

· 专栏: 政府改革 ·

公众参与的政府绩效评估是否奏效: 基于中国部分城市的多层分析

马 亮

内容提要: 越来越多的地方政府将公众参与纳入绩效评估,但是它会提升政府绩效吗?文章利用中国部分城市的公务员访谈和公民调查数据,实证研究公众参与的政府绩效评估对公共服务满意度的影响。结果显示,在公众参与的政府绩效评估实施得较好的城市,公民对政府绩效的满意度也较高。这表明公众参与政府绩效评估有利于政府更好地理解民众的需求,增强民众问责政府官员的话语权,进而推动政府绩效改善。

关键词: 公众参与 政府绩效评估 公共服务满意度 多层分析 问责

中图分类号: D630.9

文献标识码: A

文章编号: 1003-3947(2018)03-0113-12

一、引 言

新公共管理运动以来,政府越来越强调顾客至上、市场竞争、绩效标准等价值取向,这使政府绩效评价与管理问题日益凸显。随着新公共管理运动的兴起,顾客至上理念和顾客满意度指数的引入在极大程度上推动了公众参与政府绩效评估的发展。国内外各级政府都在如火如荼地开展公民调查、公民评议等公众参与的政府绩效评估活动。以美国国家研究中心开发的全国公民调查(National Citizen Survey, NCS)为例,该调查对城市管理的多个方面进行公民满意度调查和比较,迄今已被美国 45 个州的 300 多个城市采用。学者们发现,公民发起的政府绩效评价在推动政府问责与绩效改进中发挥了重要作用(Ho & Coates 2004)。

政府绩效管理的支持者认为,只有收集和利用政府绩效信息,才有可能改进政府绩效。但是,大量研究却关注政府为什么采用和实施绩效管理(Dalehite 2008; de Lancer Julnes & Holzer 2001),只有很少研究探讨绩效管理是否奏效。当然,我们也有必要研究为什么有些政府热衷于引入公众参与政府绩效评估,而有些政府则视公民参与政府绩效评估为洪水猛兽。但是,本文关注的是公众参与的政府绩效评估是否会改进公共服务绩效。本文的研究问题是,公众参与的政府绩效评估是否奏效了?如果奏效,它在哪些方面奏效了?它对哪些绩效维度奏效了?

从中国公众参与的政府绩效评估实践来看,我们不能忽视民意调查的崛起以及由此而产生的公众参与政府绩效评估热(吴建南、高小平 2006)。越来越多的政府开始引入公民调查,并将其纳入目标责任考核这一核心政府绩效评价机制中。2000 年以来,珠海、南京、沈阳等地开展的“万人评议”活动声势浩大,但其影响如何却没有得到系统研究(周志忍 2015)。有鉴于公众参与政府绩效评估的成本高昂(有形的调查经费、无形的社会成本以及对政府行为的影响),有必要研究它是否达到了预期目标,特别是对公共服务绩效产生了怎样的实际影响。

本文使用中国 19 个大城市的公务员访谈和公民抽样调查数据,采取多层模型分析方法,实证研究

作者简介: 马亮,中国人民大学公共管理学院副教授,国家发展与战略研究院研究员。

基金项目: 国家自然科学基金面上项目“第三方评估如何提升组织绩效:中国地方政府的实证研究”(项目编号:71774164)。

了公众参与的政府绩效评估对公共服务满意度的影响。结果显示,在公众参与的政府绩效评估实施得较好的城市,公民对政府绩效的满意度也较高。上述发现为深化人们对政府绩效评价与公众参与的理解提供了理论贡献,也为政府部门改善公众参与的效果提供了政策启示。

本文剩余部分做如下安排。我们首先回顾有关公众参与和政府绩效评估的研究文献,并指出现有研究的不足和本文的立足点。其次,我们具体探讨政府绩效评价中的公众参与及其对政府绩效的影响,并提出本文有待检验的研究假设。再次,我们介绍本文使用的样本和数据来源,以及本文对核心变量的操作方法和模型设定。又次,我们报告本研究的主要实证发现。最后,我们对上述发现进行讨论,总结本文的理论贡献和政策启示,并探讨本文的不足及未来研究方向。

二、文献评估与理论假设

(一) 文献综述

与本文相关的研究大致可以分为两类,一类是对公众参与的研究,另一类则是对政府绩效评价与管理的研究。公众参与是多方面的,而不局限于政府绩效评价。在政策咨询、政策评估等方面,公众参与的范围很广。对公民的参与大致可以分为政治参与和行政参与。与投票、选举、抗争等政治参与相比,学术界对行政参与的研究远远不足。尽管有关公众参与和政府绩效管理的研究为我们提供了启迪,但却并非直接探讨公众参与的政府绩效评估。

尽管公众参与政府绩效评估有很多不同的环节、方式和深度,公众发起的政府绩效评估也被寄予厚望(Ho & Coates 2004),但公民调查仍然是公众参与政府绩效评估的主要表现形式。20世纪70年代末以来,美国公共行政学界开始关注公民调查及其相关问题。直至今日,公民调查的研究仍然在持续地开展,虽然不同时期的研究主题会有所不同。研究者们关注的焦点是,公民调查能否成为一项衡量地方政府公共服务绩效的工具。一些学者将公民调查结果同政府部门收集的统计数据进行对比,发现二者的相关关系并不高,并对公民调查结果提出质疑(曾莉、李佳源 2013)。还有学者认为,公民调查的主观性较强,公民缺少判断公共服务质量的信息和知识,公民的评价很大程度上受其教育程度和政治态度的左右,很难说准确和可靠。

虽然没有研究直接考察公众参与政府绩效评估的影响,但它本身是公共管理系统的一个重要组成部分,而其研究则可以归属于一个更大的范畴,即探讨管理对绩效的影响(Andrews, Boyne & Walker, 2011; Im & Lee 2012; Meier & O'Toole 2007)。比如,王小虎考察了美国城市使用绩效评价的情况及其影响,发现政府主管在利用绩效评价,但公众的兴趣却不大(Wang 2002)。杨开峰和谢俊义发展了一套中层理论,用以解释政府绩效管理如何发挥作用,并得到台北市政府调查数据的支持(Yang & Hsieh, 2007)。

