

# 公共数据开放对企业“漂绿”行为的影响研究\*

## ——基于双重机器学习的因果推断

周 洋 方 恺 袁书强 李浩伍

**摘要:**公共数据开放是数字化治理的重要制度安排,为抑制企业“漂绿”行为提供了新的治理路径。本文基于城市公共数据开放平台上线的准自然实验,以2009—2022年沪深A股上市公司为研究样本,运用双重机器学习方法系统评估公共数据开放对企业“漂绿”行为的治理效应。研究结果表明,公共数据开放能够显著抑制企业的“漂绿”行为。机制分析显示,公共数据开放通过增大规制性压力、强化规范性压力和加剧模仿性压力三条路径形成对企业的持续约束,从而抑制企业“漂绿”行为。异质性分析显示,公共数据开放的治理效应具有显著的对象差异与制度环境依赖特征,其抑制作用在重污染行业、国有企业和短期主义倾向较强的企业中更为明显,并在市场化程度较高的地区表现得更为显著。本研究为通过公共数据开放提升环境治理效能和助力经济高质量发展提供了重要的经验证据。

**关键词:**公共数据开放;企业“漂绿”;双重机器学习;制度压力;环境治理

### 一、引言与文献综述

2025年《政府工作报告》提出要“完善支持绿色低碳发展的政策和标准体系,营造绿色低碳产业健康发展生态”。随着高质量发展的推动和“双碳”战略的推进,绿色转型的经济社会效应受到广泛关注(方恺等,2023),监管部门与市场也更加重视企业环保行为,对企业ESG(环境、社会与公司治理)报告的规范性与完整性要求也不断增强。然而,由于我国目前监管机制与披露标准尚未完善,且企业ESG信息多为自主披露,企业对于披露的广度、深度和披露

\*周洋,中共浙江省委党校“八八战略”创新发展研究院,邮政编码:311121,电子信箱:m19157787669@163.com;方恺(通讯作者),浙江大学公共管理学院,邮政编码:310058,电子信箱:fangk@zju.edu.cn;袁书强,中共浙江省委党校,邮政编码:311121,电子信箱:ys\_qiang@126.com;李浩伍,新加坡国立大学李光耀公共政策学院,邮政编码:259772,电子信箱:e1553681@u.nus.edu。

本文系国家自然科学基金重点项目“面向可持续发展目标的全球气候治理风险与韧性提升策略研究”(72534005)、国家重点研发计划项目“排放低碳治理关键技术体系”(2024YFC3712900)的阶段性成果。感谢匿名审稿专家和编辑部的宝贵修改意见。文责自负。

方式拥有较大的决定权。企业在环境信息披露与真实环保行动之间存在较大的操作空间。部分企业以环保宣传为营销工具,实际却通过模糊陈述、选择性呈现等方式粉饰其环境表现,构成“漂绿”行为(李大元等,2015)。从本质上看,企业“漂绿”是一种在信息不对称条件下产生的策略性披露行为(黄溶冰等,2020),其核心动机在于掩饰企业真实的环境绩效,并借此塑造更为正面的绿色形象以影响外部评价、获取资源与制度优势(肖红军等,2013)。“漂绿”行为使得利益相关方难以准确判断企业环境责任的履行情况,并且影响绿色转型的整体效率。当前我国正处于绿色转型向高质量发展的关键时期,“漂绿”不再仅是企业层面的信息披露偏差,而是会实质性影响企业行为与绿色发展绩效的重要制度性问题。一方面,同行企业“漂绿”会扰乱市场对企业真实环境绩效的识别机制,削弱优质企业的竞争优势,挤出企业在绿色技术创新上的资源投入与风险承担,最终抑制绿色创新产出(蔡真等,2025)。另一方面,企业实施“漂绿”行为会扭曲资本市场对其可持续能力的判断,减少耐心资本供给并削弱企业的财务柔性与长期调整能力(吕怀立等,2025)。因此,如何有效治理企业“漂绿”行为,对于规范企业环境信息披露、完善环境治理体系以及推动经济高质量发展具有重要的现实意义。

围绕“漂绿”行为的治理,现有研究从不同治理逻辑出发,形成了两类具有代表性的研究路径。第一类研究强调企业内部行为在规范环境信息披露的真实性与一致性中的作用。研究发现企业有效的内部控制和较高的数字化水平可缓解企业在ESG报告中采取策略性“漂绿”的倾向(李常青、辛立柱,2025;陈琪、李梦函,2025)。同时,高管团队国际化可通过降低管理层短视倾向、提升环境信息透明度并缓解融资约束,弱化企业实施“漂绿”的动机(李强、蒋洮,2023)。再者,高管激励机制能够通过薪酬和股权激励计划提升高管的工作积极性和企业主人翁意识,从而使管理者与企业的目标趋于一致,做出提高企业信息披露质量的决策,抑制管理层的“漂绿”行为(魏宇琦、高锦萍,2026)。这一路径从产权结构、管理层特征与激励制度层面揭示了企业内部治理在约束“漂绿”行为中的重要作用,凸显了通过改善内部制度安排规范环境信息披露的现实意义。然而,这类治理机制主要作用于企业内部决策过程,其有效性高度依赖企业自身制度基础,难以改变环境信息由企业单向生产的格局。在缺乏外部可持续核验机制的情况下,仅依赖内部治理难以从根本上缓解信息不对称,也难以支撑对“漂绿”行为进行长期和稳定的治理。

第二类研究立足外部监督对企业“漂绿”行为的治理作用。首先,政府监管被认为能够通过提高违规成本与合规压力,抑制企业环境信息披露中的选择性偏差(杨永聪、李学轩,2025)。并且,监管垂直化(金暄暄等,2025)以及相关环境法规的实行(李四海、马文琪,2025;石建勋、辛沛远,2025)均能显著抑制企业“漂绿”。其次,机构投资者实地调研与退出威胁被证实能够通过降低信息不对称、改善内部控制并约束管理层与控股股东的自利动机,对企业“漂绿”行为形成持续压力(李强等,2023;刘兴华等,2025)。另外,媒体监督可通过信息揭示与声誉机制,对企业“漂绿”行为形成外部压力(陶云清等,2024),从而推动企业真正履行环境

责任。然而,这类外部监督方式本质上仍建立在个案识别与外部介入基础之上,其治理效果高度依赖监管资源投入与监督主体的主动发现,运行逻辑具有明显的事后性与选择性(Zhang & Fang, 2019)。在信息高度分散、核验成本较高且监管对象广泛的现实情境下,仅依赖传统外部监督和单一评价体系,难以充分揭示企业的环境责任转移(Zhang & Fang, 2019),也难以从制度层面构建起稳定、低成本且可持续的“漂绿”治理机制,且容易被企业通过策略性调整行为加以消解(Huang et al., 2025)。

随着数字治理理念的深入推进,信息治理方式逐步从依赖人工识别监管转向依托数据基础设施的透明监督与过程可见模式。数字化手段使企业能够系统整合分散在生产与管理环节的能耗、排放、项目建设等数据,提高信息采集披露的效率与准确性(李恩极等, 2025)。综观当前文献,已有研究表明公共数据首先能打破信息孤岛,优化信息传递效率(于小悦、郑晓慧, 2026),并且能在区域治理能力提升(方锦程等, 2023)、绿色全要素生产率提升(郭家堂, 2025)及城市绿色发展(汪克亮、付丽翔, 2025; Wu et al., 2026)等方面发挥积极作用。这些研究从宏观与区域层面表明,公共数据开放通过改善信息供给方式与治理工具结构,正在成为数字治理体系中的重要制度基础。然而,尽管现有文献证实公共数据开放能促进企业技术创新(沈坤荣、林剑威, 2025)、提升供应链韧性(乔子轶等, 2025)以及降低制造业企业的污染排放强度(陈诗一等, 2025),但其是否能有效约束环境信息操纵行为,尤其是对“漂绿”这一典型策略性披露行为产生治理作用,尚缺乏系统而直接的经验证据。从理论上,公共数据开放通过将企业排污信息、环保处罚记录、信用评价等关键数据结构化、规范化地公开,有助于降低信息搜集成本、增强监督的一致性和可验证性(Du et al., 2024)。在这种制度化信息供给模式下,企业环境行为更容易被持续跟踪与横向比较,“漂绿”行为的隐蔽性随之下降,其被识别、放大与追责的概率显著提高,由此可推动企业面临更为持续的约束。

