

政府数据开放平台效率评估及影响因素实证研究

——基于无导向 SBM-Malmquist-Tobit 模型

吴应强¹ 李白杨² 石乐怡¹

(1. 武汉大学信息管理学院, 湖北 武汉 430072; 2. 南京大学数据管理创新研究中心, 江苏 苏州 215163)

摘要: [目的/意义] 效率评估有利于掌握平台开放效率情况。将 SBM、Malmquist 和 Tobit 模型结合起来, 形成一个综合评估框架, 不仅可以评估当前效率, 还可以分析效率变化趋势和影响因素, 为政府数据开放平台效率评估提供新的方法借鉴。[方法/过程] 以 2019—2022 年中国 10 个省市市政府数据开放平台的“投入—产出”面板数据作为数据来源, 利用无导向 SBM 模型不需要预设生产可能性边界方向的特性, 使平台效率评估更加灵活和全面; 利用 Malmquist 指数可衡量时间序列数据效率变化的特性, 帮助平台识别效率演变趋势; 利用 Tobit 模型适用于分析截尾数据的特性, 帮助识别效率背后各影响因素的作用机理。[结果/结论] 省市市政府数据开放平台效率水平整体呈上升趋势, 平均增长率为 17.1%, 其中深圳、广东表现较为突出, 平台通过优化流程、提高服务质量、增强用户体验、提升运营模式等措施, 使整体效率得到了显著提升, 为其他平台促进开放效率和成效提供了可借鉴的经验; 平台的全要素生产率(M 指数)平均增长率为 17.1%, 表明技术进步, 但技术效率平均值下降, 存在浪费或无效率问题。规模效率平均增长率为 14.5%, 但纯技术效率下降, 表明平台需更注重提升纯技术效率; 政府数据开放平台受到平台数据综合水平、数据平台建设水平、组织准备度、财政支持水平和公众外部需求的交叉正向影响。数据的多样性、质量及可用性, 以及平台的功能设计、用户体验等方面是提升效率的关键因素, 政府的重视程度、财政支持和社会公众的需求直接影响平台的综合效率水平。

关键词: 政府数据开放; 平台; 门户网站; 效率评估; 影响因素

DOI:10.3969/j.issn.1008-0821.2024.07.007

[中图分类号] D63; G203 [文献标识码] A [文章编号] 1008-0821 (2024) 07-0074-10

Empirical Study on Efficiency Evaluation and Influencing Factors of Open Government Data Portals

——Based on the Non-Directional SBM-Malmquist-Tobit Model

Wu Yingqiang¹ Li Baiyang² Shi Leyi¹

(1. School of Information Management, Wuhan University, Wuhan 430072, China;

2. Research Institute for Data Management Innovation, Nanjing University, Jiangsu 215163, China)

Abstract: [Purpose/Significance] Efficiency evaluation is beneficial for understanding the status of platform openness efficiency. Combining the SBM, Malmquist, and Tobit models to form a comprehensive evaluation framework not only assesses current efficiency but also analyzes trends in efficiency changes and influencing factors. This provides a new methodological reference for the efficiency evaluation of government data open portals. [Method/Process] Using the “input-output” panel data of government data open portals in 10 provinces and cities in China from 2019 to 2022, the direction-

收稿日期: 2023-11-09

基金项目: 国家社会科学基金重大项目“我国政府信息公开到数据开放的理论创新与实践路径研究”(项目编号: 22&ZD329)。

作者简介: 吴应强 (1995-), 男, 博士研究生, 研究方向: 政府数据开放。李白杨 (1991-), 男, 助理教授, 博士, 研究方向: 政府数据治理、信息资源管理。石乐怡 (1997-), 女, 博士研究生, 研究方向: 信息检索和信息服务。

free SBM model was utilized for its flexibility and comprehensiveness in platform efficiency evaluation without the need to preset the direction of the production possibility frontier. The Malmquist index was employed for its ability to measure efficiency changes in time-series data, aiding in identifying trends in platform efficiency evolution. The Tobit model, suitable for analyzing censored data, helped identify the mechanisms behind various influencing factors of efficiency. [Results/Conclusion] The efficiency level of provincial and municipal government data open portals generally shows an upward 17.1%. Shenzhen and Guangdong have shown particularly notable performances. Through optimizing processes, improving service quality, enhancing user experience, and upgrading operational models, the overall efficiency has been significantly enhanced, providing referential experiences for other portals to promote open efficiency and effectiveness. The total factor productivity (M index) of the platform has an average growth rate of 17.1%, indicating technological progress, but the average value of technical efficiency has decreased, pointing to waste or inefficiency issues. The average growth rate of scale efficiency is 14.5%, but pure technical efficiency has declined, indicating that portals need to focus more on improving pure technical efficiency. Government data open portals are positively influenced by the comprehensive level of platform data, data platform construction level, organizational readiness, fiscal support level, and public external demand. The diversity, quality, and availability of data, as well as the functional design and user experience of the platform, are key factors in enhancing efficiency. The level of government attention, financial support, and public demand directly affect the comprehensive efficiency level of the platform.

