

环境评比表彰能否提升公众政府绩效感知？*

杨开峰 刘 杨 文伟铭**

【摘 要】作为非经济领域激励地方政府实现职能高效配置的重要举措和“锦标赛”，“评比表彰”能否提升公众的政府绩效感知？将 2011 至 2015 年主要环境评比表彰结果与中国社会状况综合调查进行匹配，本研究发现：城市在环境评比中获奖有助于提升公众对客观环境质量和政府总体治理绩效的感知，但与公众的政府环境治理绩效感知没有显著关系。此外，获奖数量与公众对环境质量和政府总体绩效的感知呈“倒 U 形”关系，达到临界点前，获奖越多，公众绩效感知越积极；而超出临界点后获得更多奖项反而会使公众的绩效感知更消极。本研究拓展了“客观绩效—主观绩效”的理论联系，也具有现实启示，适度的评比表彰既有“自上而下”引导指挥的工具属性，又有“自下而上”提升公众满意的价值属性。

【关键词】评比表彰模式；主观绩效；政府绩效感知；环境治理；环境锦标赛

一、引言

随着中国特色社会主义进入新时代，经济、社会、文化和生态环境协调发展成为

* 本研究受国家自然科学基金管理科学部专项项目“公共数据开放利用与授权运营理论与制度设计”（项目批准号：72342010）和国家社会科学基金重大专项项目“国家治理体系和治理能力现代化指标体系研究”（项目批准号：17VZL003）的资助。感谢审稿人对本文提出的专业且细致的修改意见，文责自负。

** 杨开峰，中国人民大学公共管理学院、中国人民大学循证治理与公共绩效研究中心教授；刘杨（通信作者），中国人民大学公共管理学院博士研究生，yangliu@mail.nankai.edu.cn；文伟铭，湖南省溆浦县祖师殿镇人民政府选调生。

引文格式：杨开峰，刘杨，文伟铭. 2024. 环境评比表彰能否提升公众政府绩效感知？[J]. 公共管理评论，6(4):163-201.

Cite this article: Yang K F, Liu Y, Wen W M. 2024. Do environmental ratings and recognition enhance perceived government performance? [J]. *China Public Administration Review*, 6(4): 163-201. (in Chinese)

主线方向。增长与发展需要地方政府对激励作出反应,针对中国经济增长的激励因素和制度安排,学者们提出了“财政联邦主义”“晋升锦标赛”和“控制权”等诸多概念和理论。但是,如何调动地方政府在社会治理领域的建设热情?“评比表彰”模式是这一问题的关键解:在社会治理领域设置奖项能指挥和激励下级政府相互竞争和全面协调发展(张咏梅和李秉勤,2010;周晔,2014)。

评比表彰在环境治理领域广泛应用。国家层面典型环境评比有国家节水型城市、低碳试点城市、中国人居环境奖、国家环境保护模范城市、国家森林城市、国家园林城市和全国绿化模范城市等,主要涉及环境规章制度和生态宜居、环境质量、资源节约等定量指标^①。需要说明的是,试点和模范城市也是经历自主申报、组织推荐和专家点评的荣誉称号,旨在发挥鼓励探索和总结经验的典型示范和带头作用^②,有别于实验主义特征的“试验—推广”和先行先试的政策试点(Heilmann, 2008)。

现有研究发现评比表彰可以提升客观治理绩效(刘松瑞等,2020;张天舒和王子怡,2020),但是缺乏对公众主观绩效的研究。“自下而上”的主观绩效既能体现政府环境治理的真实效果,又可体现人民中心发展理念的价值内涵,是不容忽视的问题。本文试图回答:城市在环境评比表彰中获奖能否提升公众对政府绩效的感知?获奖数量越多公众的绩效感知越积极吗?

二、文献回顾与研究假设

(一) 文献回顾

地方政府激励理论包括经济激励、政治激励和反向激励。经济激励典型的有“财政联邦主义”,即地方政府在行政和财政分权的双重激励下塑造自身行为(Montinola et al., 1995);政治激励包括上级政府以经济增长为指挥棒设计的晋升竞赛(周黎安,2007);反向激励通过目标责任制以问责的形式进行,以威慑和压力倒逼地方政府落实上级意图、达到考核基准。

经济领域的激励方式不一定能很好地适应环境和社会等非经济领域的治理要求。随着“五位一体”发展战略的提出,地方政府需要“为和谐而竞争”,需要在社会治理方面投入更多注意力和资源。中央如何调动地方政府发展非经济领域的热情?由于非经济领域的综合性和不易测量性,我国事实上通过“评比表彰”激励和引导地方行为(周晔,2014)。上级针对特定政策领域和目标制定出一套考评下级政

① 限于篇幅,附表报告了环境评奖项目的基本情况与评选标准。

② 发展改革委网站。(2017-01-24)[2024-01-01].发展改革委关于开展第三批国家低碳城市试点工作的通知[EB/OL]. https://www.gov.cn/xinwen/2017-01/24/content_5162933.htm.

府的考核体系和考核标准,对优胜者予以表彰或授予荣誉称号(李振,2014)。新中国成立初期,毛泽东就号召采用“检查评比”的方法推进工作;20世纪90年代以来,为刺激地方政府在公共事业上的投资建设,评比竞赛得到更多应用(张咏梅和李秉勤,2010)。对于评比表彰,一种观点认为其本质上是非经济领域的“晋升锦标赛2.0”(Zhang, 2020),地方政府在评比中角逐,竞赛结果由参赛者的相对位次而不是绝对成绩决定。很多文献也将环境领域评比表彰和绩效竞争称为“生态锦标赛、污染治理锦标赛、环境保护晋升锦标赛、行政竞标制”和“自上而下的标杆竞争”等(文宏和林彬,2020;吴建祖和王碧莹,2022)。另一种观点将其视为“政治达标赛”(王哲,2018),下级政府争取达到上级要求的绩效标准,而不追求绩效最大化和在竞争中胜出。对比周黎安(2007)“锦标赛的五个前提”,本文认为地方政府在经济和非经济领域均受到锦标赛式激励和动员^①,前者按照经济表现排名,后者按照评比表彰项目排名。评比表彰本质上是一种中央政府在社会治理领域设置评比和动员竞争的激励机制。

已有研究对评比表彰模式的制度要件、功能架构、合理性和激励逻辑进行了充分的理论探讨。评比表彰可以合理利用地方资源禀赋,激励地方发挥自主性,打造不同的城市品牌,提供多元化干部晋升甄别机制,创造公众共同目标和群体认同(张咏梅和李秉勤,2010;李振,2014)。但是该模式也存在运动式治理和选择性策略应对的弊病(王哲,2018),增加地方政府超负荷任务压力,使公众处于“失语”状态(周晔,2014)。实证研究更多采用准实验等方法评估评比表彰的政策效应,一是主政官员的晋升,研究发现“创卫”官员的晋升概率高于同僚(刘松瑞等,2020)，“全国文明城市”称号对党政一把手晋升的积极作用甚至比经济绩效和政治经历更具解释力(张天舒和王子怡,2020);二是客观治理目标的实现,实证验证了卫生和文明城市的评选对提升环境质量的显著作用(徐换歌,2020;刘松瑞等,2020);也有研究发现创建环保模范城市反而抑制了环境治理绩效,在地方产业结构转型和技术创新领域未能出现预期的积极效果(王英和唐云,2020)。

总体而言,现有文献更多关注文明城市和卫生城市的评选,对环境领域评比表彰的研究较少;更多检验评比表彰的晋升激励和上级目标的实现(王哲,2018),对公众绩效感知的考察较少——无论是评比中公众“失语”“缄默”和“偏好替代”,还是

① 环境评比表彰是锦标赛的五个前提:(1)锦标赛的主持者和裁判能决定奖品分配:晋升竞争、职位生存、财政激励、声誉激励;(2)存在一种从委托人和代理人角度看都可以衡量的客观竞赛指标:环境指标、评比表彰本身;(3)各参赛主体的竞争成绩是相对可分离和可比较的;城市参评、择优获奖;(4)参赛主体能够在很大程度上控制和影响最终考核绩效;(5)参赛主体之间不容易合谋:朋辈竞争、中央设奖、多次评比。限于篇幅本文不再报告城市获奖数量的地区分布地图,有需要者可联系作者获取。该图佐证了评比表彰仍是“锦标赛”。获奖城市集中在东部和南部,这种跨越五年和七类环境奖项分布的区分度表明,至少到目前为止,评比表彰并不是最终将以多轮次覆盖全部城市的“达标赛”,也没有表现出后续评比以隐性方式降低标准和向未获奖地区倾斜的趋势。

评比后公众感知和评价。公众的绩效感知是一种主观绩效,是基于信念或态度对经历和所接受服务评估的“自我报告”,通常与他们的生活场所或社区有关(Stipak, 1979; Parks, 1984)。将客观与主观绩效联系起来,能够补充或改进客观衡量标准(Baker et al., 1994),更好地理解主观指标的效用,提升政府在公共服务过程和结果中对公众的重视程度(Van Ryzin et al., 2008),这在西方社会有民主政治的象征意义(Brudney and England, 1982),在我国则契合新时代以人民为中心的治理理念,能够避免地方追逐表彰本身而忽视提高人民幸福感和满意度的评比初衷,能够以“客观”技术评价与“主观”社会评价的融合(陈彪和贺芒,2021)推动评奖从工具属性向价值属性转变。

基于此,本文试图拓展以往锦标赛评比的工具性研究,在实证分析评比表彰时融入“人民满意”的价值维度,检验基于客观指标的环境评比是否能够提升公众的主观绩效感知。

(二) 研究假设

评比表彰是社会治理领域“锦标赛”的重要体现,主要目的是提升治理绩效。现有研究基本认为获得环境奖项有助于提升客观环境质量(徐换歌和蒋硕亮,2020;刘松瑞等,2020)。但是,提升公众获得感和幸福感的评比表彰的核心目的和“主观绩效”却很少被检验(Poister and Streib, 1999;文宏和林彬,2020)。实证研究中没有测量主观绩效的统一方式,大都仅涉及对特定服务的具体看法和总体满意度评估(Parks, 1984; Lee and Kim, 2024)。

本文借鉴单一条目和多条目测量,采用层次分析的思路(Van Ryzin, 2004)将公众对政府的绩效感知划分为:客观环境质量感知、政府环境治理绩效感知、政府总体治理绩效感知。环境评比表彰可能在这三个层次的绩效感知有不同影响(假设整体框架见图 1),也规避了本研究潜在的主观绩效共同方法偏差(Wall et al., 2004)。

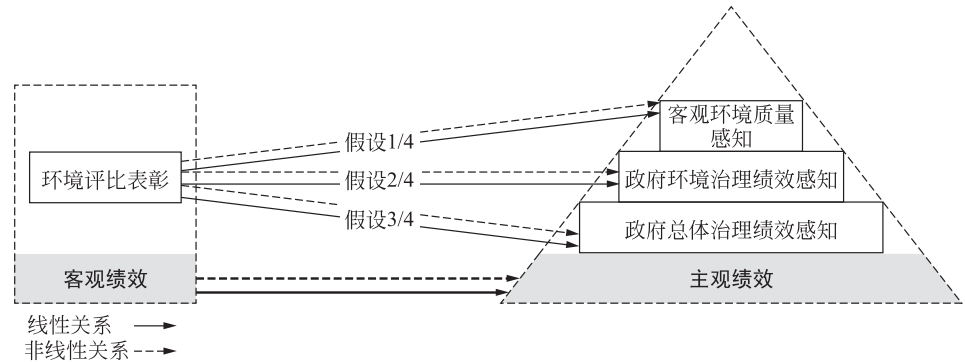


图 1 评比表彰影响主观绩效的假设整体框架

公众对客观环境质量的感知不一定受客观环境质量改善的积极影响。多数客观指标并不是显性和直观的,公众更多基于直接感受而非专业统计数据判断环境质量,而且微小的环境改善可能不足以被公众感知。虽然评选标准可能是隐性和不可感知的,但评比表彰活动本身却是显性的锦标赛。一方面,“全城动员”意味着政府机关、群团组织、企事业单位和街道社区等在统一调度和指挥管理下参与评比,任务、指标和责任也细化到各个层次和岗位(徐换歌,2020)。公众在评比中属于环保规制对象,志愿者或督导员等活跃在社区环境领域,重视和整治“脏乱差”社区与“三不管”死角(陈彪和贺芒,2021),这些如火如荼的身边举措强化了公众对环境质量的感知。另一方面,政府的大力宣传使普通市民得以掌握城市获奖的相关信息(崔慧姝和周望,2018),这种客观绩效“认证”能够为公众提供一个相当可信的质量信号,有效引导公众关注和认可服务提供者对于实现“令人满意的绩效”的明确承诺(Feng et al., 2016; Lee and Kim, 2024)。在网络信息时代,社交媒体进一步拓展并强化了象征性“信号效应”(Symbolic signaling effect)和公众对环境质量的感知(文宏和林彬,2020)。

“创城”是一种激发城市之间“标杆学习”和“标尺竞争”的锦标赛行为(徐换歌,2020),城市在这种竞争氛围和竞赛精神中持续创新工作理念和方法,认真履行环境保护等社会治理和公共服务职责,在横向比较中追赶并达到“超基准水平”(周晔,2014)。竞争胜出使公众明确感知到对邻近城市的“优越感”和“对比效应”,从而感知到并认可本市环境质量的“提升”。由此,本文提出假设1。

