

企业数字化转型的测度难题: 基于大语言模型的新方法与新发现^{*}

金星晔 左从江 方明月 李 涛 聂辉华

内容提要:社会各界关于企业数字化转型的重要性已经基本达成共识,但对企业数字化转型的效果存在严重分歧。产生这一现象的主要原因是现有研究对企业数字化转型的测度存在问题:一是测度对象不够统一明确,二是测度方法不够科学准确。这导致很多研究结论不可比较、难以复制和相互冲突。为了更好地处理上述难题,本文运用机器学习和大语言模型构造一套新的企业数字化转型指标。本文首先对2006—2020年上市公司年报中的句子进行人工打标签,然后用标记结果训练和微调包括大语言模型在内的多个机器学习模型,选择其中分类效果最好的ERNIE模型作为句子分类模型来预测全部文本中句子的标签,最终构造了企业数字化转型指标。理论分析和数据交叉验证均表明,本文构建的指标相对已有方法更准确。在此基础上,本文实证检验企业数字化转型对财务绩效的影响。研究发现:第一,企业数字化转型能够显著提高财务绩效,其中,大数据、人工智能、移动互联、云计算和物联网均有明显作用,但区块链并没有明显的作用;第二,只有在财务绩效较差的企业中,数字化转型才能够显著提高财务绩效;第三,企业数字化转型提高财务绩效的主要渠道包括改善效率和降低成本。本文研究对于推动企业数字化转型和实现经济高质量发展具有一定现实意义。

关键词:企业数字化转型 数字经济 数字技术 人工智能 大语言模型

一、引言

人类开始从工业经济时代迈入数字经济时代。在此背景下,党的二十大报告指出,“加快发展数字经济,促进数字经济和实体经济深度融合”。对于企业来说,数字经济和实体经济的融合过程本质上是一个数字化转型过程。所谓数字化转型(digital transformation),是企业借助数字技术来改造企业的生产经营系统、管理模式和核心业务流程,从而形成一个破坏性创新和变革的过程(Siebel, 2019)。伴随数字经济的不断发展,数字化转型已经成为全球企业普遍关心的热点问题。2020年春季,BCG(波士顿咨询集团)的研究报告显示,超过80%的企业正在进行数字化转型。^①对于中国企业来说,面对数字经济规模的快速增长,数字化转型已经成为它们面临的最大挑战之一。2021年,中国数字经济的规模为7.1万亿美元,位居世界第二,仅次于美国,并且数字经济与GDP的比值高达39.8%(中国信息通信研究院,2022a)。^②

实践是理论的基础。中国在数字经济领域的后发优势,催生了中国学者如火如荼的企业数字

^{*} 金星晔、左从江、李涛,中央财经大学经济学院;方明月(通讯作者),中国农业大学经济管理学院,邮政编码:100083,电子邮箱:fmingyue@163.com;聂辉华,中国人民大学经济学院。本研究是国家社科基金重大项目(22&ZD070)、国家自然科学基金项目(72002213,72273144)、中央财经大学青年科研创新团队支持计划和科教融合研究生学术新星孵化计划的阶段性成果。作者感谢匿名审稿人的有益建议,文责自负。

^① 参见BCG:“The Evolving State of Digital Transformation”,2020年9月25日。

^② 联合国对2019年数字经济的狭义和广义测算均表明,美国和中国分别位居全球第一名和第二名(Chen,2020)。蔡跃洲和牛新星(2021)提供了关于中国数字经济规模测算的详细分析。

化转型研究。中国知网收录的涉及企业数字化转型的论文数量从2018年的110篇快速增加到2022年的961篇。同期,EconLit数据库收录的相关英文论文数量也从2018年的48篇快速增加到141篇,并且其中有华人作者的文章占比从2018年的2%增加到2022年的25%。

尽管企业进行数字化转型已经成为业界和学界的共识,但是针对企业数字化转型的成败却形成了截然相反的两种观点。著名管理咨询公司麦肯锡基于1793个企业经理人的调查认为,超过80%的企业数字化转型是失败的。^①这与学术研究的结果大相径庭,因为基于中外企业的多数经验研究都表明,企业数字化转型显著提高了财务绩效(赵宸宇等,2021;Commander et al., 2011; Müller et al., 2018)。当然,在对总体效果有一定共识的同时,学术界内部对数字化转型的具体影响也存在不同判断。例如,DeStefano et al.(2018)利用1999—2005年英国企业的调查数据,发现数字化转型虽然扩大了企业规模,但并未提高全要素生产率。此外,刘淑春等(2021)基于1950家中国企业连续5年的追踪调查数据,发现企业数字化转型和效率之间并不存在线性关系,而是呈现倒U型关系。业界和学界在企业数字化转型成败方面出现严重分歧,本文认为导致这一现象的主要原因是现有研究在企业数字化转型的测度上出现了问题。这体现在两个方面:一是测度对象不统一、不明确,不同的研究对企业数字化转型的定义和测度不同,这使得不同研究的结果不可比较甚至难以复刻;二是测度方法不科学、不准确。现有文献在测度数字化转型时,大多采用了基于上市公司年报的词典法来测度企业数字化转型水平(例如,吴非等,2021;方明月等,2022)。但是,词典法存在两个明显的问题:一是词典包含的数字技术关键词不够完备,导致了一部分真实的数字化转型未被统计(第一类错误,企业应用了数字技术但未能被词典法等方法识别);二是表意不真实,错误地将一部分文本内容判定为数字化转型(第二类错误,文本中提及关键词但实际未使用数字技术)。此外,不同的数字技术可能导致了不同的财务绩效。例如,人工智能可能提高了企业的财务回报,但区块链可能只是提高了企业的生产成本,此时就不能笼统地判断数字化转型是成功还是失败。因此,如果理论要反过来指导实践,学术界需在企业数字化转型的测度上达成共识,并着力缓解测度方法不准确的问题。这样才能廓清迷雾、减少分歧,进而为当前中国企业在数字化转型过程中普遍存在的“不愿转、不敢转”的难题提供理论启迪。

在充分扬弃现有文献利弊的基础上,本文利用前沿的机器学习方法和大语言模型(large language model, LLM),^②基于2006—2020年中国上市公司年报文本,立足全面体现各种数字技术在企业中的实际使用状况,构造了4181家上市公司的一套数字化转型指标。具体来说,数字化转型的测度分五步进行:第一步,整理爬虫抓取和手动收集的上市公司年报,并将年报中的“管理层讨论与分析”和“目录、释义及重大风险提示”这两部分内容作为企业数字化转型的相关文本。第二步,将相关文本按照句号和分号全部分割为句子,构成待预测句库。第三步,对相关文本同时进行随机抽取以及抽取包含关键词的文本,形成待标记句库,对待标记句库进行人工标注,并以此判断企业是否进行了数字化转型。第四步,基于大语言模型ERNIE^③等,采取有监督的机器学习方法进行句子分类模型的训练。第五步,使用训练后的ERNIE模型对待预测句库进行逐句预测,判断上市公司是否使用以及使用哪种数字技术,并最终构建一套新的企业数字化转型指标。为了验证新指标的有效性,我们先后将其与专利数据、地区数据以及国际文献等进行六个方面的对比,均发现本文构建的数字化转型指标与现实高度吻合。与词典法相比,本文构造的指标在内容上更加完备,在表意

① “Unlocking Success in Digital Transformations”, October 2018, <http://www.mckinsey.com>.

② 大语言模型是指基于海量文本训练,并且包含极多参数的语言模型。作为一种人工智能技术,它使用深度学习算法来处理自然语言的模型。它可以根据大量数据来识别、总结、翻译、预测和生成文本及其他内容。2022年底开始风靡全球的ChatGPT就是一种大语言模型的具体应用。

③ ERNIE模型是百度公司开发的一种大语言模型,全称是Enhanced Representation through Knowledge Integration。

上更加真实。

在使用新方法构造了企业数字化转型指标的基础上,本文实证检验了企业数字化转型与企业财务绩效的关系,并得到了三点新发现:第一,总体而言,企业数字化转型能够显著提高其财务绩效(ROA和ROE),但并非所有的数字技术都有这种显著的积极作用,大数据、人工智能、移动互联、云计算和物联网均能显著提高ROA和ROE,但区块链未能显著提高ROA和ROE。第二,不同财务绩效的企业进行数字化转型的效果不同。对于财务绩效较差的企业,数字化转型能够显著提高ROA和ROE;而对于财务绩效较好尤其是非常好的企业,数字化转型对ROA和ROE的作用效果并不显著。第三,企业数字化转型提高财务绩效的主要渠道有两个,分别为改善效率和降低成本,而提高收入的渠道没有被证实。

本文的贡献主要体现在以下三个方面:第一,提供了一种测度企业数字化转型的新方法,从而为数字化转型文献奠定了坚实的实证基础。伴随数字经济在全球的兴起,“数字经济学”(digital economics)已经成为一个新兴研究领域,而企业的数字化转型属于该领域的重要内容(Goldfarb & Tucker, 2019)。近几年,国内外学者从数字技术的投入经费(刘飞和田高良, 2019; 祁怀锦等, 2020)、数字设备的使用(Bloom et al., 2014; Acemoglu & Restrepo, 2020; Brynjolfsson et al., 2021)和数字技术关键词提取(杨德明和刘泳文, 2018; 吴非等, 2021; 袁淳等, 2021; 张叶青等, 2021; 赵宸宇等, 2021; 方明月等, 2022)等多个角度研究了数字化转型对企业行为和绩效的影响。但是,由于测度对象不够统一明确,测度方法不够科学准确,导致不同文献的结果不可比较,部分文献的计量结果难以复制,并存在结论相互冲突等问题。本文基于机器学习和大语言模型提供的一套数字化转型测度指标,具有对象明确、指标齐全、准确度高和可复制的优点,为缓解上述问题提供了一种可行方案。因此,本文有助于从研究方法上推进企业数字化转型的深入研究,并且为广义的数字经济学文献提供来自中国的经验证据。考虑到中国既是发展中国家,又是全球数字经济大国,我们相信基于上市公司文本并利用机器学习方法来构造企业数字化转型指标的做法,对广大发展中国家和发达国家都具有参考价值。

