国家大数据综合试验区设立是否促进减污降碳协同增效？
——基于双重机器学习的因果推断

张洪昌[1,2]，王越[2]，王伟梅[2*]，陈瑜[3]

（1. 贵州财经大学绿色发展战略高端智库，贵州贵阳 550025；2. 贵州财经大学应用经济学院，贵州贵阳 550025；3. 贵州财经大学会计学院，贵州贵阳 550025）

【摘　要】推动减污降碳协同增效是实现中国生态文明建设与经济高质量发展的重要目标。作为数据要素市场化改革的先行试点，国家大数据综合试验区的政策效应能否破解污染减排、碳排放强度降低与经济增长之间的协同困境亟待探明。本研究基于2011—2022年中国297个地级及以上城市面板数据，以“国家大数据综合试验区”（以下简称“试验区”）设立为准自然实验，采用双重机器学习模型探究试验区设立对减污降碳协同增效的因果效应。研究结果表明：①试验区设立显著促进减污降碳协同增效，这一结论在内生性检验及稳健性检验后依然成立；②试验区设立通过产业结构升级、绿色技术创新与数字普惠金融发展三链协同机制推动减污降碳协同增效；③异质性分析表明，试验区设立在高新质生产力关注度城市、非资源型城市及大型城市中效果更为显著；④试验区设立在促进减污降碳协同增效方面具有空间协调效应，能够在全国和地区层面缩小减污降碳协同增效的相对发展差距。本研究为数字经济时代统筹污染治理、碳减排与经济增长的协同路径提供新论据，为差异化构建数据赋能的环境治理体系、推动经济社会全面绿色低碳转型提供政策启示。

【关键词】国家大数据综合试验区；减污降碳协同增效；双重机器学习；空间协调

【中图分类号】X321；X51；F123.9　　【文章编号】1674-6252（2025）05-0144-11

【文献标识码】A　　【DOI】10.16868/j.cnki.1674-6252.2025.05.144

引言

党的二十大报告明确指出，“协同推进降碳、减污、扩绿、增长，推进生态优先、节约集约、绿色低碳发展”，这一战略部署深刻揭示了中国式现代化进程中环境治理与经济发展的辩证统一关系。在全球气候变化加剧与国内生态环境承压的双重挑战下，生态环境部等七部门印发《减污降碳协同增效实施方案》，着力破解大气污染物与温室气体“同根同源”的治理困境，推动形成环境效益、气候效益与经济效益的协同发展新格局。当前，随着数字技术革命纵深演进，大数据赋能环境治理的创新实践为这一战略目标提供了全新解题思路——数据政策作为数据要素市场化改革的先行者，通过重塑环境治理的信息基础与技术范式，或为破解传统减污降碳战略中存在的“信息孤岛”与“效率衰减”难题提供关键路径。然而，既有研究多聚焦于生产端治理[1]或能源结构调整[2]，学界对数字技术

资助项目：贵州省高校人文社会科学研究项目“贵州传统产业数字化转型的路径与模式研究”（2025RW021）。

作者简介：张洪昌（1987—），男，博士，副研究员、博士生导师，研究方向为数字经济，E-mail：zhc2222@126.com。

*** 责任作者：**王伟梅（2001—），女，硕士研究生，研究方向为环境经济，E-mail：wmwang@mail.gufe.edu.cn。

驱动的环境治理模式转型及其协同效应的系统性研究仍显不足。在此背景下，深入探究数据政策对减污降碳协同增效的影响和作用机制，不仅能够拓展环境规制理论的技术演进维度，更能为数字经济时代下协同推进多维治理目标提供实践范式。

数据政策的环境影响研究长期是学术研究的核心领域，当前围绕本研究的核心议题可概括为三个方面：一是数据政策的减污效应。Wu 等 [3]、姜祎等 [4] 指出，数据政策依托现代信息技术深化城市发展模式革新，借助数字手段驱动技术迭代与产业结构优化，最终形成减污降耗的复合效能。二是数据政策的降碳效应。Song 等 [5]、Liu 等 [6]、王琴等 [7] 认为数据政策通过构建能耗实时监测平台与能效智能调控系统，在工业生产过程优化与交通运输网络运行中精准削减碳排放强度。三是数据政策的经济效应。Zhang[8]、Duygan 等 [9]、秦文晋等 [10] 认为数据政策通过构建“数据资本深化—知识溢出增强—边际收益递增”的传导机制推动经济发展。可以看出，现有研究普遍认可数据政策在减污、降碳和经济推动方面的潜力，但多将污染控制、碳排放与经济发展割裂分析，对三者动态协同关系的系统性探究不足。在政策研究的方法论层面，双重差分模型与合成控制法虽是主流政策评估工具，但其局限性显而易见：双重差分模型对数据质量要求严格，合成控制法难以有效处理多政策对象与多对照组的复杂场景。相较之下，双重机器学习模型因其无偏估计与结果正态分布的特性，为政策效应研究提供了更科学的解决方案 [11]。然而，机器学习技术在经济领域的应用仍以理论探索为主，政策评估中的实践较为匮乏。

基于此，本文以国家大数据综合试验区设立为准自然实验，考察试验区设立对减污降碳协同增效的影响，边际贡献在于：①研究视角上，试验区的设立不仅为减污降碳协同增效提供了新视角，而且超越了对污染、碳排放和经济增长的传统单一关注点；②研究对象上，既有研究大多采用省级面板数据，本研究细化到市级层面；③研究方法上，充分发挥双重机器学习在高维变量和非参数估计领域的优势，将双重机器学习模型应用于数字政策效果的评估中，以此克服传统双重差分法存在的“维度诅咒”难题和模型设定误差，提高研究结论的可信度；④研究内容上，本文从产业结构升级、绿色技术创新、数字普惠金融三个方面考察了试验区设立对减污降碳协同增效的间接影响，同时基于新质生产力关注度、资源禀赋和城市规模视角综合考察试验区政策实施效果的异质性，并剖析试验区设立对减污降碳协同增效的空间协调效应和社会效应。为决策部门构建差异化环境规制体系、优化区域协同治理机制提供了经验证据与路径参照。