假设1: 城市在环境评比表彰中获奖能够提升公众对客观环境质量的感知。

公众对政府环境治理绩效的感知是对政府环境服务质量的外部评价,通常被称为“环境满意度”(唐啸等,2020)。城市治理目标具有模糊属性,信息不完备的公众很难全面真实地感知政府绩效(王伟和王哲,2019)。参评作为一个城市的阶段性中心工作,能够在短时间内高度集中地方政府注意力,从而更容易被公众感知到。“自上而下”的评比表彰压力塑造了集中的环境治理制度框架,包括行政法规、命令、通知、意见和实施办法等公告形式,领导小组、联席会议和指挥部等组织形态(陈科霖和谷志军,2022)。市委书记等主要领导带头调研、监督和检查,频繁在“创城”相关的全体会议、工作推进会、协调会和专题会上强调和表态(陈彪和贺芒,2021),这种积极氛围使治理活动和政府工作更容易被公众感知。正如文献所证实的,强大的环境部门、环境行动主义和多元主体积极参与的环境治理结构等有助于激发公众对政府环境治理绩效的感知(Liu et al., 2018)。

然而,以政治动员、意识形态宣传和准军事化为主要特征的政策工具也可能引发公众的不满。首先,环境评比动员可能干扰当地居民的正常生活秩序,尤其是当基层公务人员迫于评比压力而强制执行或野蛮执法时(Zhang and Li, 2011)。其次,公众可能会认为政府并没有采取实质措施改善环境,而是诉诸一系列策略行为和形象工程。一方面,参评耗费大量时间和资源,需要依据评选标准在环境监察、市容整顿、垃圾处理和城市绿化等领域执行一整套政策,全面动员市民等待中央政府验收,

显性成本可能会招致公众对“形式主义”“装点门面”的批评(李振,2014),尤其是当公众在参与评比表彰中缺乏发言权时,不符合其意愿的评比可能引致“只是政府的事情”和“与自己无关”的不满情绪(崔慧姝和周望,2018)。另一方面,来自职业前景的激励可能会导致政府的治理行为扭曲走样。政府优先或仅完成指标体系范围内的任务(周晔,2014),既未致力于提升环境治理能力,也不关心评选标准之外的公众主观感受。选择性策略或许可能在信息不对称的委托代理中应付上级,但却很难应付对社会事务和公共服务有真切体验的本地民众。最后,周期性评比和限时性治理的绩效依赖短期急剧动员和政府注意力的集中,具有临时性、速成性和不可持续性(王英和唐雲,2020)。非常规形式的评比表彰缺乏长效机制,维持治理绩效的财政和时间成本非常高。一旦在评比中获胜,地方政府持续治理的积极性和注意力会迅速转移,解散“创城”指挥部,返还抽调人员(徐岩等,2015),以至于暂时解决的环境问题很快会在运动后重新出现。而获奖提高了公众对政府环境治理的期望,“预期偏差”和“阈值效应”(Zhang et al., 2022)强化了公众的负面感知。

因此,本文提出关于城市获奖与公众感知环境治理绩效的积极假设 2a 和竞争性假设 2b。

假设 2a: 城市在环境评比表彰中获奖能够提升公众对政府环境治理绩效的感知。

假设 2b: 城市在环境评比表彰中获奖反而会降低公众对政府环境治理绩效的感知。

公众对政府总体治理绩效的感知与环境评比表彰的关系很少被关注。现有对评比表彰“外溢效应”的研究主要包括引领和同侪效应,即某一城市“创卫”能够改善同侪城市的环境(刘松瑞等,2022)。环境评比表彰提升公众对政府总体治理绩效感知“外溢”的可能路径有:其一,环境评比表彰在实际绩效上的“溢出效应”,环境治理影响环境领域之外的目标对象或治理绩效(杨开峰和王璐璐,2023),有研究证实了评比表彰对地方政府其他领域的行为结果和治理绩效的正向激励(周晔,2014)。城市可以凭借参与环境评比向省级政府争取额外的资金冲抵运动开销,也能获得奖励、津贴和专项补贴等(李振,2014),这些使得城市有更强的能力和更集中的资源,可以在参评期间“顺便”解决一些平时“老大难”的问题,提升公众对政府总体治理绩效的感知。其二,城市获得奖项的“信号效应”。信号发送理论认为,接收者收到的信号会影响其对发送者的感知。有区分度的信号具有很强的筛选和“晕轮”效应^①,能够作为总体绩效感知的“启发式”^②,例如环境污染会降低公众对政府的总体满意

① 晕轮效应(Halo effect):又称光环效应,是一种社会认知偏差。原指对某个人某方面特征的认知和评价影响到对此人整体印象的一种心理现象。这里指对城市环境领域治理的评价影响对城市整体治理绩效的评价和感知。这是一种根据某种局部特征推论和扩散到整体特征的“以偏概全”认知方式。

② 启发式:指人们在复杂的决策或判断中倾向于使用启发式推断或相关替换(经验法则),大脑在处理信息时不习惯于费力思考,而是满足于相信出现在脑海中的看似合理的判断,使用相关但更容易的问题来替换信息不完全的复杂问题,这也是认知偏差的原因之一。主要有代表性启发式、可得性启发式、锚定与调整性启发式和情感性启发式。

度和信任度(李雪妮等,2020);环境安全引发的群体性事件会影响公众对政府的总体评价(王凯民和檀榕基,2014)。经过激烈竞争获得的评比荣誉具有代表国家价值分配的信号功能(韩志明和史瑞杰,2015),环境方面的获奖可能使民众对总体绩效的感知更积极(Liu et al., 2018; 张天舒和王子怡,2020)。其三,“五位一体”和多主体治理的“总体效应”。一方面,主观绩效的特点表明,公众更偏好将政府视为一个整体评价其普遍和总体倾向,即便政府只在环境治理等具体领域做出成就(Wang and Gianakis, 1999)。另一方面,虽然公众在评比表彰中参与的深度和频度不足,但参评城市政府在时间和资源有限的情况下进行的“调试性社会动员”仍有助于期待美好生活和公民意识逐渐觉醒的公众积极评价“评比表彰”所隐含的多元参与治理内涵(王诗宗和杨帆,2018)。由此,本文提出假设3。

假设3:城市在环境评比表彰中获奖能够提升公众对政府总体治理绩效的感知。

真实世界普遍存在U形关系(Sturman, 2003),但目前文献通常将市政服务客观指标对主观评价的影响预设为“单调的”(Brudney and England, 1982)。以上述三个假设为基础,本文进一步对评比表彰与主观绩效的非线性关系作出理论假设。

十八大以来,中央对社会治理创新的要求不断提高,政府职能向均衡性和多目标转变。一方面,上级政府效用函数多元化,涵盖经济、社会、文化和生态等多个领域;另一方面,社会领域本身具有对象广泛、事务复杂和需求多样的特征,评比表彰作为指向社会民生的治理工具,单一项目很难实现治理绩效。城市在多个评比活动中开展多赛道竞争、多任务竞逐和多元政绩竞赛更能提升主观绩效(文宏和林彬,2020;王磊和王兰兰,2022)。

但是,各类评比表彰活动存在冗余和重复,过量和过度评比给城市带来的额外负担和行政成本反而会降低感知绩效。改革开放以来,仅中央和各部委开展的评比就有近30项;2006至2009年全国清查各类评比表彰项目148405个,仅保留了4218个^①(王伟和王哲,2019);2009年发布的《关于公布行政等系统中央单位评比表彰达标活动保留项目的通告》在全国共撤销70461个评比表彰项目,总撤销率高达97.76%^②。虽然城市对数量繁多、开展频繁和主题交叠的评比“痴迷”与过度执行有强化综合兼顾和复合治理的意味,但密集应对不同部门检查、考核和指导的评比负担会使城市疲于应付并流于形式(周晔,2014),大量无用的评比活动甚至会挤压政府的日常运转和工作,给公众带来负面感知。基于此,本文第4个假设是:城市在环境评比表彰中的获奖数量与公众绩效感知呈“倒U形”关系,城市获奖越多,公众的绩效感知越积极;但突破临界点后,城市获得更多的奖项反而会导致消极的公众绩

① 国务院办公厅。(2006-12-24)[2023-10-01]。国务院办公厅转发监察部等部门的关于清理评比达标表彰活动意见的通知(国办发[2006]102号)[EB/OL]。 https://www.gov.cn/gongbao/content/2007/content_512702.htm。

② 新华社。(2009-12-30)[2023-10-01]。行政系统中央单位评比达标表彰活动保留项目公示[EB/OL]。 https://www.gov.cn/gzdt/2009-12/30/content_1500155.htm。

效感知。三个子假设是：

假设 4a：城市在环境评比表彰中的获奖数量与公众对客观环境质量的感知呈“倒 U 形”关系。

假设 4b：城市在环境评比表彰中的获奖数量与公众对政府环境治理绩效的感知呈“倒 U 形”关系。

假设 4c：城市在环境评比表彰中的获奖数量与公众对政府总体治理绩效的感知呈“倒 U 形”关系。

三、研究设计

(一) 样本选择与数据来源

数据来源于三方面。其一,个体数据来自 2015 年中国社会状况综合调查 (Chinese Social Survey, CSS2015)。这是 2005 年起在全国范围内(覆盖 31 个省级行政单位和 151 个地级市)采用概率抽样方式进行的大型连续调查项目,样本具有广泛代表性。其二,环境评比结果来自政策文件、政府网站公示和媒体通报等。本文收集了 2011 至 2015 年间国家层面开展的举办主体层级高、合法性权威突出、影响力和代表性强的环境评比表彰获奖城市名单^①,剔除县级及以下获奖单位并运用百度等搜索引擎查漏补缺。其三,城市特征等控制变量来自年鉴数据。人口规模、经济发展水平、产业状况、财政自主度、行政级别等来自《中国统计年鉴》;公园绿地面积、工业废水排放量、工业固体综合利用率和生活垃圾无害化处理率等客观环境指标来自《中国城市统计年鉴》;空气质量指数来自中国空气质量在线监测分析平台^②。上述变量均采用滞后一期的 2014 年数据以避免内生性问题。

(二) 变量选取

1. 因变量：公众的绩效感知

因变量是公众的政府绩效感知,这种主观绩效或满意度被认为是公共领域收集反馈、诊断政策和监测公共服务绩效的重要指标 (Kim and Kim, 2022; Lee and Kim, 2024),而公民调查是获取公民经验感知和公共服务满意度最有效的工具 (Daneke and Klobus-Edwards, 1979; Parks, 1984)。本文从 CSS2015 中获得三个层面的绩效感知因变量。

首先,公众对环境质量的满意度体现了环境评奖及其所带来的客观环境的改善

① 本文选择 2011 年至 2015 年这五年间的获奖情况,这是因为 2010 年之后评比表彰活动才逐渐规范化和制度化,2006 年至 2009 年全国清查和取消了很多评比项目,国务院也取消了很多中央层级的评比表彰项目。2010 年国务院联合中共中央办公厅印发了《评比达标表彰活动管理办法(试行)》(中办发[2010]33 号),真正建立起中央和省两级评比达标活动的审批制度。

② 数据来源于中国空气质量在线监测分析平台: <https://www.aqistudy.cn/historydata/about.php>, 查于 2022 年 3 月 12 日。

在主观层面被公众感知的情况,本文采用“您居住地的环境状况?”进行衡量,受访者评分越高表明其对环境质量越满意,1分表示“非常不满意”,10分表示“非常满意”。其次,由于公众对环境质量本身的满意度并不一定代表对当地政府环境治理的认可程度,本文选取“您认为现住地地方政府在保护环境与治理污染方面的工作做得好不好?”来直接衡量公众对政府环境治理绩效的感知,并对其反向赋值以保证因变量内涵一致,1分为“很不好”,2分为“不太好”,3分为“比较好”,4分为“很好”,分值越高表明公众越认可政府在工作方面。

除此之外,本文还关注环境改善的正外部性对政府绩效感知的“溢出”。在环境评比表彰中获奖对城市整体形象的提升可能使公众对政府治理绩效的认同感和满意度由环境领域延伸到总体治理中。本文选取“总的来说,您认为现住地地方政府工作做得好不好?”和“您认为现住地地方政府在有服务意识、能及时回应百姓诉求方面工作做得好不好?”来衡量公众对政府总体绩效的感知,其内部一致性通过了Cronbach's α ($0.8 > 0.7$) 信度检验。本文在反向赋值的基础上加总得到1至7分的等距定序变量,得分越高表明公众对政府总体治理工作的评价越好。

2. 自变量:城市获奖情况

本文参考Liu et al. (2018)的做法,选择2011至2015年具有重大影响力和代表性的七类主要国家级环境评比表彰项目^①,设置“是否获奖”哑变量作为核心自变量。如果地级市在这些评比表彰中荣获任何一个奖项或称号则取值为1,否则为0。为进一步探究获奖与绩效感知的非线性关系,本文将地级市在观察年间获得环境评比表彰奖项或称号的“获奖数量”设置为第二个自变量。

3. 控制变量

为尽可能规避遗漏变量带来的内生性问题,本文参考现有文献,在个体和城市层面设置一系列控制变量:(1)个体人口与社会经济特征,包括性别、年龄、民族、政治面貌、受教育程度、经济基础、户口性质和居住类型(Fitzgerald and Durant, 1980; Schachter, 2010);(2)个体心理状态和价值取向,包括社会总体评价、社会公平感、幸福感、生活满意度和政治参与度(Helliwell, 2003);(3)城市基本特征,包括经济水平、人口规模、产业结构、财政自主度、行政级别;(4)客观环境质量,包括空气质量^②、

① 在Liu et al. (2018)的基础上,本文选择国家节水型城市、低碳试点城市、中国人居环境奖、国家环境保护模范城市、国家森林城市、国家园林城市 and 全国绿化模范城市七类评比表彰项目;选择依据是:(1)环境领域国家级重大影响力和代表性的奖项、称号和表彰,体现为政府的大力宣传和重视,能被政府官网和人民日报等广泛报道;(2)评选主体主要是城市,而不是区县、项目、流域或园区;(3)首次评比开始的时间早于2011年,且至少持续至2015年。

② 采用空气质量指数(Air Quality Index, AQI)描述空气清洁或污染程度以及对健康的影响,主要通过五个主要污染物标准进行测算:地面臭氧、颗粒物污染(包括PM_{2.5}和PM₁₀)、一氧化碳、二氧化硫和二氧化氮。AQI为连续型变量,数值越大表示污染程度越高,通常0~50为一级,空气质量报告为优;51~100为二级,空气质量报告为良;101~150为三级,空气质量报告为轻度污染;161~200为四级,空气质量报告为中度污染;201~300为五级,空气质量报告为重度污染;300以上为六级,空气质量报告为严重污染。