第二,揭示了不同的数字技术对企业财务绩效的不同影响,并甄别了不同的影响渠道,从而为现有的数字化转型文献补充了新发现。已有文献在分析数字化转型对企业绩效的影响时,通常笼统地将所有数字技术的采用都视为企业数字化转型(何帆和刘红霞, 2019; 赵宸宇等, 2021),或者只是考虑了某种数字技术(Bloom et al., 2014),但实际上不同的数字技术可能导致不同的转型效果,而且在实践中企业可能只是采用了某些技术。本文根据政府统计口径和权威机构的定义,将数字技术分为大数据、人工智能、移动互联、云计算、物联网和区块链六种类型,并发现区块链对企业财务绩效没有显著影响,而其他五种数字技术都显著提高了企业财务绩效。此外,在影响渠道上,本文证实了效率渠道和成本渠道的成立,但收入渠道不成立。这与现有文献的结果相互验证。通过上述更细致的技术分类和渠道检验,本文丰富了企业数字化转型文献。

第三,丰富了大语言模型在经济学文献中的应用。伴随人工智能和机器学习技术的快速扩散,一些学者开始利用大语言模型进行微观经济学研究。许雪晨和田侃(2021)使用BERT模型^①进行财经新闻文本的情感分析,进而预测情感与股票市场涨跌的关系。刘青和肖柏高(2023)利用BERT模型构建了一个专利文本分类模型,并使用模型识别了专利数据库中 with 劳动节约型技术有关的专利。Acikalin et al.(2022)用已标记的23734条专利文本对Longformer模型(一种改进的BERT模型)进行训练,预测企业的专利受到美国某个判例影响的可能性。Rajan et al.(2023)使用美国上市公司1955—2020年9000封致股东的信来预训练BERT模型,并人工标注了约2%的段落来训练模型,使

^① BERT模型是谷歌公司开发的一种大语言模型,全称是Bidirectional Encoder Representations from Transformers。

得模型能够判断公司的目标,继而研究了企业发布这些目标的原因以及时机。与上述文献主要基于BERT模型进行训练不同,本文考虑到英文和中文语境的差异,使用了更适合中文语境的ERNIE模型(Sun et al., 2019)。因此,本文的研究为中文大语言模型在经济学研究中的应用开辟了一条新路。

本文余下部分安排如下:第二节对企业数字化转型指标进行一个完整的评估;第三节介绍本文基于机器学习和大语言模型构建企业数字化转型指标的方法,并从多个角度验证新指标的有效性;第四节通过分析企业数字化转型绩效揭示一些新的发现;最后是结论和政策含义。

二、现有文献测度数字化转型时存在的不足

(一)企业数字化转型的三种测度方法

现有文献在测度企业数字化转型程度时,通常使用了三种方法:第一种方法是客观数据法,这包括计算本企业与数字技术相关的软件投资或硬件投资占总资产的比重(刘飞和田高良,2019;祁怀锦等,2020;Müller et al., 2018),基于调查数据度量企业内机器人的使用(Acemoglu & Restrepo, 2020)或者分析预测工具的使用(Brynjolfsson et al., 2021),以及基于行业计算机软硬件投资额度度量行业的信息技术密度(IT intensity)(Chun et al., 2008)。

第二种方法是事件冲击法,即利用企业所属群体是否受到数字化转型政策的冲击来度量企业的数字化转型情况。常用的做法是,利用国务院确定的“宽带中国”政策(李万利等,2022)、工信部批准的“两化”(信息化和工业化)融合政策(李磊等,2022)、国家信息消费示范城市(方明月等,2022)作为政策冲击。这些政策冲击也常用于解决企业数字化转型影响企业行为和绩效时存在的内生性问题。

第三种方法是最主流的方法——词典法,即先构建一个包含各种数字技术的关键词词典,然后根据这些关键词在上市公司年报中“管理层讨论与分析”部分出现的次数或比例,^①构建企业数字化转型指标。该方法背后隐含的假设是,提及了某种数字技术的关键词就表明企业进行了数字化转型。因此,某个上市公司年报中提及数字技术的次数或者比例越高,表示企业的数字化转型程度越高。使用这一方法的文献较多,包括杨德明和刘泳文(2018)、吴非等(2021)、袁淳等(2021)、赵宸宇等(2021)、张叶青等(2021)和方明月等(2022)等。

公允地说,以上三种方法为中国学者开辟企业数字化转型这一新时代的重要研究提供了一个有效的切入点。然而,时至今日,当企业数字化转型的文献已经相当丰富时,学术界应该客观地认识到上述方法的不足,并努力寻求更加完善的测度方法,才能深入推进企业数字化转型的研究。客观数据法有两个缺点:第一,范围较窄,只适合度量某一种具体的数字技术的非人力成本投入。例如,某个企业招聘了从事数字化转型的工程师,此时支付给工程师的工资成本也应该算作企业数字化转型的投入,但是它并不能体现为数字化硬件或软件的投入。第二,度量比较粗糙。例如,笼统地统计数字技术硬件或软件的做法,无法区分不同类型数字技术(例如大数据和人工智能)的应用。

事件冲击法也有两点不足:第一,它假设试点地区所有企业都受到某项数字技术政策的同等程度影响,这显然不符合现实。事实上,即便是在试点地区,也不是所有企业都会受到该政策的影响。例如,金环等(2021)发现,“宽带中国”政策对本地区的民营企业 and 成长型企业的创新和全要素生产率(TFP)影响较大,但对国有企业和衰退期企业并没有显著影响。第二,试点地区的企业很可能会受到本地区其他政策的影响。尽管平行趋势检验可以排除其他政策在样本区间的干扰,但无法排除与试点政策同期实施的其他政策。此外,多种相关政策在短期内同时推行,并且存在政策时滞,

^① 这里的比例是这些关键词(或包含关键词的句子)占年报中“管理层讨论与分析”章节的全部词语(或句子)的比重。一些数据库(例如CSMAR)或文献(吴非等,2021)在计算企业数字化转型指标时,只统计了数字技术关键词的次数,而没有计算比例。

这导致很难区分不同政策的实际效果。

(二)词典法测度数字化转型存在的两类问题

鉴于词典法是国内多数文献使用的数字化转型测度方法,我们重点分析词典法的不足。很多研究企业数字化转型的文献直接使用CSMAR数据库自带的数字技术词频统计表作为企业数字化转型程度的指标(例如黄逵友等,2023;耀友福和周兰,2023)。因此,下面以CSMAR关键词词典作为主要分析对象。CSMAR词典总共包含了62个数字技术关键词,例如“机器学习”“数字货币”“物联网”“数据挖掘”等,并且分为四类技术:人工智能技术(包含27个关键词)、区块链(包含8个关键词)、云计算(包含17个关键词)和大数据(包含10个关键词)。

以CSMAR为代表的词典法存在的第一类问题是,词典构建不够完备,即它可能遗漏了不少属于数字技术但是未被纳入词典的关键词。我们在年报文本中找到了一些句子,这些句子能够表示企业使用了数字技术,但相关的技术词语并没有被收录到词典中。例如:(1)“二是聚焦在出行、医疗、教育等用户自然生活需求的场景构建,通过‘云+API’应用程序编程接口的方式输出金融服务能力,提高客户黏度与产品渗透率”;^①(2)“利用光学字符识别技术(OCR)赋能证件识别,识别成功率达98%以上,提升了业务审核效率”;^②(3)“研发完成通过较小代价提升方言与小语种识别可用性的技术路径,方言与维语、藏语识别效果大幅提升”;^③(4)“配合先进的图像识别算法可保证图像识别正确率>95%”。^④这些句子中的“云+API”“OCR”“小语种识别”“图像识别”都能够表示企业使用了相关的数字技术,但这些关键的技术词语均未被收录到CSMAR的词典。

之所以会出现关键词遗漏,是因为这些关键词都是研究者根据部分文献人为选定的词语,而每个人选择的标准又很难统一。在实践中,各种数字技术层出不穷,新的名词不断涌现,因此用词典法来测度数字化转型难免会存在“挂一漏万”和更新迟滞的问题。事实上,除了CSMAR的关键词词典,其他文献使用的关键词词典也存在一定的遗漏问题。

人为选择关键词导致的附带问题是,每个研究者的主观标准不同,选择范围不同,这导致不同文献使用的关键词差异很大,使得基于不同词典构造的数字化转型指标缺乏可比性。以公开了关键词词典的几篇代表性文章为例。吴非等(2021)的词典包含了76个关键词,李云鹤等(2022)包含了95个关键词,杨金玉等(2022)包含了76个关键词,方明月等(2022)包含了112个关键词。从词语范围来看,方明月等(2022)包含的关键词最多,CSMAR最少,这导致这些词典之间的重合度不高。例如,CSMAR词典与吴非等(2021)、李云鹤等(2022)均有39个关键词相同,相同的关键词占CSMAR词语总数(62个)的比重(即重合度)为63%。不过,吴非等(2021)和李云鹤等(2022)的关键词又不完全相同,后者比前者多了19个关键词。此外,词语个数最多的方明月等(2022)与词语个数次多的李云鹤等(2022)有19个关键词相同,重合度仅为17%。

第二类问题是,词典法可能会误将一些并不表示企业数字化转型实践的内容包括在内。仍以CSMAR为例,在一些上市公司的年报中,即便某个句子中包含了数字技术关键词,根据句意也不能判断该企业进行了数字化转型。具体来说,这包括三种情况:第一,句子采用了否定表述;第二,公司可能表示将在未来进行数字化转型而不是现在;第三,企业可能描述的是行业的发展背景而不是自身的数字化转型行动。这些情况都会导致词典法出现误判。我们对应地找到了如下句子来分别说明这三种情况:(1)“为了避免项目存在的不确定性与技术研发风险,公司暂时将智能教育机器人