1 政策背景与研究假设

1.1 政策背景

在全球数字经济浪潮中，我国自 2013 年起通过“宽带中国”“智慧城市”等政策推动数字技术与实体经济融合，但传统政策多聚焦技术应用层面，未能破解数据要素市场化改革中的深层次瓶颈，如政府数据开放不足、企业间数据流通交易障碍、跨领域数据融合困境等问题 [12]。为此，2015 年 8 月，《促进大数据发展行动纲要》将国家大数据综合试验区设立上升为国家战略，通过分层次、差异化的区域试点探索制度创新路径。同年 9 月，贵州率先获批首个国家级试验区，随后 2016 年国家分两批扩容试验区，形成覆盖全国多区域的“八大试验区”差异化创新网络。试验区的设立显著提升了企业数字创新水平。通过政策激励和数据资源集聚，企业加大了研发投入，提升了研发强度。同时，试验区内的知识流动和数据共享为企业数字创新提供了有力支持。在区域层面，试验区显著提升了科技创新能力，推动了区域内产业的数字化转型和升级，带动了相关产业发展，提升了区域经济竞争力。在国家层面，试验区为“数据二十条”等全国性改革提供了宝贵经验，验证了数据要素市场化改革的可行性和有效性，为我国数字经济的蓬勃发展奠定了基础。

随着我国“双碳”目标的提出，大数据综合试验区建设与减污降碳协同增效目标紧密衔接。试验区在数据资源集聚与流通方面优势显著，为环境治理提供了强大助力 [13]。通过整合和分析环境监测数据、企业生产数据等多源数据，试验区实现了对污染物排放的精准监测和预测，为制定科学合理的减排措施提供了依据。同时，试验区推动传统产业数

字化转型，使企业借助数字技术优化生产流程，降低能源消耗与污染物排放。试验区还吸引大量数字经济企业和创新人才，形成新的产业集群与创新生态。这些新兴数字经济产业本身具有低能耗、低污染的特点，为减污降碳协同增效注入新动力。在政策引导、数据驱动与产业协同的共同作用下，试验区成功实现从数字经济发展到绿色低碳转型的平稳过渡，为我国高质量发展提供重要支撑。

1.2 研究假设

在“双碳”战略与数字经济战略交会背景下，国家大数据综合试验区被赋予探索数据要素市场化改革与绿色低碳转型协同路径的使命。其制度设计不仅着眼于经济增长，更试图通过数据赋能破解环境治理与经济增长之间的结构性张力，由此形成了一套兼具减污、降碳与增长三重目标的政策逻辑。试验区通过整合环境数据资源与智能算法技术，构建全链条治理体系，破解传统环境管理难题 [14]，实现全要素动态管控与差异化激励机制 [15]，强化跨区域污染联防联控协同性，提升政策执行的减污效应；试验区以数字经济优化能源配置效率与产业结构升级为核心抓手 [16]，推动高耗能产业向低碳模式转型，降低单位产出碳排放强度 [17]。尽管存在技术替代效应与碳减排效果的时空异质性 [18]，但数字技术赋能与区域协同治理机制耦合可实现系统性碳减排目标。试验区创新数据要素市场化配置机制 [19]，引导企业优化要素配置结构，释放数据要素价值，提升企业就业水平、全要素生产率与创新水平，形成技术革新驱动的经济增长模式 [20]。京津冀、贵州等先行试验区的大数据产业规模保持较高增长，带动新型工业化示范基地建设。基于此，提出研究假设 1。

H1：国家大数据综合试验区设立能够促进减污降碳协同增效。

基于结构变迁理论，经济体的要素禀赋再配置与技术范式迭代将驱动产业结构向高级化方向演进，这一过程通过优化资源配置效率实现增长动力转换与环境负荷消减的协同平衡 [21]。试验区作为新型生产要素的体制机制创新载体，其政策效应深度嵌入区域产业结构升级的动态过程中：一方面，试验区通过产能识别机制，利用产业数据中心整合能耗、排放数据，精准识别高耗能行业的冗余产能，推动低端产能淘汰与工艺流程再造，在压缩“两高”产业规模的同时，提高资源集约化利用效率 [22]；另一方面，试验区通过流程再造机制，工业互联网平台将数据反馈嵌入生产流程，降低单位产出的资源消耗，实现“两高”产业规模压缩与效率提升的协同。这种“破旧立新”的结构调适机制，不仅能破除传统末端治理的减污降碳困境，还能依托数字经济与服务经济的融合增效开辟绿色增长新空间，最终在产业生态高级化进程中达成环境质量改善与低碳经济培育的深度协同。基于此，提出研究假设 2。

H2：国家大数据综合试验区设立能够通过产业结构升级实现减污降碳协同增效。

基于内生增长理论，技术内生变革通过知识资本积累撬动生产函数重构，成为突破环境—增长双约束的治理枢纽 [23]。试验区的设立不仅为绿色技术创新提供了数据要素支撑，更通过构建创新生态体系，实现了绿色技术从研发到应用的全链条激励机制的范式革新：其一，试验区整合企业污染排放、能源消耗与生态保护数据，建立绿色技术专利池与数字孪生验证平台，降低绿色技术研发的试错成本与市场风险，推动创新主体突破节能降耗、碳捕集等关键技术瓶颈；其二，试验区依托产业数据中心共享的污染排放与能源消耗大数据，降低绿色技术研发中的重复试验与验证成本，实现研发成本的社会化分摊。同时，数据平台将实时环境合规信息精准转化为技术需求信号，显著减少供需双方的信息摩擦，强化绿色技术市场的价格与数量信号，全面激励创新主体的研发与应用行为 [24]。绿色技术创新的外溢效应促使减污降碳的公共环境收益转化为企业竞争力提升的私有经济价值，在环境库兹涅茨曲线的改善过程中开辟高质量发展新路径。基于此，提出研究假设 3。