绿化面积、工业废水排放量、固体废弃物综合利用率与生活垃圾无害化处理率。变量定义和赋值详见表 1。

(三) 模型设定

为了检验假设,考虑到“个体嵌套于城市”的数据结构,避免不区分微观和宏观变量而将调查数据汇总到城市层面(林挺进,2015),本文构建更加准确无偏的多层模型:

$$\text{Level 1: Perceived_Performance}_{ij} = \beta_{0j} + \beta_{1j} \text{individual_CONTROL}_{ij} + \varepsilon_{ij} \quad (1)$$

$$\text{Level 2: } \beta_{0j} = \gamma_{00} + \gamma_{01} \text{PRIZE}_j + \gamma_{02} \text{city_CONTROL}_j + u_{0j} \quad (2)$$

$$\beta_{1j} = \gamma_{10} + u_{1j} \quad (3)$$

其中,Level 1 是个体层模型,表示公众个体对政府绩效的感知差异来自城市平均状况、个体特征在城市内的效应和个体独特性。Level 2 进一步确定城市层差异,即城市平均状况取决于所有城市总平均状况、城市独特性和环境评比表彰效应。Perceived_Performance_{ij} 表示绩效感知,*i* 代表不同的个体样本,*j* 代表不同的城市样本。PRIZE_j 是自变量,表示城市 *j* 在环境评比表彰中获奖的情况,包括“是否获奖”和“获奖数量”。individual_CONTROL_{ij} 和 city_CONTROL_j 分别是个体和城市的一系列控制变量。 γ_{00} 是截距项, ε_{ij} 和 u_{0j} 分别表示个体和城市层面的随机扰动项。本文采用固定效应的随机截距模型,暗含的预设一是个体绩效感知因不同城市而存在明显差异,即“组间异质和组内同质”;二是个体对相同获奖情况产生的绩效感知是一致的。多层模型能充分利用主客观相结合的数据规避内生性、共同来源偏差和“生态谬误”等问题(Diez-Roux, 1998; Robinson, 2011)。

四、实证结果

(一) 描述性统计结果

将本文收集的 285 个城市环境评比表彰数据集与 CSS2015 依据“城市—个体”的原则一对多(1:m)匹配后,本文得到 111 个城市样本和 8538 个个体样本^①。根据表 1,“是否获奖”的均值为 0.495,表明近一半的城市在 2011 至 2015 年的环境评比表彰中获奖;城市最少获奖 0 次,最多 4 次。根据图 2 对获奖分布的可视化,获得国家园林城市(18.9%)、国家森林公园城市(18%)、低碳试点城市(12.6%)称号的城市相对较多;获得全国绿化模范城市(8.1%)、中国人居环境奖(5.4%)、国家环境保护模

① 需要说明的是,城市获奖情况的观测数据为 111 个,本文在描述部分针对这 111 个获奖城市情况进行了统计分析。但由于人口规模、财政自主度、空气质量指数、公园绿地面积、工业废水排放、生活垃圾处理和工业固体废物综合利用等城市层面控制变量的部分缺失,最终进入回归模型中的城市样本为 100 个。

范城市(4.5%)和国家节水型城市(3.6%)的城市相对较少,均不足10%。

表1 主要变量定义、取值和描述性统计表

变量名	变量定义和测量	样本量	平均值	标准差	最小值	最大值
被解释变量						
客观环境质量感知	您居住地的环境状况? 1分表示非常不满意,10分表示非常满意	8530	5.947	2.289	1	10
政府环境治理绩效感知	您认为现住地地方政府在保护环境与治理污染方面的工作做得好不好? 1分表示很不好,4分表示很好	8538	2.569	0.794	1	4
政府总体治理绩效感知	总的来说,您认为现住地地方政府工作做得好不好? 您认为现住地地方政府在有服务意识、能及时回应百姓诉求方面做得好不好? 1~7分,分数越高评价越好	8538	4.126	1.327	1	7
解释变量						
是否获奖	地级市在环境评比表彰中至少获得任何一个奖项或称号取1,否则取0	111	0.495	0.502	0	1
获奖数量	地级市在环境评比表彰中获得奖项或称号的次数	111	0.712	0.898	0	4
个体层面控制变量						
性别	个体为男性取值为1,女性取值为0	8538	0.533	0.499	0	1
年龄	个体的年龄	8537	46.119	13.626	18	70
年龄的平方	同时将“年龄”和“年龄的二次项”作为控制变量	8537	2312.554	1230.696	324	4900
民族	个体是汉族取值为1,是少数民族取值为0	8526	0.921	0.270	0	1
政治面貌	个体是中共党员取值为1,其他(民主党派、无党派人士和群众)取值为0	8516	0.105	0.306	0	1
受教育程度:本科以下	个体本科以下取值为1	8524	0.930	0.256	0	1
受教育程度:本科	个体本科学历取值为2	8524	0.063	0.244	0	1
受教育程度:研究生及以上	个体研究生及以上取值为3	8524	0.007	0.083	0	1
居住类型	在城镇居住取值为1,在农村居住取值为0	8538	0.556	0.497	0	1
户口性质	非农业户口取值为1,农业户口取值为0。取消了农业和非农户口的“居民户口”按取消前的户口性质	8518	0.285	0.451	0	1
年收入水平	2014年个人年收入的自然对数	8263	8.349	3.410	0	15.611

续表						
变量名	变量定义和测量	样本量	平均值	标准差	最小值	最大值
个体层面控制变量						
主观经济水平	您认为您本人的社会经济地位现在在本地大概属于哪个层次? 1 至 5 分分别为下、中下、中、中上、上	8461	2.212	0.891	1	5
社会总体评价	请用 1~10 分来表达您现在对社会总体情况的评价。1 分为非常不好, 10 分为非常好	8526	6.430	1.643	1	10
社会公平感	您觉得当前社会生活中总体上社会公平状况如何? 1 分非常不公平, 5 分非常公平	8521	3.409	1.002	1	5
个体幸福感	总的来说,我是一个幸福的人。1 分非常不同意, 5 分非常同意	8532	3.591	1.002	1	5
生活满意度	总的来说,您对生活的满意度打分。1~10 分,分数越高越满意	8524	6.465	1.931	1	10
公众参与	自 2012 年以来,您是否参与过政治问题的讨论? 您是否愿意参与政治问题的讨论?“参与过”和“愿意参与”取值为 1,没有参与过或没有参与意愿取值为 0	8536	0.235	0.424	0	1
城市层面控制变量						
人口规模	城市常住人口与年末总人口的比值	104	605.098	442.113	77.268	2991.4
经济水平	城市地区生产总值的自然对数	111	10.759	0.346	10.182	11.564
经济增长率	地级市地区生产总值的增长率	111	8.42	1.331	4.9	10.9
产业结构	城市第一产业在地区生产总值中的比重	111	9.82	3.836	0.527	23.124
财政自主度	城市人均财政收入与人均财政支出的比值	104	0.561	0.246	0.132	1.107
行政级别	城市为省会城市或直辖市取值为 1,为一般地级市取值为 0	111	0.18	0.386	0	1
空气质量指数	以城市全年各月空气质量指数(AQI)的均值衡量空气质量,数值越高污染程度越高	108	92.72	30.829	33	194
公园绿地面积	城市绿化情况由人均公园绿地面积(公顷/万人)来衡量	103	2.521	6.6	0.129	51.831
工业废水排放	城市工业废水排放量(吨)的自然对数	104	1.75	2.545	0.101	18.112
生活垃圾处理	城市一般工业固体废物综合利用率	102	92.201	15.74	29.41	157.94
工业固体废物综合利用	城市生活垃圾无害化处理率	103	86.316	16.839	18.97	100

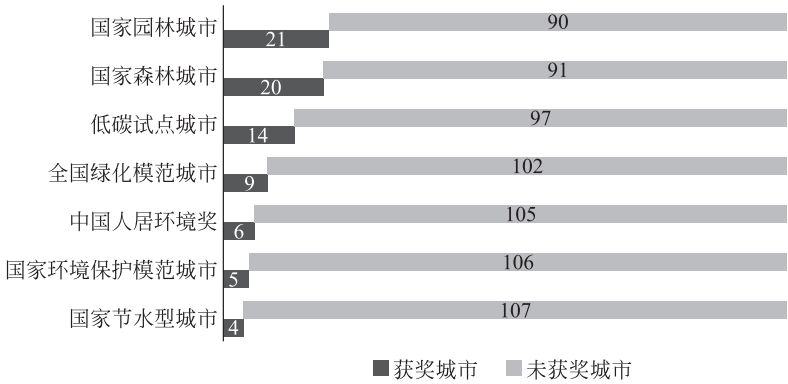


图 2 城市环境评比表彰获得奖项分布情况

(二) 多层模型的多元回归分析

表 2 报告了是否获奖影响公众绩效感知的回归结果。考虑到因变量是多值定序变量,本文同时采用基于 OLS 的多层线性模型(HLM)和基于有序逻辑回归的多层广义线性模型进行多元回归,二者结果一致。为了便于理解,本文以多层线性模型为准解释实证结果。

表 2 “是否获奖”影响公众绩效感知的多层模型

变量	客观环境质量感知		政府环境治理绩效感知		政府总体治理绩效感知	
	多层线性	广义线性	多层线性	广义线性	多层线性	广义线性
	模型 1-1	模型 1-2	模型 2-1	模型 2-2	模型 3-1	模型 3-2
是否获奖	0.198 *** (0.076)	0.212 *** (0.062)	0.069 (0.042)	0.166 (0.104)	0.155 *** (0.047)	0.221 *** (0.070)
个体层面控制变量						
性别	-0.108 ** (0.046)	-0.103 ** (0.044)	-0.001 (0.018)	-0.002 (0.048)	0.055 * (0.029)	0.086 * (0.046)
年龄	0.019 * (0.011)	0.025 ** (0.010)	0.010 ** (0.004)	0.028 ** (0.011)	0.027 *** (0.007)	0.048 *** (0.011)
年龄的平方	-0.000 (0.000)	-0.000 ** (0.000)	-0.000 (0.000)	-0.000 (0.000)	-0.000 *** (0.000)	-0.000 *** (0.000)
民族	-0.129 (0.123)	-0.054 (0.117)	-0.044 (0.050)	-0.114 (0.131)	-0.040 (0.076)	-0.017 (0.124)
政治面貌	0.085 (0.077)	0.041 (0.070)	-0.008 (0.030)	-0.046 (0.079)	0.136 *** (0.047)	0.230 *** (0.076)
本科教育水平	0.107 (0.098)	0.083 (0.090)	-0.070 * (0.039)	-0.171 * (0.100)	-0.010 (0.061)	-0.053 (0.096)

续表						
变量	客观环境质量感知		政府环境治理绩效感知		政府总体治理绩效感知	
	多层线性	广义线性	多层线性	广义线性	多层线性	广义线性
	模型 1-1	模型 1-2	模型 2-1	模型 2-2	模型 3-1	模型 3-2
个体层面控制变量						
研究生及以上教育水平	0.172 (0.258)	0.203 (0.240)	-0.119 (0.102)	-0.290 (0.261)	-0.053 (0.160)	-0.044 (0.253)
居住类型	-0.244 *** (0.054)	-0.262 *** (0.051)	-0.051 ** (0.021)	-0.133 ** (0.057)	-0.061 * (0.033)	-0.105 * (0.054)
户口性质	-0.106 * (0.058)	-0.103 * (0.055)	-0.031 (0.023)	-0.104 * (0.061)	0.021 (0.036)	0.012 (0.058)
年收入水平	-0.000 (0.007)	-0.001 (0.006)	-0.005 * (0.003)	-0.012 * (0.007)	-0.003 (0.004)	-0.005 (0.007)
主观经济水平	0.106 *** (0.027)	0.065 ** (0.026)	-0.012 (0.011)	-0.036 (0.028)	0.028 * (0.017)	0.032 (0.027)
社会总体评价	0.087 *** (0.015)	0.088 *** (0.015)	0.076 *** (0.006)	0.212 *** (0.016)	0.189 *** (0.009)	0.313 *** (0.016)
社会公平感	0.022 (0.024)	0.019 (0.023)	0.128 *** (0.009)	0.332 *** (0.025)	0.352 *** (0.015)	0.561 *** (0.024)
个体幸福感	-0.022 (0.025)	-0.032 (0.024)	-0.005 (0.010)	-0.011 (0.026)	0.024 (0.015)	0.046 * (0.025)
生活满意度	0.624 *** (0.014)	0.710 *** (0.016)	0.023 *** (0.006)	0.060 *** (0.015)	0.046 *** (0.009)	0.083 *** (0.014)
公众参与	-0.103 * (0.054)	-0.112 ** (0.051)	-0.097 *** (0.021)	-0.240 *** (0.055)	-0.175 *** (0.033)	-0.297 *** (0.053)
城市层面控制变量						
人口规模	0.000 (0.000)	-0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)
经济水平	0.176 (0.183)	0.171 (0.152)	0.199 ** (0.100)	0.520 ** (0.248)	0.330 *** (0.114)	0.520 *** (0.170)
经济增长率	0.053 ** (0.026)	0.058 *** (0.022)	0.009 (0.014)	0.029 (0.035)	0.020 (0.016)	0.031 (0.025)
产业结构	-0.022 (0.016)	-0.025 * (0.013)	0.011 (0.008)	0.026 (0.021)	0.022 ** (0.010)	0.032 ** (0.015)
财政自主度	-0.393 * (0.227)	-0.414 ** (0.187)	-0.066 (0.127)	-0.180 (0.312)	0.209 (0.140)	0.352 * (0.210)