① 上市公司(股票代码为600036)2018年年报。

② 上市公司(股票代码为000001)2020年年报。

③ 上市公司(股票代码为002230)2017年年报。

④ 上市公司(股票代码为002767)2017年年报。

研发中心项目推进节奏放缓,因此尚未对该项目有较大的投入”;^①(2)“未来公司将利用物联网行业高速发展的有利环境,努力扩大业务规模并提高公司盈利能力”;^②(3)“2021年,公司将稳步拓展大屏与专业显示器业务,进一步丰富产品品类,充分利用5G基站建设、特高压、城际轨道交通、新能源汽车充电桩、大数据中心、人工智能、工业互联网七大领域发展带来的新机遇”。^③如果说词典包括的关键词不够完备属于“苛责”,那么更严重的问题是,机械地根据文本内容是否包含关键词来判别企业是否进行数字化转型。前者属于第一类错误,即企业应用了数字技术但未能被词典法等方法识别;后者属于第二类错误,即文本中提及关键词但实际未使用数字技术。本文接下来提供的新测度方法,就试图在以上两个方面提出改进。

三、基于大语言模型的新测度方法

(一)大语言模型 ERNIE

使用词典法进行文本分析虽然操作比较简便,但是这种传统方法不能够充分提取文本中含有的信息,导致分析的准确度、指标的有效性较低。近年来,在人工智能和机器学习领域,自然语言处理(natural language processing, NLP)技术被广泛使用。它是利用计算机技术来分析、理解和处理自然语言的一门交叉学科,主要应用于机器翻译、舆情监测、自动摘要、观点提取、文本分类、问题回答、文本语义对比、语音识别和光学字符识别等。一个包含机器学习任务的自然语言处理(NLP)过程通常可以分为如下步骤:数据预处理、文本表征、目标任务模型训练。自然语言处理技术的进步又主要受益于预训练(pre-train)技术的发展。所谓预训练,即首先在一个原任务上训练一个初始模型,然后在下游任务(也称目标任务)上继续对该模型进行精调,从而达到提高下游任务准确率的目的(车万翔等,2021)。预训练技术主要在自然语言处理的文本表征和目标任务模型训练这两个阶段发挥作用,目的在于使文本能够更好地被表征从而提升下游模型效果。

预训练技术的发展可以分为三个阶段:早期的静态预训练技术、经典的动态预训练技术和近期的新式预训练技术(李舟军等,2020)。静态与动态的差别在于词语的表征是否随上下文变动。动态预训练技术主要包括GPT和BERT两类大语言模型,它开创了基于上下文的文本表征方法,解决了一词多义问题。然而,在中文表达中,知识大多以词为单元出现,BERT模型难以学会语义知识单元的完整语义表示。基于BERT模型改进的新式预训练技术ERNIE模型通过遮盖词语(知识单元)引入知识,进一步提升了模型的语义表示能力。此外,在训练数据方面,BERT仅使用百科类语料训练模型,而ERNIE使用百科类、新闻资讯、论坛对话类语料训练,且实验证明ERNIE预训练模型在五项中文文本分类任务上体现出全面超越BERT的性能(Sun et al., 2019)。^④基于以上原因,本文选择主要使用大语言模型ERNIE来完成文本分类任务。

(二)使用ERNIE模型测度企业数字化转型

1. 第一步:确定文本分析对象

由于数字化转型涉及企业组织结构、内部管理、业务流程等方方面面的变革,难以在财务指标中完整显示,但上市公司有强烈的意愿在年报中披露以获得资本市场的青睐,因此文献中通常基于年报进行文本分析来衡量数字化转型水平(方明月等,2022)。借鉴已有文献的做法,本文同样采用上市公司年报作为企业数字化转型指标的文本基础。

① 上市公司(股票代码为300010)2019年年报。

② 上市公司(股票代码为603236)2019年年报。

③ 上市公司(股票代码为000727)2020年年报。

④ 例如,BERT可以推测出“北【x】是中国的首都”,其中“【x】”为“京”,但不能对知识单元“北京”进行学习,而ERNIE 1.0可以预测“【xx】是中国的首都”中“【xx】”为“北京”,从而学习到北京与中国首都之间的关系。

我们通过爬虫和人工收集这两个途径,收集了上市公司年报,信息来源包括 Wind、巨潮资讯网和上市公司官网。2007 年 1 月 1 日起施行的新《企业会计准则》对企业财务指标要求有较大变化,而 2006 年年报实际披露于 2007 年一季度,因此我们选择 2006—2020 年间披露的上市公司年报进行分析。在年报中,“管理层讨论与分析”(MD&A)分析了企业在报告期内的经营情况、描述未来的发展战略并披露公司所面临的风险状况。因此,已有文献几乎都选择这个部分作为计算数字技术关键词的次数或比例的来源(袁淳等,2021;赵宸宇等,2021)。还有部分公司选择在“目录、释义及重大风险提示”中披露公司可能面临的风险,其中也可能包含企业数字化转型相关信息。因此本文选择“管理层讨论与分析”和“目录、释义及重大风险提示”这两个章节作为文本数据,最终我们得到 2006—2020 年 4181 家公司的 39175 份年报文本。

2. 第二步:构建待预测句库和待标记句库

本文先将全部文本按照句号和分号分割,得到待预测句库。由于年报中大多数句子与数字化转型无关,如果完全随机抽取句子进行阅读,得到的大多数标签都将与数字技术无关,为了提高人工阅读的效率并防止上下文对人工阅读产生干扰,需要使用关键词抽取具有不同代表性的年报句子,并与随机抽取的句子一起构成待标记句库。^①为此,我们先定义数字技术,并构建数字技术词典。

在定义数字技术时,我们首先考虑政策口径。国家统计局在《数字经济及其核心产业统计分类(2021)》中提到,产业数字化代表性技术为物联网、人工智能、大数据、云计算、移动互联等数字技术。国务院和工信部等部门多次出台政策文件,提出了促进大数据、人工智能、云计算、物联网、移动互联和区块链等数字技术发展的指导意见。其次是企业界的定义。中国数字技术的领军企业腾讯计算机系统有限公司董事会主席马化腾在 *The Chinese Digital Economy* 一书中指出:近年来,移动互联、云计算、大数据、人工智能、物联网、区块链等数字技术不断突破和融合发展,推动了数字经济快速发展(Ma et al., 2021)。综合以上定义,本文将数字技术分为六种类型:大数据、人工智能、移动互联、云计算、物联网和区块链。^②我们基于政策文本、研究报告和已有文献,并通过人工阅读之后的不断补充,收集整理了一个包含了 311 个数字技术关键词的词典。^③然后,我们抽取了包含 10 个及以上不同关键词的年报文本,^④并取出其中含有关键词的句子。同时,为了提高模型对不含关键词句子的预测能力,我们又随机抽取了部分年报,并分割为句子。由于上市公司总数逐年增加,如果直接在上述两部分句子中进行随机标注,将导致大部分被标注的句子靠近当前年份。为了解决年份分布不均匀的问题,在这两部分句子的基础上,我们按照年份分组,在每个年份中取出相同数量的句子,再从这部分均匀分布的句子中进行不放回的随机抽取,得到本研究的待标记句库。最终,本研究的待标记句库中包含 38994 个句子。

3. 第三步:对待标记句库的句子进行人工标注

人工标注的思路是,先判断企业使用了哪种/哪几种数字技术,进而判断企业是否进行了数字化转型。人工标注的目的是形成训练集、测试集和验证集,为后面的机器学习打下基础。

我们将 24 位研究人员分为 12 组,每组两人定期轮换。为了统一对句子打标签的标准,在正式

① 通过构建关键词词典缩小随机阅读范围的做法在类似研究中也经常使用。例如,Chen et al.(2019)使用有监督的机器学习方法对金融科技相关专利文本进行分类。在人工阅读之前,他们基于自行构造的金融词典,在原文本库的基础上筛选出符合要求的文本,然后再抽取文本进行标注。

② 因篇幅所限,具体的数字技术的定义以及例子请见本刊网站登载的附表 1。

③ 因篇幅所限,数字技术关键词词典的详细列表请见本刊网站登载的附表 2。

④ 10 个及以上这个阈值是一个经验数值,更低的阈值会包含更多的年报,但可能会导致人工阅读的效率降低。另外,我们不直接从待预测句库中找含有关键词的句子,这是因为直接抽取句子会导致句子的来源难以确定。而先抽取年报再取句子不仅可以满足随机性,还可以保证句子的来源可回溯,从而方便打标签以及排查错误。

打标签之前,我们多次讲解打标签任务的详情,并对容易混淆的标签进行了重点讲解和示范。明晰标准后,我们进行了充分的打标签训练,并对标注过程中发现的难点和疑点定期商讨。在正式标注时,待标记句库中的每个句子都会被两位研究成员标记。如果双方标注结果一致,则句子标签被记录;对存在分歧的句子,经过全部成员讨论后确定其标签;对难以确定标签的句子,不纳入训练集。最后,所有待标记句库中除了难以确定标签的句子外,剩余句子都被分类至八个标签下,包括六种新型数字技术、非新型数字技术、^①非数字技术。

4. 第四步:采用有监督的机器学习方法训练模型

度量数字化转型的关键步骤是训练机器学习模型,让人工智能技术替代人工来判别文本中包括数字技术关键词是否意味着企业真正进行了数字化转型,从而缓解在文本中提及数字技术的关键词,但实际上并未使用数字技术的难题。这需要使用分类模型。我们使用百度开源的、内嵌了ERNIE的PaddleHub框架来开展模型训练,用其内置的tokenizer函数快速将句子转为ERNIE模型训练所需要的格式。我们将全部被标记的句子按照8:1:1的比例分为训练集、测试集和验证集。同时,为了比较不同模型之间的分类性能,我们也基于PaddleHub框架训练了BERT_base_Chinese,同时基于sklearn框架训练了SVM(支持向量机)、Neural Networks(神经网络)、SVM与Neural Networks结合的Voting算法^②、KNN(K近邻)以及GaussianNB(朴素贝叶斯)共七类常见模型。