Key words: open government data; platform; portal website; efficiency evaluation; influencing factors

2009年5月,美国率先上线了政府数据开放平台网站(www.data.gov),英国、加拿大、法国、挪威、韩国、新加坡等国家和地区也相继建立了政府数据开放平台。2012年6月,上海市首先上线试运行“上海市政府数据服务网”,北京、佛山、武汉等地也相继推出政府数据开放平台网站。2021年9月1日起正式实施的《中华人民共和国数据安全法》^[1]强调,“构建政务数据开放平台,推动政务数据开放利用”。2022年9月,国务院办公厅发布《全国一体化政务大数据体系建设指南》^[2],提出“通过政务数据开放平台,推动数据安全有序开放”。截至2022年10月,我国已有208个省级和城市的地方政府上线了政府数据开放平台。可见,在数字中国背景下,国家正大力推进政府数据开放平台的建设,加强数字政府转型发展。

如雨后春笋般建立起来的政府数据开放平台,其运行效率如何?通过对过往研究的文献调查发现,绝大部分学者采用数据包络分析法(Data Envelopment Analysis, DEA)对平台的效率进行评估,且数据来源多为静态截面数据。然而,在传统的DEA模型中,对于无效率程度的测量仅仅涉及所有投入(产出)等比例缩减(增加)的部分,而并未考虑到变量松弛性以及径向问题可能带来的测量误差。这种情况可能导致无效率程度的测量结果出现偏差,影响

其准确性^[3]。

本文针对上述不足,尝试采用无导向SBM-Malmquist-Tobit模型与政府数据开放平台效率评估与影响因素的研究主题进行结合,在提升效率评估精度的基础上,继续揭示平台的效率状态和变化趋势,进一步分析影响效率的关键因素。与传统DEA方法相比,显示出更高的灵活性和精确性,更适合应对政府数据开放平台这一特殊且重要的领域。

1 文献综述

政府数据开放平台效率评估的过往研究主要集中在利用数据包络法测量平台的综合效率、纯技术效率、规模效率和规模报酬,最后综合评估出政府数据开放平台的效率。DEA具有无需任何权重假设的特点,有很强的客观性。虽然大部分学者在研究方法的选择上具有一致性,但在平台效率评价指标的选取上却不尽相同。

有的学者结合网站效率评价的特点,从内容、技术、网站结构等角度选取相应的指标,例如,输入指标包含数据量、数据内容、连通率、下载速度、网站的总页面数和网站的不重复链接量,输出指标包含下载量、访问量、网站外链接量^[4]。有的学者为了避免非DEA有效结果的情况出现,从数据和平台两个维度构建了包含9个投入指标和4个产出指标的DEA评估指标体系,但必须满足决策单元

(Decision Making Unit, DMU)个数不少于输入、输出指标总数2倍的要求^[5]。

对于省级政府数据开放平台,有的学者参考复旦大学数字与移动治理实验室发布的《中国地方政府数据开放报告》,将数据层指数和平台层指数作为投入指标,避免自己重复采集数据和“造轮子”的麻烦^[6]。除了单独研究政府数据平台资源的利用效率,有学者将创新应用的效能评价与平台效率评估进行结合,极大丰富了研究内容^[7]。因为传统DEA求解出来的最大效率值为1,此时代表决策单元达到有效,但当多个决策单元的效率值达到DEA有效时,难以进行横向的区分比较,故有学者采用超效率SBM模型对政府数据开放平台的效率进行评估,最大效率值不再仅局限为1,最终结果可以由高到低进行有效排序^[8]。

上述研究对本文提供了重要的参考与借鉴,但仍存在以下几点可以提升的空间:①传统DEA模型存在一定局限,未考虑变量松弛性以及径向问题可能带来的测量误差,可采用无径向、无导向SBM模型进行改进;②绝大部分过往研究均采取静态截面数据,鲜有学者采取面板数据作为数据来源,可采集政府数据开放平台的面板数据,利用Malmquist指数模型进行分析,弥补SBM模型在动态效率分析的不足;③评估平台效率值后,未深入研究背后的影响因素,可采用Tobit模型进行影响因素作用机理的回归分析。

无导向SBM-Malmquist-Tobit模型在方法应用上具有一定优势。首先是测算精度更高,解决了传统的DEA-BCC和DEA-CCR模型无法解决的投入产出的松弛性问题以及由投入产出的角度和径向的选择所带来的偏差问题。为了改进传统DEA模型忽视变量松弛性问题和径向问题带来的测量误差,Tone K^[9]提出了基于松弛变量的非径向、非导向型SBM模型,在某种程度上克服了传统DEA的缺点,提高了测算精度。其次是由截面数据转为分析面板数据,具有较强的动态性,Malmquist指数由Malmquist S于1953年首次提出,Caves D W等^[10]将其应用于生产效率变化的测算,Färe R等^[11]将其与DEA理论相结合,建立了用来观察两个不同时期全要素生产率增长TFPCH的Malmquist指数,弥补了静态