来自美国和英国的证据显示,政府绩效管理的确产生了正面影响。对美国佐治亚州的88个地方公共交通管理部门的研究显示,绩效管理的确提升了组织绩效(Poister, Pasha & Edwards 2013)。对纽约市公共学校的研究发现,使用绩效管理的学校明显提升了学生的标准考试成绩(Sun & Van Ryzin 2014)。对英国地方政府的调查表明,绩效管理同公民对公共服务绩效的满意度正相关(Andrews & Van de Walle 2013)。

最近的一项研究使用双重倍差法和多层模型,发现与其他欧洲国家相比,英国居民对医疗卫生的满意度显著提升,而这主要得益于英国国家医疗服务系统所推行的绩效管理举措(Grosso & Van Ryzin, 2012)。另一项对25个国家的多层分析显示,公共服务质量评定等新公共管理实践的确可以解释公民对公共服务的评价(Ariely 2011)。总之,上述研究都发现了绩效管理对政府绩效的正向影响。但是,也有研究发现,相对于私营部门而言,公共部门实施绩效管理的收效偏低(Hvidman & Andersen 2014)。对49项研究的荟萃分析显示,绩效管理会产生正面影响,但是总体来说收效甚微(Gerrish 2016)。

就中国背景下的研究而言,吴建南和华乐勤的专家调查表明,公众参与被视为政府绩效管理的未来

发展方向(Walker & Wu 2010)。在缺乏直接投票等有效参与渠道的背景下,公众参与政府绩效评价便成为他们影响政策进程的重要方式,并可能会对政府官员产生一定的激励作用(Heberer & Trappel, 2013)。对广东佛山的研究显示,市政府领导和职能部门主管对公众参与政府绩效评估的认识不同,并导致他们使用评估结果的方式也不同(Gao 2012)。对武汉的案例研究发现,政风行风测评未能形成双环学习和反馈机制,对政府绩效改进的作用也有限(So 2014)。最近对广东的访谈发现,地方政府官员忽视公众参与的政府绩效评估,因为他们并不认为公民能够评判政府绩效,而且其他方面的考评也冲淡了公众参与的作用(Li 2015)。上述研究均为理论探讨、描述性分析和案例研究,尚缺少基于大样本数据的实证研究证据。

围绕政府绩效评价与管理的研究越来越多,但已有研究存在如下不足。首先,虽然一些研究探讨了政府绩效评价与管理的影响,但却并未直接考察公众参与的政府绩效评估能否带来绩效回报,以及其对不同绩效维度的具体影响(Ma 2017)。上述研究表明,有关政府绩效评价与管理的研究,往往是对其整体绩效影响进行分析,而未具体探讨公众参与的政府绩效评估能否带来绩效影响。其次,很多的已有研究使用同一个来源和方法的数据,存在共同方法偏误,使其结果的稳健性不足(Meier & O'Toole, 2013)。最后,多数研究是在美国、英国等西方国家开展的,缺少中国等发展中国家的证据。有鉴于此,本文期望使用最近的中国调查数据,对公众参与政府绩效评估的效应进行实证分析。

(二) 公民参与的绩效影响

政府绩效管理系统包括许多方面和环节(Wholey, 1999),在此我们关注利益相关者的参与对政府绩效的影响。利益相关者的参与,尤其是公民的参与,对政府绩效管理系统优化至为重要。公众参与政府绩效评估显然是公民参与的领域之一,因此公民参与的相关研究为我们研究该问题提供了启示。这些研究发现,尽管公众参与费时费力还不讨好,但愿意尝试并合理设计的政府部门,的确从中获得了提升绩效的回报(Neshkova & Guo 2012)。更为重要的是,公众参与政府绩效评估还可以带来如下几个方面的具体好处,并进而推动政府绩效改进。

首先,公民在绩效评价系统、目标与指标设计环节的参与,有利于政府与公众对绩效指标和标准达成共识。公众参与使政府更清晰地认识到其政策使命和优先目标,为树立正确的政绩观提供了标尺(Ho & Coates 2004)。

其次,公民在绩效结果评价中提供信息、需求与反馈,为明确绩效问题提供依据。公民参与有利于强化以客为先的理念,并为政府识别绩效问题提供指引(Ma 2016b)。有时候,恰恰是来自公民的抱怨、不满、投诉等负面信息,为政府重新认识政策问题并发现绩效差距提供了启示(James & John 2007)。

与此同时,绩效结果的公开、告知与监督,强化了公民对政府的绩效问责,使政府感受到改进绩效和回应公民诉求的压力(Yu & Ma 2015)。公众参与的政府绩效评估也有助于扭转政府官员“眼睛向上”的执政观念,而更加以民为本和以客为先(Ma, Tang & Yan 2015)。当公民掌握了话语权,并能够发出“声音”时,他们就会对政府施加压力并产生影响(James & John 2007)。

由此可见,公民参与的政府绩效评估可以发挥多方面的积极作用,并可能与政府绩效产生正相关的关系。换句话说,在绩效评价中引入公民参与的政府绩效更高。基于上述理论讨论,我们提出如下研究假设。

研究假设: 公众参与政府绩效评估与政府绩效正相关,即在绩效评价中引入公民参与的政府的绩效更高。

政府绩效是一个多维度的概念,包括经济性、效率性、效果性、公平性、响应性等维度(Andrews & Van de Walle 2013)。基于上述假说的推论就是,公众参与政府绩效评估可能同这些绩效维度存在正相关关系。与此同时,公众参与政府绩效评估对上述绩效维度的影响可能是不同的,因此有必要加以区分和检验。