机器学习的主要目的是识别文本是否以及体现了哪种数字技术。对于这种分类模型,通常用精确度(Precision)、召回率(Recall)和正确率(Accuracy)来度量模型的性能。考虑到训练集各类型标签数量分布不均,通常使用F-Score来衡量模型的整体分类能力。^③在四类指标中,Precision是指模型预测为Positive类的样本中,实际为Positive类的比例。一个高Precision的模型意味着它能够准确地识别出真正为Positive类的样本,减少了误报的可能性。在本文的使用场景下,Precision度量的是模型认为的属于数字技术的句子中,其真实标签也是数字技术的句子的比例。Recall度量的是所有实际为Positive类的样本中,被预测为Positive类的比例。一个高Recall的模型意味着它能够尽可能地找出所有实际为Positive类的样本,减少了漏报的可能性。在本文的使用场景下,年报中可能会有多处句子表示企业使用了数字技术,因此正确分类的能力在本文的场景下相对更加重要。据此,本文也计算了F.8-Score,以赋予Precision更大的权重。比较不同模型在相同训练集上的表现可以发现,ERNIE模型的精确度、召回率、正确率、F1-Score、F.8-Score分别达到了81%、70%、93%、75%和76%。ERNIE仅在Recall上落后于BERT(进而导致F1-Score上落后),但考虑Precision的重要性后,ERNIE的F.8-Score是所训练模型中得分最高的,据此我们选择ERNIE作为句子分类模型,图1的Panel A对不同模型的性能进行了比较,可以发现ERNIE和BERT的综合分类能力(F1、F.8)要远高于常见的分类模型。^④

① 非新型数字技术指传统的数字技术或者数字技术的泛称,例如,互联网、平台经济、数字化、数字技术、智能化等提法。

② SVM与NN的分类性能在所选择的传统算法中表现最好,因此选择这两类算法的结合来构造Voting模型。

③ Precision度量全部被预测为Positive类(“是”)的句子中,真的是Positive类的比例;Recall是模型能够将一份年报中全部Positive类句子找出来的能力;Accuracy是度量模型的分类在多大程度上是正确的能力,包括了Positive类和Negative类(“否”),F1是Precision和Recall的调和平均数。假设一份文本中有100个句子,其中10句经人工判断被认为与数字技术相关,模型预测全文有12句与数字技术相关,比对后发现模型预测出的12句中有8句与人工判断的结果一致,则Precision值为 $8/12 = 0.75$,Recall值即为 $8/10 = 0.8$,两个指标的区别主要在于分母,此外该结果说明模型对90句Negative类的判断有4句错误,对10句Positive类的判断也有2句的错误,则Accuracy即为 $[(10-2)+(90-4)]/100=0.94$ 。 $F-Score = (1+\beta^2) * Precision * Recall / (\beta^2 * Precision + Recall)$ 。 $F1-Score = (1+1^2) * Precision * Recall / (1^2 * Precision + Recall) = 0.774$ 。F.8-Score即 β 取值为0.8。

④ 限于篇幅,不同的模型在各指标上的具体表现请见本刊网站登载的附表3。

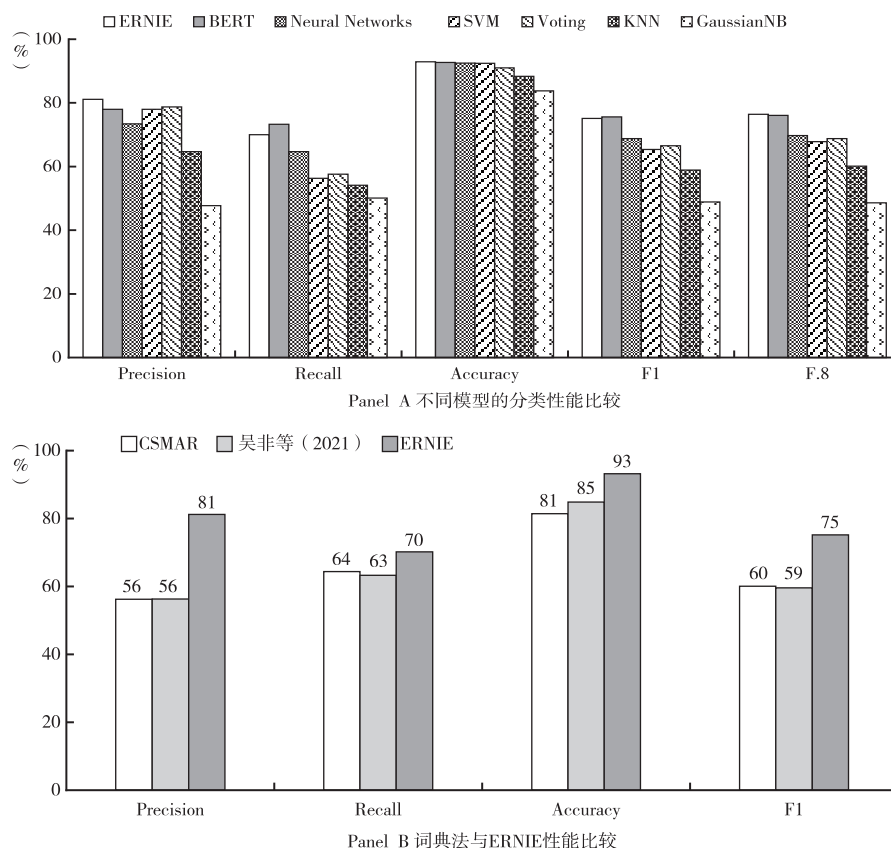


图1 各方法的分类性能比较

5. 第五步:采用ERNIE模型构造数字化转型指标

基于ERNIE大语言模型,我们对2006—2020年待预测句库中的每一句文本进行预测,判断企业是否以及使用何种数字技术。我们构造了企业数字化转型哑变量,即公司在当年只要使用大数据、人工智能、移动互联、云计算、区块链和物联网中的任意技术,则指标赋值为1,反之则为0。^①

(三)企业数字化转型指标有效性检验

虽然我们从技术上论证了使用大语言模型ERNIE构建企业数字化转型指标的合理性,但是用新方法构造的指标是否更准确、更符合现实呢?下面我们从分类性能、专利数据、时间趋势、地区差异、行业差异和国际比较六个角度,对新方法的有效性进行全面检验。

1. 分类性能

我们首先与现有文献广泛使用的词典法比较。我们分别基于CSMAR数字技术关键词(62个)和吴非等(2021)的关键词(76个)对人工阅读所用的测试集的句子进行分类。如果句子中出现了关键词,则认为句子能够表示企业使用了对应的数字技术。得到分类结果后,计算精确度(Precision)、召回率(Recall)、正确率(Accuracy)和F1-Score四个指标。图1中的Panel B表明,与词典法相比,ERNIE模型的各项指标都表现更好,而CSMAR和吴非等(2021)的分类性能差别不明显。这证明,采用机器学习方法可以更准确地判定文本是否真实体现企业数字化转型,即提高了表意的真实性。具体而言,本文的方法相对于词典法的差别是:在第二类错误上,即提及但

^① 我们没有根据数字技术相关句子的数量或比例来构造企业数字化转型的连续变量。因为一个公司在年报中多次提及数字技术关键词,在表意为真的前提下只能说明其正在进行数字化转型,而难以测度其转型程度。

实际未使用的问题方面,有近25个百分点的改进程度(Precision的变化);在第一类错误上,即句子意思能够表示使用了数字技术,但不能被模型或者关键词捕捉的问题,也有6-7个百分点的改进程度(Recall的变化)。这说明本文的方法相对于词典法在两类错误的处理上都有更好的表现。

2. 专利数据

本文将企业使用了某种数字技术作为企业数字化转型的主要判断依据。而判断企业是否使用了某种数字技术,最可靠的指标是企业的数字技术专利申请。从理论上讲,如果一个企业申请了某种数字技术的专利,那么它应该使用了这种数字技术,但反之则未必成立。因此,我们可以将词典法、ERNIE模型判断的使用了数字技术的公司,与那些确实申请了数字技术专利的上市公司进行比较。如果某种方法的判断结果与专利申请结果最接近,那么某种方法的准确度就最高。

本文先将专利数据库和上市公司数据库进行匹配。具体来说,第一步,使用大为InnoJoy专利检索平台确定上市公司专利申请记录。第二步,根据世界知识产权组织(WIPO)发布的《2022年世界知识产权报告:创新方向》,确定大数据、人工智能和云计算这三种数字技术的专利分类号,以筛选上市公司申请的三类技术专利。^①第三步,获得了吴非等(2021)、CSMAR基于词典法和本文基于ERNIE模型判断的使用了三种数字技术的公司名单,并比较了这三者与申请了这三类数字技术专利的上市公司的重合度。^②统计结果显示,ERNIE模型识别的使用数字技术的公司与实际申请数字技术专利的公司重合度最高。这表明,从专利数据来判断,本文使用的ERNIE模型对企业数字化转型的判断准确度最高。

3. 时间趋势

本文将当年上市公司是否使用某种数字技术的哑变量加总后求均值,算出某种数字技术以及任意一种数字技术在不同年份的普及度(图2)。例如,在2020年A股上市公司中,有42%的企业使用了大数据技术。从时间趋势上看,各类数字技术的普及度都明显地随时间增加,特别是在2011—2017年间。这符合直觉,也与国内外的技术扩散相吻合。以人工智能为例。在图2中,人工智能普及度在2011年以后达到了18%以上,并且在2012—2018年间快速增长。在现实中,2013年,Facebook成立了人工智能实验室,Google收购了语音和图像识别公司DN-NResearch,百度创立了深度学习研究院。2015年Google开源了机器学习平台TensorFlow,使得2015年成为人工智能技术突破之年。此外,Google人工智能AlphaGo分别于2016年、2017年击败了李世石、柯洁等围棋世界冠军,使得人工智能技术获得了大量关注。此外,从不同技术的发展趋势来看,总体上物联网和人工智能的普及度最高(60%左右),其次是大数据和移动互联(40%左右),第三是云计算(20%左右),区块链的普及度最低(7%)。本文计算了贵州省大数据技术的普及度,发现贵州省大数据技术的普及度与有关政策节点高度相关。^③此外,本文还计算了移动互联技术在2007—2020年间的普及度增速。本文发现工信部在发放3G、4G和5G牌照时,移动互联技术的普及度增速都是相近年份的峰值。^④这些证据都说明了本文所构建指标的可靠性。