DEA模型不能分析面板数据的不足。最后是研究力度更深,剖析问题更全,探究了影响效率值背后的深层次因素。Coelli T^[12]在数据包络分析基础上衍生了两步法(Two-stage Method),即在数据包络分析评测决策单元所得效率值基础上,将其作为因变量,以影响因素作为自变量建立回归模型。无导向SBM模型所得效率值在0~1的区间变动,如果直接使用最小二乘回归方法将会导致估计偏差,然而这种特殊的截尾数据却十分适用于Tobit模型^[13],通过Tobit回归分析所得自变量系数可判断影响因素对平台效率的影响方向和影响强度^[14-15]。SBM-Malmquist-Tobit模型研究成果较少,在CNKI数据库仅有16篇期刊论文与SBM-Malmquist-Tobit模型主题相关,但在生态效率^[14,16]、环境效率^[17]、绿色经济效率^[15]、创新效率^[3]、物流效率^[18]、科研效率^[19]均有成功的研究经验。

鉴于此,本文率先尝试性地采用无导向SBM-Malmquist-Tobit模型运用于政府数据开放平台效率研究,以2019—2022年中国10个省市数据开放平台的“投入—产出”面板数据作为数据来源。首先,利用无导向SBM模型不需要预设生产可能性边界方向的特性,使平台效率评估更加灵活和全面;其次,利用Malmquist指数可衡量时间序列数据效率变化的特性,帮助平台识别效率演变趋势;最后,利用Tobit模型适用于分析截尾数据的特性,帮助识别效率背后各影响因素的作用机理。

2 研究设计

2.1 研究工具与方法

DEA是一种客观衡量生产效率并利用非参数定量分析的科学效率评价方法,是可以实现评价多投入和多产出决策单元效率的一种效率评估方法。

SBM模型规划公式如式(1)所示。其中, ρ 为目标效率的评价指标, x_{ij} 为 j 个DMU的第 i 个投入变量, y_{rj}^e 为第 j 个DMU的第 r 个期望产出变量, y_{rj}^b 为第 j 个DMU的第1个非期望产出变量, λ_j 为各要素的权重。 x_{i0} 、 y_{r0}^e 、 y_{r0}^b 代表DMU的第 i 种投入、第 r 种期望产出和第1种非期望产出的数值; s_i^- 、 s_r^e 、 s_1^b 表示的是DMU的第 i 种投入、第 r 种期望产出和第1种非期望产出的松弛变量的值。

$$\begin{aligned}
 \rho = \min & \frac{1 - \frac{1}{m} \sum_{i=1}^m s_i^- / x_{i0}}{1 + \frac{1}{s_1 + s_2} \left(\sum_{r=1}^{s_1} s_r^g / y_{r0}^g + \sum_{l=1}^{s_2} s_l^b / y_{l0}^b \right)} \\
 s.t. & \begin{cases} x_{i0} = \sum_{j=1}^n x_{ij} \lambda_j + s_i^- \\ y_{r0}^g = \sum_{j=1}^n y_{rj}^g \lambda_j - s_r^g \\ y_{l0}^b = \sum_{j=1}^n y_{lj}^b \lambda_j + s_l^b \\ \sum_{j=1}^n \lambda_j = 1, s_i^- \geq 0, s_r^g \geq 0, s_l^b \geq 0, \lambda_j \geq 0 \\ i = 1, \dots, m; j = 1, \dots, n; r = 1, \dots, s_1; l = 1, \dots, s_2 \end{cases} \quad (1)
 \end{aligned}$$

为研究我国省市级政府数据开放平台效率的动态变化，本文采用 Färe R 等^[11]提出的 Malmquist 指数模型分析平台效率跨期变动情况，如式 (2) 所示：

$$M(x^{t+1}, y^{t+1}, x^t, y^t) = \sqrt{\frac{D^t(x_{t+1}, y_{t+1})}{D^t(x_t, y_t)} \times \frac{D^{t+1}(x_{t+1}, y_{t+1})}{D^{t+1}(x_t, y_t)}} \quad (2)$$

其中， $D^t(x_t, y_t)$ 代表第 t 阶段的技术水平表示当前的效率水平， $D^t(x_{t+1}, y_{t+1})$ 代表用第 $t+1$ 阶段的技术水平表示 $t+1$ 时期的效率水平。当 $M > 1$ 时，表示效率水平得到提升；当 $M = 1$ 时，表示效率水平不变；当 $M < 1$ 时，表示效率水平降低。

标准的 Tobit 模型如式 (3) 所示。式中： Y_i^* 为潜变量； y_i 为因变量； X_i 为自变量向量； δ 为相关系数向量； ε_i 为干扰项（独立，且 $\varepsilon_i : N(0, \sigma)$ ），因此 $Y_i^* : N(X_i \delta, \sigma)$ 。

$$\begin{aligned}
 Y_i^* &= X_i \delta + \varepsilon_i \quad (3) \\
 y_i &= Y_i^* \quad \text{if } Y_i^* > 0 \\
 y_i &= 0 \quad \text{if } Y_i^* \leq 0
 \end{aligned}$$