三、数据与方法

(一) 样本与数据来源

为了验证上述研究假设,我们以中国19个主要城市为研究对象,包括4个直辖市、10个省会城市和5个计划单列市。这些城市都是大城市,但在地理分布、人口规模和经济发展水平等方面存在较大差异,使我们可以观测这些状况不同的城市的公众参与政府绩效评估的情况及其对绩效的影响。

我们使用多个来源的数据,以避免共同方法偏误对研究结果的干扰。有关公民参与政府绩效评价的数据,来自中山大学的何艳玲等开展的“2013年中国城市公共服务能力调查”(何艳玲,2013)。有关公民对政府绩效的满意度的数据,来自南洋理工大学的吴伟等开展的“2013年连氏中国城市服务型政府调查”(吴伟等,2014)。

(二) 因变量

本研究有关政府绩效的数据来自抽样居民电话访谈。利用计算机辅助电话访问(CATI)系统,每个城市随机抽取700余位常住成年居民进行电话调查,了解其对公共服务质量、政府工作人员态度和能力等方面的满意度。调查工具使用11点李克特量表(0到10点),分值越高表明满意度越高。调查中还获取了居民个人的基本情况,以作为本研究的控制变量。

具体来说,我们分析了公众参与和政府绩效的多个维度的相关关系,包括的调查题项有公共服务满意度、办事手续便利度、政府服务态度、政府服务能力、政府征求民意、政府采纳民意、政府廉洁程度等,它们反映了城市政府的效果性、效率性、响应性、廉洁性等维度。

由于上述7个题项高度正相关,且都是反映政府绩效的各个维度,因此我们使用上述题项的简单算术平均值来构造政府绩效指数,作为衡量政府整体绩效的指标。

(三) 自变量

有关公民参与政府绩效评价的数据来自于样本城市政府部门工作人员的访谈。每个城市选择11个与公共服务密切相关的政府部门,并抽取一位对政府绩效评价工作较熟悉的工作人员进行访谈,请其对公民参与在政府绩效评价中的重要程度进行评价。

调查使用7点李克特量表(0到6点),并将11个受访者的回答取均值,作为衡量所在城市公众参与政府绩效评价情况的指标。调查是在2012年末进行的,总计访问了209位公务员。与单一受访者相比,多受访者的调查方式可以降低测量偏误(Walker & Enticott, 2004)。

(四) 控制变量

因为公民对政府绩效的评价会受到许多个人因素的影响,为此控制了受访者的个人统计特征变量,具体包括性别、年龄、民族、学历、居住时间、家庭收入、政治面貌、户口等。这些变量的具体定义和衡量方法可以见表1。

表1:主要变量的定义与测量

变量	定义与测量
公众参与	市民参与政府绩效评价的程度如何?
公共服务满意度	总的来说,以上公共服务领域可以达到您的期望吗?
办事手续便利度	政府部门办事手续的简单方便程度如何?
政府服务态度	政府部门工作人员的服务态度如何?
政府服务能力	政府部门工作人员的业务能力如何?
政府征求民意	政府征求老百姓的意见和建议的情况如何?
政府采纳民意	政府采纳老百姓的意见和建议的情况如何?
政府廉洁程度	政府部门工作人员的清正廉洁程度如何?

续表 1

变量	定义与测量
政府绩效指数	上述 7 个题项的简单算术平均值。
性别	以女性为参照组, 设置男性的虚拟变量。
年龄	年龄分为 6 组, 从 18 ~ 24 岁到 60 岁及以上分别编码 1 ~ 6 点。
民族	以少数民族为参照组, 设置汉族的虚拟变量。
教育程度	教育程度分为 6 组, 从小学及以下到硕士及以上学历依次编码 1 到 6 点。
居住时间	以实际居住年数分 4 组, 从 1 ~ 3 年到 10 年以上依次编码 1 ~ 4 点。
家庭收入	家庭每年总收入分为 10 组, 从无固定收入到 50 万元及以上依次赋分 1 ~ 10 分。
党员	以非中共党员为参照组, 设置中共党员的虚拟变量。
户口	以外地户口为参照组, 设置本地城镇户口和本地农村户口的虚拟变量。

(五) 分析方法

本研究使用的是多层数据, 包括第一层的公民和第二层的城市。公民嵌套在各个城市, 因此不同城市的公民之间相互不独立, 违背了普通最小二乘法 (OLS) 模型的观测点独立性的前提假设, 因此需要使用多层线性模型 (HLM) 或多层模型。在这里, 我们使用的两层模型, 第一层是公民, 第二层为城市。

各自变量都进行了中心化处理, 以使估计结果便于解释。因为考察是第二层 (城市) 变量的直接效应和交互效应, 因此对各变量进行总平均中心化 (grand mean centering) 处理 (Enders & Tofighi 2007)。本文主要关注第二层自变量的影响, 因此使用的是最基本的随机截距模型 (random intercept model) (Aguinis, Gottfredson, & Culpepper 2013)。

四、实证结果

(一) 描述性统计分析

首先, 表 2 报告了主要变量的描述性统计分析结果。总体来看, 样本城市公众参与政府绩效评估的总体情况并不理想, 平均分仅为 1.83 分, 其中最大值为 3.4 分, 而最小值则低至 0.3 分。但是, 样本城市的差异很明显, 标准差接近 1.0, 表明各个城市在公众参与政府绩效评估方面存在很大不同。

就政府绩效的各个维度而言, 受访居民对所在城市的政府总体感到满意, 其得分均接近或超过 6.0 分。得分最低的是政府廉洁程度, 得分最高的是公共服务满意度, 而其他维度的差异并不明显。从标准差来看, 各维度均超过 2.1, 表明受访者对政府绩效的认知水平差别明显。