基于此,本文选取2009—2022年沪深A股上市公司为研究样本,系统评估公共数据开放对企业“漂绿”行为的治理效应,并从制度理论出发揭示其作用路径。本文可能的边际贡献包括:(1)在研究主题上,本文为公共数据开放治理企业“漂绿”行为提供了系统的理论分析与经验证据,拓展了企业“漂绿”治理的研究边界。现有研究从企业内部治理与外部监督两个维度分别探讨企业“漂绿”的形成机理与治理路径。但对于公共数据开放这一数字化治理工具在企业“漂绿”治理中的作用关注不足。本文借助地方政府公共数据开放平台上线这一准自然实验,系统识别并验证了公共数据开放对企业“漂绿”行为的治理效应,为理解公共数据开放的环境治理功能提供了新的经验证据。(2)在研究视角上,本文从数据要素治理与企业环境信息披露行为相结合的微观视角切入,深化了公共数据开放政策效果的企业层面评估。既有研究已从宏观与区域层面对公共数据开放的经济与治理效应进行了较为充分的讨论,但对企业层面的效果评估则仍有待进一步深化。因此,本文将研究重心下沉至企业层面,聚焦公共数

据开放对企业“漂绿”的影响,从而拓展和深化了公共数据开放研究的治理内涵与应用场景。(3)在研究方法与理论解释上,本文引入双重机器学习作为主要识别手段,在高维控制变量与非线性条件下实现对公共数据开放治理企业“漂绿”效应的稳健因果识别。并基于制度理论,从规制性、规范性与模仿性压力三条路径系统刻画其治理机制,丰富了数字化背景下治理企业“漂绿”行为的理论解释。此外,本文的研究结论为政府在数字化转型背景下通过优化公共数据供给、提升环境治理效能提供了有益参考。

## 二、理论分析与研究假说

### (一)公共数据开放治理企业“漂绿”的理论基础

企业“漂绿”行为之所以能够长期存在,既源于绿色市场中的环境信息不对称(罗崇佳、段茂盛,2025),也与环境信息长期缺乏独立、可核验的第三方事实基准密切相关。当前,资本市场与公众对企业绿色表现的判断,长期主要依赖外部中介机构发布的ESG评级。短期来看,评级体系在一定程度上提升了信息透明度,降低了监督成本,并促进了内部治理与外部约束的联动(黄宏斌等,2025)。较高评级能够向市场传递企业负责任与可持续发展的积极信号,有助于企业获得利益相关者认同(蒋艺翹、姚树洁,2024),投资者也往往对ESG表现优异的公司要求更低的收益率(史永东、王淏森,2023),从而缓解融资约束(于庆瑞等,2025),并推动企业绿色转型(胡洁等,2023)。然而,在缺乏可验证事实基准的情形下,这种以评级为核心的信息体系,强化的是企业的“披露动机”,而非“履责动机”。在当前监管体系不健全、治理机制薄弱的场景下,企业绿色宣传与真实环保行动之间的偏离难以及时被识别与纠正,“漂绿”企业可以获得与“真绿”企业相近的声誉回报、融资便利与市场关注。所以企业在权衡成本与收益后,更容易选择以形式化披露或象征性举措替代实质性改进,通过较小投入换取外部认同(肖红军等,2013)。

相较于第三方评级机构主要依赖企业自报数据,公共数据来源于政府履职过程与社会运行过程的持续积累,天然具有非排他性、非竞争性与显著正外部性(方锦程等,2023)。公共数据开放则是以平台化方式,将上述分散于不同部门和系统的数据整合为面向各类社会主体的制度化数据供给体系(Nagaraj,2022)。公共数据中涵盖了环境治理中大量与企业行为直接相关的基础信息,其系统性开放,使社会主体能够通过统一平台和标准化接口,持续低成本地接入权威数据资源,扩大了基础信息要素的可及范围。公共数据开放不断缓解长期存在的信息不对称与信息碎片化问题,使政府、企业与公众能够在更为接近的认知基础上参与社会运行与公共事务。并且,依托于公共数据开放平台,社会主体获取信息的方式不再限于零散渠道,也突破了部门壁垒和封闭系统的制约,转而凭借可持续供给的公共数据体系开展分析、验证与应用,在制度层面构建起独立、可信且可验证的环境事实基准。这一事实基准的形成,使企业环境绩效不再完全依赖其自身表述,而是进入一个可以被持续比对、核验与追溯的制度化

信息环境,从而为长期内企业内部与外部行为约束的重塑提供了基础条件,从根本上压缩了企业操控披露信息以误导市场的空间。

从企业内部看,公共数据开放拓宽了信息核验渠道,改善企业内部绿色治理的决策环境(王海等,2024)。一方面,公共数据开放通过提高环境信息的可得性和透明度,使得企业“漂绿”行为更易被识别(Wu et al., 2020)。当企业排放、能耗与环保投入等数据在平台上实现可视化与长期留痕后,企业在ESG报告中的披露内容可以被直接与公共数据进行比对、核查与追责,虚假披露与选择性披露的预期监管风险与声誉风险显著上升,抑制了管理层操控披露信息的动机,使企业更倾向于提升真实环境表现(Qiang & Ruan, 2025)。另一方面,强制性环境信息披露制度的推行进一步强化了企业的数据治理要求,促使企业通过信息系统整合与分析各环节能耗与排放数据,提升环境信息的科学性 with 可追溯性(何文剑等, 2025)。从企业外部看,公共数据开放显著降低了环境信息核验门槛,重塑了外部监督的运行方式(沈斌、黎江虹, 2023)。“漂绿”行为之所以具有较强隐蔽性和反复性,根源在于相关信息分散于不同主体和渠道之中,社会监督难以获取稳定证据并形成持续扩散机制(王伟、刘传红, 2013)。随着公共数据开放平台的建设与运行,原本分散于不同部门和系统中的环境信息被持续整合与公开,使媒体、投资者与公众能够基于同一事实基础,对企业披露信息与平台数据进行持续比对,并据此开展基于证据的监督行动(张晓娟等, 2024)。此时,外部监督不再主要依赖企业自愿披露和零散线索,而逐步转向基于客观数据的常态化核验机制。信息环境的转变使“漂绿”行为更易被识别与曝光,公众与市场的评判标准也由企业主导的叙事表达,转向可验证的数据事实,从而倒逼企业通过真实履行环境责任来维系声誉、保持投资者信任并降低未来被追责风险。

综合上述分析,公共数据开放通过构建可验证的第三方事实基准,系统性破坏了企业“漂绿”所依赖的信息模糊空间。这一事实基准同时嵌入企业内部治理与外部监督过程,从而在机制上提高了“漂绿”行为的识别概率与问责风险,并对其形成持续约束。基于此,提出如下研究假说:

H1:公共数据开放能抑制企业“漂绿”行为。

## (二)公共数据开放约束企业“漂绿”的机制分析

在公共数据治理持续深化的背景下,公共数据开放不仅规范了环境信息的披露,更从根本上重构了企业所处的信息结构(解子恒、靳添全, 2026),改变了环境信息的获取方式(谢佳松等, 2025)、核验方式(杨志安等, 2026)与比较方式(金环、任仕佳, 2026)。与传统依赖零散披露和人工核查的治理模式不同,公共数据开放将分散于不同部门和系统中的企业排放、能耗与环境监管信息加以整合,使环境信息逐步具备可持续获取、可交叉核验与可横向比较的制度属性(沈坤荣、林剑威, 2025),并转化为稳定的公共事实基础。正是这种信息结构的转变,使企业环境行为逐步被纳入一个可以被持续识别、核验与比较的制度化信息环境之中(邹祥、贾永飞, 2025)。随着环境信息的可得性、可核验性与可比较性同步提升,规制执行、社会

监督与行业参照的运行基础发生改变,并由此激活和强化规制性、规范性与模仿性三类制度压力,进而系统性抑制企业“漂绿”行为。

公共数据开放通过提升信息可得性增大规制性压力。规制性压力源于制度规则是否具备可执行性,而可执行性的前提正是信息的可得性与可追溯性(汪立鑫、孙圣涛,2026)。在信息分散与核验困难的情形下,即便存在监管规则,执法往往面临取证成本高、识别难度大和选择性执法等约束,制度威慑力因此受到削弱。公共数据开放通过整合企业排放、能耗与监管记录,提高了环境信息的集中度与可追溯性,使监管部门能够更低成本地识别披露偏差与潜在违规行为,从而显著提升规制约束的可执行性(郭家堂,2025)。当企业预期监管识别概率与问责确定性上升时,其对“漂绿”边际收益与潜在成本的权衡将发生变化,更倾向于减少策略性披露与模糊陈述(程秋旺等,2025)。据此,本文提出如下假说:

H2a:公共数据开放通过增大企业面临的规制性压力,从而抑制企业“漂绿”行为。

公共数据开放通过提升信息可核验性强化规范性压力。规范性压力并不直接来自法律制裁,而依赖于社会评价机制是否能够形成稳定、可扩散的认知(习明明等,2025)。传统情境下,社会监督高度依赖企业自愿披露和媒体调查,受限于信息碎片化与事实核验困难,往往呈现出零散性与滞后性。公共数据开放为媒体、投资者与公众提供了稳定、可核验的数据来源,使企业披露内容能够被持续比对和验证,从而显著提升社会监督的事实基础与持续性(汤小莉等,2025)。当企业“漂绿”行为更易被揭示并引发负面评价时,其所面临的不再是偶发性曝光风险,而是持续存在的声誉约束压力(宫绮等,2025)。在这一机制下,企业为维持合法性认同与社会信任,更倾向于提升环境履责与披露一致性。据此,本文提出如下假说:

H2b:公共数据开放通过强化企业面临的规范性压力,从而抑制企业“漂绿”行为。

公共数据开放通过提升信息可比较性加剧模仿性压力。模仿性压力源于组织在不确定环境下对行业标杆行为的学习与跟随(Lieberman & Asaba,2006),其发生前提是行业中存在可识别、可比较且具有稳定参照意义的行为坐标。信息透明度不足时,企业对同行真实行为的认知有限,模仿往往停留在形式层面。公共数据开放提升了企业环境表现与披露取向的可观察性与可比较性,使行业内部更易形成稳定参照对象与比较框架(陆雪艳等,2025)。在横向比较强化的情境下,企业为避免在相对评价中处于不利位置,将更倾向于调整披露策略与实际行为以趋近行业规范(韩一鸣等,2025),从而压缩“漂绿”空间。据此,本文提出如下假说:

H2c:公共数据开放通过加剧企业之间的模仿性压力,从而抑制企业“漂绿”行为。

### 三、研究策略

#### (一)模型构建

基于上述理论分析,本文参考Chernozhukov等(2018)提出的方法框架,使用双重机器学习

习检验公共数据开放治理企业“漂绿”行为的有效性。相比于传统的双重差分方法,双重机器学习能够在有限样本下,通过正交化残差与交叉拟合,有效控制高位非线性混淆并获得稳健的处置效应估计(张涛、李均超,2023)。据此,本文构建部分线性双重机器学习模型如下:

$$GWS_{it} = \theta_0 D_{it} + g(X_{it}) + U_{it} \quad (1)$$

$$E[U_{it} | D_{it}, X_{it}] = 0 \quad (2)$$

其中,  $GWS_{it}$  表示企业  $i$  在年份  $t$  的“漂绿”程度;  $D_{it}$  表示企业所在城市是否为公共数据开放试点城市,  $\theta_0$  为政策对应的平均处置效应。  $X_{it}$  为控制变量集合高维特征,具体的函数形式  $g(X_{it})$  将由机器学习方法进行非参数估计,  $U_{it}$  为误差项。

若直接用机器学习对  $g(X_{it})$  回归并估计  $\theta_0$ , 在有限样本下可能因正则化而偏误。为此,构造对照方程并进行正交化处理:

$$D_{it} = m(X_{it}) + V_{it} \quad (3)$$

$$E[V_{it} | X_{it}] = 0 \quad (4)$$

利用机器学习估计  $\hat{m}(X_{it})$  与  $\hat{g}(X_{it})$ , 据此得到残差  $\hat{V}_{it} = D_{it} - \hat{m}(X_{it})$  和  $\hat{U}_{it} = GWS_{it} - \hat{g}(X_{it})$ , 随后进行正交回归, 得到  $\theta_0$  的双重机器学习估计量:

$$\hat{\theta}_0 = (\frac{1}{n} \sum_{i,t} \hat{V}_{it}^2)^{-1} (\frac{1}{n} \sum_{i,t} \hat{V}_{it} \hat{U}_{it}) \quad (5)$$

式(5)的收敛性质取决于  $\hat{m}(X_{it})$  与  $\hat{g}(X_{it})$  对真值的收敛速度。通过两次机器学习和正交残差与交叉拟合,一方面剔除了高维变量  $X_{it}$  对  $D_{it}$  的内生干扰,另一方面提升了  $\hat{\theta}_0$  的有限样本估计的准确性。另外,为提升估计稳定性,本文采用5折分割拟合的方法处理回归样本。

## (二)变量设定

### 1.被解释变量:企业“漂绿”程度

“漂绿”指企业在信息披露中刻意强调其绿色形象,通过选择性报告环境信息来展示可持续发展承诺,但实际行动与披露内容并不匹配的现象。本质上,“漂绿”反映了企业在“言行之间”的偏差,其识别关键在于如何衡量环境信息披露水平与真实环境绩效之间的差距。借鉴Hu等(2023)学者的研究方法,本文依据企业环境信息披露得分与环境绩效得分的相对偏离程度构建企业“漂绿”指数,以刻画企业在披露水平与真实绩效两者间的脱节程度。具体计算公式如下:

$$GWS_{i,t} = \left( \frac{ER_{dis,i,t} - \overline{ER}_{dis}}{\sigma_{dis}} \right) - \left( \frac{ER_{per,i,t} - \overline{ER}_{per}}{\sigma_{per}} \right) \quad (6)$$

其中,  $ER_{dis,i,t}$  和  $\overline{ER}_{dis}$  分别代表第三方评级机构彭博评定的第  $i$  家企业环境披露得分和同行业其他企业环境披露得分的均值。  $ER_{per,i,t}$  和  $\overline{ER}_{per}$  分别表示第三方评级机构华证评定的第  $i$

家企业环境绩效得分和同行业其他企业环境绩效得分的均值。 $\sigma_{dis}$  和  $\sigma_{per}$  分别表示环境披露和环境绩效的标准差。通过式(6)将企业“漂绿”指数转化为企业在环境信息披露评分相对地位 and 实际环境绩效之间的标准化差值。该差异越大,说明企业披露的环境表现相对于实际绩效有越多的掩饰成分,即“漂绿”程度越严重。

2.解释变量:公共数据开放

地方政府推动公共数据开放平台的建设与上线,为本文提供了一个具有制度冲击特征的准自然实验场景。本文所指公共数据开放,是指由地方政府统一建设并运行的一体化公共数据开放平台,其核心特征在于整合政府部门依法记录与监管形成的独立于企业披露意愿的数据资源,并以标准化方式向社会开放。平台上线意味着地方政府在数据供给方式、信息开放机制与治理工具层面的制度性调整,实质性改变了企业所处的信息环境,具有明确的制度冲击属性。从平台上线机制来看,公共数据开放平台的建设并非针对特定企业或短期环境绩效的政策选择,而是嵌入国家数字政府建设与数据要素治理的整体推进框架之中。其推进节奏通常受上级政策部署、行政程序与地方信息化基础条件等因素共同约束,具有明显的阶段性与程序性特征,难以由地方政府基于当期企业环境披露状况进行短期精确选择。同时,平台上线作为城市层面的制度安排,反映了自上而下的制度推进与技术性落地过程,单个企业对上线决策的反向影响空间有限。平台上线后,原本主要在政府系统内部流转的企业环境信息,转变为公众、媒体、投资者和第三方机构可低成本获取的公共信息,从而显著提升了企业环境行为的可观察性与可核验性。因此,基于测度口径的一致性,参考方锦程等(2023)的界定,本文仅纳入由地方政府牵头、汇聚多部门数据资源的一体化开放平台,并基于城市层面的公共数据开放平台上线时点,构造二元虚拟变量  $DID_{it}$ ,若企业  $i$  所在城市在年份  $t$  已正式上线公共数据开放平台,则该变量在当年及之后取 1,否则取 0。