续表

变量	客观环境质量感知		政府环境治理绩效感知		政府总体治理绩效感知	
	多层线性	广义线性	多层线性	广义线性	多层线性	广义线性
	模型 1-1	模型 1-2	模型 2-1	模型 2-2	模型 3-1	模型 3-2
城市层面控制变量						
行政级别	0.242 [*] (0.125)	0.256 ^{**} (0.101)	0.014 (0.070)	0.067 (0.173)	0.012 (0.077)	0.046 (0.115)
空气质量指数	0.000 (0.001)	0.000 (0.001)	0.000 (0.001)	0.001 (0.002)	-0.000 (0.001)	-0.001 (0.001)
公园绿地面积	-0.021 ^{**} (0.009)	-0.018 ^{**} (0.007)	-0.008 (0.005)	-0.018 (0.013)	-0.008 (0.006)	-0.011 (0.008)
工业废水排放	0.011 (0.024)	0.008 (0.019)	0.020 (0.014)	0.043 (0.033)	0.013 (0.015)	0.015 (0.022)
生活垃圾处理	0.003 (0.002)	0.003 (0.002)	0.001 (0.001)	0.003 (0.003)	-0.000 (0.001)	-0.000 (0.002)
工业固体废物综合利用	0.000 (0.002)	0.001 (0.002)	0.000 (0.001)	0.000 (0.003)	-0.000 (0.001)	-0.001 (0.002)
对数似然值	-15239.606	-14215.746	-8399.651	-8094.917	-11704.131	-10700.984
个体层方差分量	3.438	—	0.536	—	1.319	—
地级市层方差分量	0.066	0.034	0.029	0.168	0.025	0.049
个体层样本量	7425	7425	7429	7429	7429	7429
地级市层样本量	100	100	100	100	100	100

注:括号内为稳健标准误,* $p<0.1$; ** $p<0.05$; *** $p<0.01$ 。

根据表2,模型1-1显示了城市获奖与公众对环境质量感知的正向关系,“是否获奖”的系数在至少为1%的概率水平上显著为正,假设1得到支持,表明在控制了其他因素的情况下,与没有获奖的城市相比,获奖城市的公众感知到的城市环境质量更好,对当前居住地的环境状况更满意。

模型2-1和2-2检验了城市获奖与公众对政府环境治理绩效感知的关系,变量系数为正,表明获奖可能提升公众对政府环境治理绩效的评价。但是,并没有至少为10%的显著性水平的统计证据表明,在获奖城市居住的公民更倾向于认可政府的环境治理工作,假设2a和2b均未得证。

然而,模型3-1和3-2发现,虽然获奖城市的公众对政府的环境治理工作的满意度并不显著,但是他们对政府总体绩效却更加满意。“是否获奖”系数在至少为1%的水平上显著为正数,与没有获奖的城市相比,在获奖城市居住的公众更认可当地政府总体的治理水平,也更认可政府对公共服务需求的回应。假设3得到支持。部

分控制变量在各模型中也有至少 5% 概率水平的显著性^①。

本文在多层分析之前回归了零模型 (Null Model), 当因变量为客观环境质量感知、政府环境治理绩效感知和总体治理绩效感知时, 城市层面的组内相关系数 (ICC) 分别为 0.053, 0.062 和 0.041, 表明个体感知绩效一定程度上可以由城市方差变动来解释。各个模型均通过了似然比检验, 不同城市间组间差异明显 ($p=0.000$), 因此多层模型更适用于本研究。

进一步地, 城市在环境评比表彰中获奖并不仅是关于“是否”的一次性过程, 假设 4 认为城市在观察年间的获奖数量同样可能影响公众的绩效感知。根据表 3, “获奖数量”在模型 1-1 和 2-1 与模型 5-1 和 6-1 中系数均在至少为 10% 的水平上显著为正, 表明城市获奖数量与公众对环境质量 and 政府总体治理绩效的满意度有正相关关系。但是模型 1-2 和 2-2 与模型 5-2 和 6-2 揭示了这种正相关的临界点。将“获奖数量平方”纳入模型后, “获奖数量”仍然在至少为 1% 的概率水平上显著为正, 但平方项却显著为负, 表明获奖数量与公众绩效感知之间存在“倒 U 形”关系, 获奖数量的增加对公众的环境质量和政府总体绩效满意度的提升作用是边际递减的。到达临界点之前, 城市获奖数量越多, 公众对环境质量 and 政府总体治理绩效的感知越好; 但当获奖数量超出临界点后, 获得更多的奖项反而使公众持更不满意的负面评价和感知, 支持了假设 4a 和 4c。

(三) 从“相关”到“因果”: 关于因果推断的补充检验

为了与 CSS2015 数据库匹配后囊括更多样本城市, 本文采用自变量和因变量同年的截面数据得到相关关系。尽管基于“城市一个体”数据嵌套结构的多层模型隐含因果关系的外生性预设: 城市客观变量影响个体主观变量; 城市宏观变量影响个体微观变量。但本文仍在时间滞后、倾向值匹配、工具变量和排除干扰因素方面补充检验, 以更有信心地将结论从“相关”向“因果”拓展。

1. 因果关系的时间顺序: 自变量滞后一期

主体研究对 2011 至 2015 年环境评比表彰获奖情况的整理包括 2015 年, 微观调查数据也来自当年。本文选取了滞后一期的控制变量, 但并没有处理自变量和因变量在时间上的重合。合理性在于, 城市参与评比和获得表彰并不是“快照式”的瞬间动作, 而是包括充分酝酿、准备和宣传的历时过程。然而, 如果以更严谨和保守的因果逻辑审视“2015 年前的获奖对当年公众绩效感知的影响”, 需要剔除 2015 年的获奖情况重新进行回归。表 4 表明, 即使一定程度上考虑获奖的“时滞”效应, 原模

^① 控制变量的显著性如下。(1) 个体层面: 男性和在农村居住的人有更积极的绩效感知; 主观经济水平更高、对社会总体评价更好、社会公平感和生活满意度更高的人对环境绩效更满意; 对公共事务越关切, 越积极参与环境等治理过程, 对环境和政府工作现状可能越不满意。(2) 城市层面: 经济水平、经济增长率和行政级别更高城市的公众绩效感知更积极; 更大的人均公园绿地面积反而导致公众对环境的满意度降低, 这可能是由于城市绿地面积达到一定阈值后, 植物的挥发性有机物加剧臭氧污染, 降低蒸腾降温等生态功效, 从环境健康科学的视角看, 过度的“绿地暴露”和绿地可得性也可能提升公众的环境污染暴露风险, 从而影响他们对日常空间环境质量的感知。

表 3 “获奖数量”影响公众绩效感知的多层模型^①

变量	客观环境质量感知				政府环境治理绩效感知				政府总体治理绩效感知			
	多层线性		广义线性		多层线性		广义线性		多层线性		广义线性	
	模型 1-1	模型 1-2	模型 2-1	模型 2-2	模型 3-1	模型 3-2	模型 4-1	模型 4-2	模型 5-1	模型 5-2	模型 6-1	模型 6-2
获奖数量	0.072 [*] (0.041)	0.269 ^{***} (0.095)	0.089 ^{**} (0.034)	0.281 ^{***} (0.078)	0.036 (0.023)	0.079 (0.053)	0.089 (0.056)	0.198 (0.131)	0.078 ^{***} (0.025)	0.173 ^{***} (0.059)	0.114 ^{***} (0.038)	0.248 ^{***} (0.088)
获奖数量平方	-0.071 ^{**} (0.031)			-0.070 ^{***} (0.025)		-0.015 (0.017)		-0.039 (0.043)		-0.034 [*] (0.019)		-0.048 [*] (0.028)
个体控制变量	是	是	是	是	是	是	是	是	是	是	是	是
城市控制变量	是	是	是	是	是	是	是	是	是	是	是	是
对数似然值	-15242.04	-15242.03	-14218.01	-14214.39	-8400.30	-8403.05	-8094.89	-8094.47	-11705.43	-11706.92	-10701.32	-10699.93
个体方差分量	3.438	3.438	—	—	0.536	0.536	—	—	1.319	1.319		
城市方差分量	0.071	0.065	0.038	0.031	0.029	0.029	0.169	0.167	0.027	0.025	0.051	0.047
个体样本量	7425	7425	7425	7425	7429	7429	7429	7429	7429	7429	7429	7429
城市样本量	100	100	100	100	100	100	100	100	100	100	100	100

注：括号内为稳健标准误，^{*} $p<0.1$ ；^{**} $p<0.05$ ；^{***} $p<0.01$ 。

① 限于篇幅，本文在报告后续实证表格时均略去控制变量具体结果，可联系作者获取。

表 4 自变量滞后一期的多层模型

变量	客观环境质量感知				政府总体治理绩效感知			
	多层线性	广义线性	多层线性	广义线性	多层线性	广义线性	多层线性	广义线性
	模型 1-1	模型 1-2	模型 2-1	模型 2-2	模型 3-1	模型 3-2	模型 4-1	模型 4-2
是否获奖	0.159 ** (0.076)	0.171 *** (0.063)			0.141 *** (0.047)	0.207 *** (0.070)		
获奖数量			0.187 * (0.097)	0.200 ** (0.080)			0.155 *** (0.059)	0.228 *** (0.087)
获奖数量平方			-0.054 * (0.032)	-0.053 ** (0.027)			-0.028 (0.020)	-0.039 (0.029)
个体控制变量	是	是	是	是	是	是	是	是
城市控制变量	是	是	是	是	是	是	是	是
对数似然值	-15240.756	-14217.638	-15244.037	-14218.091	-11704.989	-10701.523	-11708.015	-10700.816
个体方差分量	3.438	—	3.438	—	1.319	—	1.319	
城市方差分量	0.069	0.037	0.071	0.038	0.026	0.050	0.026	0.049
个体样本量	7425	7425	7425	7425	7429	7429	7429	7429
城市样本量	100	100	100	100	100	100	100	100

注:括号内为稳健标准误,* $p<0.1$; ** $p<0.05$; *** $p<0.01$ 。

型结论仍然基本保持稳健^①,这从“因在前,果在后”的角度为因果关系奠定了基础。

2. 准实验因果推断的检验:倾向值匹配

虽然本文充分纳入控制变量,但多层模型仍然是建立在截面数据上的相关分析,并且混淆变量之间的相关和共线对结果的干扰仍无可避免。为尽可能探究获奖本身与公众绩效感知的“净效应”(获奖城市的平均处理效应而非样本的总平均处理效应)和因果关系,本文采用倾向值匹配在城市是否获奖的观察场域中设计准实验:在获奖城市居住的个体样本为处理组,赋值为 1;在未获奖城市居住的为控制组,赋值为 0。需要说明的是,CSS2015 对各城市个体样本的问卷调查是独立进行的,彼此

① 模型 4-1 和模型 4-2 的二项项虽然在至少为 10% 的显著度水平上并不显著,但是也保持了符号的一致性。二者的 p 值虽然不小于 0.1,但也较小,分别为 0.157 和 0.175。根据对“唯 P 值论”的批判,因此本文根据统计结果进行科学推论时,应该采用估量思考 (Estimation thinking) 而不是非黑即白的二元思考 (Dichotomous thinking),因此本文并没有放弃这个不显著的结果,而是汇报其对应的较小 p 值。

互无影响,这是构建“反事实框架”下准随机实验的重要基础。本文基于 Logit 回归和分析前 T 检验选择合适的协变量进行匹配,最大程度上实现处理组和控制组除“是否得奖”外,其余混淆变量不存在显著的系统性差异。匹配后两组非常接近,标准偏差绝对值全部大幅降低至 4% 以内(大部分小于 3%),各变量均值的组间差异均不显著,共同取值分布广泛重叠。图 3 至图 5 报告了平衡性和共同支撑假设检验结果。

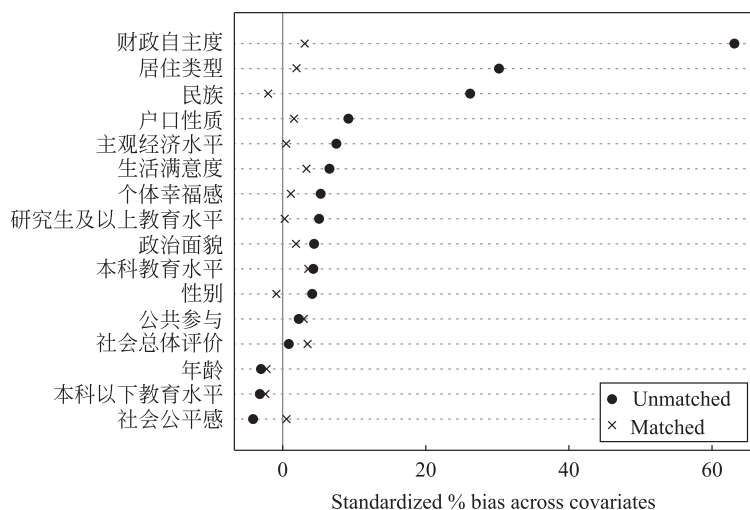


图 3 混淆变量标准化百分比示意 (PSM 匹配结果 a)

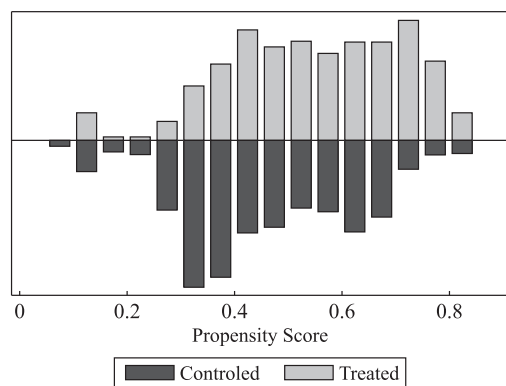


图 4 倾向值的共同取值范围 (PSM 匹配结果 b)

