

# 外部参与、绩效评估与服务型政府建设： 连氏指数的探索与实践\*

马 亮 于文轩 吴 伟\*\*

**【摘 要】** 近年来中国政府外部绩效评估作为政府内部自我绩效评估的有力补充，日益发展壮大起来，越来越得到政界、学界和社会公众的重视。在不断发展壮大的同时，中国政府外部绩效评估也面临着一系列问题和挑战。作为中国政府外部绩效评估诸项目之一，连氏中国城市服务型政府研究项目自2010年启动以来，在研究思路与研究方法上对政府外部绩效评估进行了一系列探索。本文对连氏项目过去四年的演变和发展历程进行了回顾，分析了连氏项目的发起和组织过程，介绍了连氏项目的理论框架、方法论及面临的挑战，并以连氏项目为例，探讨了中国政府外部绩效评估的相关问题与发展方向。

**【关键词】** 政府绩效评价 第三方评估 服务型政府 公共服务质量 连氏指数

以结果为导向对政府内部和外部公共管理活动的绩效进行测评，从而促使政府不断提升自身公共管理能力和改善公共服务质量，是20世纪90年代以来席卷全球的“新公共管理运动”的核心主张和重要改革议题（Osborne & Gaebler, 1992; Hood, 1991; Kettl, 2005）。20世纪80年代末以来，英国政府推行了包括“下一步行动方案”、“公民宪章”和“竞争求质量”等在内的一系列行政改革计划，确立了以结果为导向的政府绩效评估体系，不断提升和改进政府公共管理活动的效益和效果。美国于1993年通过了《政府绩效与结果法案》，以法律的形式要求美国联邦政府部门和机构进行以结果为导向的政府绩效评估。2011年，国务院批准建立政府绩效管理工作部际联席会议制度，由监察部牵头，选择北京、吉林、福建、广西、四川、新疆、杭州、深圳和发展

\* 基金项目：本文受新加坡连氏基金会“连氏中国服务型政府调查项目”资助。

\*\* 作者简介：马亮（1984-），男，江苏铜山人，南洋理工大学南洋公共管理研究生院研究员，兼任西安交通大学绩效管理研究中心研究员，研究方向：战略领导、政府创新与绩效管理；于文轩（1976-），男，河北邯郸人，新加坡南洋理工大学公共政策与全球事务系助理教授，研究方向：比较公共行政、政府绩效评估、公民参与、电子政府与政府信息公开；吴伟（1955-），男，福建仙游人，新加坡南洋理工大学公共政策与全球事务系副教授，研究方向：领导与管理、组织沟通、公共关系与危机管理。

改革委、财政部、国土资源部、环境保护部、农业部、质检总局等 14 个地区和部门开展政府绩效管理试点工作<sup>①</sup>。政府绩效评估逐渐成为国际理论界和实践界的热点, 各国政府都开始开发以结果为导向的政府绩效评估体系, 加大政府绩效评估力度, 不断提升政府绩效 (Bouckaert, 2007)。与此同时, 有别于政府自身发起和开展的绩效评估活动, 由包括国际组织、非营利机构、新闻媒体、智库、公民团体和学术机构在内的各类主体开展的政府外部绩效评估活动发展迅速, 可谓方兴未艾 (Hood et al., 2008; Coe & Brunet, 2006)。

美国锡拉丘兹大学马克斯韦尔公民与公共事务学院在美国著名的非营利组织皮尤基金会的支持下, 自 2002 年以来与新闻媒体合作对州政府和较大的市县的政府管理能力进行评价和排名, 对推动政府能力建设和绩效改进产生重要影响 (Ingraham, 2007)。英国许多非政府机构开展了形式多样的公共服务绩效评估和排名活动, 为地方政府提升公共服务质量和社会公众的参与监督提供了参考依据 (Boyne, 2002)。在国际层面, 包括联合国、世界银行、经济合作与发展组织在内的国际组织都建立了面向国家和地区公共部门的涵盖诸多治理领域的公共绩效评估系统, 定期或不定期地对公共部门绩效进行测评 (Hood et al., 2008)。除此以外, 还有大量非营利组织、学术机构、新闻媒体以及私营咨询公司开展了各种各样以城市竞争力、城市生活质量、城市经商环境等为主题的国际排名和测评活动, 尽管出自各种各样的目的, 这些政府外部的测评和排名活动受到了世界各国和地方政府的高度重视, 成为推动政府公共管理创新和持续改进的重要力量 (马亮, 2013; 马亮 & 于文轩, 2013; 马亮等, 2014)。

近年来, 尽管起步较晚, 中国政府外部绩效评估得到了迅猛发展, 政府外部评估项目越来越多, 影响力越来越大, 成为政府内部绩效评估的重要补充, 对促进政府职能转变、政府行政能力提高和政府绩效改进发挥了重要作用 (郑方辉, 2008; 包国宪等, 2010; 徐双敏, 2011; 高洪成 & 娄成武, 2012)。然而, 目前学术界对政府外部绩效评估的缘起与发展状况的介绍和研究还不够。中国政府外部绩效评估是如何产生和演进的? 中国外部绩效评估的现状如何? 评估活动是如何开展的? 评估质量如何? 评估的效果和影响力如何? 是目前非常需要回答的研究问题。