① 本文之所以只选择这三种数字技术,是因为WIPO虽然提供了人工智能、大数据、云计算、物联网等技术的分类号,但与吴非等(2021)、CSMAR重合的技术类型只有人工智能、大数据和云计算。

② 限于篇幅,重合度的比较结果请见本刊网站登载的附图1。

③ 限于篇幅,贵州省大数据技术普及度趋势图请见本刊网站登载的附图3。

④ 限于篇幅,移动互联技术的普及度增速请见本刊网站登载的附图4。

年份	大数据	人工智能	移动互联	云计算	物联网	区块链	任意一项
2006	2%	4%	3%	1%	12%	0%	15%
2007	3%	6%	4%	1%	16%	1%	20%
2008	3%	7%	5%	1%	19%	0%	23%
2009	6%	9%	6%	2%	22%	1%	28%
2010	7%	12%	7%	4%	25%	1%	31%
2011	10%	16%	10%	6%	29%	1%	37%
2012	11%	18%	13%	8%	33%	2%	42%
2013	16%	21%	19%	10%	38%	2%	50%
2014	22%	29%	25%	12%	42%	3%	57%
2015	32%	39%	32%	17%	49%	4%	65%
2016	32%	41%	32%	18%	48%	4%	66%
2017	35%	48%	34%	18%	54%	5%	70%
2018	39%	53%	36%	20%	59%	7%	73%
2019	40%	54%	39%	21%	59%	7%	75%
2020	42%	58%	41%	22%	62%	7%	78%
均值	25%	34%	25%	13%	43%	4%	

图2 数字技术普及度

注：图中展示了不同数字技术在2006—2020年间普及度的变化。图中“任意一项”指使用了任意一项数字技术的上市公司占当年全部上市公司的比重。由于不同数字技术普及度差距较大，图中条形图的长度基于每类数字技术的最大普及度标准化而来，以便于读者看出不同数字技术普及度的变化趋势，因此不同数字技术的条形图不可相互比较。

4. 地区差异

然后，本文统计了不同省份上市公司中使用了不同数字技术的比例。^① 统计显示，北京、福建、上海、浙江、江苏等地区的上市公司数字技术使用比例最高。而宁夏、西藏、青海以及内蒙古等地区的上市公司数字技术使用比例最低，即经济发达的东南沿海地区企业数字化转型程度较高，而经济水平较低的中西部地区企业数字化转型程度较低。根据中国信息通信研究院(2022b)测算，截至2022年3月底，我国以“数字经济”命名的产业园累计超过200家，在东部、中部、西部、东北地区的产业园数量占比分别是41%、28%、25%和6%。这一结果与本文的分析结果相互印证。

5. 行业差异

本文依据国民经济行业分类标准，计算了不同行业的上市公司使用数字技术的比例。^② 结果显示，信息传输软件和信息技术服务业、金融业、科学研究和技术服务业的数字化程度最高。农林牧渔、采矿以及公用事业行业的数字化程度较低。归纳为三次产业来看，服务业的数字化水平较高(35%左右)，工业其次(20%左右)，农业最低(9%左右)。根据中国信息通信研究院(2021)对中国2020年三次产业数字经济渗透率的测算，服务业、工业和农业的渗透率分别是40.7%、21%和8.9%，这一结果也说明了本文测度的准确性。

6. 国际比较

最后本文进行国际比较。Zolas et al.(2020)调查了美国85万家企业使用先进技术的情况，其中就包括了人工智能(AI)的使用情况。结果发现，规模越大的企业，使用AI的比例越高。处于规模最大的那部分企业，使用AI的比例高达60%，而那些人数少于50人的企业，使用AI的比例不超过10%。本文以2020年为例，依据企业员工人数分组，描述了企业规模与人工智能使用比例的关系。^③ 从中可以看出，本文构建的人工智能指标与美国的统计数据在规模特征上高度一致，即规模越大的企业，使用人工智能的比例越高。

① 限于篇幅，不同省份上市公司使用数字技术的比例请见本刊网站登载的附表4。

② 限于篇幅，不同行业上市公司使用数字技术的比例请见本刊网站登载的附表5。

③ 限于篇幅，不同规模的企业使用人工智能技术的情况请见本刊网站登载的附图2。

四、企业数字化转型与财务绩效:一些新发现

(一)回归模型和变量定义

对于一个好指标来说,既要准确,还要适用。所谓适用,是指运用该指标进行实证分析时,能够得到符合经济学理论的结果。因此,为了进一步表明本文采用新方法构造的企业数字化转型指标的合理性,并回应引言提到的关于企业数字化转型成败的分歧,下文分析企业数字化转型对企业财务绩效的影响。我们之所以选择企业财务绩效作为因变量,是因为财务绩效是一个客观指标,容易度量并且具有可比性。

已有文献要么从总体上分析了企业数字化转型对企业财务绩效的影响,例如杨德明和刘泳文(2018)、何帆和刘红霞(2019)以及赵宸宇等(2021);要么分析了某一种数字技术对企业财务绩效的影响,包括大数据(Müller et al., 2018)、云计算(Alali & Yeh, 2012)、区块链(林心怡和吴东, 2021)、ICT(Commander et al., 2011; DeStefano et al., 2018)、移动互联(Yang et al., 2018)等。但少有文献比较不同的数字技术是否对企业的财务绩效有不同影响。因此,要回答开头提到的数字化转型成败问题,必须既考虑总体的数字化转型,又区分不同的数字技术带来的财务回报,然后再探讨企业数字化转型影响财务绩效的渠道。

为了探究企业数字化转型对企业财务绩效的影响,本文构建如下基准模型:

$$Y_{i,t} = \alpha + \beta \cdot DT_{i,t} + \sum_n \chi_n \cdot Controls_{i,t} + \lambda_t + \mu_i + \varepsilon_{i,t} \quad (1)$$

其中因变量 $Y_{i,t}$ 表示第 t 年企业 i 的财务绩效,用 ROA (总资产收益率)和 ROE (净资产收益率)度量;关键解释变量 DT 是一组度量企业数字化转型的哑变量,包括是否进行了企业数字化转型($DigiTech$,即企业是否使用了任何一种数字技术),以及是否使用了六种新型数字技术(大数据、人工智能、移动互联、云计算、物联网和区块链)中的一种。 $Controls$ 代表一系列控制变量。参考已有文献的做法(杨德明和刘泳文,2018;赵宸宇等,2021;DeStefano et al., 2018),我们在回归方程中控制了企业年龄、企业规模、增长率(以企业营收同比增速度量)、市值账面比、第一大股东持股比例、董事长是否兼任总经理和现金流等变量。 λ_t 表示时间固定效应, μ_i 表示企业固定效应, $\varepsilon_{i,t}$ 代表随机扰动项。本文使用的是企业层面的聚类标准误。

在数据来源上,除了企业数字化转型指标,本文用到的其他变量来自Wind和CSMAR数据库。由于2008年发生了全球金融危机,2020年暴发了新冠疫情,因此本文的样本范围为2010—2019年。本文剔除了ST和ST*公司,剔除了金融业和关键变量缺失的样本,然后对连续变量进行上下1%的缩尾处理,最终得到25107个观测值。

表1是变量定义和描述性统计。表1显示,样本中60.6%的企业进行了数字化转型,即使用了六类新型数字技术中的任意一种。在六种新型数字技术中,物联网和人工智能的使用比例最高。从财务绩效上看,上市公司的 ROA 平均为6%,但营收增长速度平均值超过14%。

表1 变量定义和描述性统计

变量名称	变量定义	观测值	均值	标准差	最小值	最大值
A 栏:被解释变量						
ROA	净利润/总资产	25107	6.043	6.459	-19.819	25.415
ROE	净利润/权益	25107	6.872	12.661	-63.093	36.335
B 栏:解释变量						
$DigiTech$	是否数字化转型	25107	0.606	0.489	0	1
BD	是否使用大数据	25107	0.269	0.443	0	1
AI	是否使用人工智能	25107	0.369	0.483	0	1
MI	是否使用移动互联	25107	0.269	0.444	0	1

续表 1

变量名称	变量定义	观测值	均值	标准差	最小值	最大值
<i>CC</i>	是否使用云计算	25107	0.143	0.350	0	1
<i>Iot</i>	是否使用物联网	25107	0.476	0.499	0	1
<i>BC</i>	是否使用区块链	25107	0.0332	0.179	0	1
<i>lnAge</i>	企业年龄加 1 取对数	25107	2.879	0.306	2.079	3.555
<i>lnAsset</i>	企业资产(万元)加 1 取对数	25107	12.881	1.278	10.566	16.808
<i>Growth</i>	企业营收同比增速(%)	25107	14.270	30.283	-49.863	157.13
<i>MB</i>	市值账面比	25107	0.585	0.230	0.114	1.081
<i>Top1</i>	第一大股东持股比例(%)	25107	34.836	14.845	8.800	74.660
<i>Dual</i>	董事长是否兼任总经理(1 为是)	25107	0.272	0.445	0	1
<i>Cashflow</i>	现金流/总资产	25107	0.0433	0.0697	-0.166	0.234
C 栏：渠道变量						
<i>TFP1</i>	ACF 方法基于销售额计算的 TFP	21934	1.588	1.133	-1.322	4.623
<i>TFP2</i>	ACF 方法基于经济增加值计算的 TFP	18743	2.575	3.127	-7.424	11.820
<i>lnIncome</i>	总收入加 1 取对数	25107	12.185	1.439	9.013	16.168
<i>lnCost</i>	总支出加 1 取对数	25107	12.116	1.445	9.100	16.132
<i>Cost2Income</i>	总支出/总收入	25107	0.947	0.182	0.560	1.971