Tobit 模型是一种回归模型，用于分析截尾数据 (Censored Data)，即那些被截断或丢失的数据。它被广泛应用于各种领域，包括经济学、医学和环境科学等。Tobit 模型通常用于研究一些被限制或截断的变量，例如在某些情况下，数据的收集可能只得到部分观察值，这些值被称为截尾数据，参照本文即 SBM 模型得出的效率值 $\in [0, 1]$ 。Tobit 模型能够处理这种类型的数据，并估计出完整数据下的均值和方差。鉴于此，本文使用 Tobit 模型来分析省市级政府数据开放平台效率的主要影响因素。将平台效率值视作受截断的变动量，并基于 TOE 研

究框架将影响平台效率的因素视为解释变量，通过回归分析的途径来推断各个影响因素的作用机制。如此，不仅能得到关于平台效率的预测值，还能揭示出哪些因素影响了平台效率，并深入理解这些因素影响的具体程度。

2.2 样本与指标选取

2.2.1 样本选取

选取北京、上海、广州、深圳、贵州、哈尔滨、武汉、宣城等 10 个省市地方政府数据开放平台作为研究样本，主要基于以下考量：第一，平台至少在 2019 年之前上线且运维正常，具备 2019—2022 年的面板数据；第二，平台类型涵盖正省级、副省级和地市级，具有一定样本代表性；第三，平台具有面板数据的可获取性，历年数据便于获取和爬取。省市级政府数据开放平台网址如表 1 所示。

2.2.2 综合效率评价指标设计

为验证方法的可行性，选取的综合效率评价指标较为简单，仅作示例参考，具体综合指标还需研究者根据实际情况进行探究和考察，故本文最终沿用过往研究已有的成熟做法并综合考虑面板数据的可获取性，将投入指标分为数据集数量、数据接口数量和数据分析数量，将产出指标分为网站总访问量和注册用户数。最终构建省市级政府数据开放平台效率评估指标体系，如表 2 所示。

表1 省市级政府数据开放平台网址

Tab. 1 URLs of Provincial and Municipal Open Government Data Portals

名称	类型	政府数据开放平台网址
北京市公共数据开放平台	正省级	https://data.beijing.gov.cn/
上海市公共数据开放平台	正省级	https://data.sh.gov.cn/
广州市公共数据开放平台	副省级	https://data.gz.gov.cn/
深圳市政府数据开放平台	副省级	https://opendata.sz.gov.cn/
贵州省政府数据开放平台	正省级	http://data.guizhou.gov.cn/home
哈尔滨市公共数据开放平台	副省级	http://data.harbin.gov.cn
贵阳市政府数据开放平台	地市级	https://data.guiyang.gov.cn/city/index.htm
武汉市公共数据开放平台	副省级	https://data.wuhan.gov.cn/
宣城市人民政府数据开放网	地市级	http://sjkf.xuancheng.gov.cn/index/index.do
开放广东平台	正省级	http://gddata.gd.gov.cn/

表2 省市级政府数据开放平台效率评估指标体系

Tab. 2 Efficiency Evaluation Index System for Provincial and Municipal Open Government Data Portals

变量	一级指标	二级指标	指标依据	单位
平台效率水平	投入指标	数据集数量	全球开放数据指数(GODI) ^[20]	个
		数据接口数量	开放政府数据可行性研究 ^[21]	个
		数据应用数量	开放政府数据指数(OGD指数) ^[22]	个
	产出指标	网站总访问量	开放数据成熟度模型 ^[23]	人次
注册用户数		自拟	人	

2.2.3 影响因子确定

TOE 框架(Technology-Organization-Environment Framework)将影响技术创新实施的因素归纳为技术层面(Technology)、组织层面(Organization)与环境层面(Environment)3类,本质上是一种基于技术应用情境的综合性分析框架。经过逐步发展,在电子政务、数字政府等研究领域得到广泛的运用。

政府数据开放平台效率高低不仅受到平台自身

数据和建设水平的限制,同时受到政府机构准备情况、财政资源、外部公众的需求的影响。因此,对政府数据开放平台利用水平进行技术、组织、环境层面的组态分析显得尤为重要。为验证方法的可行性,本文仅以TOE理论作为示例分析框架,影响因素指标的选取难免存在遗漏和不足,最终拟定影响省市级政府数据开放平台效率的因素及其量化指标如表3所示。

表3 省市级政府数据开放平台效率影响因素及量化指标

Tab. 3 Efficiency Impact Factors and Quantitative Indicators for Provincial and Municipal Open Government Data Portals

影响因子	量化指标	单位
平台数据综合水平	中国开放数林指数指标体系数据层指数	分
数据平台建设水平	中国开放数林指数指标体系平台层指数	分
组织准备度	中国开放数林指数指标体系准备度指数	分
财政支持水平	政府一般公共预算支出	千万
公众外部需求	公众依申请公开次数	万次