本文选取新加坡南洋理工大学开展的连氏中国服务型政府研究项目 (下称连氏项目) 作为个案, 将之置于中国政府外部绩效评估发展的背景之下, 回顾连氏项目发起与组织的过程、介绍其评估框架和方法, 分析其经验和面临的问题与挑战, 借此对中国政府外部绩效评估的相关问题进行探讨 (Yu & Ma, 2014)<sup>②</sup>。本文共分五部分。在第一部分, 我们对中国政府绩效评估和政府外部绩效评估的发展进行介绍; 在第二部分, 我们介绍连氏项目的发起和组织过程; 在第三部分, 我们介绍连氏项目的理论基础和研究方法论; 第四部分结合中国政府外部绩效评估的实际, 我们讨论连氏项目存在的问题和面临的挑战。第五部分是结论。

<sup>①</sup> 从自行探索到开展试点 政府绩效管理制度建设提速. 人民日报, 2011 年 09 月 06 日. [http://www.gov.cn/jrzq/2011-09/06/content\\_1940916.htm](http://www.gov.cn/jrzq/2011-09/06/content_1940916.htm).

<sup>②</sup> 笔者对连氏项目 2010-2012 年的回归性总结和批判性研究可见 Yu & Ma (2014)。

## 一、政府绩效评估：从内部封闭到外部开放

沿袭中央计划经济体制的历史遗存，一直以来中国各级政府绩效评估以政府内部特别是政府行政机关自上而下的目标责任考核为主。一方面，通过强有力的目标设置、行政问责与晋升激励机制，对地方政府及其官员施加压力和激发动力（王汉生 & 王一鸽，2009），强有力地促进了中国地方经济发展（周黎安，2008）；另一方面，中国政府的内部评估存在着诸如考核指标不够科学，评估结果利用不合理，对基层公务员激励不够，绩效沟通和报告不充分，评估内容、侧重点和评估方法分散，随意性较大，考核周期不合理，机构设置不合理等制度和设计缺陷，评估开展与否以及采取何种评估形式，缺乏立法的共性约束，容易受个人意志支配，在一定程度上严重影响了绩效评估体系功能的发挥，也诱发了一些负面作用（Gao, 2010）。尚虎平等对中国政府 42 个主要政府绩效测评模式进行了研究，发现尽管有各种各样的测评体系和方法，但是地方政府的绩效仍然和期望有不小的差距（尚虎平 & 赵盼盼，2012）。

由于中国独特的政治制度和单一式的政府架构，政府自身内部（包括行政、司法、立法机关）相对缺乏强有力的监督和制约，如何通过引入政府外部评估主体，弥补政府自身内部评估的不足，成为学界和实践界关注的热点。中国学者一般称政府外部绩效评估为“第三方评估”，但是对究竟何为第三方还存在争议。郑方辉等认为，第三方有独立性、专业性和民间性等特征（郑方辉，2008）。包国宪等将第三方定义为与“政府无隶属关系和利益关系的第三部门和民间机构”（包国宪等，2010）。徐双敏认为第三方包括高校专家、专业公司、社会代表和民众，政府外部主体参与评估的都属于第三方评估（徐双敏，2011）。高洪成和娄成武提出“异体评估”的概念，认为异体评估则指包括党委、人大、政协、民主党派、社会组织、大学及科研院所、新闻媒体和公众等在内的政府系统外的异体或其联合体对政府的评估（高洪成 & 娄成武，2012）。由于中国政治制度和政府行政架构的特殊性，我们在本文中倾向于采用包国宪和郑方辉等的定义，从广义上定义政府，这也是我们使用“政府外部评估”而不使用“第三方评估”的原因。

与此同时，中国政府的治理模式经历了从“管制型政府”向“服务型政府”的转变。中国政府绩效评估从以政府机构内部运作为焦点的绩效评估模式，转向以政府公共服务提供质量和效益为主的绩效评估模式（孟华，2009）。这一转变与国际“新公共管理运动”所倡导的以结果为导向的、强调公民和社会参与的绩效评估相一致（Walker & Wu, 2010）。中国政府内部绩效评估活动开始越来越多地引入公民参与的要素。例如，福建省在机关行政效能建设中大量引入公民评价、公民满意度的成分，中国各地政府策划和主导的以政风行风评议、“万人评政府”等为代表的公民评价活动越来越多，也越来越普遍。政府外部绩效评估主体和评估活动也在过去 10 年间如雨后春笋般出现。例如：

- 兰州大学中国地方政府绩效评价中心在 2005 年对甘肃省政府部门和市政府进行绩效评估，重点调查企业对公共服务质量的评价（包国宪等，2010）。

- 华南理工大学政府绩效评价中心自 2007 年以来对广东省 21 个地级市和 121 个县级政府进行整体绩效评估（郑方辉，2008）。

- 国务院发展研究中心宏观部等在 2007 年采用客观统计数据，对省级政府公共服务

绩效进行评估（陈昌盛 & 蔡跃洲，2007）。

·上海交通大学安泰经济与管理学院在2009年开发“中国城市服务经济指数2010”，对35个城市公共服务投入和产出的客观数据进行评价（陈宪 & 康艺凡，2011）。