H3：国家大数据综合试验区设立能够通过绿色技术创新实现减污降碳协同增效。

基于新结构经济学最优金融结构理论，数字普惠金融发展通过降低环境治理的交易成本与优化绿色资源配置效率，可成为破解减污降碳支付意愿与行动能力结构性错配的制度性工具 [25]。试验区的设立通过构建数据要素和金融科技的融合创新机

制，在三个层面优化环境治理的金融支持体系：其一，试验区以统一的数据接口归集企业排污许可、碳配额及物联网监测信息，据此训练环境信用评分模型，使金融机构得以识别高潜力低碳企业并实施差异化信贷，显著降低绿色技术投资中的信息不对称与抵押品约束；其二，在数据可信共享的基础上，试验区推动供应链金融产品的环境绩效挂钩机制——将碳排放强度实时写入融资合同触发条款，形成“减排改善—利率下调—再投资绿色改造”的正反馈回路，强化产业链整体清洁化[26]；其三，依托政府开放的碳普惠数据接口，试验区支持商业银行与互联网平台合作发行碳积分数字权益，把个人低碳出行、节能消费等行为纳入可交易、可变现的激励体系，扩展绿色金融的触达范围。上述由数据要素驱动的金融创新网络，通过降低风险评估成本与资金错配概率，实现污染治理成本的社会化分摊与低碳转型动能的市场化积累，最终提升减污降碳协同增效。基于此，提出研究假设 4。

H4：国家大数据综合试验区设立能够通过促进数字普惠金融实现减污降碳协同增效。

2 研究设计

2.1 模型构建

本文针对国家大数据综合试验区政策对城市减污、降碳与经济增长协同效应的影响开展因果识别。鉴于环境治理与经济发展存在复杂的多维度交互关系，传统模型在分析过程中易受多重共线性与模型设定偏误干扰，导致政策效应评估失准。为避免传统计量方法在处理异质性样本与解析复杂机制上的局限性，本文借鉴 Xing 等[27]、张涛等[11]的做法，创新性地将双重机器学习模型引入试验区的政策效果评估，并构建以下部分线性双重机器学习模型：

$$\text{SCPCE}_{it} = \theta_0 \ \text{Bigdata}_{it} + g(X_{it}) + \varepsilon_{it} \tag{1}$$

$$E(\varepsilon_{it} | \ \text{Bigdata}_{it}, X_{it}) = 0 \tag{2}$$

式中，SRPCE_{it} 是因变量，它表示城市 i 在 t 年的减污降碳协同增效水平；Bigdata_{it} 是自变量，表示试验区政策变量；θ_0 为政策处置系数；X_{it} 表示一组高维控制变量；$g(X_{it})$ 表示自变量对控制变量的回归函数；ε_{it} 为随机误差项。

针对式（1）和式（2）直接估计可能导致参数估计值 $\hat{\theta}$ 产生偏误的问题，其根源在于在高维或复杂模型设置中，机器学习模型必须引入正则项来降低维数。这种技术处理虽有效抑制了估计量的过度拟合风险，但也会在函数中引入正则偏差，使得 θ_0 难以收敛到 $\hat{\theta}$。为了解决这个问题，我们构建了一个如下辅助回归：

$$\text{Bigdata}_{it} = m(X_{it}) + \mu_{it} \tag{3}$$

$$E(\mu_{it} | \ X_{it}) = 0 \tag{4}$$

式中，$m(X_{it})$ 为处置变量对高维控制变量的回归函数；μ_{it} 为随机误差项。此外，为了降低机器学习算法过度拟合造成的估计偏误，本文采用样本交叉验证方法，特别是K-fold交叉验证方法（1：4），以增强数据利用率，防止过度拟合，确保更稳健的参数估计，并采用随机森林的机器学习算法对主回归和辅助回归进行预测求解。

2.2 变量设定

2.2.1 被解释变量

减污降碳协同增效（SRPCE）包括 3 个方面：①污染排放（$PM_{2.5}$），以城市 $PM_{2.5}$ 浓度均值衡量；②碳排放量（$\ln CO_2$），以城市碳排放量的自然对数表征；③经济增长（ln rgdp），采用城市人均生产总值的自然对数衡量。本文参考赵燕等[28]、丛晓男[29]，使用耦合协调度模型测算减污降碳协同增效水平。

2.2.2 核心解释变量

国家大数据综合试验区政策处理变量（Bigdata），试验区建设采取分阶段推进模式：2015 年贵州省率先实施大数据产业顶层规划，形成“数据先行区”示范效应；2016 年，试验区建设进入扩容期，京津冀城市群等 7 个省级行政区被纳入第二批建设名录。基于区域集约化赋权机制，本文将满足以下条件的观测单元设为处理组：若所属省级行政区在该年度被纳入试验区名录或已完成基础建设，其下辖所有地级市即被赋值为 1，反之为 0。

2.2.3 机制变量

①产业结构升级（IS），采用城市第三产业增加值与城市第二产业增加值的比值衡量。②绿色技术创新（GTI），采用城市申请绿色专利数量与城市

申请专利总量之比进行衡量[30]。③数字普惠金融（DF），采用地级市层级的“中国数字普惠金融指数”衡量[31]。

2.2.4 控制变量

为确保政策效果估计的准确性，并基于对相关研究的参考[32]，我们进一步控制可能影响减污降碳协同增效的其他因素。控制变量设置如下：①教育水平（EL），采用教育支出占城市GDP 比重衡量；②科技水平（TL），采用科技支出占城市 GDP 比重衡量；③人口密度（PD），采用区域常住人口与城市面积之比的对数衡量；④财政分权度（FD），采用一般政府收入与城市一般政府支出之比衡量；⑤就业结构（ES），采用第三产业就业人数占城市总就业人数比重衡量；⑥社会消费水平（SCL），采用社会消费品零售额占城市 GDP 比重衡量；⑦政府干预程度（DGI），政府支出占城市 GDP 比重衡量；⑧工业化水平（IL），采用第二产业增加值占城市 GDP 比重衡量；⑨金融发展水平（FDL），采用金融机构各项贷款余额占城市 GDP 比重衡量。此外，回归分析加入了城市变量的二次项，以提高模型的准确度。