3.控制变量

为进一步控制可能影响企业“漂绿”行为的其他因素,本文设定如下控制变量:资产报酬率( $Roa$ )、总负债率( $TI$ )、无形资产比重( $Itang$ )、企业规模( $Size$ )、企业绩效( $Tobin$ )、企业年龄( $Lnage$ )、资产增长率( $Tagr$ )。另外,本文还进一步控制年份与企业固定效应。具体变量定义及描述如表 1 所示。

| 表 1   |          | 变量定义及描述 |                                                   |
|-------|----------|---------|---------------------------------------------------|
| 变量类型  | 变量名称     | 变量符号    | 测量方法                                              |
| 被解释变量 | 企业“漂绿”程度 | $GWS$   | 参考 Hu 等(2023)的方法,采用式(6)进行测算                       |
| 解释变量  | 公共数据开放   | $DID$   | 若企业 $i$ 所在城市在年份 $t$ 已正式上线公共数据开放平台,则当年及之后取 1,否则取 0 |

| 续表 1 |        | 变量定义及描述      |                |  |
|------|--------|--------------|----------------|--|
| 变量类型 | 变量名称   | 变量符号         | 测量方法           |  |
| 控制变量 | 资产报酬率  | <i>Roa</i>   | 净利润/总资产        |  |
|      | 总负债率   | <i>Tl</i>    | 总负债/总资产        |  |
|      | 无形资产比重 | <i>Itang</i> | 无形资产净额/总资产     |  |
|      | 企业规模   | <i>Size</i>  | 期末总资产的自然对数     |  |
|      | 企业绩效   | <i>Tobin</i> | 托宾 Q 值         |  |
|      | 企业年龄   | <i>Lnage</i> | 企业上市年龄的自然对数    |  |
|      | 资产增长率  | <i>Tagr</i>  | 本年资产增长额/年初资产总额 |  |

(三)数据说明

本文以 2009—2022 年沪深 A 股上市公司为研究样本,检验公共数据开放对企业“漂绿”行为的影响。样本剔除了金融保险类、ST 或 PT 以及财务数据缺失的公司,最终得到 11284 个观测值。同时,为了减小极端值的影响,对所有连续变量在上下 1%分位上进行缩尾处理。企业的“漂绿”数据来自彭博数据库和华证 ESG 数据库,各地政府数据开放平台上线时间参考国家信息中心和复旦大学联合发布的《中国地方政府数据开放报告》的相关数据及各地区政府数据开放平台来确认,其他数据来自国泰安(CSMAR)数据库。

四、实证分析

(一)描述性统计

主要变量的描述性统计结果如表 2 所示。企业“漂绿”( *GWS* )的标准差为 1.1825,最大值为 3.9004,最小值为-6.0127,说明不同企业之间的“漂绿”程度差异较大,部分企业存在较强的“漂绿”现象。公共数据开放( *DID* )的均值为 0.4312,表明约 43%的企业已处于政策实施地区。其他变量数据分布合理,整体特征与已有研究相符,满足后续实证分析的前提要求。

| 表 2          |       | 描述性统计结果 |        |         |         |         |
|--------------|-------|---------|--------|---------|---------|---------|
| 变量           | 样本数   | 平均值     | 标准差    | 最小值     | 中位数     | 最大值     |
| <i>GWS</i>   | 11284 | 0.0129  | 1.1825 | -6.0127 | 0.1358  | 3.9004  |
| <i>DID</i>   | 11284 | 0.4312  | 0.4953 | 0.0000  | 0.0000  | 1.0000  |
| <i>Roa</i>   | 11284 | 0.0454  | 0.0786 | -3.1644 | 0.0401  | 0.6640  |
| <i>Tl</i>    | 11284 | 0.4775  | 0.2037 | 0.0140  | 0.4893  | 2.4707  |
| <i>Itang</i> | 11284 | 0.0517  | 0.0762 | 0.0000  | 0.0328  | 0.7941  |
| <i>Size</i>  | 11284 | 23.1970 | 1.3791 | 18.2659 | 23.0942 | 28.6365 |
| <i>Tobin</i> | 11284 | 2.0517  | 1.7228 | 0.6338  | 1.5127  | 29.1670 |
| <i>Lnage</i> | 11284 | 2.4303  | 0.7555 | 0.0000  | 2.6391  | 3.4965  |
| <i>Tagr</i>  | 11284 | 0.1803  | 0.6492 | -0.8282 | 0.1004  | 45.4604 |

(二)基准回归

本文采用双重机器学习方法估计公共数据开放对企业“漂绿”行为的影响。为防止算法在训练过程中出现过拟合问题,样本按 1:5 比例进行分割,并使用支持向量机模型进行预测与回归估计。基准回归结果见表 3。列(1)结果显示,公共数据开放( *DID* )的回归系数

为-0.0521,并在5%水平上显著,表明公共数据开放显著抑制了企业“漂绿”程度。从经济意义的角度来看,基准回归结果表明,公共数据开放带来的政策冲击约相当于企业“漂绿”指标0.044个标准差的下降。在企业“漂绿”行为具有高度离散性与隐蔽性的背景下,这一下降幅度意味着公共数据开放已在整体层面形成具有现实约束意义的治理效应。考虑到公共数据开放从平台上线到制度运行成熟可能存在一定的传导时滞,本文进一步将被解释变量替换为滞后一期的“漂绿”指标进行检验,结果如表3列(2)所示,公共数据开放系数依然显著为负,说明公共数据开放对企业“漂绿”的抑制作用具有一定的持续性。至此,假说H1得到验证。

表3 基于部分线性模型的基准回归结果

|            | (1)                   | (2)                  |
|------------|-----------------------|----------------------|
|            | <i>GWS</i>            | <i>GWS_lag</i>       |
| <i>DID</i> | -0.0521**<br>(0.0252) | -0.0511*<br>(0.0264) |
| 控制变量       | 是                     | 是                    |
| 时间固定       | 是                     | 是                    |
| 个体固定       | 是                     | 是                    |
| 样本量        | 11284                 | 9869                 |

注: \*、\*\*、\*\*\*分别表示在10%、5%、1%的水平上显著,括号内为稳健标准误。下同。

### (三)稳健性检验

#### 1.内生性问题

考虑到遗漏变量以及反向因果关系等内生性问题可能会影响回归结果,本文采用工具变量法对公共数据开放的估计结果进行稳健性检验。一般而言,信息化水平较高的城市更可能率先建设并上线公共数据开放平台,但若地区特征同时影响政策实施条件或企业外部监管环境,则可能导致公共数据开放与企业“漂绿”之间存在潜在的内生性偏误。为解决这一问题,本文借鉴黄群慧等(2019)的研究思路,选取地区数字经济发展水平(*Digita*)作为工具变量。具体地,以“每百人中互联网宽带接入用户数”衡量各地区的数字基础设施水平。一方面,数字基础设施的完善能显著提升地方政府的信息化治理能力,从而提高公共数据开放平台的上线概率,满足工具变量的相关性要求;另一方面,地区数字经济发展水平主要反映宏观的信息基础设施状况,其作用主要为促进政策建设与执行,而非直接影响企业的信息披露决策或“漂绿”行为,因此符合外生性假设。工具变量检验结果如表4所示。第一阶段回归结果如列(1)所示,工具变量 *Digita* 与核心解释变量 *DID* 之间存在显著正相关关系,说明地区数字经济发展水平能够显著促进公共数据开放政策的落地实施。Cragg-Donald F 统计量与 Kleibergen-Paap LM 检验结果均远高于临界值,表明不存在弱工具变量问题,工具变量具有较强解释力。第二阶段回归结果如列(2)所示, *DID* 的系数仍为显著负值,表明公共数据开放对企业“漂绿”行为的抑制作用依然显著,与基准回归结果一致,进一步验证了本文核心结论的稳健性。

表 4 工具变量检验结果

|                        | (1)                   | (2)                  |
|------------------------|-----------------------|----------------------|
|                        | 第一阶段                  | 第二阶段                 |
| <i>Digita</i>          | 0.0137***<br>(0.0001) |                      |
| <i>DID</i>             |                       | -0.5164*<br>(0.2755) |
| 控制变量                   | 是                     | 是                    |
| 时间固定                   | 是                     | 是                    |
| 个体固定                   | 是                     | 是                    |
| 样本量                    | 11284                 | 11284                |
| Kleibergen-Paap LM 统计量 | 5808.37***            |                      |
| Cragg-Donald F 统计量     | 16037.85              |                      |
| Kleibergen-Paap F 统计量  | 15226.91              |                      |

2.替换被解释变量

参考 Zhang(2022)与 Avramov 等(2022)的研究,采用如下公式:  $GWS_{i,t} = \frac{|ER_{dis,i,t} - ER_{per,i,t}|}{\sqrt{2}}$

重新计算企业“漂绿”程度,变量含义同式(6)。该指标以企业环境信息披露得分与环境绩效得分之间的绝对差异衡量“漂绿”程度,差异越大代表企业绿色信息披露与实际表现偏离程度越高。基于该替代指标,本文对样本数据进行重新清洗与匹配,并重新估计主回归模型。结果如表 5 列(1)所示,公共数据开放的系数依然显著为负,表明在更换“漂绿”度量方法后,公共数据开放对企业“漂绿”行为的抑制效应仍然稳健成立。