·中国社会科学院马克思主义研究院在2011年对38个大城市的公共服务能力进行评估（侯惠勤等，2011）。

·上海财经大学、北京大学分别对省级政府的财政透明（上海财经大学公共政策研究中心，2009）和行政透明（北京大学公众参与研究与支持中心，2011）进行评估。

·《瞭望东方》周刊连续多年对中国地方政府在改善民生、创建幸福城市方面的努力进行测评和排名<sup>①</sup>。

·零点调查公司自2005年起发布《零点中国公共服务公众评价指数报告》，对不同层级的样本地方政府的公共服务进行公众满意度调查（袁岳 & 张慧，2012）。

如前文所言，尽管中国政府外部绩效评估发展方兴未艾，然而对中国政府外部绩效评估的研究还比较欠缺。外部绩效评估主体的组成是什么？外部绩效评估主体和政府的关系是什么？外部绩效评估项目开展的目的是什么？资金来源是什么？外部绩效评估如何保持其客观中立性？外部绩效评估项目的内容是什么？评估的方法论和质量如何？外部绩效评估项目的可持续性如何？社会影响力如何？能否和政府与社会形成良性互动并促进政府绩效的持续改进？对这些问题的回答对中国政府绩效管理不仅具有重要的理论意义，对中国政府外部绩效评估事业的可持续发展和中国政府绩效评估的实践推进也有着重要的实践价值。连氏项目在这些方面进行了一系列探索，对其进行审视将对理解和促进中国外部政府绩效评估的发展有相当大的参考价值。

## 二、连氏项目的发起与组织

连氏项目是2009年在新加坡连氏基金的大力赞助下，于新加坡南洋理工大学南洋公共管理研究生院成立的。新加坡连氏基金是为纪念新加坡已故著名华人企业家、外交家、社会活动家连赢洲先生而创办的。多年来，连氏基金在东南亚和中国开展了大量的以扶贫、清洁水、能源以及养老为中心的慈善援助项目<sup>②</sup>。随着中国经济的快速发展和新中两国之间关系的日益密切，连氏基金对在中国开展研究项目并促进中国社会经济发展的兴趣越来越浓厚。南洋公共管理研究生院是中国海外培训高级政府官员的重要基地之一，一直以来都以通晓西方、新加坡和中国公共管理发展为己任，促进新中两国公共管理理论和实践发展<sup>③</sup>。2009年，在时任院长吴伟教授的倡议下，连氏基金成立了连氏项目，以种子基金的形式对在中国开展政府公共服务研究提供资金支持。

2009年底和2010年初，“连氏中国公共服务质量调查”正式启动。以吴伟教授和于文轩博士为首的研究团队开始与在政府绩效评估研究领域实力雄厚、享有盛誉的厦门大学公共事务学院陈振明教授带领的团队合作研究，开发中国城市公共服务质量指数和调查方

① 对2012中国最具城市幸福感城市调查的介绍参见：<http://city.lwdf.cn/index.html>。

② 详见连氏基金的主页：<http://www.lienfoundation.org/>。

③ 详细内容参见南洋公共管理研究生院的主页：<http://www.ncpa.ntu.edu.sg/>。

案。2010年，两校研究团队选择中国大陆32个主要城市作为调查对象，对其公共服务质量进行测评。2010年9月19日，由两校专家组成的“连氏中国公共服务质量调查”项目组在广州举办新闻发布会并公布中国32个城市公共服务质量排名。这是首次由中国著名高校和海外著名高校通力合作进行的中国城市公共服务质量调查和排名，在社会上引起了广泛关注（吴伟等，2011）。

2011年，结合中国公共行政改革的新发展和政府外部绩效评估研究和实践发展的实际，项目组认为单单对政府公共服务进行研究和测评已经不能反映理论和实践的新发展和新需要，将研究主题从单纯的公共服务质量提升拓展为服务型政府建设，增加了政府效能、信息公开、公众参与和政府信任等维度，结合中外相关研究和中国政府行政改革的新理论和实践，对中国服务型政府建设进行了深入研究，对中国服务型政府概念的内涵和外延进行了提炼和总结，发展出了“中国服务型政府指数”测评体系。鉴于上海交通大学国际与公共事务学院在计算机辅助电话访问（CATI）调查方面的优势，2011年连氏项目开始委托其借助大规模随机抽样电话访谈，提高数据的可靠性和代表性（于文轩等，2012；Wu，2012）。

2012年的测评延续2011年的做法，在保持评估体系的连贯性和可比性的基础上，对评估指标进行了精炼加工，并进一步调整了调查范围，增加了西宁、银川和呼和浩特，涵盖了除乌鲁木齐和拉萨以外的中国大陆34座城市（直辖市、副省级城市、省会城市和自治区首府）（吴伟等，2013）。

2013年的测评框架和调研方法同2012年的基本保持一致，以利于跨年度比较。与此同时，在大陆地区将乌鲁木齐和拉萨纳入调查，涵盖所有直辖市、副省级城市、省会城市和自治区首府。并增加香港和台北作为观察城市纳入调查。企业样本量有所增加，从每个城市平均100家增加到150家。此外，在调查中还增加了有关公民幸福感、电子政务使用率和满意度、官方媒体评价等内容（吴伟等，2014）。