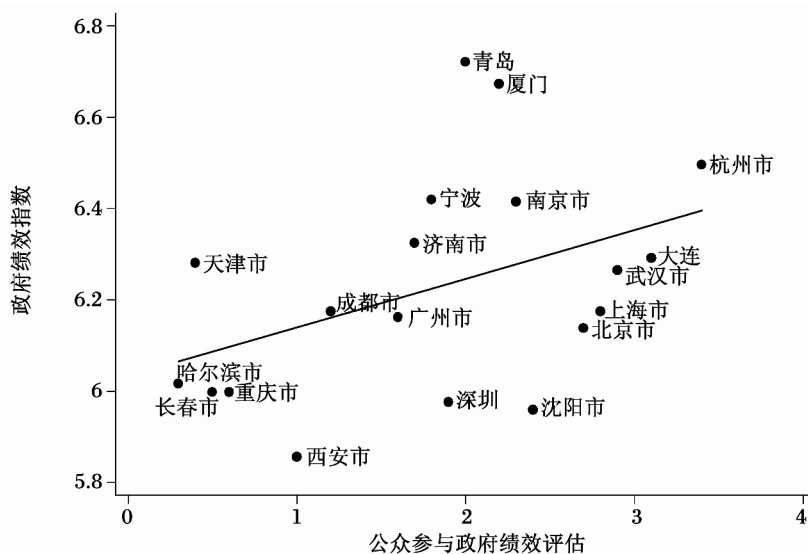
表 2: 主要变量的描述性统计分析

变量	N	均值	标准差	最小值	最大值
公众参与	19	1.832	0.956	0.3	3.4
公共服务满意度	12970	6.668	2.182	0	10
办事手续便利度	12469	6.205	2.409	0	10
政府服务态度	12665	6.166	2.427	0	10
政府服务能力	12289	6.172	2.404	0	10
政府征求民意	11819	6.179	2.506	0	10
政府采纳民意	11760	6.090	2.534	0	10
政府廉洁程度	11633	5.966	2.578	0	10
政府绩效指数	13753	6.229	1.589	0	10
性别 (男 = 1)	13756	0.521	0.500	0	1
年龄	13756	2.772	1.500	1	6

续表 2

变量	N	均值	标准差	最小值	最大值
民族(汉族=1)	13722	0.978	0.146	0	1
教育程度	11972	3.033	1.452	1	6
居住时间	13756	2.708	1.185	1	4
家庭收入	11765	4.275	1.795	1	10
党员	12979	0.140	0.347	0	1
本地城镇户口	13754	0.404	0.491	0	1
本地农村户口	13754	0.176	0.381	0	1

如图1所示,公众参与的政府绩效评估同公民对政府绩效的总体满意度之间存在明显的正相关关系。在公众参与政府绩效评估方面推进较快的城市(如杭州、青岛、厦门、大连等),其居民对政府绩效的满意度也较高。与之相比,天津、重庆、哈尔滨、西安等城市的公众参与政府绩效评估的情况不尽如人意,其居民对政府绩效的满意度也相对较低。就此而言,本文提出的研究假设得到了初步验证,但是仍有待于更精细的分析。



注: $N = 19$, $R^2 = 0.179$ 。

图1: 公众参与和政府绩效指数的二维散点图与线性拟合曲线

(二) 回归分析

在进行多层模型分析前,我们首先分析公民层面和城市层面的变量解释政府绩效的方差比例。组内相关系数(Intraclass correlation Coefficient, ICC)常常用来衡量不同层面对因变量的方差解释程度,在这里指城市层面对因变量的方差解释与公民层面与城市层面的方差解释之和的比值。

不出意外,公民层面对各个因变量的解释比例都高达90%以上,而城市层面对因变量的解释力介于0.7~1.6%。这说明各因变量的方差主要被公民层面的变量所解释,而城市层面的变量解释率较低。这是可以理解的,因为城市内部不同居民在人口统计特征方面的诸多差异,使他们对政府绩效的认知水平千差万别。

对于城市层面究竟占多大比例才使用多层模型,目前尚无共识。但一般认为,如果比例在3%以上就应使用多层模型。考虑到数据结构的特征和城市样本量远小于公民样本量,以及模型检验结果拒绝普通最小二乘法模型,因此我们仍然使用多层模型。

我们首先将 7 个政府绩效的具体指标作为因变量, 将公众参与和公民个人特征变量同时代入多层模型进行分析。表 3 报告了多层模型的主要分析结果。结果显示, 公众参与的政府绩效评估对政府绩效满意度的影响较大。