3.调整研究样本

考虑到部分企业的“漂绿”指标(*GWS*)取值小于零,即其 ESG 信息披露得分低于实际环境绩效得分,这类样本在实证分析中可能引入识别偏误。这一现象可能并非源于企业缺乏环境责任感,而是由于部分企业在信息披露中采取了相对保守的策略。企业选择在 ESG 报告中有意淡化自身的环境绩效(Strahilevitz, 2013),从而表现出“低披露、高绩效”的特征,即“漂棕”。此类“漂棕”行为可能导致“漂绿”指标的负值样本在一定程度上偏离真实意义上的“漂绿”概念,若直接纳入回归分析,可能对公共数据开放政策效应的识别造成干扰。鉴于上述情况,为进一步检验实证结果的稳健性,本文对“漂绿”指标进一步进行双侧 10%截尾处理,并据此重新估计主回归模型。回归结果如表 5 列(2)所示,公共数据开放的系数仍为显著负值,结论方向与基准模型保持一致,表明本文的核心结论不受异常样本的影响,具有较强稳健性。

4.缩短时间窗口

新冠疫情在 2019 年末爆发,并于 2020 年上半年达到高峰。期间各地陆续实施封控管理,企业生产经营活动受到不同程度限制,信息披露节奏和环境治理行为均可能出现阶段性波

动。为防止这一突发公共事件对企业绿色信息披露及“漂绿”行为造成干扰,本文在稳健性检验中进一步剔除了2020年的样本数据,并对其余年度样本重新进行回归估计。结果如表5列(3)所示,公共数据开放的回归系数仍为显著负值,与基准回归保持一致,表明研究结论未受疫情特殊时期的系统性影响。

5.改变样本分割比例

为提升模型估计效率并增强结果的代表性,本文将回归模型中训练样本与测试样本的划分比例由1:5调整为1:7,再次开展主回归分析。调整后结果如表5列(4)所示,公共数据开放的回归系数与基准回归结果保持一致,说明研究结论在不同数据分割设定下依然能保持稳健。

表5 稳健性检验

|            | (1)                    | (2)                   | (3)                   | (4)                  |
|------------|------------------------|-----------------------|-----------------------|----------------------|
|            | 替换被解释变量                | 调整研究样本                | 缩短时间窗口                | 改变分割比例               |
| <i>DID</i> | -0.0923***<br>(0.0263) | -0.0385**<br>(0.0175) | -0.0528**<br>(0.0266) | -0.0491*<br>(0.0252) |
| 控制变量       | 是                      | 是                     | 是                     | 是                    |
| 时间固定       | 是                      | 是                     | 是                     | 是                    |
| 个体固定       | 是                      | 是                     | 是                     | 是                    |
| 样本量        | 8761                   | 9140                  | 10177                 | 11284                |

6.更换双重机器学习算法

为检验算法稳定性,将主回归中采用的支持向量机算法替换为线性支持向量机、神经网络与套索回归,检验结果如表6所示。核心解释变量在不同算法下均显著为负,与主回归一致,表明公共数据开放对企业“漂绿”具有稳健的抑制效应,不依赖于单一机器学习算法的设定。

表6 更换算法检验结果

|            | (1)                    | (2)                   | (3)                  |
|------------|------------------------|-----------------------|----------------------|
|            | 线性向量机                  | 神经网络                  | 套索回归                 |
| <i>DID</i> | -0.1999***<br>(0.0242) | -0.0709**<br>(0.0358) | -0.0525*<br>(0.0272) |
| 控制变量       | 是                      | 是                     | 是                    |
| 时间固定       | 是                      | 是                     | 是                    |
| 个体固定       | 是                      | 是                     | 是                    |
| 样本量        | 11284                  | 11284                 | 11284                |

五、异质性分析

(一)污染程度异质性

在我国现行环保政策体系中,重污染行业始终是环境监管的重点对象,相关企业在生产过程中更容易成为环境执法与公众关注的焦点,一旦出现环境违规或披露不实,所面临的处罚风

险、声誉损失及融资压力都更为严峻。为了在高压监管环境下维持形象、缓解外部压力,这类企业往往具有更强的“漂绿”动机。在此背景下,公共数据开放显著提升了环境信息的可获取性与透明度,使社会公众、投资者和监管机构能够更加直接地验证企业的环境行为与披露真实性。这种信息放大强化了外部监督力量,削弱了重污染行业企业“漂绿”的可行性与隐蔽性,从而在这一群体中表现出更强的抑制效果。相较之下,非重污染行业企业在环境风险暴露度较低、社会关注度有限的情况下,其“漂绿”行为不易受到舆论或政策的及时约束,公共数据开放所带来的透明化压力尚不足以形成实质性惩戒,因此抑制作用不显著。参考刘美玉等(2025)的做法,本文对企业所属行业进行分类,进一步将样本划分为“重污染行业”组与“非重污染行业”组,并分别对主回归模型进行分组回归分析,以识别政策效果在不同污染类型行业中的差异性。具体回归结果如表7列(1)所示,公共数据开放对重污染行业企业的“漂绿”抑制作用显著,而对非重污染行业企业影响不明显,说明政策在高风险、高动机的企业群体中更能发挥治理效能。

## (二)管理层决策异质性

企业是否选择“漂绿”策略,在很大程度上取决于管理层的决策偏好与激励导向。若管理层更关注短期财务表现或资本市场反应,则可能更倾向于通过低成本手段来美化企业形象,以博取外部投资者信任。然而,此类行为一旦被揭示,不仅可能引发投资者信任危机,还可能导致企业声誉受损、股价下挫,甚至产生诉讼纠纷。在由短视管理层主导的企业中,公共数据开放政策所产生的“揭露效应”应更为敏感和显著。数据平台的公开透明属性,加大了短期操控行为被外部识别和放大的概率,从而提升“漂绿”成本,削弱此类管理层实施选择性披露的可能性。换言之,公共数据开放在这类企业中,具有更强的威慑性与约束力,对其治理效果应更为显著。

为验证上述推论,本文参考胡楠等(2021)构建的文本分析方法,以企业年报文本中反映管理层短视主义的关键词频率占比,衡量管理者短视倾向。根据指标中位数将样本划分为“短视型管理层”组与“非短视型管理层”组,并分别进行回归检验,结果如表7列(2)所示。从结果来看,在短视型管理层样本中,公共数据开放显著抑制企业“漂绿”行为,而在非短视型管理层组中,该政策效应不显著。这一发现与预期一致,表明公共数据开放在治理短期主义主导下的信息操控行为方面具有更强的针对性与干预效果。而对于本身信息披露较为规范、决策偏向长期视角的企业,政策的边际约束力则相对有限。

## (三)所有权性质异质性

企业所有权结构不仅影响其内部治理机制,也决定了其对公共政策的响应方式与执行强度。与非国有企业相比,国有企业在信息披露、社会责任履行及政策执行方面具有更高的制度约束和外部监督压力。一方面,国有企业与政府部门在治理体系上存在更紧密的制度关联,政策传导渠道更加顺畅,当环境信息通过政府平台集中公开时,其披露行为更容易受到监管部门的直接监督与反馈,从而在短期内呈现出更强的政策响应效应。另一方面,非国有企

业在制度性监管和资源获取方面约束较弱,对公共数据开放所形成的信息压力的反应相对滞后,政策效应的显现可能需要更长的时间或更具体的激励机制。因此,本文将样本企业分为国有企业和非国有企业进行分组回归,回归结果如表7列(3)所示,公共数据开放在国有企业样本中具有显著的抑制“漂绿”的作用,而在非国有企业中系数虽为负值,但不显著。这一结果说明公共数据开放更有可能在制度约束较强的主体中率先发挥对“漂绿”的抑制作用,而在中小型、民营企业中,则需要辅以更具体的规范机制和激励约束手段。