借鉴陈强(2014)的研究思路,如果分别执行不同匹配方法得到的结果较为相似,那么表示倾向值匹配的结果是稳健的,不依赖具体方式而变化。表 5 报告了绝大多数主流匹配方式和结果,平均处理效应(ATT)均在至少 5% 的概率水平上显著为正(t 值 >1.96)。这表明在没有其他因素的系统性差异的前提下,与没有获奖的城市(控制组)相比,获奖城市(处理组)公众更倾向于对环境质量和政府总体治理给予积极评价,获奖显著提升了公众对环境质量和政府总体治理的满意度。这与本文主要结论一致,佐证了稳健性并进一步将相关关系向因果推断延伸。

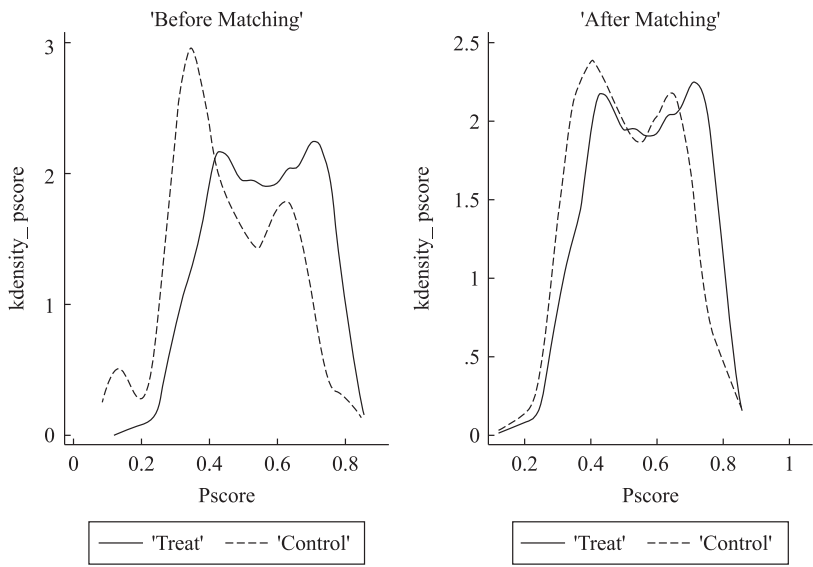


图 5 匹配前与匹配后的核密度情况

表 5 “是否获奖”的倾向值匹配结果

因变量	匹配方法	样本	处理组	控制组	差异	标准误	t 统计量
		Unmatched	6. 030	5. 832	0. 199	0. 052	3. 85***
客观环境 质量感知	Panel 1-1 一对一匹配	ATT	6. 030	5. 756	0. 274	0. 083	3. 30***
	Panel 1-2 k 近邻匹配 ($k=4$)	ATT	6. 030	5. 736	0. 294	0. 068	4. 31***
	Panel 1-3 卡尺内近邻匹配 ($k=4, cal=0. 060$)	ATT	6. 030	5. 736	0. 294	0. 068	4. 31***
	Panel 1-4 卡尺匹配 ($cal=0. 060$)	ATT	6. 030	5. 771	0. 259	0. 058	4. 46***
	Panel 1-5 核匹配	ATT	6. 030	5. 767	0. 264	0. 059	4. 47***
	Panel 1-6 局部线性回归匹配	ATT	6. 030	5. 806	0. 225	0. 083	2. 71***
	Panel 1-7 马氏匹配	ATT	6. 030	5. 880	0. 151	0. 063	2. 41**
	Panel 1-8 k 近邻马氏匹配 ($k=4$)	ATT	6. 030	5. 880	0. 151	0. 050	3. 03***
	Panel 1-9 样条匹配	ATT			0. 242	0. 049	4. 97***
政府总 体治理 绩效感知		Unmatched	4. 206	4. 010	0. 197	0. 030	6. 66***
	Panel 2-1 一对一匹配	ATT	4. 206	4. 046	0. 160	0. 047	3. 41***
	Panel 2-2 k 近邻匹配 ($k=4$)	ATT	4. 206	4. 049	0. 157	0. 039	3. 99***
	Panel 2-3 卡尺内近邻匹配 ($k=4, cal=0. 06$)	ATT	4. 206	4. 049	0. 157	0. 039	3. 99***
	Panel 2-4 卡尺匹配	ATT	4. 206	4. 034	0. 173	0. 033	5. 16***

续表

因变量	匹配方法	样本	处理组	控制组	差异	标准误	<i>t</i> 统计量
政府总体治理 绩效感知	Panel 2-5 核匹配	ATT	4.206	4.034	0.171	0.034	5.02***
	Panel 2-6 局部线性回归匹配	ATT	4.206	4.038	0.168	0.047	3.59***
	Panel 2-7 马氏匹配	ATT	4.206	4.066	0.141	0.036	3.90***
	Panel 2-8 k 近邻马氏匹配 ($k=4$)	ATT	4.206	4.066	0.141	0.030	4.64***
	Panel 2-9 样条匹配	ATT	4.206		0.133	0.033	4.10***

注: * $p<0.1$; ** $p<0.05$; *** $p<0.01$ 。

3. 互为因果的内生性问题检验: 工具变量

环境评比要求将当地居民对环境或政府的满意度作为重要参考^{①②③},这意味着城市获奖与公众的环境质量和政府总体绩效感知的共变可能存在逆向因果。一方面,获奖可能释放城市声誉的“信号”从而提升公众对环境质量和总体绩效的感知;另一方面,公众对环境和政府的积极评价本身也可能是城市获奖的前提和依据。虽然看似可以为因果倒置的内生性辩护,例如各项评比标准事实上并没有体现明确可量化的社会满意度指标,既有研究也佐证缺乏公众评价的“上评下迎”是当前评比表彰的突出问题(崔慧姝和周望,2018),但是,本文认为排除内生性更有效的方法是寻找恰当的工具变量。

借鉴已有文献(Chen et al., 2018; 陈诗一和陈登科,2018),本文选取城市政府工作报告中环境相关词汇的频率作为获奖情况的工具变量。一是与自变量高度相关,政府工作报告是全面体现政府注意力和认知集中度的“施政纲领”(吴建祖和王碧莹,2022)。环境治理在日常工作中难以进入优先序列,只有参评时才会获得足够的重视(李振,2014),对环境提及更多的城市更可能参评和获奖。二是与随机扰动项无关,政府工作报告发布于年初(1月或2月),公众绩效感知数据收集于同年中下旬,无法反向影响已经确定的工作报告,时间上的“滞后”是保证工具变量外生的重要方式(陈强,2014)。并且,政府工作报告的内容和措辞是城市层面反复酝酿和修改的结果,不直接受个体评价和感知影响。本文手工收集111个城市2015年政府工作报告文本并进行分词处理,在输出词频排名前500的名词中统计环境词汇的频数

① 国家发展改革委。(2017-01-07)[2022-02-03]. 国家发展改革委关于开展第三批国家低碳试点城市工作的通知(发改气候[2017]66号)[EB/OL]. https://www.ndrc.gov.cn/xxgk/zcfb/tz/201701/t20170124_962888.html.

② 住房和城乡建设部。(2022-01-06)[2022-02-03]. 住房和城乡建设部关于印发国家园林城市申报与评选管理办法的通知(建城[2022]2号)[EB/OL]. https://www.gov.cn/zhengce/zhengceku/2022-01/14/content_5668177.htm.

③ 山西省林业和草原局。(2021-12-27)[2022-02-03]. 全国绿化评比表彰实施办法[EB/OL]. http://lcj.shanxi.gov.cn/lczl/zthc/zllh/zhxx/202112/t20211227_4259672.html.

和比重(环保提及率高的城市见图6)。参考现有研究,本文在准确保留源文本语义的同时选取更丰富全面的词汇捕捉政府的环境注意力和治理力度^①。ROSTCM6.0 支持了文本分析。

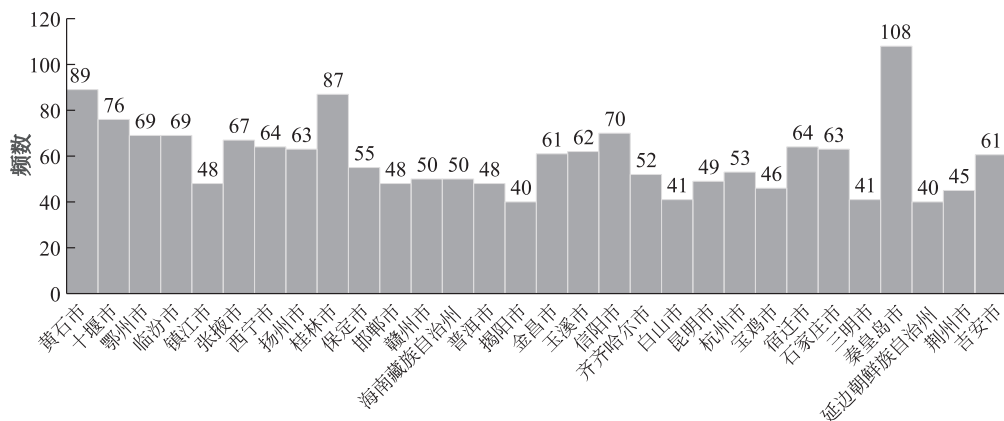


图6 政府工作报告中环境保护提及频率前30的城市频数分布情况

环境工作人员增长率是第二个工具变量(Dasgupta et al., 2000),通过《中国城市统计年鉴》中按行业分组的“水利、环境和公共设施管理业”从业人数衡量。环境人员增长的城市更可能在环境治理方面表现突出,积极参与评比并胜出,文献中也经常将环保人员数量作为政府环境治理的代理变量(Keller and Levinson, 2002)。考虑到单年度数据噪声和工具变量外生性,本文滞后一期计算2010至2014年五年间的平均增长率(图7汇集了各城市环境工作人员人数变化趋势)。“滞后”保证了环境工作人员平均增长率外生于2015年的个体绩效感知;且环境人员增长率涉及定编定岗定员等人事行政,主要由评比表彰的环境治理需要和环保注意力决定,很难被微观个体直接感知或影响。

市委书记作为总揽城市全局的“第一领导”,能在非经济领域评比表彰激励下直接影响城市准备和争取环境荣誉。他们成立临时指挥部,担任指挥长,集中管理并统筹下级部门,使其对评比工作“狠抓落实”并“确保到位”(徐换歌,2020)。年龄已被证实是影响市委书记响应激励和开展治理的重要因素(徐现祥和王贤彬,2010),有研究验证较年轻的书记职业生涯更长,更愿意积极参与节水型城市评选、节能减排和空气污染治理等(韩超等,2016;姚洋等,2020);也有研究认为更年轻的书记晋升动机更强,更愿意积极响应经济增长竞赛而忽略柔性的环境治理(王哲,2018)。

① Chen et al. (2018)采用环境、能耗、污染、减排和环保五个词语总字数占全文总字数的比例作为政府环境治理的代理变量。陈诗一和陈登科(2018)采用环境保护、环保、污染、能耗、减排、排污、生态、绿色、低碳、空气、化学需氧量、二氧化硫、二氧化碳、PM10、PM2.5等词语。本文在此基础上选取的词语包括:环境保护、环保、污染、能耗、减排、排污、节能、节水、生态、绿色、低碳、空气、化学需氧量、二氧化硫、二氧化碳、PM10和PM2.5。

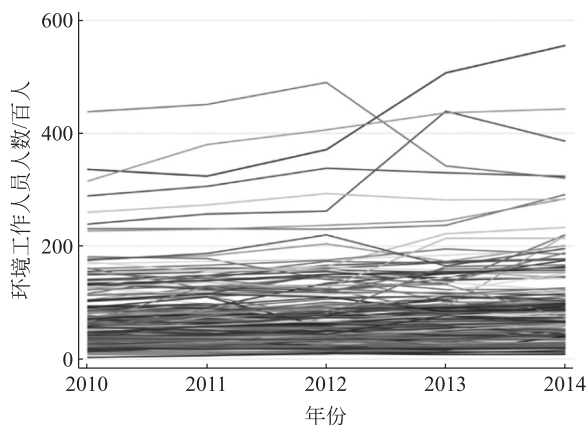


图7 五年间各城市环境工作人员人数变化趋势汇总情况

无论如何,市委书记的年龄与环境评奖情况密切相关。研究普遍选择年龄作为工具变量(Angrist and Krueger, 1991; 谭志雄等,2018),官员年龄只取决于上级综合考察任命和时间自然增长,外生于个体主观绩效感知。

表6报告了工具变量两阶段回归结果。第一阶段回归显示,政府工作报告中环境词汇占比、环境工作人员增长率与市委书记年龄均在至少1%的统计水平上显著为正。这符合本文的预期,工作报告中环境词汇比重越大,环境工作人员增长率越高,表明政府越重视环境工作和越倾向于在环境评比表彰中竞争胜出;市委书记年龄越大也越倾向于在环境评比中参与和获胜。考虑到书记年龄的影响渠道还可能是“经济绩效”,为避免工具变量出现在解释变量的扰动项中,本文也将“经济增长率和市委书记年龄的交互项”作为工具变量进行控制和排除,显著为正的结果再次印证:年龄更大的书记可能不再有发展经济以谋求晋升的强烈动机,更偏好环境治理和评比表彰,以追求地方的长远利益(林挺进,2015);经济增长率更高地区的“大龄”书记更无须集中全部力量发展经济,而是为环境治理分配更多注意力,即在经济与环境多任务竞争模式中,首要任务(经济)绩效变动较小或竞争局势较为固定(经济增长率高+书记年龄大)时,次要任务(环境治理)上的竞争才更加激烈(Holmstrom and Milgrom, 1991; List and Sturm, 2006)。这在一定程度上表明环境治理是与经济增长不同的激励,环境表彰和生态文明建设不能为经济发展劣势地区提供“弯道超车”的机会(刘松瑞等,2020)。第二阶段回归显示,使用工具变量两阶段模型控制反向因果的内生性问题后,本文主要结论仍然成立。市委书记回应环境激励、城市环境从业人员增长、政府工作报告重视时,城市更可能参与环境评比并胜出,公众的绩效感知也更积极。