表 3: 多层模型分析结果

变量	模型 1	模型 2	模型 3	模型 4	模型 5	模型 6	模型 7	模型 8
	公共服务	办事便利	服务态度	服务能力	征求意见	采纳民意	政府廉洁	政府绩效
公众参与	0.0919 (0.0732)	0.113* (0.0679)	0.106* (0.0586)	0.109* (0.0543)	0.103 (0.0704)	0.177* (0.0743)	0.114 (0.0771)	0.117* (0.0554)
性别 (男 = 1)	-0.0436 (0.0451)	-0.269*** (0.0513)	-0.0200 (0.0517)	-0.115* (0.0517)	-0.205*** (0.0551)	-0.294*** (0.0556)	-0.214*** (0.0577)	-0.152*** (0.0325)
年龄	-0.000995 (0.0153)	0.0341* (0.0174)	0.0373* (0.0175)	0.0258 (0.0175)	0.00415 (0.0184)	0.00665 (0.0186)	-0.0176 (0.0194)	0.0156 (0.0110)
民族 (汉族 = 1)	-0.0833 (0.163)	0.335* (0.188)	0.345* (0.189)	0.443* (0.189)	0.354* (0.203)	0.106 (0.208)	0.0464 (0.222)	0.239* (0.120)
教育程度	-0.0296* (0.0157)	-0.0767*** (0.0179)	-0.0489*** (0.0179)	-0.0540*** (0.0179)	-0.0648*** (0.0191)	-0.0717*** (0.0193)	-0.0736*** (0.0199)	-0.0572*** (0.0113)
居住时间	-0.00619 (0.0196)	-0.0699*** (0.0222)	-0.0207 (0.0225)	-0.0393* (0.0223)	-0.0487* (0.0238)	-0.0693*** (0.0240)	-0.123*** (0.0249)	-0.0481*** (0.0141)
家庭收入	0.00225 (0.0128)	-0.0100 (0.0144)	-0.0133 (0.0146)	0.0104 (0.0146)	-0.00843 (0.0155)	-0.0248 (0.0157)	-0.0207 (0.0163)	-0.0114 (0.00921)
党员	0.0260 (0.0654)	0.138* (0.0739)	0.00946 (0.0746)	-0.108 (0.0744)	0.0314 (0.0789)	0.102 (0.0797)	-0.0155 (0.0828)	0.0296 (0.0471)
本地城镇 户口	0.0294 (0.0511)	-0.0412 (0.0581)	-0.0610 (0.0584)	-0.0492 (0.0582)	-0.0789 (0.0621)	-0.0971 (0.0627)	-0.0230 (0.0650)	-0.0516 (0.0369)
本地农村户口	0.146* (0.0642)	-0.0503 (0.0727)	-0.0636 (0.0732)	-0.131* (0.0733)	-0.0463 (0.0778)	0.00619 (0.0785)	0.118 (0.0813)	-0.0141 (0.0462)
常数项	6.659*** (0.0683)	6.174*** (0.0635)	6.148*** (0.0549)	6.156*** (0.0509)	6.123*** (0.0659)	6.068*** (0.0695)	5.908*** (0.0722)	6.202*** (0.0517)
城市层面方差	0.0788*** (0.0285)	0.0638*** (0.0247)	0.0443*** (0.0184)	0.0362*** (0.0158)	0.0677*** (0.0267)	0.0768*** (0.0298)	0.0827*** (0.0319)	0.0457*** (0.0164)
居民层面方差	4.566*** (0.0676)	5.708*** (0.0861)	5.885*** (0.0880)	5.712*** (0.0869)	6.230*** (0.0967)	6.329*** (0.0984)	6.711*** (0.105)	2.525*** (0.0363)
Log likelihood	-19941.5	-20204.5	-20682.6	-19864.6	-19446.7	-19444.6	-19443.9	-18270.4
Wald χ^2	12.37	75.87	23.65	36.48	41.29	70.03	67.70	82.27
N	9145	8816	8966	8668	8326	8297	8194	9696

注: 组数为 19 组。括号外为回归系数, 括号内为标准误。***、**、* 分别代表在 1%、5% 和 10% 的显著性水平下统计显著。

在所有 7 个模型中, 公众参与的政府绩效评估都同因变量存在正相关关系, 且相关系数较大, 回归系数介于 0.09 到 0.18 之间。公众参与的政府绩效评估同政府办事手续便利程度、政府服务态度、政府服务能力、政府采纳民意等四个因变量的相关关系, 都通过了 0.10 水平上的统计显著性检验, 而其同公共服务满意度、政府征求意见、政府廉洁程度等的相关关系则统计不显著。

在模型 8 中, 我们将政府绩效指数作为因变量, 将公众参与代入模型。结果发现, 公众参与的政府绩效评估同其显著正相关 ($\beta = 0.117$, $p < 0.10$)。该结果表明, 公众参与的政府绩效评估如果提高 1 个单位的话, 那么政府绩效指数则提高约 0.117 点。考虑到城市之间的差别, 这个提高幅度是相对可观的。

就各个层面对因变量的方差解释力而言,分析显示城市层面的方差解释率都通过了统计显著性检验,表明公众参与的政府绩效评估同各个因变量都存在较强的相关关系。综上所述,上述结果支持了本文提出的研究假设,即公众参与的政府绩效评估的确同政府绩效显著正相关,在政府绩效评估中引入公民参与的政府的绩效相对更高。

对人口统计特征变量的分析显示,除了公共服务满意度和服务态度以外,男性相对女性对政府绩效的其他维度的满意度都显著更低。受访者的年龄对政府绩效满意度的影响并不显著,年龄较长者仅对政府办事手续和服务态度的满意度更高。就民族而言,汉族居民对政府绩效的多个维度都显示出显著较高的满意度。在教育程度方面,结果显示学历较高的受访者对政府绩效的满意度显著较低,这可能同他们较高的期望水平和批判精神有关。有趣的是,我们并未发现家庭收入对政府绩效满意度的显著影响,尽管其符号在多数模型都是负向的。身为党员的受访者对政府绩效的满意度并不比非党员的居民更高,但是在办事手续的便利程度方面则表现明显,这可能同他们多数在公职部门而“近水楼台先得月”有关。与外地户口的受访者相比,本地城镇户口的居民对政府绩效更不满意,但这个影响并不显著。本地农村户口的居民对公共服务显著更满意,但对政府业务能力的评价则显著更不满意。

五、讨论与结论

(一) 结果讨论

中国政府绩效评价的核心是自上而下的目标责任考核,表现为通过层层分解目标任务和层层包干责任,实现官僚体制的控制(Ma 2016a),但是,公众参与的政府绩效评估也在兴起,特别是民意调查、政风行风测评、万人评议被引入政府绩效评价,为反映民意和倾听民声提供渠道,使政府绩效评价多元化。比如,政风行风评议就是一种典型的公众参与政府绩效评估,虽然其关注点在于纠正行业不正之风。

本研究发现,公众参与的政府绩效评估的确正向影响政府绩效,表明各地政府应积极推动公众参与政府绩效评估。但是就政府绩效管理的具体维度而言,我们发现公众参与的政府绩效评估只对政府办事手续便利程度、政府服务态度、政府服务能力、政府采纳民意等四个维度产生显著影响,而公共服务满意度、政府征求民意、政府廉洁程度等则不受其显著影响。