(二)基准回归

表 2 提供了基准回归的结果,其中关键解释变量为企业数字化转型哑变量(*DigiTech*)。从中可以看出,无论因变量是 *ROA* 还是 *ROE*,关键解释变量的系数都在 1% 的水平上显著为正。这意味着,当使用基于 ERNIE 模型构造的新指标后,企业数字化转型显著地改善了企业的财务绩效。然而,上述基准回归可能存在反向因果关系,即财务绩效好的企业现金流充足,更有能力使用数字技术。为了缓解反向因果关系,在表 2 第(3)和第(4)列,我们将全部因变量提前一期,发现企业数字化转型的系数依然显著为正。这初步证明,企业数字化转型总体上提高了企业的财务绩效,即数字化转型总体上是成功的。这符合经济学逻辑,并且与部分已有文献(杨德明和刘泳文,2018;何帆和刘红霞,2019;赵宸宇等,2021)的发现是一致的。^①考虑到信息传输、软件和信息技术服务业的上市公司大多是数字产业化的企业,其数字技术的使用与其他行业的企业有所不同,我们将该行业剔除,重新回归的结果仍与基准回归保持一致。^②

表 2 企业数字化转型与企业财务绩效的基准回归结果^③

变量	(1)	(2)	(3)	(4)
	<i>ROA</i>	<i>ROE</i>	<i>F.ROA</i>	<i>F.ROE</i>
<i>DigiTech</i>	0.398*** (0.0896)	0.740*** (0.208)	0.183* (0.0968)	0.435** (0.218)
Controls	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes
N	25107	25107	22552	22552
R ²	0.204	0.152	0.109	0.0777

注:括号内数字为聚类在企业层面的标准误,***、**、*分别表示在 1%、5% 和 10% 的水平上显著。下表同。

① 复刻相关文献的结果显示,当因变量为 *ROA* 时,基于 CSMAR 和方明月等(2022)构造的企业数字化转型虚拟变量是显著为正的,吴非等(2021)的结果不显著,杨金玉等(2022)的结果显著为负。当因变量提前一期时,只有方明月等(2022)的结果是显著的。限于篇幅,相关结果没有在正文展示,请见本刊网站登载的附表 6。

② 限于篇幅,剔除信息传输、软件和信息技术服务业的回归结果请见本刊网站登载的附表 7。

③ 限于篇幅,控制变量的回归结果在正文中没有展示,作者留存备案。

(三)内生性问题

除了反向因果关系,还可能存在遗漏变量问题。为了更好地解决内生性问题,我们构建一个工具变量(IV)。对于使用数字技术进行数字化转型的公司来说,科技人才供给是一个关键问题。因此,在现有文献中,Babina et al.(2024)使用企业与AI强校之间的联系作为企业应用AI的IV,并研究AI对企业增长的影响。背后的逻辑是:企业从AI强校(以AI领域的论文发表来度量)雇佣的员工越多,就越可能应用AI技术。类似地,张叶青等(2021)用“珠峰计划”作为企业采用大数据技术的IV,研究大数据对企业价值(托宾Q)的影响。参考张叶青等(2021)的做法,本文使用“珠峰计划”作为数字化转型的IV。

“珠峰计划”是“基础学科拔尖学生培养试验计划”的简称,是国家为回应“钱学森之问”而推出的人才培养计划。主要内容是在高水平研究型大学和科研院所的优势基础学科建设一批国家青年英才培养基地,建立拔尖人才重点培养体制机制,吸引最优秀的学生投身基础科学研究。第一批“珠峰计划”囊括了17所高校,选择了清华、北大等高校的数学、物理、化学、生物、计算机科学等理工类相关专业作为试点。该计划提高了高校毕业生选择科技类职业的概率,有效增加了科技类人才供给(宋弘和陆毅,2020)。我们推测,上市公司办公地址距离“珠峰计划”17个试点学校的距离越远,越不可能招聘到科技类人才,就越不可能进行数字化转型。与此同时,当地上市公司数目越多,每个上市公司受到该政策的影响就越弱,也越不太可能进行数字化转型。因此,试点学校的距离和上市公司数量均与企业数字化转型的概率成反比,这满足了IV的相关性假设。定义的IV如下:

$$IV_{i,t} = mSumdis_i * mN_{i,t} * Post_t$$

其中, i 代表上市公司, t 代表年份, $mSumdis$ 表示上市公司*i*登记的办公地址到17个试点学校主校区直线距离之和除以10000(km), mN 表示*t*年公司*i*所在城市*c*中的上市公司总数除以1000。由于受“珠峰计划”影响的第一批大学生大部分是2010年入学的本科生,他们的毕业年份为2014年及其之后,因此我们乘以一个时间虚拟变量 $Post$,2014年及之前取值为0,2014年之后取值为1。^①

表3提供了工具变量法的回归结果。第(1)列是第一阶段回归结果,其中IV的系数显著为负,且F统计量大于10,这符合相关性的预期。与此同时,我们认为“珠峰计划”本身不会直接影响某个企业的数字化转型,因此该IV满足排他性假设。第(2)列和第(3)列的结果表明,使用IV之后,关键解释变量企业数字化转型的系数在5%的水平上显著为正。这说明本文的结论是稳健的。

表3 工具变量法回归结果

变量	(1)	(2)	(3)
	<i>DigiTech</i>	<i>ROA</i>	<i>ROE</i>
<i>IV</i>	-0.102*** (0.0300)		
<i>DigiTech</i>		9.455** (4.599)	21.628** (10.053)
Cragg-Donald Wald F statistic	21.487		
Kleibergen-Paap Wald rk F statistic	11.456		
Controls	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes
N	25107	25107	25107
R ²		-0.232	-0.346

① 考虑到部分学生会选择读研,从而推迟毕业年份,我们将毕业年份从2014年先后改为2015年、2016年、2017年和2018年,结果依然稳健。限于篇幅,此处省略,作者留存备索。

（四）不同数字技术的影响

为了探究企业使用某一种数字技术（大数据、人工智能、移动互联、云计算、物联网和区块链）对财务绩效的影响，本节重新选取处理组和对照组。我们将不使用任何数字技术的企业定义为对照组，赋值为0；将单独使用大数据、人工智能、移动互联、云计算、物联网和区块链的企业分别记作BD、AI、MI、CC、IoT和BC，并赋值为1。表4展示了企业单独使用某种数字技术对ROA的影响。第（1）—（5）列显示，无论企业单独使用哪一种数字技术（除区块链），都能显著提高企业的ROA。这与已有研究的结果是一致的。值得注意的是，第（6）列显示区块链技术的回归系数并不显著。原因可能是，区块链技术虽然有助于提高企业信息的安全性（Sharma et al., 2023），但对股价的影响往往是昙花一现，而且放大了企业的收益风险（Jain & Jain, 2019）。^①当我们将ROE作为被解释变量时，也得到了类似结果。^②

表4 企业使用不同数字技术对ROA的影响

变量	(1)	(2)	(3)	(4)	(5)	(6)
	ROA	ROA	ROA	ROA	ROA	ROA
BD	0.533*** (0.151)					
AI		0.544*** (0.125)				
MI			0.344** (0.150)			
CC				0.686*** (0.217)		
IoT					0.447*** (0.102)	
BC						-0.108 (0.370)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
N	16627	19150	16643	13476	21821	10714
R ²	0.193	0.198	0.198	0.189	0.202	0.187

（五）分位数回归

上文分析表明，企业数字化转型能够显著提高企业的财务绩效。但对于不同财务绩效的企业而言，使用数字技术的影响可能存在差异。同时，最小二乘法模型中残差平方和易受极端值影响，回归结果容易出现偏差。为了进一步探究数字化转型对于不同财务绩效企业的影响，我们使用分位数回归方法，分别选取10%、25%、50%、75%和90%这五个分位点进行分位数回归。

本文将ROA作为被解释变量，回归结果如表5所示。第（1）—（3）列回归结果表明，数字化转型对企业ROA有着显著的正向影响；第（4）（5）列结果表明，随着分位数的上升，在75%和90%的高分位点上，数字化转型对企业ROA的影响并不显著。一种可能的解释是，数字技术（例如云计算）为新兴的中小企业赋能，通过轻资产化减轻了它们在大额资本支出和规模方面的相对劣势（Jin & McEl-

① 我们在访谈中了解到，目前区块链技术的应用处于初级阶段，且几乎都是私链或者联盟链，真正去中心化的公链极少。

② 限于篇幅，企业使用不同数字技术对ROE的影响的结果请见本刊网站登载的附表8。

heran, 2017),因此数字化转型可以为绩效较差的企业提供“弯道超车”的机会。当本文将 ROE 作为被解释变量时,结果类似。^①

结合上一小节的内容,本文得到三点结论:第一,企业在进行数字化转型时使用的数字技术不同,数字化转型的效果就不同。目前的证据表明,使用区块链技术并不会改善企业的财务绩效。第二,企业数字化转型的效果存在异质性,即对于财务绩效较差的企业来说,数字化转型的效果较好,但对于财务绩效较好特别是很好的企业来说,数字化转型的效果并不显著。第三,使用不同的方法构造数字化转型指标,会出现不同的回归结果。这回应了本文在引言中提到的数字化转型成败之争。正是因为不同企业数字化转型的路径(数字技术)不同,基础(财务绩效)不同,以及学者们采取的测度方法不同,所以才会出现不同的结果。这也说明,研究企业数字化转型,必须区分不同的数字技术和不同的财务基础,并且构造一个统一、可比的企业数字化转型指标。

表5 企业数字化转型对 ROA 的不同分位数回归

变量	(1)	(2)	(3)	(4)	(5)
	QR_10	QR_25	QR_50	QR_75	QR_90
<i>DigiTech</i>	0.341* (0.198)	0.154* (0.0803)	0.346*** (0.107)	0.147 (0.112)	0.158 (0.192)
Controls	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes
N	25107	25107	25107	25107	25107