表1 连氏项目的演变与发展

项目	2010年	2011年	2012年	2013年
评估主题	公共服务质量	服务型政府	服务型政府	服务型政府
指标数量	106	75	68	74
调查城市	32	32	34	38
调查方法	面访	电话访问	电话访问	电话访问
数据来源	问卷调查 + 统计数据 + 独立评估	问卷调查 + 统计数据	问卷调查 + 统计数据	问卷调查 + 统计数据
公民样本	31173	25222	23923	25958
企业样本	3153	3203	3606	5552
子维度加总	主成分分析/标准化回归/算术平均	主成分分析/标准化回归/算术平均	主成分分析/算术平均	主成分分析/算术平均
维度加总	未加总	1:1:1	4:3:3	4:3:3

资料来源：根据历年调查报告的主要参数整理而得

总体来说,经过连续四年的实践与探索,目前连氏项目已经建立了较为完善的组织和管理体系,理论框架和评估体系也日趋成熟(见表1)。由于有连氏基金的大力支持,项目能够实现较为稳定和持续的运营。目前连氏项目一般每年3-4月规划项目工作,修订调查问卷并完善评估体系;5-7月执行问卷调查;8-10月统计分析并编纂研究报告;10-11月发布评估结果并总结项目工作;年底和次年年初开发和推介项目研究成果。通过这种连续一贯的项目执行和管理,为积累中国城市服务型政府的连续数据提供了基础,也为建立一种独立审慎的评估模式和评估文化提供范例。

### 三、连氏项目的理论框架、评价体系和研究方法

#### (一) 连氏项目的理论框架

自2011年起,连氏项目对中国城市公共服务提供的关注扩展到了以中国政府服务型政府建设作为研究和评价的中心。2003年爆发的“非典”(SARS)危机使中国政府面临着行政问责、危机管理和公共服务提供等诸多方面的挑战,2004年中央政府正式确立服务型政府的建设理念,将其作为行政管理体制改革和政府自身建设的重要纲领。项目组认识到建设服务型政府是当前中国政府深化行政管理体制改革的重点和核心,而责任政府、法治政府和廉洁政府的建设则是其保障手段(薄贵利,2012)。因此,服务型政府的内涵不限于公共服务提供,而是反映了人们对良好政府和善治的诸多价值取向,如透明公开、廉洁高效、参与包容、诚信公正等。从这个意义上讲,我们需要建立一种服务型政府的整体观,对其进行全方位的考察和评价,方能准确把握和定位各地区在服务型政府建设过程中的进展、不足和前景,并提供具有指导性的政策建议(吴伟等,2014)。

结合国内外服务型政府的相关研究(薄贵利,2012;燕继荣,2009;郁建兴&徐越倩,2012),项目组认为服务型政府需要强调两个视角:公民视角和整体观。强调服务型政府的公民视角,即构建公民导向的服务型政府,因为只有真正将公民置于公共服务的中心位置,才能准确把握服务型政府的核心内涵。强调服务型政府同时还要强调整体观,要构建和完善公共服务体系,实现全面均衡发展,使政府在公共服务的各个领域和服务型政府的各个维度都能够实现均衡协调可持续发展,从而推动人的全面发展和社会的整体提升。项目组把服务型政府界定为“廉洁负责、法治高效、透明公开、民主参与为基本原则,将公共服务供给置于政府职能的重心并着力构建和完善公共服务体系,从而为公民和企业发展营造良好的制度环境,不断满足公众日益增长的物质文化需求并赢得广大人民群众信任和支持的政府”(吴伟等,2014)。

#### (二) 服务型政府的评价体系

2010年首次测评时,课题组侧重于公共服务质量调查,通过德尔菲法和大量前期测评,最终确定的2010年连氏中国城市公共服务质量评价指标体系包括公共服务满意度、企业经营环境和一般公共服务3个维度共106项测评指标(吴伟等,2011)。公共服务满意度包括公民满意度和企业满意度两个方面,分别涵盖10个和7个公共服务领域,数据

来自问卷调查。企业经营环境包括7个领域，数据分别来自统计资料和问卷调查。一般公共服务包括10个领域，数据主要来自统计资料和研究人员的独立评估，统计资料主要来自政府发布的官方数据，独立评估则由经过培训的调查人员在公交站点和办税大厅观察记录公交候车时间和税务登记速度。

2011年，项目组将研究的主题从单纯的公共服务质量拓展为服务型政府，增加了政府效能、信息公开、公众参与和政府信任等维度，使之能够反映服务型政府建设的总体进展。通过德尔菲法，我们确定的服务型政府评价体系包括服务型政府公众视角、服务型政府企业视角和基本公共服务3个维度，下设10个子维度和75个测量指标（于文轩等，2012）。2012年，为了保持评估的连贯性和可比性，评价指标体系基本保持不变。

连氏项目的评价体系包括服务型政府公众视角、服务型政府企业视角和基本公共服务3个维度。前两个维度的数据主要来自面向公众和企业的问卷调查，基本公共服务的数据主要来自政府统计年鉴和工作报告中有关公共服务投入与产出的数据。