研究显示,公众参与的政府绩效评估对公共服务绩效的影响不明显。诸如公共交通、公共教育、医疗卫生、环境保护等方面的公共服务绩效,更多地取决于财政投入和政府能力等因素,而公众参与的影响可能还不够(马亮、于文轩 2013)。

研究表明,公众参与的政府绩效评估同政府征求民意的相关关系也不显著。在政府绩效评估中邀请公众参与,恰恰是政府征求民意的表现,而我们的发现则说明公众参与的程度可能还有待提高。目前来看,公民参与的浅层化和初级化色彩还较为明显,许多公众可能只是“打打分”,而没能深度参与到政府绩效管理过程(周志忍 2015)。

公众参与对政府廉洁程度的影响也是不显著的,这可能同调查期间中国正在开展的“反腐风暴”有密切关系,即由于国家对腐败行为的“零容忍”,揭露了一大批高层次领导干部的贪腐行为,动摇了民众对政府廉洁的信心。

(二) 理论贡献与政策启示

基于中国情境的实证研究,本文证明了公众参与的政府绩效评估对政府绩效的推进作用。就此而言,公众参与的政府绩效评估在改善政府绩效方面有很大的潜力。但是,当前中国公众参与的政府绩效评估也存在一些突出的问题。关注和破解这些问题,可能进一步提升公众参与对政府绩效评估与管理的贡献。

早期的公众参与政府绩效评估多数是不科学的,比如组织政府部门在广场散发问卷,没有随机抽样的概念,这样谁都可以填写,而未考虑调查对象是否接触过相关服务。再比如问卷设计质量粗糙,选项

中只有“满意”和“比较满意”,存在一定的评优摆好嫌疑。与此同时,公民参与被滥用和误用,抽样调查方法粗糙也使公民参与的作用甚微(吴建南、高小平 2006)。

更为重要的是,公众参与的政府绩效评估往往被用于上级政府控制和监督下级的工具。由于存在较强的激励(如奖励先进、末位淘汰等奖惩措施),一些被评部门组织工作人员抢问卷和自行填写去捏造数据,或者通过非正式途径影响组织部门篡改数据(如施压或贿赂)。绩效问责存在的问题较为突出,而强力激励也导致政府的功能紊乱。比如末位淘汰的异化,导致弄虚作假与“轮流坐庄”。部门之间的差异较大,可比性不足,产生评比疲劳与倦怠。与此同时,被评部门的博弈行为凸显,产生许多不利于绩效改进的“迎评”策略(Li 2015)。显然,由于这种形式的公众参与政府绩效评估严重走样,可能很难发挥其改进公共服务绩效的预期目标。

越来越多的地方政府认识到这种局限,而对公众参与政府绩效评估进行改进。比如,委托专业调查机构(如统计局的社情民意调查中心、大学社会调查研究机构、市场调查公司等),采用随机电话调查方法,使调查结果的可信度得以提升。另外,一些政府部门也愈加认识到公众参与政府绩效评估的局限性,比如需要关注部门之间的可比性,对评估结果的使用力度也趋于弱化。

与此同时,从政府绩效评价向政府绩效管理的转变,使政府不再单纯注重评价和排名(如GDP增长率),而更强调如何利用评价过程和结果,以诊断问题、学习先进并改进绩效(马亮 2017)。目前监察部开展的政府绩效管理试点与推广,为政府绩效管理的进一步推动提供了可能。但是,如何进一步创新和改进政府绩效管理,仍然是一项不容忽视的课题。

结合本文的发现及相关研究和实践进展,我们提出如下政策建议。首先,“关口前移”。公民被动参与政府绩效评估,只是发挥了提供信息的作用。实际上,公民可以更主动地参与政府绩效评价,在政府绩效评价指标的设计、权重的设置、评价方式的选择、评价结果的使用等方面,都可以发挥很重要的咨询作用。公众参与的政府绩效评估可以借鉴对话理论,以增进公众参与的深度。以新加坡政府在2013年举行的全国对话会为例,它不只是单向或冷冰冰的参与,而是多向和热情的互动,对于提升民意的影响力作用更为明显。

值得注意的是,日益流行的大数据技术为公众参与提供了更多的可能。公众无论是被动提供数据,还是主动提供数据,都为政府绩效管理提供了基础。信息技术缩短了政府绩效反馈的周期,而借助于社交媒体也使绩效信息的扩散更为充分。美国许多城市的311系统与Citistat的战略性绩效管理,都开始推动实时绩效信息披露,而公众的反馈是其主要信息来源。

其次,政府绩效信息使用至关重要。从政府绩效评价走向政府绩效管理,特别是充分和合理使用政府绩效信息,是至关重要的举措(马亮 2014)。将政府绩效信息用于奖惩和用于绩效诊断,可能会产生截然不同的影响。就目前公众参与的政府绩效评估而言,诸如上文所述的强力激励可能有一定价值,但不应操之过急和用之过度。

更为重要的是,需要重视不同部门和地区之间的可比性问题。以纽约市警察局为例,它在推行以犯罪率为依据的绩效考核时,是在范围较小的片区或警区(precinct)开展的,因为这样各个片区之间更有可比性,主客观绩效指标之间的关联性也更强(Behn 2014)。与此同时,城市之间的标杆比较,也是值得发展的方向。比如,美国全国公民调查(NCS)专门从事公民调查,利用统一量表和方法获得各城市的数据并进行标杆比对。中国许多城市之间完全可以建立类似的联盟或结对子,以基本一致的公民调查工具进行比较,以推动城市之间的相互学习和竞争。

(三) 研究不足与未来展望

本文的研究依然存在一些不足,而这也正是我们未来研究的方向。首先,本文使用的是截面数据,虽然我们滞后了因变量,但仍然无法推断变量之间的因果关系。未来研究可以设计实地实验研究,或者使用纵贯数据,以考察公众参与的政府绩效评估对公共服务满意度的作用。比如,中国各地政府试点公众参与的政府绩效评估,可以利用这些“自然实验”或“准实验”,开展更细致的研究(马亮 2015)。