2.3 数据说明

为确保样本数据的连续性和可获得性，并综合考虑行政规划与数据可得性等因素，本文最终确定的样本为我国 297 个地级及以上城市，研究时间跨度设定为 2011—2022 年。城市碳排放数据来自全球大气研究排放数据库（EDGAR），其他城市层面数据源自《中国城市统计年鉴》、各地统计年鉴、各市国民经济和社会发展统计公报及中国研究数据服务平台（CNRDS）数据库。针对少量数据缺失问题，本文采用线性插值法进行补充，各变量描述性统计见表 1。

3 实证结果与分析

3.1 基准回归结果

表 2 第（1）至第（3）列报告了试验区设立影响减污降碳协同增效的估计结果，无论选用哪种算

表1　主要变量的描述性统计

变量类型	变量名称	符号	观测值	均值	标准差	最小值	最大值
被解释变量	减污降碳协同增效	SRPCE	3 564	0.361	0.037	0.000	0.492
核心解释变量	试验区设立的虚拟变量	Bigdata	3 564	0.111	0.315	0.000	1.000
机制变量	产业结构升级	IS	3 564	1.092	0.623	0.113	5.928
	绿色技术创新	GTI	3 564	0.117	0.573	0.000	11.222
	数字普惠金融	DF	3 564	192.975	76.219	17.020	361.066
控制变量	教育水平	EL	3 564	0.018	0.099	0.0002	5.891
	科技水平	TL	3 564	0.002	0.002	0.000	0.063
	人口密度	PD	3 564	5.725	0.933	0.554	7.905
	财政分权度	FD	3 564	0.440	0.222	0.020	1.541
	就业结构	ES	3 564	0.561	0.148	0.126	1.051
	社会消费水平	SCL	3 564	0.377	0.112	0.000	1.012
	政府干预程度	DGI	3 564	0.226	0.197	0.043	2.672
	工业化水平	IL	3 564	0.448	0.113	0.088	0.893
	金融发展水平	FDL	3 564	1.089	0.686	0.000	12.774

表2　基准回归与稳健性检验结果

变量	(1) 随机森林 SRPCE	(2) 套索回归 SRPCE	(3) 神经网络 SRPCE	(4) 替换被解释变量 SRPCE1	(5) 加入省份—时间交互项 SRPCE
Bigdata	0.013*** (3.83)	0.504*** (83.13)	0.096*** (14.47)	0.018*** (5.40)	0.004** (2.09)
常数项	0.000 (0.19)	0.000 (0.00)	0.001*** (3.36)	0.000 (0.95)	−0.000 (−0.59)
控制变量一次项	是	是	是	是	是
控制变量二次项	是	是	是	是	是
年份固定	是	是	是	是	是
城市固定	是	是	是	是	是
样本量	3 564	3 564	3 564	3 564	3 564

注：***、**、* 分别表示在 1%、5%、10% 的水平上显著；括号内为 z 值。下表同。

法，试验区设立对减污降碳协同增效的影响在1%的水平下显著为正。可能的原因是，试验区通过整合环境监测大数据与区块链技术构建智能减污降碳追踪系统，显著提升污染源识别和排放核算精度，同时数字基础设施支撑的绿色创新网络促进清洁技术迭代，并借助数字金融工具优化资源配置[25]，最终实现污染减排、碳减量与经济增长的非线性协同跃迁。本文的研究假设1得到验证，即试验区设立能够促进减污降碳协同增效。为了方便起见，后文分析均选择随机森林算法进行计算。

3.2 稳健性检验

3.2.1 改变被解释变量的衡量方式

进一步用熵权法[33]替代减污降碳协同增效的测算（SRPCE1），$PM_{2.5}$、ln CO_2 和 ln rgdp 三者的权重分别为0.199、0.515和0.286。由表2第（4）列可以看出，试验区对减污降碳协同增效的影响仍在1%的水平下显著为正，说明本文基准回归的稳健性。

3.2.2 加入省份—时间固定效应

省级行政区作为我国重要的行政层级，同一省级行政区内城市的政策实施和治理逻辑存在共性，因此本文通过加入省份—时间固定效应控制各省级行政区随时间变化的潜在影响因素，以验证核心结论在不同行政层级与时间维度下的稳健性。由表2第（5）列可以看出，加入省份—时间固定效应后，试验区对减污降碳协同增效的回归系数依然显著为正，进一步验证了基准回归。

3.2.3 剔除异常值的影响

由于样本中的异常值可能通过交互作用干扰减污降碳协同增效指数的测算结果。本文在基准回归基础上，除对所有连续变量分别进行1%～99%、5%～95%双侧缩尾处理外，还对控制变量的二次项进行相同截断操作，以确保参数估计不受分布尾部异常值的随机性扰动。由表3第（1）列和第（2）列可以看出，在不同缩尾尺度下试验区的影响系数仍保持显著正向特征，验证了核心结论对数据分布条件敏感性较低且具有统计稳健性。

3.2.4 更改样本分割比例

考虑到样本分割比例可能影响双重机器学习估计器的表现，本文将训练集与估计集的样本分割比例由基准的1∶4拓展为1∶2和1∶7进行敏感性分析，探究不同数据划分方式对政策效应估计的影响。由表3第（3）列和第（4）列可以看出，在调整样本分割比例后，试验区的估计系数在1%的水平下显著为正，进一步证实了政策效应的估计结果并非源于特定的样本划分方式。

3.2.5 消除并行政策的影响

由于我国数据政策实施具有试点叠加特征（如2013年“智慧城市建设”和2014年“宽带中国”战略形成政策协同），同一时期内城市可能受到多项数字改革政策的交错影响，因此本文在基准回归中纳入“智慧城市”（Smartcity）和“宽带中国”（Broadband）的交互虚拟变量，用以控制并行政策对减污降碳协同效应的潜在扰动。由表3第（5）列可以看出，试验区的估计系数显著性未受显著影响，证实核心结论的稳健性。