(六)渠道分析

接下来本文探讨企业数字化转型提高企业财务绩效的渠道。根据已有文献,可以将数字化转型的渠道归纳为三种:第一种是效率渠道。例如,刘飞(2020)和赵宸宇等(2021)均发现,企业数字化转型能够显著提高生产率,而效率的提升将提高企业的财务绩效(包群和梁贺,2022)。第二种是收入渠道。例如,Yadav(2014)发现,数字化转型会促进企业参与国际贸易活动。此外,消费数据的积累和分析也有助于建立产品忠诚度,从而鼓励消费者重复消费(Hänninen et al., 2017)。这些都有助于企业增加营业收入,进而提高企业的财务绩效。第三种是成本渠道。例如,Shivajee et al.(2019)发现,企业数字化转型有助于降低制造成本,减少零件报废和原材料浪费。何帆和刘红霞(2019)也揭示了这一渠道。为此,本文将验证效率渠道、收入渠道和成本渠道是否成立。

为了检验效率渠道,本文使用企业的 TFP (全要素生产率)作为因变量。测算企业 TFP 的核心问题是解决生产函数估算中的内生性问题,而ACF方法能够有效解决OP和LP法在估计劳动力投入弹性时可能出现的多重共线性问题,因此被广为接受(Loecker & Warzynski, 2012)。在表6第(1)(2)列,我们采用ACF方法,先后基于销售额和经济增加值计算了 $TFP1$ 和 $TFP2$ 。

为了检验收入渠道,在表6的第(3)列,本文将总收入的对数($\ln Income$)作为因变量。为了检验成本渠道,在表6第(4)列将总成本的对数($\ln Cost$)作为被解释变量。综合成本和收入两个维度,我们在第(5)列加入了成本收入比($cost2Income$ =总成本/总收入)作为被解释变量。表6显示,企业进行数字化转型后, TFP 以及总成本分别显著提高和降低,这印证了效率渠道和成本渠道。^②但同时,总收入并未显著增加,这说明收入渠道未被证实。同时,表6的第(5)列显示每单

① 限于篇幅, ROE 为被解释变量的分位数回归结果请见本刊网站登载的附表9。

② 根据江艇(2022),经济学文献在分析渠道时,关键是验证解释变量对渠道(中介变量)的影响,因为渠道变量对因变量的影响通常是不言自明的。为此,我们省略了渠道变量对企业财务绩效的回归分析。

位的收入所需的成本下降了,这说明成本收入比下降,因此总体上数字技术的使用提高了企业的财务绩效。

表 6 渠道分析

变量	(1)	(2)	(3)	(4)	(5)
	<i>TFP1</i>	<i>TFP2</i>	<i>lnIncome</i>	<i>lnCost</i>	<i>cost2Income</i>
<i>DigiTech</i>	0.0121** (0.00516)	0.0214* (0.0121)	-0.00184 (0.00682)	-0.0187*** (0.00720)	-0.0164*** (0.00268)
Controls	Yes	Yes	Yes	Yes	Yes
YearFE	Yes	Yes	Yes	Yes	Yes
FirmFE	Yes	Yes	Yes	Yes	Yes
N	21934	18743	25107	25107	25107
R ²	0.422	0.0989	0.727	0.718	0.145

五、结论与政策启示

作为世界上最大的发展中国家,中国在数字经济领域后来居上,已经成为数字经济大国。在中国数字经济快速发展的过程中,企业的数字化转型是非常重要的微观基础。在中国企业数字化转型实践备受关注的同时,学术界对企业数字化转型的研究也如火如荼。然而,目前学术界在测度企业数字化转型指标时,存在测度对象不统一明确、测度方法不科学准确问题,导致对企业数字化转型的现状和效果存在严重分歧。为了更好地推进企业数字化转型的深入研究,同时呼应中国企业数字化转型的难题,本文基于前沿的机器学习和大语言模型,开发了一套新的企业数字化转型指标。为此,本文首先收集了2006—2020年中国上市公司的年报文本作为分析对象。然后,整理了一个包含311个数字技术关键词的词典,并将数字技术分为六种类型:大数据、人工智能、移动互联、云计算、物联网和区块链。接着,对年报文本进行人工标注,形成训练集。之后,采用有监督的机器学习方法,使用百度开发的ERNIE大语言模型,对年报文本进行预测,判断企业是否使用了数字技术以及使用了何种数字技术,在此基础上构造了中国上市公司数字化转型指标。多项交叉验证结果表明,本文开发的新指标明显优于主流的词典法,并且符合中国企业数字化转型的实践。在本文的后半部分,我们使用新方法构建的指标,证明中国上市公司的企业数字化转型显著提高了财务绩效。而且,对于财务业绩比较差的上市公司而言,数字化转型的效果更加明显。进一步,本文发现数字化转型提高财务绩效的主要渠道是提高效率和降低成本。

本文研究对于推动企业数字化转型和实现经济高质量发展具有重要的政策含义。第一,政府、企业和学界应该通力合作,为企业数字化转型提供基础数据和基础研究。当前,企业在数字化转型方面之所以面临挑战,原因之一是缺乏统一的数字化转型度量以及权威的转型效果分析,这阻碍了有关部门及时地追踪并总结企业数字化转型的成败。要构建一套全面、统一、准确的数字化转型指数和数字化转型数据库,需要掌握企业在数字技术投入、人力投入、资金投入等多方面的信息,而目前学术界主要还是依赖文本分析。本文利用前沿的大语言模型构造的数字化转型指标,为此提供了前期基础和可行方向,但仍需相关数据相互验证。因此,只有政府、企业和学界的多方参与,才能摸清中国企业数字化转型的“底数”,进而为企业进行数字化转型提供理论指导和政策支持。

第二,企业应坚定不移地推进数字化转型。本文研究表明,数字化转型通过提高效率和降低成本,能够改善企业的财务绩效。因此,中国企业应该克服“不愿转、不敢转”的心态,坚定拥抱数字化转型。特别是,在全球化时代,中国企业面临跨国巨头在规模、品牌、资金等方面的传统优势,更应

该充分利用数字经济时代的后发优势,加大力度推进企业数字化转型,以此提高国际竞争力。

第三,政府应根据不同企业的实际情况,推出差异化的数字化转型鼓励政策。本文研究表明,数字化转型对财务绩效较差和一般的企业更有作用,对财务绩效特别好的企业没有显著效果。因此,政府在推进企业数字化转型时,不能一刀切,要尊重规律,多给绩效较差的企业“雪中送炭”,少给绩效好的企业“锦上添花”,以免浪费宝贵的财政资源。这样可以更好地实现数字技术和实体经济的深度融合,促进经济高质量发展。

第四,要鼓励中西部地区的企业加快数字化转型。经济发展水平较低的中西部地区企业数字化转型程度较低,更应该高度重视数字技术的利用,加快数字化转型的步伐。考虑到数字技术所具有的后发优势,政府可以针对中西部地区出台偏向性鼓励政策。例如,加大中西部地区的数字基础设施建设,鼓励东部发达地区的企业向中西部地区提供数字化转型的技术和人才支持。这有利于避免中西部地区和东部地区的企业形成“数字鸿沟”,从而更好地实现共同富裕。

参考文献

- 包群、梁贺,2022:《下放与改制:不同国企改革路径的绩效比较》,《世界经济》第6期。
- 蔡跃洲、牛新星,2021:《中国数字经济增加值规模测算及结构分析》,《中国社会科学》第11期。
- 车万翔、崔一鸣、郭江,2021:《自然语言处理:基于预训练模型的方法》,电子工业出版社。
- 方明月、林佳妮、聂辉华,2022:《数字化转型是否促进了企业内共同富裕?——来自中国A股上市公司的证据》,《数量经济技术经济研究》第11期。
- 何帆、刘红霞,2019:《数字经济视角下实体企业数字化变革的业绩提升效应评估》,《改革》第4期。
- 黄遼友、李增福、潘南佩、倪江崑,2023:《企业数字化转型与劳动收入份额》,《经济评论》第2期。
- 江艇,2022:《因果推断经验研究中的中介效应与调节效应》,《中国工业经济》第5期。
- 金环、魏佳丽、于立宏,2021:《网络基础设施建设能否助力企业转型升级——来自“宽带中国”战略的准自然实验》,《产业经济研究》第6期。
- 李磊、刘常青、韩民春,2022:《信息化建设能够提升企业创新能力吗?——来自“两化融合试验区”的证据》,《经济学(季刊)》第3期。
- 李万利、潘文东、袁凯彬,2022:《企业数字化转型与中国实体经济发展》,《数量经济技术经济研究》第9期。
- 李云鹤、蓝齐芳、吴文锋,2022:《客户公司数字化转型的供应链扩散机制研究》,《中国工业经济》第12期。
- 李舟军、范宇、吴贤杰,2020:《面向自然语言处理的预训练技术研究综述》,《计算机科学》第3期。
- 林心怡、吴东,2021:《区块链技术与企业绩效:公司治理结构的调节作用》,《管理评论》第11期。
- 刘飞,2020:《数字化转型如何提升制造业生产率——基于数字化转型的三重影响机制》,《财经科学》第10期。
- 刘飞、田高良,2019:《信息技术是否替代了就业——基于中国上市公司的证据》,《财经科学》第7期。
- 刘青、肖柏高,2023:《劳动力成本与劳动节约型技术创新——来自AI语言模型和专利文本的证据》,《经济研究》第2期。
- 刘淑春、闫津臣、张思雪、林汉川,2021:《企业管理数字化变革能提升投入产出效率吗》,《管理世界》第5期。
- 祁怀锦、曹修琴、刘艳霞,2020:《数字经济对公司治理的影响——基于信息不对称和管理者非理性行为视角》,《改革》第4期。
- 宋弘、陆毅,2020:《如何有效增加理工科领域人才供给?——来自拔尖学生培养计划的实证研究》,《经济研究》第2期。
- 吴非、胡慧芷、林慧妍、任晓怡,2021:《企业数字化转型与资本市场表现——来自股票流动性的经验证据》,《管理世界》第7期。
- 许雪晨、田侃,2021:《一种基于金融文本情感分析的股票指数预测新方法》,《数量经济技术经济研究》第12期。
- 杨德明、刘泳文,2018:《“互联网+”为什么加出了业绩》,《中国工业经济》第5期。
- 杨金玉、彭秋萍、葛震霆,2022:《数字化转型的客户传染效应——供应商创新视角》,《中国工业经济》第8期。
- 耀友福、周兰,2023:《企业数字化影响关键审计事项决策吗?》,《审计研究》第1期。
- 袁淳、肖士盛、耿春晓、盛誉,2021:《数字化转型与企业分工:专业化还是纵向一体化》,《中国工业经济》第9期。
- 张叶青、陆瑶、李乐芸,2021:《大数据应用对中国企业市场价值的影响——来自中国上市公司年报文本分析的证据》,《经济研究》第12期。
- 赵宸宇、王文春、李雪松,2021:《数字化转型如何影响企业全要素生产率》,《财贸经济》第7期。
- 中国信息通信研究院,2021:《中国数字经济发展报告(2021年)》,中国信息通信研究院网站。
- 中国信息通信研究院,2022a:《全球数字经济白皮书(2022年)》,中国信息通信研究院网站。