此外,本文也定量检验了工具变量的有效性。Durbin-Wu-Hausman 检验的显著性表明环境获奖具有内生性,工具变量方法合理;Kleibergen-Paap LM 检验统计量 p 值均接近 0.00, Hansen J 统计量 p 值远大于 0.1,表明工具变量不存在不可识别和识别过度问题,验证了工具变量的相关性和外生性;Kleibergen-Paap Wald F 统计量均大于 10%的临界值,拒绝了弱工具变量假设。

表 6 工具变量两阶段估计的城市获奖影响公众绩效感知^①

变量	客观环境质量感知				政府总体治理绩效感知			
	模型 1	模型 2	模型 3	模型 4	模型 5	模型 6	模型 7	模型 8
是否获奖	0.592 ^{***} (0.184)		0.544 ^{***} (0.177)		0.341 ^{***} (0.104)		0.324 ^{***} (0.1002)	
获奖数量		0.267 ^{***} (0.089)		0.246 ^{***} (0.085)		0.178 ^{***} (0.0497)		0.168 ^{***} (0.048)
控制变量	是	是	是	是	是	是	是	是
样本量	7015	7015	7015	7015	7019	7019	7019	7019
调整 R 方	0.1307	0.1317	0.1321	0.1327	0.1552	0.1549	0.1558	0.1556
第一阶段回归								
环境同汇比重	39.301 ^{***} (1.494)	69.564 ^{***} (3.206)	40.354 ^{***} (1.502)	71.835 ^{***} (3.219)	39.278 ^{***} (1.493)	69.568 ^{***} (3.204)	40.330 ^{***} (1.501)	71.836 ^{***} (3.217)
环境工作人员增长率	0.178 ^{***} (0.013)	0.491 ^{***} (0.025)	0.183 ^{***} (0.012)	0.501 ^{***} (0.024)	0.178 ^{***} (0.013)	0.491 ^{***} (0.025)	0.183 ^{***} (0.012)	0.501 ^{***} (0.024)
市委书记年龄	0.029 ^{***} (0.002)	0.058 ^{***} (0.003)	0.029 ^{***} (0.002)	0.058 ^{***} (0.003)				
市委书记年龄×经济增长率			0.004 ^{***} (0.0002)	0.008 ^{***} (0.0004)			0.004 ^{***} (0.0002)	0.008 ^{***} (0.0004)
Kleibergen-Paap LM rk 统计量	641.80 ^{***}	611.67 ^{***}	676.64 ^{***}	633.71 ^{***}	641.44 ^{***}	612.13 ^{***}	676.24 ^{***}	634.16 ^{***}
Kleibergen-Paap Wald rk F 统计量	364.05	337.52	397.12	365.71	363.67	337.76	396.65	365.93

① 需要说明的是,所有进入第一阶段回归的控制变量均被纳入第二阶段回归模型中,但限于篇幅,本文在报告两阶段回归表格时均略去控制变量的具体结果。

续表

变量	客观环境质量感知				政府总体治理绩效感知			
	模型 1	模型 2	模型 3	模型 4	模型 5	模型 6	模型 7	模型 8
第一阶段回归								
Stock-Yogo 临界值(10%)	9.08	9.08	9.08	9.08	9.08	9.08	9.08	9.08
Durbin-Wu-Hausman 统计量	5.60 ^{**}	5.50 ^{**}	4.72 ^{**}	4.73 ^{**}	3.33 [*]	4.25 ^{**}	2.96 [*]	3.77 [*]
Hansen J 统计量	2.188	3.430	1.097	2.180	3.631	1.569	3.915	1.960
	(p=0.3349) (p=0.1799) (p=0.5779) (p=0.3363) (p=0.1628) (p=0.4564) (p=0.1412) (p=0.3753)							
样本量	7015	7015	7015	7015	7019	7019	7019	7019

注:括号内为稳健标准误, * p<0.1, ** p<0.05, *** p<0.01。

4. 排除干扰因素的替代解释：“创文”与“创卫”

“全国文明城市”是目前国内城市评选中规格最高、分量最重、获评难度最大的荣誉和最具价值的城市品牌和综合称号(文宏和林彬,2020;王磊和王兰兰,2022);“国家卫生城市”同样是非经济领域典型评比项目。2015 年恰好处于第四届全国文明城市^①和国家卫生城市^②名单公示的关键时间窗口,公众的绩效感知是否会受到这两大同类政策的干扰?公众对环境质量和政府总体绩效的积极感知是否出于文明城市或卫生城市称号对城市知名度和美誉度的提升?为了排除干扰政策的替代解释(elimination of spuriousness),本文在基准模型中分别增加 2015 年入选全国文明城市和国家卫生城市的哑变量,并进一步控制城市 2011 至 2015 年历届“创文”和“创卫”的情况,表 7 报告的结果与主要结论高度一致。

(四) 稳健性检验

为充分缓解样本选择和研究方法等不利影响,本文进行了更多稳健性检验^③。第一,剔除直辖市。考虑到环境评比的对象是城市,本文在与 CSS2015 匹配时将直辖市也纳入城市样本。虽然将“高行政级别”设置为控制变量,但直辖市仍然具有特殊性、可比性弱的特点和更容易率先获奖等信息资源和政治优势(刘松瑞等,2020),剔除后的重新检验并没有显著改变结果。第二,改变研究方法。本文将定序变量“降维”为二分类变量,忽略各评分间的细微差异^④,选取多层逻辑回归重新检验(Liu et al., 2018),结论不随方法而改变。

五、结论与讨论

习近平总书记强调:“我们党的执政水平和执政成效必须而且只能由人民来评判。”^⑤对客观治理实践和主观感知的关系的探究是重要的和欠缺的(Parks, 1984)。本文以城市环境评比为切口,建立 2015 年中国社会状况综合调查与城市获奖情况匹配的数据集,探讨城市获奖如何影响公众的主观绩效感知。

研究有如下发现。第一,城市在环境评比表彰中获奖显著提升该城市居民对环境质量的感知。公众可能会将获奖视为服务质量可信赖的声誉信号和“认证”(Teodoro and Hughes, 2012)。这在一定程度上验证了客观绩效与主观感知的正向关系。

① 新华社. (2015-01-30)[2023-09-01]. 中央文明委公示第四全国文明城市(区)、文明村镇、文明单位候选名单[EB/OL]. https://www.gov.cn/xinwen/2015-01/30/content_2812168.htm.

② 卫计委网站,中央政府门户网站. (2015-12-08)[2023-09-01]. 全国爱卫办公布 2015 年拟命名国家卫生城市(区)名单[EB/OL]. https://www.gov.cn/xinwen/2015-12/08/content_5021061.htm.

③ 限于篇幅,在此不再报告稳健性检验的各项具体结果,可联系作者获取。

④ 重新编码以“是否满意”为标准对定序变量降维,将“不太好”“很不好”和 6 分以下的绩效感知界定为“不满意”,赋值为 0,其余正面评价为“满意”,赋值为 1。除此之外,文献中对定序变量做因变量的处理方式还有两种:(1)近似地视为连续变量采用线性模型,尤其是 1~10 分量表时;(2)选取更适合定序变量的有序逻辑回归。这是本文主体研究中所采用的方法。

⑤ 习近平. 2013-12-27(02). 在纪念毛泽东同志诞辰 120 周年座谈会上的讲话[N]. 人民日报.

表 7 排除“创文”和“创卫”干扰因素的多层模型^①

变量	客观环境质量感知				政府总体治理绩效感知			
	模型 1	模型 2	模型 3	模型 4	模型 5	模型 6	模型 7	模型 8
是否获奖	0.188 ^{**} (0.077)		0.195 ^{**} (0.0787)		0.149 ^{***} (0.0481)		0.140 ^{***} (0.0476)	
获奖数量		0.236 ^{***} (0.098)		0.264 ^{***} (0.0998)		0.150 ^{**} (0.0610)		0.144 ^{***} (0.0603)
获奖数量平方		-0.057 [*] (0.032)		-0.0689 ^{**} (0.0325)		-0.0245 (0.0200)		-0.0247 (0.0197)
2014 年全国文明城市	-0.179 (0.110)	-0.166 (0.114)			-0.101 (0.0691)	-0.112 (0.0708)		
2015 年国家卫生城市	0.073 (0.089)	0.069 (0.089)			0.0315 (0.0555)	0.0375 (0.0554)		
2011—2015 年全国文明城市			-0.0290 (0.0975)	-0.0171 (0.0983)			-0.0982 [*] (0.0590)	-0.105 [*] (0.0593)
2011—2015 年国家卫生城市			0.0523 (0.0821)	0.0409 (0.0821)			0.0436 (0.0497)	0.0424 (0.0495)
个体控制变量	是	是	是	是	是	是	是	是
城市控制变量	是	是	是	是	是	是	是	是
对数似然值	-15240.622	-15243.362	-15242.359	-15244.89	-11706.537	-11709.071	-11706.397	-11709.063
城市方差分量	0.0626	0.0626	0.0681	0.0673	0.0250	0.0243	0.0241	0.0236
个体方差分量	3.4383	3.4384	3.4378	3.4380	1.3193	1.3193	1.3195	1.3195
个体样本量	7425	7425	7425	7425	7429	7429	7429	7429
城市样本量	100	100	100	100	100	100	100	100

注:括号内为稳健标准误,* $p<0.1$,** $p<0.05$,*** $p<0.01$ 。

① 限于篇幅,本文只报告了多层线性模型的回归结果,略去结论一致的广义多层模型,可联系作者获取。

然而,第二个结论表明城市获得环境奖项并没有显著提升公众对政府环境治理绩效的感知。这是令人困惑的发现,公众并没有将获得表彰和主观的环境质量改善归因于政府环境治理中的成就,目前评比表彰无法保证居民对政府环境工作的积极感知,主观与客观绩效呈现低一致性和相关性(Stipak, 1979; Lee and Kim, 2024),这也通常是主观测量被质疑为“错误”的原因(Licari et al., 2005)。可能的解释是,首先,高度汇总的宏观数据与具体的微观感知之间存在“温差”。评比立足整个城市辖区内的平均状况和统一指标,隐性地抑制和忽略了不同街区具体的需求与供应差异,而公众所享受和评价的服务实际上分散且有限,满足一组人口结构的环境服务可能无法满足其他人群(Schachter, 2010)。其次,主观和客观绩效测量的可能不是同一件事,这与“知觉显著性”(perceptual salience)类似,一方面,公众对政府环境治理感知的“低显著性”使其并没有意识到环境质量的改善是由于政府的治理(Stipak, 1979),政府投入与公民感知之间有许多中间步骤(Parks, 1984)。公众虽然感知到客观环境质量的提升,但只要服务质量保持在可接受范围内,他们并不关心政府的投入产出等环境治理行为。街道清洁等环境改善清晰可见,而垃圾收集吨数、巡逻频率和业务支出等客观指标并不被公众评价所关注(Brudney and England, 1982; Van Ryzin et al., 2008),尤其在不常接触有关部门时,公众对治理工作的改善缺乏了解,也未必有“感知政府努力”的能力和评价市政服务的基础知识,“外行认知”使其忽略政府在城市获奖和提供环境公共产品与服务中的作用(Rosentraub et al., 1979),以至于“政府在客观绩效上的巨大进步通常对公民的主观评价似乎没有什么影响”(Grimmelikhuijsen and Meijer, 2015)。另一方面,公众没有意识到政府环境服务的真实水平。没有直接接触政府环境治理的公民根据媒体报道等进行主观评价(Schachter, 2010),更可能体现为不分青红皂白的“负性偏见”^①或“代理受害”^②(Skogan and Maxfield, 1981),特别地,评比表彰中声势浩大的运动式治理更可能给公众带来“政绩工程”式环境工作的消极感知。最后,评比表彰的信号会提升公民期望,形成“预设满意度”(assumed satisfaction)与“表现满意度”(demonstrated satisfaction)的差距(Van Ryzin, 2004; James, 2009),使公众更容易因看到政府环境治理的短板而降低满意度。

第三,城市环境获奖显著提升公众对政府总体绩效的感知。这意味着环境治理本身的正外部性和溢出性,也符合评比表彰成为综合激励的长期愿景。政府以环境评比表彰为抓手,同步在市政设施和社会保障等与人民需求紧密相关的公共服务和

① 负性偏见(negative bias):人们对信息的处理具有“不对称性”,对所接收信息中的负面因素更为敏感,对负面信息的关注、学习和使用远远多于正面信息。负面信息更容易被感知,消极事件比非消极事件引发的反应更快更显著。

② 代理受害(vicarious victimization):指自身没有受害的公民仍然可能通过从朋友、邻居或媒体那里听到他们附近的犯罪活动而意识到他们周围的犯罪活动。本文指自身的生活没有因为城市参与环境评比表彰而受到影响的公民可能因为媒体报道和亲朋好友的经验获得关于政府环境治理的替代信息,形成负面认知。

民生环节取得总体成就(李振,2014)。本文的发现补充了主观总体绩效与客观绩效正相关关系的坚实证据(Parks, 1984)。为什么主客观绩效的低一致性在总体绩效方面失效了?为什么环境评比表彰不能提升公众对环境治理的满意度反而提升了对政府的总体满意度?这是由于客观指标能够较好地衡量环境治理等“部分的绩效”(Ailawadi et al., 2004);而主观感知是边界不清晰的、整体概括的,往往反映比决策更广泛的总体绩效(Wang and Gianakis, 1999),民众更能感知到市政府作为一个整体的普遍倾向而非其在环境治理等高度受限领域的绩效水平。尤其在缺乏了解时,环境评比表彰提供的政府整体形象和声誉的“启发式”能通过“晕轮误差”等思维捷径使公众对环境治理的感知与政府总体评估混为一谈(Borman, 1975),形成从“对环境质量满意”到“对政府总体绩效满意”的以偏概全式认知扩散。