3.2.6 内生性检验

本文参考Chernozhukov等[34]的做法，构建双重

表3 稳健性检验结果

变量	(1)	(2)	(3)	(4)	(5)	(6)
	消除异常值 SRPCE		更改样本分割比例 SRPCE		消除其他并行政策影响 SRPCE	工具变量 SRPCE
Bigdata	0.013*** (3.83)	0.006*** (3.23)	0.013*** (4.83)	0.010*** (4.01)	0.013*** (3.97)	0.495*** (3.67)
常数项	0.000 (0.19)	0.000 (0.36)	0.000 (0.84)	0.000*** (0.02)	0.000 (0.08)	0.001 (0.91)
控制变量一次项	是	是	是	是	是	是
控制变量二次项	是	是	是	是	是	是
年份固定	是	是	是	是	是	是
城市固定	是	是	是	是	是	是
样本量	3 564	3 564	3 564	3 564	3 564	3 564

机器学习的部分线性工具变量模型，具体设置如下：

$$\text{Bigdata}_{it+1} = \theta_0 \, \text{Event}_{it} + g(X_{it}) + \varepsilon_{it} \quad (5)$$

$$\text{Instrument}_{it} = m(X_{it}) + \mu_{it} \quad (6)$$

式中，Instrument 为工具变量。本文参考 Nunn 等[35]的做法，采用城市地形起伏度与时间趋势项的交互项，此变量满足工具变量的外生性与相关性假设。一方面，作为一种由长期地质演化决定的地理要素，地形起伏度在短期内既不受环境政策干预，也不直接影响减污降碳协同机制，满足工具变量外生性假设；另一方面，数字基础设施的铺设与运维成本对地形敏感，地形起伏越大，光缆、基站及数据中心的建设费用越高，从而显著降低城市获批试验区的概率，符合相关性的要求。由表 3 第（6）列可以看出，第二阶段回归估计系数均显著为正，表明在解决内生性问题困扰的情况下，试验区对减污降碳协同增效的显著促进作用依然成立，由此证明基准回归结论稳健。

3.3 机制分析

为检验试验区设立对减污降碳协同增效的影响机制，本文参考杨思莹等[36]的做法，构建部分线性机制分析模型：

$$M_{it} = \theta_0 \, \text{Bigdata}_{it} + g(X_{it}) + \varepsilon_{it} \quad (7)$$

$$E(\varepsilon_{it} \mid \text{Bigdata}_{it}, X_{it}) = 0 \quad (8)$$

本文在原模型基础上将解释变量替换为试验区虚拟变量 Bigdata。其中，M_{it} 代表机制变量，分别采用产业结构升级、绿色技术创新和数字普惠金融衡量。

3.3.1 产业结构升级

由表 4 第（1）列可以看出，试验区设立显著促进产业结构升级。此外，相关研究表明产业结构升级对减污降碳协同增效起显著促进作用[37]。可能的解释是，试验区政策加速产业从高耗能制造业向数字化服务业的结构转型，这种“二三一”向“三二一”的演进趋势通过要素再配置，在压缩传统产业污染排放空间的同时释放了数字经济边际收益递增效应，形成环境质量改善与绿色经济规模扩张的动态协同。据此，本文的研究假设 2 得到验证。

3.3.2 绿色技术创新

由表 4 第（2）列可以看出，试验区设立显著促进了绿色技术创新。此外，相关研究表明绿色技术创新对减污降碳协同增效具有显著促进作用[38]。可能的解释是，试验区通过构建跨部门数据共享平台降低了知识溢出壁垒，系统性整合了污染监测数据与碳排放核算信息，为绿色技术研发提供了精准的应用场景指向与全生命周期验证基础。这种以数据要素驱动技术创新扩散的机制设计，在降低研发试错成本的同时加速了清洁生产工艺的市场渗透率，促使环境治理技术从末端处置向源头控制的范式转移，通过技术效率的提升同步实现了污染减排、碳强度降低与绿色产业附加值扩张的三重目标。据此，本文的研究假设 3 得到验证。

3.3.3 数字普惠金融

由表 4 第（3）列可以看出，试验区设立显著促进了数字普惠金融发展。此外，相关研究表明数字普惠金融对减污降碳协同增效起显著促进作用[28]。可能的解释是，试验区借助数字金融工具的价格信号传导功能，将碳排放强度与环境风险要素嵌入信贷资产定价模型，形成资金供给与环境绩效的动态反馈机制。这种市场激励结构的重塑不仅优化了绿色项目的融资可得性，更重要的是通过引导资本跨期配置推动传统高碳产业的低碳化技改投资，在降低污染治理边际成本的同时催生数字化环保服务的新兴业态，最终形成环境治理公共效益与金融机构私人收益的激励兼容格局。据此，本文的研究假设 4 得到验证。

表4　机制检验结果

变量	(1) 产业结构升级	(2) 绿色技术创新	(3) 数字普惠金融
Bigdata	0.031*** (4.04)	0.027*** (2.66)	0.028** (2.53)
常数项	−0.000 (−0.11)	−0.000 (−0.11)	0.000 (0.11)
控制变量一次项	是	是	是
控制变量二次项	是	是	是
年份固定	是	是	是
城市固定	是	是	是
样本量	3 564	3 564	3 564

4 进一步分析

4.1 异质性分析

4.1.1 新质生产力关注度异质性

在数字经济加速重构经济增长新动能的背景下，发展新质生产力已成为推动高质量发展的战略选择。为探究地方政府对于新质生产力的关注程度是否会对国家大数据综合试验区效能的发挥产生影响，参考陈旭东等[39]的做法，本文使用 Python 工具对各城市政府工作报告中与新质生产力相关的词汇进行文本识别与词频统计，采用新质生产力词频占总词频比重作为各城市新质生产力关注度的衡量指标。按照各年份城市新质生产力关注度中位数将研究样本划分为高、低两组进行异质性回归。由表 5 第（1）列和第（2）列可以看出，高新质生产力关注度城市的试验区设立对减污降碳协同增效存在显著正向影响，而低关注度城市的影响系数不显著。这种分化可从技术创新生态差异视角阐释：高关注度城市依托政府科技支出优先保障、数字基础设施先行完善及数字人才储备体系化，能够通过数据要素市场化配置打通绿色技术供需链条，实现环保约束下经济增长的精准调控；低关注度城市则因研发投入碎片化与应用场景错配，导致数字技术难以有效弥合传统产业升级与碳排放规制间的张力。

