  
<https://doi.org/10.13253/j.cnki.ddjgl.2019.01.012>
- [12] Lee, C.C., Li, X.R., Yu, C.H. and Zhao, J.S. (2021) Does Fintech Innovation Improve Bank Efficiency? Evidence from China's Banking Industry. *International Review of Economics and Finance*, 74, 468-483.  
<https://doi.org/10.1016/j.iref.2021.03.009>
- [13] 邱晗, 黄益平, 纪洋. 数据要素对传统银行行为的影响——基于互联网理财的视角[J]. 金融研究, 2018(11): 17-29.
- [14] 郑霄鹏, 刘文栋. 互联网金融对商业银行的冲击及其对策[J]. 现代管理科学, 2014(2): 78-80.
- [15] Wang, R., Liu, J. and Luo, H.R. (2020) Fintech Development and Bank Risk Taking in China. *The European Journal of Finance*, 27, 397-418. <https://doi.org/10.1080/1351847X.2020.1805782>
- [16] Sun, F. (2021) Development and Transformation Strategies of Banking Industry under the Background of Fintech. *E3S Web of Conferences*, 253, Article ID: 03068. <https://doi.org/10.1051/e3sconf/202125303068>
- [17] 唐也然. 商业银行发展金融科技如何影响信贷业务?——基于上市银行年报文本挖掘的证据[J]. 金融与经济, 2021(2): 38-44. <https://doi.org/10.19622/j.cnki.cn36-1005/f.2021.02.005>
- [18] 李权洋. 金融科技在商业银行发展中的作用探究[J]. 科技经济导刊, 2018(31).
- [19] Muganyi, T., Yan, L., Yin, Y., et al. (2022) Fintech, Regtech and Financial Development: Evidence from China. *Financial Innovation*, 8, Article No. 29. <https://doi.org/10.1186/s40854-021-00313-6>
- [20] 顾海峰. 信贷配给、银保协作与银行信贷效率改进——基于信贷效用熵的分析视角[J]. 经济与管理研究, 2013(4): 81-87. <https://doi.org/10.13502/j.cnki.issn1000-7636.2013.04.014>
- [21] 郭品, 沈悦. 互联网金融对商业银行风险承担的影响: 理论解读与实证检验[J]. 财贸经济, 2015(10): 102-116.  
<https://doi.org/10.19795/j.cnki.cn11-1166/f.2015.10.009>
- [22] 宋献中, 禹天寒. 商业银行高管薪酬、风险承担水平与银行绩效——基于信贷集中度与损失类贷款研究[J]. 华东经济管理, 2018, 32(5): 172-176. <https://doi.org/10.19629/j.cnki.34-1014/f.171105002>
- [23] 姜世超, 刘畅, 胡永宏, 马敬元. 空间外溢性和区域差异化视角下银行数据要素的影响因素——基于某大型国有商业银行县域数据的研究[J]. 中央财经大学学报, 2020(3): 19-32.
- [24] 王秀丽, 鲍明明, 张龙天. 金融发展、信贷行为与信贷效率——基于我国城市商业银行的实证研究[J]. 金融研究, 2014(7): 94-108.
- [25] 喻微锋, 周黛. 互联网金融、商业银行规模与风险承担[J]. 云南财经大学学报, 2018, 34(1): 59-69.
- [26] 江曙霞, 陈玉婵. 货币政策、银行资本与风险承担[J]. 金融研究, 2012(4): 1-16.