2.3 数据来源与处理

数据来源上,考虑到研究样本的代表性和数据可获取性,本文选取2019—2022年中国10个省市数据开放平台的面板数据,计算无导向SBM-Malmquist-Tobit模型。2019—2021年平台效率评估指标中,数据集数量、数据接口数量和数据应用数量的数据主要来源于华中师范大学信息管理学院国家社会科学基金重点项目(17ATQ006)课题组及湖北省数据治理与智能决策研究中心共同发布的《中国政府开放数据利用研究报告》^[24](2020—2022),2022年数据主要来源于Archive.org Cache的网页历史快照记录、数据爬取和网络调查。2019—2022年网站总访问量和注册用户数来源于各政府数据平台主管部门统计的《政府网站工作年度报表》。2019—2022年平台效率影响因素的量化指标数据来源于历年《中国地方政府数据开放报告》^[25]《中国统计年鉴》和《政府信息公开工作年度报告》。

数据处理上,无导向SBM模型和Malmquist指数模型的计算采用MATLAB R2021a计算,Tobit模型采用StataMP 17计算。

3 实证结果分析

3.1 政府数据开放平台效率综合评价

采用MATLAB R2021a软件对2019—2022年我国10个省市级政府数据开放平台效率进行逐年测算,计算结果如表4所示。

表4 2019—2022年平台效率评估结果

Tab. 4 Efficiency Assessment Results for the Years 2019—2022

省市	2019年	2020年	2021年	2022年
上海	0.522	0.229	0.147	0.127
北京	1.000	1.000	0.084	0.066
广州	0.473	0.457	0.401	0.252
深圳	1.000	1.000	1.000	1.000
贵州	0.285	0.285	0.075	0.044
哈尔滨	0.339	0.265	0.176	0.136
贵阳	0.312	0.303	0.238	0.241
武汉	0.139	0.338	0.182	0.166
宣城	0.317	0.285	0.264	0.248
广东	0.390	1.000	0.264	1.000

值得注意的是,上海市数据服务网在本文的评估指标体系下,综合效率得分逐年降低。上海市数据服务网作为首个推出的政府数据开放平台,在平台建设、功能设计、用户体验等方面具备经验和基础,但在综合效率上的得分与其先发地位不匹配。北京市公共数据开放平台在2019—2020年连续达到综合效率有效,但近两年的效率值下降较为明显。

深圳市数据开放平台在2019—2022年均达到了综合效率有效,这充分显示出深圳市政府对数据开放平台网站方面的重视程度。深圳市政府对数据开放平台的建设和维护进行了大量投入,并采取了一系列措施来提高平台的综合效率。与此同时,一些省市级的政府数据开放平台综合效率却逐年降低,例如广州、贵州、贵阳、宣城,表明这些省市级数据开放平台在建设和维护方面存在一些问题,需要加强对平台的管理和优化。为了改善这种情况,这些城市需要着手在整体网站的投入进行资源配置优化,进一步合理配置、整合已有的资源。这包括加强技术投入、提高数据质量、加强用户体验等方面的工作。

2019年的武汉政务公共数据服务网的平台综合效率为0.139,2020年平台综合效率为0.338,2021年又降低至0.182,此现象主要由于2020年受新冠肺炎疫情的影响,全国公众对武汉政务公共数据服务网的数据需求激增。2019年,武汉政务公共数据服务网的网站总访问量为13.98万人次,2020年网站总访问量达到1603万人次,故导致产出指标的异常增长,带来当年平台综合效率的突然增长到次年又突然下降的现象,但与2019年相比,平台整体综合效率仍有所提升。

2020年和2022年,开放广东平台在提升综合效率方面取得了显著成效。这两年,平台通过优化流程、提高服务质量、增强用户体验等措施,使整体效率得到了显著提升。这些措施的实施,不仅使得平台在处理事务的时间上有所缩短,还提高了平台的用户满意度,进一步增强了平台的竞争力。在这两年,开放广东平台还不断加强技术创新,大力推进数字化转型。通过引入先进的技术和工具,优

化平台的运营模式，进一步提高了平台的运营效率。同时，平台还不断加强人才队伍建设，提升员工的专业素养和服务意识，为平台的长远发展打下了坚实的基础。此外，开放广东平台还注重与各方的合作与交流。通过与政府部门、企业、社会组织等各方面的合作，实现了资源共享、优势互补，推动了平台的可持续发展。这种合作模式不仅提高了平台的综合效率，还为其他平台的发展提供了可借鉴的经验。

3.2 Malmquist 指数分析结果

在分析各决策单元在不同时期的变化时，Malmquist 指数模型可以对效率进行动态研究。Malmquist 指数可分解为技术效率和技术进步，技术效率可以分解为纯技术效率和规模效率。当 $M > 1$ 时，表示效率水平得到提升；当 $M = 1$ 时，表示效率水平不变；当 $M < 1$ 时，表示效率水平降低。采用 MATLAB R2021a 软件计算，结果如表 5、表 6 所示。

表 5 2019—2022 年平台效率分年度变动指数及其分解指数

Tab. 5 Yearly Change Indices and Decomposition Indices of Portal Efficiency for the Years 2019—2022

年份	技术效率 变化指数	技术进步 变化指数	纯技术效率 变化指数	规模效率 变化指数	Malmquist 指数
2019—2020	1.205	1.165	1.156	1.298	1.392
2020—2021	0.605	1.418	0.787	0.904	0.818
2021—2022	1.129	1.102	0.986	1.233	1.302
平均值	0.980	1.228	0.976	1.145	1.171