4.1.2 资源禀赋异质性

资源型城市产业结构单一，能源消费主要以煤炭、石油等高污染和高排放的化石能源为主，同时面临资源枯竭和人才流失严重等困境，试验区设立对减污降碳协同增效影响可能存在差异。根据《全国资源型城市可持续发展规划（2013—2020 年）》，我们将城市划分为资源型城市和非资源型城市两类。由表 5 第（3）列和第（4）列可以看出，试验区设立对非资源型城市减污降碳协同增效具有显著正向推动效应，而对资源型城市的影响系数为 -0.004 且不显著。可能的解释是，资源型城市因传统能源产业链的“碳锁定”特征，其数字化转型面临深层次结构性矛盾：一方面，试验区政策虽能构建能源消费数据监测系统，但短期内难以突破化石能源技术体系的物质代谢惯性，高碳产业退出阻滞导致清洁技术替代弹性不足[38]；另一方面，资源型城市对矿产开发的路径依赖压缩了数字经济服务业的要素配置空间，削弱了数据驱动型产业对传统重化工业的绿色替代潜力。相较而言，非资源型城市凭借多元化的产业基础和灵活的创新生态系统，能够更高效地将数据要素嵌入绿色产业链重构过程，实现环境治理与经济扩张的动态协同。

4.1.3 城市规模异质性

根据国务院颁布的《关于调整城市规模划分标准的通知》，以市辖区常住人口“100 万人口”作为分界线将我国城市划分为大型城市和中小型城市。由表 5 第（5）列和第（6）列可以看出，试验区设立对大型城市减污降碳协同增效的促进效应显著强于中小型城市。可能的解释是，大型城市凭借更完备的数字基础设施和高端人才集聚效应，能够快速整合政务数据与企业能耗信息，构建全覆盖的减污

表 5 异质性检验结果

变量	(1) 高新质生产力关注度 SRPCE	(2) 低新质生产力关注度 SRPCE	(3) 资源型城市 SRPCE	(4) 非资源型城市 SRPCE	(5) 大型城市 SRPCE	(6) 中小型城市 SRPCE
Bigdata	0.016*** (2.91)	0.000 (0.20)	−0.004 (−0.33)	0.009*** (3.76)	0.012*** (3.19)	0.006** (2.32)
常数项	−0.000 (−0.68)	0.000 (0.12)	0.000 (0.10)	0.000 (1.22)	−0.000 (−0.62)	0.000 (0.50)
控制变量一次项	是	是	是	是	是	是
控制变量二次项	是	是	是	是	是	是
年份固定	是	是	是	是	是	是
城市固定	是	是	是	是	是	是
样本量	1 777	1 787	1 380	2 184	1 260	2 304

降碳协同管控平台，通过边际治理成本递减效应实现环境与经济目标的非线性协同；而中小型城市因财政约束和技术适配能力不足，数字化改革多聚焦于硬件建设层面，难以突破行政区划壁垒与部门数据孤岛，导致政策红利向产业链低碳化改造的渗透效率偏低。

4.2 空间协调分析

当前，尽管中国整体发展进入新阶段，但“省会（自治区首府）明星效应”“南快北慢”等空间分化现象依然显著，缩小区域性环境治理差距与破解“环境—经济”协同困境具有重大现实意义。国家大数据综合试验区作为数字要素市场化改革的制度载体，依托其技术渗透性、数据融合性和治理协同性特点，可通过重构环境信息基础设施、优化跨域资源配置网络，为解决上述难题提供新路径。基于此，本文参考 Griffith 等 [40] 的做法，构建空间协调效应模型，具体设置如下：

$$\frac{\text{SRPCE}_{it}}{\text{SRPCE}_{mt}}=\theta_0\ \text{Bigdata}_{it}+g(X_{it})+\varepsilon_{it} \qquad (9)$$

式中，SRPCE_{it} 为本地区减污降碳协同增效；SRPCE_{mt} 为对应地区减污降碳协同增效集合的最大值；SRPCE_{it} 与 SRPCE_{mt} 的比值为减污降碳协同增效相对差距综合指数，指数值向 1 收敛时，反映地区间协同发展差距缩小；若向 0 趋近，则意味着地区差异极化加剧。

由表 6 可以看出，试验区设立使全国城市层面、南北地区、胡焕庸线、沿海内陆及中心外围的减污降碳协同增效相对差距综合指数显著提升，验证了其空间协调效应。具体可归因于三方面机制：一是技术外溢效应。依托数字平台打破信息壁垒，加速知识与绿色技术的跨区域扩散，缩小了后发地区学习成本。二是要素重构优势。通过智能监测与数据互联优化资源跨域配置效率，化解传统发展中的“中心—外围”虹吸矛盾。三是协同网络构建。数字治理模式推动形成跨区域联动机制，促进生态治理与经济增长的均衡协同。

表 6　空间协调效应

变量	(1) 城市相对差距	(2) 南北相对差距	(3) 胡焕庸线相对差距	(4) 沿海内陆相对差距	(5) 中心外围相对差距
Bigdata	0.028*** (3.83)	0.039*** (5.06)	0.031*** (3.97)	0.030*** (4.04)	0.028*** (3.72)
常数项	0.000 (0.19)	0.000 (0.44)	0.000 (1.19)	−0.000 (−0.38)	0.000 (0.37)
控制变量一次项	是	是	是	是	是
控制变量二次项	是	是	是	是	是
年份固定	是	是	是	是	是
城市固定	是	是	是	是	是
样本量	3 564	3 564	3 564	3 564	3 564