其次,本文使用的样本是中国的大城市,而中小城市的适用性还有待拓展。规模较大的省会和副省级城市往往有较强的政府能力,在开展公众参与的政府绩效评估方面积累了较为丰富的经验,其所产生的绩效影响也更为凸显。与之相比,中小城市的公众参与能否推动政府绩效改进,仍然值得进一步研究。

此外,公众参与的形式与深度不同,可能对政府绩效产生不同的影响。受制于调查资料的限制,目前本文只是笼统地研究了公民在政府绩效管理的参与情况。与此同时,单指标测量的局限性也有必要在未来研究中加以克服。公众参与是一个多维概念,而公众参与的不同形式与深度也有待于进一步研究。

值得注意的是,公众参与的政府绩效评估可能并不是直接影响公共服务满意度,其中间过程和中介机制还有待于未来研究加以识别。公众参与的政府绩效评估在何种条件下和通过哪些机制来影响政府绩效,仍然没有得到充分研究。绩效管理的核心在于利用绩效信息,因为只有诊断和发现绩效问题并出台新举措,才能真正改进绩效。在政府绩效评价研究中,一个趋势是研究政府绩效信息是否及如何被使用(马亮,2014)。因为,单纯收集绩效信息而不加以利用,并不会产生明显的绩效影响。只有将收集的绩效信息加以充分利用,才能期望其产生绩效影响。

因而,公众参与政府绩效评估能否发挥作用,很大程度上取决于评估结果是否会被运用。政府绩效管理作为一个综合系统,需要不同元素的协调运作。因此未来研究可以进一步考察不同绩效管理元素(公众参与、绩效问责与绩效改进等)对政府绩效产生的交互效应。未来研究可以发展中层理论,以揭示公众参与的政府绩效评估的运作机理。比如发展政府绩效评价的中层理论,勾勒出政府绩效评价效果的中间过程和中介机制,并深化人们对其作用机理的认识(Yang & Hsieh, 2007)。

最后,情境至关重要。在不同的政治体制和社会文化下,公众参与的政府绩效评估可能产生不同的影响,而这有待于在其他情境进行检验和比较(Meier et al., 2015)。比如,聚焦于具体行业和领域(如公共交通、医疗卫生和公共教育等)的公众参与是否会影响公共服务绩效?再如,民主体制下的参与渠道较为多元,公众参与的政府绩效评估是否只会发挥较弱的补偿作用?政府绩效评估需要来自外部与内部的支持,未来研究可以在政治体制和社会文化不同的其他情境中对本文的研究发现加以检验和比较。

参考文献:

- 何艳玲, 2013: “中国城市政府公共服务能力评估报告(2013)”, 北京: 社会科学文献出版社。
- 马亮, 2014: “政府绩效信息使用: 理论整合、文献述评与研究展望”, 《电子科技大学学报》(社科版) 2014, 5: 1—11。
- 2015: “公共管理实验研究何以可能: 一项方法学回顾”, 《甘肃行政学院学报》2015, 4: 13—23。
- 2017: “目标治国、绩效差距与政府行为: 研究述评与理论展望”, 《公共管理与政策评论》2017, 2: 77—91。
- 马亮、于文轩, 2013: “第三方公共服务绩效评价的评价: 一项比较案例研究”, 《南京社会科学》2013, 5: 55—63。
- 吴建南、高小平, 2006: “行风评议: 公众参与的政府绩效评价研究进展与未来框架”, 《中国行政管理》2006, 4: 22—25。
- 吴伟、于文轩、马亮等, 2014: “打造服务型政府, 创建优质城市: 2013 连氏中国服务型政府调查报告”, 《电子政务》, 2014, 4: 18—33。
- 杨宏山, 2017: “超越目标管理: 地方政府绩效管理展望”, 《公共管理与政策评论》2017, 1: 55—61。
- 周志忍, 2015: “政府绩效评估中的公民参与: 中国地方政府的实践与经验”, 北京: 人民出版社。
- 曾莉、李佳源, 2013: “公共服务绩效主客观评价的契合性研究——来自 H 市基层警察服务的实证分析”, 《公共行政评论》2013, 2: 72—103。
- Aguinis, H., R. K. Gottfredson, and S. A. Culpepper, 2013. “Best – Practice Recommendations for Estimating Cross – Level Interaction Effects Using Multilevel Modeling.” *Journal of Management*. 39 (6): 1490 – 528.
- Andrews, R., G. Boyne, and R. Walker, 2011. “The Impact of Management on Administrative and Survey Measures of Organizational Performance.” *Public Management Review*. 13 (2): 227 – 55.