表 6 2019—2022 年平台效率整体变动指数及其分解指数

Tab. 6 Overall Change Indices and Decomposition Indices of Portal Efficiency for the Years 2019—2022

省市	技术效率 变化指数	技术进步 指数	纯技术效率 变化指数	规模效率 变化指数	Malmquist 指数
上海	0.649	1.116	0.617	1.244	0.714
北京	0.624	1.018	1.000	0.624	0.643
广州	0.824	1.236	1.000	0.824	1.031
深圳	1.000	1.303	1.000	1.000	1.303
贵州	0.615	1.185	0.396	1.611	0.700
哈尔滨	0.739	1.236	0.602	1.448	0.909
贵阳	0.923	1.164	1.340	0.739	1.061
武汉	1.296	1.153	1.807	0.835	1.454
宣城	0.921	1.087	1.000	0.921	1.002
广东	2.204	1.782	1.000	2.204	2.894

总的来说，平台在 2019—2022 年的全要素生产率(M 指数)为 1.171，达到了效率前沿面，但增长率仅为 17.1%。尽管技术进步的平均值为 1.228，增长率为 22.8%，表明平台在技术上有所进步，但技术效率的平均值为 0.98，增长率为 -2%，意味着还存在一定程度的浪费或无效率，导致全要素生

产率的涨幅受到影响。

进一步分析技术效率，可以发现纯技术效率的平均值为 0.976，增长率为 -2.4%，而规模效率的平均值为 1.145，增长率为 14.5%。这表明平台在纯技术效率和规模效率方面存在一定的差异，平台需要更加注重提升纯技术效率。

分省市看, 2019—2022年6个省市政府数据开放平台全要素生产率大于1, 占本次研究样本总量的60%, 说明我国政府数据开放平台效率情况整体趋势较好。其中, 广州(1.031)、深圳(1.303)、贵阳(1.061)、武汉(1.454)、宣城(1.002)、广东(2.894)均达到效率前沿面, 其中开放广东平台增长幅度最高, 其次是武汉和深圳。北京的全要素生产率最低, 仅为0.643, 北京市作为我国首都, 其政府数据开放平台网站建设应该具备得天独厚的实力支撑, 但就目前得分来看, 表现不尽人意, 明显与其地位不匹配。值得注意的是, 上海市政府数据开放平台在前人的研究中均显示出较高的效率值, 但本文却得出不一样的研究结果, 这一现象值得探讨, 究其原因: 第一, 前人研究较早, 所用数据均为当年, 数据存在大幅度更新, 且近几年各地响应国家号召, 大力发展政府数据开放, 尤其以山东省

为首, 后来居上, 在2022年《中国地方政府数据开放评估报告(省域指数)》位居榜首; 第二, 前人研究较多以下载量和浏览量作为产出指标, 而本文在保持浏览量指标不变的前提下将下载量更改为注册用户数, 进而得出新的研究结果。可见, 由于指标没有固定的标准, 基于不同的理论或框架选择不同的产出指标, 会产生不同的结果, 制定更为完善的产出指标有助于加强研究结果的可靠性和评价指标体系的鲁棒性。

3.3 政府数据开放平台效率影响因素分析

本文基于TOE理论框架选取平台数据综合水平、数据平台建设水平、组织准备度、财政支持水平和公众外部需求, 量化指标和数据来源前文已做介绍, 故不再赘述, 回归区间为2019—2022年, 回归结果如表7所示。

表7 Tobit 回归模型结果: 平台效率影响因素

Tab. 7 Tobit Regression Model Results: Factors Influencing Portal Efficiency

解释变量与常数项	系数估计值	标准误	t 值	p 值	Sig
平台数据综合水平	0.006	0.002	2.91	0.007	***
数据平台建设水平	0.007	0.001	6.64	0	***
组织准备度	0.002	0.001	2.42	0.022	**
财政支持水平	0.003	0.001	3.11	0.004	***
公众外部需求	0.024	0.004	5.68	0	***
常数项	-0.107	0.036	-2.99	0.005	***

注: ***表示 $p < 0.01$, **表示 $p < 0.05$, *表示 $p < 0.1$ 。

通过回归分析, 研究发现:

1) 平台数据综合水平对政府数据开放平台效率正向影响显著。

政府数据开放平台的效率, 与平台数据的综合水平息息相关。数据的多样性、质量及可用性, 犹如平台的灵魂, 赋予平台独特的魅力与价值。一个平台若能广泛涵盖各种数据类型与来源, 如商业数据(例如商会信息和有关公司或行业数据的官方商业信息), 注册与专利、商标和公共招标相关的数据, 地理数据(例如地形和地址数据), 法律数据(例如法庭判决、立法数据), 气象数据(例如天气和气候数据), 社会数据(例如人口、就业和公共

卫生数据), 以及交通数据(例如车辆注册、交通和公共交通数据), 便能充分满足用户的需求; 数据的质量是平台综合水平的关键因素之一, 高质量的数据意味着数据的准确性与完整性更高; 数据的可用性意味着数据是否易于获取、理解和使用, 政府数据开放平台应提供清晰易懂的数据描述和元数据, 以及便捷易用的数据接口和工具。可见, 平台综合数据水平有助于提升平台的综合效率水平。