第四,获奖与环境质量和总体绩效感知呈“倒U形”关系。城市获奖数量的增加一开始显著提升公众的绩效感知,这可能是由于不同评选在指标上各有侧重,更多获奖表明政府加强节水、森林、人居和生态等更多环境维度的投入,提升全方位环境质量并释放综合治理信号。但是,获奖的积极效用是有限度的,过度参与、重复评选和冗余表彰有适得其反和过犹不及的后果,如果城市总是热衷于各类环境评比,会形成“追求‘面子工程’”“好大喜功”和“劳民伤财”等不顾公众实际需求的形式主义形象(徐岩等,2015),公众也会长时期面临错综复杂的评比指标带来的严格环境规制、监管执法和遵从负担等(Zhang and Li, 2011),这塑造了公众的负面感知。曲线关系的发现拓展了公民评价与市政服务实际绩效线性关系的讨论(Brudney and England, 1982)。

本文对引导评比表彰符合公众期待具有理论意义和实践启示。首先,本文验证了客观与主观绩效积极但有限的关系,以中国城市的证据填补了二者的理论空隙(Van Ryzin et al., 2008),回答了“个人层面如何评估已被认可的服务提供者?”,丰富了“认证—满意度联系”(accreditation-satisfaction link)研究(Lee and Kim, 2024)。其次,评比表彰应更多纳入主观绩效和满意度指标,尊重人民主体地位并着力解决“急难愁盼”的民生问题,将民众感受和需求放在环境治理实践的第一位,以补充客观绩效(Kim and Kim, 2022),并且根据公众心理活动特征培育具有评价能力的“批判性公民”,畅通信息公开和参与渠道,真正实现“自上而下国家评比”和“自下而上社会评价”的良性互动融合。这既体现了另类公众参与和对民主规范与公共价值的承诺,也是绩效合法性的重要来源和公共问责制的组成部分(Daneke and Klobus-Edwards, 1979)。最后,评比表彰未必是完全的客观绩效,而是人口特征中立和不亚于客观指标的“上级主观评价”(Schachter, 2010),是上级发挥指挥棒激励效用的“环境治理锦标赛”,可以引导地方政府以公共服务和民生福祉而不是经济增长与政治晋升为导向,以长期改善和公众满意而不是集中作为和短期成效为目标。本文的发现在理论上将“锦标赛”在非经济领域进行了拓展和延伸(Zhang, 2020)。环境治理作为一个缩影,先进做法可以推广到社会多领域以全方位促进主观绩效和公众参与,实现公共服务的共同生产而不是政府单方面提供。

本研究也具有局限性,一是截面数据问题,虽然各种准实验思路的补充检验在较大程度上证实了因果性,但仍缺乏面板数据的“前测”,未来期待运用双重差分或调查实验等方法更准确地估计获奖净效应。二是忽视获奖时效性,全部奖项以同等权重进入模型,结合评比表彰的间歇性激励特征导致地方政府治理投入和绩效的不稳定性,不同年份获奖可能意味着不同的绩效感知。三是样本局限性,匹配后仅有约 100 个城市,虽然没有样本选择偏误的内生性问题,但仍然存在信息损失。四是对行为公共管理的关注有限,本文基于公众心理态度对主观感知进行了有益探索,区别于传统产出和效果的绩效衡量,未来可以借鉴内隐满意度测量等进一步关注公众认知和情绪体验。此外,什么类型和特征的环境评比更能提升绩效感知?评比表彰作用机制和路径是什么?评比表彰与环保督察的关系如何?这些问题有待进一步探讨。

附表

评奖项目	组织方	首次评选时间	主要指标	其他说明
国家节水型城市	住建部、国家发改委	2002	基本条件、基础管理指标、技术考核指标(包括综合节水、生活节水、工业节水、环境生态节水指标)	每两年开展一次,奇数年为申报年,偶数年为评选年,有效期五年
国家环境保护模范城市(现生态环境部)	环保部	1997	基本条件、社会经济、环境质量、环境建设、环境管理	考核跨度为近三年
国家低碳试点城市	国家发改委	2010	领导高度重视、明确试点目标、能发挥绿色低碳发展的示范带头作用、编写并提交《低碳试点工作初步实施方案》情况(编制低碳发展规划、建立控制温室气体排放目标考核制度、积极探索创新经验和做法、提高低碳发展管理能力)	
中国人居环境奖	住建部	2001	基本条件、生态宜居、健康舒适、安全韧性、交通便捷、风貌特色、整洁有序、多元包容、创新活力、资源节约、社会和谐等	每两年开展一次,偶数年为申报年,奇数年为评选年,有效期五年

续表

评奖项目	组织方	首次评选时间	主要指标	其他说明
国家园林城市	住建部	1992	基本条件、综合管理、绿地建设、规章制度、绿地规划、生态宜居、节能减排、健康舒适、安全韧性、特色风貌、特色园林文化保护传承、“全国城市园林绿化管理信息系统”更新基础工作	每两年开展一次,偶数年为申报年,奇数年为评选年
国家森林公园城市	全国绿化委员会 国家林业局	2004	组织领导、管理制度、森林建设(包括综合指标、覆盖率、森林生态网络、森林健康、公共休闲、生态文化、乡村绿化、生态福利)	
全国绿化模范城市	全国绿化委员会	2003	各级领导对城市绿化工作的重视程度、绿化委员会及其办公室机构建设情况、城市绿化规划计划制定情况、义务植树开展情况、造林绿化率和保存率情况、城乡绿化建设情况、绿色通道建设情况、单位庭院和居住区绿化情况、绿化成果保护和管理工作、古树名木保护工作、绿化宣传教育情况	

参考文献

陈彪, 贺芒. 2021. 复合治理视域下的地方“评比表彰”执行逻辑——以 S 省 D 市四城同创为例[J]. 北京理工大学学报(社会科学版), 23(6): 110-118.

Chen B, He M. 2021. Local governments’ appraisal and commendation from the perspective of compound governance—An example of co-creation of four models of city in D city of S province [J]. *Journal of Beijing Institute of Technology (Social Sciences Edition)*, 23(6): 110-118. (in Chinese)

陈科霖, 谷志军. 2022. 多元政绩竞赛: 中国地方官员晋升的新解释[J]. 政治学研究, (1): 117-128, 159-160.

Chen K L, Gu Z J. 2022. Pluralistic political achievements competition: A new explanation for promotion of Chinese local cadres[J]. *Cass Journal of Political Science*, (1): 117-128, 159-160. (in Chinese)

- 陈强. 2014. 高级计量经济学及 Stata 应用[M]. 2 版. 北京: 高等教育出版社.
- Chen Q. 2014. *Advanced econometrics with Stata applications 2nd edition*[M]. Beijing: Higher Education Press. (in Chinese)
- 陈诗一, 陈登科. 2018. 雾霾污染、政府治理与经济高质量发展[J]. 经济研究, 53(2): 20-34.
- Chen S Y, Chen D K. 2018. Air pollution, government regulations and high-quality economic development[J]. *Economic Research Journal*, 53(2): 20-34. (in Chinese)
- 崔慧姝, 周望. 2018. 城市治理的中国之道: 基于“创城”机制的一项分析[J]. 学习论坛, (8): 59-64.
- Cui H S, Zhou W. 2018. The Chinese way of urban governance: An analysis based on the mechanism of “creating cities”[J]. *Tribune of Study*, (8): 59-64. (in Chinese)
- 韩超, 刘鑫颖, 王海. 2016. 规制官员激励与行为偏好——独立性缺失下环境规制失效新解[J]. 管理世界, (2): 82-94.
- Han C, Liu X Y, Wang H. 2016. Regulatory officials' incentives and behavioral preferences—A new solution to environmental regulatory failure under the lack of independence[J]. *Journal of Management World*, (2): 82-94. (in Chinese)
- 韩志明, 史瑞杰. 2015. 国家荣誉的社会认知——基于问卷调查数据的实证分析[J]. 中国行政管理, (10): 64-68.
- Han Z M, Shi R J. 2015. Social cognition of national honor—The empirical analysis based on questionnaire data[J]. *Chinese Public Administration*, (10): 64-68. (in Chinese)
- 李雪妮, 苏杨, 周绍杰. 2020. 空气质量如何影响公众主观满意度? ——基于中国民生调查微观数据的证据[J]. 中国人口·资源与环境, 30(5): 127-136.
- Li X N, Su Y, Zhou S J. 2020. How does air quality affect the subjective satisfaction of the public? —Evidence babed on Chinese people's livelihood survey [J]. *China Population, Resources and Environment*, 30(5): 127-136. (in Chinese)
- 李振. 2014. 作为锦标赛动员官员的评比表彰模式——以“创建卫生城市”运动为例[J]. 上海交通大学学报(哲学社会科学版), 22(5): 54-62.
- Li Z. 2014. Motivating cadres by tournaments: Rating and praise in contemporary China [J]. *Journal of Shanghai Jiao Tong University (Philosophy and Social Science)*, 22(5): 54-62. (in Chinese)
- 林挺进. 2015. 城市环保绩效、市长升迁偏好与市民环保满意度——基于 2011 连氏中国城市服务型政府调查的实证研究[J]. 甘肃行政学院学报, (6): 12-21, 125.
- Lin T J. 2015. Explaining citizens' satisfaction with environment-protection service: 2012 Lien's public service survey in 32 Chinese cities[J]. *Journal of Gansu Administration Institute*, (6): 12-21, 125. (in Chinese)
- 刘松瑞, 王赫, 席天扬. 2020. 行政竞标制、治理绩效和官员激励——基于国家卫生城市评比的研究[J]. 公共管理学报, 17(4): 10-20, 164.

- Liu S R, Wang H, Xi T Y. 2020. Administrative bidding, performance, and political incentives—Evidence from national health cities [J]. *Journal of Public Management*, 17(4): 10-20, 164. (in Chinese)
- 刘松瑞, 周润人, 席天扬. 2022. 城市评比表彰中的引领效应与同侪效应——基于创建国家卫生城市的研究[J]. *公共行政评论*, 15(3): 4-20, 196.
- Liu S R, Zhou R R, Xi T Y. 2022. Leadership effect and peer effect in competition for honorary city titles: Evidence from national health cities [J]. *Journal of Public Administration*, 15(3): 4-20, 196. (in Chinese)
- 谭志雄, 张阳阳, 付佳. 2018. 预期的力量: 晋升会引致环境污染吗? ——基于工具变量法的实证研究[J]. *管理工程学报*, 32(4): 37-45.
- Tan Z X, Zhang Y Y, Fu J. 2018. The power of expectation: Does promotion cause pollution? —The empirical research based on instrumental-variable method [J]. *Journal of Industrial Engineering and Engineering Management*, 32(4): 37-45. (in Chinese)
- 唐啸, 周绍杰, 赵鑫蕊, 等. 2020. 回应性外溢与央地关系: 基于中国民众环境满意度的实证研究[J]. *管理世界*, 36(6): 120-134, 249.
- Tang X, Zhou S J, Zhao X R, et al. 2020. Central-local relations and government responsive spillover: An empirical research of Chinese citizens' satisfaction with environmental governance [J]. *Journal of Management World*, 36(6): 120-134, 249. (in Chinese)
- 王凯民, 檀榕基. 2014. 环境安全感、政府信任与风险治理——从“邻避效应”的角度分析[J]. *行政与法*, (2): 10-15.
- Wang K M, Tan R J. 2014. Feeling of environmental safety, trust in government and risk governing—From the perspective of “not-in-my-back-yard” effects [J]. *Administration and Law*, (2): 10-15. (in Chinese)
- 王磊, 王兰兰. 2022. “文明城市”评选与地方城投债规模的扩张: 一个基于评比表彰机制的分析[J]. *中央财经大学学报*, (1): 74-88.
- Wang L, Wang L L. 2022. “National civilized city” policy and expansion of municipal investment debt size: An analysis from “rating & praise” scheme [J]. *Journal of Central University of Finance & Economics*, (1): 74-88. (in Chinese)
- 王诗宗, 杨帆. 2018. 基层政策执行中的调适性社会动员: 行政控制与多元参与[J]. *中国社会科学*, (11): 135-155, 205-206.
- Wang S Z, Yang F. 2018. Adaptive social mobilization in the implementation of grassroots policy: Administrative control and multipolar involvement [J]. *Social Sciences in China*, (11): 135-155, 205-206. (in Chinese)
- 王伟, 王哲. 2019. 走向政社合作: 评比表彰计划发起主体的行政学分析[J]. *湖北行政学院学报*, (3): 61-66.
- Wang W, Wang Z. 2019. Toward government-society cooperation: An administrative analysis of the initiating body of the evaluation and recognition program [J]. *Journal of*