- Andrews, R., and S. Van de Walle, 2013. "New Public Management and Citizens' Perceptions of Local Service Efficiency, Responsiveness, Equity and Effectiveness." *Public Management Review*. 15 (5): 762 – 83.
- Ariely, G., 2011. "Why People (Dis) Like the Public Service: Citizen Perception of the Public Service and the NPM Doctrine." *Politics & Policy*. 39 (6): 997 – 1019.
- Behn, R. D., 2014. *The PerformanceStat Potential: A Leadership Strategy for Producing Results*. Washington, DC: Brookings Institution Press.
- Dalehite, E. G., 2008. "Determinants of Performance Measurement: An Investigation into the Decision to Conduct Citizen Surveys." *Public Administration Review*. 68 (5): 891 – 907.
- de Lancer Julnes, P., and M. Holzer, 2001. "Promoting the Utilization of Performance Measures in Public Organization: An Empirical Study of Factors Affecting Adoption and Implementation." *Public Administration Review*. 61 (6): 693 – 708.
- Enders, C. K., and D. Tofighi, 2007. "Centering Predictor Variables in Cross – Sectional Multilevel Models: A New Look at an Old Issue." *Psychological Methods*. 12 (2): 121 – 38.
- Gao, J., 2012. "How Does Chinese Local Government Respond to Citizen Satisfaction Surveys? A Case Study of Foshan City." *Australian Journal of Public Administration*. 71 (2): 136 – 47.
- Gerrish, E., 2016. "The Impact of Performance Management on Performance in Public Organizations: A Meta – Analysis." *Public Administration Review*. 48 (66): 48 – 66.
- Grosso, A. L., and G. G. Van Ryzin, 2012. "Public Management Reform and Citizen Perceptions of the UK Health System." *International Review of Administrative Sciences*. 78 (3): 494 – 513.
- Heberer, T., and R. Trappel, 2013. "Evaluation Processes, Local Cadres' Behaviour and Local Development Processes." *Journal of Contemporary China*. 22 (84): 1048 – 66.
- Ho, A., and P. Coates, 2004. "Citizen – Initiated Performance Assessment: The Initial Iowa Experience." *Public Performance & Management Review*. 27 (3): 29 – 50.
- Hvidman, U., and S. C. Andersen, 2014. "Impact of Performance Management in Public and Private Organizations." *Journal of Public Administration Research and Theory*. 24 (1): 35 – 58.
- Im, T., and S. J. Lee, 2012. "Does Management Performance Impact Citizen Satisfaction?" *The American Review of Public Administration*. 42 (4): 419 – 36.
- James, O., and P. John, 2007. "Public Management at the Ballot Box: Performance Information and Electoral Support for Incumbent English Local Governments." *Journal of Public Administration Research and Theory*. 17 (4): 567 – 80.
- Li, J., 2015. "The Paradox of Performance Regimes: Strategic Responses to Target Regimes in Chinese Local Government." *Public Administration*. <https://doi.org/10.1111/padm.12197>.
- Ma, L., 2016a. "Performance Feedback, Government Goal – Setting, and Aspiration Level Adaptation: Evidence from Chinese Provinces." *Public Administration*. 94 (2): 452 – 71.
- , 2016b. "Does Super – Department Reform Improve Public Service Performance in China?" *Public Management Review*. 18 (3): 369 – 91.
- , 2017. "Performance Management and Citizen Satisfaction with the Government: Evidence from Chinese Municipalities." *Public Administration*. 95 (1): 39 – 59.
- Ma, L., H. Tang, and B. Yan, 2015. "Public Employees' Perceived Promotion Channels in Local China: Merit – Based or Guanxi – Orientated?" *Australian Journal of Public Administration*. 74 (3): 283 – 97.
- Meier, K. J., and L. J. O' Toole, 2007. "Modeling Public Management – Empirical Analysis of the Management – Performance Nexus." *Public Management Review*. 9 (4): 503 – 27.
- , 2013. "Subjective Organizational Performance and Measurement Error: Common Source Bias and Spurious Relationships." *Journal of Public Administration Research and Theory*. 23 (2): 429 – 56.
- Neshkova, M. I., and H. Guo, 2012. "Public Participation and Organizational Performance: Evidence from State Agencies." *Journal of Public Administration Research and Theory*. 22 (2): 267 – 88.
- Poister, T. H., O. Q. Pasha, and L. H. Edwards, 2013. "Does Performance Management Lead to Better Outcomes? Evidence from the U. S. Public Transit Industry." *Public Administration Review*. 73 (4): 625 – 36.

- So , B. W. Y ,2014. "Civic Engagement in the Performance Evaluation of the Public Sector in China. " *Public Management Review*. 16 (3) :341 – 57.
- Sun ,R. , and G. G. Van Ryzin ,2014. "Are Performance Management Practices Associated with Better Public Outcomes? Empirical Evidence from New York Public Schools. " *The American Review of Public Administration*. 44 (3) :324 – 38.
- Walker ,R. M. , and G. Enticott ,2004. "Using Multiple Informants in Public Administration: Revisiting the Managerial Values and Actions Debate. " *Journal of Public Administration Research and Theory*. 14 (3) :417 – 34.
- Walker ,R. M. , and J. Wu ,2010. "Future Prospects for Performance Management in Chinese City Governments. " *Administration & Society*. 42 (1 suppl) :34S – 55S.
- Wang ,X. ,2002. "Assessing Performance Measurement Impact: A Study of U. S. Local Governments. " *Public Performance & Management Review*. 26 (1) :26 – 43.
- Wholey ,J. S. ,1999. "Performance – Based Management: Responding to the Challenges. " *Public Productivity & Management Review*. 22 (3) :288 – 307.
- Yang ,K. , and J. Y. Hsieh ,2007. "Managerial Effectiveness of Government Performance Measurement: Testing a Middle – Range Model. " *Public Administration Review*. 67 (5) :861 – 79.
- Yu ,W. , and L. Ma ,2015. "External Government Performance Evaluation in China: A Case Study of the 'Lien Service – Oriented Government Project' . " *Public Money & Management*. 35 (6) :431 – 37.

Does Citizen Engagement in Government Performance Measurement Matter : A Multilevel Analysis of Chinese Cities

Ma Liang

(School of Public Administration and Policy , Renmin University of China , Beijing)

Abstract: Citizen engagement has increasingly been adopted in the government performance measurement system (e. g. , citizen survey) to absorb citizen feedback and improve public service performance. The involvement of citizens in performance measurement may help the government to understand the demand of citizens , which may raise their voices in public service delivery to hold public managers accountable. In this paper , we empirically examine the effect of citizen engagement in government performance measurement on various dimensions of government performance. The data on Chinese cities are from multiple sources (e. g. , citizen interviews , public manager surveys , and archival statistics) and the multilevel modeling method is used to test the hypotheses. The findings suggest that citizens are more likely to be satisfied with government performance in cities in which governments involve citizens in performance measurement. We discuss the theoretical and policy implications of the findings , and conclude with limitations and future research avenues.

Key words: Citizen Engagement; Government Performance Measurement; Public Service Satisfaction; Multilevel Modeling; Accountability

(责任编辑: 闫 健)