(四)地区市场化程度异质性

考虑到公共数据开放作为一种制度性治理工具,其作用效果可能受到地区制度环境与市场基础条件的制约(Cao et al., 2021),本文进一步基于樊纲等(2011)构建的市场化指数,对样本地区的市场化程度进行分组检验。具体而言,按照各地区市场化指数的年度中位数,将样本划分为市场化程度较高地区与市场化程度较低地区两组,并在此基础上分别估计公共数据开放对企业“漂绿”行为的影响。回归结果如表7第(4)列所示,在市场化程度较高的地区,公共数据开放能够显著抑制企业“漂绿”行为;而在市场化程度较低的地区,公共数据开放对企业“漂绿”的治理效果并不明显。这一结果表明,公共数据开放的治理效应具有显著的制度环境依赖性,其对企业行为的约束功能更容易在市场机制较为成熟、信息环境较为完善的地区得以发挥。可能的原因在于,一方面,在高市场化地区,政府数据开放往往与较为完善的信息披露制度、媒体监督机制以及第三方数据服务市场相结合,企业绿色行为更容易被持续跟踪、交叉验证与外部评价,从而显著提高“漂绿”行为的识别概率与声誉成本。另一方面,市场化程度较高地区中投资者保护水平与要素流动效率相对更高,企业面临的资本市场约束与竞争压力更强,公共数据开放更容易转化为实质性的治理力量,从而强化对企业象征性披露与策略性“漂绿”的约束。

表7 异质性分析结果

|            | (1)                    |                     | (2)                  |                     | (3)                  |                     | (4)                    |                    |
|------------|------------------------|---------------------|----------------------|---------------------|----------------------|---------------------|------------------------|--------------------|
|            | 污染程度异质性                |                     | 管理层决策异质性             |                     | 所有权异质性               |                     | 地区市场化程度异质性             |                    |
|            | 重污染行业                  | 非重污染行业              | 短视型管理层               | 非短视型管理层             | 国有企业                 | 非国有企业               | 市场化程度高                 | 市场化程度低             |
| <i>DiD</i> | -0.1711***<br>(0.0465) | -0.0019<br>(0.0315) | -0.0630*<br>(0.0364) | -0.0294<br>(0.0368) | -0.0574*<br>(0.0347) | -0.0312<br>(0.0393) | -0.1622***<br>(0.0395) | 0.0437<br>(0.0361) |
| 控制变量       | 是                      | 是                   | 是                    | 是                   | 是                    | 是                   | 是                      | 是                  |
| 时间固定       | 是                      | 是                   | 是                    | 是                   | 是                    | 是                   | 是                      | 是                  |
| 个体固定       | 是                      | 是                   | 是                    | 是                   | 是                    | 是                   | 是                      | 是                  |
| 样本量        | 4118                   | 7166                | 5643                 | 5641                | 6555                 | 4729                | 5633                   | 5651               |

六、机制检验

(一)增大规制性压力

公共数据开放并不直接等同于监管强度的提升,其通过改善政府部门的信息基础与数据

条件,降低了环境监管的识别成本和执行成本,使环境议题更容易进入地方治理议程并转化为可操作的政策行动。因此,公共数据开放往往首先表现为地方政府在政策表述与治理重心上对生态环境问题的持续强化,并进一步转化为更高强度的制度约束。基于此,本文借鉴邵帅等(2024)的方法,以各城市政府工作报告中环保关键词所在句子字数占报告总字数的比例来衡量城市环境规制的强度。该指标能够反映地方政府在政策表述层面对生态环境问题的关注程度及其治理优先顺序,从而间接刻画公共数据开放背景下,环境议题由信息可得向制度约束转化的过程。比例越高,表明地方政府对环境治理的政策关注与制度投入越强,企业因而面临更高水平的规制性压力。检验结果如表8列(1)所示,公共数据开放对规制性压力的估计系数显著为正;在进一步采用滞后一期规制性压力指标后,结果如表8列(2)所示,仍显著为正,表明公共数据开放平台上线后城市层面的环境规制强度具有持续提升趋势。这一结果表明,公共数据开放通过推动政府部门之间的数据共享与信息可视化,改善了环境执法、行政处罚与排污许可等监管行为的信息基础与执行条件,从而系统性强化了企业所面临的政策约束。在监管识别能力与问责确定性同步提升的情形下,企业“漂绿”行为被发现与追责的概率显著上升,进而促使企业减少选择性披露和虚假陈述,降低“漂绿”行为发生的可能性。假说H2a得证。

## (二)强化规范性压力

除正式制度规制外,信息传播所塑造的社会舆论与媒体监督构成了重要的规范性约束机制。为验证假说H2b,本文参考郭檬楠等(2023)的做法,选取 $i$ 公司 $t$ 年被网络媒体报道的总次数加1的自然对数来衡量规范性制度压力。媒体报道不仅反映企业进入舆论视野的程度,更体现社会监督是否具备稳定的信息来源与事实支撑。当环境信息可获取性与可核验性不足时,媒体监督往往高度依赖线索型调查与事件驱动,难以形成持续关注。公共数据开放通过提供可核验的第三方数据基础,显著降低媒体开展基于事实监督的成本,促使舆论监督实现常态化。检验结果如表8列(3)所示,公共数据开放变量的系数显著为正,表明公共数据开放平台上线显著提升了企业在媒体中的曝光度。同时,鉴于媒体关注可能随企业当期披露行为变化而同步调整,本文继续采用滞后一期的媒体报道指标刻画规范性压力,以更清晰地界定公共数据开放引致的信息可核验性提升如何进一步强化外部舆论监督。检验结果如表8列(4)所示,公共数据开放对滞后一期媒体报道的影响仍显著为正。该结果表明,公共数据开放通过提升环境信息的可核验性,重塑了舆论监督的事实基础,使媒体能够围绕企业环保表现开展持续、基于数据的报道与核查,由此将原本零散的社会关注转化为稳定的声誉约束机制。在此背景下,企业为避免声誉损失,更倾向于提升信息披露质量、减少虚假陈述和粉饰行为。上述结果支持了公共数据开放通过规范性压力约束企业“漂绿”的机制。假说H2b得证。

## (三)加剧模仿性压力

当行业内企业面临共同的制度变革时,只有在企业行为具有持续可比性的条件下,模仿

性压力才可能形成稳定传导。为识别公共数据开放是否通过强化行业参照结构进而加剧模仿性压力,本文以企业在所属行业中的市场份额刻画其潜在标杆地位。市场份额越高,企业在行业中的可见度与影响力越强,其环境披露与治理方式越容易被同行视为可参照对象。结果如表8列(5)所示,公共数据开放系数在1%水平下显著为正。同时,为更贴近模仿性压力的后续传导过程,本文在列(6)进一步以滞后一期市场份额进行检验,结果保持一致。上述结果表明,公共数据开放通过提高行业信息的可比较性,使部分企业更容易形成稳定的行业参照地位。在此基础上,同行企业在横向比较与相对评价压力下,更倾向于对标高市场份额企业的披露方式与环境治理取向,主动调整自身行为以缩小差距。由此,公共数据开放通过强化以行业参照为核心的模仿性压力,对企业“漂绿”行为形成约束。假说H2c得证。

表 8 机制检验结果

|            | 规制性压力               |                       | 规范性压力                 |                       | 模仿性压力                 |                       |
|------------|---------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
|            | (1)                 | (2)                   | (3)                   | (4)                   | (5)                   | (6)                   |
| <i>DID</i> | 0.0081*<br>(0.0047) | 0.0181***<br>(0.0045) | 0.0639***<br>(0.0199) | 0.1017***<br>(0.0202) | 0.6801***<br>(0.1930) | 0.7075***<br>(0.1996) |
| 控制变量       | 是                   | 是                     | 是                     | 是                     | 是                     | 是                     |
| 时间固定       | 是                   | 是                     | 是                     | 是                     | 是                     | 是                     |
| 个体固定       | 是                   | 是                     | 是                     | 是                     | 是                     | 是                     |
| 样本量        | 11284               | 9869                  | 11284                 | 9869                  | 11284                 | 9869                  |

七、研究结论与政策启示

本文基于2009—2022年沪深A股上市公司样本,以地方公共数据开放平台上线为准自然实验,运用双重机器学习方法系统检验公共数据开放对企业“漂绿”行为的治理效应。研究结果表明,公共数据开放显著抑制了企业“漂绿”行为,且在多种稳健性检验后结论依然成立,表明公共数据开放能够在微观层面有效约束企业环境信息失真行为。异质性分析显示,公共数据开放的治理效应具有明显的对象差异与制度环境依赖特征,其抑制作用在重污染行业、短视型管理层主导的企业以及国有企业中更为显著,并主要集中于市场化程度较高的地区,说明公共数据开放更容易在环境风险较高、信息操纵动机更强且制度基础较为完善的情境中转化为实质性约束。机制检验表明,公共数据开放并非通过直接干预企业决策发挥作用,而是通过重构环境信息结构,提升环境信息的可得性、可核验性与可比较性,进而系统性强化了监管约束、社会监督与行业参照等多重制度压力,提高了企业“漂绿”行为被识别与被问责的概率,推动企业由象征性披露转向更真实的环境履责。