- Hubei Administration Institute*, (3): 61-66. (in Chinese)
- 王英, 唐雲. 2020. 评比表彰与城市治理——来自国家环境保护模范城市创建的经验证据[J]. 治理研究, 36(5): 38-49.
- Wang Y, Tang Y. 2020. Evaluation and recognition and urban governance: Empirical evidence from the creation of national model cities for environmental protection [J]. *Governance Studies*, 36(5): 38-49. (in Chinese)
- 王哲. 2018. 作为政治达标赛的评比表彰: 理论意义与演进逻辑——基于 A 省“省级园林县城”计划的案例研究[J]. 公共管理学报, 15(3): 16-26, 154-155.
- Wang Z. 2018. Rating and praise project as political race to the targets: Theoretical implications and evolution logic—A case study of “provincial garden county” project in A Province[J]. *Journal of Public Management*, 15(3): 16-26, 154-155. (in Chinese)
- 文宏, 林彬. 2020. “多任务竞逐”: 中国地方政府间竞争激励的整体性解释——以粤港澳大湾区区域治理为例[J]. 江苏社会科学, (1): 75-85.
- Wen H, Lin B. 2020. “Multi-tasking competition”: A holistic explanation of competitive incentives among local governments in China—The case of regional governance in Guangdong, Hong Kong and Macao [J]. *Jiangsu Social Science*, (1): 75-85. (in Chinese)
- 吴建祖, 王碧莹. 2022. 地方政府环境竞争、环境注意力与环境治理效率——基于地级市面板数据的实证研究[J]. 东北大学学报(社会科学版), 24(6): 33-40.
- Wu J Z, Wang B Y. 2022. Environmental competition, environmental attention, and environmental governance efficiency of local governments: An empirical study based on the panel data of prefecture-level cities [J]. *Journal of Northeastern University (Social Science)*, 24(6): 33-40. (in Chinese)
- 徐换歌, 蒋硕亮. 2020. 国家创新型城市试点政策的效果以及空间溢出[J]. 科学学研究, 38(12): 2161-2170.
- Xu H G, Jiang S L. 2020. Research on the policy effect and spatial spillover of the national innovative city pilot policy [J]. *Studies in Science of Science*, 38(12): 2161-2170. (in Chinese)
- 徐换歌. 2020. 评比表彰何以促进污染治理? ——来自文明城市评比的经验证据[J]. 公共行政评论, 13(6): 151-169, 213-214.
- Xu H G. 2020. How does evaluation and recognition promote pollution control? Empirical evidence from the evaluation of civilized cities [J]. *Journal of Public Administration*, 13(6): 151-169, 213-214. (in Chinese)
- 徐现祥, 王贤彬. 2010. 任命制下的官员经济增长行为[J]. 经济学(季刊), 9(4): 1447-1466.
- Xu X X, Wang X B. 2010. Growth behavior in the appointment economy [J]. *China Economic Quarterly*, 9(4): 1447-1466. (in Chinese)
- 徐岩, 范娜娜, 陈那波. 2015. 合法性承载: 对运动式治理及其转变的新解释——以 A

- 市 18 年创卫历程为例[J]. 公共行政评论, 8(2): 22-46, 179.
- Xu Y, Fan N N, Chen N B. 2015. Legitimacy loading: A new explanation for the political campaign in China—18 years of “creating a national sanitary city” in a Chinese city[J]. *Journal of Public Administration*, 8(2): 22-46, 179. (in Chinese)
- 杨开峰, 王璐璐. 2023. 政策溢出效应: 概念廓清、类型划分与机制探究[J]. 治理研究, 39(3): 29-50, 158-159.
- Yang K F, Wang L L. 2023. Policy spillover effect: Definition, classification and mechanism exploration[J]. *Governance Studies*, 39(3): 29-50, 158-159. (in Chinese)
- 姚洋, 席天扬, 李力行, 等. 2020. 选拔、培养和激励——来自 CCER 官员数据库的证据[J]. 经济学(季刊), 19(3): 1017-1040.
- Yang Y, Xi T Y, Li L X, et al. 2020. Selection, training and incentivizing—Evidence from the CCER officials dataset[J]. *China Economic Quarterly*, 19(03): 1017-1040. (in Chinese)
- 张天舒, 王子怡. 2020. 荣誉称号影响官员晋升的信号机制研究——来自全国文明城市评比的证据[J]. 中国行政管理, (9): 121-127.
- Zhang T S, Wang Z Y. 2020. A study on the signaling mechanism of honorary titles affecting the promotion of officials—Evidence from the national civilized cities[J]. *China Public Administration*, (9): 121-127. (in Chinese)
- 张咏梅, 李秉勤. 2010. 社会政策执行中的地方政府间竞赛与评奖: 地方激励理论及其在改善环境卫生中的实践[J]. 中国公共政策评论, 4: 94-110.
- Zhang Y M, Li B Q. 2010. Local government competitions and awards in social policy implementation: Local incentive theory and its practice in improving environmental sanitation[J]. *China Public Policy Review*, 4: 94-110. (in Chinese)
- 周黎安. 2007. 中国地方官员的晋升锦标赛模式研究[J]. 经济研究, (7): 36-50.
- Zhou L A. 2007. Governing China's local officials: An analysis of promotion tournament model[J]. *Economic Research Journal*, (7): 36-50. (in Chinese)
- 周晔. 2014. 合法性视域下的“评比表彰”: 行动逻辑、功能局限及治理——以中央政府及职能部门组织的评比表彰活动为例[J]. 中国行政管理, (9): 69-74.
- Zhou Y. 2014. Research on the phenomena of evaluation and awarding from the perspective of legitimacy: Action logic, functional limitation and its governance—Taking examples of the evaluation and awarding organized by the central government and its departments[J]. *China Public Administration*, (9): 69-74. (in Chinese)
- Ailawadi K L, Dant R P, Grewal D. 2004. The difference between perceptual and objective performance measures: An empirical analysis[R]. Marketing Science Institute Working Paper 04-001. Cambridge, MA: Marketing Science Institute.
- Angrist J D, Krueger A B. 1991. Does compulsory school attendance affect schooling and earnings? [J]. *The Quarterly Journal of Economics*, 106(4): 979-1014.
- Baker G, Gibbons R, Murphy K J. 1994. Subjective performance measures in optimal

- incentive contracts[J]. *The Quarterly Journal of Economics*, 109(4): 1125-1156.
- Borman W C. 1975. Effects of instructions to avoid halo error on reliability and validity of performance evaluation ratings[J]. *Journal of Applied Psychology*, 60(5): 556-560.
- Brudney J L, England R E. 1982. Urban policy making and subjective service evaluations: Are they compatible? [J]. *Public Administration Review*, 42(2): 127-135.
- Chen Z, Kahn M E, Liu Y, et al. 2018. The consequences of spatially differentiated water pollution regulation in China[J]. *Journal of Environmental Economics and Management*, 88: 468-485.
- Daneke G A, Klobus-Edwards P. 1979. Survey research for public administrators[J]. *Public Administration Review*, 39(5): 421-426.
- Dasgupta S, Hettige H, Wheeler D. 2000. What improves environmental compliance? Evidence from Mexican industry [J]. *Journal of Environmental Economics and Management*, 39(1): 39-66.
- Diez-Roux A V. 1998. Bringing context back into epidemiology: Variables and fallacies in multilevel analysis[J]. *American Journal of Public Health*, 88(2): 216-222.
- Feng N C, Neely D G, Slatten L A D. 2016. Accountability standards for nonprofit organizations: Do organizations benefit from certification programs? [J]. *International Journal of Public Administration*, 39(6): 470-479.
- Fitzgerald M R, Durant R F. 1980. Citizen evaluations and urban management: Service delivery in an era of protest[J]. *Public Administration Review*, 40(6): 585-594.
- Grimmelikhuijsen S G, Meijer A J. 2015. Does Twitter increase perceived police legitimacy? [J]. *Public Administration Review*, 75(4): 598-607.
- Heilmann S. 2008. Policy experimentation in China's economic rise [J]. *Studies in Comparative International Development*, 43(1): 1-26.
- Helliwell J F. 2003. How's life? Combining individual and national variables to explain subjective well-being[J]. *Economic Modelling*, 20(2): 331-360.
- Holmstrom B, Milgrom P. 1991. Multitask principal-agent analyses: Incentive contracts, asset ownership, and job design[J]. *Journal of Law, Economics, & Organization*, 7(Special issue): 24-52.
- James O. 2009. Evaluating the expectations disconfirmation and expectations anchoring approaches to citizen satisfaction with local public services [J]. *Journal of Public Administration Research and Theory*, 19(1): 107-123.
- Keller W, Levinson A. 2002. Pollution abatement costs and foreign direct investment inflows to U. S. states[J]. *Review of Economics and Statistics*, 84(4): 691-703.
- Kim S, Kim M. 2022. Citizen satisfaction in the public sector[M]//Schedler K. Elgar Encyclopedia of Public Management. *Edward Elgar Publishing Limited*, 205-207.
- Lee J B, Kim S. 2024. Understanding gaps between objective and subjective performance measures: Accreditation of public service organizations and citizen satisfaction[J]. *The*

- American Review of Public Administration*, 54(3): 271-286.
- Licari M J, McLean W, Rice T W. 2005. The condition of community streets and parks: A comparison of resident and nonresident evaluations[J]. *Public Administration Review*, 65(3): 360-368.
- List J A, Sturm D M. 2006. How elections matter: Theory and evidence from environmental policy[J]. *The Quarterly Journal of Economics*, 121(4): 1249-1281.
- Liu Z L, Wang J, Wang Y J. 2018. Understanding individual environmental concern in the context of local environmental governance in China: A multi-level analysis[J]. *Society & Natural Resources*, 31(11): 1283-1301.
- Montinola G, Qian Y Y, Weingast B R. 1995. Federalism, Chinese style: The political basis for economic success in China[J]. *World Politics*, 48(1): 50-81.
- Parks R B. 1984. Linking objective and subjective measures of performance[J]. *Public Administration Review*, 44(2): 118-127.
- Poister T H, Streib G. 1999. Performance measurement in municipal government: Assessing the state of the practice[J]. *Public Administration Review*, 59(4): 325-335.
- Robinson W S. 2011. Ecological correlations and the behavior of individuals[J]. *International Journal of Epidemiology*, 40(4): 1134.
- Rosentraub M S, Harlow K, Thompson L. 1979. In defense of surveys as a reliable source of evaluation data[J]. *Public Administration Review*, 39(3): 302-303.
- Schachter H L. 2010. Objective and subjective performance measures: A note on terminology[J]. *Administration & Society*, 42(5): 550-567.
- Skogan W G, Maxfield M G. 1981. *Coping with crime: Individual and neighborhood reactions* [M]. Newbury Park, CA: Sage Publications.
- Stipak B. 1979. Citizen satisfaction with urban services: Potential misuse as a performance indicator[J]. *Public Administration Review*, 39(1): 46-52.
- Sturman M C. 2003. Searching for the inverted U-shaped relationship between time and performance: Meta-analyses of the experience/performance, tenure/performance, and age/performance relationships[J]. *Journal of Management*, 29(5): 609-640.
- Teodoro M P, Hughes A G. 2012. Socializer or signal? How agency accreditation affects organizational culture[J]. *Public Administration Review*, 72(4): 583-591.
- Van Ryzin G G, Immerwahr S, Altman S. 2008. Measuring street cleanliness: A comparison of New York City's scorecard and results from a citizen survey[J]. *Public Administration Review*, 68(2): 295-303.
- Van Ryzin G G. 2004. Expectations, performance, and citizen satisfaction with urban services[J]. *Journal of Policy Analysis and Management*, 23(3): 433-448.
- Wall T D, Michie J, Patterson M, et al. 2004. On the validity of subjective measures of company performance[J]. *Personnel Psychology*, 57(1): 95-118.
- Wang X H, Gianakis G A. 1999. Public officials' attitudes toward subjective performance

- measures[J]. *Public Productivity and Management Review*, 22(4): 537-553.
- Zhang J S, Chen W N, Petrovsky N, et al. 2022. The expectancy-disconfirmation model and citizen satisfaction with public services: A meta-analysis and an agenda for best practice [J]. *Public Administration Review*, 82(1): 147-159.
- Zhang Y M, Li B Q. 2011. Motivating service improvement with awards and competitions-hygienic city campaigns in China[J]. *Environment and Urbanization*, 23(1): 41-56.
- Zhang Y X. 2020. “Promotion tournament 2.0”: Why local cadres expand health-care provision in China[J]. *Governance*, 33(4): 897-914.

Do Environmental Ratings and Recognition Enhance Perceived Government Performance?

YANG Kaifeng^{1,2} LIU Yang¹ WEN Weiming³

- (1. School of Public Administration and Policy, Renmin University of China;
2. Centre for Evidence-based Governance and Public Performance, Renmin University of China;
3. People's Government of Zushidian, Xupu County in Hunan Province)

Abstract: How can local governments be motivated to actively engage in social governance? The central government employs “ratings and recognition” to incentivize competition and comprehensive development among local governments. As a key non-economic initiative aimed at promoting efficient governance, can this approach enhance public perceptions of government performance?

Taking environmental ratings and recognition as an example, we manually collected data on typical national-level environmental awards from 2011 to 2015 and merged them with the database of the 2015 Chinese Social Survey. To avoid any ecological fallacy, we adopted a multilevel model regression analysis. Additionally, we moved beyond a mere correlation in cross-sectional data to casual inferences, adding greater confidence in strengthening our study. First, lagging the independent variable by one period establishes the foundation for a causal relationship, ensuring that “cause precedes, effect follows”. Second, we examined the net effect of the ratings and recognition and perceived performance by matching the award (experimental group) and the non-award (control group) with the propensity score matching in a quasi-experimental approach. Again, we drew on the existing literature to select instrumental variables, including the frequency of environment-related terms in city government work reports, the growth rate of environmental staff, and the age of municipal secretaries (an interaction term between the age of municipal secretaries and the rate of economic growth). These variables help address potential endogeneity issues of a causal inversion between city awards and public

perceptions of government performance. Finally, we included the Sanitary City and Civilized City ratings-two typical awards similar to, but distinct from, environmental ratings and recognition-as controls to account for alternative explanations and confounding factors.

It is found that municipal winning of awards in environmental competitions, such as ratings and recognition, contributes to improve public perceptions of objective environmental quality and overall government performance. However, these awards are not significantly related to perceived government environmental performance. In addition, the relationship between the number of awards and public perceptions of performance follows an inverted U-shaped. As the number of awards increases, public perceptions of performance become more positive up to a certain threshold. Beyond this threshold, additional awards lead to more negative perceptions. Our findings confirm the positive but limited relationship between subjective and objective performance, filling the theoretical gap with evidence from Chinese cities. This research also contributes to answering the classic theoretical question of how to evaluate recognized service providers at the individual level. Moreover, our study suggests that evaluations and recognition should integrate more subjective performance and satisfaction. By prioritizing people's feelings and needs in environmental governance practices, these evaluations can complement objective performance. This approach enhances both the instrumental role of guiding and directing from the "top-down" and the value attribute of boosting public satisfaction from the "bottom-up".

Keywords: rating and recognition patterns; subjective performance; perceived government performance; environmental governance; environmental tournament

投稿日期: 2023/11/3 送外审日期: 2023/11/21 首轮外审完成日期: 2023/12/11
录用日期: 2024/3/3 最终修回日期: 2024/9/10