服务型政府指数的公众视角包括5个维度：公众公共服务满意度、政府效能、政府信息公开、公众参与和政府信任。公众公共服务满意度指公众对公共教育、医疗卫生、住房与社会保障、公共安全、基础设施、文体设施、环境保护、公共交通8项与老百姓生活息息相关的重要公共服务领域的主观感知和满意度。政府效能指公众在与政府打交道和获取公共服务的过程中对政府工作效率和能力的感知和评价。政府信息公开指公众对政府政策制定和执行等相关信息的可获得性的认知和评价。公众参与指公众对政府在多大程度上允许公众参与和影响公共政策制定和执行的认知与评价。政府信任指公众对政府是否真心实意为自己谋福利及在没有监督的情况下也可以做正确的事的信心。

服务型政府指数的企业视角包括企业公共服务满意度、企业经营环境、企业参与以及政府效能4个维度。企业公共服务满意度包括企业经营者对其经营和发展所需要的基础设施、交通运输、公共安全、就业服务、环境保护、信息服务6项公共服务质量的满意度。企业经营环境指企业对工商登记、年检服务、税收服务、知识产权保护等10项重要的经营环境要素的满意度。企业参与是企业对在多大程度上可以参与政府制定和执行商业政策的感知和评价。政府效能指企业在与政府打交道的过程中对政府提供公共服务的能力和效率的认知和评价。

服务型政府指数的基本公共服务维度涵盖就业服务、住房保障、公共安全、公共教育、医疗卫生、环境保护、社会保障、基础设施、公共交通和文体休闲10个公共服务领域的投入、产出和结果，通过客观统计数据反映城市公共服务的基本状况。项目组希望通过引入客观统计数据指标，纠正主观调查数据的偏差，全面反映各城市服务型政府建设的整体状况，也符合政府绩效评估的发展趋势（Shingler et al., 2008）。

### （三）连氏项目的调研方法

#### 1. 服务型政府的调查范围

连氏项目选择中国36个城市（直辖市、副省级城市、省会城市和自治区首府）作为测评对象，是因为服务型政府建设主要是在中国城市开展的。随着中国城镇化速度的加快，中国已经有超过一半人口居住在城市里。在进行样本城市选择时，选择最能代表中国城市发展水平和面貌、最引人关注、也最能反映中国城市公共服务提供现状的主要城市，

是自然的选择,对其进行评价也比较容易引起公众和政府相关部门的兴趣和重视。

同时,连氏项目也希望通过测评可以引起国内外企业界的关注和重视,为企业投资选址提供参考。此外,选择这些城市也与研究项目经费的限制和城市数据的可获取性有关。对这些城市的研究,为在条件成熟时将调查对象拓展到重要的地级市、县级市乃至乡镇提供了理论和技术基础,也为连氏项目的进一步发展提供了空间。

连氏项目从最初的32座城市开始,不断扩大调查范围,至今已涵盖大陆所有36座城市<sup>①</sup>。与此同时,将调查范围扩大到大中华区的香港、台北等城市,也有利于选取标杆城市进行对比,有利于城市之间的知识转换和交流学习。

## 2. 样本量与抽样技术

从目前的政府外部绩效评估项目来看,评估者通常会在主观绩效调查时选择企业或公民进行调查,而很少同时将二者纳入调查范围。基于项目组对服务型政府概念的界定(政府服务对象不限于公民,也不止于企业)和对政府公共管理实践发展(虽然一些地方在早期的服务型政府建设过程中强调服务企业和地方经济,但越来越多的地方政府认识到服务型政府需要兼顾服务公民和推动社会发展)的观察,我们在调查中同时涵盖公民和企业,突出以民为本和均衡发展的服务型政府理念,纠正以往重商轻民的倾向,在样本数量和评估重点等方面都侧重于公民。

本质上来说中国的所有调查都是“地方性”的,在进行推断性的描述时,可能存在偏颇,但对于研究变量之间的相关关系而言则并无大碍(Manion, 1994)。在项目执行过程中,项目组综合考虑到项目的可操作性、质量控制、时间、预算、样本代表性和抽样精确度等因素,保证公众调查在99%的置信水平下抽样误差不超过 $\pm 5\%$ 。

2010年,连氏项目的调查采用了面对面访谈的调查方法。市民问卷主要在商业区、居民区等活动场所随机抽取,每个城市约1000个样本,共计收集31173份问卷。企业问卷则根据当地企业的类型和规模在每个城市随机抽取约100家,共计3153份问卷。在问卷调查的同时,项目组还培训访问员在公交站点和办税大厅观察记录公交候车时间和税务登记速度,进行实地的独立评估。尽管项目组对样本的代表性进行了各种各样的控制,但是由于实践上的困难和限制,上述调查方法的代表性仍有一定问题。

2011年以来,出于问卷调查的随机性、准确性和可控性等方面的考虑,为进一步提高样本的代表性和降低调研成本,我们采用计算机辅助电话访问(CATI)系统进行问卷调查。由于目前固定电话和移动电话的普及率较高,这种根据电话号段随机抽取的方式更加精确。2011年在9个特大城市分别抽取1000个左右的公众样本,其他23个城市700个左右,总共25222个。企业调查仍然按照每个城市100个的额定样本抽取,总共3203个。2012年在调查样本量方面,基本实现每个城市调查700个左右的公民和100个以上的企业,公民和企业问卷分别共计23923和3606份,一定程度上确保调查结论的统计推断性。2013年的调查样本分布基本延续此前做法,只是在企业调查方面的样本量有所增加,以观察其可能带来的变化。