5 结论与政策建议

本文基于 2011—2022 年中国 297 个地级及以上城市面板数据，以国家大数据综合试验区设立为准自然实验，构建耦合协调度模型测算减污降碳协同增效水平，并采用双重机器学习模型识别政策因果效应。主要结论如下：①试验区设立显著促进减污降碳协同增效，在进行一系列稳健性检验后该结论依然成立；②作用机制检验表明，试验区通过“产业结构升级—绿色技术创新—数字普惠金融”三链协同的传导路径实现环境、气候与经济目标的统筹优化；③异质性分析表明，试验区的减污降碳协同增效效应在新质生产力关注度高、非资源型及大型城市中更为显著；④空间形态分析表明，试验区设立能够在全国和地区层面缩小减污降碳协同增效的相对发展差距。

基于以上结论，本文提出以下政策建议：①强化顶层设计，构建“数据要素 × 减污降碳”国家级标准体系。以国家大数据综合试验区为“压力测试场”，尽快制定并动态迭代《城市级减污降碳数据要素流通技术规范》这一强制性国家标准，破解当前跨域碳排放因子不统一、数据接口碎片化难题。完善基于监测数据与评估结果的污染减排、碳配额分配和交易政策的即时反馈与调整机制，形成“监测—治理—评估—调整”的闭环管理。②深化三大核心机制联动，构建绿色转型长效路径。推动数字经济与传统产业融合，建设工业互联网平台，提供“云—边—端”一体化节能降碳方案。加强绿色技术创新，布局

开源平台与数字孪生试验场，设立“卡脖子”技术研发基金。构建嵌入 ESG 因子与碳足迹数据的普惠金融产品体系与个人碳账户激励机制，引导资源优化配置，将减污降碳目标转化为经济效益。③实施差异化区域策略，释放城市禀赋优势与治理潜能。资源型城市需破解资源“碳锁定”问题，将数字技术应用于传统能源企业清洁生产改造、伴生资源利用及接续产业培育。非资源型城市需深化大数据应用，打造低碳转型标杆。高新质生产力关注度城市可加大对“工业元宇宙”等前沿领域的投入，构建减污降碳评估体系。大型城市需拓展数字孪生技术应用，培育高端绿色产业集群；中小城市可推动区域协同，弥补技术短板。

参考文献

[1] 孙伟增，毛宁，兰峰，等. 政策赋能、数字生态与企业数字化转型——基于国家大数据综合试验区的准自然实验 [J]. 中国工业经济，2023(9): 117-135.

[2] XUE H, CAI M, LIU B L, et al. Sustainable development through digital innovation: Unveiling the impact of big data comprehensive experimental zones on energy utilization efficiency[J]. Sustainable development, 2025, 33(1): 177-189.

[3] WU D S, XIE Y, LYU S. Disentangling the complex impacts of urban digital transformation and environmental pollution: Evidence from smart city pilots in China[J]. Sustainable cities and society, 2023, 88: 104266.

[4] 姜祎，张庆红. 协同推进降碳减污扩绿增长路径研究——基于国家大数据综合试验区视角 [J]. 技术经济与管理研究，2025(7): 15-21.

[5] SONG T, DIAN J, CHEN H W. Can smart city construction improve carbon productivity?—A quasi-natural experiment based on China's smart city pilot[J]. Sustainable cities and society, 2023, 92: 104478.

[6] LIU M G, LI G Y. Research on the impact of the digital economy on carbon pollution based on the national big data comprehensive pilot zone in China[J]. Sustainability, 2023, 15(21): 15390.

[7] 王琴，郭莉滨，丁可可. 数字基础设施建设与城市减污降碳协同——基于准自然实验的研究 [J]. 城市发展研究，2025, 32(6): 100-106.

[8] ZHANG X Q. Broadband and economic growth in China: an empirical study during the COVID-19 pandemic period[J]. Telematics and informatics, 2021, 58: 101533.

[9] DUYGAN M, FISCHER M, PÄRLI R, et al. Where do smart cities grow? The spatial and socio-economic configurations of smart city development[J]. Sustainable cities and society, 2022, 77: 103578.

[10] 秦文晋，刘鑫鹏. 网络基础设施建设对数字经济发展的影响研究——基于“宽带中国”试点政策的准自然实验 [J]. 经济问题探索，2022(3): 15-30.

[11] 张涛，李均超. 网络基础设施、包容性绿色增长与地区差距——基于双重机器学习的因果推断 [J]. 数量经济技术经济研究，2023, 40(4): 113-135.

[12] 张修凡，范德成. 数字经济发展赋能我国低碳经济转型研究——基于国家级大数据综合试验区的分析 [J]. 科技进步与对策，2023, 40(19): 118-128.

[13] 陈洪波，张骁潇. 国家大数据综合试验区如何促进城市绿色转型：基于碳排放量增速和碳生产率增幅的双重视角 [J]. 中国软科学，2024(11): 172-188.

[14] 杨玉琪，王小华. 数字化赋能与城市经济高质量发展——基于国家级大数据综合试验区的准自然实验 [J]. 经济问题探索，2023(12): 105-123.

[15] WEI X H, JIANG F, YANG L S. Does digital dividend matter in China's green low-carbon development: environmental impact assessment of the big data comprehensive pilot zones policy[J]. Environmental impact assessment review, 2023, 101: 107143.

[16] WU Y C, DU H F, WEI X Y, et al. Big data development and agricultural carbon emissions: Exacerbation or suppression? A quasi-natural experiment based on the establishment of the National Big Data Comprehensive Pilot Zone[J]. Journal of environmental management, 2024, 368: 122178.

[17] SHEN X B, WANG Z C. Can digital industrialization promote energy conservation development in China? Empirical evidence based on national big data comprehensive pilot zone policy[J]. Journal of environmental management, 2024, 368: 122125.

[18] 常皓亮，金碚，薛飞. 大数据战略对电力消费碳排放的影响——基于国家级大数据综合试验区的准自然实验 [J]. 经济与管理研究，2023, 44(5): 93-109.

[19] 王智勇，杨金娇. 国家级大数据综合试验区的就业效应研究 [J]. 劳动经济研究，2024, 12(2): 95-118.