- 中国信息通信研究院, 2022b:《中国数字经济发展报告(2022年)》, 中国信息通信研究院网站。
- Acemoglu, D., and P. Restrepo, 2020, “Robots and Jobs: Evidence from US Labor Markets”, *Journal of Political Economy*, 128(6), 2188—2244.
- Acikalin, U.U., T. Caskurlu, G. Hoberg, and G.M. Phillips, 2022, “Intellectual Property Protection Lost and Competition: An Examination Using Machine Learning”, National Bureau of Economic Research(2022—11).
- Alali, F.A., and C.-L. Yeh, 2012, “Cloud Computing: Overview and Risk Analysis”, *Journal of Information Systems*, 26(2), 13—33.
- Babina T., A. Fedyk, A. He, and J. Hodson, 2024, “Artificial Intelligence, Firm Growth, and Product Innovation”, *Journal of Financial Economics*, 151, 103745.
- Bloom, N., L. Garicano, R. Sadun, and J.V. Reenen, 2014, “The Distinct Effects of Information Technology and Communication Technology on Firm Organization”, *Management Science*, 60(12), 2859—2885.
- Brynjolfsson, E., W. Jin, and K. McElheran, 2021, “The Power of Prediction: Predictive Analytics, Workplace Complements, and Business Performance”, *Business Economics*, 56(4), 217—239.
- Chen, Y., 2020, “Improving Market Performance in the Digital Economy”, *China Economic Review*, 62, 101482.
- Chen, M.A., Q. Wu, and B. Yang, 2019, “How Valuable is FinTech Innovation?”, *Review of Financial Studies*, 32(5), 2062—2106.
- Chun, H., J.-W. Kim, R. Morck, and B. Yeung, 2008, “Creative Destruction and Firm-Specific Performance Heterogeneity”, *Journal of Financial Economics*, 89(1), 109—135.
- Commander, S., R. Harrison, and N. Menezes-Filho, 2011, “ICT and Productivity in Developing Countries: New Firm-Level Evidence from Brazil and India”, *Review of Economics and Statistics*, 93(2), 528—541.
- DeStefano, T., R. Kneller, and J. Timmis, 2018, “Broadband Infrastructure, ICT Use and Firm Performance: Evidence for UK Firms”, *Journal of Economic Behavior & Organization*, 155, 110—139.
- Goldfarb, A., and C. Tucker, 2019, “Digital Economics”, *Journal of Economic Literature*, 57(1), 3—43.
- Hänninen, M., A. Smedlund, and L. Mitronen, 2017, “Digitalization in Retailing: Multi-Sided Platforms as Drivers of Industry Transformation”, *Baltic Journal of Management*, 13(2), 152—168.
- Jin, W., and K. McElheran, 2017, “Economies before Scale: Survival and Performance of Young Plants in the Age of Cloud Computing”, SSRN working paper, No.3112901.
- Jain, A. WP, and C. Jain, 2019, “Blockchain Hysteria: Adding ‘Blockchain’ to Company’s Name”, *Economics Letters*, 181, 178—181.
- Loecker, J.D., and F. Warzynski, 2012, “Markups and Firm-Level Export Status”, *American Economic Review*, 102(6), 2437—2471.
- Ma, H., M. Zhao, D. Yan, and H. Wang, 2021, *The Chinese Digital Economy*, Singapore: Springer Singapore.
- Müller, O., M. Fay, and J. vom Brocke, 2018, “The Effect of Big Data and Analytics on Firm Performance: An Econometric Analysis Considering Industry Characteristics”, *Journal of Management Information Systems*, 35(2), 488—509.
- Rajan, R., P. Ramella, and L. Zingales, 2023, “What Purpose Do Corporations Purport? Evidence from Letters to Shareholders”, National Bureau of Economic Research(2023—03).
- Sharma, P., D.M. Shukla, and A. Raj, 2023, “Blockchain Adoption and Firm Performance: The Contingent Roles of Intangible Capital and Environmental Dynamism”, *International Journal of Production Economics*, 256, 108727.
- Shivajee, V., R.K. Singh, and S. Rastogi, 2019, “Manufacturing Conversion Cost Reduction Using Quality Control Tools and Digitization of Real-Time Data”, *Journal of Cleaner Production*, 237, 117678.
- Siebel, T.M., 2019, *Digital Transformation: Survive and Thrive in an Era of Mass Extinction*, New York: Rodin Books.
- Sun, Y., S. Wang, Y. Li, S. Feng, X. Chen, H. Zhang, X. Tian, D. Zhu, H. Tian, and H. Wu, 2019, “ERNIE: Enhanced Representation Through Knowledge Integration”, arXiv:1904.09223 [Cs].
- Yadav, N., 2014, “The Role of Internet Use on International Trade: Evidence from Asian and Sub-Saharan African Enterprises”, *Global Economy Journal*, 14(2), 189—214.
- Yang, S., Z. Li, Y. Ma, and X. Chen, 2018, “Does Electronic Banking Really Improve Bank Performance? Evidence in China”, *International Journal of Economics and Finance*, 10(2), 82—94.
- Zolas, N., Z. Kroff, E. Brynjolfsson, K. McElheran, D.N. Beede, C. Buffington, N. Goldschlag, L. Foster, and E. Dinlersoz, 2020, “Advanced Technologies Adoption and Use by U.S. Firms: Evidence from the Annual Business Survey”, National Bureau of Economic Research(2020—12).

Measurement Problem of Enterprise Digital Transformation: New Methods and Findings Based on Large Language Models

JIN Xingye^a, ZUO Congjiang^a, FANG Mingyue^b, LI Tao^a and NIE Huihua^c

(a: School of Economics, Central University of Financial and Economics;

b: College of Economics and Management, China Agricultural University;

c: School of Economics, Renmin University of China)

Summary: The rapidly growing digital economy in China has become the second-largest digital economy in the world. The importance of digital transformation for enterprises is widely acknowledged, but there are serious disagreements about the effects of enterprise digital transformation. The main reason for this phenomenon is that existing research has problems in measuring enterprise digital transformation, which is reflected in two aspects: Firstly, the measurement objects are neither unified nor clearly defined; secondly, the measurement methods are not scientific and accurate enough. This results in many research conclusions being incomparable, difficult to replicate, and often contradictory. Therefore, to ensure that theory effectively informs practice, academia needs to reach a consensus on the measurement of enterprise digital transformation and work hard to alleviate the problem of inaccurate measurement methods. Only in this way can the confusion be clarified, and the way for theoretical insights be paved.

This paper uses an advanced machine learning method, the large language models, to construct a set of digital transformation indicators based on the annual report texts of Chinese listed companies from 2006 to 2020. Specifically, the indicators are measured in five steps. The first step is to sort out the annual reports of listed companies, and use the two parts of the annual report, “Management Discussion and Analysis” and “Table of Contents, Definitions, and Key Risk Indicators” as relevant texts for enterprise digital transformation. The second step is to divide all texts into sentences to form a predictive sentence pool. In the third step, a set of sentences is randomly selected and extracted based on the presence of keywords to constitute a pending tagged sentence library. This library is manually annotated to determine whether companies are using digital technologies including big data, artificial intelligence (AI), mobile Internet, Internet of Things, blockchain, and cloud computing. The fourth step uses supervised machine learning methods with models such as ERNIE and BERT for the training of sentence classifiers. The fifth step is to use the trained ERNIE model to predict sentence by sentence in the predictive sentence pool, to assess whether and which digital technologies are utilized by the listed companies, thereby constructing a new set of digital transformation indicators for enterprises. To verify the effectiveness of the new indicators, this paper conducts comparisons with patent data, regional data and international literature in six aspects, and finds that the digital transformation indicators constructed are highly consistent with reality. With these indicators, this paper empirically tests the relationship between enterprise digital transformation and corporate financial performance, and obtains some new findings. Digital transformation generally improves financial performance (measured with ROA and ROE). Big data, AI, mobile Internet, cloud computing, and Internet of Things improve ROA and ROE, while blockchain has no such effect. For companies with poor financial performance, digital transformation can significantly improve financial performance, while for companies with good financial performance, digital transformation has no significant effect on financial performance. There are two main channels for enterprise digital transformation to improve financial performance, namely, improving efficiency and reducing costs.

The contributions of this paper can be summarized in three aspects. Firstly, this paper provides a new method for measuring enterprise digital transformation. It proposes a novel approach to construct digital transformation indicators based on the annual reports of listed companies in China. The new indicators promote in-depth research on enterprise digital transformation in terms of research methods, and provide empirical evidence from China for the general digital economics literature. Secondly, this paper reveals impacts of different digital technologies on corporate financial performance and identifies different channels. Thirdly, this paper enriches the application of large language models in economic literature, for there are limited studies using large language models. This article is of great significance for promoting enterprise digital transformation and achieving high-quality economic development.

Keywords: Enterprise Digital Transformation; Digital Economy; Digital Technology; Artificial Intelligence; Large Language Models

JEL Classification: C89, C43, C45

(责任编辑:刘洪愧)(校对:何伟)