<sup>①</sup> 连氏项目在2010-2011年间曾将苏州纳入调查,但由于苏州不属于副省级城市,鉴于可比性的考虑,在此后调查中未予考察。



### 3. 问卷设计

区别于完全依赖客观统计数据的外部评估，连氏项目注重来自公众和企业的声音，因为他们是公共服务的消费者，也是服务型政府建设的直接受益者，他们的意见至关重要。国际公共管理理论界和实践界对这种主观认知和看法的重视甚至超过了对客观现实的关注。

2010年的调查项目组采取面访的方式，因此可以设置较多的题项，以综合反映公众和企业的感知水平。但此后的调查由于是电话访问，问卷长度难以涵盖许多方面，否则将导致拒答率过高，影响调查的效度。2011年的调查证实了这一点。在反复测试的基础上，2012年项目组对问卷中问题的提问方式和方法进行了微调，精简了市民问卷的题项，有效降低了问卷的拒答率，提高了问卷回答的有效性和数据质量。此外，为了考察公共服务的期望与满意度的关系，我们还增加了公众总体期望和满意度的题项。2013年的调查工具与此前相似，但我们仍然进行了取舍。在确保调查长度稳定的情况下，我们精简了若干题项，并增加了一些反映政策发展趋势的题项。

### 4. 加权与标准化

政府绩效评估都可以归为多指标综合评价体系问题，即通过诸多评估指标来反映多维度的政府绩效。为了能够获得一个整体性的评价，以形成类似于“成绩单”、“排行榜”的结果，需要通过一定的加权方式进行降维和聚合（Nardo et al., 2005）。如何对子指标体系进行加权，是连氏项目面临的方法论上最严峻的挑战之一。

2010年连氏项目的城市公共服务总体评价包括公众和企业两个视角，而没有将其加总，以避免两者相加可能抵消不同差距（吴伟等，2011）。但随后的两年调查都尝试进行一定的加总，以反映服务型政府的整体状况。结合国内外的相关理论和实践，项目组采用德尔菲法咨询专家决策和计算机模拟技术，获得一个相对可以接受和较为稳健的加权方法，以期客观反映服务型政府建设的成果。2011年，项目组采取等权重的方式加总，即公众视角、企业视角和基本公共服务的比重为1:1:1。

2012年，项目组在同专家的沟通咨询的过程中认识到，由于公众调查的样本量和代表性都较强，我们将原来的等权重加总修改为4:3:3的加权比例，增加了公民评价的权重，以强化服务型政府的公民导向。为了使排名结果不至于过度受到权重的影响，项目组还进行了多种仿真测算，发现权重的微小调整对排名结果的整体影响可以忽略不计，排名的结果较为稳健。2013年也遵循这一赋权原则，对三个维度进行汇总。

与此同时，为了便于公众理解和形象展示，项目组还将最终的标准化结果进行了逆向处理，使之以公众熟悉的十分制表示。显然，分数的高低只是相对的，最终加总的得分不具有明确的意义指向，而只是作为公众评判政府优劣的一个相对坐标系。

## 四、从连氏项目看中国政府外部绩效评估

由于中国复杂的国情和独特的政治、社会和经济制度和发展状况，中国政府外部绩效评估和国际经验相比呈现出非常独特的风貌。由于公民社会和非营利组织发展不足，尽管有零星项目是由私营企业（如零点调查公司）和大众传媒（如《瞭望东方周刊》）进行的，中国政府外部绩效评估项目基本上主要是高等教育机构和科研机构完成的。因此，从

项目的资金来源和治理模式等诸方面来看，政府外部绩效评估和政府主导的内部绩效评估呈现出非常复杂多样的联系和表现形式。中国高等教育机构和科研机构要么接受政府委托进行政府绩效测评，要么申请国家和地方政府的社会科学或自然科学基金资助进行调研和测评。这在一定程度上对外部政府绩效评估的独立性、公信力和项目的可持续性产生了一定的影响。

连氏项目得到了新加坡连氏基金会的鼎力赞助，以种子基金的形式为研究项目提供支持，并对研究过程不进行任何干预，是慈善性质的援助，是项目持续开展的保证。南洋公共管理研究生院是中国重要的海外培训高级政府官员的基地之一，一直都受到中国各级政府的高度重视。连氏项目组的专家都受过国际著名高校公共管理和政府绩效评估的专业训练，又有在中国生活和工作的背景，既通晓国际先进的理念、理论和技术，又了解中国国情。连氏项目选择的中国合作伙伴都是在中国享有盛誉，在公共部门绩效管理研究领域实力雄厚的著名大学，他们为研究项目提供了非常强有力的调研支持，为研究项目在中国的开展扫除了政策障碍、文化障碍、行政障碍和技术障碍。这样的组合突出了连氏项目的独立性，强调了项目的本土化和针对性，提高了连氏项目被接受和受重视的程度。

政府外部绩效评估的另一重大挑战是研究的质量和政府对公共管理活动的影响。高质量的评估能为政府公共管理活动的改进和提升发挥实质性作用，是政府外部评估项目可否成功的关键。连氏项目经过四年的不断调整、充实和提高，已经形成了自己的研究体系和框架，反映了目前最先进的理念和方法。连续四年的研究表明，研究结果呈现出相当的稳定性，说明了研究体系的科学性和有效性。尽管如此，连氏研究项目在调查范围、城市样本的选取、问卷的样本代表性、理论框架、指标体系设计、数据收集、加权汇总等方面还有很多可以改进的空间。不断提升研究的信度和效度，不断提升测评质量，是政府外部绩效评估项目的生命力的源泉。