[20] 沈坤荣，乔刚，谭睿鹏. 国家级大数据综合试验区设立与就业增长 [J]. 中国工业经济，2024(12): 5-23.

[21] 郭凯明，龚六堂. 减税降费与产业结构变迁 [J]. 经济学（季刊），2023, 23(5): 1723-1740.

[22] 巩红禹，万艳，曹昱亮. 数字经济赋能减污降碳效应——基于国家级大数据综合试验区的准自然实验 [J]. 调研世界，2025(1): 29-41.

[23] 严成樑，赵扶扬，牛欢. 环境目标责任制、环境治理与内生经济增长 [J]. 经济研究，2024, 59(4): 133-152.

[24] 唐金环，吴宛亭，陈曦. 区块链技术在碳交易体系中的应用研究 [J]. 中国环境管理，2023, 15(5): 78-86.

[25] 刘敏楼，黄旭，孙俊．数字金融对绿色发展的影响机制 [J]. 中国人口·资源与环境，2022, 32(6): 113-122.

[26] 王志刚，项猛．数实融合的路径分析：基于大数据综合试验区的准自然实验 [J]. 经济问题探索，2024(9): 95-112.

[27] XING L, HAN D H, HUI X. The impact of carbon policy on corporate risk-taking with a double/debiased machine learning based difference-in-differences approach[J]. Finance research letters, 2023, 58: 104502.

[28] 赵燕，孙慧，马点圆．城市绿色金融改革对减污降碳协同增效的影响研究——基于双重机器学习的因果推断 [J]. 产业经济研究，2024(3): 15-28, 58-58.

[29] 丛晓男．耦合度模型的形式、性质及在地理学中的若干误用 [J]. 经济地理，2019, 39(4): 18-25.

[30] 谢思，房克雷．国家大数据综合试验区对城市环境污染的影响及机制——基于双重机器学习的因果推断 [J]. 资源科学，2024, 46(7): 1433-1445.

[31] 郭峰，王靖一，王芳，等．测度中国数字普惠金融发展：指数编制与空间特征 [J]. 经济学 (季刊), 2020, 19(4): 1401-1418.

[32] ZENG J T, YANG M R. Digital technology and carbon emissions: evidence from China[J]. Journal of cleaner production, 2023, 430: 139765.

[33] ZHONG W L, LIU Y, DONG K Y, et al. Assessing the synergistic effects of artificial intelligence on pollutant and carbon emission mitigation in China[J]. Energy economics, 2024, 138: 107829.

[34] CHERNOZHUKOV V, CHETVERIKOV D, DEMIRER M, et al. Double/debiased machine learning for treatment and structural parameters[J]. The econometrics journal, 2018, 21(1): C1-C68.

[35] NUNN N, QIAN N. US food aid and civil conflict[J]. American economic review, 2014, 104(6): 1630-1666.

[36] 杨思莹，白桦，李政．供应链金融提高了企业供应链效率吗——基于双重机器学习的因果推断 [J]. 财经科学，2025(2): 57-73.

[37] 杨晓军，薛洪畅．创新驱动政策是否促进城市减污降碳协同增效？——来自国家创新型城市试点政策的经验证据 [J]. 产业经济研究，2024(3): 1-14.

[38] 宋德勇，陈梁，王班班．环境权益交易如何实现减污降碳协同增效：理论与经验证据 [J]. 数量经济技术经济研究，2024, 41(2): 171-192.

[39] 陈旭东，沈利芸．数字经济、财政激励与地区税收竞争力——基于双重机器学习的因果推断 [J]. 中央财经大学学报，2025(2): 21-37.

[40] GRIFFITH R, REDDING S, VAN REENEN J. Mapping the two faces of R&D: productivity growth in a panel of OECD industries[J]. Review of economics and statistics, 2004, 86(4): 883-895.

Does the Establishment of National Big Data Comprehensive Pilot Zone Promote Synergistic Reduction of Pollution and Carbon Emissions? —A Causal Inference Analysis Based on Double Machine Learning

ZHANG Hongchang[1,2], WANG Yue[2], WANG Weimei[2*], CHEN Yu[3]

(1. Guizhou Green Development Strategy High-End Think Tank, Guizhou University of Finance and Economics, Guiyang 550025, China; 2. School of Applied Economics, Guizhou University of Finance and Economics, Guiyang 550025, China; 3. The School of Accounting, Guizhou University of Finance and Economics, Guiyang 550025, China)

Abstract: Synergizing the reduction of pollution and carbon emissions represents an important goal in achieving China's ecological civilization and high-quality economic development. As pioneering pilot zones for market-oriented reform of data elements, it remains to be explored whether the policy effects of the national big data comprehensive pilot zones can solve the synergistic dilemma among pollution abatement, carbon emission intensity reduction and economic growth. Based on the panel data of 297 prefecture-level and above cities in China from 2011 to 2022, this study treats the establishment of the national big data comprehensive pilot zone (hereinafter referred to as the pilot zone) as a quasi-natural experiment and employs a double machine learning model to investigate the causal effects of the pilot zone on the reduction of pollution and carbon emissions. The study finds that: ① the establishment of the pilot zone significantly promotes the synergistic reduction of pollution and carbon emissions, and this conclusion remains valid after endogeneity and robustness tests; ② The establishment of the pilot zone promotes synergizing the reduction of pollution and carbon emissions through a three-chain synergistic mechanism of industrial structure upgrading, green technology innovation and digital inclusive finance development; ③ Heterogeneity analysis shows that the establishment of the pilot zone is more effective in high attention to new quality productvity ctics non-resource-based cities, and large cities; ④ The establishment of pilot zones has a spatial coordination effect in promoting the synergistic reduction of pollution and carbon emissions, narrowing the relative development gap at both the national and regional levels. This study provides new arguments for integrating the synergistic paths of pollution management, carbon reduction and economic growth in the era of digital economy, and provides policy insights for differentiating the construction of a data-enabled environmental governance system and promoting a comprehensive green and low-carbon transformation of the economy and society.

Keywords: national big data comprehensive pilot zone; synergistic reduction of pollution and carbon emissions; dual machine learning; spatial coordination