2) 数据平台建设水平对政府数据开放平台效率正向影响显著。

数据平台建设在政府数据开放中扮演着至关重要的角色, 它直接影响了平台的运行效率。政府数

据开放平台是政府数据在线访问及获取的载体，数据平台建设水平的高低表现在其功能设计是否完善、用户使用体验是否满意、平台互动反馈是否及时、平台体系的构建是否完善等方面，涉及社会公众使用平台过程中的方方面面。因此，数据平台建设水平的高低与否，直接影响到平台效率水平。

3) 组织准备度对政府数据开放平台效率正向影响显著。

组织准备度反映各省级政府对政府数据开放平台建设的重视程度，重视程度越高，越能带动相关法规政策和标准规范的制定与落实，有利于政府数据开放平台的建设与使用，从而提升平台综合效率。

组织准备度是衡量各省级政府对政府数据开放平台建设投入程度的重要标志。准备度的高低直接影响到平台建设与使用的整体效能。当政府高度重视数据开放平台建设时，会积极推动相关法规政策和标准规范的制定与执行，从而为平台的顺利建设与使用提供坚实的保障。同时，政府通过加强对平台建设和使用的监管，能够及时发现并解决存在的问题，确保平台的稳定运行和持续优化。因此，组织准备度对于提升政府数据开放平台的综合效率具有至关重要的作用。

4) 政府财政支持水平对政府数据开放平台效率正向影响显著。

当地经济发展水平越发达，则政府财政收入情况越好，分配到公共预算支出的财政资金越多。只有当政府充分认识到数据开放平台建设的重要性，并给予足够的支持和投入，才能够确保平台的顺利建设与使用，功能完善、体验感舒适的平台更有利于得到社会公众的接纳与利用，从而更好地服务于社会公众。

5) 公众外部需求对政府数据开放平台效率正向影响显著。

政府数据开放平台的产出指标是浏览量和用户注册数，当社会公众对政府数据产生需求时，才会对政府数据开放平台进行搜索引擎检索、注册、站内检索、浏览、筛查并下载满足需求条件的数据集，这一系列操作均由需求驱动。可见，公众外部需求对政府数据开放平台效率具有直接的正向影响。

4 结语

4.1 结论

本文探究了2019—2022年中国10个省市数据开放平台效率综合水平及其影响因素，利用无导向SBM-Malmquist模型对这10个省市数据开放平台效率进行综合评估，并基于TOE理论框架，构建政府数据开放平台效率影响因素Tobit模型，分析了省市数据开放平台效率影响因素的作用机理，完成了无导向SBM-Malmquist-Tobit模型在政府数据开放平台效率评估及影响因素实证研究主题上的尝试性运用，为后续研究者提供新的方法借鉴。

本文基于新方法的运用，得出以下主要结论：省市数据开放平台效率水平整体呈上升趋势，平均增长率为17.1%，其中深圳、广东表现较为突出，平台通过优化流程、提高服务质量、增强用户体验、提升运营模式等措施，使整体效率得到了显著提升，为其他平台促进开放效率和成效提供了可借鉴的经验；平台的全要素生产率(M指数)平均增长率为17.1%，表明技术进步，但技术效率平均值下降，存在浪费或无效率问题。规模效率平均增长率为14.5%，但纯技术效率下降，表明平台需更注重提升纯技术效率；政府数据开放平台受到平台数据综合水平、数据平台建设水平、组织准备度、财政支持水平和公众外部需求的交叉正向影响。数据的多样性、质量及可用性，以及平台的功能设计、用户体验等方面是提升效率的关键因素，政府的重视程度、财政支持和社会公众的需求直接影响平台的综合效率水平。

4.2 研究局限与未来展望

本文力图将无导向SBM-Malmquist-Tobit模型引入到政府数据开放平台效率评价及影响因素研究的具体实践中，重点在验证该方法的可行性，在指标设置上以参考和示例作用为主。因此，本文在平台效率指标及影响因素指标的选取上存在遗漏和不足；其次，本文选取的时间周期为2019—2022年，时间序列跨度较短，且研究样本对象数量较少，导致所收集的面板数据容量存在不足；最后，本文在确定影响因素时仅参考TOE理论框架，理论支撑

较为单一，影响因素的选取较为简单。

鉴于此，本文建议未来研究者从以下几方面对无导向SBM-Malmquist-Tobit模型在政府数据开放平台效率评价及影响因素研究的实践进行完善。第一，制定科学严谨、体系全面的平台效率指标，指标的制定必须经过专业调研和专家问询，最终赋予指标以科学性和代表性。第二，面板数据的容量需要进一步扩充，时间序列的跨度需要拉长，研究对象的数量需要增加，为效率演变和发展态势的分析提供充足的数据量支撑。第三，通过专业调研的方式确定影响因素，例如通过扎根理论分析提炼访谈内容，梳理确立初步的影响因素框架，通过多轮德尔菲法对其进行剔除、补充、更新，或是综合多个已有成熟的理论框架，从多元化角度对影响因素进行确定，为Tobit模型分析奠定科学严谨的影响因素指标。

参 考 文 献

[1] 全国人大常委会. 中华人民共和国数据安全法 [EB/OL]. <http://www.npc.gov.cn/npc/c30834/202106/7c9af12f51334a73b56d7938f99a788a.shtml>, 2024-02-06.