政府绩效评估的目的在于将其产生的绩效信息予以充分利用，借此促进组织学习和组织绩效的不断改进（Behn, 2003）。政府内部的绩效评估通常会将评估结果用于内部奖惩、学习和差距识别等方面，通过强力激励推动政府部门改进绩效。与政府自身组织开展的绩效评估不同，政府外部绩效评估的结果如果能受到重视，成为政府持续改进的动力，将关系到政府外部绩效评估项目的成败。连氏项目每年都于11月份召开大型新闻发布会，邀请国内数十家主流媒体（传统和网络）参加<sup>①</sup>。每年新华社、人民网等权威新闻机构都推出专稿对排名和研究结果进行发布，新加坡《联合早报》和《海峡时报》也对发布会进行全程报道。部分获奖城市（如成都、长春等）也以此为契机，在本地媒体进行长篇连载专题报道，总结和反思本地经验及不足。在2012年和2013年的新闻发布会上，项目组还专门邀请获奖城市代表出席发布会，分享城市建设服务型政府的心得和经验。2011 - 2012年连续两年荣登榜首的厦门市不仅在《厦门日报》头版头条报道连氏项目的测评成果，还将之写入政府工作报告，指导本市服务型政府建设工作的进一步开展。从一定意义上说，厦门连续两年获得服务型政府十佳城市之首的桂冠，跟政府高度重视连氏项目，积极推进服务型政府建设不无关系。

目前，连氏项目界定了服务型政府的概念并确立了成熟的评价体系，定期进行城市服

<sup>①</sup> 2013年的新闻发布专题网详见：<http://sg.xinhuanet.com/zb/4/>。

务型政府建设排名。从长远来看，连氏项目作用的进一步发挥还有赖于深入政府内部管理过程，将政府管理过程与结果联系起来，识别和发现推动政府绩效改进的关键影响因素，为城市政府管理不断改善和优化提供决策咨询。这有赖于更为深入的实地调查和案例研究，通过与城市政府合作将连氏项目的研究发现转化为改进服务型政府建设的政策建议，真正发挥连氏项目的研究价值。当然，在与城市互动合作的过程中，如何保持测评活动的客观和中立将是连氏项目的另一个新的挑战。

## 五、结论

本文回顾了中国政府外部绩效评估的兴起和发展的过程，介绍了连氏中国服务型政府调查项目的发起、组织和在探索中国政府外部绩效评估方面进行的努力和尝试，尽管连氏项目具有一定的特殊性，但是它的经验和教训对于发展中国政府外部绩效评估具有普遍意义。连氏项目的发展表明，中国政府外部绩效评估面临复杂的政治、文化、社会环境，面临着本土化、理论构建、方法论完善、独立性保持、可持续性发展以及影响力等方方面面的挑战。连氏项目的发展在这些方面对中国政府外部绩效评估项目的发展是有一定启发和借鉴意义的。

随着时代的发展，中国政府外部绩效评估迎来了难得的历史机遇，政府越来越重视外部绩效评估的结果（Gao, 2012）。政府外部绩效评估日益形成规模并逐渐产生越来越重要的社会影响。更为重要的是，当越来越多的“连氏项目”加入政府外部绩效评估的行列中，通过独立、科学和持续的绩效评估形成对政府运作与管理的有效外部监督和激励时，政府外部绩效评估对中国行政体制改革、行政管理创新以及政府行政管理水平和能力提升的作用和影响将是深远和不容忽视的。

### 参考文献

1. 包国宪,董静,郎玫,王浩权,周云飞,何文盛. 第三方政府绩效评价的实践探索与理论研究——甘肃模式的解析 [J]. 行政论坛, 2010 (4).
2. 薄贵利. 准确理解和深刻认识服务型政府建设 [J]. 行政论坛, 2012 (1).
3. 北京大学公众参与研究与支持中心. 中国行政透明度观察报告 (2009 年度) [M]. 北京: 法律出版社, 2011.
4. 陈昌盛,蔡跃洲. 中国政府公共服务: 体制变迁与地区综合评估 [M]. 北京: 中国社会科学出版社, 2007.
5. 陈宪,康艺凡. 中国城市公共服务指数·2010 [J]. 科学发展, 2011 (2).
6. 高洪成, 娄成武. 异体评估: 我国政府绩效评价的路径选择及理论建构 [J]. 中国行政管理, 2012 (9).
7. 侯惠勤,辛向阳,易定宏. 公共服务蓝皮书: 中国城市基本公共服务力评价 (2010-2011) [M]. 北京: 社会科学文献出版社, 2011.
8. 马亮. 城市排行榜: 流行、问题与展望 [J]. 甘肃行政学院学报, 2013 (3).
9. 马亮,于文轩. 第三方公共服务绩效评价的评价: 一项比较案例研究 [J]. 南京社会科学, 2013 (5).
10. 马亮,于文轩,吴伟. 公共服务提供、绩效测评与政府信任: 2013 连氏公共管理国际会议综述