基于上述研究结论,本文提出如下政策建议:

第一,围绕高风险领域与关键主体,实施差异化的数据开放与监管协同策略。实证结果表明,公共数据开放在重污染行业与国有企业中具有更为显著的“漂绿”抑制效应。这一发现

意味着,公共数据开放更容易在环境风险集中、制度约束较强的领域率先转化为治理能力。政策上,应将重污染行业作为公共数据开放与环境治理深度融合的优先场景,重点推进排污许可、在线监测、执法处罚、环境信用等数据的高频更新与跨部门贯通,并同步嵌入监管流程与问责机制,强化数据开放对执法识别与精准监管的支撑作用。对于国有企业,应进一步推动其在公共数据平台中的信息整合与规范披露,将其作为公共数据驱动环境治理的制度接口,发挥其在政策传导与行业示范中的枢纽作用。

第二,将公共数据开放系统嵌入金融与市场约束体系,重点强化对短视型企业的长期约束。异质性结果显示,公共数据开放对短视型管理层企业的治理效应尤为显著,机制检验也表明规范性压力与规制性压力是重要传导路径。这表明,公共数据的治理价值不仅体现在政府监管层面,更在于其能否转化为稳定的外部约束。政策方面,应推动公共数据开放与绿色信贷、债券审核、ESG评级、投资尽调等市场机制深度对接,将开放数据作为核验企业环境信息的基础,降低金融机构对企业自报信息的依赖,使环境数据更直接进入融资决策与风险定价过程。通过将数据开放嵌入资本配置与信用评价体系,可使“漂绿”行为持续体现为融资约束与声誉后果,从而对短期主义导向企业形成更具持续性的外部约束。

第三,依托公共数据开放重构行业参照结构,发挥龙头企业的示范扩散效应。机制检验表明,公共数据开放通过提高行业可比较性,强化了以行业标杆为核心的模仿性压力。在此基础上,可引导市场份额高、环境治理基础较好的企业,基于公共数据开展标准化、可核验的环境绩效披露与对标展示,在行业内形成稳定参照坐标。通过鼓励龙头企业发布基于公共数据核验的高质量ESG报告、参与绿色治理标准建设与信息平台共建,使公共数据不仅成为监管工具,也成为行业自我约束与规范扩散的基础设施,从而通过模仿性压力推动绿色披露规范在行业内扩散。

## 参考文献:

- [1] 蔡真,万兆,霍盈辰,等.“漂绿”对绿色创新的挤出效应[J].世界经济,2025,48(11):97-125.
- [2] 陈琪,李梦函.企业数字化转型能抑制ESG“漂绿”行为吗?[J].财贸研究,2025,36(10):96-110.
- [3] 陈诗一,卢娟,栾雨晴.产业链协同视角下公共数据开放与制造业污染减排[J].经济学动态,2025(12):34-51.
- [4] 程秋旺,蔡雪雄,文守逊.数字化转型、绿色创新与企业漂绿行为[J].北京工商大学学报(社会科学版),2025,40(01):37-50.
- [5] 樊纲,王小鲁,马光荣.中国市场化进程对经济增长的贡献[J].经济研究,2011,46(9):4-16.
- [6] 方锦程,刘颖,高昊宇,等.公共数据开放能否促进区域协调发展?——来自政府数据平台上线的准自然实验[J].管理世界,2023,39(9):124-142.
- [7] 方恺,毛梦圆,刘潇,等.双碳政策工具的共同富裕效应——基于中国核证自愿减排项目的县域数据

研究[J]. 浙江大学学报(人文社会科学版), 2023, 53(2): 101-115.

[8] 宫绮,董秀成,孔朝阳. 企业 ESG“漂绿”对就业规模的影响及作用机制[J]. 中国人口·资源与环境, 2025, 35(11): 115-126.

[9] 郭家堂. 公共数据开放与中国绿色全要素生产率:数据要素的视角[J]. 经济研究, 2025, 60(2): 56-72.

[10] 郭檬楠,贺一凡,牛建业. 内部控制、网络媒体报道与企业 ESG 表现[J]. 管理学报, 2023, 36(3): 103-119.

[11] 何文剑,李学尧,苗妙. 环境信息规制、信息基础设施与企业数字化转型[J]. 南开经济研究, 2025(7): 168-188.

[12] 胡洁,于宪荣,韩一鸣. ESG 评级能否促进企业绿色转型?——基于多时点双重差分法的验证[J]. 数量经济技术经济研究, 2023, 40(7): 90-111.

[13] 胡楠,薛付婧,王昊楠. 管理者短视主义影响企业长期投资吗?——基于文本分析和机器学习[J]. 管理世界, 2021, 37(5): 139-156, 11, 19-21.

[14] 黄宏斌,张娴婧,张玥杨. ESG 表现对企业绿色协同创新的影响研究[J]. 统计与信息论坛, 2025, 40(12): 88-102.

[15] 黄群慧,余泳泽,张松林. 互联网发展与制造业生产率提升:内在机制与中国经验[J]. 中国工业经济, 2019(8): 5-23.

[16] 黄溶冰,谢晓君,周卉芬. 企业漂绿的“同构”行为[J]. 中国人口·资源与环境, 2020, 30(11): 139-150.

[17] 韩一鸣,胡洁,于宪荣. 企业 ESG 表现与产业链供应链韧性——基于信号传递的视角[J]. 经济与管理研究, 2025, 46(09): 3-20.

[18] 蒋艺翹,姚树洁. ESG 信息披露、外部关注与企业风险[J]. 系统管理学报, 2024, 33(1): 214-229.

[19] 金暄暄,黄鑫,王俊豪. 环境监管垂直化与企业策略性信息披露[J]. 南开经济研究, 2025(5): 250-270.

[20] 金环,任仕佳. 公共数据开放、知识搜寻与企业合作创新[J]. 华东经济管理, 2026, 40(04): 53-63.

[21] 李常青,辛立柱. 股权质押背景下的 ESG 漂绿行为研究[J]. 管理科学, 2025, 38(3): 52-67.

[22] 李大元,贾晓琳,辛琳娜. 企业漂绿行为研究述评与展望[J]. 外国经济与管理, 2015, 37(12): 86-96.

[23] 李恩极,谢康,李晶. 公共数据开放对企业高质量发展的影响研究——基于城市公共数据开放平台上线的准自然实验分析[J]. 工业技术经济, 2025, 44(12): 69-78.

[24] 李强,蒋洮. 高管团队国际化能抑制企业“漂绿”行为吗——兼论信任文化的影响[J]. 中国地质大学学报(社会科学版), 2023, 23(4): 37-55.

[25] 李强,王睿,何子纯. 机构投资者退出威胁具有绿色治理作用吗?——基于企业“漂绿”行为的考察[J]. 经济与管理, 2023, 37(04): 72-82.

[26] 李四海,马文琪. 环境规制压力与企业“多言寡行”的“漂绿”行为[J]. 环境经济研究, 2025, 10(1): 46-73.

[27] 刘美玉,段红莉,黄速建. 双化协同对企业高质量发展的影响[J]. 经济与管理研究, 2025, 46(12): 39-56.

[28] 刘兴华,李文圳,文杰. 机构投资者实地调研与企业“漂绿”行为:信息约束与内部治理视角[J]. 中央财经大学学报, 2025(8): 92-107.