[2] 国务院办公厅. 国务院办公厅关于印发全国一体化政务大数据体系建设指南的通知 [EB/OL]. https://www.gov.cn/zhengce/content/2022-10/28/content_5722322.htm, 2024-02-06.

[3] 赖一飞, 谢潘佳, 叶丽婷, 等. 我国区域科技创新效率测评及影响因素研究——基于超效率SBM-Malmquist-Tobit模型 [J]. 科技进步与对策, 2021, 38 (13): 37-45.

[4] 马海群, 王今. 基于DEA的政府开放数据网站效率评价 [J]. 数字图书馆论坛, 2016, (6): 2-7.

[5] 韩磊, 胡广伟. 政府数据开放平台建设效率评估及其启示 [J]. 数字图书馆论坛, 2018, (9): 52-59.

[6] 海伦, 邓崧. 基于数据包络法的城市政府数据开放平台效率评估 [J]. 电子政务, 2018, (8): 112-118.

[7] 赵龙文, 洪逸飞, 赵雪琦. 政府开放数据资源利用与创新应用效能评价研究 [J]. 现代情报, 2023, 43 (7): 133-144.

[8] 刘园园, 段尧清, 俞享. 基于超效率SBM-fsQCA的政府数据开放有效路径研究 [J]. 情报科学, 2023, 41 (9): 146-154.

[9] Tone K. A Slacks-based Measure of Efficiency in Data Envelopment Analysis [J]. European Journal of Operational Research, 2001, 130 (3): 498-509.

[10] Caves D W, Christensen L R, Diewert W E. The Economic The-

ory of Index Numbers and the Measurement of Input, Output, and Productivity [J]. Econometrica, 1982, 50 (6): 1393-1414.

[11] Färe R, Grosskopf S, Norris M, et al. Productivity Growth, Technical Progress, and Efficiency Change in Industrialized Countries [J]. The American Economic Review, 1994, 84 (1): 66-83.

[12] Coelli T. A Multi-stage Methodology for the Solution of Orientated DEA Models [J]. Operations Research Letters, 1998, 23 (3): 143-149.

[13] Tobin J. Estimation of Relationships for Limited Dependent Variables [J]. Econometrica, 1958, 26 (1): 24-36.

[14] 李青松, 徐国劲, 邓素君, 等. 基于DEA-Malmquist-Tobit模型的河南省生态效率研究 [J]. 环境科学与技术, 2016, 39 (4): 194-199.

[15] 孙金岭, 朱沛宇. 基于SBM-Malmquist-Tobit的“一带一路”重点省份绿色经济效率评价及影响因素分析 [J]. 科技管理研究, 2019, 39 (12): 230-237.

[16] 马晓君, 李煜东, 王常欣, 等. 约束条件下中国循环经济发展中的生态效率——基于优化的超效率SBM-Malmquist-Tobit模型 [J]. 中国环境科学, 2018, 38 (9): 3584-3593.

[17] 白永平, 张晓州, 郝永佩, 等. 基于SBM-Malmquist-Tobit模型的沿黄九省(区)环境效率差异及影响因素分析 [J]. 地域研究与开发, 2013, 32 (2): 90-95.

[18] 何景师, 王术峰, 徐兰. 碳排放约束下我国三大湾区城市群绿色物流效率及影响因素研究 [J]. 铁路运输与经济, 2021, 43 (8): 30-36.

[19] 查道林, 陈思, 杨茜. “双一流”建设高校科研效率及影响因素实证研究——基于超效率SBM-Malmquist-Tobit模型 [J]. 教育与经济, 2022, 38 (3): 9-16.

[20] Open Knowledge. Global Open Data Index [EB/OL]. <http://index.okfn.org/>, 2024-02-06.

[21] World Wide Web Foundation. Open Government Data Feasibility Studies [EB/OL]. <https://webfoundation.org/our-work/projects/open-government-data-feasibility-studies/>, 2024-02-06.

[22] OECD. Government at a Glance 2015 [EB/OL]. https://www.oecd-ilibrary.org/governance/government-at-a-glance-2015_gov_glance-2015-en, 2024-02-06.

[23] Open Data Institute. Open Data Maturity Model [EB/OL]. <https://theodi.org/insights/guides/open-data-maturity-model-2/>, 2024-02-06.

[24] 华中师范大学信息管理学院. 中国政府开放数据利用研究报告(2022) [EB/OL]. <https://www.fxbaogao.com>, 2024-02-06.

[25] 复旦大学数字与移动治理实验室. 2022中国地方政府数据开放报告(省域) [R]. <http://ifopendata.fudan.edu.cn/report>, 2022.

(责任编辑: 杨丰侨)