- [J]. 甘肃行政学院学报, 2014 (1).
11. 孟华. 推进以公共服务为主要内容的政府绩效评估——从机构绩效评估向公共服务绩效评估的转变 [J]. 中国行政管理, 2009 (2).
  12. 上海财经大学公共政策研究中心. 2009 中国财政透明度报告——省级财政信息公开状况评估 [M]. 上海: 上海财经大学出版社, 2009.
  13. 尚虎平, 赵盼盼. 绩效评估模式泛滥与绩效不彰困境 [J]. 中国行政管理, 2012 (11).
  14. 王汉生, 王一鹤. 目标管理责任制: 农村基层政权的实践逻辑 [J]. 社会学研究, 2009 (2).
  15. 吴伟, 于文轩, 林挺进, 马亮, 盛明科, 王君. 完善服务型政府体系, 实现全面均衡发展: 2012 连氏中国服务型政府调查报告 [J]. 经济研究参考, 2013 (10).
  16. 吴伟, 于文轩, 林挺进, 王君. 提升城市公共服务质量, 打造服务型政府——2010 连氏中国城市公共服务质量调查 [J]. 城市观察, 2011 (1).
  17. 吴伟, 于文轩, 马亮, 孟羽, 王君, 林挺进. 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 [J]. 电子政务, 2014 (4).
  18. 徐双敏. 政府绩效管理中的“第三方评估”模式及其完善 [J]. 中国行政管理, 2014 (1).
  19. 燕继荣. 服务型政府的研究路向——近十年来国内服务型政府研究综述 [J]. 学海, 2009 (1).
  20. 于文轩, 林挺进, 吴伟. 提升政府治理水平, 打造服务型政府——2011 连氏中国服务型政府指数及中国城市服务型政府调查报告 [J]. 华东经济管理, 2012 (7).
  21. 郁建兴, 徐越倩. 服务型政府研究的理论进路与出路 [J]. 行政论坛, 2012 (1).
  22. 袁岳, 张慧. 2011 年中国城市居民生活质量指数报告 [A]. 汝信, 陆学艺, 李培林. 社会蓝皮书: 2012 年中国社会形势分析与预测 [C]. 北京: 社会科学文献出版社, 2012.
  23. 郑方辉. 第三方评价地方政府整体绩效的实证研究——以广东省市、县两级政府为例 [J]. 中国行政管理, 2008 (5).
  24. 周黎安. 转型中的地方政府: 官员激励与治理 [M]. 上海: 格致出版社, 2008.
  25. Behn, R. D. Why Measure Performance? Different Purposes Require Different Measures. *Public Administration Review*, 2003. 63 (5).
  26. Bouckaert, G., Halligan, John. *Managing Performance: International Comparisons*. New York: Routledge, 2007.
  27. Boyne, G. A. Concepts and Indicators of Local Authority Performance: An Evaluation of the Statutory Frameworks in England and Wales. *Public Money and Management*, 2002. 22 (2).
  28. Coe, C. K., & Brunet, J. R. Organizational Report Cards: Significant Impact or Much Ado about Nothing? *Public Administration Review*, 2006. 66 (1).
  29. Gao, J. Hitting the Target But Missing the Point: The Rise of Non - Mission - Based Targets in Performance Measurement of Chinese Local Governments. *Administration Society*, 2010. 42 (1 suppl).
  30. Gao, J. How Does Chinese Local Government Respond to Citizen Satisfaction Surveys? A Case Study of Foshan City. *Australian Journal of Public Administration*, 2012. 71 (2).
  31. Hood, C. A Public Management for All Seasons? *Public Administration*, 1991. 69 (1).
  32. Hood, C., Dixon, R., & Beeston, C. Rating the Rankings: Assessing International Rankings of Public Service Performance. *International Public Management Journal*, 2008. 11 (3).
  33. Ingraham, P. W. *In Pursuit of Performance: Management Systems in State and Local Government*. Baltimore: John Hopkins University Press, 2007.
  34. Kettl, D. F. *The Global Public Management Revolution: A Report on the Transformation of Governance*. Washington: Brookings Institution Press, 2005.
  35. Manion, M. Survey Research in the Study of Contemporary China: Learning from Local Samples. *The*

China Quarterly (139), 1994.

36. Nardo, M. , Saisana, M. , Saltelli, A. , Tarantola, S. , Hoffman, A. , & Giovannini, E. Handbook on Constructing Composite Indicators: Methodology and User Guide. In OECD Statistics Working Papers. Paris: OECD Publishing, 2005.

37. Osborne, D. , & Gaebler, T. Reinventing Government: How the Entrepreneurial Spirit is Transforming the Public Sector. Reading, Mass. : Addison – Wesley, 1992.

38. Shingler, J. , Van Loon, M. E. , Alter, T. R. , & Bridger, J. C. The Importance of Subjective Data for Public Agency Performance Evaluation. Public Administration Review 68 (6) , 2008.

39. Walker, R. M. , & Wu, J. Future Prospects for Performance Management in Chinese City Governments. Administration & Society 42 (1 suppl) , 2010.

40. Wu, W. , ed. Building Service – Oriented Government in China: Lessons, Challenges and Prospects. Singapore: World Scientific, 2012.

41. Yu, W. , & Ma, L. External Government Performance Evaluation in China: A Case Study of the Lien Service – Oriented Government Project. Public Money and Management forthcoming, 2014.