- [29] 陆雪艳,潘海英,新夫. 基于ESG信息披露同群模仿的企业“漂绿”动因与治理机制[J]. 中国人口·资源与环境,2025,35(05):116-128.
- [30] 吕怀立,高逸菲,刘诚芸. ESG漂洗行为对财务柔性的影响研究[J]. 管理学报,2025,38(5):32-47.
- [31] 罗崇佳,段茂盛. 中小企业环境信息披露:现状、难点及对策[J]. 环境经济研究,2025,10(3):178-198.
- [32] 乔子轶,滕曼茹,戴明禹. 公共数据资源共享能否提升供应链韧性? 基于公共数据开放平台的准自然实验[J]. 世界经济研究,2025(12):88-102,133-134.
- [33] 邵帅,葛力铭,朱佳玲. 人与自然何以和谐共生:地理要素视角下的环境规制与环境福利绩效[J]. 管理世界,2024,40(8):119-146.
- [34] 沈坤荣,林剑威. 链“岛”成“陆”:公共数据开放的技术创新效应研究[J]. 管理世界,2025,41(2):83-104.
- [35] 沈斌,黎江虹. 论公共数据的类型化规制及其立法落实[J]. 武汉大学学报(哲学社会科学版),2023,76(1):67-77.
- [36] 石建勋,辛沛远. 市场激励型环境规制能抑制上市企业“漂绿”行为吗?——来自《环境保护税法》实施的准自然实验证据[J]. 华东经济管理,2025,39(9):86-96.
- [37] 史永东,王湜森. 企业社会责任与公司价值——基于ESG风险溢价的视角[J]. 经济研究,2023,58(6):67-83.
- [38] 汤小莉,王莽,张涵诗,等. 公共数据开放能否降低ESG评级分歧?——基于信息效应与监督效应的分析[J]. 金融发展研究,2025(10):3-15.
- [39] 陶云清,侯婉玥,刘兆达,等. 公众环境关注如何提升企业ESG表现?——基于外部压力与内部关注的双重视角[J]. 科学学与科学技术管理,2024,45(7):88-109.
- [40] 汪克亮,付丽翔. 公共数据开放与城市绿色发展效率——来自政府数据平台上线的准自然实验[J]. 科技进步与对策,2025,42(21):106-118.
- [41] 王海,叶帅,尹俊雅. 公共数据开放如何提振企业有效投资——基于产能利用视角[J]. 中国工业经济,2024(8):137-153.
- [42] 汪立鑫,孙圣涛. 公共数据开放的绿色价值——基于企业实质性绿色技术创新视角[J]. 财经科学,2026(01):124-140.
- [43] 王伟,刘传红. “漂绿广告”监管需要建立引爆机制[J]. 中国地质大学学报(社会科学版),2013,13(6):70-75,134.
- [44] 魏宇琦,高锦萍. 高管激励对企业ESG“漂绿”行为的影响研究[J]. 科学学与科学技术管理,2026,47(01):153-167.
- [45] 习明明,李婷,张卢千漪. 网络媒体监督与企业环境责任行为的“多言寡行”——基于大语言模型的人工智能技术分类测度[J]. 兰州大学学报(社会科学版),2025,53(03):159-176.
- [46] 谢佳松,樊嘉诚,林建浩. 信息摩擦、公共数据与企业创新边界[J]. 经济研究,2025,60(11):51-73.
- [47] 解子恒,靳添全. 环境信息公开与企业绿色采购——基于制度同构理论的视角[J]. 经济动态,2026(01):135-158.

- [48] 肖红军,张俊生,李伟阳. 企业伪社会责任行为研究[J]. 中国工业经济,2013(6):109-121.
- [49] 杨永聪,李学轩. 从象征性披露到实质性投入:中央环保督察何以抑制企业“漂绿”行为[J]. 环境经济研究,2025,10(3):126-151.
- [50] 杨志安,侯耀威,任珊珊. 数据要素视角下公共数据开放的产学研合作创新效应[J]. 华东经济管理,2026,40(01):32-41.
- [51] 于庆瑞,秦梦柯,李晓阳. 融资约束、ESG表现与调节效应——基于党组织参与治理视角的分析[J]. 西南大学学报(社会科学版),2025,51(2):182-193,312-313.
- [52] 于小悦,郑晓慧. 打通优惠政策“最后一公里”:公共数据开放与企业政策信息关注[J]. 财贸经济,2026,47(1):125-141.
- [53] 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.
- [54] 张晓娟,部闫若,莫富传. 政策工具视角下我国地方政府公共数据治理政策研究[J]. 情报理论与实践,2024,47(6):21-30.
- [55] 邹祥,贾永飞. 政府数据开放与企业绿色创新效率:理论机制与实证检验[J]. 统计与决策,2025,41(03):164-169.
- [56] Avramov, D., S. Cheng, A. Lioui, et al. Sustainable Investing with ESG Rating Uncertainty[J]. Journal of Financial Economics, 2022, 145(2): 642-664.
- [57] Cao, L., Y. Tang, B. Cai, et al. Was It Better or Worse? Simulating the Environmental and Health Impacts of Emissions Trading Scheme in Hubei Province, China[J]. Energy, 2021, 217: 119427.
- [58] Chernozhukov, V., D. Chetverikov, M. Demirer, et al. Double/Debiased Machine Learning for Treatment and Structural Parameters[J]. The Econometrics Journal, 2018, 21(1): C1-C68
- [59] Du, J., H. Gao, H. Wen, et al. Public Data Access and Stock Price Synchronicity:Evidence from China[J]. Economic Modelling, 2024, 130: 106591.
- [60] Qiang, G., J. Ruan. Can Government Public Data Openness Help Enterprises Achieve Sustainable Development? Empirical Evidence Based on Corporate ESG Greenwashing[J]. Finance Research Letters, 2025, 86: 108875.
- [61] Hu, X., R. Hua, Q. Liu, et al. The Green Fog: Environmental Rating Disagreement and Corporate Greenwashing[J]. Pacific-Basin Finance Journal, 2023, 78: 101952.
- [62] Huang, Y., K. Fang, G. Liu, et al. Has the Carbon Emission Trading Scheme Induced Investment Leakage in China? Firm-Level Evidence from China's Stock Market. [J]. Energy Economics, 2025, 141: 108091.
- [63] Lieberman, M. B., S. Asaba. Why Do Firms Imitate Each Other?[J]. Academy of Management Review, 2006, 31(2): 366-385.
- [64] Nagaraj, A. The Private Impact of Public Data: Landsat Satellite Maps Increased Gold Discoveries and Encouraged Entry[J]. Management Science, 2022, 68(1): 564-582.
- [65] Strahilevitz, M. The Effects of Prior Impressions of a Firm's Ethics on the Success of a Cause-Related Marketing Campaign: Do the Good Look Better While the Bad Look Worse?[M]. London: Routledge, 2013.

- [66] Wu, Y., K. Zhang, J. Xie. Bad Greenwashing, Good Greenwashing: Corporate Social Responsibility and Information Transparency[J]. *Management Science*, 2020, 66(7): 3095–3112.
- [67] Wu, Z., X. Wang, Y. Tan, et al. A Typology of Energy Scaling in Chinese Cities and Its Implication for Sustainable Urban Transitions[J]. *Sustainable Cities and Society*, 2026, 138: 107168.
- [68] Zhang, D. Environmental Regulation and Firm Product Quality Improvement: How Does the Greenwashing Response?[J]. *International Review of Financial Analysis*, 2022, 80: 102058.
- [69] Zhang, Q., K. Fang. Comment on “Consumption-based versus Production-based Accounting of CO<sub>2</sub> Emissions: Is There Evidence for Carbon Leakage?”[J]. *Environmental Science and Policy*, 2019, 101: 94–96.

## The Impact of Public Data Openness on Corporate Greenwashing: Causal Inference Based on Double Machine Learning

Zhou Yang<sup>a</sup>, Fang Kai<sup>b</sup>, Yuan Shuqiang<sup>a</sup>, Li Haowu<sup>c</sup>

(a: “Double-Eight Strategy” Innovation and Development Institute, Party School of Zhejiang Provincial Committee of C.P.C; b: School of Public Affairs, Zhejiang University; c: Lee Kuan Yew School of Public Policy, National University of Singapore)

**Abstract:** Public data openness is an important institutional arrangement in digital governance and provides a new pathway for curbing corporate greenwashing. Exploiting the launch of city-level public data open platforms as a quasi-natural experiment, this paper uses a sample of Chinese A-share listed firms from 2009 to 2022 and applies the double machine learning approach to estimate the governance effect of public data openness on corporate greenwashing. The results show that public data openness significantly restrains corporate greenwashing. Mechanism analyses indicate that public data openness constrains corporate greenwashing by intensifying regulatory pressure, strengthening normative pressure, and amplifying mimetic pressure, thereby reducing corporate greenwashing. Heterogeneity analyses further reveal that the governance effect of public data openness exhibits pronounced firm-level differences and institutional-environment dependence: the inhibitory effect is more evident among heavily polluting firms, state-owned enterprises, and firms characterized by stronger managerial short-termism, and is significantly stronger in regions with a higher degree of marketization. This study provides important empirical evidence for improving environmental governance and promoting high-quality economic development through public data openness.

**Key words:** Public Data Openness; Corporate Greenwashing; Double Machine Learning; Institutional Pressure; Environmental Governance

**JEL Classification:** O30, P48

(责任编辑:卢 玲)